Machine Translated by Google

Enciclopedia

Autor y editor

Elzein Salah Elzein

Miembro del IEEE

• Artículos sobre Matemática Aplicada

Editor IEEE, Reino Unido

Este contenido es confidencial

Machine Translated by Google

AppliedMath

MDPI

Artículo

Modelando el impacto de la desinformación en la transmisión Dinámica del COVID-19

Ziyi Su † y Efraín Agyingi *,†

Escuela de Matemáticas y Estadística, Instituto de Tecnología de Rochester, Rochester, NY 14623, EE. UU.; zs4768@rit.edu
* Correspondencia: eoasma@rit.edu
† Estos autores contribuyeron igualmente a este trabajo.

Resumen: La amenaza que plantea la pandemia de COVID-19 ha ido acompañada de una epidemia de desinformación, que ha provocado confusión y desconfianza entre el público. La información errónea sobre la COVID-19, ya sea intencional o no, adopta muchas formas, incluidas teorías de conspiración, tratamientos falsos e información inexacta sobre los orígenes y la propagación del virus. Aunque la pandemia ha sacado a la luz el importante impacto de la desinformación en la salud pública, los modelos matemáticos surgieron como una herramienta valiosa para comprender la propagación de la COVID-19 y el impacto de las intervenciones de salud pública. Sin embargo, ha habido investigaciones limitadas sobre el modelado matemático de la difusión de información errónea relacionada con COVID-19. En este artículo, presentamos un modelo matemático de la difusión de información errónea relacionada con COVID-19. El modelo destaca los desafíos que plantea la desinformación, ya que en lugar de centrarse únicamente en el número de reproducción que genera nuevas infecciones, existe un parámetro de umbral adicional que impulsa la difusión de información errónea. Se analizan los equilibrios del modelo para determinar la estabilidad local y global y se presentan simulaciones numéricas. También discutimos el potencial del modelo para desarrollar estrategias efectivas para combatir la desinformación relacionada con COVID-19.

Palabras clave: COVID-19; desinformación; modelo matemático

check for updates

Cita: Su, Z.; Agyingi, E. Modelado del impacto de la desinformación en la dinámica de transmisión de COVID-19. Matemáticas aplicadas 2024, 4, 544–560. https://doi.org/10.3390/appledmath4020029

Editor académico: Thomas Woolley

Recibido: 1 de marzo de 2024
Revisado: 24 de abril de 2024
Aceptado: 26 de abril de 2024
Publicado: 30 de abril de 2024

Copyright: © 2024 por los autores. Licenciatario MDPI, Basilea, Suiza. Este artículo es un artículo de acceso abierto. distribuido bajo los términos y condiciones de los Creative Commons Licencia de atribución (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

1. Introducción

La COVID-19, también conocida como enfermedad por coronavirus, causada por el nuevo coronavirus SARS-CoV-2, ha tenido un profundo impacto en la salud mundial desde su aparición a finales de 2019. El primer caso se informó en Wuhan, China, y el virus Se extendió rápidamente por todo el mundo y se observaron múltiples oleadas de infecciones en diferentes regiones. Después de más de 118.000 casos en 114 países y 4291 muertes, la Organización Mundial de la Salud (OMS) declaró la COVID-19 pandemia el 11 de marzo de 2020 [1,2]. El COVID-19 se transmite principalmente a través de gotitas respiratorias cuando una persona infectada tose, estornuda, habla o respira con dificultad cerca de otras personas [3]. Presenta un amplio espectro de manifestaciones clínicas, que van desde síntomas leves similares a los de la gripe hasta dificultad respiratoria grave e insuficiencia multiorgánica [4,5]. Varios factores, incluido el comportamiento social, han influido en la dinámica de transmisión de la enfermedad. El COVID-19 continúa evolucionando, con variaciones en la prevalencia y gravedad de la enfermedad debido a variantes emergentes, campañas de vacunación y medidas de salud pública implementadas por gobiernos y organizaciones de salud. A nivel mundial , hasta el 27 de septiembre de 2023, se han notificado a la OMS 770.875.433 casos confirmados de COVID-19, incluidas 6.959.316 muertes [6].

Durante la pandemia, muchos países implementaron una serie de medidas para contener la propagación de COVID-19, como distanciamiento social, enmascaramiento, autocuarentena y, finalmente, vacunas cuando estuvieron disponibles, lo que esencialmente ayudó a regular la transmisión de la enfermedad. la enfermedad del coronavirus. Sin embargo, la falta de mensajes claros y concisos por parte de los funcionarios de salud pública también contribuyó a la difusión de información errónea relacionada con el COVI
La naturaleza en constante evolución de la pandemia y la falta de comprensión del virus

Machine Translated by Google

Desde el principio generó información y mensajes contradictorios por parte de los funcionarios de salud pública, lo que generó confusión y desconfianza entre el público. Esto llevó a algunas personas a recurrir a fuentes alternativas de información, incluida la información errónea, en busca de orientación.

La información errónea relacionada con la COVID-19 ha adoptado diversas formas [7–9]. Las teorías de conspiración relacionadas con la COVID-19 han circulado ampliamente, y a menudo afirman que el virus fue creado y propagado intencionalmente como un arma biológica o parte de un complot global [10]. Por ejemplo, la teoría de que la tecnología 5G es responsable de la propagación de la COVID-19 se ha compartido ampliamente en las plataformas de redes sociales, a pesar de la falta de evidencia científica que respalde esta afirmación [11]. También se han promovido tratamientos y curas falsos para la COVID-19 , lo que tiene consecuencias peligrosas para las personas que los siguen. Por ejemplo, se ha informado ampliamente sobre el uso de desinfectantes, como la lejía, para tratar la COVID-19 , lo que ha provocado lesiones graves y muertes [12]. También han circulado afirmaciones falsas de que ciertos medicamentos, como la hidroxicloroquina, son eficaces para tratar el COVID-19 , lo que ha provocado acaparamiento y escasez del medicamento, que en realidad no es eficaz para tratar el virus [13]. Los rumores sobre los orígenes y la propagación de la COVID-19 también han contribuido al panorama de desinformación que rodea a la enfermedad. Por ejemplo, algunos medios de comunicación y políticos han promovido la teoría de que el virus fue creado y propagado intencionalmente por China, a pesar de la falta de evidencia que respalde esta afirmación [10]. Además, la información errónea en torno a las vacunas incluye acusaciones infundadas de efectos adversos, dudas sobre su eficacia y teorías infundadas sobre agendas secretas. Esta información errónea socava la confianza del público en las vacunas, lo que genera dudas sobre las vacunas y pone en peligro el logro de una cobertura de vacunación generalizada. La politización de la pandemia también contribuyó a la difusión de información errónea relacionada con la COVID-19. Los líderes políticos y los medios de comunicación a menudo han adoptado posturas polarizadoras sobre la pandemia, lo que ha llevado a la amplificación de la desinformación y las teorías de conspiración [14]. Esto provocó una falta de consenso sobre la gravedad de la pandemia y la eficacia de las medidas de salud pública, lo que contribuyó aún más a la difusión de información errónea.

Los modelos matemáticos se han utilizado ampliamente para comprender la propagación de COVID-19 e informar las intervenciones de salud pública [15]. Varios de estos modelos suelen implicar un conjunto de ecuaciones diferenciales que describen la dinámica de la propagación del virus, incluido el número de personas infectadas, la tasa de transmisión y la eficacia de intervenciones como el distanciamiento social y la vacunación [16,17]. Se pueden utilizar modelos matemáticos para estimar parámetros clave de la pandemia, como el número de reproducción básico (R0), que mide el número promedio de personas infectadas por cada individuo infectado en una población susceptible [18]. Al estimar R0, se pueden utilizar modelos matemáticos para predecir el impacto potencial de diferentes intervenciones y guiar las políticas de salud pública.

Hay muchos desafíos asociados con el modelado de COVID-19, incluida la incertidumbre de parámetros clave, como la tasa de transmisión, la duración de la infecciosidad y la duración de la inmunidad después de la recuperación o la vacunación [19,20]. Además, la compleja dinámica de la pandemia, incluida la aparición de nuevas variantes y cambios en el comportamiento del público debido a la desinformación, dificultó predecir con precisión el curso de la pandemia. En este estudio, presentaremos un modelo matemático del papel de la información engañosa en la dinámica de transmisión de COVID-19, ya que es probable que las personas mal informadas ignoren medidas como el distanciamiento social, el enmascaramiento, las vacunas y otras, implementadas para salvaguardar a la población. . El modelo se basa en el supuesto de que una proporción de la población es vulnerable a la desinformación. Nos referimos a la desinformación como información falsa o inexacta que se difunde, ya sea intencionalmente o no, dando lugar a una distorsión de los hechos o una mala comprensión de la realidad. Puede difundirse a través de diversos canales, incluidos los medios sociales, los medios de comunicación, el boca a boca o las plataformas en línea. La información errónea puede adoptar diferentes formas, como declaraciones engañosas, datos fabricados, teorías de conspiración o interpretaciones erróneas de información legítima. A menudo pretende engañar o manipular a las personas, lo que tiene consecuencias perjudiciales, incluida la erosión de la confianza, la promoción del miedo y el pánico, el obstáculo a la toma de decisiones informadas y efectos potencialmente perjudiciales para la salud pública, la cohesión social y la democracia. procesos.

Machine Translated by Google

El propósito de este estudio es evaluar el impacto que tiene la cantidad de información errónea circulante tiene sobre la dinámica de transmisión de COVID-19 en una población. para lograr Para lograr este objetivo, formulamos un modelo tipo SIR con un compartimento adicional que representa personas que han sido engañadas a través de información errónea. porque abundaba información errónea (intencionada o no) asociada con COVID-19, el modelo presentado también trata la cantidad de información errónea como un estado denominado M basado en la suposición de que es sintetizada por individuos engañosos. Este enfoque es similar al objetivo modelos celulares donde los viriones son producidos por células infectadas [21]. Las soluciones de equilibrio de el modelo se obtiene utilizando parámetros de umbral calculados y su estabilidad establecida tanto analítica como numéricamente. Las simulaciones del modelo indican que la COVID- 19 afectar significativamente negativamente a una población si la desinformación no se aborda adecuadamente.

2. Modelo Matemático

El modelo que se presenta a continuación es un modelo SIR por naturaleza con compartimentos adicionales. de personas mal informadas, muertes y la densidad de información engañosa que circula. Suponemos que a medida que la población susceptible (S) entra en contacto con virus falsos circulantes o información inexacta (M) que se difunde entre una comunidad, algunos serán engañados en un tasa de ρ y pasar al compartimento mal informado (Sm). Independientemente del medio a través del cual se difunde, suponemos que en cualquier momento (t), circulando La información errónea (M) es la suma total de todas las formas, incluidas declaraciones engañosas, datos inventados, teorías de conspiración o interpretaciones erróneas de información legítima. asumimos que la desinformación es generada a un ritmo μ0 por personas que ya están mal informadas o malos actores a los que también catalogamos como mal informados, y suponemos que los engañosos la densidad de información decae a razón de μ1. Por lo tanto, el modelo supone que La información sólo la generan los individuos del compartimento Sm . También suponemos que personas engañadas a través de campañas de concientización u otras fuentes rechazan información falsa recuperarse a una velocidad κ y regresar al compartimento susceptible. Interacciones entre individuos susceptibles (S) y mal informados (Sm) con individuos infectados por COVID-19 (I) conducen a infecciones, respectivamente, a tasas α y σ. Supusimos que los individuos infectados se recuperó (R) a una tasa β o murió (D) por una complicación relacionada con COVID-19 en el tasa γ. Se supone además que los nuevos individuos son reclutados sólo como susceptibles en algún momento. tasa Λ y que la tasa de muerte natural en todos los compartimentos es θ. Un diagrama esquemático de la El modelo se proporciona en la Figura 1, y un resumen de los parámetros y su descripción se proporciona en Tabla 1.

Teniendo en cuenta los supuestos establecidos anteriormente e ilustrados en el esquema diagrama, el sistema de ecuaciones diferenciales que gobierna el modelo es el siguiente:

$$\frac{dS}{dt} = \Lambda + \kappa Sm - \alpha SI - \rho SM - \theta S \quad (1)$$

$$\frac{dSm}{dt} = \rho SM - \sigma Sm\, I - \kappa Sm - \theta Sm \quad (2)$$

$$\frac{dI}{dt} = \alpha SI + \sigma Sm\, I - \beta I - \gamma I - \theta I \quad (3)$$

$$\frac{dR}{dt} = \beta I - \theta R \quad (4)$$

$$\frac{dd}{dt} = \gamma I \quad (5)$$

$$\frac{dM}{dt} = \mu 0 Sm - \mu 1 M, \quad (6)$$

donde hemos asumido que todos los valores de los parámetros son positivos y las condiciones iniciales $S(0) = S0$, $Sm(0) = Sm0$, $I(0) = I0$, $R(0) = R0$, $D(0) = D0$ y $M(0) = M0$ son todos no negativo. Observamos que al inicio del brote, $Sm(0) \geq 0$ y $R(0) = D(0) = 0$.

Machine Translated by Google

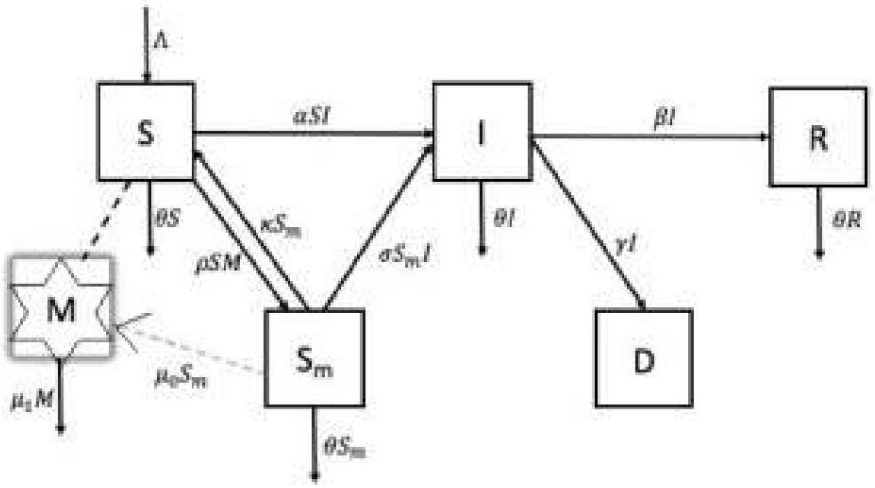


Figura 1. Diagrama esquemático de transmisión de COVID-19 con información errónea. Transición de individuos de los compartimentos S y Sm al compartimento I como resultado de su interacción con individuos en compartimento I. Los individuos hacen la transición entre los compartimentos S y Sm como resultado de la interacción con desinformar a M (ya sea consumiéndolo o recuperándose de él). La densidad de desinformación se produce por individuos en el compartimento Sm.

Tabla 1. Descripción de los parámetros del modelo.

Parámetro	Descripción
Λ	Tasa de reclutamiento de susceptibles
α	Tasa de infección de los susceptibles
β	Tasa a la que las infecciones se recuperan
ρ	Tasa de engaño a los susceptibles
κ	Tasa de recuperación de los engañados de la desinformación
σ	Tasa de contagio de los desinformados
γ	Tasa de mortalidad de enfermedades infecciosas.
θ	Tasa de mortalidad natural
μ0	Tasa de crecimiento de la información engañosa
μ1	Tasa de decadencia de información engañosa

Antes de continuar, observamos que la ecuación de muertes por COVID-19 (D), es decir La ecuación (5) está desacoplada del resto del sistema y, por lo tanto, se despreciará en todos los análisis posteriores. Comenzamos explorando las características dinámicas básicas del modelo. En particular, establecemos la positividad de las soluciones del modelo y un espacio de invariancia. El siguiente resultado muestra que el modelo siempre admite soluciones positivas.

Teorema 1. Sean las condiciones iniciales del modelo S(0) > 0, Sm(0) ≥ 0, I(0) ≥ 0, R(0) ≥ 0, y M(0) ≥ 0. Entonces, las soluciones (S(t), Sm(t), I(t), R(t), M(t)) del modelo con valores positivos condiciones iniciales, permanecerá positivo durante todo el tiempo t > 0.

Prueba. Comenzamos dividiendo las ecuaciones (1) a (4) y (6), respectivamente, por las variables S, Sm, I, R y M. Integrando las ecuaciones diferenciales resultantes en el intervalo de tiempo [0, t], y aplicando las condiciones iniciales obtenemos las siguientes expresiones:

$$S(t) = S0 \exp - \alpha I(\eta) - \frac{\Lambda}{S(\eta)} - \rho M(\eta) + \kappa \frac{Sm(\eta)}{S(\eta)} - \theta \, d\eta,$$

$$Sm(t) = Sm0 \exp \int_0^t \rho \frac{S(\eta)M(\eta)}{} - \sigma I(\eta) - \kappa - \theta \, d\eta,$$

$$I(t) = I0 \exp \int_0^t Sm(\eta)\,(\alpha S(\eta) + \sigma Sm(\eta) - \beta - \gamma - \theta) d\eta,$$

$$R(t) = R0 \exp \int_0^t \frac{yo(\eta)}{R(\eta)} - \theta \, d\eta, \text{ y } M(t) = M0 \exp \beta \int_0^t \mu 0 \frac{Sm(\eta)}{METRO(\eta)} - \mu 1 \, d\eta.$$

Machine Translated by Google

Dado que las condiciones iniciales no son negativas, es obvio a partir de las expresiones anteriores que S(t), Sm(t), I(t), R(t) y M(t) siempre son positivas para todos los tiempos donde t > 0. Por tanto, la solución del modelo con la condición inicial establecida es siempre positiva. □

A continuación demostramos que la solución del modelo es acotada. Con la condición dada, consideramos la población viva (L) definida por L(t) = S(t) + Sm(t) + I(t) + R(t) y demostramos que L(t) y M(t) están acotados en el siguiente resultado.

Teorema 2. Si las condiciones iniciales del modelo satisfacen L(0) ≤ la $\frac{\Lambda}{\theta}$ y M(0) ≤ $\frac{\mu 0 \Lambda}{\theta \mu 1}$, entonces el solución del modelo es acotada.

Prueba. Dado que S(t) > 0, Sm(t) ≥ 0, I(t) ≥ 0 y R(t) ≥ 0 del Teorema 1, es obvio que L(t) está acotado desde abajo y, por lo tanto, L (t) ≥ 0. Sólo queda establecer que L(t) está acotado desde arriba. Derivando L(t) = S(t) + Sm(t) + I(t) + R(t) y sustituyendo las derivadas del modelo, obtenemos la ecuación diferencial

$$\frac{dL}{dt} = \Lambda - \theta(S + Sm + I + R) - \gamma I.$$

de donde obtenemos la desigualdad diferencial

$$\frac{dL}{dt} \leq \Lambda - \theta L$$

cuya solución satisface

$$L(t) \leq \frac{\Lambda}{\theta} - \left(\frac{\Lambda}{\theta} - L(0)\right)e^{-\theta t}.$$

De la desigualdad anterior se desprende claramente que la población viva del sistema siempre está limitada desde arriba. Además, observamos de la desigualdad que si la condición inicial L(0) ≤ Para concluir que el modelo es $\frac{\Lambda}{\theta}$, entonces tenemos que L(t) ≤ $\frac{\Lambda}{\theta}$ ya que t → ∞ es un límite superior estricto.

acotado, también establecimos que la densidad de información errónea M(t) está acotada. Ya mostramos en el Teorema 1 que M(t) es siempre positivo, es decir, M(t) ≥ 0, por lo que está acotado desde abajo. Para la acotación desde arriba, consideramos el caso en el que L(0) ≤ θ . Como L(t) ≤ debemos tener que Sm(t) ≤ θ . Usando la ecuación diferencial para la densidad de desinformación, obtenemos la desigualdad diferencial

$$\frac{dM}{dt} \mu 0\Lambda \leq -\mu 1 M \theta$$

cuya solución M(t) satisface

$$M(t) \leq -\left(\frac{\mu 0\Lambda - \mu 1\theta}{\mu 1\theta} - M(0)\right)e^{-\mu 1 t}.$$

Claramente, la densidad de la información está limitada desde arriba. De manera similar, si M(0) ≤ $\frac{\mu 0 \Lambda}{\theta \mu 1}$, entonces tenemos que M(t) ≤ $\frac{\mu 0 \Lambda}{\theta \mu 1}$ como límite superior estricto. □

Lema 1. El conjunto cerrado

$$\Pi = \{(S(t), Sm(t), I(t), R(t), M(t)) \in \mathbb{R}^5_+ : L(t) \leq \frac{\Lambda}{\theta}, M(t) \leq \frac{\mu 0 \Lambda}{\theta \mu 1}\}$$

es positivamente invariante y atractivo para el modelo.

Prueba. La prueba se deriva directamente de los teoremas 1 y 2 , donde se supone la existencia trivial y la unicidad de la solución. □

Machine Translated by Google

3. Análisis del modelo En

esta sección calculamos los equilibrios y proporcionamos las condiciones para su existencia. También proporcionamos el número de reproducción del modelo y lo utilizamos para estudiar la estabilidad de los equilibrios calculados. Tenga en cuenta también que la ecuación para los recuperados (R) está desacoplada del resto del sistema, además de la de las muertes (D). Por lo tanto, ignoraremos estas ecuaciones en el análisis posterior dado que su comportamiento puede deducirse fácilmente dado el comportamiento del subsistema restante.

3.1. Existencia de

equilibrios Al desacoplar las ecuaciones de R y D, los equilibrios del modelo vienen dados por la solución del subsistema

$$\begin{aligned} &\Lambda + \kappa S_{metro} - \alpha S\, I - \rho S\, M - \theta S = 0\\ &\rho S\, M - \sigma S_{metro} I - \kappa S_{metro} - \theta S_{metro} = 0\\ &\alpha S\, I + \sigma S_{metro} I - \beta I - \gamma I - \theta I = 0\\ &\mu 0 S_{metro} - \mu 1 M = 0. \end{aligned}$$

Dado que la última ecuación anterior solo depende de la variable $S_{metro,}\ M$ se puede representar como

$$M = \frac{\mu 0 S_{metro}}{\mu 1}.$$

De la tercera ecuación del subsistema, vemos que I = 0, lo que conduce al equilibrio libre de enfermedades (DFE) del modelo, o αS, que conduce al equilibrio endémico. $+ \sigma S_{metro} - \beta - \gamma - \theta = 0$,

Supongamos que E0 = (S , $S_{metro,}$ I , M) denota un DFE, luego un cálculo adicional produce dos EFD,

$$mi_0^1 = \frac{\Lambda,}{0,\ 0,\ 0\ y\ E\ 0\ \theta} \qquad ^2 = \frac{(\theta + \kappa)\mu 1}{\rho\mu 0}, \frac{\Lambda}{\theta} - \frac{(\theta + \kappa)\mu 1\ \Lambda\mu 0\ ,\ 0,}{\theta\mu 1\ \rho\mu 0} - \frac{(\theta + \kappa)\ \rho}{}\ .$$

Observamos aquí que el segundo DFE E 0 ^{2} sólo existirá si $\frac{\Lambda}{\theta} > \frac{(\theta+\kappa)\mu 1}{\rho\mu 0}$.

Antes de continuar con el análisis, hacemos una pausa y calculamos el número de reproducción, R0, del modelo. El número de reproducción es el número de infecciones secundarias que ocurren cuando una persona infecciosa se coloca en una población completamente susceptible. Usando el método matricial de próxima generación (ver por ejemplo Duatel et al. [22]) y los DFE, E y 2 E 0 , respectivamente, obtenemos los números de reproducción. $^1_0$

$$R_0^1 = \frac{\alpha\Lambda}{\theta(\beta + \gamma + \theta)} \quad y\ r \qquad ^2_0 = \frac{1}{\beta + \gamma + \theta}\ \rho\mu 0\frac{(\theta + \kappa)\mu 1\ \Lambda\alpha(\ ) + \sigma(\ \theta}{} - \frac{(\theta + \kappa)\mu 1\ )}{\rho\mu 0}\ .$$

Observamos aquí que R 0 ^{1} regula el modelo en ausencia de información errónea, mientras que R_0^2 impulsa el modelo cuando la información errónea es persistente.

Ahora centramos nuestra atención en establecer los equilibrios endémicos del modelo. Supongamos + que EE = (S , $S_{metro,}$ I , M) denota un equilibrio endémico, entonces usando la ecuación αS $\sigma S_{metro} - \beta - \gamma - \theta = 0$ y las ecuaciones necesarias en el subsistema, obtenemos dos equilibrios endémicos, a saber,

$$EE1 = \frac{A\ \theta\ ,\ 0,}{\alpha} \quad \frac{}{\alpha}(R_0^1 - 1),\ 0$$

y

$$EE2 = \frac{(\theta + \kappa + \sigma I\ )\mu 1}{\rho\mu 0}, \frac{A}{\sigma} - \frac{\alpha(\theta + \kappa + \sigma I\ )\mu 1}{\sigma\rho\mu 0}, I, \frac{\mu 0}{\mu 1}\ \frac{A}{\sigma} - \frac{\alpha(\theta + \kappa + \sigma I\ )\mu 1}{\sigma\rho\mu 0}\ ,$$

Machine Translated by Google

donde A = β + γ + θ y donde

$$I = (R_0^2 - 1)\frac{\theta A}{\sigma}\left[A + \left(\frac{\theta\alpha\mu 1}{\rho\mu 0}\right)\frac{\sigma}{\alpha} - 1\right]^{-1}.$$

También observamos aquí que el primer equilibrio endémico EE1 existe si $R_0^1 > 1$ y $A - \frac{\alpha(\theta+\kappa+\sigma I)}{\rho\mu 0}\mu 1 > 0$, y el segundo equilibrio endémico EE2 existe si $R_0^1 > 1$ y $A - \frac{\alpha(\theta+\kappa+\sigma I)}{\rho\mu 0}\mu 1 > 0$.

3.2. Estabilidad de los

equilibrios En esta sección, establecemos los resultados de estabilidad local para las EFD y los equilibrios endémicos obtenidos anteriormente en condiciones apropiadas. Comenzamos observando que la matriz jacobiana, J(S, Sm, I, M), del subsistema está dada por

$$J(\cdot) = \begin{matrix} -\alpha I - \rho M - \theta & k & -\alpha S & -\rho S \\ \alpha I & -\sigma I - \kappa - \theta & -\sigma Sm & \rho S \\ 0 & \sigma yo & \alpha S + \sigma Sm - \beta - \gamma - \theta & 0 \\ & \mu 0 & & -\mu 1 \end{matrix}.$$

Teorema 3. Si $\mu 1\theta(\theta + \kappa) > \mu 0\rho\Lambda$ y R en caso $R_0^1 < 1$, entonces el DFE E_0^1 es localmente asintóticamente estable, contrario es inestable.

Prueba. Evaluación de la matriz jacobiana en el DFE $E_0^1 = \left(\frac{\Lambda}{\theta}, 0, 0, 0\right)$, obtenemos

$$J(E_0^1) = \begin{matrix} -\theta & k & -\alpha\theta\frac{\Lambda}{\theta} & -\rho\frac{\Lambda}{\theta} \\ 0 & -\kappa - \theta & 0 & \rho\frac{\theta}{\Lambda\theta} \\ 0 & 0 & \alpha\frac{\Lambda}{\theta} - \beta - \gamma - \theta & 0 \\ 0 & \mu 0 & & -\mu 1 \end{matrix}$$

cuyo polinomio característico es

$$P(\lambda) = \det(J(E_0^1) - \lambda I)$$
$$= (\alpha - \beta - \gamma - \theta - \lambda)[(-\theta - \lambda)(-\theta - \kappa - \lambda)(-\mu 1 - \lambda) - \mu 0(\rho)(-\theta - \lambda)],$$

obteniendo los valores propios

$$\lambda 1 = -\theta$$
$$\lambda 2 = \alpha - \beta - \gamma - \theta$$
$$\lambda 3 = \frac{-(\theta + \kappa + \mu 1) - \sqrt{(\mu 1 + \theta + \kappa)^2 - 4((\theta + \kappa)\mu 1 - \mu 0\rho\frac{\Lambda}{\theta})}}{2}$$
$$\lambda 4 = \frac{-(\theta + \kappa + \mu 1) + \sqrt{(\mu 1 + \theta + \kappa)^2 - 4((\theta + \kappa)\mu 1 - \mu 0\rho\frac{\Lambda}{\theta})}}{2}.$$

De lo anterior, es obvio que λ1 y λ3 tienen partes reales negativas, ya que $R_0^1 = \frac{\alpha\Lambda}{\theta(\beta+\gamma+\theta)}$ y λ2 se pueden reescribir como

$$\lambda 2 = \alpha\theta\frac{1}{\Lambda} - (\beta + \gamma + \theta)$$
$$= R_0(\beta + \gamma + \theta) - (\beta + \gamma + \theta)$$
$$= (R_0^1 - 1)(\beta + \gamma + \theta).$$

Machine Translated by Google

Por lo tanto, si $R_0^1 < 1$, entonces $\lambda 2 < 0$. $\lambda 4$ tiene una parte real negativa si

$$\theta + \kappa + \mu 1 > \sqrt{(\mu 1 + \theta + \kappa)^2 - 4((\theta + \kappa)\mu 1 - \mu 0\rho(\frac{\Lambda}{\theta})}$$

$$(\theta + \kappa + \mu 1)^2 > (\mu 1 + \theta + \kappa)^2 - 4((\theta + \kappa)\mu 1 - \mu 0\rho(\frac{\Lambda}{\theta})$$

$$0 > -4((\theta + \kappa)\mu 1 - \mu 0\rho(\frac{\Lambda}{\theta})$$

$$\mu 1\theta(\theta + \kappa) > \mu 0\rho\Lambda$$

es cierto. Por lo tanto, si $\mu 1\theta(\theta + \kappa) > \mu 0\rho\Lambda$ y R es asintóticamente $R_0^1 < 1$, entonces el DFE E_0^1 es localmente estable, en caso contrario es inestable si $R > 1$. R_0^1 □

Concluimos ahora nuestro análisis del primer DFE E_0^1 estableciendo un resultado de estabilidad global. Para este propósito, recordamos establecer en la Sección 2 que en cualquier momento t, la población viva $L(t) = S(t) + Sm(t) + I(t) + R(t)$ está acotada arriba y satisface

$$L(t) \le \frac{\Lambda}{\theta} - (\frac{\Lambda}{\theta} - L0)e^{-\theta t},$$

donde $L0 = L(0)$. Los siguientes resultados son consecuencia de la desigualdad anterior.

Lema 2. La población viva $L(t) \le L0$ en todo momento t.

El siguiente resultado proporciona estabilidad global para el primer DFE E sin $E_{0,}^1$ es decir, cuando hay COVID-19 y tampoco hay desinformación.

Teorema 4. En vista del Lema 2, si $L0 = 1$ E 0 $\frac{\alpha}{\theta}\frac{\Lambda}{,\sigma}$ y si $r_0^1 < 1$, entonces el primer equilibrio libre de enfermedad, es globalmente asintóticamente estable; de lo contrario, es inestable.

Prueba. Para demostrar el Teorema 4, consideramos la función candidata de Lyapunov $V = \frac{1}{y_o}\frac{2}{2}$. Tomando su derivada obtenemos,

$$\frac{dV}{dt} = \frac{dil}{quiero}^{= yo}$$

$$= (\alpha SI + \sigma Sm\, I - \beta I - \gamma I - \theta I)I = [\alpha S + \sigma Sm - (\beta + \gamma + \theta)]I$$

$$\alpha S + \sigma Sm = (\frac{(\beta + \gamma + \theta)(}{\beta + \gamma + \theta} - 1)I^2.$$

Dado que el modelo supone que $\alpha < \sigma$, entonces la ecuación anterior se convierte en

$$\frac{dV}{dt} \le (\beta + \gamma + \theta)(\frac{\sigma(S + Sm)}{\beta + \gamma + \theta} - 1)I^2.$$

Si $L0 = \frac{\alpha}{\sigma}\frac{\Lambda}{\theta}$ luego por el Lema 2, obtenemos que $S + Sm \le \frac{\alpha}{\sigma}\frac{\Lambda}{\theta}$. Así, obtenemos la desigualdad diferencial

$$\frac{dV}{dt} \le (\beta + \gamma + \theta)(\frac{\alpha\Lambda}{\theta(\beta + \gamma + \theta)} - 1)I^2$$

$$\le (\beta + \gamma + \theta)(R_0^1 - 1)y_o^2.$$

Está claro que si $R_0^1 < 1$, entonces $\frac{dV}{dt} \le 0$ y por lo tanto, por el principio de invariancia de LaSalle el equilibrio libre de enfermedad E_0^1 es globalmente asintóticamente estable [23]. □

A continuación, consideramos la estabilidad local del segundo DFE E_0^2 y enuncie el siguiente resultado.

Machine Translated by Google

Teorema 5. Si ρΛμ0 > θ(θ + κ)μ1 y R en caso contrario $_0^2$ < 1, entonces el DFE E $_0^2$ es localmente asintóticamente estable, es inestable.

Prueba. Evaluando la matriz jacobiana en el segundo DFE, es decir, E (θ+κ)μ1 (θ+κ) Λμ0), obtenemos la matriz $_0^2$ = ((θ+κ)μ1 ρμ0 , Λ θ −

jacobiana , 0, θμ1 ρμ0 ρ

$$J(E)^{0}_{2}) = \begin{matrix} -\frac{\rho\Lambda\mu 0}{\theta\mu 1}+\kappa & k & -\alpha\frac{(\theta+\kappa)\mu 1}{\rho\mu 0} & -\rho\frac{(\theta+\kappa)\mu 1}{\rho\mu 0} \\ \rho\Lambda\mu 0-\theta-\kappa & -\kappa-\theta & -\sigma(\frac{\Lambda}{\theta}-\frac{(\theta+\kappa)\mu 1}{\rho\mu 0}) & \frac{(\theta+\kappa)\mu 1\,\rho}{\rho\mu 0} \\ 0 & 0 & \alpha+\frac{(\theta+\kappa)\mu 1}{\rho\mu 0}+\sigma(\frac{\theta}{\Lambda\theta}-\frac{(\theta+\kappa)\mu 1}{\rho\mu 0}) & 0 \\ 0 & \mu 0 & -\beta-\gamma-\theta & -\mu 1 \end{matrix}.$$

Claramente, vemos en la tercera fila de la matriz anterior que

$$\rho\mu 0\,{}^{2}_{0}\;\frac{\theta\mu 1\,\Lambda}{\theta\mu 1}\,\lambda 1=\alpha+\sigma(\frac{\theta}{\Lambda}-\frac{}{\theta\mu 1})-\beta-\gamma-\theta\,\rho\mu 0$$

$$= R_0(\beta+\gamma+\theta)-(\beta+\gamma+\theta)$$

$$= (R^{2}_{0}-1)(\beta+\gamma+\theta)<0 \text{ si } R^{2}_{0}<1.$$

El factor restante del polinomio característico es de grado tres y viene dado por

$$\lambda^{3}+b1\lambda^{2}+b2\lambda+b3=0,$$

dónde

$$b1=\mu 1++\theta\frac{\rho\Lambda\mu 0}{\theta\mu 1}-\rho\Lambda\mu 0$$

$$\theta\,\mu 1\;b3=\rho\Lambda\mu 0-\frac{\rho\Lambda\mu 0\;b2=+-\kappa\mu 1}{(\theta+\kappa)\mu 1\theta}.$$

Para establecer que los valores propios restantes (es decir, los ceros de la ecuación cúbica anterior) tienen partes reales negativas, utilizamos el criterio de Routh-Hurwitz. Es obvio a partir de los coeficientes b1, b2 y b3 anteriores que si se satisface la condición ρΛμ0 > (θ + κ)μ1θ, entonces b1 > 0, b2 > 0 y b3 > 0. Además, también vemos que la desigualdad

$$(-\kappa\mu 1)>\rho\Lambda\mu 0-\frac{\rho\Lambda\mu 0\;\rho\Lambda\mu 0\;b1b2=(\mu 1++\theta)}{(\theta+\kappa)\mu 1\theta=b3+\theta\mu 1\;\theta\;\mu 1}$$

es cierto. Dado que b1, b2 y b3 son todos positivos y b1b2 > b3, aplicando el criterio de Routh-Hurwitz concluimos que los valores propios restantes tienen partes reales negativas. Por lo tanto, si R es localmente asintóticamente estable, en caso contrario, es inestable.

$_0^2$ < 1, el DFE, E $_0^2$, □

Observación 1. Sólo como recordatorio, recordamos aquí que la condición ρΛμ0 > (θ + κ)μ1θ se impuso para garantizar que el DFE E2 sea$_0$ válido.

Teorema 6. Si R es $_0^1$ > 1, entonces el Equilibrio Endémico, EE1, es localmente asintóticamente estable, contrario, es inestable.

Machine Translated by Google

Prueba. Evaluando la matriz jacobiana en el equilibrio endémico EE1, es decir, en EE1 = ($\frac{A}{\alpha}$, 0, $\frac{\theta}{\alpha}(R_0^1 - 1)$, 0) donde A = β + γ + θ, obtenemos la matriz jacobiana

$$J(EE1) = \begin{matrix} 1-\theta R_0 & k & -A-\rho-1)) \frac{A}{\alpha} \\ 0 & -\sigma(& \frac{\theta}{\alpha}(R_0^1 - \kappa 0 - 1)) 0 0 & -\mu 1 & Un\overline{\alpha} \\ 1\theta R & \sigma(\frac{\theta}{\alpha}(R^1 & & & 0 \\ 0 0 & 0 \mu 0 & & & \end{matrix}.$$

El polinomio característico es

$$P(\lambda) = \det(J(EE1) - \lambda I)$$
$$= (-\sigma(\frac{\theta}{\alpha}(R_0^1 - 1)) - \lambda)[(-\mu 1 - \lambda)(\theta R\, 0\lambda + \lambda^{1\ 2} + A\theta R\, 0^{\ 1})],$$

y los valores propios son

$$\theta\, \lambda 1 = -\sigma(\frac{}{\alpha}(R_0^1 - 1)) < 0 \text{ si } R_0^{\ 1} > 1$$
$$\lambda 2 = -\mu 1 < 0$$
$$\lambda 3 = \frac{-\theta R_0^{\ 1} - (\theta R_0^{\ 1})^2 - 4A\theta R\, 1_0}{2}$$
$$\lambda 4 = \frac{1-\theta R_0 + (\theta R_0^{\ 1})^2 - 4A\theta R\, 1_0}{2}.$$

De los resultados anteriores, los valores propios λ3 y λ4 claramente son siempre reales negativos números o tienen partes reales negativas, ya que $(\theta R_{10})^2 - 4A\theta R\, 1\, 0 \quad 1 < \theta R_0$. Por lo tanto, podemos > 1, Concluyo que si R 0 1 e inestable si R 0 entonces el equilibrio endémico, EE1, es localmente asintóticamente estable ≤ 1. □

Observación 2. Observamos aquí que de manera similar al Teorema 6, se puede establecer que si $R_0^2 > 1$, entonces el equilibrio endémico

$$EE2 = S\ , \frac{A-\alpha S}{\sigma}, I\ , \frac{\mu 0}{(\mu 1}\ \frac{A-\alpha S}{\sigma})\ ,$$

donde S = $\frac{(\theta+\kappa+\sigma I\ )\mu 1}{\rho\mu 0}$, I = $(R_0^{\ 2} - 1)\ \frac{\theta A}{\sigma}\ A + \frac{\theta\alpha\mu 1}{\rho\mu 0}\ (\overline{\sigma} - 1)^{-1}$, y A = β + γ + θ es localmente estable.

4. Resultados y Discusión

En esta sección presentamos simulaciones numéricas del modelo bajo diferentes circunstancias centrándonos en la población de Estados Unidos. Simulamos el número de infecciones para diferentes tasas relacionadas con la desinformación y, en algunos casos, el número de muertes acumuladas. Los valores de los parámetros utilizados en las simulaciones se eligieron en función de los valores de la literatura publicada sobre COVID-19 [20,24–26] cuando estaban disponibles y se supusieron otros valores de parámetros para explorar el comportamiento del modelo. Las condiciones iniciales se eligieron para reflejar la población de Estados Unidos de aproximadamente 335 millones estableciendo S(0) = 3,35 × 108 . Considerando que se estima que la población real de Estados Unidos oscila entre 331 y 341 millones, llevamos a cabo varias simulaciones con pequeñas perturbaciones en las condiciones iniciales y el comportamiento a largo plazo de las predicciones del modelo no cambió.

En la sección anterior establecimos varios resultados cualitativos relacionados con los equilibrios del modelo. Comenzamos confirmando estos resultados numéricamente en las Figuras 2 a 4.

Machine Translated by Google

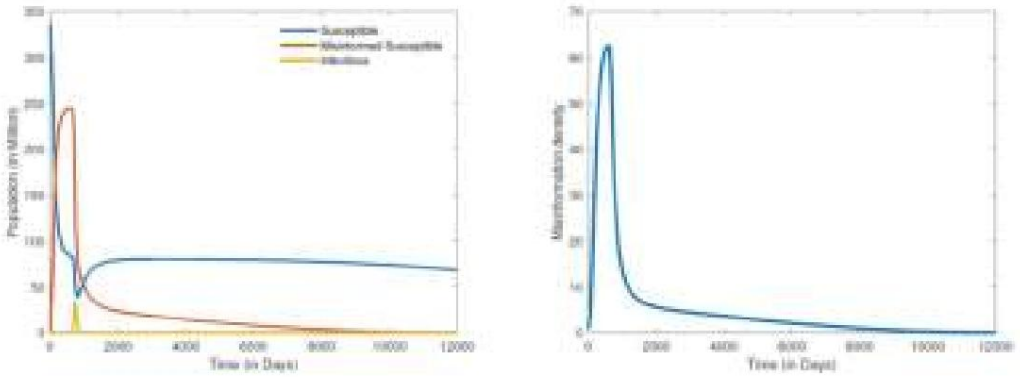


Figura 2. Simulación numérica de la primera DFE: distribución de poblaciones (izquierda) y densidad de información (derecha), para valores de parámetros Λ = 0, α = 0,25 × 10−4 , β = 1/14, ρ = 0,1, κ = 20ρ , σ = 0,5 × 10−3 , γ = 1,10−4 θ = 4 × 10−5 , μ0 = 0,0025 y μ1 = 0,01.

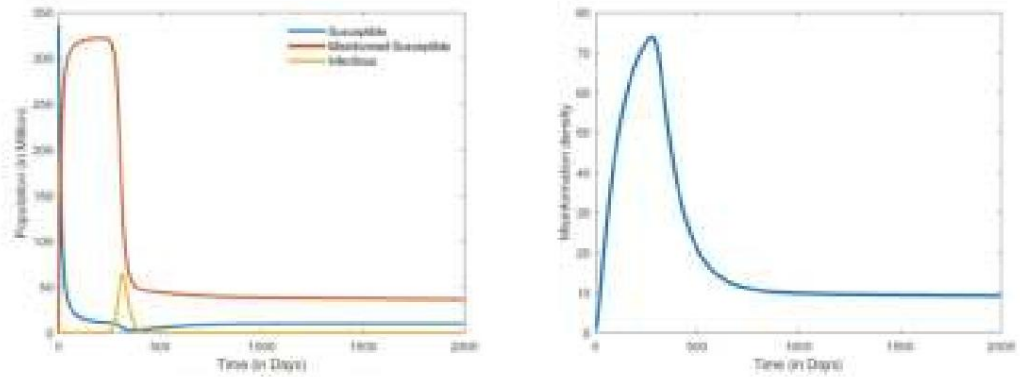


Figura 3. Simulación numérica del segundo DFE: distribución de poblaciones (izquierda) y densidad de información (derecha), para valores de parámetros Λ = 0, α = 0,25 × 10−4 , β = 1/14, ρ = 0,1, κ = 2,5 ρ, σ = 0,5 × 10−3 , γ =,10−4 θ = 4 × 10−5 , μ0 = 0,0025 y μ1 = 0,01.

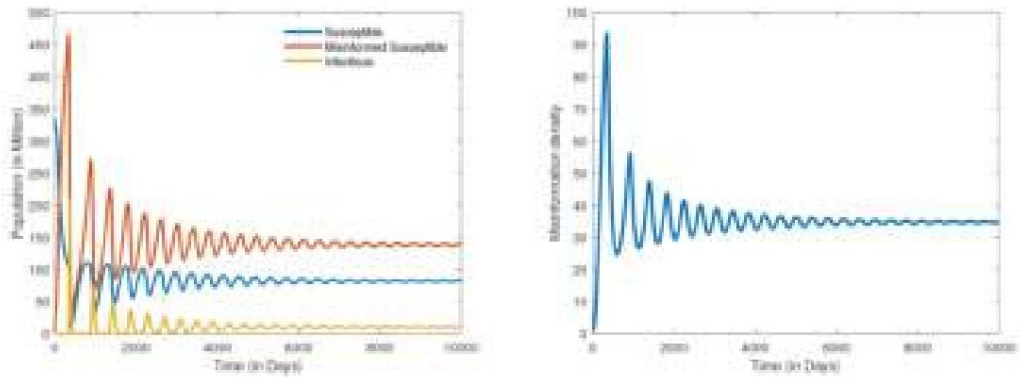


Figura 4. Simulación numérica de un equilibrio endémico estable: distribución de poblaciones (izquierda) y densidad de información (derecha), para valores de parámetros Λ = 0,75, α = 0,25 × 10−4 , β = 1/14, ρ = 0,1, κ = 20ρ, σ = 0,5 × 10−3 , γ = 10−4 θ = 4 × 10−5 μ0 = 0,0025 y μ1 = 0,01.

Los resultados de la Figura 2 demuestran la estabilidad global del primer DFE. Las simulaciones muestran que la enfermedad, los desinformados y la densidad de la desinformación acaban desapareciendo con el paso del tiempo. Es importante observar a partir de estos resultados que, si bien la COVID-19 puede desaparecer mucho antes, se necesita mucho más tiempo para erradicar tanto la desinformación como la densidad de información errónea. Observamos aquí que los cálculos numéricos para el primer DFE se obtuvieron estableciendo la tasa de recuperación de individuos mal informado

Machine Translated by Google

a aproximadamente 20 veces la tasa a la que se engaña a las personas susceptibles (ρ). Cuando la tasa de recuperación de individuos mal informados se establece en aproximadamente dos veces la tasa a la que los susceptibles son engañados, obtuvimos las simulaciones dadas en la Figura 3. Claramente, los resultados en la Figura 3 demuestran el segundo DFE que también establecimos en la sección anterior. .
Estos resultados muestran que incluso si la COVID-19 desapareciera por completo, la información errónea sobre la enfermedad nunca desaparecerá. Es probable que las teorías de conspiración como el origen de la COVID-19, que la enfermedad es un arma biológica, conceptos erróneos sobre las vacunas contra la COVID- 19 y quienes la niegan persistan durante mucho tiempo, dada la facilidad y velocidad con la que actualmente se difunde la información errónea.

Los resultados para los equilibrios libres de enfermedades discutidos anteriormente se obtuvieron asumiendo una población cerrada (Λ = 0), es decir, no ingresan nuevos individuos a la comunidad. Este es un escenario en el que hay un bloqueo estricto o una restricción de movimiento. De hecho, este fue el caso de las primeras variantes de COVID-19 , donde las ciudades/países pudieron frenar la propagación de la enfermedad mediante cierres estrictos; sin embargo, no se pudo contener la información errónea. Al cambiar el paradigma de una población cerrada a una abierta donde hay libertad de movimiento con personas entrando y saliendo de la comunidad, la dinámica cambia rápidamente hacia simulaciones de un equilibrio endémico. Los resultados de la Figura 4 muestran un equilibrio endémico globalmente estable. Recuerde que establecer la tasa de recuperación (κ) de personas mal informadas en aproximadamente 20 veces la tasa a la que las personas susceptibles están mal informadas (ρ) condujo a una DFE. Sin embargo, para una población abierta donde la tasa de reclutamiento Λ = 0,75 y κ = 20ρ, obtenemos un equilibrio endémico. Figura 4. Un aumento adicional en la tasa de recuperación de individuos mal informados, digamos en un múltiplo de 50 veces la tasa a la que las personas susceptibles son mal informado, es decir, κ = 50ρ, conduce a una bifurcación de hopf que vuelve inestable el equilibrio endémico como se demuestra en la Figura 5. Vale la pena señalar que la bifurcación aquí transforma un equilibrio estable en uno inestable que contiene una amplitud muy grande de oscilaciones. Por lo tanto, podemos inferir de los resultados de las Figuras 4 y 5 que no basta con refutar información engañosa mediante la verificación de hechos por parte de los medios. Es necesario refutar y educar correctamente al público sobre el COVID-19 al mismo tiempo.

La alfabetización mediática también es esencial para combatir la desinformación relacionada con la COVID-19. Una mayor alfabetización mediática puede ayudar a las personas a identificar y distinguir entre fuentes creíbles de información y desinformación. Un público educado que sea menos susceptible a ser engañado conducirá a mejores resultados en la erradicación de la COVID-19 o en el control de la propagación de la enfermedad en un estado endémico. Observamos aquí que todas las simulaciones que llevamos a cabo para una tasa de reclutamiento positiva, es decir Λ > 0 (independientemente de cuán pequeño sea el valor de Λ), condujo a un equilibrio endémico. Como consecuencia, considerando que vivimos en una población abierta y con libertad de movimiento, el COVID-19 se volverá endémico.

A continuación, centramos nuestra atención en simular el efecto que tendrá variar la tasa de recuperación de individuos mal informados (κ) sobre el número de casos de COVID-19 y muertes acumuladas. Los resultados para el número de casos de COVID-19 para una población cerrada se muestran en la Figura 6. Observamos aquí que un aumento en la tasa de recuperación de los mal informados (κ) conduce a menos casos de COVID-19. Cuanto mayor es la tasa de recuperación (κ), más tarda el número de casos en alcanzar su punto máximo, y la magnitud del pico también es menor. También observamos que debido a que tarda más en alcanzar su punto máximo dada una mayor tasa de recuperación (κ), se brinda un tiempo precioso que los funcionarios de salud pública pueden utilizar para implementar otras medidas a fin de reducir aún más el impacto de la enfermedad en toda la población. Los resultados de la Figura 7 muestran las muertes acumuladas por COVID-19 para diferentes tasas de recuperación mal informadas para una población cerrada. También es evidente a partir de estas simulaciones que una mayor tasa de recuperación κ conducirá a un menor número de muertes acumuladas por COVID-19 . Según la OMS [6], se notificaron alrededor de 1.127.252 muertes por COVID-19 en los Estados Unidos entre el 3 de enero de 2020 y el 21 de septiembre de 2023. Observamos en las simulaciones de muertes acumuladas en la Figura 7 que la tasa a la que las personas mal informadas se recuperan durante el Estados Unidos es menos de cinco veces la tasa de engaño a la gente. Teniendo en cuenta que las personas mal informadas contribuyen al crecimiento de la información errónea, la tasa de recuperación de personas engañadas en Estados Unidos parece ser pequeña.

Machine Translated by Google

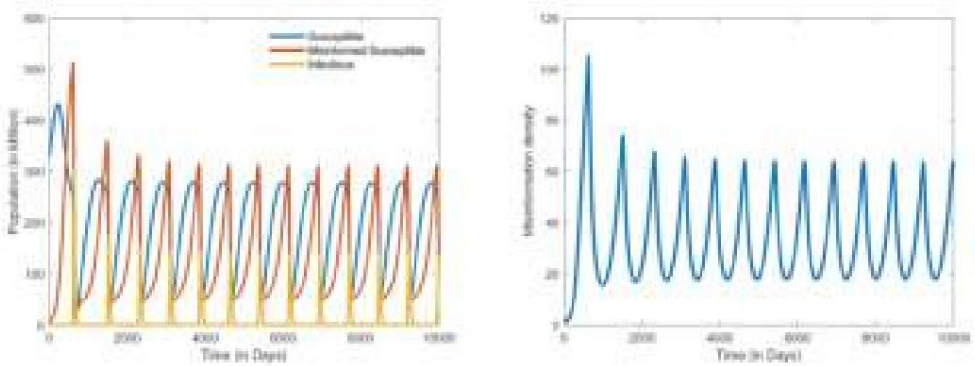


Figura 5. Simulación numérica de un equilibrio endémico inestable: distribución de poblaciones (izquierda) y densidad de información (derecha), para valores de parámetros Λ = 0,75, α = 0,25 × 10−4 , β = 1/14, ρ = 0,1, κ = 50ρ, σ = 0,5 × 10−3 , γ = 10−4 , θ = 4 × 10−5 , μ0 = 0,0025 y μ1 = 0,01.

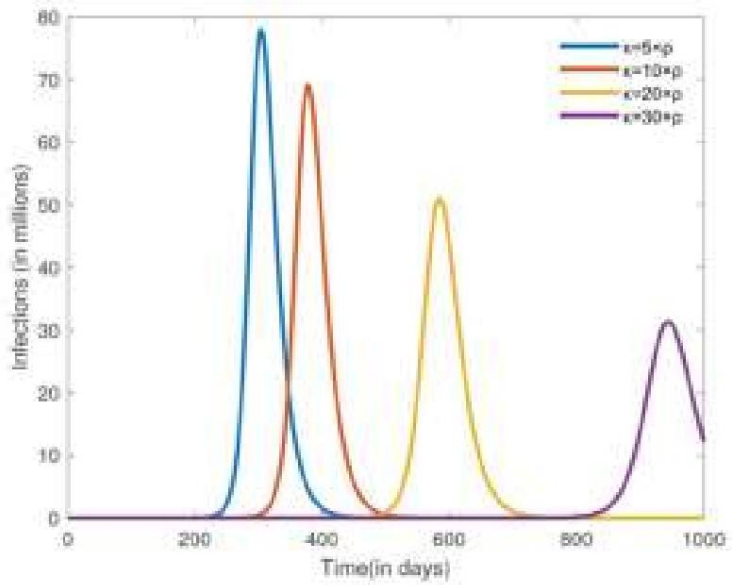


Figura 6. Simulación numérica del número de infecciones para diferentes tasas de recuperación mal informadas κ. Los otros valores de los parámetros se fijan en Λ = 0, α = 0,25 × 10−4 , β = 1/14, ρ = 0,1, σ = 0,5 × 10−3 , γ = 10−4 , θ = 4 × 10−5 , μ0 = 0,0025 y μ1 = 0,01.

Finalmente, consideramos los efectos del crecimiento de la densidad de información errónea y las tasas de disminución en el número de casos de COVID-19 , centrándonos en una población cerrada. Los resultados de la Figura 8 son para diferentes tasas de crecimiento μ0, mientras que los de la Figura 9 son para diferentes tasas de decaimiento μ1. Podemos ver claramente en los resultados de la Figura 8 que reducir la tasa de crecimiento de la densidad de información errónea conduce a menos casos de COVID-19. De hecho, mantener baja la tasa de crecimiento también conduce a un pico de infección más pequeño que ocurre más adelante, lo que brinda tiempo que puede usarse para implementar otras estrategias de contención y, en consecuencia, minimizar la carga de enfermedad. Promover información precisa es esencial para combatir la desinformación relacionada con el COVID-19. Los funcionarios de salud pública y las fuentes creíbles de información deben brindar mensajes claros y concisos sobre el virus y las medidas de salud pública, incluida Los mensajes positivos frenan la tasa de crecimiento de la desinformación. Por otro lado, aumentar la tasa de decadencia de la desinformación también conduce a un pico de infección más pequeño que ocurre mucho más tarde. Se pueden implementar mecanismos de verificación de datos para combatir la información errónea relacionada con el COVID-19. Las plataformas de televisión, radio y redes sociales pueden implementar mecanismos de verificación de datos para detectar y eliminar información errónea relacionada con la enfermedad.

Machine Translated by Google

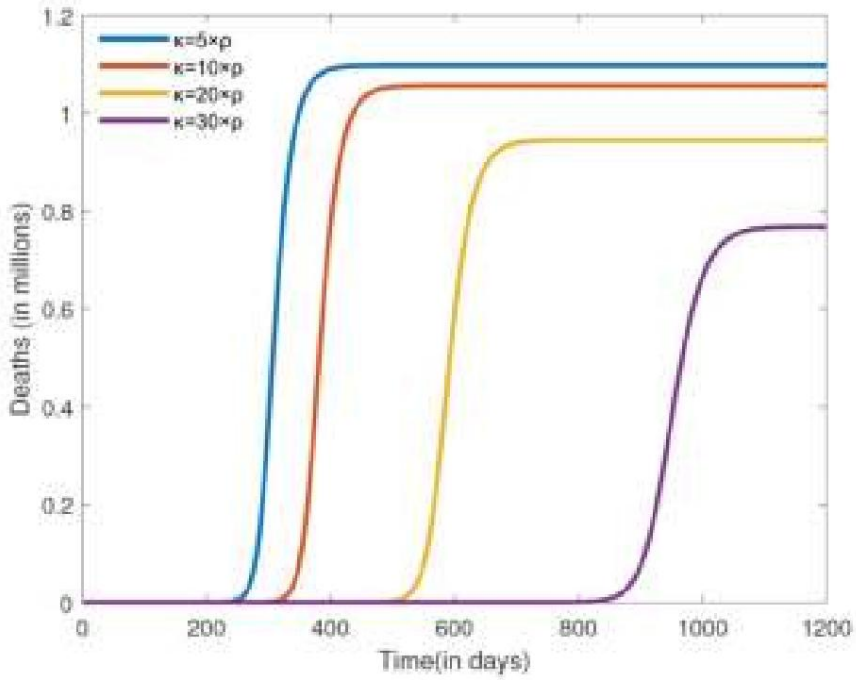


Figura 7. Muertes acumuladas de simulación numérica para diferentes tasas de recuperación κ. El otro parámetro los valores se fijan en Λ = 0, α = 0,25 × 10−4, β = 1/14, ρ = 0,1, σ = 0,5 × 10−3, γ = 10-4, θ = 4 × 10−5, μ0 = 0,0025 y μ1 = 0,01.

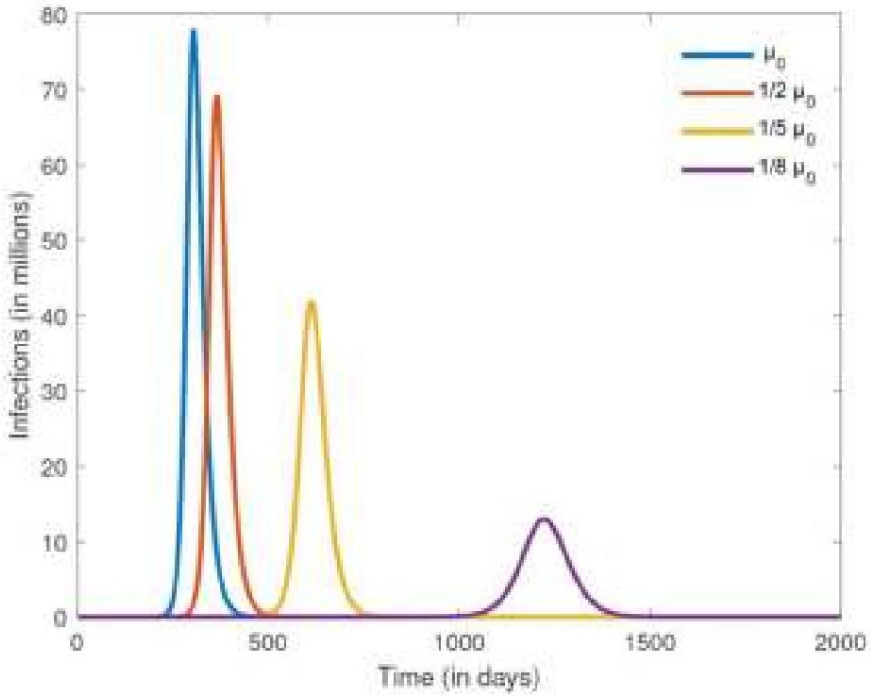


Figura 8. Simulación numérica del número de infecciones para diferentes crecimientos de desinformación tasas μ0. Los otros valores de los parámetros se fijan en Λ = 0, α = 0,25 × 10−4, β = 1/14, ρ = 0,1, κ = 5ρ, s = 0,5 × 10−3, γ = 10-4 , θ = 4 × 10−5, y μ1 = 0,01.

Machine Translated by Google

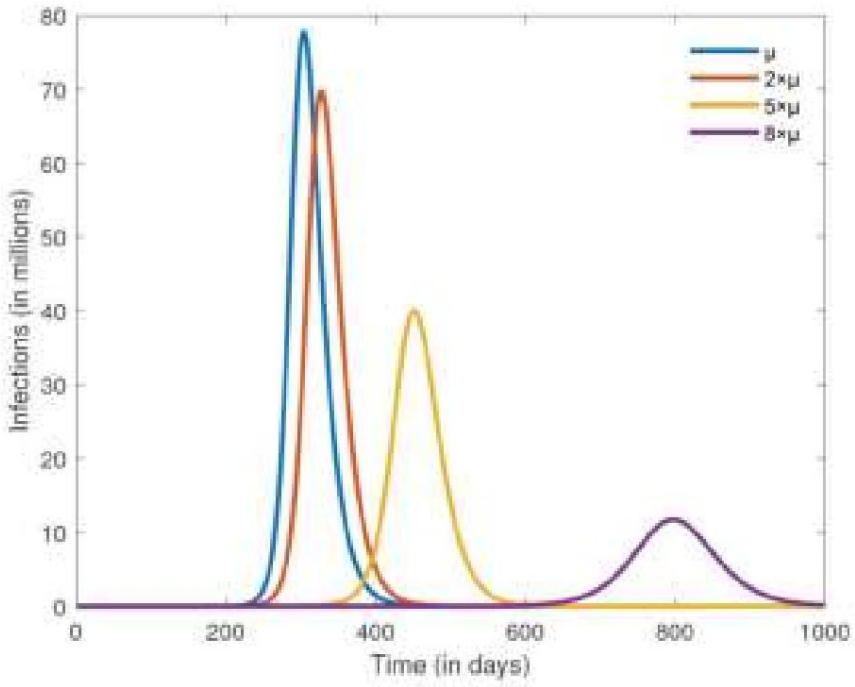


Figura 9. Simulación numérica del número de infecciones para diferentes tasas de disminución de la desinformación $\mu_1 = \mu$. Los otros valores de los parámetros se fijan en $\Lambda = 0$, $\alpha = 0{,}25 \times 10^{-4}$, $\beta = 1/14$, $\rho = 0{,}1$, $\kappa = 5\rho$, $s = 0{,}5 \times 10^{-3}$, $\gamma = 10\text{-}4$, $\theta = 4 \times 10^{-5}$, y $\mu_0 = 0{,}0025$.

5. Conclusiones

La pandemia de COVID-19 ha puesto en primer plano el tema de la desinformación, destacando el importante impacto en la salud pública. La desinformación puede socavar los esfuerzos para controlar la propagación del virus y, por lo tanto, tener una serie de consecuencias negativas, que incluyen una mayor renuencia a vacunarse, el incumplimiento de las medidas de salud pública y la desconfianza. de los funcionarios de salud pública. En este artículo, formulamos un modelo que captura la dinámica de transmisión de COVID-19 simultáneamente con la difusión de información errónea en un comunidad. El análisis del modelo muestra que tiene dos equilibrios libres de enfermedad, uno que contiene información errónea y el otro que no. El DFE sin desinformación rinde un número de reproducción R0 y es estable local y globalmente si R0 < 1. El DFE contener información errónea produce un parámetro umbral que hace que el equilibrio sea estable si el valor del parámetro es menor que uno.

Los factores que contribuyen a la difusión de información errónea relacionada con COVID-19 incluyen la rápida difusión de información a través de las redes sociales, el efecto cámara de eco y la Politización de la pandemia. Cuando el modelo se aplica a la población de los Estados Unidos Estados Unidos, las simulaciones de muertes acumuladas sugieren que Estados Unidos hizo un mal trabajo para combatir la desinformación relacionada con el COVID-19. Simulaciones numéricas del modelo. sugieren que se necesita un enfoque multifacético para combatir la difusión de información errónea y promover información veraz sobre la pandemia. Estrategias para combatir la desinformación relacionados con COVID-19 incluyen promover información precisa, aumentar la alfabetización mediática, e implementar mecanismos de verificación de hechos. Además, las simulaciones del modelo predicen que mientras las personas puedan entrar y salir de una comunidad, el virus COVID-19 siguen siendo endémicas en dicha comunidad.

En conclusión, hemos demostrado que el modelado matemático puede desempeñar un papel importante. para comprender la difusión de información errónea relacionada con COVID-19 y desarrollar estrategias efectivas para combatir la información errónea. Combatir la desinformación sobre la COVID-19 es esencial para promover la salud pública y controlar la propagación del virus. Además, es esencial abordar las causas subyacentes de la difusión de información errónea, así como promover información precisa, aumentar la alfabetización mediática e implementar mecanismos de verificación de hechos.

Machine Translated by Google

Contribuciones de los autores: Conceptualización, ZS y EA; metodología, ZS y EA; software, ZS y EA; validación, ZS y EA; análisis formal, ZS y EA; investigación, ZS y EA; recursos, ZS y EA; curación de datos, ZS y EA; redacción: preparación del borrador original, ZS y EA; redacción: revisión y edición, ZS y EA; visualización, ZS y EA; supervisión, EA; administración de proyectos, EA; adquisición de financiación, EA Todos los autores han leído y aceptado la versión publicada del manuscrito.

Financiamiento: Esta investigación no recibió financiamiento externo.

Declaración de la Junta de Revisión Institucional: No aplicable.

Declaración de Consentimiento Informado: No aplicable.

Declaración de disponibilidad de datos: los datos están contenidos en el artículo.

Conflictos de intereses: Los autores declaran no tener conflictos de intereses.

Referencias

1. Palabras de apertura del Director General de la OMS en la rueda de prensa sobre la COVID-19, 11 de marzo de 2020. Disponible en línea: https://www.who.int/director-general/speeches/detail/who-director-general-s- palabras-de-apertura-en-la- rueda-de-prensa-sobre- COVID-19---11-de-marzo-2020 (consultado el 1 de enero de 2023).
2. Cucinotta, D.; Vanelli, M. La OMS declara al COVID-19 una pandemia. Acta Biomed. 2020, 91, 157–160. [Referencia cruzada] [PubMed]
3. Centros para el Control y la Prevención de Enfermedades. Disponible en línea: https://www.cdc.gov/coronavirus/2019-ncov/science/science-resúmenes/sars-cov-2-transmission.html (consultado el 1 de enero de 2023).
4. Centros para el Control y la Prevención de Enfermedades. Disponible en línea: https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/síntomas.html (consultado el 1 de enero de 2023).
5. de Rosa Mesquita, R.; Francelino Silva Júnior, LC; Santos Santana, FM; Farias de Oliveira, T.; Campos Alcántara, R.; Monteiro Arnozo, G.; Rodrigues da Silva Filho, E.; Galdino dos Santos, AG; Oliveira da Cunha, EJ; Salgueiro de Aquino, SH; et al. Manifestaciones clínicas de COVID-19 en la población general: Revisión sistemática. Viena Klin Wochenschr. 2021, 133, 377–382. [Referencia cruzada] [PubMed]
6. Organización Mundial de la Salud. Quién Panel de control del coronavirus (COVID-19) . Disponible en línea: https://covid19.who.int/ (accedido en 1 de octubre de 2023).
7. Suárez-Lledo, V.; Álvarez-Gálvez, J. Prevalencia de información errónea sobre salud en las redes sociales: revisión sistemática. J. Med. Internet Res. 2021, 23, e17187. [Referencia cruzada] [PubMed]
8. Greene, CM; Murphy, G. Cuantificación de los efectos de las noticias falsas en el comportamiento: evidencia de un estudio sobre la desinformación sobre el COVID-19 . J. Exp. Psicólogo. Aplica. Adv. Publicación en línea. 2021, 27, 773. [Referencia cruzada]
9. Agley, J.; Xiao, Y. Información errónea sobre COVID-19: evidencia de perfiles latentes diferenciales y una fuerte asociación con la confianza en ciencia. BMC Public Health 2021, 21, 89. [CrossRef]
10. Lewowsky, S.; Jacobs, P.; Neil, S. La hipótesis de la fuga de laboratorio hizo que a los científicos les resultara más difícil buscar la verdad. Ciencia. Soy. 2022, 326, 72–77.
11. Flaherty, E.; Sturm, T.; Farries, E. La conspiración de COVID-19 y 5G: falacias del análisis espacial en la era de la democratización de los datos. Soc. Ciencia. Medicina. 2022, 293, 114546. [Referencia cruzada] [PubMed]
12. Litman, L.; Rosen, Z.; Hartman, R.; Rosenzweig, C.; Weinberger-Litman, SL; musgo, AJ; Robinson, J. ¿La gente realmente bebió lejía para prevenir el COVID-19? Una guía para proteger los datos de las encuestas contra encuestados problemáticos. PLoS ONE 2023, 18, e0287837 [CrossRef] [PubMed]
13. Comán, C.; Bularca, MC; Repanovici, A.; Rogozea, L. Desinformación sobre los medicamentos durante la pandemia de COVID-19: una perspectiva del personal médico. MÁS UNO 2022, 17, e0276693. [Referencia cruzada] [PubMed]
14. Hart, PS; Chinn, S.; Soroka, S. Politización y polarización en la cobertura noticiosa de COVID-19 . Ciencia. Comunitario. 2020, 42, 679–697. [Referencia cruzada] [PubMed]
15. Enserink, M.; Kupferschmidt, K. Con COVID-19, el modelaje adquiere una importancia de vida o muerte. Ciencia 2020, 367, 1414–1415. [Referencia cruzada] [PubMed]
16. Musa, S.; Qureshi, S.; Zhao, S.; Yusuf, A.; Mustafa, UT; Él, D. Modelado matemático de la epidemia de COVID-19 con efecto de programas de concientización. Infectar. Dis. Modelo. 2021, 6, 448–460. [Referencia cruzada] [PubMed]
17. Moore, S.; colina, EM; Tildesley, MJ; Dyson, L.; Keeling, MJ Vacunación e intervenciones no farmacéuticas para COVID-19: A estudio de modelación matemática. Infección por lanceta. Dis. 2021, 21, 793–802. [Referencia cruzada] [PubMed]
18. Acayaah, Carolina del Norte; Subbarajasetty, SB; Shetty, RM R0 y Re de COVID-19: ¿Podemos predecir cuándo será el brote pandémico? ¿Contenido? Indio J. Crit. Cuidado médico. 2020, 24, 1125–1127. [Referencia cruzada] [PubMed]
19. Olivares, A.; Staffetti, E. Cuantificación de la incertidumbre de un modelo matemático de la dinámica de transmisión de COVID-19 con estrategia de vacunación masiva . Fractales de Solitones del Caos 2021, 146, 110895. [CrossRef] [PubMed]
20. Zhu, M.; Agyingi, E. Modelado de infecciones irruptivas por COVID-19 en una población vacunada. Wseas Trans. Sistema. 2023, 22, 584–592. [Referencia cruzada]

Machine Translated by Google

21. Nguyen, VK; Hernandez-Vargas, Estimación de parámetros de EA en modelos matemáticos de infecciones virales utilizando R. In Influenza Virus. Métodos en biología molecular, volumen 1836; Yamauchi, Y., Ed.; Humana Press: Nueva York, NY, EE. UU., 2018.
22. Dautel, K.; Agyingi, E. Modelización del impacto de una campaña educativa sobre la dinámica de transmisión del Ébola. J. Biol. Sistema. 2020, 29, 1–23. [Referencia cruzada]
23. LaSalle, JP La estabilidad de los sistemas dinámicos; SIAM: Filadelfia, PA, EE. UU., 1976.
24. Eikenberry, SE; Mancuso, M.; Iboi, E.; Phan, T.; Eikenberry, K.; Kuang, Y.; Kostelich, E.; Gumel, AB Usar o no usar mascarilla: modelar el potencial del uso de mascarillas por parte del público en general para reducir la pandemia de COVID-19 . Infectar. Dis. Modelo. 2020, 5, 293–308. [Referencia cruzada] [PubMed]
25. Tang, B.; Bragazzi, Países Bajos; Li, Q.; Tang, S.; Xiao, Y.; Wu, J. Una estimación actualizada del riesgo de transmisión del nuevo coronavirus (2019-nCoV). Infectar. Dis. Modelo. 2020, 5, 248255.
26. Xu, C.; Pang, Y.; Liu, Z.; Shen, J.; Liao, M.; Li, P. Información sobre el modelado estocástico de COVID-19 con efectos de diversas transmisiones Tarifas: Simulaciones con datos estadísticos reales de Reino Unido, Australia, España e India. Física. scr. 2024, 99, 025218 [Referencia cruzada]

Descargo de responsabilidad/Nota del editor: Las declaraciones, opiniones y datos contenidos en todas las publicaciones son únicamente de los autores y contribuyentes individuales y no de MDPI ni de los editores. MDPI y/o los editores renuncian a toda responsabilidad por cualquier daño a personas o propiedad que resulte de cualquier idea, método, instrucción o producto mencionado en el contenido.

Machine Translated by Google

AppliedMath

MDPI

Artículo

Aproximación de una función con una discontinuidad de salto: la Caso de alto ruido

Qusay Muzaffar 1,*, David Levin 2 y Michael Werman 1

1 Departamento de Ciencias de la Computación, Universidad Hebrea de Jerusalén, Jerusalén 9190401, Israel; michael.werman@mail.huji.ac.il
2 Departamento de Matemáticas Aplicadas, Universidad de Tel Aviv, Tel Aviv 6997801, Israel; levin@tauex.tau.ac.il
* Correspondencia: qusay.muzaffar@mail.huji.ac.il

Resumen: Este artículo presenta una novedosa red de aprendizaje profundo diseñada para detectar intervalos de discontinuidades de salto en funciones suaves por partes de una sola variable a partir de sus muestras ruidosas. Se puede mejorar la precisión de las estimaciones de discontinuidad de salto para encontrar una aproximación general más precisa de la función, ya que los métodos de aproximación tradicionales a menudo producen errores significativos cerca de las discontinuidades. Detectar intervalos de discontinuidades es relativamente sencillo cuando se trabaja con datos de funciones exactas, ya que las diferencias finitas en los datos pueden servir como indicadores de suavidad. Sin embargo, estos indicadores de suavidad se vuelven poco fiables cuando se trata de datos muy ruidosos. En este artículo, proponemos una red de aprendizaje profundo para identificar la ubicación de una discontinuidad de salto incluso en presencia de ruido sustancial.

Palabras clave: teoría de la aproximación; aprendizaje profundo; aproximación de función suave por partes; datos muy ruidosos

MSC: 65D15

check for updates

Cita: Muzaffar, Q.; Levin, D.; Werman, M. Aproximación de una función con una discontinuidad de salto: el caso de alto ruido. Matemáticas aplicadas 2024, 4, 561–569. https://doi.org/10.3390/appliedmath4020030

Editor académico: Claude Chaudet

Recibido: 29 de febrero de 2024
Revisado: 2 de abril de 2024
Aceptado: 15 de abril de 2024
Publicado: 2 de mayo de 2024

Copyright: © 2024 por los autores. Licenciatario MDPI, Basilea, Suiza.
Este artículo es un artículo de acceso abierto. distribuido bajo los términos y condiciones de los Creative Commons Licencia de atribución (CC BY) (https://creativecommons.org/licenses/by/4.0/).

1. Introducción

Las discontinuidades de salto ocurren cuando una función experimenta un cambio repentino y finito en su valor en un punto específico. Las funciones que contienen discontinuidades de salto prevalecen en varios campos, incluidos la física, la ingeniería eléctrica, la visión por computadora, la geología, la biología y otros dominios aplicados [1–3]. En consecuencia, surge con frecuencia la tarea de aproximar estas funciones a partir de datos muestreados. Estimar la ubicación de la discontinuidad es crucial ya que los métodos de aproximación estándar a menudo introducen errores significativos debido a la pérdida de regularidad cerca de las discontinuidades de salto.

Varios artículos han investigado algoritmos para aproximar funciones univariadas suaves por partes con un enfoque particular en la detección del intervalo de discontinuidades, incluido [4,5]. En un artículo más reciente [6], se utilizó una función de las diferencias en los datos, denominada firma, para aproximar la función. Sin embargo, los métodos antes mencionados fallan cuando se trata de datos ruidosos.

En otros artículos se han propuesto y discutido algoritmos alternativos diseñados para datos ruidosos, como en [7]. Sin embargo, estos métodos se basan en el conocimiento previo de la función, en encontrar un tipo de función, llamada función de promedio, y en parámetros de regularización. Además, no se proporciona ninguna información sobre la cantidad o el tipo de ruido aditivo. En [8], se presenta otro método basado en la segunda derivada numérica, aunque limitado a muestras en una cuadrícula uniforme. Requiere conocer el supremo de la segunda derivada de la función y se limita a un ruido inferior a una sexta parte de la longitud del salto. En [9], se describe un tercer método que utiliza convolución con una función de filtro. Sin embargo, un análisis comparativo demuestra una tasa de éxito menor en comparación con el método propuesto.

Machine Translated by Google

El aprendizaje profundo es una metodología de aprendizaje automático en la que se utilizan redes neuronales artificiales con múltiples capas de procesamiento para extraer progresivamente características de nivel superior de los datos entre la entrada y la salida. Los modelos de aprendizaje profundo se componen de múltiples capas de neuronas que procesan jerárquicamente los datos de entrada. Cada capa aprende a transformar los datos en una representación más abstracta, y la capa final proporciona el resultado del modelo (por ejemplo, clasificación, regresión, generación). El proceso de aprendizaje implica ajustar iterativamente los parámetros de la red neuronal (por ejemplo, pesos y sesgos) utilizando algoritmos de optimización como el descenso de gradiente [10].

Los modelos de aprendizaje profundo se pueden entrenar en grandes conjuntos de datos para aprender automáticamente patrones y características directamente a partir de los datos, eliminando la necesidad de ingeniería de características manual. Esta capacidad de aprender representaciones complejas a partir de datos sin procesar ha ganado una amplia aceptación como el enfoque preferido en diversas aplicaciones, incluido el reconocimiento de voz, el reconocimiento visual de objetos, la detección de objetos, el procesamiento del lenguaje natural, el descubrimiento de fármacos y la genómica [11].

En este artículo, proponemos una nueva red neuronal totalmente convolucional [10] para estimar la posición de la discontinuidad de salto de una función univariada a partir de un conjunto de muestras ruidosas. Posteriormente, empleamos dos splines cúbicos para aproximar la función, uno para los datos del lado izquierdo de la discontinuidad del salto y el otro para el lado derecho. Nuestro método ha demostrado ser superior a los métodos anteriores demostrados a través de una serie de experimentos.

2. Nuestro enfoque

En esta sección, presentamos nuestro enfoque para aproximar la ubicación de la discontinuidad de una función univariante suave por tramos f a partir de una muestra con ruido aditivo iid $\{(x_i , f(x_i) + \omega_i)\}_{i=0}^{n}$, donde

$$\omega_j \quad n\ 0, \sigma^2 \quad .$$

2.1. El modelo

Nuestro modelo de aprendizaje profundo consta de un total de diez capas convolucionales, como se ilustra en la Figura 1. A través de la experimentación, determinamos que un modelo de diez capas logra el mejor equilibrio entre la simplicidad en términos de características que se pueden aprender y la precisión de la salida. Estas diez capas constan de una capa de entrada, una capa de salida y ocho capas intermedias. Cada capa está equipada con dos operadores de convolución que utilizan tamaños de filtro de 3 y 5. Elegimos dos tamaños de filtro diferentes para que el modelo pueda capturar el comportamiento de la función a través de sus muestras de dos maneras diferentes. Vale la pena mencionar que elegimos filtros con números impares para que el tamaño de la salida después de la operación de convolución sea el mismo que el de la entrada después del relleno. Por ejemplo, si tenemos una entrada de tamaño N, agregamos ceros al principio y al final de la entrada, y luego aplicamos una convolución con un tamaño de filtro de 3 y conservamos el tamaño original de la entrada. que es N. Si tuviéramos que elegir un filtro de tamaño 2, entonces el tamaño de salida sería N + 1 en este caso. Mantener el tamaño de la entrada ayuda a capturar los detalles del modelo. Los canales de salida de estos dos operadores se concatenan como se muestra en la Ecuación (1). La capa inicial acepta 8 canales de entrada, ecuación (3), y genera 64 canales. Para la i-ésima capa intermedia, la entrada es de 64 × i canales y genera 64 canales. La salida de la i-ésima capa intermedia se concatena con sus canales de entrada y se pasa a la (i + 1)-ésima capa.

$$\boxed{\quad} \quad \boxed{v \quad Cf5} \quad . \qquad (1)$$

$$v \quad Cf3$$

La última capa tiene una entrada de 64 canales y genera un vector de probabilidad.

$$v(x) = \begin{cases} 0 \le x < 1\ 1 \\ 2 - x, \quad \le x \le 2 \end{cases} \qquad (2)$$

Machine Translated by Google

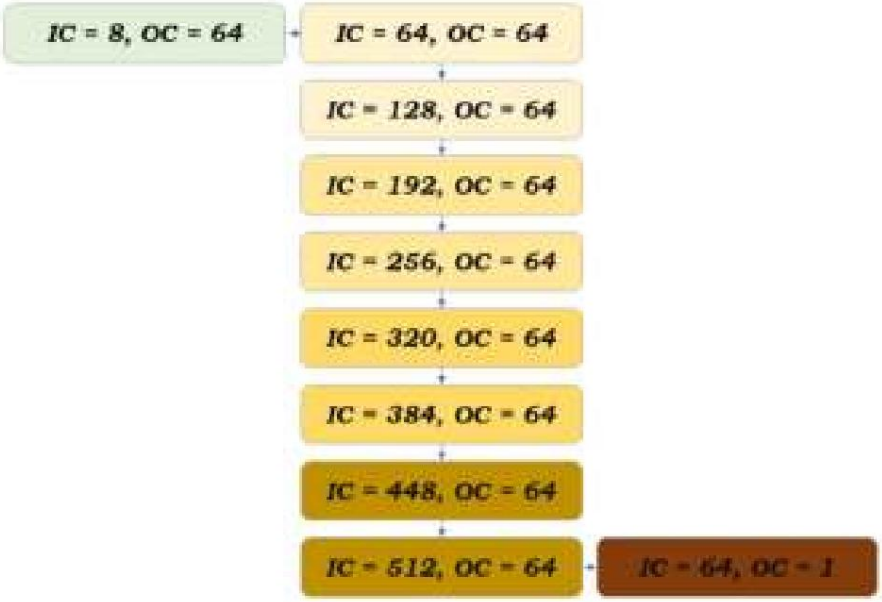
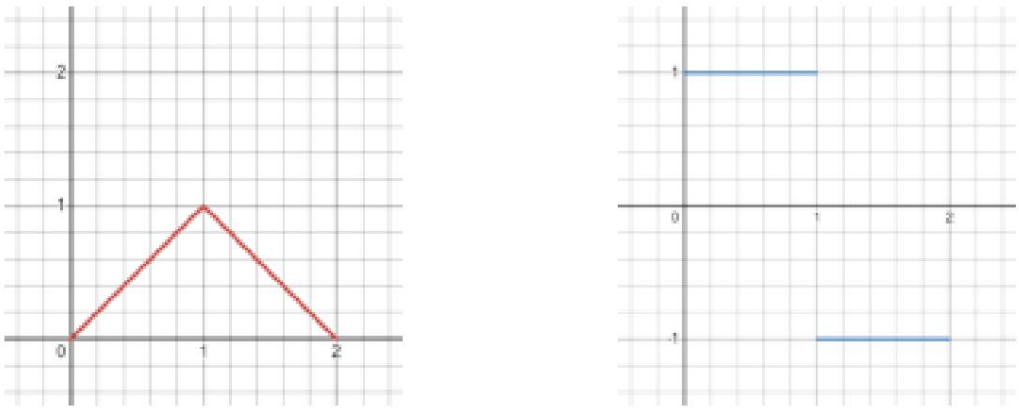


Figura 1. La arquitectura del modelo. Cada rectángulo representa una capa y los dos números indican el número de canales de entrada y de salida. En cada capa, hay dos operaciones convolucionales como en la Ecuación (1). En las capas intermedias, cada salida se concatena a la entrada de la misma capa y se pasa a la siguiente capa.

La función de activación, Ecuación (2) (ver Figura 2), se aplica a las salidas de los dos operadores de convolución después de que se concatenan antes de pasar a la siguiente capa.



(a) (b)

Figura 2. La gráfica de la función de activación (a) y la gráfica de su derivada (b).

Las muestras xey de tamaño N + 1 , donde N es arbitrario, son parte de los datos de entrada del modelo. Aumentamos los datos de entrada con diferencias directas de primer, segundo y tercer orden, como se presenta en la Ecuación (3).

$$(x, y) \rightarrow x, y, \Delta x, \Delta y, \Delta \quad {}^{2}_{x,\,\Delta y}{}^{2}, {}^{3}_{x,\,\Delta y}{}^{3} \tag{3}$$

La salida del modelo es un vector de probabilidades, de longitud N, donde la k-ésima entrada es la probabilidad de que haya una discontinuidad entre la k-ésima y la k + 1-ª muestra. La función de pérdida es de entropía cruzada [10] ya que el problema es la clasificación.

$$L(x) = -\sum_{n=1}^{norte}\sum_{c=1}^{do} \text{iniciar sesión}\frac{e^{x_{n,c}}}{\sum_{i=1}^{do} e^{x_{n,i}}} \tag{4}$$

Machine Translated by Google

Sea M el modelo y sea p = M x, y, ∆x, ∆y, ∆ el índice de $^{2}_{x,\ ^{2}\Delta y,\Delta^{3}_{x,\ \Delta^{3}}}$ y . Entonces, ˜i = arg max {p} la muestra más cercana al punto de discontinuidad desde el lado izquierdo. Nosotros definimos

$$\tilde{\xi} = \frac{\tilde{x}i + \tilde{x}i+1}{2} \tag{5}$$

como el punto de discontinuidad aproximado.

2.2. Datos de

entrenamiento En esta sección, describimos el proceso de generación de datos sintéticos para entrenar el modelo, tanto como entrada como salida. Recordamos que un número aleatorio entre a y b de una distribución uniforme se denota como

$$c \quad U[a, b]$$

Comenzamos generando funciones univariadas por partes, cada una de las cuales contiene una discontinuidad de salto , de la cual tomamos muestras aleatoriamente de puntos. Para crear un escenario realista, introducimos ruido blanco en estas muestras y las utilizamos como datos de entrada. Los datos de salida consisten en el índice de la muestra más cercana al lado izquierdo de la discontinuidad del salto. Dejar

$$t = -3\Delta t + i\Delta t \quad {}^{t}_{yo=0} \tag{6}$$

ser el conjunto de nudos donde $\Delta t \quad U[\frac{3}{10}\ ^{5}_{10,}]$ y $T = \Delta \quad {}^{-1}_{t} + 7$, entonces hay suficientes nudos dentro y fuera del intervalo, y dejemos

$$g(x) = \begin{matrix} 1 & , x < 0 \\ 0 & , 0 \le x \end{matrix} .$$

Ahora, dejamos

$$f(x) = ag(x - \xi) + bxg(x - \xi) + \sum_{yo=0}^{T-k-1} ciBi,k,t(x). \tag{7}$$

Aquí, Bi,k,t son funciones de base B-splines cúbicas (k = 3). Los coeficientes de los B-splines cúbicos $\{ci\}_{yo=0}^{T-k-3}$ se seleccionan aleatoriamente de una distribución uniforme:

$$\{ci\}_{yo=0}^{T-k-3} \quad U[-2, 2] .$$

De manera similar, se seleccionan los coeficientes a y b .

$$a, b \quad U - 2, -10 \quad \frac{4}{} \quad \frac{4}{2\ 10,}$$

Finalmente, la ubicación del punto de discontinuidad es la siguiente:

$$\xi \quad \frac{6}{}_{norte}, 1 - \frac{5}{}_{norte}$$

Los datos de entrada constan de puntos muestreados de una función f(x) de la forma como en Ecuación (7):

$$S = (xi , f(xi) + \omega i)_{norte\ yo = 0\ ,} \tag{8}$$

dónde

$$\omega i \quad U\ 0, 3 \quad \frac{a}{} .$$

Machine Translated by Google

Aquí, {ωi} $_{yo=0}^{norte}$ es el ruido aditivo. La salida es un vector unitario de longitud N − 1 donde 1 está ubicado en la muestra más cercana desde la izquierda de la discontinuidad del salto. Entrenamos el modelo con el algoritmo de descenso de gradiente estocástico ADAM [12] .

2.3. Aproximación de la función a partir de sus

muestras Después de ubicar el intervalo de discontinuidad y aproximar la ubicación del punto de discontinuidad, utilizamos dos aproximaciones spline cúbicas para aproximar la función con su discontinuidad de salto. Usamos la aproximación spline cúbica porque aproximar con splines cúbicos es ventajoso debido a su suavidad, flexibilidad, precisión, facilidad de implementación y estabilidad numérica. Estas propiedades hacen de los splines cúbicos una herramienta versátil y ampliamente utilizada para la interpolación de datos, ajuste de curvas y modelado en diversos campos, incluidas las matemáticas, la ingeniería, los gráficos por computadora y la informática científica [13 Primero, dividimos los datos en los grupos de la izquierda y los grupos de la derecha de la siguiente manera:

$$\begin{aligned} x\ell &= x \text{ es un componente de,} (x, y) \quad S : x < \tilde{\xi} \\ xr &= x \text{ es componente de,} (x, y) \quad S : x > \tilde{\xi} \\ y\ell &= y \text{ es un componente de,} (x, y) \quad S : x < \tilde{\xi} \\ yr &= y \text{ es componente de,} (x, y) \quad S : x > \tilde{\xi} \end{aligned} \tag{9}$$

donde S se define como en la Ecuación (8) y ˜ξ proviene de la Ecuación (5).

Luego, usando una aproximación de mínimos cuadrados, encontramos dos splines cúbicos, Γℓ (x), que aproxima la función del lado izquierdo de la discontinuidad de salto usando los datos de xℓ e yℓ y el segundo, Γr(x), que aproxima la función desde el lado derecho de la discontinuidad del salto utilizando los datos de xr e yr . Antes de continuar, comenzamos construyendo los nudos para ambas ranuras. Dejar

$$\Delta\ell = \left(\frac{\tilde{\xi}}{5}\right)^{-1} \quad \text{y} \quad \Delta r = \left(\frac{1-\tilde{\xi}}{5}\right)^{-1},$$

y sean $T\ell = \Delta + 7$ y $Tr = \Delta^{-1}_{r} + 7$. Luego, los dos conjuntos de nudos para la izquierda y la son

$$t\ell = -3\Delta\ell + i\Delta\ell \quad {}_{T\ell i=0} \quad \text{y} \quad tr = -3\Delta r + \xi + j\Delta r \quad {}_{Trj=0}$$

Denotamos el conjunto de los coeficientes de Γℓ (x) y Γr(x) por cℓ y cr , respectivamente. Para encontrar los conjuntos óptimos de coeficientes cℓ y cr , encontramos una solución en forma cerrada para la siguiente ecuación:

$$\arg\min_{\substack{c\ell \ R \\ T\ell-4\ cr \ \ RTr-4}} \left\| \frac{M\ell \ 0}{0 \ \text{señor}} \frac{c\ell}{cr} - \frac{y\ell}{\text{año}} \right\|_{L2}^{2} + \left\| \frac{\lambda\ell D\ell \ 0 \ 0}{\lambda r Dr} \frac{c\ell}{cr} \right\|_{L2}^{2}.$$

Aquí, las matrices de Mℓ Rxℓ×cℓ y Mr Rxr×cr son las matrices de datos y se definen de la siguiente manera:

$$M\ell = \begin{bmatrix} \ddots & \vdots & \iddots \\ \dots B_{j,k=3,t\ell}(x\ell,i) & \cdots & \text{, señor} = \end{bmatrix} \begin{bmatrix} \ddots & \vdots & \iddots \\ \dots B_{j,k=3,tr}(xr,i) \dots \\ \iddots & \vdots & \ddots \end{bmatrix}$$

$$cr-2 \times cr$$

Las matrices de Dℓ R cℓ−2 ×cℓ y Dr R son las matrices de segunda diferencia . Esas matrices se utilizan en la ecuación anterior de modo que la solución óptima es aquella que tiene una segunda derivada mínima.

Machine Translated by Google

3. Resultados numéricos

En esta sección, ilustramos los dos aspectos de nuestro enfoque: aproximar la ubicación de una discontinuidad de salto en una función dada y aproximar la función misma. Además, realizamos una prueba estadística para mostrar la efectividad de nuestro método en comparación con dos métodos alternativos. Por último, realizamos otra prueba estadística para presentar el error promedio de nuestra aproximación final.

3.1. Detectar el intervalo de la discontinuidad y aproximar la función

En esta sección, presentamos nuestro enfoque para aproximar el punto de discontinuidad de una función generada aleatoriamente y posteriormente aproximar la función misma. Para lograr esto, generamos dos funciones siguiendo el formato descrito en la Ecuación (7).
Tomamos 40 muestras aleatorias de [0, 1]. Luego, estas muestras fueron sometidas a la adición de ruido blanco con una variación igual a un tercio de la altura del salto.

A continuación, empleamos nuestro modelo para determinar el índice de la muestra más cercana a la izquierda de la discontinuidad. Posteriormente, utilizamos el procedimiento descrito en la Sección 2.3 para derivar una aproximación de la función. Los resultados de estos experimentos se presentan en la Figura 3.

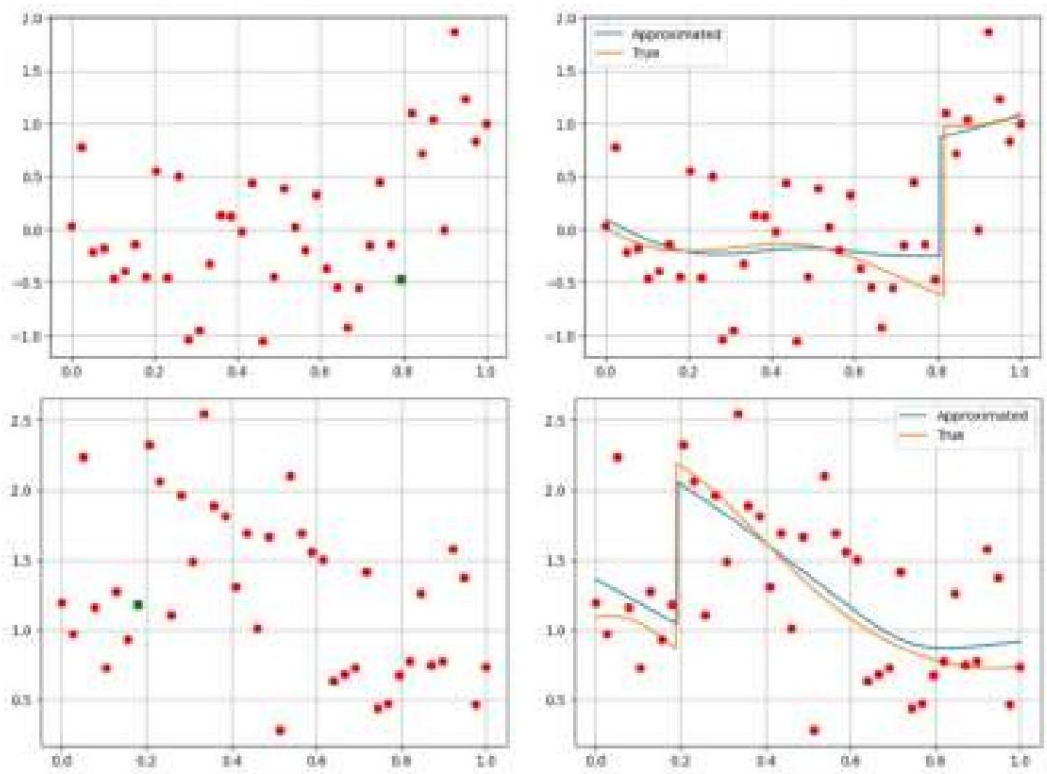


Figura 3. Demostramos nuestro método visualmente aproximando la ubicación del punto de discontinuidad de dos funciones suaves univariadas por partes de sus muestras, a las que se agregó ruido blanco. Las figuras superior izquierda e inferior izquierda muestran las muestras con ruido agregado. Todas las muestras son de color rojo excepto el punto que indica la muestra más cercana desde la derecha a la discontinuidad del salto, que es de color verde. Nuestro enfoque encontró el punto verde. Las figuras superior derecha e inferior derecha muestran las aproximaciones para las dos funciones, las cuales se lograron utilizando dos splines cúbicos para cada una de las funciones.

3.2. Comparando nuestro enfoque con otros dos enfoques

En [6], la firma, de grado k, de g, es σk (g) RN+k, que es la diferencia directa de orden k de las muestras de g rellenadas con k ceros en cada lado. En [9], se desarrolló un método basado en el uso de una función de filtro con la que se calcula una convolución para detectar el punto de discontinuidad. En este experimento, comparamos nuestro método con

Machine Translated by Google

los otros dos métodos antes mencionados en una gran cantidad de muestras con diversos ruidos niveles. Las muestras se toman de funciones suaves univariadas por partes con varios
Alturas de salto de discontinuidad. Cada una de las tablas 1 a 3 representa los resultados de uno de los enfoques. Las columnas representan la altura del salto y las filas representan la varianza en el ruido blanco aditivo. Como se muestra en las tablas, nuestro enfoque tiene un aumento de entre dos y diez veces en comparación con los otros métodos.

Tabla 1. La probabilidad de detectar el intervalo de la ubicación de la discontinuidad utilizando el método en [6].
Las columnas y filas representan la altura del salto y el nivel de ruido, respectivamente. Cada celda contiene el promedio éxito de 1000 ensayos.

	0,30	0,60	0,90	1.20	1,50	1,80	2.10	2.40
0.00	0,38	0,47	0,50	0,49	0,48	0,49	0,48	0,49
0,10	0,10	0,28	0,45	0,50	0,50	0,53	0,48	0,51
0,20		0,10	0,21	0,29	0,42	0,46	0,46	0,50
0,30			0,12	0,14	0,24	0,29	0,37	0,42
0,40				0,11	0,14	0,19	0,25	0,31
0,50					0,11	0,13	0,17	0,22
0,60						0,11	0,13	0,17
0,70							0,10	0,12
0,80								0,12

Tabla 2. La probabilidad de detectar el intervalo de la ubicación de la discontinuidad utilizando el método en [9].
Las columnas y filas representan la altura del salto y el nivel de ruido, respectivamente. Cada celda contiene el promedio éxito de 1000 ensayos.

	0,30	0,60	0,90	1.20	1,50	1,80	2.10	2.40
0.00	0,24	0,24	0,28	0,24	0,26	0,25	0,24	0,23
0,10	0,30	0,27	0,25	0,25	0,27	0,24	0,25	0,27
0,20		0,30	0,26	0,28	0,24	0,25	0,28	0,25
0,30			0,33	0,28	0,29	0,26	0,28	0,28
0,40				0,34	0,30	0,30	0,26	0,30
0,50					0,34	0,32	0,30	0,28
0,60						0,32	0,29	0,33
0,70							0,34	0,31
0,80								0,33

Tabla 3. La probabilidad de detectar el intervalo de la ubicación de la discontinuidad utilizando nuestro método.
Las columnas y filas representan la altura del salto y el nivel de ruido, respectivamente. Cada celda es el promedio éxito de 1000 ensayos.

	0,30	0,60	0,90	1.20	1,50	1,80	2.10	2.40
0.00	0,89	0,99	1.00	1.00	1.00	1.00	1.00	1.00
0,10	0,53	0,93	1.00	1.00	1.00	1.00	1.00	1.00
0,20		0,67	0,92	0,98	1.00	1.00	1.00	1.00
0,30			0,73	0,92	0,97	0,99	1.00	1.00
0,40				0,76	0,90	0,95	0,98	0,99
0,50					0,79	0,89	0,94	0,97
0,60						0,79	0,88	0,95
0,70							0,79	0,88
0,80								0,79

3.3. Medición de errores de aproximaciones de funciones

En este experimento, comparamos nuestro enfoque con el enfoque de usar la firma [6] al aproximar una función suave univariante por partes a partir de sus muestras ruidosas. Como antes, Realizamos una prueba en muchas muestras con varios niveles de ruido. Medimos el error de la aproximación utilizando la métrica L^∞ en el conjunto $0, \min(\xi, \tilde{\xi}) \quad \max(\xi, \tilde{\xi}), 1$, donde ξ y $\tilde{\xi}$ son el punto de discontinuidad real y aproximado. Además, medimos

Machine Translated by Google

˜ξ − ξ . Creamos dos gráficos, en las Figuras 4 y 5, para los dos enfoques y los dos medidas. El eje x representa la varianza del ruido blanco aditivo, mientras que el El eje y representa la altura del salto. Cada punto del gráfico representa el promedio error de 100 intentos. En la primera prueba, el error de nuestro enfoque fue menos de la mitad del error del otro enfoque, y en la segunda prueba, el error de nuestro enfoque fue el quinto.

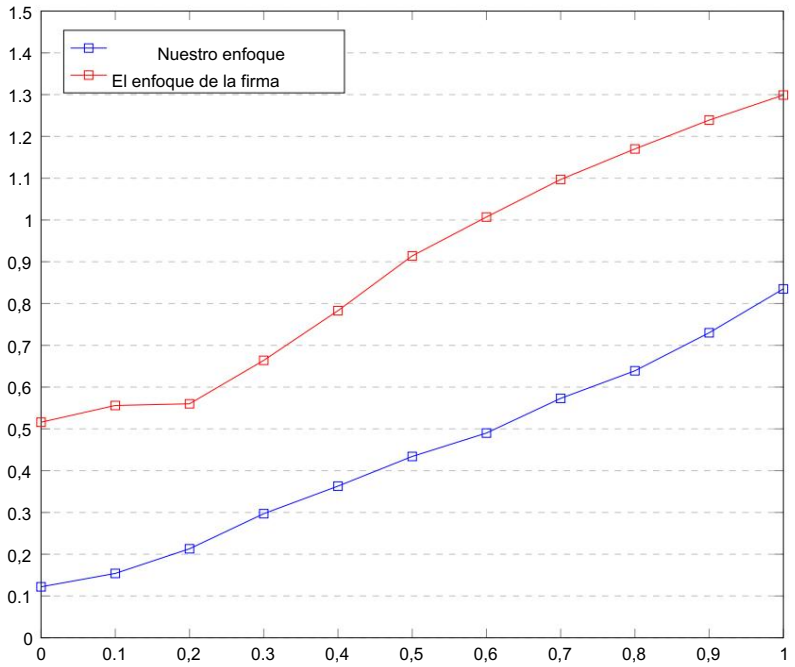


Figura 4. Este gráfico muestra el error L∞ de la aproximación en el conjunto 0, min (ξ, ˜ξ) máx (ξ, ˜ξ), 1 , donde ξ y ˜ξ son el punto de discontinuidad real y aproximado. Aquí, ξ y ˜ξ son el punto de discontinuidad real y aproximado.

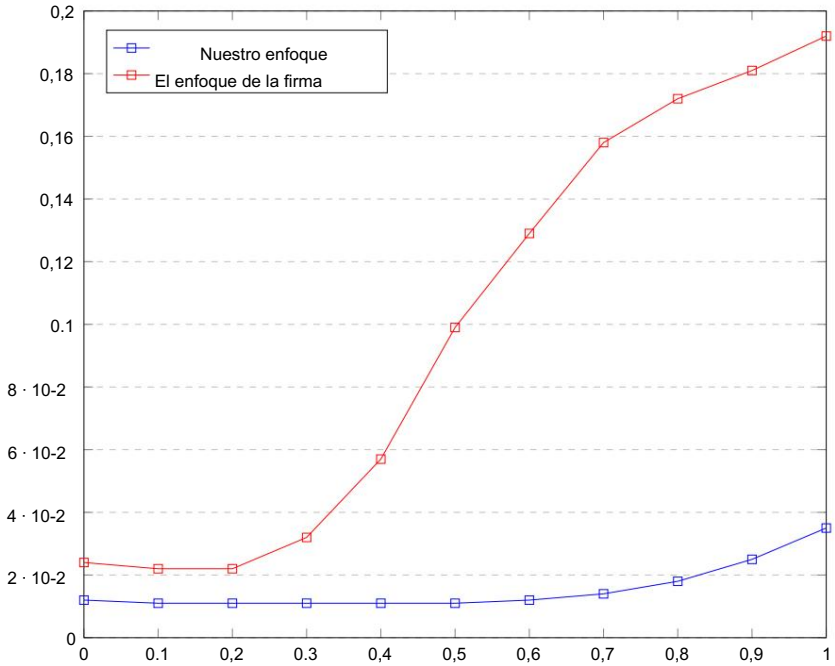


Figura 5. Este gráfico muestra el error de ˜ξ − ξ . Aquí, ξ y ˜ξ son el punto real y aproximado de discontinuidad.

Machine Translated by Google

4. Conclusiones

Presentamos un nuevo enfoque, que se basa en el aprendizaje profundo, para aproximar la ubicación del punto de discontinuidad de salto a partir de muestras dadas, con ruido altamente aditivo, de una función suave univariante por partes con una discontinuidad de salto. Comparamos nuestro enfoque con otros dos métodos e ilustramos la tasa de éxito significativamente mayor de nuestro método con ellos. Además, demostramos nuestro enfoque para aproximar la ubicación de la discontinuidad en datos sintéticos de funciones univariadas suaves por partes generadas. Luego, aproximamos estas funciones y mostramos las med

Nuestro trabajo futuro se centrará en extender este trabajo a funciones bivariantes suaves por partes, en las que el salto no ocurre en un punto sino en una curva unidimensional. En caso de que los datos bidimensionales estén igualmente distanciados a lo largo del eje y, podemos utilizar este enfoque para localizar los puntos más cercanos a la curva unidimensional. Sin embargo , aproximar la curva unidimensional no es intuitivo y será un tema para trabajos futuros.

Contribuciones de los autores: Metodología, Gestión de la Calidad; investigación, gestión de calidad; redacción: preparación del borrador original, gestión de calidad; redacción: revisión y edición, MW y QM; supervisión, MW y DL Todos los autores han leído y aceptado la versión publicada del manuscrito.

Financiamiento: Fundación Científica Israelí ISF.

Declaración de disponibilidad de datos: No se utilizan datos específicos en este documento.

Conflictos de intereses: Los autores declaran no tener ningún conflicto de intereses.

Referencias

1. Ageev, AL; Antonova, TV Sobre una nueva clase de problemas mal planteados. Izv. Ural. Vaya. Univ. Estera. Mej. Informar. 2008, 58, 24–42.
2. Cherepashchuk, AM; Goncharskii, AV; Yagola, AG Problemas inversos de la astrofísica. En Problemas mal planteados en ciencias naturales, Actas de la conferencia internacional celebrada en Moscú del 19 al 25 de agosto. 1991; De Gruyter: Berlín, Alemania, 1992; págs. 472–481.
3. Terebizh, VY Introducción a la teoría estadística de problemas inversos; Fimzmatlit: Moscú, Rusia, 2005.
4. Arándiga, F.; Donat, R. Descomposiciones multiescala no lineales: el enfoque de A. Harten. Número. Algoritmo 2000, 23, 175–216. [Referencia cruzada]
5. Shu, C.-W. Esquemas esencialmente no oscilatorios y ponderados esencialmente no oscilatorios para leyes de conservación hiperbólicas. En Aproximación numérica avanzada de ecuaciones hiperbólicas no lineales: conferencias impartidas en la segunda sesión del Centro Internazionale Matematico Estivo (CIME), celebrada en Cetraro, Italia, del 23 al 28 de junio. 1997; Springer: Berlín/Heidelberg, Alemania, 1998; págs. 325–432.
6. Amat, S.; Levin, D.; Ruiz-Álvarez, J. Una estrategia de aproximación en dos etapas para funciones suaves por partes en dos y tres dimensiones. IMA J. Número. Anal. 2021, 42, 3330–3359. [Referencia cruzada]
7. Antonova, TV Nuevos métodos para localizar discontinuidades de una función ruidosa, Numer. Anal. Aplica. 2010, 3, 306–316.
8. Mbakop, E. Reconstrucción de discontinuidades en datos ruidosos; Citador: Pensilvania, Estados Unidos. Disponible en línea: https://citeseerx.ist.psu. edu/document?repid=rep1&type=pdf&doi=0c1b69c38e1df16c7c939ad07705ee6bb53a6bd2 (consultado el 31 de marzo de 2024).
9. Lee, D. Detección, clasificación y medición de discontinuidades. SIAM J. Ciencias. Estadística. Computadora. 1991, 12, 311–341. [Referencia cruzada]
10. Buen compañero, I.; Bengio, Y.; Courville, A. Aprendizaje profundo; Prensa del MIT: Cambridge, MA, EE. UU., 2016.
11. LeCun, Y.; Bengio, Y.; Hinton, G. Aprendizaje profundo. Naturaleza 2015, 521, 436–444. [Referencia cruzada] [PubMed]
12. Kingma, DP; Ba, J. Adam: Un método para la optimización estocástica. arXiv 2014, arXiv:1412.6980.
13. Chapra, Carolina del Sur; Canale, RP Métodos numéricos para ingenieros, 8ª ed.; Educación McGraw-Hill: Nueva York, NY, EE. UU., 2021.

Descargo de responsabilidad/Nota del editor: Las declaraciones, opiniones y datos contenidos en todas las publicaciones son únicamente de los autores y contribuyentes individuales y no de MDPI ni de los editores. MDPI y/o los editores renuncian a toda responsabilidad por cualquier daño a personas o propiedad que resulte de cualquier idea, método, instrucción o producto mencionado en el contenido.

Machine Translated by Google

AppliedMath

MDPI

Artículo

Mezclas cuánticas y pérdida de información en sistemas de muchos cuerpos

Diana Monteoliva [1], Ángelo Plastino [2,*] y Ángel Ricardo Plastino [3]

[1] UNLP-Comisión de Investigaciones Científicas, Provincia de Buenos Aires, La Plata 1900, Argentina; monteoli@fisica.unlp.edu.ar
[2] Instituto de Física La Plata–CCT-CONICET, Universidad Nacional de La Plata, CC 727, La Plata 1900, Argentina CeBio-
[3] Departamento de Ciencias Básicas, Universidad Nacional del Noroeste, Prov. de Buenos Aires (UNNOBA), CONICET, Junín 1988, Argentina; arplastino@unnoba.edu.ar
* Correspondencia: plastino@fisica.unlp.edu.ar

Resumen: En nuestro estudio, investigamos el fenómeno de la pérdida de información, medida por la divergencia Kullback-Leibler , en un sistema de muchos fermiones, como el modelo de Lipkin. La pérdida de información se introduce a medida que aumenta el número N de partículas, particularmente cuando el sistema está en un estado mixto. Encontramos que hay una pérdida significativa de información en estas condiciones. Sin embargo, observamos que esta pérdida casi desaparece cuando el sistema está en estado puro. Nuestro análisis emplea herramientas de la teoría de la información para cuantificar y comprender estos efectos.

Palabras clave: sistemas de muchos fermiones; grado de mezcla; temperatura finita; simetría SU2;

MSC: 81-82

1. Introducción

En este artículo estudiamos, utilizando herramientas de la teoría de la información, las peculiaridades de las mezclas cuánticas [1–5] en el contexto de un modelo exactamente soluble de N fermiones interactuantes de masa m llamado modelo de Lipkin. Nos centramos en las pérdidas de información, medidas por la divergencia Kullback-Leibler (KL), que tienen lugar cuando N aumenta desde un número de fermión de referencia anterior, inferior, No.

Presentamos a continuación los principales contenidos de nuestro estudio.

El modelo de Lipkin [6-16] es un modelo bien conocido de N fermiones de masa m que interactúan. Es bien conocido en física nuclear y en sistemas cuánticos de muchos cuerpos, particularmente en el estudio de la interacción entre correlaciones de emparejamiento y transiciones de fase cuánticas [6]. El uso de herramientas de la teoría de la información para explorar las correlaciones entre el grado de mezcla cuántica (QMx) y algunos rasgos del modelo es una dirección de investigación interesante y válida. Aquí, prestamos especial atención a la relación QMx-número de partícula N. La teoría de la información proporciona un formalismo para cuantificar y comprender las correlaciones, la entropía y el contenido de la información en los sistemas físicos. Los conceptos clave incluyen entropía, información mutua y entropía condicional. Estas herramientas se pueden aplicar para analizar las relaciones entre diferentes parámetros en un sistema cuántico y proporcionar información sobre la complejidad, las correlaciones y el entrelazamiento presentes en el sistema. Este enfoque interdisciplinario, que combina el modelo de Lipkin, la teoría de la información y la mecánica estadística cuántica, puede proporcionar un marco rico para explorar las intrincadas conexiones entre los rasgos del modelo y los cuantificadores estadísticos especiales.

1.1. Mezclas cuánticas

En esta contribución, el concepto de mezcla cuántica (QMx) [1–4] se estudia desde el punto de vista de la interacción de sistemas de muchos fermiones. QMx es importante en la mecánica cuántica porque refleja la naturaleza probabilística inherente de los sistemas cuánticos y proporciona una forma de describir su comportamiento estadístico. La mecánica cuántica es fundamentalmente diferente.

check for updates

Cita: Monteoliva, D.; Plastino, A.; Plastino, AR Mezclas Cuánticas y pérdida de información en sistemas de muchos cuerpos . Matemáticas aplicadas 2024, 4, 570–579. https:// doi.org/10.3390 /appliedmath4020031

Recibido: 4 de marzo de 2024
Revisado: 13 de abril de 2024
Aceptado: 20 de abril de 2024
Publicado: 2 de mayo de 2024

Copyright: © 2024 por los autores. Licenciatario MDPI, Basilea, Suiza. Este artículo es un artículo de acceso abierto. distribuido bajo los términos y condiciones de los Creative Commons Licencia de atribución (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

Machine Translated by Google

de la mecánica clásica y una de sus características clave es la existencia de estados de superposición y el concepto asociado de entrelazamiento cuántico. El concepto de mezclas cuánticas es importante

porque [1–5]: • En la mecánica cuántica, los sistemas pueden existir en múltiples estados simultáneamente, un fenómeno conocido como superposición. Una mezcla cuántica es una combinación probabilística de estos estados y nos permite describir el estado general de un sistema

- cuando no está en estado puro; Predicciones probabilísticas: la mecánica cuántica predice resultados de forma probabilística. El estado de un sistema se describe mediante una función de onda, y las probabilidades de diferentes resultados están determinadas por las magnitudes al cuadrado de las amplitudes de probabilidad. Una mezcla cuántica encapsula

estas predicciones probabilísticas para sistemas compuestos; • Mediciones y observables: cuando se realiza una medición en un sistema cuántico, el sistema normalmente colapsa en uno de sus posibles estados y el resultado de la medición es probabilístico. Las mezclas cuánticas proporcionan una forma de expresar la distribución estadística de resultados de mediciones en un

conjunto de sistemas preparados de manera idéntica; • Teoría de la información cuántica: los conceptos de mezcla cuántica son fundamentales en la teoría de la información cuántica, donde se estudia la manipulación y transmisión de información cuántica. Comprender cómo evolucionan e interactúan los estados mixtos es esencial para desarrollar

- algoritmos cuánticos y protocolos de comunicación cuánticos; Mecánica estadística: en el contexto de la mecánica estadística, las mezclas cuánticas se utilizan para describir los conjuntos estadísticos de sistemas cuánticos. El gran conjunto canónico, por ejemplo,

implica una mezcla de estados con diferentes números de partículas; • En resumen [1–5], el concepto de mezcla cuántica es un aspecto fundamental de la mecánica cuántica que nos permite manejar la naturaleza probabilística y estadística de los sistemas cuánticos. Es una herramienta crucial para hacer predicciones, comprender correlaciones y desarrollar tecnologías

Las consideraciones anteriores dejan muy claro que los estados mixtos desempeñan un papel crucial en la mecánica cuántica, y que su importancia radica en proporcionar una descripción más completa y realista de los sistemas físicos en comparación con los estados puros solos. Como se indicó anteriormente, el estado mixto representa un conjunto estadístico de estados puros. Este conjunto puede incluir diferentes estados puros con determinadas probabilidades. Matemáticamente, un estado mixto se describe mediante una matriz de densidad, que es un operador hermitiano positivo semidefinido que actúa sobre el espacio de Hilbert. Los estados mixtos son esenciales cuando se trata de conjuntos estadísticos, equilibrios térmicos o sistemas cuánticos abiertos que están sujetos a interacciones con su entorno. En muchas situaciones prácticas, es posible que no tengamos un conocimiento completo sobre el estado exacto de un sistema cuántico. En cambio, podríamos tener información estadística sobre las probabilidades de diferentes estados puros en los que podría estar el sistema. Los estados mixtos proporcionan una manera natural de describir tales conjuntos estadísticos. En la teoría de la información cuántica, los estados mixtos desempeñan un papel central en la caracterización del desempeño de diversos algoritmos cuánticos y protocolos de comunicación cuánticos. De hecho, hay que destacar que los estados mixtos son fundamentales para una comprensión integral de los sistemas cuánticos, especialmente cuando se considera la naturaleza estadística de las mediciones, los efectos térmicos y el impacto de las interacciones con el medio ambiente. Proporcionan un puente entre el concepto idealizado de estados puros y los aspectos prácticos de los sistemas cuánticos del mundo real [1-5]. Finalmente, es interesante para nuestros propósitos actuales observar que la pérdida de comportamientos característicos cuánticos a medida que aumenta el número de partículas es un fenómeno bien conocido [5]. De hecho, sabemos que cuando el número de fermiones en un colectivo se vuelve muy grande, ciertos sistemas cuánticos pueden exhibir un comportamiento que se aproxima al comportamiento clásico. A esto se le suele denominar principio de correspondencia. El comportamiento estadístico de un gran número de partículas cuánticas puede entonces parecerse al comportamiento clásico a escala macroscópica [5].

Machine Translated by Google

1.2. Naturaleza de nuestros resultados

Destaquemos que en la presente contribución cada resultado es exacto. Utilizamos un modelo de muchos fermiones fácilmente solucionable llamado modelo de Lipkin (LM) [6] que es de gran relevancia en el campo de la física cuántica y la física de la materia condensada [6-16]. Este modelo, aunque simplificado, ofrece información valiosa sobre el comportamiento de sistemas cuánticos complejos y proporciona entornos simplificados en los que se pueden explorar analíticamente los principios cuánticos fundamentales. Los investigadores suelen utilizar el LM para desarrollar y probar nuevos marcos teóricos, como las técnicas de muchos cuerpos y la teoría cuántica de campos [6-16].

Usaremos aquí temperaturas muy bajas. Esta es una técnica de mecánica estadística bien conocida para aproximar las propiedades del estado fundamental de un sistema. Da como resultado un enfoque común y poderoso en física de la materia condensada y mecánica cuántica. Este enfoque explota la relación entre las propiedades de baja temperatura y el estado fundamental de un sistema [15,16]. Tomamos como nuestras bajas temperaturas T los valores correspondientes a la temperatura inversa β = 20 o 10, siendo β = 1/kT y k la constante de Boltzmann.

1.3. Objetivo actual y organización

Nuestro objetivo es investigar algunas pistas sobre las consideraciones precedentes en el contexto de un célebre modelo de física nuclear llamado Modelo Lipkin [7]. Hay cientos de referencias pertinentes que se podrían citar. Por supuesto, no hay espacio para tal tarea. Nos contentamos con [6-16]. El modelo nos ofrece un entorno puramente cuántico. Deseamos investigar aquellos de los rasgos del modelo que representan mezclas cuánticas. El resto del documento está organizado de la siguiente manera. Comenzamos en la Sección 2 describiendo las matemáticas del modelo de Lipkin [6-16]. La sección 3 recuerda materiales relevantes para la divergencia Kullback-Leibler, una de las dos principales herramientas de información que empleamos aquí. La segunda herramienta avanza en este esfuerzo en el contexto del modelo Likpin (es bien conocido en otros lugares), y se aborda en la Sección 4, donde mostramos que constituye un cuantificador válido para evaluar el grado de mezcla. Nuestro tema central sobre la pérdida de información se analiza en la Sección 5 y las conclusiones se extraen en la Sección 6.

2. El formalismo de Lipkin [15,16]

Consideramos N fermiones de masa m que interactúan y definimos Ω = N/2. El modelo de Lipkin [6-16] consta de N = 2Ω fermiones de masa m que se acomodan en dos distintos

niveles de energía de una sola partícula (sp), cada uno de ellos degenerado N veces. Una brecha energética separa los dos niveles. Esta brecha se llama . Nos encontramos así ante un total de microestados de 4Ω sp, etiquetados por dos números cuánticos (denominados μ y p). El primer número cuántico μ alcanza los valores μ = −1 (nivel inferior) y μ = +1 (nivel superior). El número cuántico p, llamado cuasi espín, pertenece a una degeneración 2N veces. Este par p, μ se considera habitualmente como un “sitio”. Este sitio puede estar ocupado (por un fermión) o estar vacío. Conjuntos Lipkin [6]

$$norte = 2J, \tag{1}$$

donde J actúa como una especie de momento angular. Además, Lipkin [6] emplea operadores particulares similares al momento angular denominados operadores de cuasi-espín, que son

$$Jz = \sum_{p,\,\mu} \mu\, C^{+}_{p,\mu} C_{p,\mu}, \tag{2}$$

$$J+ = \sum_{p\mu} do^{+}_{p,+} Cp,-, \tag{3}$$

$$J- = \sum_{p\mu} do^{+}_{p,-} Cp,+, \tag{4}$$

Machine Translated by Google

junto con el operador Casimir

$$j^2 = J_z^2 + \frac{1}{2}(J_+ J_- + J_- J_+). \quad (5)$$

Los valores propios de la ^{2} tomar la forma J(J + 1) y el hamiltoniano de Lipkin toma la forma forma J (v es una constante de acoplamiento) [6]

$$H = J_z + \frac{v}{4}(j_+^2 + J_-^2). \quad (6)$$

La matriz de nuestro hamiltoniano dice [6]

$$\begin{aligned} {}_{norte}\langle HL|n\rangle_{norte} = {} & \frac{norte}{2}\,norte - norte + 1 - norte\frac{}{2} - norte^2 + 2n - 1\,\omega\,\delta n', \\ & -\frac{v}{2}\sqrt{(N-n)(N-n+1)(n+1)n}\,\delta n', n+2 \\ & -\frac{v}{2}\sqrt{(norte - norte')(norte - norte + 1)(n'+1)n'}\,\delta n', n-2 \end{aligned} \quad (7)$$

con norte = 0, 1 ,. . . , N para J = N/2. Una diagonalización (numérica) da los valores propios energéticos En(v, J) .

2.1. Entorno externo del sistema: baño de calor del conjunto canónico de Gibb a temperatura T

Se supone que nuestra estructura Lipkin interactúa con un depósito de calor. Con los valores propios anteriores [15,16] construimos una función de partición de conjunto canónico Z [15] de la cual obtenemos todos los cuantificadores térmicos asociados que emergen [15,16].

Construimos Z empleando probabilidades que entran en Z [16]. Nuestros indicadores térmicos y Z se derivan de distribuciones de probabilidad [16] pn(v, J, β). β es, como vimos anteriormente, la temperatura inversa. Las fórmulas pertinentes se encuentran en [16]. Si denominamos la energía media U, la entropía S y la energía libre F, obtenemos [15,16]

$$pn(v, J, \beta) = \frac{1}{Z(v, J, \beta)}\, mi^{-\beta En(v,J)}, \quad (8)$$

$$Z(v, J, \beta) = \sum_{norte=0}^{norte} mi^{-\beta En(v,J)}, \quad (9)$$

$$\begin{aligned} U(v, J, \beta) = \langle E \rangle &= -\frac{\partial \ln Z(v, J, \beta)}{\partial \beta} \\ &= \sum_{norte=0}^{norte} En(v, J) Pn(v, J, \beta) \\ &= \frac{1}{Z(v, J, \beta)} \sum_{norte=0}^{norte} En(v, J) e^{-\beta En(v,J)}, \end{aligned} \quad (10)$$

$$S(v, J, \beta) = -\sum_{norte=0}^{norte} Pn(v, J, \beta) \ln[Pn(v, J, \beta)] \quad (11)$$

$$F(v, J, \beta) = U(v, J, \beta) - T\,S(v, J, \beta). \quad (12)$$

Estos cuantificadores térmicos producen mucha más información que la obtenida mediante los cuantificadores cuánticos de temperatura cero T [16]. Como se indicó anteriormente, utilizando una T suficientemente baja, nuestros indicadores proporcionan una representación razonablemente buena del panorama T = 0 [16]. A continuación, normalmente tomamos β = 10 o incluso menos.

2.2. Grado de mezcla de Lipkin para un estado cuántico ρ

La pureza y la mezcla son nociones básicas que desempeñan un papel importante en la descripción de sistemas cuánticos como el modelo de Lipkin. Las dos nociones son muy importantes en el análisis de

Machine Translated by Google

información cuántica, computación, entrelazamiento y mediciones. Comprender la diferencia entre estados puros y mixtos requiere comprender la coherencia, la superposición y el comportamiento estadístico. Como es bien sabido, el grado de mezcla Cf de un estado ρ es, en términos de las probabilidades pi [17],

$$Cf = 1 - Tr\rho^2 = 1 - \sum_i p_i^2, \quad (13)$$

donde $Tr\rho^2$ es la pureza Py (P mayúscula). Para un estado puro, se tiene Cf = 0 y Py = 1. Cf es un cuantificador muy importante en este estudio. $Py = \sum_{norte=0}^{norte} (Pn(v, J, \beta))2$ y $Cf = S2 = 1 - P_y^2$.

Reiteramos también, a riesgo de redundancia, que para un estado de Lipkin a la temperatura fija T , se utilizan, para las probabilidades y la función de partición, todas las cantidades importantes

$$pn(v, J, \beta) = \frac{1}{Z(v, J, \beta)} e^{-\beta En(v,J)} Z(v, J, \beta) = \sum_{norte=0}^{norte} m_i^{-\beta En(v,J)}. \quad (14)$$

3. La divergencia Kullback-Leibler (KL) [18]

Esta es una herramienta importante de la teoría de la información. La divergencia Kullback-Leibler (KL) cuantifica la distinguibilidad entre dos estados. Para los estados cuánticos, la divergencia KL puede capturar diferencias relacionadas con la interferencia y el entrelazamiento, proporcionando información valiosa sobre las propiedades del sistema cuántico. Se encuentran aplicaciones de KL en la teoría de la información cuántica, la computación cuántica y la comunicación cuántica, donde comprenderla y cuantificarla produce una medida de la diferencia entre dos distribuciones de probabilidad. En la teoría de la información clásica, la divergencia KL entre las distribuciones de probabilidad P(x) y Q(x) viene dada por:

$$DKL(P||Q) = \sum_x P(x)\log(P(x)/Q(x)). \quad (15)$$

La divergencia KL cuantifica cuánta información se pierde cuando se utiliza Q para aproximar P. No es negativa e igual a cero si y sólo si las dos distribuciones son iguales . En el contexto de la mecánica cuántica, el concepto de divergencia KL se puede ampliar para comparar dos estados cuánticos. Para dos estados cuánticos descritos por los operadores de densidad ρ y σ, la divergencia cuántica de Kullback-Leibler FF viene dada por

$$D(\rho||\sigma) = Tr(\rho\log(\rho) - \rho\log(\sigma)). \quad (16)$$

4. Nuevo indicador térmico cuántico para el modelo Lipkin

Esto puede sonar extraño al principio, pero considere la cantidad (con unidades de longitud).

$$\lambda = \frac{2\pi\hbar}{\sqrt{2\pi mkBT}}, \quad (17)$$

llamada longitud térmica de De Broglie. Por supuesto, no hay una dimensión de longitud en el modelo de Lipkin, pero nuestro escenario contiene fermiones de masa m (masa del sistema Nm) a la temperatura T, de modo que se puede calcular un valor de λ y depende de nosotros mostrar su relevancia para el modelo. Modelo de Lipkin, que hacemos a continuación. Según la convención de De Broglie, λ muy pequeños deberían indicar clasicidad. Esto sucede obviamente para T alto y N grande como lo indica (17), como debería esperarse intuitivamente.

La masa del protón es pf = 1,67262192 × 10−27 kg. El radio del átomo de hidrógeno rH es 10-10 metros. Se puede trazar λ versus Nmp. Como $\hbar$ = 6,582119569 × 10−16 Js, se obtiene

$$\lambda = \frac{2\,2\pi\hbar}{\text{diputado}} \quad \beta = 0{,}0403419 \text{ norte} \quad \beta = 1{,}0857 \times 1020 \text{ norte} \quad \frac{1}{\text{Nuevo Testamento}} \text{ Rh}, \quad (18)$$

donde el centro considera kB = 1 y se expresa en metros, mientras que la última igualdad se expresa en unidades rH y tomamos β = 1/kBT con kB = 1.380648 × 10−23 J/K.

Machine Translated by Google

4.1. Relación entre λ y N

La Figura 1 muestra λ versus el número de fermión N. Este último disminuye, como se esperaba, cuando N crece. Esperamos este comportamiento porque es bien sabido que la clasicidad comienza a insinuarse para N grande [5].

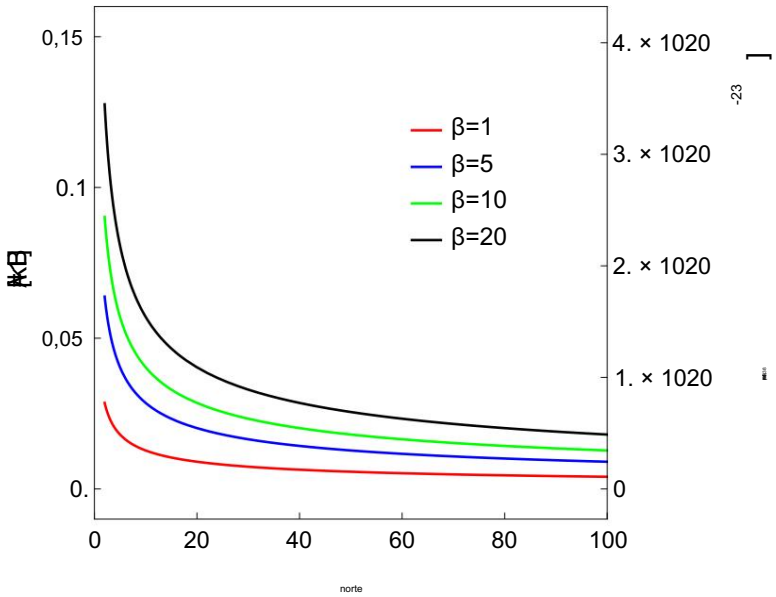


Figura 1. Cuantificador de longitud térmica λ vs. N para varios valores de β = 1, 5, 10 y 20 (en metros para la escala vertical izquierda y en términos del radio del átomo de hidrógeno en la escala vertical derecha) tomando kB = 1 (izquierda) y kB = 1,380648 × 10−23 (derecha). Vemos que la longitud térmica disminuye a medida que crece el número de fermiones en el sistema, a tasas reguladas por la temperatura. Las tasas aumentan a medida que disminuye la temperatura. La longitud se vuelve más pequeña a medida que N crece, lo que puede leerse como un vestigio de clasicidad. Que los efectos sean más perceptibles con temperaturas más bajas que con temperaturas más altas es algo que uno debería esperar intuitivamente.

Repitamos que el gráfico indica claramente que nuestro cuantificador λ funciona de la manera que cabría esperar. De hecho, como se mencionó anteriormente, el conocido surgimiento de la clasicidad a medida que crece el número de partículas N en un sistema es un aspecto fascinante de la física cuántica [5]. NO encontraremos más este fenómeno en este trabajo.

4.2. Relación entre pureza, temperatura y λ

La Figura 2 representa Py = 1 − Cf versus λ.

Vemos que, como se prometió anteriormente, la longitud térmica muestra las propiedades esperadas de crecer cuando la pureza es grande y disminuir cuando el grado de mezcla es grande. Por tanto, λ queda validado como indicador cuántico.

La limitación que observamos en el modelo de Lipkin, en el que el grado de mezcla cuántica no puede exceder la mitad, probablemente surge de la combinación del principio de exclusión de Pauli y las características específicas de los niveles degenerados en el sistema. El principio de exclusión de Pauli restringe la ocupación de sitios por fermiones. No pueden dos fermiones ocupar el mismo sitio simultáneamente. En nuestro caso, cada sitio p de cuasi-espín puede acomodar dos fermiones (uno con giro hacia arriba y otro con giro hacia abajo).

En el modelo de Lipkin, cada nivel puede albergar múltiples fermiones. Sin embargo, el número máximo de fermiones que pueden ocupar los niveles degenerados pertinentes es 2N. esto es porque Hay 2N sitios distintos disponibles.

Cuando el sistema está medio lleno, como es el caso aquí, tenemos N fermiones en total. Esta situación corresponde a un fermión en cada uno de los espines p dos veces degenerados y representa un escenario de mezcla cuántica máxima.

Machine Translated by Google

El principio de exclusión de Pauli todavía se aplica independientemente a cada estado-sitio de espín, permitiendo un máximo de dos fermiones (giro hacia arriba y hacia abajo) en cada espín p degenerado.

En resumen, el principio de exclusión de Pauli, combinado con la degeneración de los niveles, da como resultado una mezcla cuántica máxima cuando el sistema está lleno hasta la mitad con N fermiones. Esta situación corresponde a un fermión en cada cuasi-espín p dos veces degenerado, lo que lleva a un estado de superposición máxima. La estructura específica del modelo de Lipkin, considerando la degeneración quais-spin, determina las restricciones sobre la ocupación y la mezcla cuántica resultante.

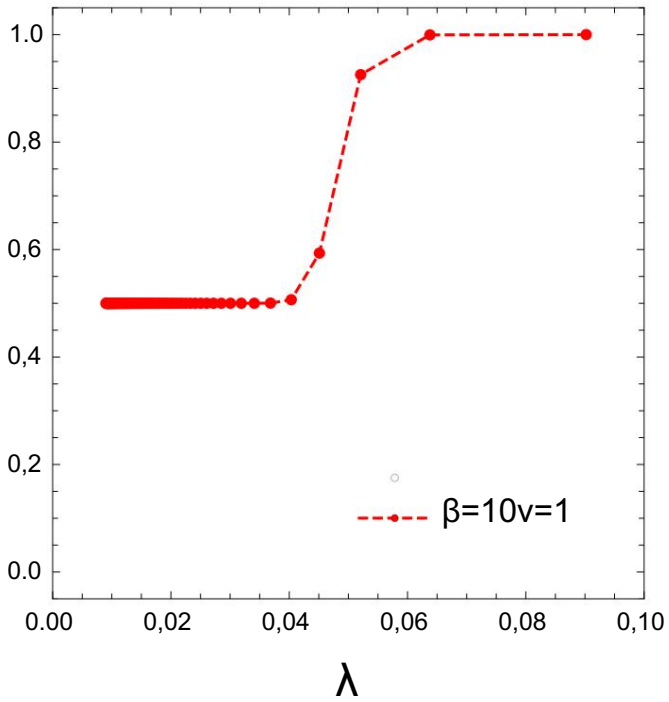


Figura 2. Pureza versus longitud de onda térmica de De Broglie (en metros) para β = 10, N = 20 y v = 1. Cuando la pureza es pequeña, la longitud de onda también lo es, y viceversa cuando la pureza se acerca a la unidad. Se detecta una pureza mínima igual a la mitad (ver texto).

5. Pérdida de información y λ

5.1. La relación KL-pérdida de información

Reformulamos aquí nuestro material de Kullback-Leibler en términos del número de fermiones N que contiene nuestro sistema Lipkin. La divergencia de Kullback-Leibler (KL) entre dos distribuciones de probabilidad P1 y P2 proporciona, como vimos, una medida de cómo una distribución de probabilidad diverge de otra. Específicamente, KL(P1||P2) cuantifica la información perdida cuando se usa P2 para aproximar P1. En otras palabras, la divergencia de KL mide la cantidad adicional esperada de información necesaria para codificar datos de P1 usando una distribución de probabilidad P2 , es decir, cuantifica cuánta información se pierde cuando la verdadera distribución P1 se aproxima a P2. En este trabajo, la primera distribución será la que describe un sistema de N-fermiones, mientras que P2 hace lo mismo para un sistema idéntico pero que contiene 2N partículas.

5.2. El grado de pérdida de información cuando N cambia

Ahora abordamos una cuestión relacionada con el grado de pérdida de información a medida que N cambia, utilizando KL como una medida de cómo una distribución de probabilidad (PD) diverge de una segunda PD en el contexto de comparar sistemas con diferentes números de partículas. Encontraremos que la divergencia KL disminuye al comparar un sistema con un número de partículas N relativamente pequeño con uno con un N mayor para sistemas mixtos cuánticos. El efecto de disminución es drásticamente

Machine Translated by Google

menor si los sistemas están en estado puro. Esta cuestión se aborda en la Figura 3, que relaciona λ (en metros) con la divergencia KL entre los fermiones N y 2N (N = 100). Se pierde información cuando N aumenta, pero lo hace de forma muy diferente si el estado es mixto o puro.
Este hecho constituye el principal hallazgo del presente trabajo.

El gráfico muestra que este proceso de pérdida también ocurre a longitudes térmicas mayores, cuanto mayor se vuelve v . La característica principal aquí es que el grado de pérdida de información es bastante grande para longitudes térmicas pequeñas. Sostenemos que la pérdida continúa existiendo en gran medida. No vemos que esto ocurra en nuestro gráfico por razones de escala. Lo que queda claro es que cuanto mayor es la pureza, menor es la pérdida de información.

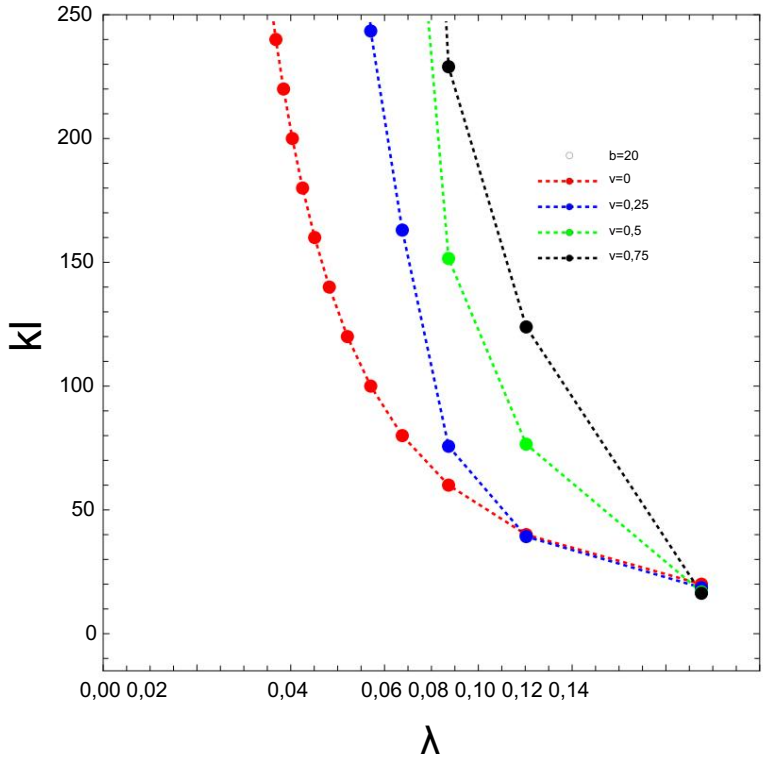


Figura 3. KL versus λ para varios valores de v , obtenidos al comparar distribuciones de probabilidad pertenecientes a escenarios N versus 2N . Estos valores de KL miden nuestra pérdida de información. Tomamos β = 20. Cuando alcanzamos λ 13 cm, estamos en el rango de estado puro de nuestro indicador, como vimos anteriormente. En este régimen de estado puro , la pérdida de información es tan pequeña que no puede verse en la escala vertical que hemos elegido aquí. Las diferencias son menores que los puntos utilizados allí. Vemos que la pérdida de información cuando aumentamos N es inmensamente mayor para los estados mixtos que para los puros.

5.3. Posibles interpretaciones

En la mecánica cuántica, el entrelazamiento es una característica única que puede existir en estados puros. Cuando un sistema está en estado puro, el entrelazamiento entre partículas puede conducir a distribuciones de probabilidad más complejas y correlacionadas. Por el contrario, los estados mixtos surgen de conjuntos de sistemas cuánticos, potencialmente con un entrelazamiento reducido. La presencia de estados mixtos podría conducir a un cambio más gradual en la distribución de probabilidad a medida que aumenta el número de partículas, lo que resultaría en una divergencia KL más pequeña.

Los estados mixtos pueden verse como conjuntos estadísticos de estados puros. Al comparar un sistema con un número de partículas pequeño con uno con un número de partículas grande en un escenario de estado mixto, en realidad se está considerando un promedio sobre una variedad de posibles estados puros dentro del conjunto. Este promedio podría conducir a un cambio más suave y menos pronunciado en la distribución de probabilidad, contribuyendo a una divergencia KL más pequeña.

Las mezclas cuánticas pueden exhibir correlaciones reducidas entre partículas en comparación con los estados puros entrelazados. A medida que aumenta el número de partículas, la independencia estadística

Machine Translated by Google

entre partículas podría volverse más pronunciada, dando lugar a una distribución de probabilidad menos compleja y más predecible.

Los estados mixtos suelen tener una entropía mayor que los estados puros. El aumento de entropía puede contribuir a una distribución de probabilidad más uniforme y menos informativa, particularmente a medida que aumenta el número de partículas. Esto puede dar como resultado una divergencia KL más pequeña.

6. Conclusiones

Hemos encontrado diferentes formas de visualizar la diferencia entre estados mixtos y puros y también hemos introducido un nuevo cuantificador para el grado de mezcla cuántica, llamado λ, que puede expresarse como una longitud.

Para los sistemas considerados en este trabajo, los valores de λ mayores que 12 cm indican una pureza igual a la unidad, mientras que los valores más pequeños indican grados crecientes de mezcla de estados (ver Figura 3). Cabe señalar que, según la ecuación (18), valores de λ menores que, digamos, 0,001 mm, podrían ser evidencia de vestigios del comportamiento clásico. Por lo tanto, podemos llamar apropiadamente a nuestra λ una longitud pseudotérmica que es un indicador de pureza útil para el modelo de Lipkin. Para alta pureza, produce valores grandes y viceversa para un alto grado de mezcla de estados.

Vimos que, a una temperatura finita T, la pureza del sistema Py disminuye de manera más rápida. moda cuando el grado de interacción fermión-fermión v aumenta.

Descubrimos que, en el modelo de Lipkin a una temperatura finita T, el crecimiento de N aumenta el grado de mezcla. Si se aumenta N, se pierde información, medida por la divergencia KL . Cuanto mayor sea el grado de mezcla, mayor será la pérdida de información que la acompaña. En el caso de estados puros (a temperatura cero), la pérdida se vuelve extremadamente pequeña.

Contribuciones de los autores: Investigación, DM, AP y ARP; Administración de proyectos, AP; Redacción: borrador original, DM, AP y ARP. Todos los autores han leído y aceptado la versión publicada del manuscrito.

Financiamiento: Esta investigación fue apoyada parcialmente por CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas de Argentina).

Declaración de disponibilidad de datos: Todos los datos están contenidos en el manuscrito.

Conflictos de intereses: Los autores declaran no tener conflictos de intereses.

Referencias

1. Bengtsson, I.; Zyczkowsi, K. Geometría de estados cuánticos: una introducción al entrelazamiento cuántico; Cambridge University Press: Cambridge, MA, EE. UU., 2006.
2. Jaeger, G. Información cuántica: descripción general; Springer: Berlín/Heidelberg, Alemania, 2007.
3. Giulini, D.; Joos, E.; Kiefer, C.; Kupsch, J.; Stamatescu, I.-O.; Zeh, HD Decoherencia y la aparición de un mundo clásico en la teoría cuántica; Springer: Berlín/Heidelberg, Alemania, 1996.
4. Facchi, P.; Mariano, A.; Pascazio, S. Decoherencia versus entropía en interferometría de neutrones. Física. Rev. A 2001, 63, 052108. [Referencia cruzada]
5. Zurek, WH La decoherencia y la transición de lo cuántico a lo clásico: revisada. Física. Hoy 1991, 44, 1–31. [Referencia cruzada]
6. Lipkin, HJ; Meshkov, N.; Glick, AJ Validez de métodos de aproximación de muchos cuerpos para un modelo soluble: (I) Soluciones exactas y teoría de perturbaciones. Núcleo. Física. 1965, 62, 188. [Referencia cruzada]
7. Lipkin, HJ; Meshkov, N.; Glick, AJ Validez de métodos de aproximación de muchos cuerpos para un modelo soluble: (II). Linealización procedimientos. Núcleo. Física. 1965, 62, 199. [Referencia cruzada]
8. Lipkin, HJ; Meshkov, N.; Glick, AJ Validez de métodos de aproximación de muchos cuerpos para un modelo soluble: (III). Soluciones exactas y teoría de la perturbación. Núcleo. Física. 1965, 62, 188. [Referencia cruzada]
9. Lipkin, HJ; Meshkov, N.; Glick, AJ Validez de métodos de aproximación de muchos cuerpos para un modelo soluble: (IV). La solución Hartree-Fock deformada . Núcleo. Física. 1965, 62, 220. [Referencia cruzada]
10. Rowe, DJ El modelo Lipkin. Mod. Rev. Física. 1968, 40, 153. [Referencia cruzada]
11. Schmid, KW; Bauer, A. El modelo Lipkin. Física. Representante 1982, 88, 251–334.
12. Isacker, PV; Heinze, S. El modelo Lipkin: un paradigma para una teoría autoconsistente de las transiciones de formas. Física. Rep. 1991, 199, 67.
13. Dukelsky, J.; Pittel, S.; Sierra, G. Transiciones de fase cuántica en el modelo Lipkin-Meshkov-Glick. Mod. Rev. Física. 2004, 76, 643. [Referencia cruzada]
14. Li, J. Transición de fase cuántica y diagrama de fase en el modelo Lipkin-Meshkov-Glick. Física. Rev. B 2017, 95, 075139.

Machine Translated by Google

15. Plastino, AR; Monteoliva, D.; Plastino, A. Características teóricas de la información de muchos sistemas de fermiones: una exploración basada en modelos exactamente solucionables. Entropía 2021, 23, 1488. [CrossRef] [PubMed]
16. Monteoliva, D.; Plastino, A.; Plastino, AR Los cuantificadores estadísticos resuelven una controversia sobre la teoría nuclear. Q. Rep. 2022, 4, 127. [Referencia cruzada]
17. Anillo, P.; Schuck, P. El problema nuclear de muchos cuerpos; Springer: Berlín/Heidelberg, Alemania, 1980.
18. Sánchez-Moreno, P.; Zarzo, A.; Dehesa, JS Jensen divergencia basada en la información de Fisher. J. Física. Una Matemática. Teor. 2012, 45, 125305. [Referencia cruzada]

Descargo de responsabilidad/Nota del editor: Las declaraciones, opiniones y datos contenidos en todas las publicaciones son únicamente de los autores y contribuyentes individuales y no de MDPI ni de los editores. MDPI y/o los editores renuncian a toda responsabilidad por cualquier daño a personas o propiedad que resulte de cualquier idea, método, instrucción o producto mencionado en el contenido.

Machine Translated by Google

AppliedMath

MDPI

Artículo

Una ecuación de Helmholtz modificada por G con nuevas expansiones para el El perturbador potencial gravitacional de la Tierra, sus funcionales y el estudio de las superficies isogravitacionales

Gerassimos Manoussakis

Departamento de Matemáticas, Facultad de Ciencias Físicas y Matemáticas Aplicadas, Universidad Técnica Nacional de Atenas, Iroon Polytechneiou 9, 15780 Zografos, Grecia; gmanous@math.ntua.gr

Resumen: La ecuación de Helmholtz modificada con G es una ecuación diferencial parcial que nos permite expresar la intensidad de la gravedad g como una serie de armónicos esféricos que tienen una distancia radial r en potencias irracionales. La ecuación de Laplace en el espacio tridimensional (en coordenadas cartesianas, es la suma de las derivadas parciales de segundo orden de la cantidad desconocida igual a cero) se utiliza para expresar el potencial de gravedad de la Tierra (potencial perturbador y normal) con el fin de representar otros cantidades útiles , que también se conocen como funcionales del potencial perturbador, como la perturbación de la gravedad , la anomalía de la gravedad y la ondulación del geoide como una serie de armónicos esféricos. Demostramos que al utilizar la ecuación de Helmholtz modificada con G, no solo la intensidad de la gravedad sino también el potencial perturbador y sus funcionales pueden expresarse como una serie de armónicos esféricos. Tener la intensidad de la gravedad representada como una serie de armónicos esféricos nos permite crear nuevos Modelos de Gravedad Global . Además, se lleva a cabo un examen más detallado de las superficies isogravitacionales de la Tierra.
Finalmente, tabulamos nuestros resultados, lo que deja en claro que los nuevos modelos de gravedad global para la intensidad de la gravedad g serán muy útiles para muchas aplicaciones geofísicas y geodésicas.

Palabras clave: gravedad; anomalía de la gravedad; perturbación de la gravedad; gradiente vertical de gravedad; potencial perturbador ; armónicos esféricos; superficies isogravitacionales

MSC: 35A09; 35C10; 35J25; 33C05; 33C75

check for updates

Cita: Manoussakis, G. Una ecuación de Helmholtz modificada por G con nuevas expansiones para el potencial gravitacional perturbador de la Tierra , sus funcionales y el estudio de superficies isogravitacionales. Matemáticas aplicadas 2024, 4, 580–595. https://doi.org/10.3390/appliedmath4020032

Editor académico: Thomas Woolley

Recibido: 7 de marzo de 2024
Revisado: 24 de abril de 2024
Aceptado: 26 de abril de 2024
Publicado: 4 de mayo de 2024

Copyright: © 2024 por el autor. Licenciatario MDPI, Basilea, Suiza.

Este artículo es un artículo de acceso abierto. distribuido bajo los términos y condiciones de los Creative Commons Licencia de atribución (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

1. Introducción

Los armónicos esféricos son importantes en muchas aplicaciones teóricas y prácticas (campos electromagnéticos, configuraciones electrónicas, campos gravitacionales, geoides, gráficos por computadora, radiación cósmica de fondo de microondas, etc.). Algunas de estas aplicaciones se mencionan brevemente en este trabajo. Los armónicos esféricos han sido ampliamente estudiados y son ampliamente aplicables debido a sus numerosas propiedades convenientes. En gráficos por computadora [1], son bastante útiles ya que comparten muchas de las mismas fortalezas que las wavelets de Haar, y existen técnicas que cierran la brecha entre ellas. Esto permite la representación de funciones de baja y alta frecuencia. Además, en [2–4], se utilizan en la formación de un modelo global realista de rayos.

El potencial de gravedad es una cantidad expresada como una serie de armónicos esféricos. Esta serie se utiliza para formular modelos de gravedad global y modelos de gravedad local. Un modelo de gravedad global [5], también conocido como modelo geopotencial global, sirve como una representación matemática del potencial gravitacional externo de un cuerpo celeste. Aquí nos centramos en el escenario en el que la Tierra es el cuerpo atractivo. El uso de modelos de gravedad global nos permite determinar todas las funciones del campo de gravedad relacionadas, por ejemplo, vectores de gravedad y gravedad.

Es importante señalar que estos modelos [6–9] son cruciales ya que ofrecen información valiosa sobre las características de la Tierra, incluida la ondulación geoide, anomalías de gravedad, perturbaciones de gravedad,

Machine Translated by Google

y deflexión vertical. Vale la pena mencionar que estos modelos tienen una amplia gama de aplicaciones , incluida la determinación de órbitas, la navegación inercial para aviones y misiles, estudios oceanográficos de la circulación oceánica e investigaciones geofísicas sobre distribuciones de densid

El campo vectorial magnético de la Tierra [10] está relacionado con el potencial geomagnético escalar que obedece a la ecuación de Laplace y puede representarse como una serie de armónicos esféricos. Las aplicaciones del potencial geomagnético [11,12] son increíblemente amplias; van desde estudiar la propagación de ondas hidromagnéticas y las corrientes ionosféricas hasta investigar las mareas, monitorear los cambios en el clima troposférico y explorar la corteza terrestre.

La intensidad de la gravedad y las fuerzas electrostáticas se puede expresar matemáticamente como una serie de armónicos esféricos. Esta serie surge de la solución de la ecuación de Helmholtz modificada con G. La intensidad de la electrostática desempeña un papel fundamental en una amplia gama de aplicaciones [13], que van desde la transferencia de calor y masa hasta la síntesis de nanomateriales y la reducción de las emisiones de CO2. Además, se han empleado esferoides [14-16] para abordar problemas de dispersión electromagnética y analizar interacciones electrostáticas entre partículas esferoidales en diferentes escenarios. La intensidad de la gravedad es muy valiosa en diversas aplicaciones. Desempeña un papel crucial [17,18] en la detección de fallas sísmicas, estimando estructuras esféricas simples dentro de la Tierra, determinando geoide y cuasigeoide (que definen la forma física de la Tierra y la superficie de referencia de las alturas físicas), [19] analizando topografía de la superficie del mar , comprensión de la estructura de la litosfera, inferir el espesor del hielo flotante e incluso detectar estructuras arqueológicas durante la exploración.

En este estudio, presentamos un método novedoso para determinar el potencial perturbador que es independiente de la ecuación de Laplace. Al expresar [21] la anomalía de la gravedad y la perturbación de la gravedad como una serie de armónicos esféricos utilizando la ecuación de Helmholtz modificada (con la distancia radial, r, elevada a potencias irracionales), podemos determinar sin esfuerzo el potencial perturbador mediante una simple ecuación diferencial parcial. Este nuevo enfoque nos permite representar cantidades relativas asociadas con el potencial gravitacional (es decir, sus funcionales) en una nueva serie de armónicos esféricos. Además, al expresar la intensidad de la gravedad como una serie de armónicos esféricos, podemos realizar un análisis más exhaustivo de las superficies isogravitacionales. Es importante destacar que, en este contexto, se considera a la Tierra como un cuerpo que no gira.

2. Ecuación de Laplace, potencial perturbador y cantidades relativas Centrándonos

en el caso de la Tierra, el potencial perturbador T [22] (que es la diferencia entre el potencial de gravedad real W de la Tierra y el potencial normal U generado por un elipsoide de revolución adecuado) es una solución de la ecuación de Laplace.

$$2T = 0 \tag{1}$$

El concepto de potencial perturbador tiene una importancia significativa, ya que es vital para calcular la ondulación N del geoide, que representa la distancia entre la superficie del geoide y la superficie de un elipsoide de revolución adecuado. Además, juega un papel crucial en la determinación de otras cantidades importantes, como la perturbación de la gravedad δg, la anomalía de la gravedad Δg y las deflexiones de la vertical. Hay tres problemas clásicos [5] para determinar el potencial perturbador T a partir de valores límite en una esfera: (a) el problema de Poisson, (b) el problema de Hotine y (c) el problema de Stokes.

El primer problema es un problema de Dirichlet sobre una esfera con un radio de R, donde los valores de T están dados en la esfera. El segundo problema, como aproximación esférica, implica resolver un problema de Neumann. En este problema, se nos proporcionan los valores de la perturbación de la gravedad δg (que representa la derivada radial del potencial perturbador T). El tercer problema (como aproximación esférica) es un problema de Robin en el que los valores de la anomalía de gravedad Δg se dan sobre el geoide cumpliendo la siguiente relación:

$$\frac{\partial T}{\partial r} + \frac{2}{R} T = -\Delta g \tag{2}$$

Machine Translated by Google

En lugar de resolver problemas de valores límite para determinar el potencial perturbador T, utilizamos Modelos Geopotenciales Globales (GGM). Un Modelo Geopotencial Global [5,23] es una expresión matemática que permite calcular el potencial perturbador mediante expansión armónica esférica. La serie armónica de un GGM es finita y también proporciona una colección de expresiones matemáticas, valores numéricos y algoritmos. Con estas herramientas, los usuarios pueden realizar fácilmente cálculos para determinar valores numéricos de cantidades asociadas con el potencial perturbador (como la perturbación de la gravedad y la anomalía de la graveda Además, estos cálculos permiten la evaluación de errores específicos relacionados con estas cantidades.

Presentamos algunas expresiones de series finitas [22-24], en coordenadas esféricas (r, θ, λ) (θ se mide desde el eje z (coeficientes normalizados).

$$T(r,\theta,\lambda) = \sum_{n=2}^{N\text{máx}} \frac{R}{r}^{n+1} Tn(\theta,\lambda) = \\ = \frac{GM}{r} \sum_{n=2}^{N\text{máx}} \sum_{m=0}^{n} \frac{R}{r}^{n+1} (\overline{Cnm}\cos m\lambda + \overline{Snm}\,\text{sen}\, m\lambda)\overline{Pnm}(\cos\theta) \tag{3}$$

$$\delta g(r,\theta,\lambda) = -\frac{\partial T}{\partial r} = \frac{GM}{2r} \sum_{n=0}^{N\text{máx}} (n+1) \frac{R}{r}^{n+1} Tn(\theta,\lambda) \tag{4}$$

$$\Delta g(r,\theta,\lambda) = \frac{GM}{2r} \sum_{n=0}^{N\text{máx}} (n-1) \frac{R}{r}^{n+1} Tn(\theta,\lambda) \tag{5}$$

La ondulación del geoide se calcula mediante

$$norte(\theta,\lambda) = \frac{1}{\gamma elispoide(\theta,\lambda)} T(R,\theta,\lambda) \tag{6}$$

donde γ es el valor de la gravedad normal [22] en el elipsoide, G es la constante gravitacional y M es la masa de la Tierra. Los componentes de la deflexión de la vertical [23] son

$$\xi(r,\theta,\lambda) = -r\gamma(r,\frac{1}{\theta,\lambda)}\frac{\partial T}{\partial\theta} \tag{7}$$

$$\eta(r,\theta,\lambda) = -r\gamma(r,\frac{1}{\theta,\lambda)\,\text{sen}\,\theta}\frac{\partial T}{\partial\lambda} \tag{8}$$

$$\frac{\partial T}{\partial\theta} = \frac{GM}{r} \sum_{n=2m}^{N\text{máx}} \sum_{=0}^{n} \frac{R}{r}^{n+1} (\overline{Cnm}\cos m\lambda + \overline{Snm}\,\text{sen}\, m\lambda)(\overline{Pnm+1}(\cos\theta) - m\tan\theta\cdot\overline{Pnm}(\cos\theta)) \tag{9}$$

$$\frac{\partial T}{\partial\lambda} = -\frac{GM}{r} \sum_{n=2m}^{N\text{máx}} \sum_{=0}^{n} \frac{R}{r}^{n+1} m(\overline{Cnm}\,\text{sen}\, m\lambda - \overline{Snm}\cos m\lambda)\overline{Pnm}(\cos\theta) \tag{10}$$

En las ecuaciones (7) y (8), la gravedad normal [22] se determina en un punto adecuado que está ubicado en una superficie equipotencial apropiada del campo de gravedad normal. El potencial perturbador , T, es una cantidad que no se puede medir directamente. Por tanto, se requieren otras cantidades mensurables para determinar su valor. Sin embargo, demostraremos un método alternativo para determinar T en el siguiente párrafo.

3. Ecuación de Helmholtz modificada con G, potencial perturbador y cantidades relativas

Suponiendo que la perturbación de la gravedad δg es una función continua en la superficie del geoide, podemos determinarla resolviendo el problema de Dirichlet que involucra la ecuación de Helmholtz modificada con G.

$$\begin{aligned} &2(\delta g) - \frac{2}{r^2}\delta g = 0 \\ &\delta g|S = f1(\theta',\lambda) \end{aligned} \tag{11}$$

Machine Translated by Google

La expresión de la perturbación de la gravedad en armónicos esféricos [21] es la siguiente:

$$\delta g(r, \theta', \lambda) = \sum_{n=0}^{+\infty} \sum_{m=0}^{norte} r^{-\frac{1+\sqrt{9+4n(n+1)}}{2}} \delta g\,[a_{nm} P_{nm}(\sin\theta') \cos m\lambda + b\,\delta g_{nm} P_{nm}(\sin\theta') \sin m\lambda] \quad (12)$$

Según [22], sobre la superficie del geoide, que como aproximación esférica corresponde a una esfera de radio R, podemos deducir

$$f1(\theta', \lambda) = \sum_{n=0}^{+\infty} \sum_{m=0}^{norte} R^{-\frac{1+\sqrt{9+4n(n+1)}}{2}} \delta g_{nm} P_{nm}(\sin\theta') \cos m\lambda + b\,\delta g\,[a \quad (13)$$

Puesta

$$A_{\text{Nuevo México}}^{\delta g} = R^{-\frac{1+\sqrt{9+4n(n+1)}}{2}} \delta g_{un\,nm} \quad (14)$$

$$B_{\text{Nuevo México}}^{\delta g} = R^{-\frac{1+\sqrt{9+4n(n+1)}}{2}} b\,\delta g \quad (15)$$

La serie en la Ecuación (13) se convierte en

$$f1(\theta', \lambda) = \sum_{n=0}^{+\infty} \sum_{m=0}^{norte} [A_{nm}^{\delta g} \delta g P_{nm}(\sin\theta') \cos m\lambda + B_{nm}^{\delta g} P_{nm}(\sin\theta') \sin m\lambda] \quad (16)$$

De este modo,

$$\int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f1(\theta', \lambda) P_{jk}(\sin\theta') \sin\theta' d\lambda d\theta' =$$

$$= \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sum_{n=0}^{+\infty} \sum_{m=0}^{norte} [A_{nm}^{\delta g} \delta g P_{nm}(\sin\theta') \cos m\lambda + B_{nm}^{\delta g} \delta g P_{nm}(\sin\theta') \sin m\lambda] \cdot \quad (17)$$

$$\cdot P_{jk}(\sin\theta') \sin\theta' d\lambda d\theta'$$

y los coeficientes [22] de la serie (17) son iguales a

$$U_{n0}^{\delta g} = \frac{2n+1}{4\pi} \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f1(\theta', \lambda) P_n(\sin\theta') \sin\theta' d\lambda d\theta' \quad (18)$$

$$A_{\text{Nuevo México}}^{\delta g} = \frac{\text{metro})!}{4\pi} \frac{(\text{norte} -}{2n+1\,(n+m)!} \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f1(\theta', \lambda) P_{nm}(\text{sen}\,\theta') \text{ porque } m\lambda \text{ sen } \theta' d\lambda d\theta' \quad , \quad \text{metro} = 0 \quad (19)$$

$$B_{\text{Nuevo México}}^{\delta g} = \frac{\text{metro})!}{4\pi} \frac{(\text{norte} -}{2n+1\,(n+m)!} \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f1(\theta', \lambda) P_{nm}(\text{sen}\,\theta') \text{ sen } m\lambda \text{ sen } \theta' d\lambda d\theta' \quad (20)$$

Por lo tanto, la expresión deseable para la perturbación de la gravedad es

$$\delta g(r, \theta', \lambda) = \sum_{n=0}^{+\infty} \sum_{m=0}^{norte} \frac{R}{r}^{\frac{1+\sqrt{9+4n(n+1)}}{2}} [A_{nm}^{\delta g} \delta g P_{nm}(\sin\theta') \cos m\lambda + B_{nm}^{\delta g} P_{nm}(\sin\theta') \sin m\lambda] \quad (21)$$

La anomalía de la gravedad también se puede identificar mediante un problema de Dirichlet similar si se conocen los valores de la anomalía de la gravedad Δg en la superficie del geoide.

$$\begin{aligned} &2\,(\Delta g) - \frac{2}{r} 2\,\Delta g = 0 \\ &\Delta g S = f2(\theta', \lambda) \end{aligned} \quad (22)$$

La expresión de la anomalía de la gravedad es

Machine Translated by Google

$$\Delta g(r,\theta',\lambda) = \sum_{n=0}^{+\infty}\sum_{m=0}^{n}\left(\frac{R}{r}\right)^{\frac{1+\sqrt{9+4n(n+1)}}{2}}\left[A_{nm}^{\Delta g}P_{nm}(\sin\theta')\cos m\lambda + B_{nm}P_{nm}(\sin\theta')\sin m\lambda\right] \quad (23)$$

Ahora estamos listos para encontrar el potencial perturbador. La función T será determinada extraído [22] de la siguiente ecuación diferencial parcial:

$$\frac{\partial T}{\partial r} = -\delta g$$

$$\frac{\partial T}{\partial r} = -\sum_{n=0}^{+\infty}\sum_{m=0}^{n}\left(\frac{R}{r}\right)^{\frac{1+\sqrt{9+4n(n+1)}}{2}}\delta g_{nm}\left[A_{nm}P_{nm}(\sin\theta')\cos m\lambda + B_{nm}P_{nm}(\sin\theta')\sin m\lambda\right] \quad (24)$$

La solución de esta ecuación diferencial parcial es

$$T(r,\theta',\lambda) = \frac{2R}{\sqrt{9+4n(n+1)}-1}\left(\frac{R}{r}\right)^{\frac{\sqrt{9+4n(n+1)}-1}{2}}\sum_{m=0}^{n}\left[A_{nm}^{\delta g}P_{nm}(\sin\theta')\cos m\lambda + B_{nm}^{\delta g}P_{nm}(\sin\theta')\sin m\lambda\right] + f_{\delta g}(\theta',\lambda) \quad (25)$$

La función $f_{\delta g}$ es una función arbitraria, pero se puede determinar a partir de la siguiente condición de contorno [22]:

$$2\,\delta g = \Delta g + \frac{}{R}\,t \quad (26)$$

Sustituyendo las cantidades necesarias de las ecuaciones (21), (23) y (25) tenemos

$$\sum_{n=0}^{+\infty}\sum_{m=0}^{n}\left[A_{nm}^{\delta g}P_{nm}(\sin\theta')\cos m\lambda + B_{nm}^{\delta g}P_{nm}(\sin\theta')\sin m\lambda\right] = \sum_{m=0}^{+\infty}\sum_{n=0}^{n}\left[A_{nm}^{\Delta g}P_{nm}(\sin\theta')\cos m\lambda + B_{nm}^{\Delta g}P_{nm}(\sin\theta')\sin m\lambda\right] + \frac{4}{\sqrt{9+4n(n+1)}-1}\sum_{m=0}^{n}\left[A_{nm}^{\delta g}P_{nm}(\sin\theta')\cos m\lambda + B_{nm}^{\delta g}P_{nm}(\sin\theta')\sin m\lambda\right] + \frac{2}{R}f_{\delta g}(\theta',\lambda) \quad (27)$$

Por tanto, la función $f_{\delta g}$ se expresa como una serie de armónicos esféricos.

$$f_{\delta g}(\theta',\lambda) = \sum_{n=0}^{+\infty}\sum_{m=0}^{n}\left[A_{nm}P_{nm}(\sin\theta')\cos m\lambda + B_{nm}P_{nm}(\sin\theta')\sin m\lambda\right] \quad (28)$$

Los coeficientes de la serie anterior son iguales a

$$A_{nm} = \frac{R}{2}\,\frac{\sqrt{9+4n(n+1)}-5}{9+4n(n+1)-1}\left(A_{\text{Nuevo Méjico}}^{\delta g} - UN_{\text{Nuevo Méjico}}^{\Delta g}\right) \quad (29)$$

$$B_{nm} = \frac{R}{2}\,\frac{\sqrt{9+4n(n+1)}-5}{segundo\ 9+4n(n+1)-1}\left(\delta g_{\text{Nuevo Méjico}} - B_{\text{Nuevo Méjico}}^{\Delta g}\right) \quad (30)$$

Finalmente, el potencial perturbador T viene dado como

$$T(r,\theta',\lambda) = \frac{2R}{\sqrt{9+4n(n+1)}-1}\left(\frac{R}{r}\right)^{\frac{\sqrt{9+4n(n+1)}-1}{2}}\sum_{m=0}^{n}\left[A_{nm}^{\delta g}P_{nm}(\sin\theta')\cos m\lambda + B_{nm}P_{nm}(\sin\theta')\sin m\lambda\right] + \sum_{m=0}^{n=0}\sum_{n=0}^{+\infty}\frac{R}{2+\sqrt{9+4n(n+1)}-5}\left(A_{\text{Nuevo Méjico}}^{\delta g} - UN_{nm}^{\Delta g}\right)P_{nm}(\sin\theta')\cos m\lambda + \frac{\sqrt{9+4n(n+1)}-5}{\sqrt{9+4n(n+1)}-1}\left(B_{\text{Nuevo Méjico}} - B_{nm}\right)P_{nm}(\sin\theta')\,\text{sen}\,m\lambda \quad (31)$$

Machine Translated by Google

El procedimiento para encontrar la ecuación (31) puede denominarse gravimétrico alternativo. determinación del potencial perturbador.

La expresión del potencial perturbador dada por la Ecuación (31) se puede reemplazar con una serie finita de armónicos esféricos para desarrollar un nuevo Modelo Geopotencial Global. Siguiendo la terminología geodésica, en este modelo el potencial perturbador es una función de las perturbaciones gravitatorias. Para n = Nmax, la Ecuación (31) se convierte en una serie finita y nos permite determinar las derivadas parciales de tercer orden del potencial perturbador en coordenadas esféricas. Estos derivados pueden expresarse en diferentes sistemas cartesianos [25] y utilizarse eficazmente en diversos estudios y exploraciones geofísicas.

En la Ecuación (31), para r = R, el potencial perturbador se determina en la superficie del geoide. El resultado es el siguiente:

$$T(R, \theta', \lambda) \equiv T(\theta', \lambda) = \frac{R}{2} \sum_{n=2}^{n_{max}} \sum_{m=0}^{n} [A_{nm}\delta g P_{nm}(\sin\theta') \cos m\lambda + B_{nm}\delta g P_{nm}(\sin\theta') \sin m\lambda] - 2\Delta g \Delta g \\ -[A_{nm} P_{nm}(\sin\theta') \cos m\lambda + B_{nm} P_{nm}(\sin\theta') \sin m\lambda] \quad (32)$$

La fórmula anterior se puede escribir como

$$T(\theta', \lambda) = \frac{R}{2} [\delta g(\theta', \lambda) - \Delta g(\theta', \lambda)] \quad (33)$$

Por lo tanto, en la superficie del geoide, el potencial perturbador es igual a la diferencia entre la perturbación de la gravedad y la anomalía de la gravedad, multiplicada por R/2. Utilizando la ecuación (6), la ondulación del geoide se puede expresar en una nueva forma alternativa.

$$N(\theta', \lambda) = \frac{R}{2} [\delta g(\theta', \lambda) - \Delta g(\theta', \lambda)] \, 2\gamma(\theta', \lambda) \quad (34)$$

Las ecuaciones (33) y (34) también representan la importancia de la perturbación de la gravedad y la anomalía de la gravedad (determinadas por mediciones). Se produce un resultado teórico interesante si δg y Δg se sustituyen por sus iguales.

$$T(r, \theta', \lambda) = \frac{R}{2} [\gamma_{ellepsoide}(\theta', \lambda) - \gamma_{geoide}(r, \theta', \lambda)] \quad (35)$$

$$norte(r, \theta', \lambda) = \frac{R}{2} \frac{\gamma_{ellepsoide}(\theta', \lambda) - \gamma_{geoide}(r, \theta', \lambda)}{\gamma_{ellepsoide}(\theta', \lambda)} \quad (36)$$

Por lo tanto, si se conoce el valor de la gravedad normal γ en la superficie del geoide, entonces se conocen el potencial perturbador (en esta superficie) y la ondulación del geoide (nuevamente como una serie de armónicos esféricos).

La intensidad de la gravedad g es una cantidad muy significativa que puede expresarse como serie de armónicos esféricos.

$$g(r, \theta', \lambda) = \sum_{n=0}^{+\infty} \sum_{m=0}^{n_{max}} \frac{R}{r}^{\frac{1+\sqrt{9+4n(n+1)}}{2}} [A_{nm} P_{nm}(\sin\theta') \cos m\lambda + B_{nm} P_{nm}(\sin\theta') \sin m\lambda] \quad (37)$$

Esta fórmula es mucho más manejable que la fórmula típica.

$$g(r, \theta', \lambda) = \text{gradVE} \quad (38)$$

donde VE representa el potencial de gravedad newtoniano de la Tierra. La ecuación (33) también es valiosa para diversas conversiones entre diferentes tipos de alturas, como la corrección de altura ortométrica y la corrección dinámica.

Machine Translated by Google

Como aproximación esférica, el gradiente vertical de la gravedad es igual a la derivada radial de la intensidad de la gravedad; así [22] (recordamos que estamos ante una Tierra que no gira),

$$\frac{\partial g}{} = -2gJ \quad \partial r \tag{39}$$

donde J es la curvatura media de la superficie equipotencial relativa.

$$\frac{\partial g}{\partial r} = \sum_{n=0}^{+\infty} \sum_{m=0}^{n} - \frac{1+\sqrt{9+4n(n+1)}}{2r} \left(\frac{R}{r}\right)^{\frac{1+\sqrt{9+4n(n+1)}}{2}} [A_{nm} P_{nm}(\sin\theta') \cos m\lambda + B_{nm} P_{nm}(\sin\theta') \sin m\lambda], \quad r > r \tag{40}$$

Combinando las ecuaciones (39) y (40), podemos determinar la curvatura media de las superficies equipotenciales del campo de gravedad de la Tierra.

4. Estudio de las superficies isogravitacionales de la Tierra.

La ecuación (37) nos permite investigar las superficies isogravitacionales del campo de gravedad de la Tierra. En estas superficies, el valor de la gravedad permanece constante. Como se indica en [26], las superficies isogravitacionales se parecen mucho a las esferas. Para realizar un estudio localizado de las superficies isogravitacionales, iniciamos la siguiente transformación:

$$\begin{matrix} X = r \cos\theta' \cos\lambda \\ Y = r \cos\theta' \sin\lambda \\ Z = r \sin\theta' \end{matrix}, \quad \theta' = \frac{\pi}{2} - \theta \tag{41}$$

El sistema cartesiano (X, Y, Z) [22] tiene su centro en el centro de gravedad de la Tierra, el eje Z coincide con el eje medio de rotación de la Tierra, el eje X está en el plano meridiano de Greenwich apuntando hacia afuera, y el eje Y hace que el sistema sea diestro. Entonces,

$$\frac{\partial(X,Y,Z)}{\partial(r,\theta',\lambda)} = \begin{matrix} \cos\theta'\cos\lambda & \cos\theta'\sin\lambda & \sin\theta' \\ -r\sin\theta'\cos\lambda & -r\sin\theta'\sin\lambda & r\cos\theta' \\ -r\cos\theta'\sin\lambda & r\cos\theta'\cos\lambda & 0 \end{matrix} \tag{42}$$

Los vectores unitarios normales son

$$\bar{\varepsilon}1 = (\cos\theta'\cos\lambda, \cos\theta'\sin\lambda, \sin\theta') \tag{43}$$

$$\bar{\varepsilon}2 = (-\sin\theta'\cos\lambda, -\sin\theta'\sin\lambda, \cos\theta') \quad \varepsilon3 = (-\sin \tag{44}$$

$$\bar{\lambda}, \cos\lambda, 0) \tag{45}$$

Sea P un punto arbitrario con coordenadas esféricas $(r_P, \theta_P', \lambda_P)$. Definimos un segundo sistema cartesiano (x, y, z), tal que su centro está en el punto P y el eje z tiene la misma dirección que el vector ε1 (Ecuación (43)). El eje x es tangente al meridiano de una esfera de radio rP y centro de gravedad de la Tierra, y tiene la misma dirección que el vector ε2. El eje y hace que el sistema sea diestro. La transformación entre (X, Y, Z) y (x, y, z) [27] es

$$\begin{matrix} \text{incógnita} \\ Y \\ z \end{matrix} = \begin{matrix} -\operatorname{sen}\theta'\cos\lambda\operatorname{sen}\lambda & \cos\theta'\cos\lambda & \text{—cógnita} \\ -\operatorname{sen}\theta'\operatorname{sen}\lambda & -\cos\lambda\cos\theta'\operatorname{sen}\lambda & \\ \text{porque } \theta' & 0 & \text{pecado } \theta' \end{matrix} \begin{matrix} \\ \\ \text{PAG} \end{matrix} \begin{matrix} \\ \\ yz + rP \end{matrix} \tag{46}$$

Las derivadas parciales de primer y segundo orden de la intensidad de la gravedad expresadas en el sistema (x, y, z) [27] se dan como sigue:

$$\frac{\partial g}{\partial x} = - \quad \frac{1}{} \frac{\partial g}{\partial \theta' r} \tag{47}$$

Machine Translated by Google

$$\frac{\partial g}{\partial y} = -\frac{1}{\text{porque } \theta' \, \partial \lambda} \frac{\partial gr}{} \tag{48}$$

$$\frac{\partial g}{\partial z} = \frac{\partial g}{\partial r} \tag{49}$$

$$\frac{\partial\, 2g}{\partial x\, 2} = -\frac{1}{}\frac{\partial g}{\partial r r} + \frac{1}{2\, r}\frac{\partial\, 2g}{\partial \theta'\, 2} \tag{50}$$

$$\frac{\partial}{2g\ \partial x \partial y} = \frac{1}{{}^{2}{}_{\partial \theta' \partial \lambda \cos \theta' \, r}}\frac{\partial\, 2g}{} - \frac{\text{pecado } \theta'}{r\, 2 \cos 2 \, \theta' \, \partial \lambda}\frac{\partial g}{} \tag{51}$$

$$\frac{\partial\, 2g}{\partial x \partial z} = \frac{1}{{}^{2}{}_{\partial \theta' \, r}}\frac{\partial g}{} - \frac{1}{}\frac{\partial g}{\partial r \partial \theta' r} \tag{52}$$

$$\frac{\partial\, 2g}{\partial y^{\,2}} = -\frac{1}{}\frac{\partial g}{\partial r r} + \frac{1}{r \text{ cuna } \theta' \, \partial \theta'}\frac{\partial g}{} + 2\,\frac{1}{r\, 2g\ 2 \cos 2 \, \theta' \, \partial \lambda 2}\frac{\partial}{} \tag{53}$$

$$\frac{\partial}{2g\ \partial y \partial z} = \frac{1}{{}^{2}{}_{\partial \lambda \cos \theta' \, r}}\frac{\partial g}{} - \frac{1}{\text{porque } \theta' \, \partial r \partial \lambda}\frac{\partial gr}{} \tag{54}$$

$$\frac{\partial\, 2g}{\partial z\, 2} = \frac{\partial\, 2g}{\partial r\, 2} \tag{55}$$

Las ecuaciones anteriores (Ecuación (47) a Ecuación (55)) se conocen como punto P. Debido a la derivada parcial distinta de cero de g con respecto a r, se requiere una parametrización de la superficie isogravitacional [21] para una pequeña región cercana al punto P es

$$\bar{s} : (-\varepsilon,\varepsilon) \times (-\varepsilon,\varepsilon) \rightarrow \quad 3 : (x, y) \rightarrow s(x, y) = (x, \bar{y}, z(x, y)) \tag{56}$$

La función z no se conoce, pero se pueden determinar sus derivadas parciales. Por tanto, los vectores de coordenadas son

$$\frac{\partial \bar{s}}{\partial x} = 1, 0, \partial x \frac{\partial z}{} = 1, 0, -\partial g \, \partial z \frac{\frac{\partial g}{\partial x}}{} \tag{57}$$

$$\frac{\partial \bar{s}}{\partial y} = 1, 0, \partial y \frac{\partial z}{} = 1, 0, \frac{\frac{\partial g}{\partial y}}{\frac{\partial g}{\partial z}} \tag{58}$$

El vector normal es

$$\overline{NC} = \frac{\partial \bar{s}}{\partial x} \times \frac{\partial \bar{s}}{\partial y} = \frac{\frac{\partial g}{\partial x}}{\frac{\partial g}{\partial z}}, \frac{\frac{\partial g}{\partial y}}{\frac{\partial g}{\partial z}}, 1 \tag{59}$$

El vector normal unitario es

$$\overline{\text{norte}} = \frac{\partial g}{\partial z}\frac{1}{\text{gradg}}\overline{NC} \tag{60}$$

Los elementos fundamentales [21] del primer tipo son

$$mi = \frac{\frac{\partial g}{\partial x}\ 2 + \frac{\partial g}{\partial z}^{2}}{\frac{\partial g}{\partial z}^{2}} \tag{61}$$

$$F = \frac{\frac{\partial g}{\partial x}\frac{\partial g}{\partial y\, 2}}{\frac{\partial g}{\partial z}} \tag{62}$$

Machine Translated by Google

$$\text{GRAMO} = \frac{\left(\frac{\partial g}{\partial y}\right)^2 + \left(\frac{\partial g}{\partial z}\right)^2}{\left(\frac{\partial g}{\partial z}\right)^2} \tag{63}$$

Los elementos fundamentales del segundo tipo son

$$L = \overline{N}, \frac{\partial^2 \overline{s}}{\partial x^2} \frac{\partial g}{\partial z} \frac{-\frac{\partial^2 g}{\partial x^2}\frac{\partial g}{\partial z} \frac{2\partial^2 g}{\partial z}\frac{\partial g}{\partial x}\frac{\partial g}{\partial z} + 2\frac{\partial^2 g}{\partial x \partial z}\frac{\partial g}{\partial x}\frac{\partial g}{\partial z} - \frac{\partial^2 g}{\partial z^2}\left(\frac{\partial g}{\partial x}\right)^2}{\left(\frac{\partial g}{\partial z}\right)^3} \text{ grado} \tag{64}$$

$$M = \overline{N}, \frac{\partial^2 \overline{s}}{\partial x \partial y} \frac{\partial g}{\partial z} \frac{\frac{\partial^2 g}{\partial x \partial z}\frac{\partial g}{\partial y}\frac{\partial g}{\partial z} + \frac{\partial^2 g}{\partial y \partial z}\frac{\partial g}{\partial x}\frac{\partial g}{\partial z} + \frac{\partial^2 g}{\partial x \partial y}\left(\frac{\partial g}{\partial z}\right)^2 - \frac{\partial^2 g}{\partial x \partial y}\frac{\partial g}{\partial z}\frac{\partial g}{\partial z}}{\left(\frac{\partial g}{\partial z}\right)^3} \text{ grado} \tag{65}$$

$$\text{norte} = \overline{\text{norte}, \partial y}^{\partial^2 \overline{s}} = \frac{\partial g}{\partial z} \frac{-\frac{\partial^2 \text{gramos}}{\partial y^2}\frac{\partial g}{\partial z} + \frac{\partial^2 g}{\partial y \partial z}\frac{\partial g}{\partial y}\frac{\partial g}{\partial z} - \frac{\partial^2 g}{\partial y \partial z}\frac{\partial g}{\partial y}\frac{\partial g}{\partial z} + \frac{\partial^2 g}{\partial z^2}\left(\frac{\partial g}{\partial y}\right)^2}{\left(\frac{\partial g}{\partial z}\right)^3} \text{ grado} \tag{66}$$

Sea P un punto alejado de la Tierra, es decir, rP toma un valor grande. A medida que nos acercamos lentamente a la Tierra a lo largo del eje polar r = rP, gx y gy son cantidades pequeñas. Por lo tanto, la ecuación (61) a la ecuación (66) se convierte en

$$PE = 1 \tag{67}$$

$$PF = 0 \tag{68}$$

$$PG = 1 \tag{69}$$

$$LP = \frac{\frac{\partial^2 g}{\partial x^2}}{\text{gradg}}_{PAG} = -\frac{\frac{\partial^2 g}{\partial x^2}}{\frac{\partial g}{\partial z}}_{PAG} = -\frac{1}{RP} - \frac{1}{r^2_{PAG}}\frac{\frac{\partial^2 g}{\partial \theta'}}{\frac{\partial g}{\partial r}}_{PAG} \tag{70}$$

$$MP = \frac{\frac{\partial^2 g}{\partial x \partial y}}{\text{grado}}_{PAG} = -\frac{\frac{\partial^2 g}{\partial x \partial y}}{\frac{\partial g}{\partial z}}_{PAG} = -\frac{1}{r^2_{PAG}\,\text{porque}\,\theta P'}\frac{\frac{\partial^2 g}{\partial \theta' \partial \lambda}}{\frac{\partial g}{\partial r}}_{PAG} + \frac{\text{pecado}\,\theta P'}{r^2_{PAG}\cos 2\,\theta P'\,\partial r\,P}\frac{\frac{\partial g}{\partial \lambda}}{\frac{\partial g}{\partial r}} \tag{71}$$

$$NP = \frac{\frac{\partial^2 \text{gramos}}{\partial y^2}}{\text{gradg}}_{PAG} = -\frac{\frac{\partial^2 \text{gramos}}{\partial y^2}}{\frac{\partial g}{\partial z}}_{PAG} = -\frac{1}{RP} + \frac{1}{r^2_{PAG}\,\text{cuna}\,\theta P'}\frac{\frac{\partial g}{\partial \theta'}}{\frac{\partial g}{\partial r}\,P} + \frac{1}{r^2_{PAG}\cos 2\,\theta P'}\frac{\frac{\partial^2 g}{\partial \lambda^2}}{\frac{\partial g}{\partial r}}_{PAG} \tag{72}$$

En el punto P, las curvaturas principales, k1 y k2, de la superficie isogravitacional [28] son las raíces de la siguiente ecuación:

$$\begin{aligned}(EG - F^2)P\,k^2 - (EN - 2FM + GL)P\,k + (LN - M2)P &= 0\\ k^2 - (L + N)P\,k + (LN - M2)P &= 0\end{aligned} \tag{73}$$

Por eso,

$$k1 = \frac{L + N + (L - N)\sqrt{\quad^2 - 4M2}}{2}_{PAG} \tag{74}$$

$$k2 = \frac{L + N - (L - N)\sqrt{\quad^2 - 4M2}}{2}_{PAG} \tag{75}$$

Machine Translated by Google

Las curvaturas media y gaussiana de la superficie isogravitacional en el punto P son iguales a

$$H = \frac{L+N}{2}_{PAG} = \frac{1}{2}\left[\frac{1}{RP} - \frac{1}{2\,r_{PAG}}\frac{\partial^2 g}{\partial\theta'^2}\Big/\frac{\partial g}{\partial r}_{PAG} - \frac{1}{RP} + \frac{1}{2\,r_{PAG}\,\text{cuna}\,\theta P'}\frac{\partial g}{\partial\theta'}\Big/\frac{\partial g}{\partial r}P + \frac{1}{2\,r_{PAG}\cos 2\,\theta P'}\frac{\partial^2 g}{\partial\lambda 2}\Big/\frac{\partial g}{\partial r}_{PAG}\right] \quad (76)$$

$$\text{kilos} = \frac{\text{LN-M2}}{\text{GE-F 2}}_{PAG} = -\left[\frac{1}{rp} - \frac{1}{2r_{PAG}}\frac{\partial^2 g}{\partial\theta'^2}\Big/\frac{\partial g}{\partial r}_{PAG} - \frac{1}{rp} + \frac{1}{2\,r_{PAG}\,\text{cuna}\,\theta P'}\frac{\partial g}{\partial\theta}\Big/\frac{\partial g}{\partial r}P\,2 + \frac{1}{2\,r_{PAG}\cos 2\,\theta P'}\frac{\partial^2 g}{\partial\lambda 2}\Big/\frac{\partial g}{\partial r}_{PAG} - \frac{1}{r^2_{PAG}\,\text{porque}\,\theta P'}\frac{P\partial^2 g}{\partial\theta'\partial\lambda}\Big/\frac{\partial g}{\partial r}_{PAG} + \frac{\text{pecado}\,\theta P'}{r^2_{PAG}\cos 2\,\theta P'}\frac{\partial g}{\partial\lambda}\Big/\frac{\partial g}{\partial r}P\right] \quad (77)$$

Si el punto P está lejos de la Tierra, las ecuaciones (76) y (77) se convierten en

$$H = -\frac{1}{rP} \quad (78)$$

$$\text{kilos} = \frac{1}{2\,r_{PAG}} \quad (79)$$

Por lo tanto, para valores grandes de r, las superficies isogravitacionales son esferas de radio rP.

Acercándose a la Tierra a lo largo del eje polar r = rP, la curvatura media viene dada por la ecuación (76) y la curvatura de Gauss se deriva de la siguiente relación:

$$k^{r>}_{GNAG} = \frac{1}{2\,r_{PAG}} + \frac{1}{3r_{PAG}}\left[\frac{1}{\text{cuna}\,\theta P'}\frac{\partial g}{\partial\theta'}\Big/\frac{\partial g}{\partial r}P - \frac{1}{\cos 2\,\theta P'}\frac{\partial^2 g}{\partial\lambda 2}\Big/\frac{\partial g}{\partial r}_{PAG} + \frac{\partial^2 g}{\partial\theta'^2}\Big/\frac{\partial g}{\partial r}_{PAG}\right] \quad (80)$$

La curvatura de Gauss en el punto P caracteriza la forma de la superficie isogravitacional alrededor de ese punto. El segundo término de la ecuación (80) describe la desviación de la forma de la superficie isogravitacional de la de una esfera. A medida que el punto P se acerca gradualmente a la Tierra (con un pequeño δr), las superficies isogravitacionales forman superficies cerradas con una curvatura de Gauss positiva. Alrededor del punto P, la forma de la superficie isogravitacional es un paraboloide elíptico.

La Figura 1 muestra la forma de la superficie isogravitacional alrededor del punto P, que es un paraboloide. La forma de las superficies isogravitacionales se vuelve más complicada cuando el punto P está cerca de la Tierra. Suponiendo que el geoide representa la verdadera forma de la Tierra, sea P un punto en el geoide, particularmente en el Polo Norte (θ' = π/2). La superficie isogravitacional que pasa por el punto P no puede cruzar el geoide. Por tanto, ambas superficies son tangentes en el punto P, es decir, tienen un plano tangente común. Alrededor del punto P, las coordenadas z de los puntos vecinos tienen signo positivo.

Sea γ una curva en la superficie isogravitacional que se forma por la intersección de la superficie isogravitacional y un plano meridiano. Dado que las superficies isogravitacionales son superficies cerradas, esto significa que a medida que nos alejamos del punto P a lo largo de la curva γ, llegamos a un punto A donde la curvatura de Gauss es positiva y los puntos vecinos de la superficie isogravitacional tienen coordenadas z negativas. En consecuencia, en la curva γ entre el punto P y A, hay un punto, Q, donde la curvatura de Gauss es igual a cero y las coordenadas z de los puntos vecinos también son iguales a cero. Esto significa que el punto Q es un punto singular, es decir, LQ = MQ = NQ = 0.

La superficie isogravitacional que pasa por el punto P intersecta el plano ecuatorial y se encuentra fuera del geoide, sin puntos singulares. El conjunto de puntos que tienen una curvatura de Gauss cero se define como un anillo singular de la superficie isogravitacional. Cerca de la superficie del geoide, cada superficie isogravitacional tiene al menos un anillo singular.

La figura 2 muestra una representación simplificada de un anillo singular. A medida que el punto P se mueve a lo largo del eje z, rP aumenta. Hemos demostrado que a medida que la distancia polar r alcanza valores elevados, las superficies isogravitacionales se convierten en esferas. Esto indica que la curvatura de Gauss en el punto P se acerca a cero. Simultáneamente, a medida que los anillos singulares se reducen lentamente,

Machine Translated by Google

alcanzan un punto crítico cuando la curvatura de Gauss en el punto P se vuelve cero. El punto P representa la ubicación del anillo singular en esta superficie. Si rP supera el valor crítico, las superficies isogravitacionales exhibirán una curvatura de Gauss positiva en todas partes, lo que resultará en la ausencia de anillos singulares. Finalmente, a medida que rP adquiere valores grandes, las superficies isogravitacionales se convierten en esferas con un radio de rP.

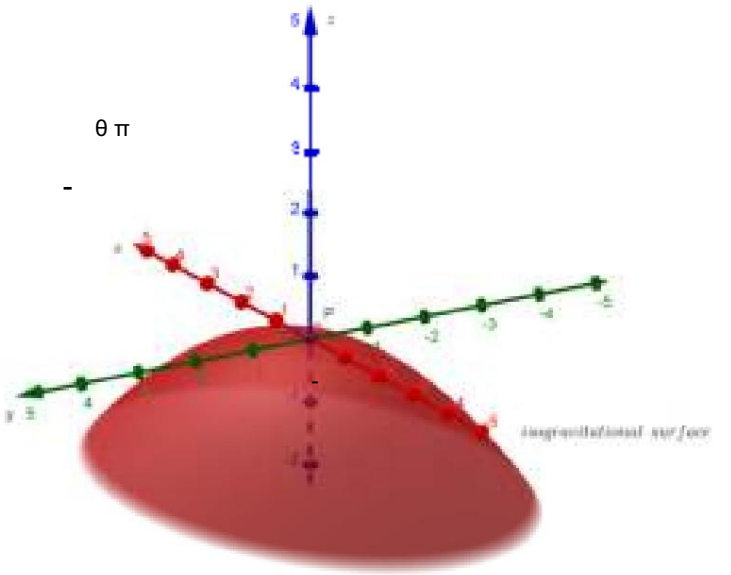


Figura 1. Superficie isogravitacional.

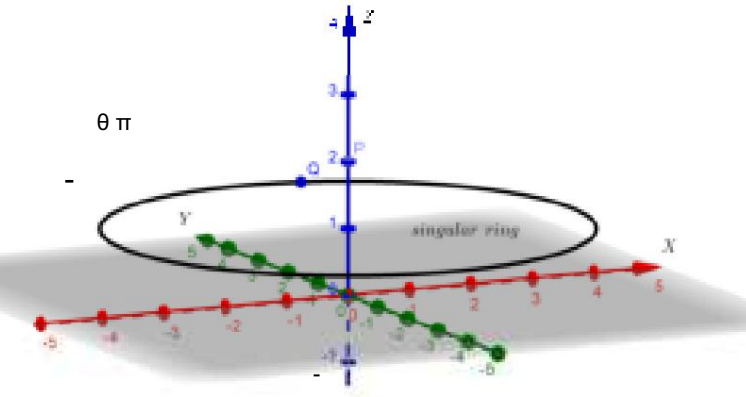


Figura 2. Anillo Singular.

La Figura 3 muestra una imagen simplificada de una familia de anillos singulares.

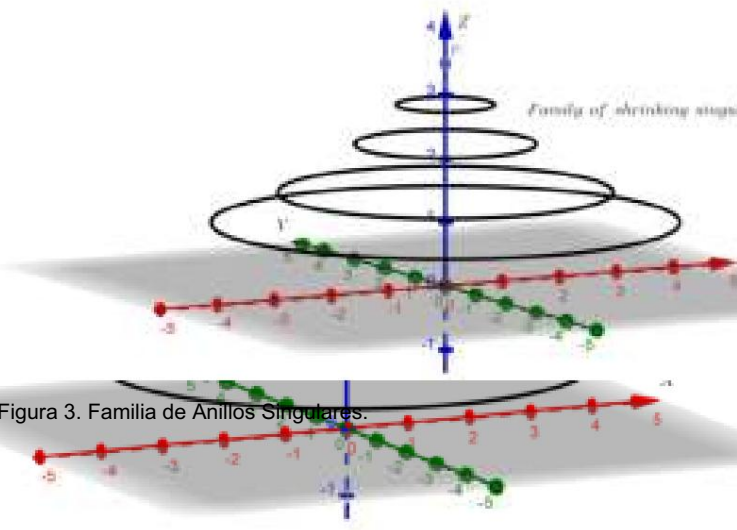


Figura 3. Familia de Anillos Singulares.

Machine Translated by Google

5. Superficies isogravitacionales perturbadoras, superficies isogravitacionales anómalas y una Comparación entre los dos métodos para determinar el potencial perturbador

Una superficie isogravitacional perturbadora se caracteriza por un valor constante de gravedad. perturbación δg. De manera similar, una superficie isogravitacional anómala se caracteriza por un valor constante de anomalía de gravedad ∆g. Este estudio local es similar al del estudio isogravitacional. superficies, donde las curvaturas media y gaussiana están dadas por las ecuaciones (76) y (77), respectivamente, con g reemplazado por δg (o ∆g).

En la siguiente tabla, presentamos un resumen de los hallazgos de las Secciones 3 y 4. La primera columna enumera las cantidades gravitacionales, mientras que la segunda columna indica si cada cantidad se expresa como una serie de armónicos esféricos proporcionada por Laplace ecuación. Usamos "sí" para indicar cuando la cantidad se expresa explícitamente como tal y "no" cuando no lo es. En la tercera columna, se selecciona "sí" cuando se representa la cantidad como una serie de armónicos esféricos obtenidos de la ecuación de Helmholtz modificada con G. El El término "directo" se utiliza para describir una cantidad que satisface el diferencial parcial elegido. ecuación y permanece mensurable o constante en el límite. Por el contrario, nosotros utilizar el término "indirecto" cuando no se cumplan las condiciones mencionadas anteriormente. Además, Los símbolos "Ve" y "VE" representan la gravitación newtoniana de la Tierra y el elipsoide. potencial, respectivamente.

De la Tabla 1, podemos observar que la ecuación de Laplace ofrece una ventaja en la análisis del potencial gravitacional, mientras que la ecuación de Helmholtz modificada con G tiene un ventaja en el estudio de la intensidad de la gravedad, la perturbación de la gravedad y la anomalía de la gravedad. Es Vale la pena señalar que el proceso de determinación del potencial perturbador utilizando el G-modificado La ecuación de Helmholtz es bastante sencilla, ya que sólo implica resolver una ecuación parcial simple. ecuación diferencial (ver Ecuación (24)).

Tabla 1. Comparación entre Laplace y G – ecuación de Helmholtz modificada.

Cantidad	Ecuación de Laplace esférica Armónicos, distancia radial en potencias racionales	Helmholtz modificado con G Ecuación, esférica Armónicos, distancia radial en poderes irracionales
Potencial perturbador T	Sí, determinación indirecta	Sí, determinación indirecta
Potencial de gravedad del elipsoide Ve	Sí, determinación directa *	No
Potencial de gravedad de la Tierra VE	Sí, determinación indirecta	No
Perturbación de la gravedad δg	Sí, determinación indirecta	Sí, determinación directa.
Anomalía de gravedad ∆g	Sí, determinación indirecta	Sí, determinación directa.
Intensidad de gravedad g	No	Sí, determinación directa.
Intensidad de gravedad del elipsoide γ	No	Sí, determinación directa.
gradiente vertical de g	No	Sí, determinación indirecta
gradiente vertical de γ	No	Sí, determinación indirecta
Componente ξ	Sí, determinación indirecta	Sí, determinación indirecta
Componente η	Sí, determinación indirecta	Sí, determinación indirecta
Ondulación geoide N	Sí, determinación indirecta	Sí, determinación indirecta

* Para más detalles, nos referimos a [22].

6. Resumen de los resultados

En esta sección, tabulamos los nuevos resultados que se presentaron en la Sección 3.

La Tabla 2 representa la importancia de la ecuación de Helmholtz modificada con G, que nos permite estudiar cantidades gravitacionales independientemente del potencial de gravedad y expresar ellos como un nuevo tipo de serie de armónicos esféricos. Elegimos expresar lo más significativo. fórmulas en detalle, que están relacionadas con la gravedad g, la perturbación de la gravedad δg, la anomalía de la gravedad ∆g, potencial perturbador T, ondulación geoide N, gradiente vertical de gravedad y normal

Machine Translated by Google

gravedad en el geoide (marcado en negrita). Además, la Tabla 2 también incluye los componentes de la deflexión de la vertical.

Tabla 2. Se produjeron nuevas expresiones armónicas esféricas a partir de G: ecuación de Helmholtz modificada.

Cantidad	Expresión
gravedad g	$g(r, \theta', \lambda) = \sum_{n=0}^{+\infty} \sum_{m=0}^{n} \left(\frac{r}{r}\right)^{\frac{1+\sqrt{9+4n(n+1)}}{2}} [A_{nm} P_{nm}(\text{sen } \theta') \cos m\lambda + B_{nm}^{g} P_{nm}(\sin \theta') \sin m\lambda]$
Perturbación de la gravedad δg	$\delta g(r, \theta', \lambda) = \sum_{n=0}^{+\infty} \sum_{m=0}^{n} \left(\frac{R}{r}\right)^{\frac{1+\sqrt{9+4n(n+1)}}{2}} [A_{nm}^{\delta g} P_{nm}(\sin \theta') \cos m\lambda + B_{nm}^{\delta g} P_{nm}(\sin \theta') \sin m\lambda]$
Anomalía de gravedad Δg	$\Delta g(r, \theta', \lambda) = \sum_{n=0}^{+\infty} \sum_{m=0}^{n} \left(\frac{R}{r}\right)^{\frac{1+\sqrt{9+4n(n+1)}}{2}} [A_{nm}^{\Delta g} P_{nm}(\sin \theta') \cos m\lambda + B_{nm}^{\Delta g} P_{nm}(\sin \theta') \sin m\lambda]$
Potencial perturbador T en el espacio tridimensional	$T(r, \theta', \lambda) = \sum_{n=0}^{+\infty} \frac{2R}{\sqrt{9+4n(n+1)-1}} \left(\frac{R}{r}\right)^{\frac{\sqrt{9+4n(n+1)-1}}{2}} \{ \sum_{m=0}^{n} [A_{nm}^{\delta g} P_{nm}(\sin \theta') \cos m\lambda + B_{nm}^{\delta g} P_{nm}(\sin \theta') \sin m\lambda] + \sum_{m=0}^{n} \frac{\sqrt{9+4n(n+1)-5}}{\sqrt{9+4n(n+1)-1}} A_{nm}^{\delta g} - UN_{nm}^{\Delta g} P_{nm}(\text{sen } \theta') \cos m\lambda + \frac{\sqrt{9+4n(n+1)-5}}{\sqrt{9+4n(n+1)-1}} B_{nm}^{\delta g} - B_{nm}^{\Delta g} P_{nm}(\text{sen } \theta') \text{ sen } m\lambda \}$
Potencial perturbador en el geoide	$T(\theta', \lambda) = \frac{R}{2} [A_{nm} P_{nm}(\sin \theta') \cos m\lambda + B_{nm} P_{nm}(\sin \theta') \sin m\lambda] - \Delta g - [U n_{nm}^{\Delta g} P_{nm}(\sin \theta') \text{ porque } m\lambda + B_{nm} P_{nm}(\sin \theta') \sin m\lambda] = \frac{R}{2} [\delta g(\theta', \lambda) - \Delta g(\theta', \lambda)], r = R$
Ondulación geoide N	$N(\theta', \lambda) = [\delta g(\theta', \lambda) - \Delta g(\theta', \lambda)] \frac{R}{2\gamma_{elipsoide}(\theta', \lambda) - \gamma_{geoide}(\theta', \lambda)}$ o $norte(r, \theta', \lambda) = \frac{R}{2} \frac{T(r, \theta', \lambda)}{\gamma_{elipsoide}(\theta', \lambda)}$
gradiente vertical de gravedad	$\frac{\partial g}{\partial r} = \sum_{n=0}^{+\infty} \sum_{m=0}^{n} - \frac{1+\sqrt{9+4n(n+1)}}{2r} \left(\frac{r}{r}\right)^{\frac{1+\sqrt{9+4n(n+1)}}{2}} [\bar{A}_{nm} P_{nm}(\sin \theta') \cos m\lambda + B_{nm} \bar{P}_{nm}(\sin \theta') \sin m\lambda], \quad r > r$
Componente ξ	$\lambda) = -r\gamma(r,\theta',\lambda) \frac{1}{\partial \theta'} \frac{\partial T}{\partial \xi(r, \theta',}$
Componente η	$= -r\gamma(r,\theta',\lambda) \text{ sen } \theta' \frac{1}{\partial \lambda} \frac{\partial T}{\eta(r, \theta', \lambda)}$
Gravedad normal en el geoide.	$\gamma_{geoide}(r, \theta', \lambda) = \sum_{n=0}^{n0} a_{2nr}^{-\frac{1+\sqrt{9+4n(n+1)}}{2}} \cdot \cdot P_{2n}(\text{sen } \theta') + e^{2} (2n + 1) \sin \theta' [P_{2n}(\sin \theta') \sin \theta' - P_{2n+1}(\sin \theta')]$

Junto con la ecuación de Laplace, la ecuación de Helmholtz modificada con G también puede ser clasificada como una ecuación diferencial parcial fundamental de la geodesia física.

7. Conclusiones

En este trabajo, presentamos nuevas fórmulas alternativas para varias cantidades relacionadas con el campo de gravedad de la Tierra.

La ecuación de Helmholtz modificada con G es una ecuación diferencial parcial que nos permite calcular la intensidad de la gravedad como una serie de armónicos esféricos. La intensidad de la gravedad juega un papel crucial en una amplia gama de aplicaciones. La serie armónica de intensidad de la gravedad es una serie en la que las potencias de la distancia radial son números irracionales. Esta serie nos permite representar el gradiente vertical de gravedad como una serie de armónicos esféricos. El gradiente vertical de gravedad es muy importante para numerosas aplicaciones geofísicas.

La perturbación de la gravedad y la anomalía de la gravedad son muy importantes. Se expresan como una serie de armónicos esféricos, como lo proporciona la ecuación de Laplace o la ecuación de Helmholtz modificada con G. En el primer escenario, su determinación no es sencilla, como

Machine Translated by Google

Tenemos que resolver un problema de valor límite de Neumann (problema de Hotine) o de Robin (problema de Stokes). En el segundo caso, sus determinaciones ocurren como soluciones a problemas de valores en la frontera de Dirichlet, que son más simples que los dos casos anteriores.

La ecuación de Helmholtz modificada con G ofrece una ventaja adicional: nos permite calcular el potencial perturbador de la Tierra utilizando una ecuación diferencial parcial simple. La solución de esta ecuación implica sólo la integración de la serie armónica de anomalía de la gravedad con respecto a la distancia radial, r. Además, la ecuación de Helmholtz modificada con G también se puede utilizar para encontrar las componentes ξ y η de la deflexión de la vertical y la ondulación del geoide.

Además, en la superficie del geoide, el potencial perturbador y la ondulación del geoide toman la forma elegante de una serie de armónicos esféricos, lo que muestra la dependencia de la diferencia entre la perturbación de la gravedad y la anomalía de la gravedad en esta superficie. Un resultado teórico interesante que se produce es que tanto el potencial perturbador como la ondulación del geoide se pueden determinar si se conoce el valor de la gravedad normal en la superficie del geoide.

Todas las nuevas fórmulas que surgen de la ecuación de Helmholtz modificada con G se tabulan en la Sección 6.

La serie armónica de intensidad gravitacional nos brinda una ventaja significativa en el estudio de las superficies isogravitacionales de la Tierra. Las derivadas parciales de primer y segundo orden del potencial gravitacional de la Tierra se pueden expresar convenientemente en sistemas cartesianos. Esto simplifica las expresiones de las derivadas parciales y las aproximaciones requeridas. Se muestra que las superficies isogravitacionales son superficies cerradas con una curvatura de Gauss no negativa.

En altitudes muy elevadas, cuando el valor de la distancia radial “r” es grande, las superficies isogravitacionales son esferas con un radio igual a la distancia radial seleccionada. A medida que nos acercamos a la Tierra, su apariencia se vuelve progresivamente más compleja. Cada superficie isogravitacional cercana a la Tierra tiene al menos un conjunto de puntos definidos como un anillo singular. Este nombre se atribuye a que todos los puntos pertenecientes a este conjunto son puntos singulares. La familia de anillos singulares converge en un único punto situado en el eje medio de rotación de la T

El trabajo futuro podría implicar el desarrollo de nuevos modelos de gravedad global para la intensidad de la gravedad g, la perturbación de la gravedad δg y la anomalía de la gravedad Δg. Un modelo de gravedad global para la intensidad de la gravedad representaría un avance significativo con una amplia gama de aplicaciones, como en los campos de la geofísica, la arqueología, la geodesia, la determinación de órbitas y la navegación inercial. Además, sería ventajoso desarrollar nuevos modelos de gravedad global para anomalías y perturbaciones de la gravedad, separados de los modelos geopotenciales globales. Esta evolución puede presentarnos oportunidades para mejorar las técnicas actuales para manipular datos gravitacionales, lo que conducirá a una comprensión más profunda del campo gravitatorio de la Tierra.

Financiamiento: Esta investigación no recibió financiamiento externo.

Declaración de la Junta de Revisión Institucional: El estudio no requirió aprobación ética.

Declaración de Consentimiento Informado: No aplicable.

Declaración de disponibilidad de datos: No se utilizaron datos.

Conflictos de intereses: El autor declara no tener conflictos de intereses.

Referencias

1. Jarosz, W.; Carr, NA; Jensen, HW Importancia del muestreo de armónicos esféricos. Computadora. Gráfico. Foro 2009, 28, 577–586. Disponible en línea: http://graphics.ucsd.edu/~henrik/papers/importance_sampling_spherical_harmonics.pdf (consultado el 2 de febrero de 2024). [Referencia cruzada]

2. Silvennoinen, A.; Sloan, PP Descomposición de bases móviles para transporte ligero precalculado. Computadora. Gráfico. Foro 2021, 40, 127–137. Disponible en línea: https://www.ppsloan.org/publications/mbd.pdf (consultado el 2 de febrero de 2024). [Referencia cruzada]

3. Wang, J.; Xu, K.; Zhou, K.; Lin, S.; Hu, S.; Guo, B. Escalado de armónicos esféricos. Vis. Computadora. 2006, 22, 713–720. Disponible en línea: https://link.springer.com/article/10.1007/s00371-006-0057-8 (consultado el 2 de febrero de 2024). [Referencia cruzada]

Machine Translated by Google

4. Annen, T.; Kautz, J.; Durand, F.; Seidel, gradientes armónicos esféricos HP para iluminación de rango medio. En Actas del 15º Taller de Eurografía sobre Técnicas de Renderizado, Norrköping, Suecia, 21 a 24 de junio de 2004; págs. 331–336. Disponible en línea: https://diglib.eg.org/bitstream/handle/10.2312/EGWR.EGSR04.331-336/331-336.pdf?sequence=1&isAllowed=y (consultado el 2 de febrero de 2024).

5. Sansó, F.; Sideris, M. Teoría y métodos de determinación de geoides; Springer Science & Business Media: Berlín, Alemania, 2013; Capítulo 6; Volumen 262, pág. 131. Disponible en línea: https://link.springer.com/book/10.1007/978-3-540-74700-0 (consultado el 2 de febrero de 2024).
6. Shouni, A.; Khalil, R.; Kamel, A.; Miky, Y. Evaluaciones de modelos geopotenciales globales recientes basados en datos de gravedad y nivelación GPS a lo largo de las zonas costeras de Egipto. Open Gosciences 2023, 15, 20220450. Disponible en línea: https://www.degruyter.com/document/doi/10.1515/geo-2022-0450/html (consultado el 2 de febrero de 2024). [Referencia cruzada]
7. Wu, Y.; Él, X.; Luo, Z.; Shi, H. Una evaluación de modelos geopotenciales globales de alto grado publicados recientemente basados en datos geodésicos y oceánicos heterogéneos. Frente. Ciencia de la Tierra. 2021, 9, 749611. Disponible en línea: https://www.frontiersin.org/ artículos/10.3389/feart.2021.749611/full (consultado el 2 de febrero de 2024). [Referencia cruzada]
8. Isik, MS; Çevikalp, MR; Erol, B.; Erol, S. Mejora de los modelos geopotenciales basados en GOCE para el modelado de geoides gravimétricos en Turquía. Geociencias 2022, 12, 432. Disponible en línea: https://www.mdpi.com/2076-3263/12/12/432 (consultado el 2 de febrero de 2024). [Referencia cruzada]
9. Elizondo, NA; López, RVG; Carillo, XGT; Becerra, GEV Combinación de modelos geopotenciales globales, modelos de elevación digitales y GNSS/nivelación para una determinación precisa de geoides locales en algunas áreas urbanas de México: estudio de caso. ISPRS Int. J. Geo-Inf. 2021, 10, 819. Disponible en línea: https://www.mdpi.com/2220-9964/10/12/819 (consultado el 2 de febrero de 2024). [Referencia cruzada]
10. Roithmayr, CM Contribuciones de los armónicos esféricos a los campos magnéticos y gravitacionales; NASA TM-2004-213007; Administración Nacional de Aeronáutica y del Espacio: Washington, DC, EE. UU., 2004. Disponible en línea: https://ntrs.nasa.gov/api/citations/2004 0047194/descargas/20040047194.pdf (consultado el 2 de febrero de 2024).
11. Campbell, WH Aplicaciones del geomagnetismo; Sircular 1109 del Servicio Geológico de Estados Unidos; Imprenta del Gobierno de los Estados Unidos: Washington, DC, EE. UU., 1995. Disponible en línea: https://pubs.usgs.gov/circ/1995/1109/report.pdf (consultado el 2 de febrero de 2024).
12. Prada, KO; Antelo, MJ; Murillo, NA; Cózar, JR; Contreras–de–Villar, F.; Pérez, JJM Un Nuevo Método para la Recolección de Información Geomagnética Marina: Aplicación de Levantamientos en el Caribe Colombiano. J. Mar. Ciencias. Ing. 2021, 9, 10. Disponible en línea: https://www.mdpi.com/2077-1312/9/1/10 (consultado el 2 de febrero de 2024). [Referencia cruzada]
13. Zigan, L. Descripción general de las aplicaciones del campo eléctrico en ingeniería de procesos y energía. Energías 2018, 11, 1361. Disponible online: https://www.mdpi.com/1996-1073/11/6/1361 (consultado el 15 de febrero de 2024). [Referencia cruzada]
14. Tognolatti, L.; Ponti, C.; Santarsiero, M.; Schettini, G. Una técnica computacional eficiente para la dispersión electromagnética por esferoides prolatos. Matemáticas 2022, 10, 1761. Disponible en línea: https://www.mdpi.com/2227-7390/10/10/1761 (consultado el 15 de febrero de 2024). [Referencia cruzada]
15. Vafeas, P.; Perrusson, G.; Lesselier, D. Dispersión de baja frecuencia de cuerpos esferoidales perfectamente conductores en un medio conductor con excitación dipolar magnética. Int. J. Ing. Ciencia. 2009, 47, 372–390. Disponible en línea: https: //www.researchgate.net/publication/222923079_Lowfrequency_scattering_fro_perfectly_conducting_spheroidal_bodies_in_ un_medio_conductivo_con_excitación_dipolo_magnética (consultado el 15 de febrero de 2024). [Referencia cruzada]
16. Perrusson, G.; Vafeas, P.; Chatjigeorgiou, IK; Lesselier, D. Identificación in situ de baja frecuencia de un cuerpo altamente conductor enterrado en la Tierra a partir de un modelo de elipsoide. IMA J. Aplicación. Matemáticas. 2015, 80, 963–980. Disponible en línea : https://www.researchgate.net/publication/274162167_Low-frequency_onsite_identification_of_a_highly_conductive_body_ enterrado_en_la_tierra_desde_un_modelo_elipsoide (consultado el 15 de febrero de 2024). [Referencia cruzada]
17. Ming, Y.; Niu, X.; Xie, X.; Han, Z.; Li, Q.; Sun, S. Aplicación a la exploración gravitacional en la detección de fallas activas urbanas. En La novena Conferencia Internacional sobre Geofísica Ambiental y de Ingeniería 2021, Actas de la serie de conferencias del IOP: Ciencias de la Tierra y el Medio Ambiente, Changchun, China, 11 a 14 de octubre de 2020; Publicación IOP: Bristol, Reino Unido, 2021; Volumen 660, pág. 012057. Disponible en línea: https://www.researchgate.net/publication/349532168_Application_of_gravity_exploration_in_urban_active_ detección_falla (consultado el 21 de febrero de 2024).
18. Sjoberg, LE; Bagherbagdi, M. Inversión e integración de la gravedad. En Teoría y Aplicaciones en Geodesia y Geofísica; Springer: Basilea, Suiza, 2017; pag. 251. Disponible en línea : https://www.geokniga.org/bookfiles/geokniga-gravity-inversion-and-integration.pdf (consultado el 21 de febrero de 2024).
19. Scheinert, M.; Ferranccioli, F.; Schwabe, J.; Studinger, M.; Damasco, D.; Jokat, W.; Aleshkova, N.; Jordán, T.; Leitchenkov, G.; Blankenship, DD; et al. Nueva cuadrícula de gnomalías gravitacionales antárticas para estudios geodésicos y geofísicos mejorados en la Antártida. Geofis. Res. Letón. 2016, 43, 600–610. Disponible en línea: https://www.researchgate.net/publication/289983658_New_Cuadrícula_de_anomalía_de_gravedad_antártica_para_estudios_geodésicos_y_geofísicos_mejorados_en_la_Antártida (consultado el 21 de febrero de 2024). [Referencia cruzada] [PubMed]
20. Ismail, N.; Yanís, M.; Abdullah, F.; Irfansyam, A.; Atmojo, BSW Mapeo de estructuras antiguas enterradas utilizando el método de gravedad: un estudio de caso de Cot Sidi Abdullah North Aceh. J. Física. 2018, 1120, 012035. Disponible en línea: https://iopscience.iop.org/article/10.1088/1742-6596/1120/1/012035/pdf (consultado el 21 de febrero de 2024). [Referencia cruzada]
21. Manoussakis, G. La fuerza de gravedad generada por un elipsoide de revolución de nivel no giratorio con baja excentricidad como una serie de armónicos esféricos. Matemáticas 2023, 11, 1974. Disponible en línea: https://www.mdpi.com/journal/mathematics/special_ problemas/Aplicaciones_Ecuaciones_Diferenciales_Partiales_Física_Matemática (consultado el 21 de febrero de 2024). [Referencia cruzada]

Machine Translated by Google

22. Heiskanen, W.; Moritz, H. Geodesia física. WH Freeman Co. 1967, 82–112. Disponible en línea: https://archive.org/details/ heiskanen_morits_1967_physical_geodesy/page/n89/mode/2up?view=teatro (consultado el 3 de marzo de 2024).
23. Bagherbandi, M.; Jouybari, A.; Niflorousan, F.; Ägren, J. Desviación del efecto vertical en la georreferenciación directa en sistemas de mapeo aéreo móvil : un estudio de caso en Suecia. Fotograma. Rec. 2022, 37, 285–305. Disponible en línea: https://www.diva-portal.org/ smash/get/diva2:1684428/FULLTEXT01.pdf (consultado el 3 de marzo de 2024). [Referencia cruzada]
24. Barzaghi, R.; Carrión, D.; Pepe, M.; Prezioso, G. Cálculo de la deflexión de la vertical para mejorar los reconocimientos aéreos: una comparación entre las estimaciones de EGM2008 e ITALGEO05. Sensores 2016, 16, 1168. Disponible en línea: https://www.ncbi.nlm.nih. gov/pmc/articles/PMC5017334/pdf/sensors-16-01168.pdf (consultado el 3 de marzo de 2024). [Referencia cruzada] [PubMed] 25. Šprlák, M.; Novak, P. Fórmulas integrales para calcular un tensor gravitacional de tercer orden a partir de la densidad de masa volumétrica, el potencial gravitacional perturbador, la anomalía de la gravedad y la perturbación de la gravedad. J. Geod. 2015, 89, 141-157. Disponible en línea: https://www.researchgate.net/publication/271211475_Integral_formulas_for_computing_a_thirdorder_gravitational_tensor_from_volumetric_mass_density_disturbing_gravitational_potential_gravity_anomaly_and_gravity_disturbance (consultado el 5 de marzo de 2024). [Referencia cruzada]
26. Gorshkov, GP Superficies isogravitacionales. Geofis. Abstr. 1934, 82. Disponible en línea: https://books.google.gr/books?id=1RxUCPOWOtEC&pg=PA1738&lpg=PA1738&dq=isogravitational+surfaces+of+the+Earth&source=bl&ots=9VnOhvB6Cs&sig=ACfU3U2hlSJxufGjWTktf6yfI13mwKuylA&hl=el&sa=X&ved=2ahUKEwjx8pD85tqEAxVh9LsIHbuUA_E4ChDoAXoECAY-QAw#v=onepage&q=isogravitational%20surfaces%20of%20the%20Earth&f=false (consultado el 5 de marzo de 2024). [Referencia cruzada]
27. Koop, R. Modelado del campo de gravedad global mediante gradiometría de gravedad satelital; Nederlandse Commissie voor Geodesie: Delft, Países Bajos, 1993; Volumen 38, págs. 32–183. Disponible en línea : https://ncgeo.nl/index.php/en/publicatiesgb/publications-on-geodesy/item/2544-pog-38-r-koop-global-gravity-modelling-using-satellite-gravity-gradiometry (consultado el 5 de marzo de 2024).
28. Giblin, P. Geometría diferencial de curvas y superficies; Departamento de Ciencias Matemáticas, Universidad de Liverpool: Liverpool, Reino Unido, 1999; págs. 32–39. Disponible en línea: https://www.liverpool.ac.uk/~pjgiblin/papers/diff-geom-book.pdf (consultado el 5 de marzo de 2024).

Descargo de responsabilidad/Nota del editor: Las declaraciones, opiniones y datos contenidos en todas las publicaciones son únicamente de los autores y contribuyentes individuales y no de MDPI ni de los editores. MDPI y/o los editores renuncian a toda responsabilidad por cualquier daño a personas o propiedad que resulte de cualquier idea, método, instrucción o producto mencionado en el contenido.

Machine Translated by Google

AppliedMath

MDPI

Artículo

Problema electromagnético 3D-Kepler de tres cuerpos: existencia de Soluciones periódicas

Vasil Georgiev Angelov

Departamento de Matemáticas e Informática, Universidad de Minería y Geología, “St. I. Rilski”, 1700 Sofía, Bulgaria; angelov@mgu.bg

Resumen: El objetivo principal del presente artículo es demostrar la existencia de soluciones periódicas del problema de los tres cuerpos en la formulación 3D de Kepler. Hemos resuelto el mismo problema en el caso de que las tres partículas se consideren en un sistema inercial externo. Comenzamos con las ecuaciones de movimiento de tres cuerpos, que son un subconjunto de las ecuaciones de movimiento (que hemos obtenido anteriormente) para cualquier número de cuerpos. En el espacio de Minkowski existen 12 ecuaciones de movimiento. Está demostrado que tres de ellas son consecuencia de las otras nueve, por lo que su número pasa a ser nueve, tanto como lo son las trayectorias desconocidas. La formulación de Kepler supone que una partícula (el núcleo) está situada en el origen de coordenadas. El movimiento de las otras dos partículas se describe mediante un sistema neutro con respecto a velocidades desconocidas. Los retrasos dependientes del estado surgen como consecuencia de la velocidad finita de la luz en el vacío. Obtenemos las ecuaciones de movimiento en coordenadas esféricas y las dividimos en dos grupos. En el primer grupo, todos los argumentos de las funciones desconocidas son retrasos. Tomamos sus soluciones como funciones iniciales. Luego, las ecuaciones de movimiento de las dos partículas restantes deben resolverse a la derecha del punto inicial. Para probar la existencia-unicidad de una solución periódica, elegimos un espacio que consta de funciones periódicas infinitamente suaves que satisfacen algunas condiciones suplementarias. Luego utilizamos un operador adecuado que actúa sobre estos espacios y cuyos puntos fijos son soluciones periódicas. Aplicamos el teorema del punto fijo para los operadores que actúan sobre los espacios de funciones periódicas. De esta manera mostramos la estabilidad del átomo de He en el marco de la electrodinámica clásica. En un artículo anterior nuestro, demostramos la existencia de funciones de espín para el movimiento plano. Por tanto, confirmamos la hipótesis de Bohr y S

Palabras clave: electrodinámica clásica; problema de los tres cuerpos; soluciones periódicas; método de punto fijo; sistema neutral

MSC: 78A35; 34K40; 47H10

check for updates

Cita: Angelov, VG Electromagnético 3D-Kepler de tres cuerpos
Problema: existencia de periódicos
Soluciones. Matemáticas aplicadas
2024, 4, 612–640. https://doi.org/10.3390/
appledmath4020034

Editor académico: Claude Chaudet

Recibido: 25 de marzo de 2024
Revisado: 19 de abril de 2024
Aceptado: 30 de abril de 2024
Publicado: 16 de mayo de 2024

Copyright: © 2024 por el autor.
Licenciatario MDPI, Basilea, Suiza.
Este artículo es un artículo de acceso abierto.
distribuido bajo los términos y
condiciones de los Creative Commons
Licencia de atribución (CC BY)
(https:// creativecommons.org/licenses/by/
4.0/).

1. Introducción

El objetivo principal del presente artículo es investigar el problema de los tres cuerpos de la electrodinámica clásica en la formulación 3D de Kepler. En [1], formulamos el mismo problema de tres cuerpos externamente para el sistema inercial de partículas, y en [2], demostramos la existencia-unicidad de una solución periódica. El problema de los tres cuerpos es un caso particular de la formulación general del problema de los N cuerpos (cf. [3]), donde se ha demostrado la existencia-unicidad de trayectorias de escape de N partículas cargadas. Aquí sumamos términos de radiación para cada una de las tres partículas. Usamos la forma relativista invariante del término de radiación de Dirac introducido en [4]. Partimos del sistema de [1] y pasando a las coordenadas esféricas formulamos el problema de Kepler 3D para tres partículas cargadas. Los resultados obtenidos nos permiten ampliar las investigaciones cualitativas sobre el átomo de He y átomos similares a He, como Li+ y otros. El método para resolver ecuaciones de movimiento se basa en la elección de espacios adecuados de funciones que consisten en una periódica infinitamente suave que satisface algunas condiciones adicionales e introduce operadores que actúan sobre estos espacios.

Sus puntos fijos son soluciones periódicas de las ecuaciones de movimiento. no es superfluo

Machine Translated by Google

Cabe señalar que, utilizando la formulación 3D de Kepler, colocamos los núcleos en el origen y buscamos trayectorias de dos partículas en movimiento que giran alrededor de los núcleos. Desde un punto de vista matemático es una operación formal, pero desde un punto de vista físico esto significa que pasamos del caso relativista al cuasi clásico, porque el problema de Kepler es en realidad un problema clásico (Newton). . Esto también lo confirma el tipo de término de radiación, que resulta estar en la forma de Lorentz.

Este artículo consta de siete secciones, cinco apéndices y conclusiones. La sección 1 es la introducción. La sección 2 contiene la formulación tridimensional de Kepler para las ecuaciones de movimiento en coordenadas cartesianas. Esto lleva a la separación del sistema en dos grupos. Una de ellas consta de tres ecuaciones que contienen funciones desconocidas en el intervalo inicial. Las llamamos ecuaciones iniciales. Seguimos el enfoque de RD Driver (cf. [5]) para describir la interacción retardada entre partículas, que se basa en la idea de W. Pauli [6] y JL Synge [7]. Se obtienen los retrasos dependientes del estado. De hecho, respecto a las velocidades desconocidas, el sistema obtenido es de tipo neutro, y los retrasos dependen de las trayectorias desconocidas. En la Sección 3, derivamos las ecuaciones básicas de movimiento en coordenadas esféricas. Las ecuaciones obtenidas son nuevamente de tipo neutral, con retrasos que dependen de las trayectorias desconocidas. En la Sección 4, presentamos los espacios funcionales que consisten en funciones periódicas infinitamente diferenciables que satisfacen algunas condiciones adicionales. También se prueban algunas proposiciones auxiliares. En la Sección 5, presentamos operadores cuyos puntos fijos son soluciones periódicas de las ecuaciones de movimiento y demostramos el Lema principal. La sección 6 contiene el resultado principal, que garantiza la existencia y unicidad de la solución periódica de las ecuaciones de movimiento para el problema de los tres cuerpos en la forma 3D de Kepler. La demostración se basa en un teorema del punto fijo [8]. Se da un ejemplo numérico. En la sección 7 se analizan las hipótesis de Bohr y Sommerfeld.

Mencionamos algunos resultados en los que se aplican diferentes enfoques [9-18].

Recordamos algunas denotaciones y resultados de [1,2] referentes al problema de los tres cuerpos en coordenadas cartesianas. Consideramos el sistema de 12 ecuaciones de movimiento usando términos de radiación para el sistema de tres cuerpos [1]:

$$\begin{aligned} m1\,\frac{d\lambda\, r^{(1)}}{ds1} &= \frac{e1}{2c}\left(F\lambda^{(12)}(1) + F\lambda^{(13)}(1)(1)_{rad}(1) + F\lambda\, rl\, l\, rl\, l\right), \\ m2\,\frac{d\lambda\, r^{(2)}}{ds2} &= \frac{e2}{2c}\left(F\lambda^{(21)}(2)(23)(2)(2)_{rad}(2) + F\lambda\, rl\, l\, rl\, l\, rl\, l\right), \\ m3\,\frac{d\lambda\, r^{(3)}}{ds3} &= \frac{e3}{2c}\left(F\lambda^{(31)}(3)(32)(3)(3)_{rad}(3) + F\lambda\, rl\, l\, rl\, l\, rl\, l\right), \end{aligned} \tag{1}$$

donde r = 1, 2, 3, 4; c es la velocidad de la luz en el vacío; mk son las masas; y ek (k = 1, 2, 3) son las cargas de las partículas. Recuerde que hay una sumatoria al repetir l en la mano derecha lados del sistema anterior. Los elementos de los tensores electromagnéticos F. $F_{rl}^{(kn)} = \frac{\partial \Phi_l^{(n)}}{\partial x_r^{(k)}} - \frac{\partial \Phi_r^{(n)}}{(k)\partial xl}$ (n) puede calcularse mediante los potenciales retardados de Lienard-Wiechert, $\Phi^{(n)} = \frac{e_n}{\text{en } \lambda \text{ mientras } \lambda(n), \xi(kn)4,}$ Los términos de radiación se pueden calcular como la mitad de la diferencia entre los potenciales retardado y avanzado, de acuerdo con el supuesto de Dirac:

$$F^{(k)rad}_{\text{Minnesota}} = \frac{1}{2}\left(\frac{\partial A^{(k)ret}_{\text{norte}}}{\partial x^{(k)ret}_{\text{metro}}} - \frac{\partial A^{(n)ret}_{\text{metro}}}{\partial x^{(k)ret}_{\text{norte}}} - \frac{\partial A^{(k)adv}_{\text{norte}}}{\partial x^{(k)adv}_{\text{metro}}} - \frac{\partial A^{(n)adv}_{\text{metro}}}{\partial x^{(k)adv}_{\text{norte}}}\right),$$

donde $A^{(k)ret}_{\text{norte}} = -\frac{ek\lambda}{\lambda(k)ret, \xi(k)ret4, (k)}$ $A^{(k)adverbio}_{\text{norte}} = -\frac{ek\lambda}{\lambda(k)adv, \xi(k)adv}$

Aquí, x1 (k) (t), x (k) (t), x (t) = tic 2 3 4 4 ., (k = 1, 2, 3) son coordenadas espacio-temporales de las partículas cargadas. El producto escalar en el espacio de Minkowski es $a, b4 = anbn = \sum_{norte=1}^{4} anbn,$

Machine Translated by Google

mientras que en el subespacio tridimensional, $a, b = a_\alpha b_\alpha = \sum_{\alpha=1}^{3} a_\alpha b_\alpha$. Los elementos de la adecuada los tiempos son dsk (k = 1, 2, 3) y los vectores unitarios tangentes a las líneas mundiales son

$$\lambda^{(k)} \quad \lambda^{\lambda^{1}, 2, 3, 4} = \lambda^{\lambda} \quad {}^{(k)}_{(k)} \quad {}^{(k)} = \frac{tu^{(k)}_{1}(t)}{1\,\Delta k,} \quad \frac{tu^{(k)}_{2}(t)}{2\,\Delta k,} \quad \frac{u^{(k)}_{i}(t)}{\Delta k,} \quad \vec{tú}^{(k)} = tu^{1} \quad {}^{(k)}_{(t),} \quad {}^{(k)}_{2} \quad {}^{(k)}_{(t),} \quad tu^{(t),}_{3}$$

$$\Delta k = c \sqrt{2 - \vec{tú}^{(k)}, \vec{tú}^{(k)}}.$$

Hemos demostrado que tres ecuaciones de (1) son consecuencia del resto [1].

Las ecuaciones de movimiento pasan a ser nueve, al igual que las funciones desconocidas (k) (t), u2 (k) (t), u3 (k) (t) , (k = 1, u1 2, 3).

Los vectores isotrópicos ξ (km) (k = 1, 2, 3 ; m = k) se obtienen de la siguiente manera: fija cualquier evento en la línea mundial de la k-ésima partícula y dibuja el cono de luz hacia el pasado. Este cono cruza la línea mundial de la m-ésima partícula en cualquier otro evento (pasado). Entonces,

$$= \xi^{(km)}_{1}, \xi^{(km)}_{2}, \xi^{(km)}_{3}, \xi^{(km)}_{4} = \vec{\xi}^{(km)}, \xi^{(km)}_{4} =$$

$$= x1^{(k)}_{(t)} - x^{(m)}_{2}(t - \tau km(t)), x^{(k)}_{2}(t) - x^{(m)}_{2}(t - \tau km(t)), x^{(k)}_{3}(t) - x^{(m)}_{3}(t - \tau km(t)), ic\tau km(t).$$

ξ (km) son cuatro vectores isotrópicos ξ (km) , las $_4 = 0$, lo que implica que el retrasado funciones ξ (km) τkm(t) deben satisfacer las siguientes ecuaciones funcionales:

$$\tau km(t) = \frac{1}{do} \sqrt{\vec{\xi}^{(kilómetros)}, \vec{\xi}^{(kilómetros)}} \equiv \frac{1}{do} \sqrt{\sum_{a=1}^{3} {}_{cuyo}{}^{(k)}_{(m)} \alpha^{(t)} - x^{\alpha}(t - \tau km(t))^{2}}.$$

Las ecuaciones anteriores son seis porque (km) = (12), (13), (23), (21), (31), (32)

y $\vec{tú}^{(m)} \equiv tu^{(m)}(t - \tau km(t))$, para m = k, $\Delta km = c\sqrt{2 - \vec{tú}^{(m)}_{(t - \tau km),} \vec{tú}^{(m)}_{(t - \tau km),}}$

$$\lambda^{(metro)} = \lambda^{1} \quad {}^{(m)}_{(m)}, {}^{4}_{(m)} \quad {}^{(m)}_{\lambda} \lambda^{\lambda} \lambda^{2, 3,} = \frac{tu^{(m)}_{1}(t - \tau km)\,\Delta km}{}, \frac{tu^{(m)}_{2}(t - \tau km)\,\Delta km}{}, \frac{tu^{(m)}_{3}(t - \tau km)\,\Delta km}{}, \frac{ic}{\Delta km} = \frac{\vec{tú}^{(m)}(t - \tau km)\,\Delta km}{}, \frac{ic}{\Delta km}.$$

2. Separación de las ecuaciones de movimiento y derivación de las ecuaciones iniciales

El supuesto (C): $\vec{tú}^{(k)} = \vec{tú}^{(k)}_{(t),} \vec{tú}^{(k)}(t) \le c < \bar{c}$ (k = 1, 2, 3; c = constante)

implica un $\sqrt{2 - \vec{tú}^{(k)}, \vec{tú}^{(k)}} \ge c\sqrt{2 - \bar{c}^{2}} > 0$. Esta condición nos permite obtener de (1) un sistema c que contiene tres grupos de ecuaciones (cf. [1]):

$$\begin{aligned}
\dot{tu}^{(1)}_{1} &= \frac{2 - c^{2}\, tu^{(1)}_{1}}{\cdot} {}_{cuyo}{}^{(12)}_{1}{}^{(13)}_{1}{}^{(1)rad} + G^{1} - \frac{tu^{(1)}_{1} tu^{(1)}_{2}}{2c} {}_{cuyo}{}^{(12)}_{2}{}^{(13)}_{2}{}^{(1)rad} + G - \frac{tu^{(1)}_{1} tu^{(1)}_{3}}{2c} {}_{cuyo}{}^{(12)}_{3}{}^{(13)}_{3}{}^{(1)rad} + G^{3} \equiv U1_{1} \\
\dot{tu}^{(1)}_{2} &= -\frac{tu^{(1)}_{1} tu^{(1)}_{2}}{do^{2}} {}_{cuyo}{}^{(12)}_{1}{}^{(13)}_{1}{}^{(1)rad} + G^{1} \quad \frac{c^{2} - (u^{(1)})^{2}}{2} {}_{cuyo}{}^{(12)}_{2}{}^{(13)}_{2}{}^{(1)rad} + G^{2}{}_{2} - \frac{tu^{(1)}_{2} tu^{(1)}_{3}}{do^{2}} {}_{cuyo}{}^{(12)}_{3}{}^{(13)}_{3}{}^{(1)rad} + G^{3} \equiv U1_{2} \\
\dot{tu}^{(1)}_{3} &= -\frac{tu^{(1)}_{1} tu^{(1)}_{3}}{2c} {}_{cuyo}{}^{(12)}_{1}{}^{(13)}_{1}{}^{(1)rad} + G^{1} - \frac{tu^{(1)}_{2} tu^{(1)}_{3}}{2c} {}_{cuyo}{}^{(12)}_{2}{}^{(13)}_{2}{}^{(1)rad} + G \quad \frac{c^{2} - (u^{(1)})^{2}}{3\,2c} {}_{cuyo}{}^{(12)}_{3}{}^{(13)}_{3}{}^{(1)rad} + G^{3} \equiv U1_{3},
\end{aligned} \tag{2}$$

$$\begin{aligned}
\dot{tu}^{(2)}_{1} &= \frac{2 - c^{2}\, tu^{(2)}_{1}}{2c} {}_{cuyo}{}^{(21)}_{1}{}^{(23)}_{1}{}^{(2)rad} + G^{1} - \frac{tu^{(2)}_{1} tu^{(2)}_{2}}{2c} {}_{cuyo}{}^{(21)}_{2}{}^{(23)}_{2}{}^{(2)rad} + G - \frac{tu^{(2)}_{1} tu^{(2)}_{3}}{2c} {}_{cuyo}{}^{(21)}_{3}{}^{(23)}_{3}{}^{(2)rad} + G^{3} \equiv U2_{1} \\
\dot{tu}^{(2)}_{2} &= -\frac{tu^{(2)}_{1} tu^{(2)}_{2}}{2c} {}_{cuyo}{}^{(21)}_{1}{}^{(23)}_{1}{}^{(2)rad} + G^{1} \quad \frac{(2)\, c^{2} - (u^{(2)})^{2}}{2\,2c} {}_{cuyo}{}^{(21)}_{2}{}^{(23)}_{2}{}^{(2)rad} + G - \frac{tu^{(2)}_{2} tu^{(2)}_{3}}{2c} {}_{cuyo}{}^{(21)}_{3}{}^{(23)}_{3}{}^{(2)rad} + G^{3} \equiv U2_{2} \\
\dot{tu}^{(2)}_{3} &= -\frac{tu^{(2)}_{1} tu^{(2)}_{3}}{2c} {}_{cuyo}{}^{(21)}_{1}{}^{(23)}_{1} + G^{1}{}^{(2)rad} - \frac{tu^{(2)}_{2} tu^{(2)}_{3}}{2c} {}_{cuyo}{}^{(21)}_{2}{}^{(23)}_{2}{}^{(2)rad} + G \quad \frac{c^{2} - (u^{(2)})^{2}}{3\,2c} {}_{cuyo}{}^{(21)}_{3}{}^{(23)}_{3}{}^{(2)rad} + G^{3} \equiv U2_{3},
\end{aligned} \tag{3}$$

Machine Translated by Google

$$\dot{tu}_1^{(3)} = \frac{2\text{-}c\ \dot{tu}_1^{(3)}{}^2}{do^2}\ {}_{GRAD}(31)\ (32)\ (3)rad + G\ 1 - \frac{tu_1^{(3)}\ tu_2^{(3)}}{do^2}\ {}_{GRAD}(31)\ (32)\ (3)rad + G\ 2\ {}^{(3)rad\ 2} - \frac{tu_1^{(3)}\ tu_3^{(3)}}{do^2}\ {}_{GRAD}(31)\ (32)\ (3)rad + G \equiv U3_1$$

$$\dot{tu}_2^{(3)} = -\frac{tu_1^{(3)}\ tu_2^{(3)}}{2c}\ {}_{GRAD}(31)\ (32)\ (3)rad + G\ 1 + \frac{(3)\ c\ 2-(u)^2}{2\ 2c}\ {}_{GRAD}(31)\ (32)\ (3)rad + G - \frac{tu_2^{(3)}\ tu_3^{(3)}}{2c}\ {}_{GRAD}(31)\ (32)\ (3)rad + G \equiv U3_2 \quad (4)$$

$$\dot{tu}_3^{(3)} = -\frac{tu_1^{(3)}\ tu_3^{(3)}}{2c}\ {}_{GRAD}(31)\ (32)\ (3)rad + G\ 1 - \frac{tu_2^{(3)}\ tu^{(3)}}{3\ 2c}\ {}_{GRAD}(31)\ (32)\ (3)rad + G + \frac{(3)\ c\ 2-(u)^2}{3\ 2c}\ {}_{GRAD}(31)\ (32)\ (3)rad + G \equiv U3_3,\ t \geq 0,$$

dónde

$$G_\alpha(12) = \frac{e1e2}{m1}\frac{\Delta 1}{3c}\ \frac{2\text{-}c\ \vec{tú}^{(1)},\vec{tú}^{(2)}\ \alpha^{(12)} - c\ 2r12 - \vec{\xi}^{(12)},\vec{tú}^{(1)}\ tu_\alpha^{(2)}}{c\ 2r12 - \vec{\xi}^{(12)},\vec{tú}^{(2)}} \frac{\Delta^4 12 + D12\ \Delta_{12}^2\ \vec{\xi}^{(12)},\dot{\vec{tú}}^{(2)} + \vec{\xi}^{(12)},\vec{tú}^{(2)}\ -c\ 2r12\ u^{(2)},\dot{tú}^{(2)}}{\Delta 2_{12}} -$$

$$- \frac{D12}{\vec{\xi}^{(12)},\vec{tú}^{(2)}\ -c\ 2r122} \frac{{}_{12}{}^{2}\ \dot{\alpha}^{(2)}(2)\ \Delta + uu\ \vec{tú}^{(2)},\dot{\vec{tú}}^{(2)}\ c\ 2r12 - \vec{\xi}^{(12)},\vec{tú}^{(1)} + \Delta 12^2\ \vec{tú}^{(1)},\dot{\vec{tú}}^{(2)} + \vec{tú}^{(1)},\vec{tú}^{(2)}\ 2\text{-}c\ \vec{tú}^{(2)},\dot{\vec{tú}}^{(2)}\ \xi_\alpha^{(12)}}{\Delta 2_{12}};$$

$${}_{GRAD}\alpha^{(13)} = \frac{e1e3}{m1}\frac{\Delta 1}{3c}\ \frac{2\text{-}c\ \vec{tú}^{(1)},\vec{tú}^{(3)}\ \alpha^{(13)} - c\ 2r13 - \vec{\xi}^{(13)},\vec{tú}^{(1)}\ tu_\alpha^{(3)}}{c\ 2r13 - \vec{\xi}^{(13)},\vec{tú}^{(3)}} \frac{\Delta^4 13 + D13\ \Delta_{13}^2\ \vec{\xi}^{(13)},\dot{\vec{tú}}^{(3)} + \xi^{(13)},\vec{tú}^{(3)}\ -c\ 2r13\ \vec{tú}^{(3)},\dot{\vec{tú}}^{(3)}}{\Delta 2_{13}} -$$

$$- \frac{D13}{\vec{\xi}^{(13)},\vec{tu}^{(3)}\ -c\ 2r132} \frac{2\ \dot{\alpha}^{(3)}(3)\ \Delta u + u\ 13\ \alpha\ \alpha\ \vec{tú}^{(3)},\dot{\vec{tú}}^{(3)}\ c2r13 - \vec{\xi}^{(13)},\vec{tú}^{(1)} + \Delta 13^2\ \vec{tú}^{(1)},\dot{\vec{tú}}^{(3)} + \vec{tú}^{(1)},\vec{tú}^{(3)}\ 2\text{-}c\ \vec{tú}^{(3)},\dot{\vec{tú}}^{(3)}\ \xi_\alpha^{(13)}}{\Delta 2_{13}};$$

$${}_{GRAD}\alpha^{(21)} = \frac{e2e1}{m2}\frac{\Delta 2}{3c}\ \frac{2\text{-}c\ \vec{tú}^{(2)},\vec{tú}^{(1)}\ \alpha^{(21)} - c\ 2r21 - \vec{\xi}^{(21)},\vec{tú}^{(2)}\ tú_\alpha^{(1)}}{c\ 2r21 - \vec{\xi}^{(21)},\vec{tú}^{(1)}} \frac{\Delta^4 21 + D21\ \Delta_{21}^2\ \vec{\xi}^{(21)},\dot{\vec{tú}}^{(1)} + \vec{\xi}^{(21)},\vec{tú}^{(1)}\ -c\ 2r21\ \vec{tú}^{(1)},\dot{\vec{tú}}^{(1)}}{\Delta 2_{21}} -$$

$$- \frac{D23}{\vec{\xi}^{(21)},\vec{tu}^{(1)}\ -c\ 2r212} \frac{2\ \dot{\alpha}^{(1)}(1)\ \Delta u + u\ 21\ \alpha\ \vec{tú}^{(1)},\dot{\vec{tú}}^{(1)}\ c\ 2r21 - \vec{\xi}^{(21)},\vec{tú}^{(1)} + \Delta 21^2\ \vec{tú}^{(2)},\dot{\vec{tú}}^{(1)} + \vec{tú}^{(2)},\vec{tú}^{(1)}\ 2\text{-}c\ \vec{tú}^{(1)},\dot{\vec{tú}}^{(1)}\ \xi_\alpha^{(21)}}{\Delta 2_{21}};$$

$${}_{GRAD}\alpha^{(23)} = \frac{e2e3}{m2}\frac{\Delta 2}{{}^{2}}\ \frac{2\text{-}c\ \vec{tú}^{(2)},\vec{tú}^{(3)}\ \alpha^{(23)} - c\ 2r23 - \vec{\xi}^{(23)},\vec{tú}^{(2)}\ tú_\alpha^{(3)}}{c\ 2r23 - \vec{\xi}^{(23)},\vec{tu}^{(3)}} \frac{\Delta^4 23 + D23\ \Delta_{23}^2\ \vec{\xi}^{(23)},\dot{\vec{tú}}^{(3)} + \vec{\xi}^{(23)},\vec{tú}^{(3)}\ -c\ 2r23\ \vec{tú}^{(3)},\dot{\vec{tú}}^{(3)}}{\Delta 2_{23}} -$$

$$- \frac{D23}{\vec{\xi}^{(23)},\vec{tu}^{(3)}\ ,-,c\ 2,r232} \frac{2\ \dot{\alpha}^{(3)}(3)\ \Delta u + u\ 23\ \alpha\ \vec{tú}^{(3)},\dot{\vec{tú}}^{(3)}\ c\ 2r23 - \vec{\xi}^{(23)},\vec{tú}^{(2)} + \Delta 23^2\ \vec{tú}^{(2)},\dot{\vec{tú}}^{(3)} + \vec{tú}^{(2)},\vec{tú}^{(3)}\ 2\text{-}c\ \vec{tú}^{(3)},\dot{\vec{tú}}^{(3)}\ \xi_\alpha^{(23)}}{\Delta 2_{23}};$$

$${}_{GRAD}\alpha^{(31)} = \frac{e3e1}{m3}\frac{\Delta 3}{{}^{2}}\ \frac{2\text{-}c\ \vec{tú}^{(3)},\vec{tu}^{(1)}\ \alpha^{(31)} - c\ 2r31 - \vec{\xi}^{(31)},\vec{tú}^{(3)}\ tu_\alpha^{(1)}}{c\ 2r31 - \vec{\xi}^{(31)},,,\vec{tú}^{(1)}} \frac{\Delta^4 31 + D31\ \Delta_{31}^2\ \vec{\xi}^{(31)},\dot{\vec{tú}}^{(1)} + \vec{\xi}^{(31)},\vec{tú}^{(1)}\ -c\ 2r31\ \vec{tú}^{(1)},\dot{\vec{tú}}^{(1)}}{\Delta 2_{31}} -$$

$$- \frac{D31}{\vec{\xi}^{(31)},\vec{tu}^{(1)}\ -c\ 2r312} \frac{\Delta 2_{31}\ d^{\dot{tu}}(1)\ (1) + u_\alpha\ \vec{tu}^{(1)},\dot{\vec{tú}}^{(1)}\ c\ 2r31 - \vec{\xi}^{(31)},\vec{tú}^{(3)} + \Delta_{31}^2\ \vec{tú}^{(3)},\dot{\vec{tú}}^{(1)} + \vec{tú}^{(3)},\vec{tu}^{(1)}\ 2\text{-}c\ \vec{tu}^{(1)},\dot{\vec{tú}}^{(1)}\ \xi_\alpha^{(31)}}{\Delta 2_{31}};$$

$${}_{GRAD}\alpha^{(32)} = \frac{e3e2}{m3}\frac{\Delta 3}{{}^{2}}\ \frac{2\text{-}c\ \vec{tu}^{(3)},\vec{tu}^{(2)}\ \alpha^{(32)} - c\ 2r32 - \vec{\xi}^{(32)},\vec{tu}^{(3)}\ tu_\alpha^{(2)}}{c\ 2r32 - \vec{\xi}^{(32)},\vec{tu}^{(2)}} \frac{\Delta^4 32 + D32\ \Delta_{32}^2\ \vec{\xi}^{(32)},\dot{\vec{tu}}^{(2)} + \vec{\xi}^{(32)},\vec{tu}^{(2)}\ -c\ 2r32\ \vec{tu}^{(2)},\dot{\vec{tu}}^{(2)}}{\Delta 2_{32}} -$$

$$- \frac{D32}{\vec{\xi}^{(32)},\vec{tu}^{(2)}\ -c\ 2r322} \frac{\Delta 2_{32}\ d^{\dot{tu}}(2)\ (2) + u_\alpha\ \vec{tu}^{(2)},\dot{\vec{tu}}^{(2)}\ c\ 2r32 - \vec{\xi}^{(32)},\vec{tu}^{(3)} + \Delta_{32}^2\ \vec{tu}^{(3)},\dot{\vec{tu}}^{(2)} + \vec{tu}^{(3)},\vec{tu}^{(2)}\ 2\text{-}c\ \vec{tu}^{(2)},\dot{\vec{tu}}^{(2)}\ (3\emptyset)\ \xi}{\Delta 2_{32}};$$

$${}_{GRAD}\alpha(k)rad = -\frac{2}{e\ k}\ \frac{tu_\alpha^{(k)}(t)}{2\ mk\ c}\ \frac{\vec{tú}^{(k)}(t),\ \ddot{\vec{tú}}^{(k)}(t) + \ddot{tu}_\alpha^{(k)}(t)\ \Delta k}{3\Delta k},(k = 1, 2, 3), \quad (5)$$

Machine Translated by Google

donde Dkm = $\frac{\sqrt{\xi^{(\text{kilómetros})}_{\text{do}} \cdot \xi^{(\text{kilómetros})}_{\text{do}} - \xi^{(\text{kilómetros})}_{\text{do}} \cdot tú^{(\text{metro})}_{\text{do}}}}{\sqrt{\xi^{(\text{kilómetros})}_{\text{do}} \cdot \xi^{(\text{kilómetros})}_{\text{do}} - \xi^{(\text{kilómetros})}_{\text{do}} \cdot tú^{(k)}_{\text{do}}}}$ para m = k.

De acuerdo con la formulación de Kepler, la primera partícula P1 está en el origen O(0,0,0), y por tanto,

$$tu^{(1)}_{\alpha}(t) = 0, u^{(1)}_{\alpha}(t) = 0, x_{\alpha}(t) = 0 \ (\alpha = 1, 2, 3). \quad (6)$$

Entonces obtenemos $0 = G^{(12)}_{1} + G^{(13)}_{1}$, $0 = G^{(12)}_{2} + G^{(13)}_{3}$, $0 = G^{(12)}_{3} + G^{(13)}_{2}$.

Es fácil ver que la ecuación $0 = G^{(12)}_{1} + G^{(13)}_{1}$ se convierte en 0 = 0.

Observación 1. Notamos que en la ecuación $0 = G^{(12)}_{2} + G^{(13)}_{2}$, las funciones desconocidas y sus derivadas tienen argumentos retardados t − τ12, mientras que en la ecuación 0 = G + G 3 los argumentos son t − τ13. Esto implica que debemos considerar estas ecuaciones en el intervalo inicial [−T, 0]. Por eso llamamos ecuaciones iniciales al primer grupo de tres ecuaciones (2). Consideramos el segundo y tercer grupo de ecuaciones (3) y (4), respectivamente (seis en total), en el intervalo [0, ∞).

Para simplificar las ecuaciones de movimiento, observamos que para la primera órbita de Bohr, la constante de estructura fina de Sommerfeld es β = c/c ≈ 1/137 β ≠ 0. En consecuencia, $\Delta k = \sqrt{\ldots} = c\sqrt{1 - \ldots/c^2} \approx c$, $\Delta km = c\sqrt{\ldots} = c(t - \tau km)$, $tú^{(metro)}_{\text{do}}(t - \tau km) \approx c$, y

$$\text{do}^2 \sqrt{1 - \frac{tú^{(2)}_{\text{do}} \cdot tú^{(3)}_{\text{do}}}{\ldots^2}} \approx c.$$

Si ρkm es la distancia entre la k-ésima y la m-ésima partícula vía ξ (km) ≈ ρkm ≈ 10−11 m (cf. [14]), obtenemos

$$c^2\tau 23 - \xi^{(23)}_{\text{do}} \cdot tú^{(2)}_{\text{do}} \approx c^2\tau 23, D23 = c^2\tau 23 - \xi^{(23)}_{\text{do}} \cdot tú^{(3)}_{\text{do}} \approx 1 \text{ y}$$

$$H23 = \frac{\Delta^4_{23} + D23\,\Delta^2_{23}}{\Delta^2_{23}} \approx c^2 + \xi^{(23)}_{\text{do}} \cdot \ldots - \tau 23\, tu^{(3)}_{\text{do}} \cdot \ldots .$$

Podemos simplificar las ecuaciones anteriores de acuerdo con las últimas relaciones:

$$G^{(12)}_{\alpha} = \frac{e1\,e2}{m1}\frac{1}{6c}\frac{c^2 + \xi^{(12)}_{\alpha} \cdot u^{(2)}_{\alpha} - \tau 12\, u^{(2)}_{\alpha} - c^2\tau 12^2 \cdot tu^{(2)}_{\alpha}}{\tau 123};$$

$$G^{(13)}_{\alpha} = \frac{e1\,e3}{m1}\frac{1}{6c}\frac{c^2 + \xi^{(13)}_{\alpha} \cdot u^{(3)}_{\alpha} - \tau 13\, u^{(3)}_{\alpha} - c^2\tau 13^2 \cdot tu^{(3)}_{\alpha}}{\tau 133};$$

$$G^{(21)}_{\alpha} = \frac{e2e1}{c}\frac{\Delta 2}{3\,m2}\frac{\ldots^2 \xi^{(21)}_{\alpha}}{(c\,2\tau 21)}\frac{\Delta^4_{21}}{3\Delta 2_{21}} \approx \frac{\ldots}{4\tau 213\,m2}\frac{(21)}{\alpha}(23)\,\xi\, e2e1 = \frac{e2e3}{3m2c}\frac{(23)}{\alpha}\frac{\xi}{\xi\tau 233}\frac{\xi^{(23)}_{\ldots} \cdot u^{(3)}_{\ldots}}{\ldots^3} - \frac{\dot{tu}^{(3)}_{\alpha}}{c\tau 23};$$

$$G^{(31)}_{\alpha} = \frac{e3e1}{m3}\frac{\Delta 3}{\ldots^3}\frac{\ldots^2 \xi^{(31)}_{\alpha}}{3}\frac{\Delta^4_{31}}{(c\,2\tau 31)\,\Delta 2_{31}} \approx \frac{e3e1}{m3}\frac{\xi^{(31)}_{\ldots}}{c\,4\tau 313}; G^{(32)}_{\alpha} = \frac{e2e3}{m3c^3}\frac{(32)}{\alpha}\frac{\xi}{\xi\tau 323}\frac{\xi^{(32)}_{\ldots} \cdot u^{(3)}_{\ldots}}{\ldots^3} - \frac{\dot{tu}^{(2)}_{\alpha}}{c\tau 32};$$

$$G^{(2)rad}_{\alpha} = -\frac{2}{2\,mi^2c^2}\frac{tu^{(2)}_{\alpha} \cdot tú^{(2)}_{\ldots} \cdot \ddot{u}^{(2)}_{\ldots}}{\ldots^3} + \frac{\ddot{tu}^{(2)}_{\alpha}}{\text{do}} \qquad G^{(3)rad}_{\alpha} = -\frac{y^2_3}{m3c^2}\frac{tu^{(3)}_{\alpha} \cdot tú^{(3)}_{\ldots} \cdot \ddot{u}^{(3)}_{\ldots}}{\ldots^3} + \frac{\ddot{tu}^{(3)}_{\alpha}}{\text{do}}$$

; GRAMO

Machine Translated by Google

Para la partícula estacionaria, tenemos G = 0. Por lo tanto, la ecuación (vectorial) inicial (13) (1)rad
ciones, $G_1^{(12)} + G_1^{(13)} = 0$, $G_2^{(12)} + G_2^{(13)} = 0$, $G_3^{(12)} + G_3^{(13)} = 0$, convertido en:

$$\begin{aligned}
&\frac{e1}{e2\,m1}\,\ddot{\xi}_1^{(12)} + \frac{e3}{e1\,m1}\,\ddot{\xi}_1^{(13)} = -c^2\tau_{12}^2\,\ddot{u}_1^{(2)} - c^2\tau_{13}^2\,\ddot{u}_1^{(3)},\\
&\frac{e1}{e2m2}\,\ddot{\xi}_2^{(12)} + \frac{e3}{e1m2}\,\ddot{\xi}_2^{(13)} = -c^2\tau_{12}^2\,\ddot{u}_2^{(2)} - c^2\tau_{13}^2\,\ddot{u}_2^{(3)}, \qquad (7)\\
&\frac{e1}{e2m3}\,\ddot{\xi}_3^{(12)} + \frac{e3}{e1\,m3}\,\ddot{\xi}_3^{(13)} = -c^2\tau_{12}^2\,\ddot{u}_3^{(2)} - c^2\tau_{13}^2\,\ddot{u}_3^{(3)}.
\end{aligned}$$

Suponemos que

$$\begin{aligned}
&\ddot{\xi}_1^{(12)}(t) = -c^2\tau_{12}^2\,\ddot{u}_1^{(2)}(t-\tau_{12}),\\
&\ddot{\xi}_2^{(12)}(t) = -c^2\tau_{12}^2\,\ddot{u}_2^{(2)}(t-\tau_{12}), \qquad (8)\\
&\ddot{\xi}_3^{(12)}(t) = -c^2\tau_{12}^2\,\ddot{u}_3^{(2)}(t-\tau_{12}).
\end{aligned}$$

y

$$\begin{aligned}
&\ddot{\xi}_1^{(13)}(t) = -c^2\tau_{13}^2\,\ddot{u}_1^{(3)}(t-\tau_{13}) = 0\\
&\ddot{\xi}_2^{(13)}(t) = -c^2\tau_{13}^2\,\ddot{u}_2^{(3)}(t-\tau_{13}) = 0, \qquad (9)\\
&\ddot{\xi}_3^{(13)}(t) = -c^2\tau_{13}^2\,\ddot{u}_3^{(3)}(t-\tau_{13}) = 0.
\end{aligned}$$

El sistema (8) contiene funciones con el argumento $t - \tau_{12}(t) \in [-T, 0]$ para $t \in [0, T]$.
El sistema (9) contiene funciones con el argumento $t - \tau_{13}(t) \in [-T, 0]$ para $t \in [0, T]$. Si (8) y (9) se cumplen, entonces (7) también se cumple.

Observación 2. La elección de $t - \tau_{km}(t) \in [-T, 0]$ se desprende del siguiente razonamiento: Sea T el período de la solución. Observamos que todos los argumentos de las funciones desconocidas son $t - \tau_{km}$, (m) $(t - \tau_{km})$. $0\,]$, $\vec{u}(t) = \vec{u}(t)$ Por tanto, debemos buscar una solución en el conjunto inicial, es decir, para $t - \tau_{km}(t) \in [\tau 0;$ donde $\tau 0 = \min\{t - \tau_{km}(t) : t \in [0, T]\}$. Dado que $1 - d\tau_{km}(t)/dt > 0$, entonces $t - \tau_{km}(t)$ es una función creciente. Hemos demostrado que si las trayectorias son T-periódicas, entonces $\tau_{km}(t)$ también es T-periódica.

Se deduce que $-T - \tau_{km}(-T) \le 0 - \tau_{km}(0)$ $-T - \tau_{km}(0) \le -\tau_{km}(0)$ $\tau_{km}(0) \le T - \tau_{km}(0)$ y luego $0-$ $\tau_{km}(0) \le t - \tau_{km}(t) \le T - \tau_{km}(T) = T - \tau_{km}(0) = 0$. En consecuencia, $-T \le t - \tau_{km}(t) \le 0$. Por lo tanto, las funciones $t - \tau_{km}(t) : [0, T] \to [-T, 0]$, lo que significa que $\tau 0 = -T$, y ponemos $t - \tau_{km}(t) = \theta \in [-T, 0]$. Los argumentos de las funciones desconocidas (2) (θ). Después de las transformaciones dadas en el Apéndice A, obtenemos de (8) (9) las ecuaciones iniciales : $\vec{u}^{(2)}(\theta) = \vec{u}(\theta)$ y $\vec{u}^{(3)}(\theta) = \vec{u}(\theta)$ (3) (θ) pertenecen a $[-T, 0]$.

$$\begin{aligned}
&\ddot{\rho}2\,(t-\tau_{23}) = \rho 2(t-\tau_{23})\cdot\dot{\lambda}2^2(t-\tau_{23})\cos 2\,\lambda 2(t-\tau_{23}) + \rho 2(t-\tau_{23})\,\dot{\lambda}2^2(t-\tau_{23}) - \frac{\mu}{\rho 2(t-\tau_{23})^2},\\
&\ddot{\lambda}2\,(t-\tau_{23}) = 2\,\dot{\lambda}2(t-\tau_{23})\,\lambda 2(t-\tau_{23})\,tg\lambda 2(t-\tau_{23}) - \frac{2\dot{\rho}2(t-\tau_{23})\cdot\dot{\lambda}2(t-\tau_{23})}{\rho n(t-\tau_{23})}, \qquad (10\text{-}1)\\
&\ddot{\lambda}2(t-\tau_{23}) = -\frac{2\dot{\rho}2(t-\tau_{23})\,\lambda 2(t-\tau_{23})+\rho 2(t-\tau_{23})\cdot\ddot{\lambda}2(t-\tau_{23})}{(t-\tau_{23})\,\mathrm{sen}\,\lambda 2(t-\tau_{23})\cos\lambda 2(t-\tau_{23})}.
\end{aligned}$$

Machine Translated by Google

para la segunda partícula y

$$\ddot{\rho}3\,(t-\tau 32)=\rho 3(t-\tau 32)\cdot\dot{\lambda}_3^{2}(t-\tau 32)\cos^2\lambda 3(t-\tau 32)+\rho 3(t-\tau 32)\,\dot{\lambda}_3^{2}(t-\tau 32)-\frac{\mu}{\rho 3(t-\tau 32)^2},$$

$$\ddot{\theta}_3\,(t-\tau 32)=2\,\dot{\rho}3\,(t-\tau 32)\,\dot{\lambda}3(t-\tau 32)\operatorname{tg}\lambda 3(t-\tau 32)-\frac{2\dot{\rho}3\,(t-\tau 32)\cdot\dot{\theta}_3\,(t-\tau 32)}{\rho 3(t-\tau 32)}, \tag{10-2}$$

$$\ddot{\lambda}3(t-\tau 32)=-\frac{2\dot{\rho}3\,(t-\tau 32)\,\dot{\lambda}3(t-\tau 32)+\rho 3(t-\tau 32)\cdot\dot{\theta}_3\,(t-\tau 32)^2\operatorname{sen}\lambda 3(t-\tau 32)\cos\lambda 3(t-\tau 32)}{\rho 3(t-\tau 32)^2}.$$

para la tercera partícula.

El razonamiento de la Observación 2 implica que buscamos una solución (10-1) y (10-2) en el intervalo [−T, 0].

El resultado de existencia-unicidad para (10 − n) (n = 1, 2) se puede obtener como en [19] en el espacio funcional $M_{r2}^{0}\times M_{\varphi 2}^{0}\times M_{\varphi 3}^{0}$ donde $\times M_{\eta 2}^{0}\times m_{r3}^{0}\times M_{\eta 3}^{0}$,

$$m_{r2}^{0}\ \mu(t-T_k)=r20(.)\quad CT\in[-T,0]\ (m)(t)\le r20(t)+R2e\ r20(0)=0\ ,$$

$$M_{\varphi 2}^{0}=\varphi 20(.)\quad CT\in[-T,0]:\varphi 20\ (m)(t)\le\omega_2 m\Phi_2^{\mu(t-T_k)},\ \varphi 20(-T)=\varphi 20(0)=0\ e\ ,$$

$$M_{\eta 2}^{0}=\eta 20(.)\quad CT\in[-T,0]:\eta 2\ (m)(t)\le\omega_2 m\Phi_2^{\mu(t-T_k)},\ \eta 20(-T)=\eta 20(0)=0\ e\ ,$$

$$m_{r3}^{0}\ \mu(t-T_k)=r30(.)\quad CT\in[-T,0]\ (m)(t)\le r30(t)+R3e\ r30(0)=0\ ,$$

$$M_{\varphi 3}^{0}=\varphi 30\quad CT\in[-T,0]:\varphi 30\ (m)(t)\le\omega 3\ m\Phi 3e^{\mu(t-T_k)},\ \varphi 30(-T)=\varphi 30(0)=0\ ,$$

$$M_{\eta 3}^{0}\ \mu(t-T_k)=\eta 30\quad CT\in[-T,0]\ (m)(t)\le\eta 30(t)+T3e\ \eta 30(0)=0\ .$$

como puntos fijos de los operadores

$$Br2\ (r2,\varphi 2,\eta 2,r3,\varphi 3,\eta 3)(t):=\int_{-T}^{t}Gr2\ (s)ds-\frac{t+t}{t}\int_{-T}^{0}Gr2\ (s)ds,\ t\quad[-T,0],$$

$$Br3\ (r2,\varphi 2,\eta 2,r3,\varphi 3,\eta 3)(t):=\int_{-T}^{t}Gr3\ (s)ds-\frac{t+t}{t}\int_{-T}^{0}Gr3\ (s)ds,\ t\quad[-T,0],$$

$$B\varphi 2\ (r2,\varphi 2,\eta 2,r3,\varphi 3,\eta 3)(t):=\int_{-T}^{t}G\varphi 2\ (s)ds-\frac{t+t}{t}\int_{-T}^{0}G\varphi 2\ (s)ds,\ t\quad[-T,0],$$

$$B\varphi 3\ (r2,\varphi 2,\eta 2,r3,\varphi 3,\eta 3)(t):=\int_{-T}^{t}Gr3\ (s)ds-\frac{t+t}{t}\int_{-T}^{0}Gr3\ (s)ds,\ t\quad[-T,0],$$

$$B\eta 2\ (r2,\varphi 2,\eta 2,r3,\varphi 3,\eta 3)(t):=\int_{-T}^{t}G\eta 2\ (s)ds-\frac{t+t}{t}\int_{-T}^{0}G\eta 2\ (s)ds,\ t\quad[-T,0],$$

$$B\eta 3\ (r2,\varphi 2,\eta 2,r3,\varphi 3,\eta 3)(t):=\int_{-T}^{t}G\eta 3\ (s)ds-\frac{t+t}{t}\int_{-T}^{0}G\eta 3\ (s)ds,\ t\quad[-T,0].$$

Es fácil comprobar que

$$Br2\ (.)(-T)=\int_{-T}^{-T}Gr2\ (s)ds-\frac{-T+T}{t}\int_{-T}^{0}Gr2\ (s)ds=0=\int_{-T}^{0}Gr2\ (s)ds-\frac{t}{t}\int_{-T}^{0}Gr2\ (s)ds=Br2\ (.)(0)$$

etcétera. Luego, tomamos las soluciones como las funciones iniciales r20(t), φ20(t), η20(t),r30(t), φ30(t), η30(t) para las ecuaciones básicas siguientes.

3. Ecuaciones Básicas en Coordenadas Esféricas

Observamos que $\lambda_{23}=\lambda_2(t)-\lambda_3(t-\tau 23)$, $\lambda_{32}=\lambda_3(t)-\lambda_2(t-\tau 32)$.

De (2) a (4), se obtiene, en el Apéndice B, la forma final de las ecuaciones de movimiento de la segunda y tercera partículas:

Machine Translated by Google

$$\ddot{\rho}2(t) = \frac{e2e1}{m2}\frac{}{1\ c\rho 2\ 2(t)} - \frac{}{e2e3m2}\frac{\rho 2(t)-\rho 3(t-\tau 23)(\cos\ \ 23\cos\lambda 2(t)\cos\lambda 3(t-\tau 23)+\sin\lambda 2(t)\sin\lambda 3(t-\tau 23))}{^{2}}\frac{\vec{\xi}^{(23)\ldots u}}{c\ 3\tau 23\ 3}^{(3)} -$$

$$- \frac{}{e2e3m2}\frac{\ddot{\rho}3\ (\cos\ \ 23\cos\lambda 2(t)\cos\lambda 3(t-\tau 23)+\sin\lambda 2(t)\sin\lambda 3(t-\tau 23))+\rho 3(t-\tau 23)\ldots\ 3\ (t-\tau 23)\sin\ \ 23\cos\lambda 2(\ t)\cos\lambda 3(t-\tau 23)}{c\ 4\tau 23} -$$

$$- \frac{}{}\frac{\ddot{\lambda}3(\operatorname{sen}\lambda 2(t)\cos\lambda 3(t-\tau 23)-\cos\ \ 23\ \operatorname{sen}\lambda 3(t-\tau 23)\cos\lambda 2(t))\ e2e3\ \rho 3\ c\ 4\tau 23\ m2}{} - \frac{y_2^2}{m2}\frac{\dddot{\rho}\ (t)}{_{do}\ 2\ 3},$$

$$\ddot{}\ 2\ (t) = \frac{}{e2e3m2}\frac{\rho 3(t-\tau 23)\ \operatorname{sen}\ \ 23\cos\lambda 3(t-\tau 23)}{\tau 23\ 3\rho 2(t)\cos\lambda 2(t)}\frac{\vec{\xi}^{(23)\ldots}}{_{tu}.}^{(3)}}{6c} +$$

$$\frac{\ddot{\rho}3\ (t-\tau 23)\cos\lambda 3(t-\tau 23)\ \operatorname{sen}\ \ 23+\rho 3(t-\tau 23)\ldots\ 3\ (t-\tau 23)\cos\lambda 3(t-\tau 23)\cos\ \ 23\ +\ c}{4\tau 23\rho 2(t)\cos\lambda 2(t)} -$$

$$- \frac{\ddot{\rho}3(t-\tau 23)\ \ddot{\lambda}3(t-\tau 23)\ \operatorname{sen}\lambda 3(t-\tau 23)\ \operatorname{sen}\ \ 23\ c}{4\tau 23\rho 2(t)\cos\lambda 2(t)} - \frac{\dddot{}\ 2\ (t)}{c^3}\ ;$$

$$\ddot{\lambda}2(t) = \frac{3\ m2}{}\frac{\rho 3(t-\tau 23)\cos\ \ 23\ \operatorname{sen}\lambda 2(t)\cos\lambda 3(t-\tau 23)\ e2e3\ c\ 3\tau 23}{}\frac{\vec{\xi}^{(23)\ldots}}{tu}^{(3)}}{c3\rho 2} +$$

$$+ \frac{e2e3}{m2}\frac{\ddot{\rho}3\ (\cos\ \ 23\cos\lambda 3(t-\tau 23)\ \operatorname{sen}\lambda 2(t)-\operatorname{sen}\lambda 3(t-\tau 23)\cos\lambda 2(t))+\rho 3\ \ldots\ \ 3\ \operatorname{sen}\ \ 23\ \operatorname{sen}\lambda 2(t)\cos\lambda 3(t-\tau 23)\ +\ c4\tau 23\rho 2}{}(t)$$

$$m2\frac{e2e3}{}\frac{-\rho 3\ +\ \ddot{\lambda}3(\cos\ \ 23\ \operatorname{sen}\lambda 2(t)\ \operatorname{sen}\lambda 3(t-\tau 23)+\cos\lambda 2(t)\cos\lambda 3(t-\tau 23)\ c}{4\tau 23\rho 2(t)} - \frac{y_2^2}{m2}\frac{\dddot{\lambda}2\ (t)}{_{do}\ 3},$$

dónde

$$\ddot{\rho}3\ (t) = \frac{e3e1}{m3}\frac{1}{c\rho 3\ 2(t)} - \frac{e2e3}{m3}\frac{\rho 3(t)-\rho 2(t-\tau 32)(\cos\ \ 32\cos\lambda 2(t-\tau 32)\cos\lambda 3(t)+\sin\lambda 2(t-\tau 32)\sin\lambda 3(t))}{^{2}}\frac{\vec{\xi}^{(32)\ldots u}}{c\ 3\tau 32\ 3}^{(2)}$$

$$- \frac{e3e2}{m3}\frac{\ddot{\rho}2\ (\cos\ \ 32\cos\lambda 2(t-\tau 32)\cos\lambda 3(t)+\sin\lambda 2(t-\tau 32)\sin\lambda 3(t))+\rho 2(t-\tau 32)\ldots\ 2\ (t-\tau 32)\sin\ \ 32\cos\lambda 2(\ t-\tau 32)\cos\lambda 3(t)\ c\ 4\tau 23}{}$$

$$- \frac{e3e2}{m3}\frac{\rho 2(t-\tau 32)\ \ddot{\lambda}2(t-\tau 32)(\sin\lambda 3(t)\cos\lambda 2(t-\tau 32)-\cos\ \ 32\ \sin\lambda 2(t-\tau 32)\cos\lambda 3(t))}{c4\tau 23} - \frac{y_3^2}{m3}\frac{\dddot{\rho}\ (t)}{_{do}\ 3\ 3},$$

$$\ddot{}\ 3 = \frac{e3e2}{m3}\frac{\rho 2(t-\tau 32)\ \operatorname{sen}\ \ 32\cos\lambda 2(t-\tau 32)}{\tau 32\ 3\rho 3(t)\cos\lambda 3(t)}\frac{\vec{\xi}^{(32)\ldots u}}{}^{(2)}}{6c} +$$

$$\frac{\ddot{\rho}2\ (t-\tau 32)\cos\lambda 2(t-\tau 32)\ \operatorname{sen}\ \ 32+\rho 2(t-\tau 32)\ldots\ 2\ (t-\tau 32)\cos\lambda 2(t-\tau 32)\cos\ \ 32\ +\ c}{4\tau 32\rho 3\cos\lambda 3(t)} -$$

$$- \frac{\rho 2(t-\tau 32)\ \ddot{\lambda}2(t-\tau 32)\ \operatorname{sen}\lambda 2(t-\tau 32)\ \operatorname{sen}\ \ 32\ c}{4\tau 32\rho 3\cos\lambda 3(t)} - \frac{\dddot{}\ 3\ (t)}{c^3}\ ;$$

$$\ddot{\lambda}3(t) = \frac{e3e2}{m3}\frac{\rho 2(t-\tau 32)\cos\ \ 32\ \operatorname{sen}\lambda 3(t)\cos\lambda 2(t-\tau 32)\ c}{3\tau 32\ 3}\frac{\vec{\xi}^{(32)\ldots u}}{}^{(2)}}{c3\rho 3\ (t)} +$$

$$+ \frac{e3e2}{3m3c}\frac{\ddot{\rho}3\ (\cos\ \ 32\cos\lambda 2(t-\tau 32)\sin\lambda 3(t)-\sin\lambda 2(t-\tau 32)\cos\lambda 3(t))+\rho 2(t-\tau 32)\ldots\ 2\ (t-\tau 32)\sin\ \ 32\ \sin\lambda 3(\ t)\cos\lambda 2(t-\tau 32)\ c\tau 32\rho 3(t)}{}$$

$$+ \frac{e3e2}{3m3c}\frac{-\rho 2(t-\tau 32)\ \ddot{\lambda}2(t-\tau 32)(\cos\ \ 23\ \operatorname{sen}\lambda 2(t-\tau 32)\ \operatorname{sen}\lambda 3(t)+\cos\lambda 2(t-\tau 32)\cos\lambda 3(t)}{c\tau 32\rho 3(t)} - \frac{y_3^2}{m3}\frac{\dddot{\lambda}3\ (t)}{3\ .\ do}$$

y

Machine Translated by Google

$$\vec{\xi}^{(32)}_{,\,(2)\,.\to tu} =$$

$$= [\rho 3(t)\cos(\quad 3(t) - \quad 2(t-\tau 32)\cos\lambda 3(t)\cos\lambda 2(t-\tau 32) + \rho 3(t)\,\mathrm{sen}\,\lambda 3(t)\,\mathrm{sen}\,\lambda 2(t-\tau 32) - \rho 2\,(t-\tau 32)] \ldots \rho 2\,(t-\tau 32) +$$

$$+ \sin(\quad 3(t) - \quad 2(t-\tau 32))\cos\lambda 3(t)\cos\lambda 2(t-\tau 32)\rho 2(t-\tau 32)\rho 3(\ t) \ldots \quad 2\,(t-\tau 32) +$$

$$+ [\sin\lambda 3(t)\cos\lambda 2(t-\tau 32) - \cos(\quad 3(t) - \quad 2(t-\tau 32))\sin\lambda 2(t-\tau 32)\cos\lambda 3(t)]\rho 2(t-\tau 32)\rho 3(\ t)\,\lambda 2(t-\tau 32).$$

Después de la sustitución,

$$r_n, \quad \varphi_n, \ \rho_n = \quad \lambda_n = \eta_n\ (n = 2, 3) \qquad \rho_n(t) = \rho_{n0} + \int_0^t r_n(s)ds, \quad \varphi_n = \varphi_{n0} + \int_0^t \varphi_n(s)ds, \ \lambda_n = \lambda_{n0} + \int_0^t \eta_n(s)ds$$

el sistema de ecuaciones de movimiento de la segunda y tercera partículas se puede reescribir como

$$\dot{r}_2 = \frac{e_2e_1}{m_2}\frac{}{1\ c\rho_2^2} - \frac{}{e_2e_3m_2}\frac{\rho_2-\rho_3(\cos\quad 23\cos\lambda_2\cos\lambda_3+\mathrm{sen}\,\lambda_2\,\mathrm{sen}\,\lambda_3)}{\ldots}\frac{\vec{\xi}^{(23)\,.\to u\,(3)}_{,}}{c\ 3\tau_{23}\ 3}$$

$$- \frac{}{e_2e_3m_2}\frac{.\ r_3(\cos\quad 23\cos\lambda_2\cos\lambda_3+\sin\lambda_2\sin\lambda_3)+\rho_3\,\varphi_3\sin\quad 23\cos\lambda_2\cos\lambda_3+\rho_3\,\eta_3\,(\sin\lambda_2\cos\lambda_3-\cos\quad 23\sin\lambda_3\cos\lambda_2)\ c\ 4\tau_{23}}{} - \frac{y_{r2}^2\ddot{r}_{2\,2\,3}}{m_2\,c} \equiv Gr_2\ ;$$

$$\dot{\varphi}_2 = \frac{}{e_2e_3m_2}\frac{\rho_3\,\mathrm{sen}\quad 23\cos\lambda_3}{\tau_{23}\ 3\rho_2\cos\lambda_2}\frac{\vec{\xi}^{(23)\,.\to\ (3)}_{,\,tu}}{6c} + \frac{\dot{r}_3\cos\lambda_3\,\mathrm{sen}\quad 23+\rho_3\,\varphi_3\cos 3\cos\quad 23-\rho_3\,\eta_3\,\mathrm{sen}\,\lambda_3\,\mathrm{sen}\quad 23}{c\ 4\tau_{23}\rho_2\cos\lambda_2} - \frac{\ddot{\varphi}_2}{\ldots} \equiv G\varphi_2\ ;$$

$$\dot{\eta}_2 = \frac{e_2e_3}{c\ 3\tau_{23}\ 3\ m_2}\frac{\rho_3\cos\quad 23\,\mathrm{sen}\,\lambda_2\cos\lambda_3}{c3\rho_2}\frac{\vec{\xi}^{(23)\,.\to u\,(3)}_{,}}{} +$$

$$+ \frac{e_2e_3}{m_2}\frac{.\ r_3(\cos\quad 23\cos\lambda_3\,\mathrm{sen}\,\lambda_2-\mathrm{sen}\,\lambda_3\cos\lambda_2)+\rho_3\,\varphi_3\,\mathrm{sen}\quad 23\,\mathrm{sen}\,\lambda_2\cos\lambda_3-\rho_3\,\eta_3\,(\cos\quad 23\,\mathrm{sen}\,\lambda_2\,\mathrm{sen}\,\lambda_3+\cos\lambda_2\cos\lambda_3\ c}{4\tau_{23}\rho_2} - \frac{y_2^2\ddot{\eta}_2}{m_2\ 3c} \equiv G\eta_2\ ;$$

$$\dot{r}_3 = \frac{e_3e_1}{m_3}\frac{1}{2c\rho_3} - \frac{e_2e_3}{m_3}\frac{\rho_3-\rho_2(\cos\quad 32\cos\lambda_2\cos\lambda_3+\mathrm{sen}\,\lambda_2\,\mathrm{sen}\,\lambda_3)}{\ldots}\frac{\vec{\xi}^{(32)\,.\to u\,(2)}_{,}}{c\ 3\tau_{32}\ 3}$$

$$- \frac{e_3e_2}{m_3}\frac{.\ r_2(\cos\quad 32\cos\lambda_2\cos\lambda_3+\mathrm{sen}\,\lambda_2\,\mathrm{sen}\,\lambda_3)+\rho_2\,\eta_2(\mathrm{sen}\,\lambda_3\cos\lambda_2-\cos\quad 32\,\mathrm{sen}\,\lambda_2\cos\lambda_3)\ \varphi_2\,\mathrm{sen}\quad 32\cos\lambda_2\cos\lambda_3+\rho_2\ c}{} - \frac{y_3^2\ddot{r}_3}{m_3\ 3c} \equiv Gr_3\ ;$$

$$\dot{\varphi}_3 = \frac{e_3e_2}{m_3}\frac{\rho_2\,\mathrm{sen}\quad 32\cos\lambda_2}{\tau_{32}\ 3\rho_3\cos\lambda_3}\frac{\vec{\xi}^{(32)\,.\to\ (2)}_{,}}{6c}\quad \frac{\dot{r}_2\cos\lambda_2\,\mathrm{sen}\quad 32+\rho_2\,\eta_2\,\mathrm{sen}\,\lambda_2\cos\lambda_3\,\mathrm{sen}\,\lambda_3\,\varphi_2\cos 2\cos\quad 32-\rho_2}{} - \frac{\ddot{\varphi}_3}{3c} \equiv G\varphi_3\ ;$$

$$\dot{\eta}_3 = \frac{e_3e_2}{c\ 3\tau_{32}\ 3\ m_3}\frac{\rho_2\cos\quad 32\,\mathrm{sen}\,\lambda_3\cos\lambda_2}{c3\rho_3}\frac{\vec{\xi}^{(32)\,.\to\ (2)}_{,\,tu}}{} +$$

$$+ \frac{e_3e_2}{3m_3c}\frac{.\ r_3(\cos\quad 32\cos\lambda_2\,\mathrm{sen}\,\lambda_3-\mathrm{sen}\,\lambda_2\cos\lambda_3)+\rho_2\,\eta_2\,(\cos\quad 23\,\mathrm{sen}\,\lambda_2\,\mathrm{sen}\,\lambda_3+\cos\lambda_2\cos\lambda_3\,\varphi_2\,\mathrm{sen}\quad 32\,\mathrm{sen}\,\lambda_3\cos\lambda_2-\rho_2\,c\tau_{32}\rho_3}{} - \frac{y_3^2\ddot{\eta}_3}{c\quad 3\ m_3} \equiv G\eta_3\ .$$

Más adelante nos referiremos al sistema anterior como (BS): sistema básico. Es un sistema neutral de primer orden con respecto a velocidades esféricas.

4. Introducción a los espacios funcionales y algunas proposiciones auxiliares

Se demuestra que para trayectorias periódicas T continuas de Lipschitz, cada una de las ecuaciones funcionales,

Machine Translated by Google

$$\tau 23(t) = \frac{1}{d_o}\sum_{a=1}^{3}\alpha(t-\tau 23(t))2,\ \tau 32(t) = \alpha(t) - x \quad \frac{1}{d_o}\sum_{a=1}^{3}\alpha(t) - x \quad \alpha(t-\tau 32(t))2 \quad (3)\ (2)$$

tiene una solución periódica T única.

Por CT ∞[0, ∞), denotamos el conjunto de todas las funciones infinitamente diferenciables, que condición r(−T) = r(0) = 0. Introducimos los siguientes conjuntos, donde 0 < δ < (m) son constantes satisfacen 2 , μ estrictamente positivas, r significa la m-ésima derivada de r y r Luego, introducimos lo siguiente > ω > 0 (0) (t) = r(t). espacios funcionales:

$$Srn = rn \quad CT\ \infty[0,\infty) : rn \quad (m)(t) \le \omega n\, mRne^{\mu(t-Tk)};\ rn \quad (0)(Tk) = 0;\ rn(s)ds = 0 \quad ,\ t \quad [Tk, Tk+1];$$

$$d(k,m)(rn,\overline{rn}) = \sup \quad \frac{|rn(m)(t)-\overline{rn}(m)(t)|}{\omega n}\ {}^{mi}-\mu(t-Tk);\ t \quad [Tk, Tk+1];$$

$$m\varphi\ norte = \varphi n \quad CT\ \infty[0,\infty) : \varphi n(m)(t) \le \omega n\, m\Phi ne^{\mu(t-Tk)} \quad ,\ Tk+1\ \varphi n(s)ds = T\varphi n0,\ t \quad [Tk, Tk+1];$$

$$d(k,m)(\varphi n, \overline{\varphi n}) = \sup \quad \frac{|\varphi n(m)(t)-\overline{\varphi n}(m)(t)|}{\omega n}\ {}^{mi}-\mu(t-Tk);\ t \quad [Tk, Tk+1] \quad ;$$

$$M\eta n = \eta n \quad CT\ \infty[0,\infty) : \eta n;\ \eta n \quad (m)(t) \le \omega mYne^{\mu(t-Tk)} \quad (0)(Tk) = 0;\ Tk+1\ \eta n(s)ds = 0;\ \int_{tk}^{t}\eta(s)ds \le \frac{\pi}{2} - \delta \quad ,\ t \quad [Tk, Tk+1];$$

$$d(k,m)(\eta n, \overline{\eta n}) = \sup \quad \frac{|\eta n(m)(t)-\overline{\eta n}(m)(t)|}{\omega n}\ {}^{mi}-\mu(t-Tk);\ t \quad [Tk, Tk+1]$$

$(k = 0, 1, 2, \ldots), (m = 0, 1, 2, \ldots), (n = 2, 3).$

Aquí, ωn , Rn , Φn, Yn (n = 2, 3) y μ son constantes estrictamente positivas, tales que μ > ωn.

Consideramos el producto de espacios uniformes, Mr2 × Mφ2 × Mη2 × Mr3 × Mφ3 × Mη3 , dotado de la familia contable de pseudométricas:

$$d(k,m)((r2,\varphi 2,\eta 2,r3,\varphi 3,\eta 3),(r2,\overline{\varphi 2},\overline{\eta 2},r3,\overline{\varphi 3},\overline{\eta 3})) = = d(k,m)(r2,r2)$$
$$+ d(k,m)(\varphi 2,\overline{\varphi 2}) + d(k,m)(\eta 2,\overline{\eta 2}) + d(k,m)(r3,r3) + d(k,m)(\varphi 3,\overline{\varphi 3}) + d(k,m)(\eta 3,\overline{\eta 3})$$

donde el conjunto de índices A consta de todos los pares ordenados (k, m).

Lema 1 ([20]). Si $rn(.) \quad CT\ \infty[0,\infty)$ y $\int_0^t rn(t)dt = 0$, entonces, $\rho n(t) = \rho n0 + \int_0^t rn(s)ds$ $(n = 2, 3)$ es una función T-periódica. Además,

$$2(t+T) = \quad 20 + \int_0^{t+t}\varphi 2(s)ds = \quad 20 + \int_0^t \sum_{k+1}^{T}\varphi 2(s)ds + \varphi 2(s)ds = \quad 2(t) + \varphi 20T = \quad 2(t) + 2\pi$$

$$\cos \quad 2(t+T) = \cos(\quad 2(t) + 2\pi) = \cos \quad 2(t)$$

$$T - \tau 23(t)) = \quad 30 + \varphi 3(s)ds = \quad 30 + \varphi 3(s)ds + \varphi 3(s)ds = \int_0^{t+T-\tau 23(t)} \int_0^{t-\tau 23(t)} \int_{t-\tau 23(t)}^{t-\tau 23(t)+T} 3(t+T-\tau 23(t+T)) = \quad 3(t+$$

$$= \quad 3(t-\tau 23) + \varphi 30T = \quad 3(t-\tau 23) + 2\pi \quad \cos$$

$$3(t+T-\tau 23(t+T)) = \cos(\quad 3(t-\tau 23) + 2\pi) = \cos \quad 3(t-\tau 23)$$

De ello se deduce que los lados derechos de las ecuaciones de (BS) son funciones T-periódicas.

Lema 2. Las distancias ρn(t), (n = 2, 3) son funciones acotadas.

Prueba. Para μ > 0 y μT = const. > 0, obtenemos

$$\rho n(t) \ge \rho n0 - \int_0^t |r\alpha(s)|ds \ge \rho n0 - Rn \int_0^t {}^{mi}\mu s\, ds = \rho n0 - Rn \quad \frac{{}^{mi}\mu t - 1}{\mu} \ge \rho n0 - \rho n0 - \frac{Rn(e\ \mu T - 1)}{\mu} \quad \frac{Rne\ \mu T}{\mu} \equiv {}_{\rho\ norte} > 0 \quad ,(n = 2, 3).$$

Machine Translated by Google

De manera similar, inferimos

$$\rho n(t) \le \rho n0 + \int_0^t |rn(s)|ds \le \rho n0 + Rn \int_0^t {}^{mi}\mu s\, ds \le \rho n0 + Rn \frac{{}^{mi}\mu T - 1}{\mu} \equiv \rho_{\quad} > 0 \ , (n = 2, 3).$$

Finalmente, $\rho_{\quad} < {}_{\quad}$ desde e $\mu T - 1 > 0$.

Así queda demostrado el lema 2. □

Lema 3. Si $\int_0^t \varphi n(t)dt = \varphi n0T$ $(\varphi n0 > 0)$,$(n = 2, 3)$, entonces $\lim_{t\to\infty} n(t) = \infty$.

Prueba. Pongamos $s + T = p$. Entonces, obtenemos $\int_0^t \varphi n(s)ds = \int_t^{2T} \varphi n(p)dp = \cdots = \int_{tk}^{T} {}_{k+1}\, \varphi n(p)dp.$

Por lo tanto, $\lim_{t\to\infty} n(t) = \lim_{t\to\infty} n0 + \int_0^{\infty} \varphi n(s)ds = n0 + \sum_{k=0}^{\infty} \int_{tk}^{Tk+1} \varphi n(s)ds = n0 + T\varphi n0 \sum_{k=1}^{\infty} 1 = \infty.$

Así queda demostrado el lema 3. □

Lema 4. Los conjuntos Mrn , $M\varphi n$, $M\eta n$ $(n = 2, 3)$ son cerrados.

Para la prueba, cf. [19].

Las siguientes afirmaciones dan una estimación cualitativa de las órbitas de los electrones.

Lema 5. La siguiente desigualdad es válida:

$A = |\cos \quad 2 \cos \quad 3 \cos \lambda 2 \cos \lambda 3 + \operatorname{sen} \quad 2 \operatorname{sen} \quad 3 \cos \lambda 2 \cos \lambda 3 + \operatorname{sen} \lambda 2 \operatorname{sen} \lambda 3| \le 1.$

Prueba.

$$A = |\cos \lambda 2 \cos \lambda 3 \cos(\quad 2 - \quad 3) + \operatorname{sen} \lambda 2 \operatorname{sen} \lambda 3| =$$
$$= \frac{\cos(\lambda 2 - \lambda 3) + \cos(\lambda 2 + \lambda 3)}{2} \cos(\lambda 2 - \lambda 3) - \cos(\lambda 2 + \lambda 3) \cos(\frac{\quad 2 - \quad 3)}{2} + =$$
$$= \frac{\text{porque}(\lambda 2 - \lambda 3)[1 + \cos(\quad 2 - \quad 3)] - \cos(\lambda 2 + \lambda 3)[1 - \cos(\quad 2 - \quad 3)]}{2} =$$
$$= \frac{1}{2} 2 \cos 2 \frac{2 - 3}{2} \cos(\lambda 2 - \lambda 3) - 2 \operatorname{sen} 2 \frac{2 - 3}{2} \cos(\lambda 2 + \lambda 3) \le \cos 2 \frac{2 - 3}{2} + \operatorname{sen} 2 \frac{2 - 3}{2} = 1.$$

□

Lema 6. Son válidas las siguientes desigualdades:

$$|\rho 20 - \rho 30| - (R2 + R3) e \frac{\mu T - 1}{\mu} \le \rho 23(t), \rho 32(t) \le \rho 20 + \rho 30 + (R2 + R3) e \frac{\mu T - 1}{\mu}.$$

Prueba. Dado que $\rho 2(t) = \rho 20 + \int_0^t r2(s)ds$, $\rho 3(t) = \rho 30 + \int_0^t r3(s)ds$, entonces, en vista del Lema 2,

$$\rho 23(t) = \sqrt{(\rho 2 \text{ porque } \quad 2 \text{ porque } \lambda 2 - \rho 3 \text{ porque } \quad 3 \text{ porque } \lambda 3)^2 + (\rho 2 \operatorname{sen} \quad 2 \cos \lambda 2 - \rho 3 \operatorname{sen} \quad 3 \cos \lambda 3)^2 + (\rho 2 \operatorname{sen} \lambda 2 - \rho 3 \operatorname{sen} \lambda 3)^2} \le$$
$$\le \rho 2 \sqrt{{}^2 + \rho 3^2 + 2\rho 2\rho 3A} \le \rho 2 \sqrt{{}^2 + \rho 3^2 + 2\rho 2\rho 3} = \rho 2(t) + \rho 3(t) \le \rho 20 + \rho 30 + (R2 + R3) e \frac{\mu T - 1}{\mu}$$

y

$$\rho 23(t) = \sqrt{(\rho 2 \text{ porque } \quad 2 \text{ porque } \lambda 2 - \rho 3 \text{ porque } \quad 3 \text{ porque } \lambda 3)^2 + (\rho 2 \operatorname{sen} \quad 2 \cos \lambda 2 - \rho 3 \operatorname{sen} \quad 3 \cos \lambda 3)^2 + (\rho 2 \operatorname{sen} \lambda 2 - \rho 3 \operatorname{sen} \lambda 3)^2} \ge$$
$$\ge \rho 2 \sqrt{{}^2 + \rho 3^2 - 2\rho 2\rho 3A} \ge \rho 2 \sqrt{{}^2 + \rho 3^2 - 2\rho 2\rho 3} = |\rho 2 - \rho 3| \ge |\rho 20 - \rho 30| - (R2 + R3) e \frac{\mu T - 1}{\mu}$$

lo que completa la prueba. □

Lema 7. Las siguientes desigualdades son válidas:

Machine Translated by Google

$$|\rho 2(t)-\rho 3(t-\tau 23)|\ \tau 23 \geq = \frac{\rho 02-\rho 03+\int_0^t r2(s)ds-\int_0^{t-\tau 23} r3(s)ds-\int_t^{t-\tau 23} r3(s)ds}{d_0} \geq \frac{|\rho 02-\rho 03|-(R2+R3)\int_0^t e^{\mu s}ds-\int_t^{t-\tau 23} r3(s)ds}{d_0} \geq$$

$$\frac{\rho 02-\rho 03|-(R2+R3)\, e^{\frac{|\mu T-1|}{\geq}-R3\,\mu\frac{e^{\mu(t-\tau 23)}-e^{\mu t}}{\mu}}}{d_0} \quad \frac{\rho 02-\rho 03|-(R2+R3)\, e^{\mu\geq \frac{\mu T-1-\mu\tau 23-1}{\mu}-R3e^{\mu t}e|}}{d_0} \geq$$

$$\frac{|\rho 02-\rho 03|-R2\geq \frac{m_1\mu T-1}{d_0}-R3\,\mu\frac{m_1\mu T-1}{d_0}-R3\,\mu\frac{m_1\mu T}{\mu}}{d_0} = \frac{|\rho 02-\rho 03|-R2\frac{m_1\mu T-1}{d_0}-R3\,\mu\frac{m_1\mu T}{\mu}}{d_0} \quad \frac{\rho 02-\rho 03|-(R2+R3)\, e\geq \frac{\mu T}{\mu}}{d_0} \equiv \frac{\Delta}{d_0} > 0$$

y

$$|\rho 3(t)-\rho 2(t-\tau 32)|\ \tau 32 \geq = \frac{\rho 03-\rho 02+\int_0^t r3(s)ds-\int_0^{t-\tau 32} r2(s)ds-\int_t^{t-\tau 32} r2(s)ds}{d_0} \geq \frac{|\rho 03-\rho 02|-(R2+R3)\int_0^t e^{\mu s}ds-\int_t^{t-\tau 32} r2(s)ds}{d_0} \geq$$

$$\frac{\rho 03-\rho 02|-(R3+R2)\, e^{\frac{|\mu T-1|}{\mu\geq}-R2\,\mu\frac{e^{\mu(t-\tau 32)}-e^{\mu t}}{\mu}}}{d_0} \quad \frac{\rho 03-\rho 02|-(R3+R2)\, e\,\mu\,\mu\geq c\,\frac{\mu T-1-\mu\tau 32-1}{-R2e^{\mu t}e|}}{d_0} \geq$$

$$\frac{|\rho 03-\rho 02|-R3\geq \frac{m_1\mu T-1}{d_0}-R2\,\mu\frac{m_1\mu T-1}{d_0}-R2\,\mu\frac{m_1\mu T}{\mu}}{d_0} = \frac{|\rho 03-\rho 02|-R3\frac{m_1\mu T-1}{d_0}-R2\,\mu\frac{m_1\mu T}{\mu}}{d_0} \quad \frac{\rho 02-\rho 03|-(R2+R3)\, e\geq \frac{\mu T}{\mu}}{d_0} \equiv \frac{\Delta}{d_0} > 0$$

para μ > 0 suficientemente grande y μT = const. > 0. Las desigualdades anteriores implican que | ρ2(t) − ρ3(t − τ23)|, |ρ3(t) − ρ2(t − τ32)| son estrictamente positivos, lo que completa la prueba.

Observación 3. Claramente, $\frac{}{1\ \tau 23} \leq \frac{c}{\Delta}$, $\frac{}{1\ \tau 32} \leq \frac{}{c\,\Delta}$.

5. Presentación de operadores y el lema principal

Antes de la formulación y demostración del Lema Principal, damos la condición de compatibilidad (CC), que implica la continuidad de la solución para ecuaciones neutrales: las funciones iniciales satisfacen las siguientes relaciones:

$$. r20(0) = \frac{e2e1}{m2}\frac{1}{c\rho 20\ 2(0)} -$$

$$- \frac{\rho 20(0)-\rho 30(-\tau 23(0))(\cos\ 23\cos\lambda 20(0)\cos\lambda 30(-\tau 23(0))+\sin\lambda 20(0)\sin\lambda 30(-\tau 23(0)))\ e2e3\ m2}{\xi^{\rightarrow(23)}, \rightarrow u^{(3)}} \frac{}{c\ 3\tau 23\ 3}$$

$$- \frac{. r30(-\tau 23(0))(\cos\ 23\cos\lambda 20(0)\cos\lambda 30(-\tau 23(0))+\operatorname{sen}\lambda 20(0)\operatorname{sen}\lambda 30(-\tau 23(0)))\ c\ 4\tau 23}{e2e3m2}$$

$$- \frac{\rho 30(-\tau 23(0))\ \varphi 30(-\tau 23(0))\operatorname{sen}\ 23\cos\lambda 20(0)\cos\lambda 30(-\tau 23(0))}{e2e3m2\qquad c4\tau 23}$$

$$- \frac{\rho 30(-\tau 23(0)).\ \eta 30(-\tau 23(0))(\operatorname{sen}\lambda 20(0)\cos\lambda 30(-\tau 23(0))-\cos\ 23\operatorname{sen}\lambda 30(-\tau 23(0))\cos\lambda 20(0))\ c\ 4\tau 23}{e2e3m2} - \frac{y_2^2}{m2} \quad r20(0) \equiv Gr2\ (0);$$

donde 23 = 2(0) − 3(−τ23(0)) y así sucesivamente.

Defina un operador B que actúa sobre Mr2 × Mφ2 × Mη2 × Mr3 × Mφ3 × Mη3 mediante las fórmulas

Machine Translated by Google

$$B^{(k)}_{r2}(r2, \varphi 2, \eta 2, r3, \varphi 3, \eta 3)(t) := \int_{tk}^{t} Gr2(s)ds - \frac{t-Tk}{t} - \frac{1}{2} \int_{tk}^{Tk+1} Gr2(s)ds - \frac{1}{t} \int_{tk}^{Tk+1} \int_{tk}^{pag} Gr2(s)dsdp, \quad t \in [Tk, Tk+1];$$

$$r3(t) = r30(t), \varphi 3(t) = \varphi 30(t), \eta 3(t) = \eta 30(t), \quad t \in [-T, 0]$$

$$B^{(k)}_{\varphi 2}(r2, \varphi 2, \eta 2, r3, \varphi 3, \eta 3)(t) := \varphi 20 + \int_{tk}^{t} G\varphi 2(s)ds - \frac{t-Tk}{t} - \frac{1}{2} \int_{tk}^{Tk+1} G\varphi 2(s)ds - \frac{1}{t} \int_{tk}^{Tk+1} \int_{tk}^{p} G\varphi 2(s)dsdp, \quad t \in [Tk, Tk+1];$$

$$r3(t) = r30(t), \varphi 3(t) = \varphi 30(t), \eta 3(t) = \eta 30(t), \quad t \in [-T, 0]$$

$$B^{(k)}_{\eta 2}(r2, \varphi 2, \eta 2, r3, \varphi 3, \eta 3)(t) := \int_{tk}^{t} Gr3(s)ds - \frac{t-Tk}{t} - \frac{1}{2} \int_{tk}^{Tk+1} Gr3(s)ds - \frac{1}{t} \int_{tk}^{Tk+1} \int_{tk}^{pag} Gr3(s)dsdp, \quad t \in [Tk, Tk+1];$$

$$r3(t) = r30(t), \varphi 3(t) = \varphi 30(t), \eta 3(t) = \eta 30(t), \quad t \in [-T, 0]$$

$$B^{(k)}_{r3}(r2, \varphi 2, \eta 2, r3, \varphi 3, \eta 3)(t) := \int_{tk}^{t} Gr3(s)ds - \frac{t-Tk}{t} - \frac{1}{2} \int_{tk}^{Tk+1} Gr3(s)ds - \frac{1}{t} \int_{tk}^{Tk+1} \int_{tk}^{pag} Gr3(s)dsdp, \quad t \in [Tk, Tk+1];$$

$$r2(t) = r20(t), \varphi 2(t) = \varphi 20(t), \eta 2(t) = \eta 20(t), \quad t \in [-T, 0]$$

$$B^{(k)}_{\varphi 3}(r2, \varphi 2, \eta 2, r3, \varphi 3, \eta 3)(t) := \varphi 30 + \int_{tk}^{t} G\varphi 3(s)ds - \frac{t-Tk}{t} - \frac{1}{2} \int_{tk}^{Tk+1} G\varphi 3(s)ds - \frac{1}{t} \int_{tk}^{Tk+1} \int_{tk}^{p} G\varphi 3(s)dsdp, \quad t \in [Tk, Tk+1];$$

$$r2(t) = r20(t), \varphi 2(t) = \varphi 20(t), \eta 2(t) = \eta 20(t), \quad t \in [-T, 0]$$

$$B^{(k)}_{\eta 3}(r2, \varphi 2, \eta 2, r3, \varphi 3, \eta 3)(t) := \int_{tk}^{t} G\eta 3(s)ds - \frac{t-Tk}{t} - \frac{1}{2} \int_{tk}^{Tk+1} G\eta 3(s)ds - \frac{1}{t} \int_{tk}^{Tk+1} \int_{tk}^{pag} G\eta 3(s)dsdp, \quad t \in [Tk, Tk+1]$$

$$r2(t) = r20(t), \varphi 2(t) = \varphi 20(t), \eta 2(t) = \eta 20(t), \quad t \in [-T, 0]$$

donde $t \in [Tk, Tk+1]$, $Tk = kT$ $(k = 0, 1, 2, \ldots)$, $G\varphi 2 = G\varphi 2(r2, \varphi 2, \eta 2, r3, \varphi 3, \eta 3)$, $G\eta 2 = G\eta 2(r2, \varphi 2, \eta 2, r3, \varphi 3, \eta 3)$, $Gr3 = Gr3(r2, \varphi 2, \eta 2, r3, \varphi 3, \eta 3)$, $G\varphi 3 = G\varphi 3(r2, \varphi 2, \eta 2, r3, \varphi 3, \eta 3)$, $G\eta 3 = G\eta 3(r2, \varphi 2, \eta 2, r3, \varphi 3, \eta 3)$ y las funciones iniciales $rn0(t)$, $\varphi n0(t)$, $\eta n0(t)$, $(n = 2, 3)$ son soluciones del sistema de ecuaciones iniciales.

Observación 4. Observamos que en vista de la periodicidad de Gr2 (s), el operador Br2 (.)(t) es T-periódico en [0, ∞) (cf. [20]). La condición de conformidad (CC) . r20(0) = Gr2 (0) implica

$$. r2(t) = Gr2(t) - \frac{1}{T} \int_{0}^{t} Gr2(s)ds \quad . r20(0) = Gr2(0) - \frac{1}{T} \int_{0}^{t} Gr2(s)ds \int_{0}^{t} Gr2(s)ds = 0$$

y

$$= B(.)(0) \quad r2(0) = (0) \; (0) \; r2(t) = B(.)(t) \quad r2(0)$$

$$\frac{1}{2} \int_{0}^{t} Gr2(s)ds - \frac{1}{T} \int_{0}^{t} \int_{0}^{pag} Gr2(s)dsdp \int_{0}^{tp} Gr2(s)dsdp = 0$$

$$(0)_{(.)(0) = B\, r2} \quad \frac{1}{2} \int_{0}^{t} Gr2(s)ds - \frac{1}{T} \int_{0}^{t} \int_{0}^{pag} Gr2(s)dsdp = 0; \; B^{(0)}_{r2}(.)(T) = \frac{1}{2} \int_{0}^{t} Gr2(s)ds - \frac{1}{T} \int_{0}^{t} \int_{0}^{0p} Gr2(s)dsdp = 0.$$

Pero Gr2 (s) es una función T-periódica. Por lo tanto, se puede ampliar el resultado para cada $k = 1, 2, \ldots, T$ $k+1$ T $k+1$ $Gr2(s)ds =$ es decir, $\int_{tk}^{pag} Gr2(s)dsdp = 0$.

Necesitamos las siguientes desigualdades:

Machine Translated by Google

$$\left\| \xi^{(23)}, (3).\to tu \right\| \le \left\| \xi^{(23)}, \xi^{(23)} \right\| \left\| (3).\to tu, (3).\to tu \right\| \le c\tau 23.. \rho 3 \sqrt{\rho 2 + \rho 3^2 + 2.. \rho 3^2 2\cos 2\lambda 3 + \rho 3^2 3^2 \lambda 3^2} =$$

$$= c\tau 23. r3 \sqrt{\rho 2 + \rho 3^2 + 2. \rho 3^2 2\cos 2\lambda 3 + \rho 3^2 \varphi 3^2} \cdot \eta 3^2 c\tau 23\omega 3e \mu(t-Tk) R3 \le \sqrt{\rho 3^2 R3^2 2 + \rho 3^2 Y3^2 2} \le cc\tau 23\omega 3e - \mu(t-Tk) +$$

$$\left\| \xi^{(32)}, (2).\to tu \right\| \le \left\| \xi^{(32)}, \xi^{(32)} \right\| \left\| (2).\to tu, (2).\to tu \right\| \le c\tau 32.. \rho 2 \sqrt{\rho 2 + \rho 2^2 + 2.. \rho 2^2 2\cos 2\lambda 2 + \rho 2^2 2\lambda 2^2} =$$

$$= c\tau 23. r3 \sqrt{\rho 2 + \rho 3^2 + 2. \rho 3^2 2\cos 2\lambda 3 + \rho 3^2 \varphi 3^2} \cdot \eta 3^2 c\tau 32\omega 2e \mu(t-Tk) R2 \le \sqrt{\rho 2^2 R2^2 2 + \rho 2^2 Y2^2 2} \le cc\tau 32\omega 2e - \mu(t-Tk) +$$

Además, la condición (C) implica $\rho n 2\Phi n 2 + \rho n 2Yn 2 \le \left\| \tilde{tu}^{(norte)}, \tilde{tu}^{(norte)} \right\|^2 = rn^2 + \rho n 2\varphi n 2 \cos 2\lambda n + \rho n 2\eta n^2 \le Rn^2 c^2 - 2\,2 < c^2$, (norte = 2, 3).

Lema 8. Las siguientes desigualdades son válidas:

$$\int_{tk}^{t} Gr2(s)ds \le \frac{|e2e1|}{m2c} \frac{1}{\sqrt{2 \times 2}} \sqrt{m2} \sqrt{\frac{|e2e3|\,|e2e3|\,\rho 2+\rho 3\; c c\tau 23\omega 3e\; e\; 2 + + + 6\; c\; 4\tau 23^2 3\sigma 3|\rho 3^2 \rho 3 + |.\ \eta 3\ \rho 3|\ \mu(t-Tk)\ 2\ |}{3\ \tau 23}} \frac{\sqrt{m2}}{r2}\, ds \le$$

$$\le \frac{|e2e1|}{m2} \frac{1}{\sqrt{do\ 2 \times 2}} c\omega 3\sqrt{(+\ \rho\ 2+\ \ \rho\ 3+3\Delta)\ \omega 2\ 2R2 + c\ 3\Delta 2}\ \sqrt{mi}\mu(s-Tk)ds \le \frac{|e2e1|}{m2} \frac{1}{\sqrt{do\ 2 \times 2}} c\omega 3\sqrt{(\ \rho\ 2+\ \ \rho\ 3+3\Delta)\ \omega 2\ 2R2 + c\ 3\Delta 2}\ \frac{\sqrt{mi}\mu(t-Tk)}{\mu};$$

$$\int_{tk}^{Tk+1} Gr2(s)ds \le \frac{|e2e1|}{m2} \frac{1}{\sqrt{do\ \rho\ 2}} c\omega 3\sqrt{(+\ \rho\ 2+\ \ \rho\ 3+3\Delta)\ \omega 2\ 2R2 + c\ 2\Delta 2}\ \frac{mi\ \mu T-1}{\mu};$$

$$\int_{tk}^{t} G\varphi 2\, ds \le \frac{|e2e3|}{m2} \frac{\rho 3 c\omega 3}{c\ 5\tau 23\ 2\rho 2\cos\lambda 2} + \frac{|\ .\ r3|+\ \rho 3\ \varphi 3\ +|\rho 3\ .\ \eta 3|\ c}{4\tau 23\rho 2\cos\lambda 2} + \frac{\varphi 2}{\sqrt{mi}}\mu(s-Tk)ds \le \frac{|e2e3|}{m2} \frac{c\omega 3\sqrt{(\ \omega 2\ 2R2 + c\ 3\Delta 2\cos(\pi/2-\delta)\ \rho\ 2}}{\sqrt{mi}}\frac{\mu(t-Tk)}{\mu};$$

$$\int_{tk}^{Tk+1} G\varphi 2\, ds \le \frac{|e2e3|}{m2} \frac{c\omega 3\sqrt{(\ \omega 2\ 2R2 + c\ 3\Delta 2\cos(\pi/2-\delta)}}{\rho\ 2} \frac{mi\ \mu T-1}{\mu};$$

$$\int_{tk}^{t} G\eta 2 ds \le \frac{|e2e3|}{e2e3m2} \frac{\mu(s-Tk)\ \omega 2\ 2Y2\ \rho 3 c\omega 3e\ +\ e\ 5\tau 23\ 2\rho 2\ c\ 4\tau 23\rho 2}{\sqrt{mi}}\mu(s-Tk)ds \le \frac{|e2e3|}{m2} \frac{e\ \mu T+3\Delta\ \omega 2\ 2Y2\ c\omega 3\ \rho\ 3\ +\ 3\ c}{\rho 2\Delta 2} \frac{\sqrt{mi}\mu(t-Tk)}{\mu};$$

$$\int_{Tk+1} G\eta 2\, ds \le m2 \frac{|e2e3|}{\rho 2\Delta 2} \frac{+\ 3\ c\ c\omega 3\ e\ \mu T+3\Delta\ \rho\ 3}{\sqrt{\ }} \frac{\omega 2\ 2Y2}{\sqrt{\ }} \frac{mi\ \mu T-1}{\mu};$$

$$\int_{tk}^{t} Gr3(s)ds \le \frac{|e2e1|}{m3} \frac{1}{\sqrt{do\ 2 \times 3}} \frac{c\omega 2(\ \rho\ 3+3\Delta)\ \omega 3\ 2R3\ \rho\ 2+ + + c\ 2\Delta 2}{\sqrt{\ }} \frac{\sqrt{mi}\mu(t-Tk)}{\mu};$$

$$\int_{tk}^{Tk+1} Gr3(s)ds \le \frac{|e2e1|}{m3} \frac{1}{\sqrt{do\ \rho\ 3}} \frac{c\omega 2(\ \rho\ 3+3\Delta)\ \omega 3\ 2R3\ \rho\ 2+ + + c\ 2\Delta 2}{\sqrt{\ }} \frac{mi\ \mu T-1}{\mu};$$

$$\int_{tk}^{t} G\varphi 3\, ds \le \frac{|e2e3|}{m3} \frac{c\omega 2(\ \omega 2\ 2R2 + c\ 3\Delta 2\cos(\pi/2-\delta)}{\rho\ 3} \frac{\sqrt{mi}\mu(t-Tk)}{\mu};$$

$$\int_{tk}^{Tk+1} G\varphi 3\, ds \le \frac{|e2e3|}{m3} \frac{c\omega 2(\ \omega 2\ 2R2 + c\ 3\Delta 2\cos(\pi/2-\delta)\ \rho\ 3}{\sqrt{\ }} \frac{mi\ \mu T-1}{\mu};$$

$$\int_{tk}^{t} G\eta 3 ds \le \frac{|e2e3|}{m3} \frac{e\ \mu T+3\Delta\ \omega 3\ 2Y3\ \rho\ 2\ c\omega 2\ +\ 3\ c}{\rho\ 3\Delta 2} \frac{\sqrt{mi}\mu(t-Tk)}{\mu};$$

$$\int_{Tk+1} G\eta 3 ds \le \frac{|e2e3|}{m3} \frac{e\ \mu T+3\Delta\ \omega 3\ 2Y3\ \rho\ 2\ c\omega 2\ +\ 3\ c}{\rho\ 3\Delta 2} \frac{mi\ \mu T-1}{\mu}.$$

Lema 9 (Lema principal). Deje que se cumpla la condición de compatibilidad (CC). El problema periódico para el sistema (BS) tiene solución (r2, φ2, η2,r3, φ3, η3) $Mr2 \times M\varphi 2 \times M\eta 2 \times Mr3 \times M\varphi 3 \times M\eta 3$ si el operador B tiene un punto fijo perteneciente a $Mr2 \times M\varphi 2 \times M\eta 2 \times Mr3 \times M\varphi 3 \times M\eta 3$.

Prueba. Sea el sistema (BS) una solución que pertenece a $Mr2 \times M\varphi 2 \times M\eta 2 \times Mr3 \times M\varphi 3 \times M\eta 3$.

□

Machine Translated by Google

Demostramos que B = Br2 , Bφ2 , Bη2 , Br3 , Bφ3 , Bη3 posee un punto fijo perteneciente a Mr2 × Mφ2 × Mη2 × Mr3 × Mφ3 × Mη3 . De hecho, integrando la primera ecuación de (BS), tenemos

obtener $r2(t) = \int_{tk}^{t} Gr2(s)ds$, y por lo tanto, en vista de la Observación 4, se obtiene

$$r2(t) = \int_{tk}^{t} Gr2(s)ds - \frac{t - Tk}{t} \left[\frac{1}{2} \int_{tk}^{Tk+1} Gr2(s)ds - \frac{1}{t} \int_{tk}^{Tk+1} \int_{tk}^{p} Gr2(s)dsdp \right],$$

es decir, r2 = Br2 (k) (r2, φ2, η2,r3, φ3, η3).

Para el segundo componente φ2, tenemos $\int_{tk}^{Tk+1} Gφ2(s)ds = 0$. Por lo tanto,

$$φ2(Tk+1) = φ2(Tk) + \int_{tk}^{Tk+1} Gφ2(s)ds \quad φ02 = φ02 + \int_{tk}^{Tk+1} Gφ2(s)ds \quad \int_{tk}^{Tk+1} Gφ2(s)ds = 0.$$

Por lo tanto, $φ2(t) = φ02 + \int_{tk}^{t} Gφ2(s)ds - \left[\frac{t-Tk}{t} \right] \frac{1}{2} \int_{tk}^{Tk+1} Gφ2(s)ds$. Del mismo modo,

$$\int_{tk}^{Tk+1} \int_{tk}^{p} Gφ2(s)dsdp = \int_{tk}^{Tk+1} \int_{tk}^{s} Gφ2(s)dpds = \int_{tk}^{Tk+1} (Tk+1 - s)Gφ2(s)ds = Tk+1 \int_{tk}^{Tk+1} Gφ2(s)ds - \int_{tk}^{Tk+1} sGφ2(s)ds = -\int_{tk}^{Tk+1} s\frac{dφ2(s)}{ds}ds = -\left[Tk+1 \, sdφ2(s) = -Tk+1φ2(Tk+1) + Tk\, φ2(Tk) + \int_{tk}^{Tk+1} φ2(s)ds = -Tφ02 + T\, φ02(Tk) + \int_{tk}^{Tk+1} φ2(s)ds = -Tφ02 + Tφ02 = 0. \right]$$

Por lo tanto,

$$φ2(t) = φ02 + \int_{tk}^{t} Gφ2(s)ds - \frac{t - Tk}{t} \left[\frac{1}{2} \int_{tk}^{Tk+1} Gφ2(s)ds - \frac{1}{t} \int_{tk}^{Tk+1} \int_{tk}^{p} Gφ2(s)dsdp \right]$$

lo que significa φ2 = Bφ2 (r2, φ2, η2,r3, φ3, η3).

Repitiendo el mismo razonamiento obtenemos

η2 = Bη2 (r2, φ2, η2,r3, φ3, η3), r3 = Br3 (r2, φ2, η2,r3, φ3, η3), φ3 = Bφ3 (r2, φ2, η2,r3, φ3, η3) , η3 = Bη3 (r2, φ2, η2,r3, φ3, η3)).

Por el contrario, sea (r2, φ2, η2,r3, φ3, η3) Mr2 × Mφ2 × Mη2 × Mr3 × Mφ3 × Mη3 un punto fijo de B. Entonces,

$$r2(t) = \int_{tk}^{t} Gr2(s)ds - \frac{t - Tk}{t} \left[\frac{1}{2} \int_{tk}^{Tk+1} Gr2(s)ds - \frac{1}{t} \int_{tk}^{Tk+1} \int_{tk}^{s} Gr2(θ)dθds. \right]$$

da para t = Tk

$$r2(Tk) = \int_{tk}^{tk} Gr2(s)ds - \frac{Tk - Tk}{t} \left[\frac{1}{2} \int_{tk}^{Tk+1} Gr2(s)ds - \frac{1}{t} \int_{tk}^{Tk+1} \int_{tk}^{s} Gr2(θ)dθds \right] \quad 0 = \frac{1}{2} \int_{tk}^{Tk+1} Gr2(s)ds - \frac{1}{t} \int_{tk}^{Tk+1} \int_{tk}^{s} Gr2(θ)dθds.$$

Pero la observación 4 implica $\int_{Tk}^{Tk+1} Gr2(s)ds = 0 \quad \frac{1}{t} \int_{tk}^{Tk+1} \int_{tk}^{s} Gr2(θ)dθds = 0.$

En consecuencia, $r2(t) = \int_{tk}^{t} Gr2(s)ds$. r2(t) = Gr2 (t).

Para el resto de componentes de B se procede del mismo modo.

Así queda demostrado el lema 9.

Machine Translated by Google

6. El resultado principal

El principal resultado del trabajo es el siguiente:

Teorema 1 (Teorema principal). Deje que se cumpla la condición de compatibilidad (CC) y que los datos iniciales sean suficientemente pequeños (Apéndice C). Si se satisfacen las siguientes desigualdades:

$$(1)\quad \frac{|e2e1|}{m2}\,\frac{1}{\rho_{\infty}\,2^{2}}\,\frac{\bar{c}\omega 3(\rho\,2+\ \rho\,3+3|\Delta|)\,\omega 2\,2R2+}{c\,3\Delta 2}\,\frac{m\mu T}{\mu}\le R2\ (\mu>\omega 2,\ \omega 3);$$

$$\varphi 20+\ \frac{|e2e1|}{m2}\,\frac{\bar{c}\omega 3(\ \rho\,2+\rho\,3+3|\Delta|)3\Delta 2\cos(\pi/2-\delta)\,\rho\,2}{}\,\frac{m\mu T}{\mu}\le \Phi 2\ (\varphi 20>0);$$

$$\frac{|e2e1|}{m2}\,\frac{\bar{c}\,\omega 3}{c^{3}}\,\frac{\rho\,3\ mi\ \mu T+3|\Delta|\ \omega 2\ 2Y2+}{\rho\ 2\Delta 2}\,\frac{m\mu T}{\mu}\le Y2;$$

$$\frac{|e3e1|}{m3}\,\frac{1}{\rho_{\infty}\,3^{2}}\,\bar{c}\omega 2(\ \rho\ 3+3)\ \omega 3\ 2R3\ \rho\ 2+\ +\ c\ 3\Delta 2\,\frac{m\mu T}{\mu}\le R3;$$

$$\varphi 30+\ \frac{|e3e1|}{3m3c}\,\frac{\beta\omega 2\ \rho\ 2+3c\Delta+\omega 3}{3\ e^{\ c\Delta 2\ \rho}}\,2\Phi 3\ e\cos(\pi/2-\delta)\,\frac{\mu T}{\mu}\le \Phi 3\ (\varphi 30>0);$$

$$\frac{|e3e1|}{3m3c}\,\bar{}\,\frac{\mu T+3\Delta\ \omega 3\ 2Y3}{\rho\ 3\Delta 2\ |}\,c\omega 2\ \rho\ 2+3\ c\,\frac{m\mu T}{\mu}\le Y3;$$

$$(2)\quad 1+\frac{e\ \mu T-1}{\mu T}\,\frac{|e2e1|}{m2}\,\frac{1}{\rho_{\infty}\,2\ x\ 2}\,\frac{\bar{c}\omega 3(\ \rho\ 2+\ \rho\ 3+3|\Delta|)\ c}{3\Delta 2}+\frac{\omega 2\ 2R2}{}\,\frac{m\mu T}{\mu}\le \omega 2R2;$$

$$1+\frac{e\ \mu T-1}{\mu T}\,\frac{|e2e1|}{3m2c}\,\frac{\bar{c}\omega 3\quad \rho\ 3+3|\Delta|+}{\rho_{\infty}^{2}\ \rho\ 2}\,\omega 2\ 2\Phi 2\le \omega 2\Phi 2;\ \cos(\pi/2-\delta)\Delta 2$$

$$1+\frac{e\ \mu T-1}{\mu T}\,\frac{|e2e1|}{3m2c}\,\frac{\bar{c}\omega 3\quad \rho\ 3+3|\Delta|}{^{2}\ \rho^{s}2\Delta 2}+\omega 2\ 2Y2\le \omega 2Y2;$$

$$1+\frac{e\ \mu T-1}{\mu T}\,\frac{|e3e1|}{m3c}\,\frac{1}{2\ x\ 3}+\frac{\bar{c}\omega 2\quad 2+\ c\ \ \rho\ 3+3|\Delta|\ \rho}{3\Delta 2}+\frac{\omega 3\ 2R3}{}\le \omega 3R3\ ;$$

$$1+\frac{e\ \mu T-1}{\mu T}\,\frac{|e3e1|}{3m3c}\,\frac{\bar{c}\omega 2\quad \rho\ 2+3|\Delta|+}{\rho_{\infty}}\,\omega 3\ 2\Phi 3\le \omega 3\Phi 3;\ 2\quad \cos(\pi/2-\delta)\Delta 2\ \rho\ 3$$

$$1+\frac{e\ \mu T-1}{\mu T}\,\frac{|e3e1|}{3m3c}\,\frac{\bar{c}\omega 2\quad \rho 2\ +3|\Delta|}{\rho_{\infty}^{2}\ \rho\ 3\Delta 2}+\omega 3\ 2Y3\le \omega 3Y3$$

entonces, existe una solución T-periódica única (r2, φ2, η2,r3, φ3, η3)

Mr2 × Mφ2 × Mη2 × Mr3 × Mφ3 × Mη3 de (BS).

Observación principal. Observemos que tenemos un número infinito de desigualdades en el teorema. Las desigualdades anteriores implican que las funciones operadoras y sus primeras derivadas pertenecen a Mr2 × Mφ2 × Mη2 × Mr3 × Mφ3 × Mη3 . Es fácil comprobar que para las terceras derivadas obtenemos 3 y así sucesivamente. respectivamente ω2 y ω2
4

Prueba. Primero, mostramos que B mapea el conjunto Mr2 × Mφ2 × Mη2 × Mr3 × Mφ3 × Mη3 dentro de sí mismo. □

Observación 4, tenemos $B_{r2}(k)(.)(Tk) = De\ la\ \bar{1}_2 \int_{tk}^{Tk+1} Gr2(s)ds - \frac{1}{t}\int_{tk}^{Tk+1}\int_{tk}^{s} Gr2(\theta)d\theta ds = 0$,(k = +1 0, 1, 2, . . .) Y desde $\int_{tk}^{Tk} \frac{t-Tk}{t} - \bar{1}_2\, dt = 0$, obtenemos

$$B_{r2}(k)(.)(t)\,dt = \int_{tk}^{Tk+1}\int_{tk}^{t} Gr2(s)ds\,dt - \int_{tk}^{Tk+1}\frac{t-Tk}{t}-\frac{1}{2}\,dt\int_{tk}^{Tk+1} Gr2(s)ds - \int_{tk}^{Tk+1}\int_{tk}^{\theta} Gr2(s)ds d\theta = 0\ .$$

Además, en vista de $\rho \le c$ y las $\bar{}$ desigualdades de la Observación 3, tenemos

Machine Translated by Google

$$B^{(k)}_{(.)}(t) \le |Gr2(s)|ds + r2\,Tk\,\mu(t-Tk)\,|e2e1| \le mi\,\frac{1}{\bar{1}2}\int_{tk}^{Tk+1} Gr2(s)ds + \frac{1}{t}\int_{tk}^{Tk+1}\int_{tk}^{t} Gr2(s)dsdt = \int_{tk}^{t}|Gr2(s)|ds + \int_{tk}^{Tk+1} Gr2(s)ds \le$$

$$\frac{1}{m2c}\,\frac{1}{2\,x\,2}\,\frac{\beta\omega3(\rho\,3+3\Delta)\,\omega2\,2R2\,\rho\,2+}{}\,\frac{+\,c\,2\Delta2}{}\,\frac{}{1\,\mu}+\frac{mi\,\mu T-1}{\mu} \le$$

$$m2c\;\mu(t-Tk)\,|e2e1| \le e\,\frac{1}{2\,x\,2}\,\frac{\beta\omega3(+p\,2+\;\rho\,3+3\Delta)\,c}{2\Delta2}+\frac{\omega2\,2R2}{}\,\frac{}{\mu}\le R2e\,\mu(t-Tk)\,.$$

Para el segundo componente de B, tenemos

$$\varphi^{(k)}_{(.)}(Tk) = \varphi02 + B\,\varphi2\,\frac{1}{\bar{1}2}\int_{tk}^{Tk+1} G\varphi2(.)(s)ds - \frac{1}{t}\int_{tk}^{Tk+1}\int_{tk}^{s} G\varphi2(.)(\theta)d\theta ds = \varphi02,\ (k=0,1,2,\ldots);$$

$$\int_{tk}^{T}\varphi^{k+1\,(k)}_{(.)}(t)\,dt = T\varphi02 + B\,\varphi2\int_{tk}^{Tk+1t} G\varphi2(s)ds - \int_{tk}^{Tk+1}\frac{t-Tk}{t}\,\frac{}{\bar{1}2}\,{}^{k+1}\,dt\int_{tk}^{Tk+1}G\varphi2(s)ds - \int_{tk}^{Tk+1}\int_{tk}^{t} G\varphi2(s)dsdt = T\varphi02$$

y

$$B^{(k)}_{(.)}(t) \le \varphi02 + \varphi2\int_{tk}^{t} G\varphi2(s)\,ds + \int_{tk}^{Tk+1} G\varphi2(s)ds \le e^{\mu(t-Tk)}\,\varphi02 + \frac{|e2e3|}{3m2c}\,\frac{\beta\omega3 + 3c\Delta\,\rho\,3 + \omega2}{\rho\,2^{c\Delta2}}\,\frac{2\Phi2\cos(\pi/2-\delta)}{}\,\frac{e^{\mu T}}{\mu} \le \Phi2e\,\mu(t-Tk)\,.$$

Para el tercer componente de B, obtenemos

$$\eta^{(k)}_{(.)}(t) \le B\,\eta2\int_{tk}^{t} G\eta2(s)\,ds + \int_{tk}^{Tk+1} G\eta2(s)ds \le e^{\mu(t-Tk)}\,\frac{|e2e1|}{m2}\,\frac{\bar{c}\omega3}{3c}\,\frac{\rho\,3\;{}^{mi}\mu T+3|\Delta|}{\rho2\Delta2}+\frac{\omega2\,2Y2}{}\,\frac{e^{\mu T}}{\mu} \le Y2e\,\mu(t-Tk)\,.$$

Para los tres componentes restantes de B, tenemos

$$\mu(t-Tk)\,|e3e1|\,\eta^{(k)}_{(.)}(t) \le e\,B\,r3\,\frac{1}{m3c}\,\frac{}{2\,x\,3}\,\frac{\beta\omega2(+p\,2+\;\rho\,3+3)\,\omega3\,2R3+}{c\,2\Delta2}\,\frac{}{}\,\frac{}{\mu}\le R3e\,\mu(t-Tk)\,;$$

$$\mu(t-Tk)\,|e3e1|\,\mu(t-Tk)\,\varphi03 + e\,\eta^{(k)}_{(.)}(t) \le e\,B\,\varphi3\,\frac{}{3m3c}\,\frac{\beta\omega2\,\rho\,2+3c\Delta+\omega3}{c\Delta2\,\rho\,3}\,2\Phi3\,e\cos(\pi/2-\delta)\,\frac{\mu T}{\mu} \le \Phi3e\,\mu(t-Tk)\,;$$

$$B\eta3^{(k)}_{(.)}(t) \le \mu(t-Tk)\,|e3e1|\,\frac{}{3m3c}\,\frac{\bar{c}\omega2\;\;\rho2+3|\Delta|}{{}^{\omega}2\;\rho\,3\Delta2}+\omega3\,2Y3\,e\,\frac{\mu T}{\mu} \le Y3e\,\mu(t-Tk)\,.$$

Para las derivadas de primer orden de los componentes de B, obtenemos

$$\frac{dB^{(k)}_{(.)}(t)}{dt} \le |Gr2(.)(t)| + \frac{1}{t}\int_{tk}^{T}\varphi^{k+1}\,e^{\mu T-1}\,\mu(t-Tk)\,1 + Gr2(.)(s)ds \le e^{\mu T}\,\frac{}{m2c}\,\frac{|e2e1|}{}\,\frac{1}{2\,x\,2}+\frac{\bar{c}\omega3\;\;\rho\,2+\;\rho\,3+3\Delta}{c\,3\Delta2}\,\frac{\omega2\,2R2+}{}\le \omega2R2e\,\mu(t-Tk)\,;$$

$$\frac{dB^{(k)}_{\varphi2}(.)(t)}{dt} \le G\varphi2(.)(t) + \frac{1}{t}\int_{tk}^{Tk+1}(.)(s)ds \le e^{\mu T}\,\mu(t-Tk)\,1 + G\varphi2\,\frac{e^{\mu T-1}}{}\,\frac{|e2e3|}{3m2c}\,\frac{\bar{c}\omega3\;\;\rho\,3+3\Delta}{2-\delta)\Delta2\,\rho\,2}+\omega2\,2\Phi2 \le \omega2\Phi2e\,\mu(t-Tk)\,\cos(\pi/\;;$$

$$\frac{dB^{(k)}_{\eta2}(.)(t)}{dt} \le G\eta2(.)(t) + \frac{1}{t}\int_{tk}^{Tk+1}(.)(s)ds \le e^{\mu T}\,\mu(t-Tk)\,1 + G\eta2\,\frac{e^{\mu T-1}}{}\,\frac{|e2e3|}{3m2c}\,\frac{\bar{c}\omega3\;\;\rho\,3+3\Delta e^{\mu T}}{\rho2\Delta2}+\omega2\,2Y2 \le \omega2Y2e\,\mu(t-Tk)\;;$$

$$\frac{dB^{(k)}_{r3}(.)(t)}{dt} \le mi\mu(t-Tk)\,1 + \frac{mi\,\mu T-1}{\mu T}\,\frac{|e3e1|}{m3c}\,\frac{1}{\rho\,3^{2}}+\frac{\bar{c}\omega2\;\;\rho\,2+\;\rho\,3+3\Delta}{c\,3\Delta2}\,\frac{\omega3\,2R3+}{}\le \omega3R3e\,\mu(t-Tk)\,;$$

$$\frac{d\eta^{(k)}_{(.)}(t)\,\varphi3dt}{}\;\mu(t-kT)\,1+\le e^{\mu T}\,\frac{e^{\mu T-1}}{}\,\frac{|e2e3|}{3m3c}\,\frac{\bar{c}\omega2\;\;\rho\,2+3\Delta}{\cos(\pi/2-\delta)\Delta2\,\rho\,3}+\omega3\,2\Phi3 \le \omega3\Phi3e\,\mu(t-Tk)\;;$$

$$\frac{dB^{(k)}_{\eta3}(.)(t)}{dt} \le mi\,\mu(t-Tk)\,|e2e3|\,\frac{}{3m3c}\,\frac{\bar{c}\omega2\;\;\rho2+3\Delta}{}\,\frac{}{\rho\,3\Delta2}+\omega3\,2Y3 \le \omega3Y3e\,\mu(t-Tk)\;.$$

Las estimaciones de las derivadas de orden superior de los componentes de B se obtienen en el Apéndice D.

A continuación, mostramos que el operador B es contractivo en el sentido de [8].

Machine Translated by Google

En vista de la Observación 2, las funciones retardadas actúan de la siguiente manera:

$$t - \tau 23(t) : [Tk , Tk+1] \to [Tk-1 , Tk], t - \tau 32(t) : [Tk , Tk+1] \to [Tk-1 , Tk].$$

El conjunto de índices de Mr2 × Mφ2 × Mη2 × Mr3 × Mφ3 × Mη3 es (k, m), donde k, m = 0, 1, 2, . . .

Introducimos las siguientes asignaciones del conjunto de índices en sí mismo: j1(k, m) = (k, m), j2(k, m) = (k − 1, m + 1), j3(k, m) = (k, metro + 2). $\partial Gr2$ $\partial Gr2$ $\partial Gr2$ $\partial Gr2$ $\partial Gr2$ $\partial Gr2$ $\partial Gr2$ $\partial \rho 2$, ∂ 2 , $\partial \lambda 2$, $\partial \rho 3$, ∂ 3 , $\partial \lambda 3$, ∂

Denotemos por $\overline{\partial Gr2} = \max$ (cf. $r3 \overline{\quad}, \overline{\quad}, \overline{\quad}, \overline{\quad}, \overline{\quad}, \overline{\quad}, \overline{\quad}, \frac{\partial Gr2}{\partial \varphi 3}, \frac{\partial Gr2}{\partial \eta 3}, \frac{\partial Gr2}{\partial r2}$

Apéndice E). Entonces,

$$\overline{Br2}\,(r2, \varphi 2, \eta 2, r3, \varphi 3, \eta 3) - \overline{Br2}\,(r2, \varphi 2 , \eta 2 , r3, \varphi 3 , \eta 3) \le$$

$$\le \frac{\partial Gr2}{\partial \rho 2} \int_{tk}^{t} \partial Gr2\, |\rho 2(s) - \rho 2\,(s)|ds + \partial \overline{2} \int_{tk}^{t} |\lambda 2(s) - \lambda 2(s)|ds + \partial \lambda 2\, \partial \rho 3 \int_{tk}^{t} \partial Gr2\, \partial Gr2\, |\rho 3(s - \tau 23(s)) - \rho 3\,(s - \tau 23(s))|ds +$$

$$+ \frac{\partial Gr2}{\partial 3} \int_{tk}^{t} \partial Gr2\, |3(s - \tau 23(s)) - 3\,(s - \tau 23(s))|ds + \partial \lambda 3 \int_{tk}^{t} |\lambda 3(s - \tau 23(s)) - \lambda 3(s - \tau 23(s))|\, ds + \partial\, r3 \int_{tk}^{t} |r3(s - \tau 23(s)) - r3(s - \tau 23(s))|\, ds +$$

$$+ \frac{\partial Gr2}{\partial \varphi 3} \int_{tk}^{t} |\varphi 3\,(s - \tau 23(s)) - \varphi 3\,(s - \tau 23(s))|\, ds + \partial\, \eta 3 \int_{tk}^{t} |\eta 3\,(s - \tau 23(s)) - \eta 3\,(s - \tau 23(s))|\, ds + \partial\, r2 \int_{tk}^{t} |r2(s) - r2(s)|\, ds \le Tk$$

$$\le \frac{\partial Gr2}{\partial \rho 2} \int_{tk}^{t} \int_{tk}^{s} |r2(p) - r2(p)| e^{-\mu(p-Tk)}\, m_i \mu(p-Tk)dpds + \partial\, 2 \int_{tk}^{t} \int_{tk}^{s} |\varphi 2(p) - \varphi 2\,(p)|\, e^{-\mu(p-Tk)}\, \mu(p-Tk)dpds +$$

$$+ \frac{\partial Gr2}{\partial \lambda 2} \int_{tk}^{t} \int_{tk}^{s} |\eta 2(p) - \eta 2\,(p)| e^{-\mu(p-Tk)}\, m_i \mu(p-Tk)dpds + \partial \rho 3 \int_{tk}^{t} \int_{tk}^{s-\tau 23(s)} |r3(p) - r3(p)| e^{-\mu(p-Tk)}\, m_i \mu(p-Tk)dpds +$$

$$+ \frac{\partial Gr2}{\partial 3} \int_{tk}^{t} \int_{tk}^{s-\tau 23(s)} |\varphi 3(p) - \varphi 3\,(p)|\, e^{-\mu(p-Tk)}\, \mu(p-Tk)dpds + \partial \lambda 3 \int_{tk}^{t} \int_{tk}^{s-\tau 23(s)} |\eta 2(p) - \eta 2\,(p)| e^{-\mu(p-Tk)}\, m_i \mu(p-Tk)dpds +$$

$$+ \frac{\partial Gr2}{\partial r3} \int_{tk}^{t} -Tk)\, .\, r3(s - \tau 23(s)) - r3(s - \tau 23(s))\, e^{-\mu(s-\tau 23(s)}\, m_i \mu(s-\tau 23(s)-Tk)ds + \frac{\partial Gr2}{\partial \varphi 3} \int_{tk}^{t} \varphi 3 \, -Tk)\, (s - \tau 23(s)) - \varphi 3\,(s - \tau 23(s))\, e^{-\mu(s-\tau 23(s)}\, m_i \mu(s-\tau 23(s)-Tk)ds +$$

$$+ \frac{\partial Gr2}{\partial \eta 3} \int_{tk}^{t} \tau 23(s)) - \eta 3\,(s - \tau 23(s))\, e^{-\mu(s-\tau 23(s)-Tk)}\, \eta 3\,(s - \mu(s-\tau 23(s)-Tk)ds + \partial\, r2 \int_{tk}^{t} r2(s) - r2(s)\, e^{-\mu(s-Tk)}\, .. \mu(s-Tk)ds \le$$

$$\partial Gr2\, e^{\mu(t-Tk)}\, d(k,0)\, \overline{(r2 ,r2) \le} + \frac{d(k,0)(\varphi 2 ,\varphi 2)}{\mu 2} + \frac{d(k,0)(\eta 2 ,\eta 2)}{\mu 2} + \frac{d(k,0)(r3 ,r3)}{\mu} + \frac{d(k,0)(\varphi 3 ,\varphi 3)}{\mu} + \frac{d(k,0)(\eta 3 ,\eta 3)}{\mu} \int_{tk}^{t} m_i \mu(s-\tau 23-Tk) - 1\, ds +$$

$$+ \frac{\omega d(k-1,1)\,(r3 ,r3)}{\mu} + + \frac{\omega d(k-1,1)(\varphi 3 ,\varphi 3)\, \omega d(k-1,1)(\eta 3 ,\eta 3)}{\mu} \int_{tk}^{t} m_i \mu(s-\tau 23(s)-Tk)ds + \frac{\omega 2 d(k,2)\,(r2 ,r2)}{\mu} \int_{tk}^{t} m_i \mu(s-Tk)ds \le$$

$$\le \partial Gr2\, e^{\mu(t-Tk)}\, \frac{d(k,0)\,(r2 ,r2)}{\mu 2} + \frac{d(k,0)(\varphi 2 ,\varphi 2)}{\mu 2} + \frac{d(k,0)(\eta 2 ,\eta 2)}{\mu 2} + \frac{d(k,0)\,(r3 ,r3)}{\mu} + \frac{d(k,0)(\varphi 3 ,\varphi 3)}{\mu} + \frac{d(k,0)(\eta 3 ,\eta 3)}{\mu} \int_{tk}^{t} m_i \mu(s-Tk)\,(1 - e^{-\mu \tau 23})ds +$$

$$+ e^{\mu(t-Tk)}\, \omega\, d(k-1,1)\,(r3,r3) + d(k-1,1)\,\varphi 3, \varphi 3 + d(k-1,\, 1)\,(\eta 3,\, \eta 3) + e\, \mu 2 \quad \mu(t-Tk)\, \omega 2 d(k,2)\,(r2 ,r2)\, \mu 2 \le$$

$$\partial Gr2\, e + \mu 2\, \mu(t-Tk)\, 2 \le \frac{d(k,0)\,(r2 ,r2) + d(k,0)(\varphi 2 ,\varphi 2) + d(k,0)\,(\eta 2 ,\eta 2)}{\omega\, d(k-1,1)\,(r3 ,r3) + d(k-1,1)(\varphi 3 ,\varphi 3) + d(k-1,1)(\eta 3 ,\eta 3)\, \mu 2} + \frac{\omega 2 d(k,2)\,(r2 ,r2)}{\mu 2}\, .$$

Resulta que

$$\mu 2\, d(k,0) \quad {}^{(k)}_{r2}(.), B\,(.)_{r2} \le d(k,0)\, B^{(k)}$$

$$\overline{\quad} \quad \frac{2}{\overline{\quad}} \quad \overline{(r2, \varphi 2, \eta 2, r3, \varphi 3, \eta 3),(r2, \varphi 2 , \eta 2 , r3, \varphi 3 , \eta 3)} + \mu 2\, d(k-1,1)\, \overline{(r2, \varphi 2, \eta 2, r3, \varphi 3, \eta 3),(r2, \varphi 2 , \eta 2 , r3, \varphi 3 , \eta 3)} + \le \partial Gr2\, \omega 2 + \frac{\omega}{\overline{\quad}}$$

$$\overline{\mu 2}\, d(k,2)\, \overline{(r2, \varphi 2, \eta 2, r3, \varphi 3, \eta 3),(r2, \varphi 2 , \eta 2 , r3, \varphi 3 , \eta 3)}\, .$$

De manera similar se obtiene

Machine Translated by Google

$$(.), B(.) \le \le \frac{\partial G r3}{\mu 2}\, d(k,0)\,(r2,\varphi 2,\eta 2,r3,\varphi 3,\eta 3),(r2,\varphi 2,\eta 2,r3,\varphi 3,\eta 3) + \mu 2\, d(k-1,1)\,(r2,\varphi 2,\eta 2,r3,\varphi 3,\eta 3),(r2,\varphi 2,\eta 2,r3,\varphi 3,\eta 3) + \omega 2 + \mu 2\, d(k,2)\,(r2,\varphi 2,\eta 2,r3,\varphi 3,\eta 3),(r2,\varphi 2,\eta 2,r3,\varphi 3,\eta 3).$$

Más adelante tenemos

$$B\varphi 2\,(r2,\varphi 2,\eta 2,r3,\varphi 3,\eta 3) - B\varphi 2\,(r2,\varphi 2,\eta 2,r3,\varphi 3,\eta 3) \le$$

$$\le \frac{\partial G\varphi 2}{\partial \rho 2}\Big|_{tk} \partial G\varphi 2\,|\rho 2(s) - \rho 2(s)|ds + \frac{\partial}{\partial 2}\Big|_{tk} \partial G\varphi 2\,|2(s) - 2(s)|ds + \frac{\partial}{\partial \lambda 2}\Big|_{tk} \partial G\varphi 2\,\lambda 2(s) - \lambda 2(s)\,ds + \frac{\partial}{\partial \rho 3}\Big|_{tk} \rho 3(s-\tau 23(s)) - \rho 3(s-\tau 23(s))\,ds +$$

$$\frac{\partial G\varphi 2}{\partial 3}\Big|_{tk} \partial G\varphi 2\quad 3(s-\tau 23(s)) - 3(s-\tau 23(s))\,ds + \partial\lambda 3\Big|_{tk} \partial G\varphi 2\,\lambda 3(s-\tau 23(s)) - \lambda 3(s-\tau 23(s))\,ds + \frac{\partial}{\partial r3}\Big|_{tk} .\, r3(s-\tau 23(s)) - r3(s-\tau 23(s))\,ds +$$

$$\frac{\partial G\varphi 2}{\partial \varphi 3}\Big|_{tk} \varphi_3 \partial G\varphi 2\,(s-\tau 23(s)) - \varphi 3(s-\tau 23(s))\,ds + \frac{\partial}{\partial \eta 3}\Big|_{tk} \partial G\varphi 2\,\eta 3(s-\tau 23(s)) - \eta 3(s-\tau 23(s))\,ds + \frac{\partial}{\partial r2}\Big|_{tk} \varphi 2(s) - \varphi 2(s)\,ds \le$$

$$\frac{\partial G\varphi 2}{\partial \rho 2} \le \frac{|e2e3|c}{m2\cos(\pi/2-\delta)} \frac{\rho^3 c}{3(\Delta-(R3/\mu)e\,\mu T)^3} \frac{2\omega}{\rho 2^2} + \frac{6c}{c\,4\Delta-(R3/\mu)e\,\mu T)\,\rho 2^2} + \frac{c\omega}{c\,4\Delta-(R3/\mu)e\,\mu T)\,\rho 2} = \frac{\partial G\varphi 2}{\partial \rho 2}$$

etcétera.

Denotando por $Br2 = Br2\,(r2,\varphi 2,\eta 2,r3,\varphi 3,\eta 3),\ldots,B\eta 3 = B\eta 3\,(r2,\varphi 2,\eta 2,r3,\varphi 3,\eta 3)$, obtenemos

$$d(k,0)\,(Br2,B\varphi 2,B\eta 3,Br3,B\varphi 3,B\eta 3),(Br2,B\varphi 2,B\eta 3,Br3,B\varphi 3,B\eta 3) \le$$

$$\frac{2(\partial Gr3+\ldots+\partial G\eta 3)}{\mu 2}\,\omega(\partial Gr3+\ldots+\partial G\eta 3)\,d(k,0)\,(r2,\varphi 2,\eta 2,r3,\varphi 3,\eta 3),(r2,\varphi 2,\eta 2,r3,\varphi 3,\eta 3) + \le \mu 2\,\mu 2 + \frac{\omega 2\,(\partial Gr3+\ldots+\partial G\eta 3)}{\mu 2}\,d(k,2)\,(r2,\varphi 2,\eta 2,r3,\varphi 3,\eta 3),(r2,\varphi 2,\eta 2,r3,\varphi 3,\eta 3).\,\mu 2$$

Si $\frac{2(\partial Gr3+\ldots+\partial G\eta 3)}{\mu 2} < \frac{\omega(\partial Gr3+\ldots+\partial G\eta 3)}{\mu 2}\,1\,3\,; < \frac{\omega 2\,(\partial Gr3+\ldots+\partial G\eta 3)}{\mu 2}\,1\,3\,; <$ no se satisfacen, entonces aplicamos las desigualdades del Apéndice A y obtenemos las desigualdades deseadas para un $m \quad N$ suficientemente grande :

$$\frac{\omega m}{2(\partial Gr3+\ldots+\partial G\eta 3)\,\mu 2\,\mu m} < \frac{1}{3};\ \frac{\omega m}{\mu m}\,\frac{\omega(\partial Gr3+\ldots+\partial G\eta 3)}{\mu 2} < \frac{1}{3};\ \frac{\omega m}{\omega 2\,(\partial Gr3+\ldots+\partial G\eta 3)\,\mu 2\,\mu m} < \frac{1}{3}.$$

Es fácil ver que el espacio $Mr2 \times M\varphi 2 \times M\eta 2 \times Mr3 \times M\varphi 3 \times M\eta 3$ está $(j1,j2,j3)-$ acotado y las aplicaciones $js : A \to A\ (s = 1,2,3)$ son conmutativas. Entonces, el operador $(B2r, B2\varphi, B3r, B3\varphi)$, en vista del teorema del punto fijo de [8], tiene un único punto fijo, que, en vista del Lema Principal, es una solución T-periódica de (BS).

Así queda demostrado el teorema 1.

Resultados numéricos

Comprobamos las desigualdades a partir del teorema principal:

$$\frac{|e2e1|}{m2}\,\frac{1}{\infty\,\rho 2^2}\,\frac{c\omega 3(\rho 23\Delta)\,\omega 2\,2R2\,\rho 2 + + c}{\mu T}\,\frac{m\mu T}{\mu} \le R2\,;\ \varphi 02 + \frac{|e2e3|}{m2}\,\frac{c\omega 3(\varphi 23\Delta)\,\omega 2\,2\Phi 2\,\rho 2 + c\,3\Delta 2\cos(\pi/2-\delta)\,\rho}{\mu T}\,\frac{m\mu T}{\mu} \le \Phi 2\,;$$

$$\frac{|e2e3|}{m2}\,\frac{c\omega 3}{3c}\,\frac{e\,\mu T+3\Delta\,\omega 2\,2Y2\,\rho\,3 +}{\rho 2\Delta 2}\,\frac{mi}{\mu} \le Y2\,.$$

Se sabe (cf. [21]) que en $= 1{,}6\times 10-19C$; $mn = 9{,}11\times 10-19kg$ $(n = 2,3)$; $c = 3\times 108 m/$ seg; $\omega 2 = \omega 3 = 4 \times 1016 \approx \frac{2\pi}{\omega}, 57 \times 10-16$ Eligiendo $\mu = \frac{6{,}28}{4\times 1016}$, $\rho\,\Delta \approx c$, $\delta = \frac{\pi}{6}$, $\varphi 20 < \Phi 0$. $\approx \rho = 5{,}3\times 10-11m$;

$6 \times 1016 > \omega = 4 \times 1016$ y $\mu T = 6 \times 106 \times 1,\ 0{,}57 \times 10-16 = 9{,}42$ $> \omega T = 2\pi$, se obtiene $e^{9.42} \approx 1{,}2 \times 104$, $\frac{e\,\mu T}{\mu} = \frac{1{,}2\times 104}{6\times 1016} \approx 2 \times 10-13$.

Machine Translated by Google

Tenemos (R2 + R3) e $\frac{\mu T}{\mu} \le \frac{2c}{\mu} = \frac{2{,}2\times106}{6\times1016} \approx 0 \quad \Delta \approx |\rho02 - \rho03|;$

$2{,}81 \times 10{-}8 \quad \frac{1}{\rho03|\ 4\times1016(5.3\times10{-}11{+}3|\rho02{-}\rho03|\ 4\times1032 + +} \quad \frac{4\times1016(2\times5{,}3\times10{-}11{+}3|\rho02{-}\rho03|)\ 16\times1032 + +}{\rho02{-}\rho03|\ 2\times0{,}5\ 27\times1024} \quad R2\ 2 \times 10{-}13 \le R2;$

$5{,}62 \times 10{-}21 \quad \frac{5{,}3\times10{-}11\times4{,}8{+}3|\rho02{-}\rho03|\ 16\times1032 + 137\times9\times1016}{27\times1024\ 5.3\times10{-}11|\rho02{-}\rho03|} \quad \Phi2 \le \Phi2 - \varphi02;$

$2{,}81 \times 10{-}8\ 4\times1016 \quad \frac{}{2} \quad Y2\ 2 \times 10{-}13 \le Y2.$

Dado que $|\rho02 - \rho03| \approx 10{-}11$, entonces

$$5.62 \quad \frac{103}{+ 6 \times 107R2\ 10{-}21 \le R2. + + 4{,}11 \times 1012 \times 10{-}22\ 4{,}11 \times 1018 \times 10{-}11\ 15{,}9}$$

Para la segunda desigualdad, obtenemos

$$5{,}62 \times 10{-}21\ 6{,}5 \times 1019 + 3{,}7 \times 1019 + 0{,}6 \times 108\Phi2 \le \Phi2 - \varphi02.$$

Finalmente, $5{,}62\ 1{,}6 \times 1020 + 1{,}8 \times 1019 + 6{,}107Y2\ 10{-}21 \le Y2.$

7. Discusión sobre las hipótesis de Bohr y Sommerfeld

Según la primera hipótesis de Bohr (cf. [22,23]), ambos electrones del átomo de He se mueven en órbitas circulares en un plano en lados opuestos. Esto significa que tenemos el problema de Kepler con las condiciones $\lambda2 = 0$, $\lambda3 = 0$, $\eta3(t) = \eta2(t) + \pi$, lo que da como resultado

$\eta2(t) - \eta2(t - \tau23) \approx \varphi2\tau23 \approx \omega \quad \frac{\rho23}{2c} \approx 1016\ 5{,}3\times10{-}11 \approx 10{-}3;$

$\eta3(t - \tau32) - \eta3(t) \approx \varphi3\tau32 \approx \omega \quad \frac{\rho23}{2c} \approx 1016\ 5{,}3\times10{-}11 \approx 10{-}3$; y luego, 6×108

$$\cos\ \eta23 = \cos(\ \eta2(t) - \eta3(t - \tau23)) = \cos(\ \eta2(t) - \eta2(t - \tau23) - \pi) = -\cos(\ \eta2(t) - \eta2(t - \tau23)) \quad -1;$$

$$\text{sen}\ \eta23 = -\sin(\ 2 - \eta3(t - \tau23)) = -\sin(\ \eta2(t) - \eta2(t - \tau23) - \pi) = \sin(\ \eta2(t) - \eta2(t - \tau23)) \quad 0\ .$$

Desde $\vec{\xi}^{(23)}, {}^{(3)}\!\rightarrow tu = \rho2[\cos(\ 2 - \eta3) - 1] \ldots \rho3 = -2\rho2\ \rho3$, $\vec{\xi}^{(32)}, {}^{(2)}\!\rightarrow tu = -2\rho3$ la ecuación- $\rho2$,

Las condiciones de movimiento se vuelven

$$\dot{r}2 = \frac{e2e1}{m2}\frac{1}{c\rho2^2} + \frac{e2e3}{m2}\frac{4\rho2\ r3\ c^2}{6\tau23\ 3} + \frac{\dot{r}3}{c\ 4\tau23} - \frac{y_2^2}{m2}\frac{\ddot{r}2}{\omega^3}; \quad \dot{\varphi}2 = -\frac{e2e3m2}{c\ 4\tau23\rho2}\frac{\dot{\rho}3\ \varphi3}{} + \frac{\ddot{\varphi}2}{};$$

$$\dot{r}3 = \frac{e3e1}{m3}\frac{1}{2c\rho3} - \frac{e2e3}{m3}\frac{4\rho3\ \rho2 + c^2}{3\ c\ 4\tau23}\frac{\ddot{r}2}{6\tau32} - \frac{y_3^2}{m3c}\frac{\ddot{r}3}{3}; \quad \dot{\varphi}3 = \frac{e3e2}{m3}\frac{\varphi2\ \text{porque}}{32\ \rho2\ c\ 4\tau32\rho3} - \frac{\ddot{\varphi}3}{3}.\ \text{do}$$

Este sistema es un caso particular de (BS) y tiene una solución periódica única.

La segunda hipótesis de Bohr (cf. [22,23]) es la siguiente: ambos electrones se mueven en dos planos diferentes formando un ángulo de 60 grados. Esto significa que se puede elegir $\lambda2 = 0$, $\lambda3 = \pi/3$ $\eta3 = 0$, $\eta3(t) = \eta2(t) + \alpha > 0$, donde α se definirá a continuación.

De (BS), el sistema de ecuaciones de movimiento es

$$\dot{r}2 = \frac{e2e1}{m2}\frac{1}{c\rho2^2} - \frac{}{e2e3m2}\frac{\rho2{-}\rho3\ \text{porque}\ \ 23\ 3\ c}{\vec{\xi}^{(23)}, \vec{tú}^{(3)}} - \frac{\dot{r}3\ \cos\ \ 23{+}\rho3\ c}{c\ 3\tau23\ 3} - \frac{\varphi3\ \text{sen}\ \ 23}{e2e3m2}\frac{}{4\tau23} - \frac{y_2^2}{m2}\frac{\ddot{r}2}{\omega^3};$$

$$\dot{\varphi}2 = \frac{}{e2e3m2}\frac{\rho3\ \text{sen}\ \ 23}{\vec{\xi}^{(23)}, \vec{tú}^{(3)}} \quad \frac{\dot{r}3\ \text{sen}\ \ 23{+}\rho3\ \varphi3\ \cos\ \ 23 + c}{\tau23\ 3\rho2} - \frac{\ddot{\varphi}2}{3c}$$

$$\dot{r}3 = \frac{e3e1}{m3}\frac{1}{c\rho3^2} - \frac{e2e3}{m3}\frac{\rho3{-}\rho2\ \text{porque}\ \ 32\ 3\ c}{\vec{\xi}^{(32)}, \vec{tú}^{(2)}} - \frac{\dot{r}2\ \cos\ \ 32{+}\rho2\ c}{c\ 3\tau32\ 3} - \frac{\varphi2\ \text{sen}\ \ 32}{e3e2}\frac{}{4\tau23} - \frac{y_3^2}{m3}\frac{\ddot{r}3}{\omega^3}; \quad (11)$$

$$\dot{\varphi}3 = \frac{e3e2}{m3}\frac{\rho2\ \text{sen}\ \ 32}{\vec{\xi}^{(32)}, \vec{tú}^{(2)}} \quad \frac{\dot{r}2\ \text{sen}\ \ 32{+}\rho2\ \varphi2\ \cos\ \ 32 + c}{\tau32\ 3\rho3} - \frac{\ddot{\varphi}3}{}$$

$$0 = \frac{}{(\rho3\ \text{porque}\ \ 32{-}2\rho2)\ .\ r2{+}\text{sen}\ \ 32\rho2\rho3\ \varphi2\ \text{sen}\ \ 32\ \varphi2\ r3\ \cos\ \ 32{+}\rho2\ \rho2\ \cos\ \ 32 + 2\ 2c\ 6\ 4\ \tau32\ c}.$$

Machine Translated by Google

La ecuación (11) puede considerarse como una condición adicional para la existencia de una solución periódica. En efecto,

$$23 = 2(t) - 3(t - \tau 23), \quad 3(t) = 2(t) + \alpha, \quad 23 = 2(t) - 2(t - \tau 23) - \alpha \approx -\alpha;$$
$$32 = 2(t - \tau 32) - 3(t), \quad 3(t) = 2(t) + \alpha, \quad 32 = 2(t - \tau 32) - 2(t) - \alpha \approx -\alpha;$$
$$\cos 32 \approx \cos \alpha, \ \operatorname{sen} 32 \approx -\operatorname{sen} \alpha.$$

Entonces, la ecuación (11) se convierte en

$$\frac{2\,2c\,6\,\tau 32}{} \quad (\rho 3 \text{ porque } \alpha - 2\rho 2) \cdot r2 + \rho 2 \rho 3\, \varphi 2 \operatorname{sen} \alpha\, \rho 2 \cos \alpha\, 0 = \quad \frac{\cdot \quad \cdot}{r3 \cos \alpha + \rho 2\, \varphi 2 \operatorname{sen} \alpha +} .$$

Para ρ2 = ρ3 = ρ = const., se obtiene

$$\frac{\text{porque } \alpha\, \rho \quad 2}{\tau 32} \quad \frac{\cdot}{\varphi 2 \operatorname{sen} \alpha} = 0 \quad \operatorname{sen} \alpha = 0 \quad \alpha = 0 \quad \alpha = \pi,$$

lo que confirma la hipótesis de Bohr para α = π. Entonces, las ecuaciones de movimiento se convierten en

$$\cdot \quad \cdot \quad \cdot \quad \

Machine Translated by Google

(2) donde $\dot{\rho} = d\rho(2)/d\theta$. Para las aceleraciones tenemos

$$\dot{u}_1^{(n)} \approx \ddot{\rho}n\cos\quad n\cos\lambda n - \rho n\quad n\,\text{sen}\quad n\cos\lambda n - \rho n\,\lambda n\cos\quad n\,\text{sen}\,\lambda n\quad \ddot{}$$
$$\dot{u}_2^{(n)} \approx \ddot{\rho}n\,\text{sen}\quad n\cos\lambda n + \rho n\ddot{}\,n\cos\quad n\cos\lambda n - \rho n\,\lambda n\,\text{sen}\quad n\,\text{sen}\,\lambda n$$
$$\dot{u}_3^{(n)} \approx \ddot{\rho}n\,\text{sen}\,\lambda n + \rho n\,\lambda n\cos\lambda n.$$

De la misma manera obtenemos las expresiones para las segundas derivadas (n = 2, 3):

$$\ddot{u}_1^{(norte)} \approx \dddot{\rho}___\text{porque}\quad n\,\text{porque}\,\lambda n - \dddot{}___\rho n\,\text{sen}\quad n\cos\lambda n - \lambda\,n\dddot{\rho}n\cos\quad n\,\text{sen}\,\lambda n$$
$$\ddot{u}_2^{(n)} \approx \dddot{\rho}___\,\text{sen}\quad n\cos\lambda n + \dddot{}___\rho n\cos\quad n\cos\lambda n - \lambda\,n\dddot{\rho}n\,\text{sen}\quad n\,\text{sen}\,\lambda n$$
$$\ddot{u}_3^{(n)} \approx \dddot{\rho}___\,\text{sen}\,\lambda n + \lambda\,n\dddot{\rho}n\cos\lambda n$$

La condición (C) implica

$$\vec{\dot{u}}^{(2)}, \vec{\dot{u}}^{(2)} = \dot{\rho}2 + \rho 2 \quad 2\,\dot{\rho}2\quad 2\,\dot{\rho}2\cos\lambda 2 + \rho 2 \quad 2\,\dot{\lambda}2 \quad 2\,2\,2 \le c < c \quad \left|\frac{d}{d\theta}\rho 2\right| \approx 0, \quad \left|\frac{d}{d\theta}\dot{\rho}2\right| \approx 0, \quad \frac{d}{d\theta}\dot{\rho}2\dot{\lambda}2 \approx 0;$$

$$\vec{\dot{u}}^{(2)}, \vec{\dot{u}}^{(2)}/d\theta^2 \approx 0;$$

$$(2)\,\dot{\rightarrow} tu, (2)\,\dot{\rightarrow} tu = \ddot{\rho}2 + \rho 2 \quad 2\,\ddot{\rho}2\quad 2\,\ddot{\rho}2\cos\lambda 2 + \rho 2 \quad 2\,\ddot{\lambda}2\ 2\,; \quad \vec{\ddot{u}}^{(2)}, \vec{\ddot{u}}^{(2)} = \dddot{\rho}2 \times 2 + \rho 2 \quad 2\,\dddot{\rho}2\quad 2\,\dddot{\rho}2\cos\lambda 2 + \rho 2 \quad 2\,2\,\dddot{\lambda}2\ .$$

Más adelante tenemos

$$\vec{\xi}^{(12)} = (\xi_1^{(12)}, \xi_1^{(12)}, \xi_1^{(12)}, \xi_1^{(1)}, \xi_2^{(1)}, \xi_1^{(1)}, \xi_2^{(1)}, \xi_1^{(2)}) = x(t) - x(t - \tau 12(t)), x(t) - x(t - \tau 12(t)), x(t) - x(t - \tau 12(t)))$$

$$= -x_1^{(2)}(\theta), -x_2^{(2)}(\theta), -x_3^{(2)}(\theta) = (-\rho 2\cos\quad 2\cos\lambda 2, -\rho 2\,\text{sen}\quad 2\cos\lambda 2, -\rho 2\,\text{sen}\,\lambda 2);$$

$$\vec{\xi}^{(13)} = (\xi_1^{(13)}, \xi_1^{(13)}, \xi_1^{(13)}, \xi_1^{(1)}, \xi_3^{(1)}, \xi_1^{(1)}, \xi_3^{(1)}, \xi_1^{(3)}) = x(t) - x(t - \tau 13(t)), x(t) - x(t - \tau 13(t)), x(t) - x(t - \tau 13(t)))$$

$$= -x_1^{(3)}(_), -x_2^{(3)}(_), -x_3^{(3)}(_) = (-\rho 3\cos\quad 3\cos\lambda 3, -\rho 3\,\text{sen}\quad 3\cos\lambda 3, -\rho 3\,\text{sen}\,\lambda 3);$$

$$\tau 12(t) = \frac{1}{d\theta}\sqrt{\sum_{a=1}^{3}\left[\alpha^{(1)}(t) - x^{(2)}(t - \tau 12(t))\right]^2} = \frac{1}{d\theta}\sqrt{\sum_{a=1}^{3}\left[-x^{(2)}\alpha(t - \tau 12(t))\right]^2} = \frac{\rho 2(t-\tau 12(t))}{d\theta} = \frac{\rho 2(\theta)}{d\theta};$$

$$\tau 21 = \frac{1}{d\theta}\sqrt{\left(\vec{\xi}^{(21)}, \vec{\xi}^{(21)}\right)} = \frac{1}{d\theta}\sqrt{\left(\vec{x}^{(2)}, \vec{x}^{(2)}\right)} = \overline{\rho 2};$$

$$\tau 13(t) = \frac{1}{d\theta}\sqrt{\sum_{a=1}^{3}\left[\alpha^{(1)}(t) - x^{(3)}(t - \tau 13(t))\right]^2} = \frac{1}{d\theta}\sqrt{\sum_{a=1}^{3}\left[-x^{(3)}\alpha(t - \tau 13(t))\right]^2} = \frac{\rho 3(t-\tau 13(t))}{d\theta} = \frac{\rho 3(_)}{d\theta};$$

$$(21) = \xi_1^{(21)}, \xi_2^{(21)}, \xi_3^{(21)}, \xi_4^{(21)} = \vec{\xi}^{(21)}, \xi_4^{(21)} = x_1^{(2)}(t) - x_2^{(2)}(t - \tau 21(t)), x_2^{(2)}(t) - x_2^{(2)}(t - \tau 21(t)), x_3^{(2)}(t) - x_3^{(2)}(t - \tau 21(t)), ic\tau 21 =$$

$$= x_1^{(2)}(t), ic\tau 21, x_3^{(2)}(t), x_3^{(2)}(t), ic\tau 21;$$

$$(23) = \xi_1^{(23)}, \xi_2^{(23)}, \xi_3^{(23)}, \xi_4^{(23)} = \vec{\xi}^{(23)}, \xi_4^{(23)} = x_1^{(2)}(t) - x_2^{(3)}(t - \tau 23), x_2^{(2)}(t) - x_2^{(3)}(t - \tau 23), x_3^{(2)}(t) - x_3^{(3)}(t - \tau 23), ic\tau 23 =$$

$$= (\rho 2\cos\quad 2\cos\lambda 2 - \rho 3\cos\quad 3\cos\lambda 3, \rho 2\,\text{sen}\quad 2\cos\lambda 2 - \rho 3\,\text{sen}\quad 3\cos\lambda 3, \rho 2\,\text{sen}\,\lambda 2 - \rho 3\,\text{sen}\,\lambda 3);\ c^2\tau 23^2 = (\rho 2\,\text{porque}$$
$$2\,\text{porque}\,\lambda 2 - \rho 3\,\text{porque}\quad 3\,\text{porque}\,\lambda 3)^2 + (\rho 2\,\text{sen}\quad 2\cos\lambda 2 - \rho 3\,\text{sen}\quad 3\cos\lambda 3)^2 + (\rho 2\,\text{sen}\,\lambda 2 - \rho 3\,\text{sen}\,\lambda 3)^2 =$$
$$2 = \rho 2^2 + \rho 3^2 - 2\rho 2\rho 3[\cos\lambda 2\cos\lambda 3\cos(\quad 2 - \quad 3) + \text{sen}\,\lambda 2\,\text{sen}\,\lambda 3];$$

$$(31) = \xi_1^{(31)}, \xi_2^{(31)}, \xi_3^{(31)}, \xi_4^{(31)} = \vec{\xi}^{(31)}, \xi_4^{(31)} = x_1^{(3)}(t) - x_2^{(1)}(t - \tau 31(t)), x_2^{(3)}(t) - x_2^{(1)}(t - \tau 31(t)), x_3^{(3)}(t) - x_3^{(1)}(t - \tau 31(t)), ic\tau 31 =$$

$$= x_1^{(3)}(t), ic\tau 21, x_3^{(3)}(t), x_3^{(3)}(t), ic\tau 21;$$

$$(32) = \xi_1^{(32)}, \xi_2^{(32)}, \xi_3^{(32)}, \xi_4^{(32)} = \vec{\xi}^{(32)}, \xi_4^{(32)} = x_1^{(3)}(t) - x_2^{(2)}(t - \tau 32), x_2^{(3)}(t) - x_2^{(2)}(t - \tau 32), x_3^{(3)}(t) - x_3^{(2)}(t - \tau 32), ic\tau 32 =$$

$$= (\rho 3\cos\quad 3\cos\lambda 3 - \rho 2\cos\quad 2\cos\lambda 2, \rho 3\,\text{sen}\quad 3\cos\lambda 3 - \rho 2\,\text{sen}\quad 2\cos\lambda 2, \rho 3\,\text{sen}\,\lambda 3 - \rho 2\,\text{sen}\,\lambda 2);\ c^2\tau 32$$
$$= (\rho 3\,\text{porque}\quad 3\,\text{porque}\,\lambda 3 - \rho 2\,\text{porque}\quad 2\,\text{porque}\,\lambda 2)^2 + (\rho 3\,\text{sen}\quad 3\cos\lambda 3 - \rho 2\,\text{sen}\quad 2\cos\lambda 2)^2 + (\rho 3\,\text{sen}\,\lambda 3 - \rho 2\,\text{sen}\,\lambda 2)^2 =$$
$$2 = \rho 2^2 + \rho 3^2 - 2\rho 2\rho 3[\cos\lambda 2\cos\lambda 3\cos(\quad 2 - \quad 3) + \text{sen}\,\lambda 2\,\text{sen}\,\lambda 3].$$

Machine Translated by Google

Pero

$$\cos\lambda_2\cos\lambda_3\cos(\varphi_2-\varphi_3)+\sin\lambda_2\sin\lambda_3=$$

$$=\frac{\cos(\lambda_2-\lambda_3)+\cos(\lambda_2+\lambda_3)\cos(\lambda_2-\lambda_3)-\cos(\lambda_2+\lambda_3)\cos(\varphi_2-\varphi_3)+\sin^2\varphi_2-\sin^2\varphi_3}{2}=$$

$$=\frac{\cos(\lambda_2-\lambda_3)[1+\cos(\varphi_2-\varphi_3)]-[\cos(\lambda_2+\lambda_3)-\cos(\lambda_2+\lambda_3)\cos(\varphi_2-\varphi_3)]}{2}=$$

$$=\frac{\cos(\lambda_2-\lambda_3)2\cos^2\frac{\varphi_2-\varphi_3}{2}-\cos(\lambda_2+\lambda_3)2\sin\frac{\varphi_2-\varphi_3}{2}\sin\frac{\varphi_2-\varphi_3}{2}}{2}\le\cos^2\frac{\varphi_2-\varphi_3}{2}+\sin^2\frac{\varphi_2-\varphi_3}{2}=1.$$

Por lo tanto,

$$c^2\tau_{23}^2=\rho_2^2+\rho_3^2-2\rho_2\rho_3[\cos\lambda_2\cos\lambda_3\cos(\varphi_2-\varphi_3)+\sin\lambda_2\sin\lambda_3]\ge(\rho_2-\rho_3)^2$$

$$\tau_{23}\ge\frac{|\rho_2(t)-\rho_3(t-\tau_{23})|}{c}$$

y $\tau_{32}\ge\frac{|\rho_3(t)-\rho_2(t-\tau_{32})|}{c}$.

Para las aceleraciones obtenemos (n = 2, 3)

$$\ddot{u}_1(n)=\ddot{\lambda}_n\rho_n\frac{\cos\varphi_n\cos\lambda_n-\dot{\rho}_n\dot{\varphi}_n\sin\varphi_n\cos\lambda_n-\rho_n\dot{\lambda}_n\cos\varphi_n\sin\lambda_n-c\cos\varphi_n\cos\lambda_n=c\,\dot{\rho}_n\dot{\varphi}_n\sin\varphi_n\cos\lambda_n-\rho_n\dot{\lambda}_n\sin\varphi_n\sin\lambda_n-c\sin\varphi_n\cos\varphi_n\dot{\lambda}_n}{2c\approx-\cos\varphi_n\cos\lambda_n}\equiv Q1(n)\,\rho_n,$$

$$\ddot{u}_2(n)=c\frac{\dot{\varphi}_n\sin\varphi_n\sin\lambda_n-\rho_n\dot{\varphi}_n\cos\varphi_n\sin\lambda_n-\rho_n\dot{\lambda}_n\sin\varphi_n\cos\lambda_n}{2c\approx-\sin\varphi_n\cos\lambda_n}\equiv Q2(n)\,\rho_n,$$

$$\ddot{u}_3(n)=\frac{\dot{\rho}_n\sin\lambda_n+\rho_n\dot{\lambda}_n\cos\lambda_n-c\sin\lambda_n=c}{\rho_n}\approx-c\frac{\dot{\lambda}^2\text{pecado}}{\lambda_n\rho_n}\equiv Q3(n),\ (n=2,3).$$

Pasando a las coordenadas esféricas, obtenemos

$$\ddot{\rho}_n\cos\varphi_n\cos\lambda_n-\rho_n\ddot{\varphi}_n\sin\varphi_n\cos\lambda_n-\rho_n\ddot{\lambda}_n\cos\varphi_n\sin\lambda_n=$$

$$=2\dot{\rho}_n\sin\varphi_n\cos\lambda_n+2\dot{\lambda}_n\cos\varphi_n\sin\varphi_n\sin\lambda_n-2\rho_n\dot{\varphi}_n\sin\varphi_n\sin\lambda_n+\rho_n\dot{\varphi}_n\rho_n\varphi_n\cos\varphi_n\cos\lambda_n+\rho_n\ddot{\lambda}_n\cos\varphi_n\cos\lambda_n-\rho_n\dot{\varphi}_n^2\cos\varphi_n\sin\varphi_n\cos\lambda_n-\rho_n\dot{\lambda}_n^2\cos^2\varphi_n\cos\lambda_n;$$

$$\ddot{\rho}_n\sin\varphi_n\cos\lambda_n+\rho_n\ddot{\varphi}_n\cos\varphi_n\cos\lambda_n-\rho_n\dot{\lambda}_n\sin\varphi_n\sin\lambda_n=$$

$$=-2\dot{\rho}_n\cos\varphi_n\cos\lambda_n+2\dot{\lambda}_n\sin\varphi_n\sin\lambda_n+2\rho_n\dot{\varphi}_n\sin\varphi_n\cos\lambda_n+\rho_n\dot{\varphi}_n\rho_n\varphi_n\sin\varphi_n\sin\lambda_n-\rho_n\dot{\lambda}_n^2\cos^2\varphi_n\sin\lambda_n;$$

$$\ddot{\rho}_n\sin\lambda_n+\rho_n\ddot{\lambda}_n\cos\lambda_n=-2\dot{\lambda}_n\cos\lambda_n+\rho_n\dot{\lambda}_n^2\rho_n\text{pecado}\lambda_n-\frac{\ddot{\lambda}}{\lambda_n}\text{pecado}\lambda_n\rho_n.$$

(A1)

Introducimos denotaciones:

$$P1(n)=2\dot{\rho}_n\sin\varphi_n\cos\lambda_n+2\dot{\lambda}_n\cos\varphi_n\sin\varphi_n\sin\lambda_n-2\rho_n\dot{\lambda}_n\sin\varphi_n\sin\varphi_n\sin\lambda_n+\rho_n\dot{\varphi}_n\rho_n\varphi_n\cos\varphi_n\cos\lambda_n+\rho_n\dot{\lambda}_n^2\cos^2\varphi_n\cos\lambda_n+\rho_n\dot{\varphi}_n^2\cos\varphi_n\sin\varphi_n\cos\lambda_n+\rho_n\dot{\lambda}_n^2\cos^2\varphi_n\cos\lambda_n;$$

$$P2(n)=\ddot{\rho}_n\cos\varphi_n\cos\lambda_n+2\dot{\lambda}_n\sin\varphi_n\sin\lambda_n+2\rho_n\dot{\lambda}_n\cos\varphi_n\sin\varphi_n\sin\lambda_n+\rho_n\ddot{\varphi}_n\rho_n\varphi_n+\rho_n\dot{\varphi}_n^2\sin\varphi_n\cos\varphi_n\cos\lambda_n+\rho_n\dot{\lambda}_n^2\cos^2\varphi_n\sin\varphi_n\cos\lambda_n=-2\dot{\rho}_n\dot{\varphi}_n\rho_n\varphi_n+\rho_n\dot{\lambda}_n^2\sin\varphi_n\cos\varphi_n\cos\lambda_n;$$

$$P3(n)=-2\dot{\lambda}_n\text{pecado}\lambda_n+\rho_n\ddot{\lambda}_n\rho_n\text{pecado}^2\lambda_n.$$

Entonces (A1) toma la siguiente forma:

$$\ddot{\rho}_n\cos\varphi_n\cos\lambda_n-\rho_n\ddot{\varphi}_n\sin\varphi_n\cos\lambda_n-\rho_n\ddot{\lambda}_n\cos\varphi_n\sin\lambda_n=P1(n)+Q1(n)\,\rho_n\cos\varphi_n$$

$$\ddot{\rho}_n\sin\varphi_n\cos\lambda_n+\rho_n\ddot{\varphi}_n\cos\varphi_n\cos\lambda_n-\rho_n\ddot{\lambda}_n\sin\varphi_n\sin\lambda_n=P2(n)+Q2(n)\,\rho_n\cos\varphi_n\cos\lambda_n;$$

$$\ddot{\rho}_n\sin\varphi_n\lambda_n+\rho_n\ddot{\varphi}_n\sin\varphi_n\lambda_n+\rho_n\ddot{\lambda}_n\cos\varphi_n\sin\varphi_n\lambda_n=P3(n)+Q3(n)$$

Para resolver el último sistema con respecto a $\ddot{\rho}_n$, $\ddot{\varphi}_n$, $\ddot{\lambda}_n$, notamos que

$$\Delta=\begin{vmatrix}\cos\varphi_n\cos\lambda_n & -\rho_n\sin\varphi_n\cos\lambda_n & -\rho_n\cos\varphi_n\sin\lambda_n\\ \sin\varphi_n\cos\lambda_n & \rho_n\cos\varphi_n\cos\lambda_n & -\rho_n\sin\varphi_n\sin\lambda_n\\ \text{pecado}\lambda_n & 0 & \rho_n\cos\varphi_n\cos\lambda_n\end{vmatrix}=\rho_n^2\cos\varphi_n\ \text{porque}\ \lambda_n>0$$

y luego obtenemos las ecuaciones iniciales (n = 2, 3):

$$\ddot{\rho}_n=\rho_n\dot{\varphi}_n^2+\dot{\lambda}_n^2\rho_n\cos^2\varphi_n-\frac{\ddot{\rho}^2}{\rho_n},$$

$$\ddot{\varphi}_n=-2\frac{\dot{\rho}_n\dot{\varphi}_n}{\rho_n}-\dot{\lambda}_n^2\sin\varphi_n\cos\varphi_n,$$

$$\ddot{\lambda}_n=-\frac{2\dot{\rho}_n\dot{\lambda}_n+\rho_n\ddot{\lambda}_n\rho_n}{\rho_n\cos\varphi_n}\sin\varphi_n\cos\varphi_n.$$

Machine Translated by Google

Apéndice B. Forma final de las ecuaciones básicas (n) .→ u

Primero calculamos $\xi^{(n)} .\to u$, $= \ddot{\rho}_n + \rho_n \dot{\lambda}_n^2 - 2\cos 2\lambda_n + \rho_n \ddot{\lambda}_n^2$;

$\to (23)$

$\xi^{(3)} .\to tu = [\rho 2 \cos(\ddot{\alpha}_2 - \alpha_3) \cos\lambda 2 \cos\lambda 3 + \rho 2 \operatorname{sen}\lambda 2 \operatorname{sen}\lambda 3 - \rho 3] .. \rho 3 +$

$+\rho 2 \rho 3 \operatorname{sen}(\ddot{\alpha}_2 - \alpha_3) \cos\lambda 2 \cos\lambda 3 \ddot{\alpha}_3 + \rho 2 \rho 3[\operatorname{sen}\lambda 2 \cos\lambda 3 - \cos(\ddot{\alpha}_2 - \alpha_3) \operatorname{sen}\lambda 3 \cos\lambda 2] \ddot{\lambda} 3;$

$\to (32)$

$\xi^{(2)} .\to tu = [\rho 3 \cos(\ddot{\alpha}_3 - \alpha_2) \cos\lambda 3 \cos\lambda 2 + \rho 3 \operatorname{sen}\lambda 3 \operatorname{sen}\lambda 2 - \rho 2] .. \rho 2 +$

$+ \operatorname{sin}(\ddot{\alpha}_3 - \alpha_2) \cos\lambda 3 \cos\lambda 2 \rho 2 \rho 3 \ddot{\alpha}_2 + [\sin\lambda 3 \cos\lambda 2) - \cos(\ddot{\alpha}_3 - \alpha_2) \operatorname{sin}\lambda 2 \cos\lambda 3] \rho 2 \rho 3 \ddot{\lambda} 2.$

Introducir denotaciones

$$\alpha_{PAG} = \alpha_{GRAMO}^{(2)} + G^{(21)} \alpha^{(23)} + G^{(2)} \alpha^{rad}; \quad \alpha_{PAG} = \alpha_{GRAMO}^{(31)} + G^{(32)} \alpha^{(3)} + G^{rad} \alpha \quad (a = 1, 2, 3).$$

Entonces, el sistema de ecuaciones de movimiento se convierte en

$$\begin{aligned}
\dot{tu}_1^{(norte)} &= \frac{2\text{- c}}{\text{tau}^2}\, tu_1^{(norte)} P_1^{(norte)} - \frac{tu_1^{(norte)} tu_2^{(norte)}}{\text{tau}^2} P_2^{(norte)} - \frac{tu_1^{(norte)} tu_3^{(norte)}}{2c} P \equiv U\,;\, 3\,1^{(n)}{}_{(n)}{}_{(n)} \\
\dot{tu}_2^{(norte)} &= -\frac{tu_1^{(norte)} tu_2^{(norte)}}{2c} P + 1^{2norte}{}_{c\,2-(u\,(norte)} \frac{2(n))}{\text{tau}^2} P_2^{(norte)} - \frac{tu_2^{(norte)} tu_3^{(norte)}}{2c} P \equiv U\,2\,;\,3^{(n)}{}_{2}{}^{,(n=2,} \\
\dot{tu}_3^{(norte)} &= -\frac{tu_1^{(norte)} tu_3^{(norte)}}{2c} P_1^{(n)}{}_{(n)} - \frac{tu_2^{(norte)} tu_3^{(norte)}}{2c} P + 2^{(n)}{}_{(n)}{}_{c\,2-(u\,(n)\,)} \frac{3}{\text{tau}^2} P_3^{(norte)} \equiv U3
\end{aligned}$$

Pero

$$\begin{aligned}
\dot{u}1^{(n)} &\approx \ddot{\rho}_n \cos\lambda n - \rho_n \dot{\lambda}_n^2 \operatorname{sen}\lambda n \cos\lambda n - \rho_n \ddot{\lambda} n \cos\lambda n \operatorname{sen}\lambda n \\
\dot{u}2^{(n)} &\approx \ddot{\rho}_n \operatorname{sen}\lambda n \cos\lambda n + \rho_n \dot{\lambda}_n^2 \cos\lambda n \cos\lambda n - \rho_n \ddot{\lambda} n \operatorname{sen}\lambda n \operatorname{sen}\lambda n \\
\dot{u}3^{(n)} &\approx \ddot{\rho}_n \operatorname{sen}\lambda n + \rho_n \ddot{\lambda} n \cos\lambda n
\end{aligned}$$

y desde $\dot{\vec{tu}}^{(norte)}, \vec{tu}^{(norte)} 2/c \approx 0$, inferimos

$$\begin{aligned}
\dot{tu}_1^{(norte)} &= \frac{2\text{- c}}{\text{tau}^2}\, tu_1^{(norte)} P_1^{(norte)} - \frac{tu_1^{(norte)} tu_2^{(norte)}}{2c} P_2^{(n)}{}_{(n)} - \frac{tu_1^{(norte)} tu_3^{(norte)}}{2c} P \approx P\,;\,3\,1^{(n)}{}_{(n)}{}_{(n)} \\
\dot{tu}_2^{(norte)} &= -\frac{tu_1^{(norte)} u_2}{2} P + 1^{(n)}{}_{(n)}{}_{c\,2-(u\,(n)\,)} \frac{2}{2} P_2^{(norte)} - \frac{tu_2^{(norte)} u_3}{2} P \approx P\,;\,3^{(n)}{}_{2}{}^{,(n=} \\
\dot{tu}_3^{(norte)} &= -\frac{c\,(n)\,(n)}{2c}\, uu\,1\,3 - \frac{tu_2^{(norte)} tu_3^{(n)}}{2c} P + 2^{c}{}_{c\,2-tu} \frac{c\,(n)}{\text{tau}^2}\, P_3^{(norte)} \quad {}^{2,3)} \approx P\,3
\end{aligned}$$

y por lo tanto,

$$\begin{aligned}
&\ddot{\rho}_n \cos\lambda n - \rho_n \dot{\lambda}_n^2 \operatorname{sen}\lambda n \cos\lambda n - \rho_n \ddot{\lambda} n \operatorname{sen}\lambda n = P_1^{(n)} \\
&\ddot{\rho}_n \operatorname{sen}\lambda n + \rho_n \dot{\lambda}_n^2 \cos\lambda n \operatorname{sen}\lambda n + \rho_n \ddot{\lambda} n \cos\lambda n \operatorname{sen}\lambda n = P_2^{(n)} \\
&\ddot{\rho}_n \operatorname{sen}\lambda n + \rho_n \ddot{\lambda} n \cos\lambda n = P_3^{(n)}.
\end{aligned}$$

Resolviendo el sistema anterior con respecto a $\ddot{\lambda}n$, obtenemos $\ddot{\rho}_n$, $\dot{\lambda}_n^2$,

$$\ddot{\rho}_n = P_1^{(n)} \text{porque}\ \lambda n + P_2^{(n)}{}_{(n)} \operatorname{sen}\lambda n \cos\lambda n + P \operatorname{sen}\lambda n;$$

$$\dot{\lambda}_n^2 = \frac{-P_{\text{sen}}^{(n)}{}_{(n)} n + P \cos n\,1\,;\,2}{\rho_n \text{porque}\ \lambda n}, \tag{A2}$$

$$\ddot{\lambda}n = \frac{-P_{\cos}^{(n)}{}_{(n)}{}_{(n)} n \operatorname{sen}\lambda n - P \operatorname{sen} n \operatorname{sen}\lambda n + P \cos\lambda n\,1\,2\,3\,\rho n}{} \quad (\text{norte} = 2, 3)$$

Machine Translated by Google

Recordemos que 23 = 2(t) − 3(t − τ23) y 32 = 2(t − τ32) − 3(t), y, en vista de

$$\overset{(23)}{{}_{\text{GRAMO}}\alpha} \approx \frac{e2e3}{3m2c}\left(\frac{\xi_\alpha^{(23)}}{\tau 23^3}\frac{\vec{\xi}^{(23)},\vec{tu}^{(3)}}{\|\cdot\|^2}-\frac{\dot{tu}_\alpha^{(3)}}{c\tau 23}\right) \qquad \overset{(32)}{\alpha} \approx \frac{e2e3}{3m3c}\left(\frac{\xi_\alpha^{(32)}}{\tau 32^3}\frac{\vec{\xi}^{(32)},\vec{tu}^{(3)}}{\|\cdot\|^2}-\frac{\dot{tu}_\alpha^{(2)}}{c\tau 32}\right)$$

$$\overset{(2)\,(21)\,(23)}{{}_{PAG}\alpha_{=\,GRAMO}\alpha} + G_\alpha + G_\alpha^{(2)rad} \approx \frac{e2e1}{4m2c}\frac{\xi_\alpha^{(21)}}{\tau 21^3} + \frac{e2e3}{3m2c}\left(\frac{\xi_\alpha^{(23)}}{\tau 23^3}\frac{\vec{\xi}^{(23)},\vec{tu}^{(3)}}{\|\cdot\|^2}-\frac{\dot{tu}_\alpha^{(3)}}{c\tau 23}\right) - \frac{{}^2y}{3\,m2c^2}\ddot{tu}_\alpha^{(2)}$$

$$\overset{(3)\,(31)\,(32)}{{}_{PAG}\alpha_{=\,GRAMO}\alpha} + G_\alpha + G_\alpha^{(3)rad} \approx \frac{e3e1}{m3}\frac{\xi_\alpha^{(31)}}{+\,c\,4\tau 31\,3} \frac{e2e3}{3m3c}\left(\frac{\xi_\alpha^{(32)}}{\tau 32^3}\frac{\vec{\xi}^{(32)},\vec{tu}^{(2)}}{\|\cdot\|^2}-\frac{\dot{tu}_\alpha^{(2)}}{c\tau 32}\right) - \frac{{}^2y}{3\,m3c^3}\ddot{tu}_\alpha^{(3)}$$

(a = 1, 2, 3)

transformamos los lados derechos de las ecuaciones de (A2).

En primer lugar,

$$\ddot{\rho}2 = P\cos\left(\frac{2}{1}\right)\cos\lambda 2 + P\,\text{sen}\left(\frac{2}{2}\right)\cos\lambda 2 + P\,\text{sen}\left(\frac{3}{2}\right)\lambda 2 =$$

$$= \cos\left(\frac{2}{1}\right)\cos\lambda 2\left[\frac{e2e1}{4\,m2c}\frac{\xi_1^{(21)}}{\tau 21^3} + \frac{e2e3}{3m2c}\left(\frac{\xi_1^{(23)}}{\tau 23^3}\frac{\vec{\xi}^{(23)},\dot{tu}^{(3)}}{\|\cdot\|^2}-\frac{\dot{tu}_1^{(3)}}{c\tau 23}\right) - \frac{{}^2y_2}{m2c}\ddot{u}_1^{(2)}\right]_3 +$$

$$+ \text{sen}\left(\frac{2}{1}\right)\cos\lambda 2\left[\frac{e2e1}{4\,m2c}\frac{\xi_2^{(21)}}{\tau 21^3} + \frac{e2e3}{3m2c}\left(\frac{\xi_2^{(23)}}{\tau 23^3}\frac{\vec{\xi}^{(23)},\dot{tu}^{(3)}}{\|\cdot\|^2}-\frac{\dot{tu}_2^{(3)}}{c\tau 23}\right) - \frac{{}^2y_2}{3\,m2c}\ddot{u}_2^{(2)}\right]_3 +$$

$$+ \text{pecado}\left(\frac{2}{1}\right)\lambda 2\left[\frac{e2e1}{4m2c}\frac{\xi^{(21)}}{3\,\tau 21^3} + \frac{e2e3}{3m2c}\left(\frac{\xi_3^{(23)}}{\tau 23^3}\frac{\vec{\xi}^{(23)},\dot{tu}^{(3)}}{\|\cdot\|^2}-\frac{\dot{tu}_3^{(3)}}{c\tau 23}\right) - \frac{{}^2y_{2\,\tau}u_3}{m2c}\ddot{u}_3^{(2)}\right] =$$

$$= \frac{e2e1}{4\,m2c}\frac{\lambda 2^{(21)\,(21)\,(21)}\,\xi\cos\left(\frac{2}{2}\right)\cos\lambda 2+\xi\,\text{sen}\left(\frac{2}{3}\right)\cos\lambda 2+\xi\,\text{sen}\left(\frac{3}{2}\right)\lambda 2}{\tau 21^3} + \frac{e2e3}{m2c}\left(\frac{\xi^{(23)\,(23)\,(23)}\,4\,\xi\cos\left(\frac{2}{2}\right)\cos\lambda 2 + \xi\,\text{sen}\left(\frac{2}{3}\right)\cos\lambda 2}{c\,2\tau 23\,3}\frac{\vec{\xi}^{(23)},\dot{tu}^{(3)}}{\|\cdot\|^2}\right) -$$

$$- \frac{e2e3}{m2c\,3\tau 23}\left[\dot{tu}_1^{(3)}\cos\left(\frac{2}{1}\right)\cos\lambda 2 + \dot{tu}_2^{(3)}\,\text{sen}\left(\frac{2}{1}\right)\cos\lambda 2 + \dot{tu}_3^{(3)}\,\text{sen}\left(\frac{2}{1}\right)\lambda 2\right] - \frac{{}^2y}{3\,m2c^2}\left[\ddot{tu}_1^{(2)}\cos\left(\frac{2}{1}\right)\cos\lambda 2 + \ddot{tu}_2^{(2)}\,\text{sen}\left(\frac{2}{1}\right)\cos\lambda 2 + \ddot{u}^{(2)}\,\text{sen}\left(\frac{2}{1}\right)\lambda 2 \approx u_3\right]$$

$$\approx \frac{e2e1}{m2c\,\rho 2}\left(\frac{1}{2}\right) - \frac{\rho 2-\rho 3(\cos\left(\frac{2}{3}\right)\cos\lambda 2\cos\lambda 3+\text{sen}\,\lambda 2\,\text{sen}\,\lambda 3)}{e2e3m2}\frac{\vec{\xi}^{(23)},\dot{tu}^{(3)}}{\|\cdot\|^2}}{c\,3\tau 23\,3} -$$

$$- \frac{\ddot{\rho}3\,(\cos\left(\frac{2}{3}\right)\cos\lambda 2\cos\lambda 3+\sin\lambda 2\,\sin\lambda 3)+\rho 3\,\ddot{\lambda}3(\sin\lambda 2\cos\lambda 3-\cos\left(\frac{2}{3}\right)\sin\lambda.3\cos\lambda 2) - 3\,\ddot{\rho}\sin\left(\frac{2}{3}\right)\cos\lambda 2\cos\lambda 3+\rho 3\,c\,4\tau 23}{e2e3m2} - \frac{{}^2y_2\dddot{\rho}\,2}{m2c^3}.$$

Para la segunda ecuación de (A2), tenemos

$$\overset{(2)\,(21)\,(23)}{{}_{PAG}\alpha_{=\,GRAMO}\alpha} + G_\alpha + G\alpha\alpha^{(2)rad} \approx \frac{e2e1}{4\,m2c}\frac{x\alpha^{(2)}}{3\,\tau 21} + \frac{e2e3}{3m2c}\left(\frac{\xi_\alpha^{(23)}}{3\,\tau 23}\frac{\vec{\xi}^{(23)},\dot{tu}^{(3)}}{\|\cdot\|^2}-\frac{\dot{tu}_\alpha^{(3)}}{c\tau 23}\right) - \frac{2}{e\,2\,u\,3\,\alpha\,m2c}\ddot{u}^{(2)};$$

$$\ddot{\rho}2 = \frac{2P\,1\,2^{(2)\,(2)-\text{sen}\left(\frac{2P}{2}\right)+\cos\left(\frac{2}{1}\right)}}{\rho 2\,\text{porque}\,\lambda 2} = \frac{1}{\rho 2\,\text{porque}\,\lambda 2}\left[\frac{e2e1}{4m2c}\frac{-x\,1^{(2)\,(2)\,\text{sen}\left(\frac{2}{1}\right)+x\cos\left(\frac{2}{2}\right)}}{\tau 21^3}\right.$$

$$\left. + \frac{e2e3}{4m2c}\left(\frac{2}{1}\right)\left(\frac{3}{2}\right)\text{sen}\left(\frac{2}{1}\right)\frac{\xi\cos\left(\frac{2}{2}\right)}{c\,2\tau 23\,3}\frac{\vec{\xi}^{(23)},\dot{tu}^{(3)}}{\|\cdot\|^2} + \frac{\dot{tu}_1^{(3)}\,\text{sen}\left(\frac{2}{1}\right)-\tau 23\,\dot{tu}_2^{(3)}\,\text{porque}\left(\frac{2}{1}\right)}{{}^2y}\ddot{\alpha}^{(2)}\,\text{sen}\left(\frac{2}{1}\right)-u\,\ddot{u}^{(2)}\cos\left(\frac{2}{1}\right)\right] = 2 \qquad \text{(A3)}$$

$$= \frac{1}{\rho 2\,\text{porque}\,\lambda 2}\left[\frac{e2e3}{3m2c}\frac{-\xi\,\rho 3\,\text{sen}\left(\frac{2}{3}\right)\cos\lambda 3\,\tau 23}{3}\frac{\vec{\xi}^{(23)},\dot{tu}^{(3)}}{\|\cdot\|^2}\right.$$

$$\left. \ddot{\rho}3\cos\lambda 3\,\text{sen}\left(\frac{2}{3}\right)+\rho 3\,\ddot{\lambda}3\,\text{sen}\,\lambda 3\,\text{sen}\left(\frac{2}{3}\right)-3\cos 3\cos\left(\frac{2}{3}\right)-\rho 3+c\tau 23\right] - \frac{\dddot{\rho}}{2\,\rho 2\,\text{porque}\,\lambda 2}.$$

Para la tercera ecuación de (A2), tenemos

Machine Translated by Google

$$\ddot{\lambda}2 = \frac{-\cos\ 2\,\text{sen}\,\lambda2}{\rho2} P_1^{(2)}(2)(2) \frac{-\text{sen}\ 2\,\text{sen}\,\lambda2\cos\lambda2 +}{2\,3\,\rho2\,\rho2\,(3)} P =$$

$$= \frac{-\cos\ 2\,\text{sen}\,\lambda2\,\rho2}{4m2c}\frac{x_1^{(2)}}{\tau21^3} + \frac{e2e3}{3m2c}\frac{\xi_1^{(23)}}{\tau23^3}\frac{\vec{\xi}^{(23)},\vec{tu}}{\cdots} - \frac{\dot{u}^{(3)}}{1\,c\tau23}\frac{2_y}{3\,m2c^2}\ddot{tu}_1^{(2)}$$

$$\frac{e2e1}{} - +$$

$$\frac{-\text{sen}\ 2\,\text{sen}\,\lambda2+\rho2}{4m2c}\frac{x_2^{(2)}}{\tau21^3} + \frac{e2e3}{3m2c}\frac{\xi_2^{(23)}}{\tau23^3}\frac{\vec{\xi}^{(23)},\vec{tu}^{(3)}}{\cdots} - \frac{\dot{u}^{(3)}}{2\,c\tau23}\frac{2_y}{3\,m2c^2}\ddot{tu}_2^{(2)}$$

$$\frac{e2e1}{} - +$$

$$\lambda2\frac{\text{porque}}{+\rho2}\frac{x_3^{(2)}}{4m2c\ \tau21^3} + \frac{e2e3}{3m2c}\frac{\xi_3^{(23)}}{\tau23^3}\frac{\vec{\xi}^{(23)},\vec{tu}^{(3)}}{\cdots} - \frac{\dot{u}^{(3)}}{3\,c\tau23}\frac{2_y}{3\,m2c^2}\ddot{tu}_3^{(2)}$$

$$\frac{e2e1}{} - =$$

$$= \frac{e2e3}{3m2c}\frac{\rho3\cos\ 23\,\text{sen}\,\lambda2\cos\lambda3}{\tau23^3}\frac{\vec{\xi}^{(23)},\vec{tu}^{(3)}}{c} +$$

$$+\frac{e2e3}{3m2c}\ddot{3}\rho2\,\rho3\,(\cos\ 23\cos\lambda3\,\text{sen}\,\lambda2-\text{sen}\,\lambda3\cos\lambda2)+\rho3c\tau23\rho2\cos\ 23\,\text{sen}\,\lambda2\,\text{sen}\,\lambda3+\cos\lambda2\cos\lambda3\ \ 3\,\text{sen}\ 23\,\text{sen}\,\lambda2\cos\lambda3-\rho3 - \frac{2_y}{3\,m2c^2}\dddot{\lambda}2.$$

Apéndice C

Dado que, en el teorema principal, (m) (m = 1, 2, . . .) debe ser suficientemente pequeña, tomamos $r0^{(metro)} = r0\ 1/\mu(m = 1, 2, \ldots)$, entonces,

$$|rn(t)| = rn0 + rn\int_{tk}^{t}(1)(t1)dt1 = rn0 + \int_{tk}^{t} rn0\ (1) + rn\int_{tk}^{t1}(2)(t2)dt2\,dt1 = \ldots =$$

$$= \int_{tk}^{t} rn0\,(1)dt1 + \int_{tk}^{t}\int_{tk}^{t1} rn0\,(2)dt2dt1 + \ldots + \int_{tk}^{t}\int_{tk}^{t1}\cdots\int_{tk}^{Tk} tm{-}1\,rn(m)\,(tm)dtm\ldots dt1 \le$$

$$\le \frac{|rn0\,(1)|}{\mu}{}_{mi}\mu(t{-}Tk) + \frac{|rn0\,(2)|}{\mu2}{}_{mi}\mu(t{-}Tk) + \ldots + \frac{\omega m}{\mu m}\text{Rne}\,\mu(t{-}Tk) \le$$

$$1 \le e^{-\mu2}\ 1\,\mu T + e\,\mu3\ \ \mu T + \ldots + \frac{\omega m}{\mu m}\text{Rne}\,\mu(t{-}Tk) \approx \frac{\omega m}{\mu m}\text{Rne}\,\mu(t{-}Tk),$$

Eso significa $|rn(t)| \le \frac{\omega m}{\mu m}\text{Rne}\,\mu(t{-}Tk)$ (n = 2, 3), porque $\frac{1}{\mu} \approx 0$.

De la misma manera, $rn\ \ (1)(t) \le \frac{\omega m}{\mu m}\omega\text{Rne}\,\mu(t{-}Tk)$, $rn\ \ (2)(t) \le \frac{\omega m}{\mu m}\omega2\text{Rne}\,\mu(t{-}Tk)$, etcétera.

Apéndice D. Estimaciones para las derivadas superiores

Suponiendo $\beta \approx 0$, obtenemos

$$\frac{d\tau pq\,(t)}{dt} = \frac{\vec{\xi}^{(pq)},\vec{tu}^{(pag)} - \vec{\xi}^{(pq)},\vec{tu}^{(q)}}{\text{do}\ \vec{\xi}^{(pq)},\vec{\xi}^{(pq)} - \vec{\xi}^{(pq)},\vec{tu}^{(q)}} = \frac{\vec{\xi}^{(pq)},\vec{tu}^{(pag)} - \vec{tu}^{(q)}}{c\,2\tau pq - \vec{\xi}^{(pq)},\vec{tu}^{(q)}} \approx \frac{2\beta}{1-\beta} < 1 \qquad \beta = \frac{}{1\ 137} < \frac{}{1\ 3} < 1$$

$$\frac{d\xi^{(23)}}{2\,DT} = \frac{-x\,(t_2\tau23)^{(2)}(3)\,dx\,(t)}{dt^2} \qquad (2)(3)(t) - u\,(t - \tau23) \qquad -$$
$$(1 - .\ \tau23) \le 2c. = tu\ 2$$

Entonces,

$$\vec{tu}^{(norte)},\vec{tu}^{(norte)} = \dot{2}\,\rho n + \rho n^{2}\,\dot{n}\ \ 2\cos2\,\lambda n + \rho n\ \ \dot{\lambda}n\ \ 2\,2 \le c -\ \ 2 < c;\ \ \vec{tu}^{(2)},\vec{tu}^{(2)}/do^2 \approx 0;$$

$$(n) \dashrightarrow u, (n) \dashrightarrow u = \dot{}_{vueltas^2} + \rho n^{2}\,\dot{}\ \ 2\cos2\,\lambda n + \rho n\,\varphi n\ \ 2\,\dot{}\,2\,\eta n\ \ \rho n\,2\Phi n\,2 + \rho n\,2Yn\,2)e\,2\mu(t{-}Tk) \le \omega nce\mu(t{-}Tk) \le \omega n\,2(Rn\,2 +;$$

$$\ddot{\vec{tu}}^{(norte)},\ddot{\vec{tu}}^{(norte)} = \ddot{}_{vueltas^2} + \rho n^{2}\,\ddot{\varphi}_{vueltas}\ \ 2\cos2\,\lambda n + \rho n\ \ \ddot{\eta}n\ \ 2\ \ \omega n\ \ 2\,\overline{ce}\mu(t{-}Tk) \le ;$$

$$\dddot{\vec{tu}}^{(norte)},\dddot{\vec{tu}}^{(norte)} \le \dddot{}_{2m} + \rho n^{2}\,\dddot{\varphi}_{vueltas}\ \ 2\cos2\,\lambda n + \rho n\ \ \dddot{\eta}_{vueltas}\ \ 2\ \ \omega n\ \ 3\,\overline{ce}\mu(t{-}Tk) \le, \frac{}{1\ \tau23} \le \Delta; \frac{\text{do}}{} \frac{}{1\ \tau32} \le \overline{c}\,\Delta.$$

Además,

Machine Translated by Google

$$\frac{d\,2B(k)}{r2\,dt2} = \frac{dGr2}{dt} = \frac{\overset{(2)}{1}(2)(2)\,d(P\,\text{sen}\,\lambda2)\cos \quad 2\cos\lambda2+P\,\text{sen} \quad 2\cos\lambda2+P\,2\,3\,dt}{} =$$

$$= \overset{\cdot}{P}{}^{(2)}_{PAG}(2) + \overset{\cdot}{P}{}_{2}(2)(2)\,1\cos \quad 2\cos\lambda2 - P\,\varphi2\,\text{sen} \quad 2\cos\lambda2 - P\,1\,\eta2\cos \quad 2\,\text{sen}\,\lambda2 + \overset{\cdot}{P}{}^{(2)}_{3}(2)\,\text{sen}\,\lambda2 + P\,3\,\eta2\cos\lambda2 =$$

$$= \frac{}{dtt}\left[\frac{e2e1}{\tau21}\frac{x1^{(2)}}{4\,3\,m2c} + \frac{e2e3}{3m2c}\right]\frac{\xi_1^{(23)}}{3\,\tau23}\frac{\vec{\xi}^{\,(23)},\vec{tu}^{\,(3)}}{} - \frac{\dot{tu}_1^{(3)}}{c\tau23}\left[\frac{{}^{2}\gamma_{Z0\,3}}{m2c}\ddot{}^{(2)}\right]\cos \quad 2\cos\lambda2 -$$

$$- \left[\frac{e2e1}{\tau21}\frac{x1^{(2)}}{4\,3\,m2c} + \frac{e2e3}{3m2c}\right]\frac{\xi_1^{(23)}}{3\,\tau23}\frac{\vec{\xi}^{\,(23)},\vec{tu}^{\,(3)}}{} - \frac{\dot{tu}_1^{(3)}}{c\tau23}\left[\frac{{}^{2}\gamma_{Z0\,3}}{m2c}\ddot{}^{(2)}\right](\varphi2\,\text{sen} \quad 2\cos\lambda2 + \eta2\cos \quad 2\,\text{sen}\,\lambda2) +$$

$$+ \frac{d}{dt}\left[\frac{e2e1}{4\,m2c\,\tau21}\frac{x2^{(2)}}{3} + \frac{e2e3}{3m2c}\right]\frac{\xi_2^{(23)}}{3\,\tau23}\frac{\vec{\xi}^{\,(23)},\vec{tu}^{\,(3)}}{} - \frac{\dot{tu}_2^{(3)}}{c\tau23}\left[\frac{{}^{2}\gamma_{Z0\,3}}{2\,m2c}\ddot{}^{(2)}\right]\text{sen} \quad 2\cos\lambda2 +$$

$$+ \left[\frac{e2e1}{\tau21}\frac{x2^{(2)}}{4\,3\,m2c} + \frac{e2e3}{3m2c}\right]\frac{\xi_2^{(23)}}{3\,\tau23}\frac{\vec{\xi}^{\,(23)},\vec{tu}^{\,(3)}}{} - \frac{\dot{tu}_2^{(3)}}{c\tau23}\left[\frac{{}^{2}\gamma_{Z0\,3}}{2\,m2c}\ddot{}^{(2)}\right](\varphi2\cos \quad 2\cos\lambda2 - \eta2\,\text{sen} \quad 2\,\text{sen}\,\lambda2) +$$

$$+ \frac{d}{dt}\left[\frac{e2e1}{\tau21}\frac{x3^{(2)}}{4\,3\,m2c} + \frac{e2e3}{3m2c}\right]\frac{\xi_3^{(23)}}{3\,\tau23}\frac{\vec{\xi}^{\,(23)},\vec{tu}^{\,(3)}}{} - \frac{\dot{tu}_3^{(3)}}{c\tau23}\left[\frac{{}^{2}\gamma_{Z0\,3}}{3\,m2c}\ddot{}^{(2)}\right]\text{pecado}\,\lambda2 +$$

$$+ \left[\frac{e2e1}{4\,m2c\,\tau21}\frac{x3^{(2)}}{3} + \frac{e2e3}{3m2c}\right]\frac{\xi_3^{(23)}}{3\,\tau23}\frac{\vec{\xi}^{\,(23)},\vec{tu}^{\,(3)}}{} - \frac{\dot{tu}_3^{(3)}}{c\tau23}\left[\frac{{}^{2}\gamma_{Z0\,3}}{3\,m2c}\ddot{}^{(2)}\right]\eta2\,\text{sen}\,\lambda2 \le$$

$$\le 4\,3\,\frac{e2e1}{m2c\,\tau21}\frac{tu_1^{(2)}}{} - 3\,e2e1\frac{}{4\,m2c}\frac{x1^{(2)}\dot{}}{\tau21}\frac{\tau21+}{4} + \frac{e2e3}{3m2c}\frac{\dot{\xi}_1^{(23)}}{3\,\tau23}\frac{\vec{\xi}^{\,(23)},\vec{tu}^{\,(3)}}{} - \frac{(23\eta)\,3\dot{\xi}\tau23}{4\,\tau23}\frac{\vec{\xi}^{\,(23)},\vec{tu}^{\,(3)}}{} + 3\,\xi_1^{(23)}\frac{\dot{\vec{\xi}}^{\,(23)},\vec{tu}^{\,(3)}}{} + \frac{\vec{\xi}^{\,(23)},\vec{\ddot{tu}}^{\,(3)}}{} - \frac{\ddot{tu}_1^{(3)}}{c\tau23} + \frac{\dot{u}_{\tau23}^{(3)}\,1\,2}{c\tau23}\left[\frac{{}^{2}\gamma_{3\,m2c}}{}\right]\dddot{tu}_1^{(2)} +$$

$$+ \frac{|e2e1|}{m2c}\left[\frac{1}{2\,\rho} + \frac{\bar{c}\omega3}{3\Delta2}\frac{2+c \quad \rho\,3+3\Delta\,\rho}{}\right]\frac{\omega2\,2R2+}{d\omega}2\,(\Phi2 + Y2)e\,\mu(t-Tk)+$$

$$+ \frac{e2e1}{4\,m2c\,\tau21}\frac{tu_2^{(2)}}{3} - 3\,e2e1\frac{}{4\,m2c}\frac{x2^{(2)}\dot{}}{\tau21}\frac{\tau21\,4}{} + \frac{e2e3}{3m2c}\frac{\dot{\xi}_2^{(23)}}{3\,\tau23}\frac{\vec{\xi}^{\,(23)},\vec{tu}^{\,(3)}}{} - \frac{3\dot{\xi}\tau23\,4\,\tau23}{}\frac{\vec{\xi}^{\,(23)},\vec{tu}^{\,(3)}}{} + 3\,\xi_2^{(23)}\frac{\dot{\vec{\xi}}^{\,(23)},\vec{tu}^{\,(3)}}{} + \frac{\vec{\xi}^{\,(23)},\vec{\ddot{tu}}^{\,(3)}}{} - \frac{\ddot{tu}_2^{(3)}}{c\tau23} + \frac{\dot{u}_{\tau23}^{(3)}\,2\,2}{c\tau23}\left[\frac{{}^{2}\gamma_{3\,m2c}}{}\right]\dddot{tu}_1^{(2)} + 2$$

$$+ \frac{e2e1}{4\,m2c}\frac{x2^{(2)}}{3\,\tau21} + \frac{e2e3}{3m2c}\frac{\xi_2^{(23)}}{3\,\tau23}\frac{\vec{\xi}^{\,(23)},\vec{tu}^{\,(3)}}{} - \frac{\dot{tu}_2^{(3)}}{c\tau23}\left[\frac{{}^{2}\gamma_{Z0\,3}}{2\,m2c}\ddot{}^{(2)}\right](\Phi2 + Y2)e\,\mu(t-Tk)+$$

$$+ \frac{e2e1}{4\,m2c\,\tau21}\frac{tu_3^{(2)}}{3} - 3\,e2e1\frac{}{4\,m2c}\frac{x3^{(2)}\dot{}}{\tau21}\frac{\tau21\,4}{} + \frac{e2e3}{3m2c}\frac{\dot{\xi}_3^{(23)}}{\tau23\,3}\frac{\vec{\xi}^{\,(23)},\vec{tu}^{\,(3)}}{} - \frac{(23)\,3\dot{\xi}\tau23}{3\,4\,\tau23}\frac{\vec{\xi}^{\,(23)},\vec{tu}^{\,(3)}}{} + \frac{\xi_3^{(23)}}{3\,\tau23^{3}}\frac{\dot{\vec{\xi}}^{\,(23)},\vec{tu}^{\,(3)}}{} + \frac{\vec{\xi}^{\,(23)},\vec{\ddot{tu}}^{\,(3)}}{} - \frac{\ddot{u}_3^{(3)}}{+\,c\tau23}\frac{\dot{u}_{\tau23}^{(3)}\,3\,c\tau23}{2}\left[\frac{{}^{2}\gamma_{3\,m2c}}{}\right]\dddot{tu}_1^{(2)} + 3$$

$$+ \frac{e2e1}{4m2c}\frac{x^{(2)}}{3\,\tau21^{3}} + \frac{e2e3}{3m2c}\frac{\xi_3^{(23)}}{\tau23\,3}\frac{\vec{\xi}^{\,(23)},\vec{tu}^{\,(3)}}{} - \frac{\dot{u}_3^{(3)}}{c\tau23}\left[\frac{{}^{2}\gamma_{Z0\,3}}{m2c}\ddot{}^{(2)}\right]Y2e\,\mu(t-Tk)\,;$$

$$\frac{d\,2B(k)}{r2\,dt2} = \frac{dGr2}{dt} \le \frac{|e2e1|}{m2}\left[\frac{c\,4\Delta2\,c\,4\Delta}{} \quad \frac{2\,3p2\,8\omega3c\,2\omega3\,cc+\bar{+}+c\,4\Delta3\,c\,4\Delta4}{}\right]\frac{\omega2\,3R2+}{\omega}\,{}_{mi}e2e1|\,\mu(t\text{-}Tk)\Big|_{m2c}^{+}\left[\frac{1}{2\,x\,2} + \frac{\bar{c}\omega3}{3\Delta2}\frac{\rho\,2+ \quad \rho\,3+3\Delta\,c}{}\right]\frac{\omega2\,2R2+}{d\omega}2\,(\Phi2 + Y2)e\,\mu(t-Tk)+$$

$$+ \frac{|e2e1|}{m2}\left[\frac{\bar{\omega}}{}\,4\Delta4\,\frac{3p2\,8\omega3c\,2\omega3+\bar{+}}{c\,4\Delta2\,c\,4\Delta} + c\,4\Delta3\,c\,\frac{\bar{}}{\omega}\right]\frac{\omega2\,3R2+}{\omega}\,{}_{mi}e2e1|\,\mu(t\text{-}Tk)\Big|_{m2c}^{+}\left[\frac{1}{2\,x\,2} + \frac{\bar{c}\omega3}{3\Delta2}\frac{2+c \quad \rho\,3+3\Delta\,\rho}{}\right]\frac{\omega2\,2R2+}{d\omega}2\,(\Phi2 + Y2)e\,\mu(t-Tk)+$$

$$+ \frac{|e2e1|}{m2}\left[\frac{\omega}{}\,4\Delta4\,\frac{3p2\,8\omega3c\,2\omega3+\bar{+}}{c\,4\Delta2\,c\,4\Delta} + c\,4\Delta3\,c\,\frac{\bar{}}{\omega}\right]\frac{\omega2\,3R2}{\omega}\,{}_{mi}e2e1|\,\mu(t\text{-}Tk)\Big|_{m2c}^{+}\left[\frac{1}{2\,x\,2} + \frac{\bar{c}\omega3}{3\Delta2}\frac{2+c \quad \rho\,3+3\Delta\,\rho}{}\right]\frac{\omega2\,2R2+}{\omega}Y2e\,\mu(t-Tk) \le \omega2\,2R2e\,\mu(t-Tk)\ .$$

Machine Translated by Google

Apéndice E. Límites superiores para derivados parciales

Dado que $\Delta_{23} = \Delta_2(t) - \Delta_3(t - \tau_{23})$ y $\vec{\xi}^{(23)}$, $\vec{\xi}^{(3)} \to t_u$ $\leq c\tau_{23}\bar{\omega}_{3e}$ μ_T, uno obtiene

$$\frac{\partial Gr_2}{\partial \rho_2} \leq \frac{|e_2 e_1|}{m_2}\left(\frac{2}{3c\rho_2} + \frac{1}{3c}\frac{\bar{c}\tau_{23}\bar{\omega}_{3e}}{c\,3\tau_{23}\,3}\right)\mu(t-T_k) \leq \frac{|e_2 e_1|}{m_2}\left(\frac{2}{d_0 3\rho} + \frac{\bar{c}\omega_3}{c\,5\Delta_2}\right)m_i\mu_T \equiv \frac{\partial \bar{Gr}_2}{\partial \rho_2};$$

$$\frac{\partial Gr_2(r_2,\varphi_2,\eta_2,r_3,\varphi_3,\eta_3)}{\partial r_2} = 0; \quad \frac{\partial Gr_2}{\partial \rho_3} \frac{2|e_2 e_3|\omega_3 c+\rho(\Phi_3+Y_3)\Delta}{m_2} \leq \frac{\bar{c}}{c6\Delta_2} m_i\mu_T \equiv \frac{\partial \bar{Gr}_2}{\partial \rho_3};$$

$$\frac{\partial Gr_2}{\partial r_3} \leq \frac{|e_2 e_3|}{m_2}\frac{2(\quad)\rho_3+\rho\,3\,c}{2} + \Delta_2 \equiv \frac{\partial \bar{Gr}_2}{\partial r_3};$$

$$\frac{\partial Gr_2}{\partial \lambda_2} \leq \frac{|e_2 e_3|}{m_2}\bar{\omega}_{3c} \frac{\rho\,3ec}{5\Delta_2}\mu_T + \frac{\rho_2+\rho\,3\,3\,c}{3\Delta_3\,c\,4\Delta}\rho\,2\,1 + c \equiv \frac{\partial \bar{Gr}_2}{\partial \lambda_2};$$

$$\frac{\partial Gr_2}{\partial \lambda_3} \leq \frac{|e_2 e_1|}{m_2}\left(\rho\,2+\rho\,3\,c\bar{\omega}_{3e} + \frac{\mu_T}{c\,2\Delta_2}\right)\left(\rho_2+\rho\,3+\right)\left(\rho\,2\omega_{3c}\bar{\omega}_{3c} + c\,3\bar{\Delta}_3\right) \equiv \frac{\partial \bar{Gr}_2}{\partial \lambda_3};$$

$$\frac{\partial Gr_2(r_2,\varphi_2,\eta_2,r_3,\varphi_3,\eta_3)}{\partial \varphi_2} = 0; \quad \frac{\partial Gr_2(r_2,\varphi_2,\eta_2,r_3,\varphi_3,\eta_3)}{\partial \varphi_3} = 0;$$

$$\frac{\partial Gr_2}{\partial \lambda_2} \leq \frac{|e_2 e_3|}{m_2} m_i\mu_T \bar{c}\omega_3 \frac{\rho_2+\rho}{3c}\frac{\leftarrow_3}{4\Delta}\frac{\Delta+\rho\,2\,1}{c}+c\,3\Delta_3 = \frac{\partial \bar{Gr}_2}{\partial \lambda_2}; \quad \frac{\partial Gr_2(r_2,\varphi_2,\eta_2,r_3,\varphi_3,\eta_3)}{\partial \eta_2} = 0;$$

$$\frac{\partial Gr_2}{\partial \lambda_3} \leq \frac{|e_2 e_1|}{m_2}\frac{\bar{c}\omega_3}{c4\Delta}\frac{(\rho\,2+\rho\,3\quad)(\Delta e\,\mu T+2\,c\quad\rho\,2)}{2\Delta_2}\partial Gr_2 + 1 \equiv \partial\bar{\lambda}_3;$$

$$\frac{\partial Gr_2(r_2,\varphi_2,\eta_2,r_3,\varphi_3,\eta_3)}{\partial \eta_3} = 0; \quad \frac{\partial Gr_2}{\partial r_2} \leq \frac{y_2^2}{3m_2c} \equiv \frac{\partial \bar{Gr}_2}{\partial r_2}.$$

límites superiores de $\frac{\partial Gr_3}{\partial \rho_2}, \ldots, \frac{\partial Gr_3}{\partial r_3}\frac{\partial Gr_3}{\partial r_3}$ Los $, \ldots, \frac{\partial G\eta_3}{\partial r_3}$ se puede obtener de forma similar.

Referencias

1. Angelov, VG El problema electromagnético de los tres cuerpos con términos de radiación: derivación de ecuaciones de movimiento (I). Resultados Anal no lineal. 2020, 3, 45–58.
2. Angelov, VG El problema electromagnético de los tres cuerpos con términos de radiación: existencia-singularidad de la órbita periódica (II). Resultados Anal no lineal. 2020, 3, 137-158.
3. Angelov, VG El problema de los N-cuerpos en la electrodinámica clásica. Física. Ensayos 1993, 6, 204–211. [Referencia cruzada]
4. Angelov, VG Sobre las ecuaciones originales de Dirac con término de radiación. Lib. Matemáticas. 2011, 31, 57–86. 5. Conductor, RD Un sistema neutral con retardo dependiente del estado. J. Dif. Ecuación. 1984, 54, 73–86. [Referencia cruzada]
6. Pauli, W. Relativitatstheorie. En Encyklopedie der Mathematischen Wissenschaften; Banda. 5, peso 4, art. 19; Springer: Berlín/Heidelberg, Alemania, 1921.
7. Synge, JL Sobre el problema electromagnético de los dos cuerpos. Proc. Roy. Soc. 1940, A177, 118-139.
8. Angelov, VG Teoremas del punto fijo para aplicaciones contractivas (Φ, j1 , j2 ,..., jm) en espacios uniformes. Toro. Matemáticas de Allahabad. Soc. 2023, 2. en prensa.
9. Bauer, G.; Deckert, DA; Durr, D.; Hinrichs, G. Sobre la irreversibilidad y la radiación en la electrodinámica clásica de partículas puntuales. J. Estadística. Física. 2014, 154, 610. [Referencia cruzada]
10. Deckert, DA; Durr, D.; Vona, N. Ecuaciones de retardo del tipo Wheeler-Feynman. J. Matemáticas. Ciencia. 2014, 202, 623–636. [Referencia cruzada]
11. Bauer, G.; Deckert, DA; Durr, D. Sobre la existencia de dinámica en el electromagnetismo de Wheeler-Feynman. J. Aplica. Matemáticas. Física. 2012, 64, 1087–1124. [Referencia cruzada]
12. de Luca, J. Electrodinámica variacional de átomos. Prog. Electromagn. Res. B 2013, 53, 1-11. [Referencia cruzada]
13. Simaciu, CC Modelo clásico de electrón en electrodinámica estocástica. Rev. Méx. Fis. 2001, 47, 392–394.
14. Abouelmagd, EI; García Guirao, JL; Llibre, J. Órbitas periódicas del problema de tres cuerpos restringido y cuantificado. Universo 2023, 9, 149. [Referencia cruzada]
15. Pan, S.; Hou, X. Análisis de órbitas periódicas de transición de resonancia en el problema circular restringido de tres cuerpos. Aplica. Ciencia. 2022, 12, 8952. [Referencia cruzada]
16. Abouelmagd, EI; Guirao, JLG; Llibre, J. Sobre las órbitas periódicas de los problemas de dos y tres cuerpos perturbados. Galaxias 2023, 11, 58. [Referencia cruzada]
17. Abouelmagd, EI; Alhowaity, S.; Diab, Z.; Guirao, JLG; Shehata, MH Sobre las soluciones periódicas para el problema de las colinas cuantificadas espacialmente perturbadas. Matemáticas 2022, 10, 614. [CrossRef]
18. Cseh, J.; Dang, P.; Szilágyi, S.; Lévai, G. Regularidad orbital de exoplanetas y simetrías del problema de Kepler. Simetría 2023, 15, 2114. [Referencia cruzada]
19. Angelov, VG Trayectorias periódicas, de transición y de escape para el problema de 2 cuerpos 3D-Kepler de la electrodinámica clásica. Física. Ciencia. Int. J. 2023, 27, 31–72. [Referencia cruzada]
20. Angelov, VG Un método para el análisis de líneas de transmisión terminadas por cargas no lineales; Nova Science: Nueva York, NY, EE. UU., 2014.
21. Pohl, RW Optik und Atomphysik; Springer: Berlín/Heidelberg, Alemania, 1963.
22. Bohr, N. Teoría atómica y descripción de la naturaleza; Cambridge University Press: Teddington, Reino Unido, 1934.

Machine Translated by Google

23. Sommerfeld, A. Atombau und Spektrallinien; Friedrich Vieweg & Sohn, Braunschweig: Berlín, Alemania, 1951.
24. Angelov, VG Problema de espín de tres cuerpos de la electrodinámica clásica con términos de radiación (I): derivación de ecuaciones de espín. Resultados Anal no lineal. 2021, 4, 1–20. [Referencia cruzada]

Descargo de responsabilidad/Nota del editor: Las declaraciones, opiniones y datos contenidos en todas las publicaciones son únicamente de los autores y contribuyentes individuales y no de MDPI ni de los editores. MDPI y/o los editores renuncian a toda responsabilidad por cualquier daño a personas o propiedad que resulte de cualquier idea, método, instrucción o producto mencionado en el contenido.

Machine Translated by Google

AppliedMath

MDPI

Artículo

Nuevas Cópulas Bivariadas vía Distribución Lomax Distorsiones generadas

Fadal Abdullah Ali Aldhufairi 1,* y Jungsywan H. Sepanski 2

1 Departamento de Matemáticas, Universidad King Khalid, Asir-Abha 61421, KSA, Arabia Saudita
2 Departamento de Estadística, Ciencias Actuariales y de Datos, Universidad Central Adicional de Michigan, Mt. Pleasant, MI 48859, EE. UU.; sepan1jh@cmich.edu
* Correspondencia: aldhu1fa@cmich.edu

Resumen: Desarrollamos un marco para crear funciones de distorsión que se utilizan para construir nuevas cópulas bivariadas. Se logra transformando variables aleatorias no negativas con distribuciones relacionadas con Lomax. En este artículo aplicamos las distorsiones a las cópulas base de la independencia, las cópulas de Clayton, Frank y Gumbel. Se exploran las propiedades del coeficiente de dependencia de la cola, el orden de la cola y el orden de concordancia para las nuevas familias de cópulas distorsionadas. Realizamos un estudio empírico utilizando los rendimientos netos diarios de las acciones de Amazon y Google desde enero de 2014 hasta diciembre de 2023. Comparamos los populares modelos de cópula de Clayton, Gumbel, Frank y Gauss con sus correspondientes modelos de cópula distorsionados inducidos por la unidad-Lomax y la unidad. -distorsiones de Pareto inversas . Las nuevas familias de cópulas de distorsión están equipadas con parámetros adicionales inherentes a la función de distorsión, lo que proporciona más flexibilidad y se ha demostrado que funcionan mejor que las cópulas base. Después de analizar los datos, hemos descubierto que los extremos conjuntos de las acciones de Amazon y Google tienen más probabilidades de obtener rendimientos netos diarios elevados que de rendimientos netos diarios bajos.

Palabras clave: Arquímedes; concordancia; cópula de Clayton; distorsión; Cópula franca; cópula de Gumbel; τ de Kendall ; distribución Lomax; coeficiente de dependencia de la cola; orden de cola

check for updates

Cita: Aldhufairi, FAA; Sepanski, JH Nuevas Cópulas Bivariadas vía Distribución Lomax generada Distorsiones. Matemáticas aplicadas 2024, 4, 641–665. https://doi.org/10.3390/appliedmath4020035

Recibido: 29 de marzo de 2024
Revisado: 7 de mayo de 2024
Aceptado: 8 de mayo de 2024
Publicado: 17 de mayo de 2024

Copyright: © 2024 por los autores. Licenciatario MDPI, Basilea, Suiza. Este artículo es un artículo de acceso abierto. distribuido bajo los términos y condiciones de los Creative Commons Licencia de atribución (CC BY) (https://creativecommons.org/licenses/by/4.0/).

1. Introducción

Las descripciones y mediciones de la correlación y dependencia entre riesgos y pérdidas han sido importantes en diversos campos, como las finanzas, la gestión de riesgos y los estudios actuariales. Cópulas bivariadas de diferentes estructuras de dependencia pueden modelar la dependencia de la cola en riesgos extremos Ref. [1] pero ser independiente de las distribuciones marginales de los riesgos. Árbitro. [2] ilustró numérica y gráficamente los tipos de relaciones que varias medidas de asociación basadas en cópulas pueden detectar. Árbitro. [3] abordó las matemáticas de las funciones de cópula ilustradas con una aplicación financiera a temas financieros en fijación de precios de derivados y análisis de riesgo crediticio. El artículo fundamental de [4] demostró sus aplicaciones prácticas, como la estimación de la mortalidad en vidas conjuntas y los modelos multidecrementales.
Las pérdidas extremas o de cola tienden a ocurrir juntas; ver [5]. Árbitro. [6] emplearon cópulas para estudiar el efecto de la dependencia de la cola y la cola mediante la cuantificación de riesgos extremos. Árbitros. [7,8] aplicaron modelos de cópula para investigar los crecientes extremos hidroclimáticos asociados con un clima más cálido; ver también [9] para modelos hidrológicos y de inundaciones. En general, el modelado de cópulas ha demostrado ser una herramienta eficaz para analizar estructuras dependientes entre variables.

El teorema de Sklar [10] establece que la función de distribución acumulativa conjunta (cdf) se puede expresar como el producto de los marginales y la cópula y, a la inversa, la cópula se puede determinar de forma única si conocemos la cdf y los marginales. Por ejemplo, la cópula gaussiana se deriva de la distribución gaussiana bivariada y también se puede utilizar para generar nuevas distribuciones de probabilidad bivariadas mediante (A1) en el apéndice; ver [11] para resúmenes de los métodos de construcción de cópulas. Los preliminares necesarios sobre cópulas bivariadas recopiladas en el Apéndice A se pueden encontrar en [11,12].

Machine Translated by Google

Más recientemente, la Ref. [13] ampliaron las cópulas tradicionales de Arquímedes de un parámetro integrando los márgenes generados por log-gamma. Árbitro. [14] propusieron una clase de transformación de cópula de independencia bivariada de la forma C(u, v) = uv f((1 − u),(1 − v)), donde f es dos veces diferenciable. Sea Π(u, v) = uv la cópula de independencia. Generalizando la cópula de MGF dada por π(u, v) + αxy(1 − u)(1 − v), Refs. [15,16] construyeron nuevas cópulas de la forma Π(u, v) + Q(u, v; α), donde Q es una función de perturbación que involucra funciones trigonométricas, hiperbólicas, logarítmicas o exponenciales. Árbitro. [7] utilizó un truncamiento de la función de distribución semilogística log-cóncava F como generador multiplicativo de Arquímedes para F(u)F(v) . construir las cópulas de la forma F En este artículo, nos ocupamos de la construcción de nuevas cópulas bivariadas mediante funciones de distorsión. Una función T se llama función de distorsión si es continua y creciente en el intervalo unitario I = = 0 y T(1) = 1. Una nueva familia de cópulas nacida $^{-1}$ [0, 1], con T(0)

de la distorsión está dado por

$$CT(u, v) = TCT^{-1}(u), t^{-1}(v), tu, v \quad [0, 1], \tag{1}$$

T se denomina distorsión admisible si (1) es una cópula. Si la cópula inicial es de Arquímedes con generador ψ, entonces CT es de Arquímedes con generador ψ ◦ T en $^{-1}$; ver [17,18]. Teorema 3.3.3 [11] (p. 96) muestra que T es admisible si y sólo si T es creciente y convexo; ver también [19]. Este resultado clave dicta el requisito de convexidad y abre la puerta a exploraciones de funciones de distorsión admisibles. Árbitro. [20] demostró que T es admisible si T ◦ exp : (−∞, 0) → [0, 1] es log-convexo y sugirió varias funciones de distorsión.
Árbitro. [21] propusieron aplicar la distorsión solo a la función de cópula, solo a los marginales o a ambos. La cópula inducida en (1) es el resultado de distorsiones tanto de la cópula como de los marginales.

Árbitros. [22,23] construyeron nuevas familias de cópulas mediante distorsiones beta y cdf de Kumaraswamy.Ref. [24] emplearon la distorsión unitaria Lomax. Árbitro. [25] estudiaron familias de cópulas generadas por una distorsión unitaria de Weibull. Árbitro. [26] investigaron las propiedades de las cópulas distorsionadas de Gompertz unitario y las aplicaron para analizar los datos antropocéntricos. Las distorsiones unitarias de Lomax, unitarias de Weibull y unitarias de Gompertz se derivan de una transformación exponencial de las variables aleatorias de Lomax, Weibull y Gompertz.

Motivados por el hecho de que la CDF de una variable aleatoria continua con soporte de intervalo unitario cumple con la definición de función de distorsión, proponemos una transformación que convierte una variable aleatoria no negativa en una con soporte de intervalo unitario, lo que, en consecuencia, establece una función de distorsión. Luego, la distorsión se puede utilizar para generar nuevas familias de cópulas distorsionando las existentes. Al igual que en otras construcciones de nuevas cópulas en la literatura, el objetivo es obtener nuevas familias de cópulas que puedan explicar una gama más amplia de valores de dependencia de la cola. Con los parámetros en la cdf de distorsión, las cópulas distorsionadas tienen parámetros adicionales además de los de las cópulas existentes, haciéndolas más flexibles.

El documento está organizado de la siguiente manera. La sección 2 comienza con el mecanismo propuesto para generar nuevas funciones de distorsión y espacios de parámetros de distorsión admisibles que se estudiarán más a fondo. La sección 3 proporciona generadores de Arquímedes para las nuevas familias de cópulas distorsionadas cuando las cópulas base son la independencia, Clayton, Frank y Gumbel. Se presenta la familia de cópulas de independencia distorsionadas para que sirva como validación de los resultados obtenidos en este artículo. Las secciones 4 y 5 investigan las propiedades de la dependencia de la cola, el orden de la cola, el orden de concordancia y la tau de Kendall. La sección 6 contiene los resultados numéricos de un estudio de simulación y una aplicación empírica, seguidos de observaciones finales. Los preliminares de cópula y las derivaciones de órdenes de cola y ordenamiento de concordancia se incluyen en los Apéndices A y B.

2. Método propuesto

Sea Y una variable aleatoria continua no negativa con CDF F. Considere lo siguiente transformación de la variable aleatoria Y:

Machine Translated by Google

$$X = \frac{Y}{Y\theta + \dot{Y}}, \ \theta > 0. \tag{2}$$

La variable aleatoria X tiene un soporte del intervalo unitario y su CDF está dada por

$$G(x) = P(X \le x) = PY \le \frac{\theta x}{1 - x} = F\left(\frac{\theta x}{1 - x}\right). \tag{3}$$

La cdf G y su cuantil o función inversa pueden servir como funciones de distorsión para desarrollar nuevas cópulas. Si la variable Y asume cualquier valor en la recta numérica real, se puede considerar una transformación de su valor absoluto; es decir, |Y|/(1 + |Y|), cuya CDF no es tan sencilla como (3).

A continuación demostramos el método siendo F las distribuciones Lomax no negativas y Lomax inversa, y derivamos el espacio de parámetros en el que cada una de las distorsiones es convexa. Se supone que todos los parámetros de la función de distribución generadora F son positivos. El símbolo primo, como G $'$ o G ", denota la derivada de una función.

Ejemplo 1. Distorsión Unit-Lomax (UL) y su cuantil (QUL). Sea Y una variable aleatoria Lomax o Pareto Tipo II con una CDF dada por $F(y) = 1 - [\beta/(\beta + y)]$ donde $y > 0$. En este caso, la transformación en (2) produce la distorsión dado por

$$\text{GRAMO}(x) = 1 - \left(\frac{\beta(1 - x)}{\beta(1 - x) + \theta x}\right)^{\alpha}, \tag{4}$$

donde $0 < x < 1$. Tenga en cuenta que, por ejemplo,

$$\left(\frac{2(1 - x)}{2(1 - x) + 1x}\right)^{\alpha} = \left(\frac{4(1 - x)}{4(1 - x) + 2x}\right) = \left(\frac{8(1 - x)}{4(1 - x) + 4x}\right)^{\alpha}. \tag{5}$$

La ecuación (5) demuestra que múltiples valores de los parámetros (β, θ) dan lugar a la misma distorsión en (4) y, por tanto, los parámetros no pueden identificarse de forma única. Por tanto, para resolver este problema, con reparametrización, consideramos la distorsión GL (UL) y su inversa QL (QUL) dada por

$$\text{GL}(x) = 1 - \left(\frac{(1 - x)}{(1 - x) + \theta x}\right)^{\alpha}; \ \text{QL}(x) = 1 + \frac{1}{\theta[(1 - x)^{-1/\alpha} - 1]^{-1}}.$$

Lema 1. La distorsión $\text{GL}(x) = 1 - (1 - x)^{\alpha} (1 - x) + \theta x$ y $0 < \alpha \le 1$. $^{-1}$ es convexo en I si $0 < \theta \le 1$

Prueba. Sea $A(x) = (1 - x) + \theta x$, $B(x) = (1 - x)[A(x)]^{-1}$. Por y luego $\text{GL}(x) = 1 - [B(x)]^{\alpha}$. simplicidad, se elimina la notación del argumento de (x) . Tenga en cuenta que $(B(x))^{-1} = (1 - x)^{-1}$. Las derivadas están dadas por

$$\begin{aligned} B' &= -\theta A^{-2}; \ \text{GRAMO}' = \alpha\theta B^{\alpha-1} A^{-2} \\ \text{GRAMO}'' &= -\alpha(\alpha - 1)\theta^2 B^{\alpha-2} A^{-4} - 2\alpha\theta B^{\alpha-1} A^{-3} (\theta - 1) (\alpha - 1)\theta = \\ &\alpha\theta B^{\alpha-1} A^{-3} + 2(\theta - 1) . 1 - x \end{aligned} \tag{6}$$

Si $0 < \theta \le 1$ y $0 < \alpha \le 1$, entonces $G''(x) \ge 0$ para todo x I. □

Lema 2. La distorsión $\text{QL}(x) = 1/ 1 + \theta (1 - x)^{\alpha} \ge 1$. $^{-1/\alpha - 1}$ es convexo en I si $\theta \ge 1$ y

Prueba. Dado que QL es la función inversa de GL, y GL es una función creciente, por (6), tenemos obtener este lema. □

Machine Translated by Google

Ejemplo 2. Distorsión de Pareto inversa unitaria (UIP) y su cuantil (QUP). Considere la variable aleatoria de Pareto inversa, definida como el recíproco de una variable aleatoria de Lomax, con una CDF dada por $F(y) = y/(\beta + y)$ donde $y > 0$. En este caso, (3) da

$$G(x) = \left(\frac{\theta x}{\beta(1-x) + \theta x}\right)^{\alpha},$$

donde $0 < x < 1$. Por la misma razón explicada por (5), proponemos la distorsión GP (UIP) y su inversa QP (QUP) dada por

$$GP(x) = \left(\frac{\theta x}{(1-x) + \theta x}\right)^{\alpha}; \quad QP(x) = 1 + \frac{1}{\theta(x^{-1/\alpha} - 1)}.$$

Lema 3. La distorsión $GP(x) = \left(\frac{\theta x}{(1-x) + \theta x}\right)^{\alpha}$ es convexo en I si $0 < \theta \le 1$ y $0 < \alpha \le 1$.

Prueba. Sean $A(x) = (1-x) + \theta x$ y $B(x) = \theta x[A(x)]^{-1}$. Entonces $[B(x)A(x)]^{-1} = 1/(\theta x)$ y $GP(x) = [B(x)]^{\alpha}$. Por simplicidad, se elimina la notación del argumento de (x). Las derivadas relevantes están dadas por

$$\begin{aligned} B' &= \theta(A) = {}^{-1} - \theta x A^{-2}(\theta - 1) = \theta A^{-2}; \; GRAMO_{PAG}' = \alpha\theta B^{\alpha-1} A^{-2} \\ GRAMO_{PAG}'' &\; \alpha(\alpha - 1)\theta^2 B^{\alpha-2} A^{-4} - 2\alpha\theta B^{\alpha-1} A^{-3}(\theta - 1) = \alpha\theta B^{\alpha-1} A^{-3} \\ & x^{-1}[(\alpha - 1) - 2(\theta - 1)x]. \end{aligned} \tag{7}$$

Si $0 < \theta \le 1$ y $\alpha \ge 1$, entonces la segunda derivada $G_{PAG}'' \ge 0$ para todo x I. □

Lema 4. $QP(x) = [1 + \theta(x^{-1/\alpha} - 1)]^{-1}$ es convexo en I si $\theta \ge 1$ y $0 < \alpha \le 1$.

Prueba. Este lema se sigue ya que QP es la función inversa de GP y GP es creciente y cóncava por (7) cuando $\theta \ge 1$ y $0 < \alpha \le 1$. □

En resumen, al aplicar la transformación propuesta en (2) a variables aleatorias no negativas relacionadas con Lomax, obtenemos cuatro nuevas distorsiones admisibles tabuladas en la Tabla 1. Se muestran dos formas funcionales que serían útiles para los cálculos.

Tenga en cuenta que los espacios de parámetros admisibles son sólo una condición suficiente para que CT en (1) sea una cópula. La distorsión de potencia dual es un caso especial de GL con $\theta = 1$. La distorsión de potencia es un caso especial de GP con $\theta = 1$. Cuando $\theta = \alpha = 1$, es importante señalar que todas las distorsiones de la Tabla 1 se convierten en función de identidad. Esto significa que cuando estas distorsiones se aplican a una cópula base, la familia de cópulas resultante ofrece una mayor flexibilidad cuando se ajusta a los datos porque incluye la cópula base como un caso especial.

Observación 1. Sea Y una variable aleatoria Lomax. También consideramos la CDF de $Y^{1/\tau}$, es decir, la distribución de Burr, dada por $F(y) = 1 - (1 + y^{\tau})^{-\alpha}$, como CDF generadora. Sin embargo, no existe un espacio de parámetros en el que la distorsión resultante es convexa. Otro candidato generador de CDF es el CDF Lomax exponencial de la forma $1 - F^{\gamma}/(\gamma + y)^{\alpha\gamma}$, que es más complejo y se investigará en el futuro.

Observación 2. En lugar de (2), la transformación de $X = 1/(1 + Y)$ también da lugar a una CDF de distorsión. La CDF de $X = 1/[1 + Y]$ viene dada por, para $0 < x < 1$,

$$G(x) = P(X \le x) = 1 - F_Y\left(\frac{1 - X}{\text{incógnita}}\right).$$

En este caso, cuando Y tiene una distribución Lomax en el Ejemplo 1, derivamos el GP. Cuando Y tiene una distribución Lomax en el Ejemplo 2, derivamos el GL.

Machine Translated by Google

Tabla 1. Distorsiones propuestas y espacios de parámetros admisibles.

Distorsión	Formulario de función		Espacio de parámetros convexo
U-Lomax (UL)	GL(x)	$= 1 - \left(\frac{1-x}{1-x+\theta x}\right)^{\alpha}$	$0 < \theta \le 1,\ 0 < \alpha \le 1$
Cuantil U-Lomax (QL)	CL(x)	$= 1 - \frac{(1-x)^{-1/\alpha}}{\{1+\theta[(1-x)^{-1}-1]\}^{\alpha}-1}$ $= \frac{(1-x)^{-1/\alpha}}{\theta+(1-x)^{-1/\alpha}-1}$ $= \frac{1}{1+\theta[(1-x)^{-1/\alpha}-1]-1}$	$\theta \ge 1,\ \alpha \ge 1$
U-Pareto inverso (UIP)	médico de cabecera(x)	$= \frac{\theta x}{1-x+\theta x}^{\alpha}$ $= \frac{1}{[1+\theta^{-1}(x^{-1}-1)]^{\alpha}}$	$0 < \theta \le 1,\ \alpha \ge 1.$
Cuantil U-Pareto inverso (QUP)	QP(x)	$= \frac{1}{x^{-1/\alpha}+\theta(1-x^{1/\alpha})}$ $= \frac{1}{1+\theta(x^{-1/\alpha}-1)}$	$\theta \ge 1,\ 0 < \alpha \le 1$

Observación 3. Hay una multitud de CDF que pueden usarse como CDF generadora F en (3). Por ejemplo, una distribución exponencial con media β. En este caso, (3) da una distorsión candidata dada por $1 - \exp\{-x/[\beta(1-x)]\}$. Otro ejemplo, un Pareto generalizado con una PDF dada por

$$f(y) = \frac{\Gamma(\alpha+\tau)}{\Gamma(\alpha)\Gamma(\tau)} \frac{\theta^{\tau} y^{\tau-1}}{(y+\theta)^{\alpha+\tau}}$$

donde $x > 0$, $\alpha > 0$, $\tau > 0$ y $\theta > 0$.

3. Familias de nuevas cópulas distorsionadas

La forma general de cópulas construidas mediante una distorsión T está en (1). Las cópulas de Arquímedes se utilizan ampliamente [27] y son reconocidas por su flexibilidad para modelar la dependencia dentro de variables aleatorias multivariadas. Formulamos los generadores de Arquímedes resultantes de las distorsiones de la Sección 3.1 cuando las cópulas base son independencia, Clayton, Frank y Gumbel; consulte la Tabla A1 en el Apéndice A. Las cópulas base, excepto la cópula de independencia, tienen un solo parámetro, y las familias de las cópulas distorsionadas contienen más parámetros y son menos restrictivas. Sea r el parámetro en la cópula base.

3.1. Generadores de Arquímedes

Ejemplo 3 (Cópula de la Independencia). Los generadores de Arquímedes son

– Generador de independencia UL: $\log\{\theta(1-u)^{-1/\alpha} - 1 + 1\}^{-1}$.

– Generador de independencia QUL: $-\log\{1 - 1 + \theta(1-u)^{-1} - 1\}^{-\alpha}$.

– Generador de independencia UIP: $\log\{1 + \theta(u^{-1/\alpha} - 1)\}$.

– Generador de independencia QUP: $\log\{1 + \theta^{-1}(u^{-1} - 1)^{\alpha}\}$.

Ejemplo 4 (Cópula Clayton). En este caso obtenemos lo siguiente: –

Generador UL-Clayton: $\theta\{(1-u)^{-1/\alpha} - 1\}^{-1} + 1\}^{-r} - 1/r$.

– Generador QUL-Clayton: $1 - 1 + \theta((1-u)^{-1} - 1)^{-\alpha}\}^{-r} - 1/r$.

– Generador UIP-Clayton: $1 + \theta\{u^{-1/\alpha} - 1\}^{-r} - 1/r$.

– Generador QUP-Clayton: $1 + \theta\{(u^{-1} - 1)^{\alpha}\}^{r} - 1/r$.

Ejemplo 5 (Cópula franca). Cuando la base de la cópula C es Frank, obtenemos lo siguiente:

Machine Translated by Google

– Generador UL-Frank: − log 1 + e $^{-r[\theta((1-u)^{-1/\alpha}-1)^{-1}+1]}$ $_{-1}$ − 1 /(e $^{-r}$ − 1) .

– Generador QUL-Frank: − log 1 + e $^{-r[1-(1+\theta((1-u)^{-1}-1))^{-\alpha}]}$ − 1 /(e $^{-r}$ − 1) .

Generador UIP-Frank: − log 1 + e − 1)r[1+θ(u−1)]α + 1 /(e −1r/(e1)r.–

– Generador QUP-Frank: − log 1 + e $^{-r[1+\theta^{-1}}$

Ejemplo 6 (Cópula de Gumbel). Para la cópula de Gumbel obtenemos lo siguiente:

– Generador UL-Gumbel: − log θ((1 − u) $^{-1/\alpha}$ − 1) $^{-1}$ + 1 ^{r} .

– Generador QUL-Gumbel: − log 1 − (1 + θ((1 − u) $^{-1}$ − 1))−α ^{r} .

– Generador UIP-Gumbel: log 1 + θ(u − 1) $^{-1/\alpha}$ ^{r} .

– Generador QUP-Gumbel: log 1 + θ $^{-1}$ (tú$^{-1}$ − 1) $^{\alpha}$ ^{r} .

3.2. Distorsiones de la cópula de la Independencia

Presentamos aquí las familias de cópulas construidas mediante distorsiones de la cópula independiente con el objetivo de demostrar la versatilidad de la distorsión y validar los resultados presentados en este artículo. Las distorsiones propuestas distorsionan la cópula independiente libre de parámetros en diversas familias de cópulas adornadas con dos parámetros y mejor flexibilidad.

Tenga en cuenta que G_l^{-1} = CL; consulte la Tabla 1. Para la cópula de independencia de UL, C(u, v) = uv,

$$\mathrm{CGL}(tu, v) = 1 - \frac{1}{1 + \theta\,1 - \mathrm{CG}_l^{-1}{}_{(u),\,\mathrm{GRAMO}}{}_l^{-1}(tú)^{-1} - 1}^{\alpha} \quad (8)$$

$$= 1 - \frac{\theta + (1-tu)^{-1/\alpha} + (1-v)^{-1/\alpha} - 2}{\theta + (1-u)^{-1/\alpha}(1-v)^{-1/\alpha} - 1}^{\alpha}.$$

Cuando α = θ = 1, CGL (u, v) = uv. Sean $u^- = (1-u)$ y $v^- = (1-v)$. Cuando θ = 1,

$$\mathrm{CGL}(u, v) = 1 - (u^{-} \quad {}^{1/\alpha\,1/\alpha\,1/\alpha\,\alpha} + v^{-} - u^{-}v^{-}, \quad (9)$$

que es la cópula Joe o B5. La cópula de Joe está dada por C(u, v) = 1 − (1 − u) r + (1 − v) r − (1 − u) r (1 − v) r 0 < α ≤ 1, lo que respalda la proposición 1. Teorema 2 y Sección 4.2. La cópula de Joe, en (9) está ordenada negativamente por α. Cuando α = 1, CGL (u, v) = uv/[1 − (1 − θ)u^-v^-], que es la cópula Ali–Mikhail–Haq (AMH) [11] y está ordenada negativamente por el parámetro θ. La familia de cópulas AMH tiene coeficientes de dependencia de cola inferior y superior de cero, y está dada por C(u, v) = uv/[1 − r(1 − u) (1 − v)],r [−1, 1].

Para la cópula de independencia QUL, de la Tabla 1,

$$\mathrm{CQL}(u, v) = 1 - \theta + \frac{1}{[1 - \mathrm{CQ}_l^{-1}(u), Q_l^{-1}(u)]^{-1/\alpha} - 1}, \text{ donde} \quad (10)$$

$$\mathrm{CQ}_l^{-1}(u), Q_l^{-1}(tu) = 1 - \frac{1-tu}{1-u+\theta u}\,\alpha 1 - \frac{1-v}{1-v+\theta v}^{\alpha}.$$

Cuando θ = 1, (10) produce la cópula Joe en (9) con una parametrización diferente a la el de [12]. Cuando α = 1, (10) produce

$$1 - \frac{u^-v^- + \theta uv^- + \theta uv^-}{\theta\, u^-v^- + \theta uv^- + \theta uv^- + \theta uv}.$$

Machine Translated by Google

Para la cópula UIP-independencia, de la Tabla 1,

$$\text{CGP}(u, v) = \frac{1}{1 + \theta^{-1}\,\text{CG}_{\text{PAG}}^{-1}(u), \text{GRAMO}_{\text{PAG}}^{-1}(tú)^{-1} - 1}^{\alpha} \tag{11}$$

$$= 1 + (u^{-1/\alpha} - 1) + (v^{-1/a} - 1) + \theta(u^{-1/\alpha} - 1)(v^{-1/\alpha} - 1)^{-\alpha}$$

$$= uv\,1 - (1 - \theta)\,(1 - u^{1/\alpha})(1 - v^{1/a})^{-\alpha}, \tag{12}$$

que es la cópula BB10 [12] con κL = κU = 2. La cópula BB10 está dada por C(u, v) = 1/r2)(1 − v 1/ − u donde r1 [0, 1]) y r2 ∞) [1, ∞). Como se muestra en la Proposición 3, las cópulas distorsionadas por UIP tienen el mismo orden de cola superior e inferior que la cópula base.

Para la cópula QUP-independencia, de la Tabla 1,

$$\text{PCC}(u, v) = \frac{1}{1 + \theta\text{CG}_{\text{PAG}}^{-1}(u), \text{GRAMO}_{\text{PAG}}^{-1}(u)^{-1/\alpha} - 1} \tag{13}$$

$$\frac{(1 - u)(1 - v) + \theta u(1 - v) + \theta v(1 - u) = 1 + \theta uv}{}^{-1}$$

$$= uv\,1 - (1 - u)(1 - v)\frac{\theta - 1}{\theta}^{-1},$$

que es la cópula AMH o un caso especial de la cópula BB10.

4. Coeficientes de dependencia de cola y órdenes de cola

En esta sección, investigamos los coeficientes de dependencia de cola, órdenes de cola y concordancias para las nuevas familias de cópulas que surgen de las cuatro distorsiones en la Tabla 1. Las largas derivaciones de órdenes de cola están estacionadas en Apéndice B.

Sea t(u) = dT(u)/du y supongamos que existe el coeficiente de dependencia de cola inferior (ltd) λL de la cópula base C y limτ→0+ t C(u, u) /t(u) . Por definición en (A3) y la regla de L'Hopital , el coeficiente ltd para una cópula inducida por distorsión T viene dado por

$$\lambda_{T,L} = \lim_{u\to 0+} \frac{T\,C(T^{-1}(u), t^{-1}(tú))}{tu}$$

$$= \lim_{u\to 0+} \frac{T\,C(tú, tú)}{T(u)} = \lim_{u\to 0+} \frac{t\,C(u, u)}{t(u)} \frac{dC(u, u)}{du}. \tag{14}$$

Dado que limu→1− T(u) = 1, con la sustitución de v = T el $^{-1}(u)$, la cola superior depende coeficiente de dencia (utd) de CT está dado por

$$\lambda_{T,U} = 2 - \lim_{u\to 1-} \frac{1 - T\,C(T^{-1}(u), t^{-1}(tú))}{1 - tu}$$

$$= 2 - \lim_{v\to 1-} \frac{1 - T(C(v, v)) = 2}{1 - T(v)} - \lim_{v\to 1-} \frac{t\,C(v, v)\,dC(v, v)\,t(v)\,dv}{}. \tag{15}$$

A continuación, asumimos que el coeficiente ltd λL = 0 cuando κL > 1 y el coeficiente utd λU = 0 cuando κU > 1 para la cópula base C. Además, asumimos que C(u, u) u κL ℓ(u) como u → 0 + y C(1 − u, 1 − u) u κU ℓ (u) como u → 0 + para algunas funciones que varían lentamente ℓ y ℓ en 0+. Sea el subíndice el propietario de una propiedad, por ejemplo, el subíndice T en λT,U se utiliza para denotar el coeficiente ltd de una cópula T−distorsionada.

Proposición 1 (Distorsión Unitaria-Lomax). Sea CGL (u, v) la cópula GL distorsionada definida en (8), donde 0 < θ ≤ 1 y 0 < α ≤ 1. Entonces,

(i) κGL,L = κL y λGL,L = λL. (ii)
κGL,U = κU cuando α = 1 y κGL,U = 1 cuando 0 < α < 1. Y λGL,U = 2 − (2 − λU) α .

Machine Translated by Google

Prueba. Los órdenes de cola se muestran en (A9) y (A17) en el Apéndice B. El coeficiente ltd, por (14) y la regla de L'Hopital, viene dado por

$$\lambda_{GL,L} = \lim_{u\to 0+} \left(1- \frac{1-C(tu,tu)}{1-C(u,u)+\theta C(u,u)}\right)^{\alpha} \left(1- \frac{1-tu}{1-u+\theta u}\right)^{\alpha}$$

$$= \lim_{u\to 0+} \frac{[1-C(u,u)]\,[(1-u)+\theta u]}{[1-C(u,u)+\theta C(u,u)]\,(1-u)\,(u,} \left(\frac{1-u}{1-u+\theta u}\right)^{\alpha-1}$$

$$\times \frac{do'u)\,[(1-u)+\theta u]}{[1-C(u,u)+\theta C(u,u)]2} \frac{}{1}^{2} = \lambda_L$$

ya que limu→0+ C(u, u)/u = limu→0+ dC(u, u)/du = λL. El coeficiente utd, por (15) y la regla de L'Hopital, obtenemos que

$$\lambda_{GL,U} = 2 - \lim_{u\to 1-} \left(\frac{1-C(tu,tu)}{1-C(u,u)+\theta C(u,u)}\right)^{\alpha} \left(\frac{1-tu}{1-u+\theta u}\right)^{\alpha}$$

$$= 2 - \lim u\to 1- \frac{1-u+\theta u}{[1-C(u,u)+\theta C(u,u)]\alpha} \left(\frac{1-C(u,u)}{1-u}\right)^{\alpha} = 2-(2-\lambda_U)\,\alpha$$

ya que limu→1−[1 − C(u, u)]/(1 − u) = 2 − λU. □

Proposición 2 (Distorsión Unidad Cuantil-Lomax). Sea CQL (u, v) la cópula QL distorsionada definida en (10), donde θ ≥ 1 y α ≥ 1. Entonces, (i) κQL,L = κL y λQL,L = λL. (ii) κQL,U = κU cuando α = 1 y κQL,U = 1 cuando α > 1. Y λQL,U = 2 − (2 − λU) 1/α .

Prueba. Los órdenes de cola se muestran en (A10) y (A18) en el Apéndice B. Según la regla de L'Hopital,

$$\lambda_{QL,L} = \lim_{u\to 0+} \frac{1-C(u,u)^{-1/\alpha}-1}{1-C(tu,tu)^{\alpha}-1+\theta} \frac{(1-u)^{-1/\alpha}(1-u)^{-1/\alpha}}{-1+\theta}$$

$$= \lim_{u\to 0+} \frac{[(1-u)-1+\theta]^{-1/\alpha}}{[\,1-C(u,u)-1+\theta]}^{2} \frac{1-C(u,u)^{-1/\alpha-1}}{(1-u)^{-1/\alpha-1}} do'(tu, tu) = \lambda_L$$

ya que limu→0+ C(u, u)/u = limu→0+ dC(u, u)/du = λL. Por (15) y la regla de L'Hopital, el coeficiente ltd está dado por

$$\lambda_{QL,U} = 2 - \lim_{u\to 1-} \frac{\theta}{1-C(tu,tu)^{-1/\alpha}-1+\theta} \frac{\theta}{(1-u)^{-1/\alpha}-1+\theta}$$

$$= 2 - \lim u\to 1- \frac{1-C(tu,tu)}{1-tu}^{1/\alpha+1} \frac{dC(u,u)}{du}^{-1} = 2-(2-\lambda_U)^{1/\alpha}$$

ya que limu→1−[1 − C(u, u)]/[1 − u] = 2 − λU. □

Proposición 3 (Distorsión de Pareto inversa unitaria). Sea CGP (u, v) la cópula distorsionada por GP definida en (11), donde 0 < θ ≤ 1 y α ≥ 1. Entonces, (i) κGP,L = κL y λGP,L = (λL) a (ii) κGP,U = κU y λGP,U = λU.

Machine Translated by Google

Prueba. Los pedidos de cola se muestran en (A11) y (A19) en el Apéndice B. Para ltd, con la ayuda de la regla de L'Hopital, obtenemos que

$$\lim_{u\to 0+} \{u/[1-u+\theta u]\}^{\alpha} \{C(u,u)/[1-C(u,u)+\theta C(u,u)]\}^{\alpha} \lambda_{GP,L} = \lim_{u\to 0+} \left(\frac{C(tú,tú)}{tu}\right)^{\alpha} \left(\frac{1-C(u,u)+\theta C(u,u)}{\theta u}\right)^{-\alpha} (1-u+)^{-\alpha} = (\lambda_L)^{\alpha}.$$

El coeficiente utd, por (15) y la regla de L'Hopital, viene dado por

$$\lim_{u\to 1-} 1-\{u/[1-u+\theta u]\}^{-\alpha} \frac{1-\{C(u,u)/[1-C(u,u)+\theta C(u,u)]\}^{\alpha}\lambda_{GP,U}}{(\alpha-1)\,1-C(u,u)+\theta C(u,u)} = 2-\lim_{u\to 1-} \left(\frac{C(tú,tú)}{tu}\right)^{-\alpha-1} \left(\frac{1-C(u,u)+\theta C(u,u)}{1-u+\theta u}\right)^{-\alpha-1} \frac{dC(u,u)}{du} = \lambda_U$$

ya que $\lim_{u\to 1-} C(u,u) = 1$ y $\lim_{u\to 1-} dC(u,u)/du = 2-\lambda_U$. □

Proposición 4 (Distorsión de Pareto inversa unitaria cuantil). Sea $C_{QP}(u,v)$ la cópula QP distorsionada definida en (13), donde $\theta \geq 1$ y $0 < \alpha \leq 1$. Entonces, (i) $\kappa_{QP,L} = \kappa_L$ y $\lambda_{QP,L} = (\lambda_L)^{1/\alpha}$. (ii) $\kappa_{QP,U} = \kappa_U$. Si $\kappa_U = 1$, entonces $\lambda_{QP,U} = \lambda_U$.

Prueba. Los pedidos finales se derivan en (A13) y (A20) en el Apéndice B. Para el coeficiente ltd,

$$= \lim_{u\to 0+} 1/[1+\theta(u-1/\alpha-1)]^{-1/\alpha} \frac{1/\,1+\theta\,C(u,u)\,\lambda_{QP,L}^{-1/\alpha}-1}{} = \lim_{u\to 0+} \frac{1+\theta(u-1)^{-1/\alpha}+\theta\,C(u,u)^{-1/\alpha}-1}{} = \lim_{u\to 0+} \left(\frac{C(tú,tú)}{tu}\right)^{1+1/\alpha} \frac{du}{} = (\lambda_L)^{1/\alpha}\,dC(u,u).$$

El coeficiente utd, por (15) y la regla de L'Hopital, viene dado por

$$\lambda_{QP,U} = 2-\lim_{u\to 1-} \frac{\theta[C(u,u)-1]^{-1/\alpha}\,1+\theta[C(u,u)}{-1/\alpha-1]} \frac{\theta(u-1)^{-1/\alpha}\,1+\theta(u-1/}{\alpha-1)}$$

$$= 2-\lim_{u\to 1-} \frac{C(u,u)-1\,1+\theta[C(u,}{u)^{-1/\alpha}-1]\,(u^{-1/\alpha}-1)} \frac{(u-1)^{-1/\alpha-1/\alpha}\,1+\theta(u-1)}{}$$

$$= 2-\lim_{u\to 1-} \frac{[C(u,u)]^{-1/\alpha-1}\,dC(u,u) = \lambda_U}{u^{-1/\alpha-1}\,du} \frac{1+\theta(u-1)^{-1/\alpha}\,1+\theta[C(u,}{u)^{-1/\alpha}-1]}$$

ya que $\lim_{u\to 1-} C(u,u)/u = 1$ y $\lim_{u\to 1-} dC(u,u)/du = 2-\lambda_U$. □

Los resultados de las proposiciones se resumen en la Tabla 2. Los coeficientes utd de las cópulas distorsionadas UL y QUL difieren de los de las cópulas base cuando $\alpha = 1$, mientras que los coeficientes ltd permanecen sin cambios. Las distorsiones UL y QUL generan nuevas cópulas con dependencia de la cola superior independientemente de si la cópula base la tiene o no. Por el contrario, las distorsiones UIP y QUP forman cópulas con diferentes coeficientes ltd cuando $\alpha = 1$ de la cópula base, mientras que los coeficientes utd permanecen iguales. La cópula de Clayton tiene cero dependencia de la cola superior. Sin embargo, con base en la Tabla 2, al aplicar distorsiones UL y QUL a una cópula de Clayton, creamos una nueva familia de cópulas que son más flexibles en el sentido de que pueden acomodar valores de dependencia de la cola superior que van de 0 a 1. Esta misma conclusión puede aplicarse a la cópula de Frank.

Machine Translated by Google

Tabla 2. Coeficientes de dependencia de la cola de las cópulas distorsionadas.

UL ($0 < \alpha \leq 1$)	QUL ($\alpha \geq 1$)	UIP ($\alpha \geq 1$)	QUP ($0 < \alpha \leq 1$)
$\lambda_{T,L}$ λ_L	λ_L	(λ_L)	$(\lambda_L)^{1/\alpha}$
$\lambda_{T,U}$ $2 - (2 - \lambda_U)^{\alpha}$	$1/\alpha\ 2 - (2 - \lambda_U)$	$\alpha\ \lambda_U$	λ_U

A continuación, construimos los gráficos de contorno de densidad para cópulas distorsionadas por UL y UIP en las Figuras 1 y 2. Los gráficos de contorno y las observaciones para cópulas distorsionadas por QUL y QUP se muestran en las Figuras A1 y A2 en el Apéndice B.1, respectivamente. La función de densidad de probabilidad conjunta de una cópula es $\partial C(u, v)/\partial u \partial v$. Tenga en cuenta que si un gráfico de contorno se alarga en una dirección, indica una fuerte dependencia en esa dirección. Una forma de contorno circular indica independencia entre variables. Por ejemplo, el Clayton tiene dependencia de cola inferior y no dependencia de cola superior, por lo tanto, uno espera ver una línea de contorno alargada o apretada en el lado inferior izquierdo. Sea r el parámetro en la cópula base. Cuando $\theta = \alpha = 1$, tanto la distorsión UL como la UIP nos entregan las cópulas base que se muestran en la primera columna de ambas figuras.

Mientras que la cópula de Frank es independiente de la cola, las cópulas de Clayton y Gumbel se caracterizan por una dependencia asimétrica de la cola solo con utd y ltd, respectivamente. Como se resume en la Tabla 2, las distorsiones de UL pueden producir nuevas familias de cópulas con dependencia de la cola superior. Las cópulas construidas utilizando la distorsión UL exhiben dependencia de la cola superior y mantienen los mismos patrones de comportamiento de la cola inferior que la cópula base. En particular, las cópulas Clayton distorsionadas por UL tienen colas tanto inferior como superior.

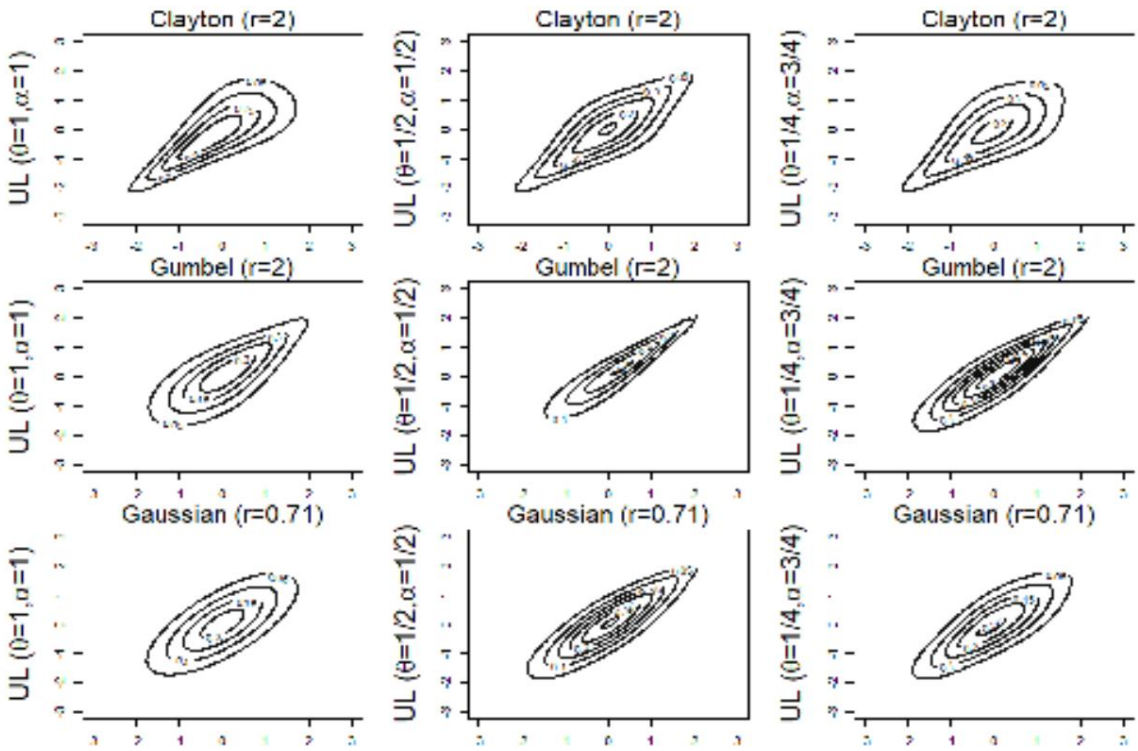


Figura 1. Gráficos de contorno de densidad de las cópulas UL-Clayton, UL-Gumbel y UL-Gaussianas con márgenes normales estándar y parámetros $(\theta, \alpha) = (1, 1), (1/2, 1/2)$, y $(1/4, 3/4)$. Para Clayton, Gumbel y Gauss, el parámetro r se selecciona de modo que el valor de tau de Kendall sea 1/2.

Como se revela en la Tabla 2 y la Figura 2, las cópulas distorsionadas por UIP tienen características similares a las cópulas base en la primera columna. Si la cópula base tiene una dependencia de la cola inferior, la familia introdujo un parámetro adicional al coeficiente ltd pero conserva la dependencia de la cola superior de la cópula base.

Machine Translated by Google

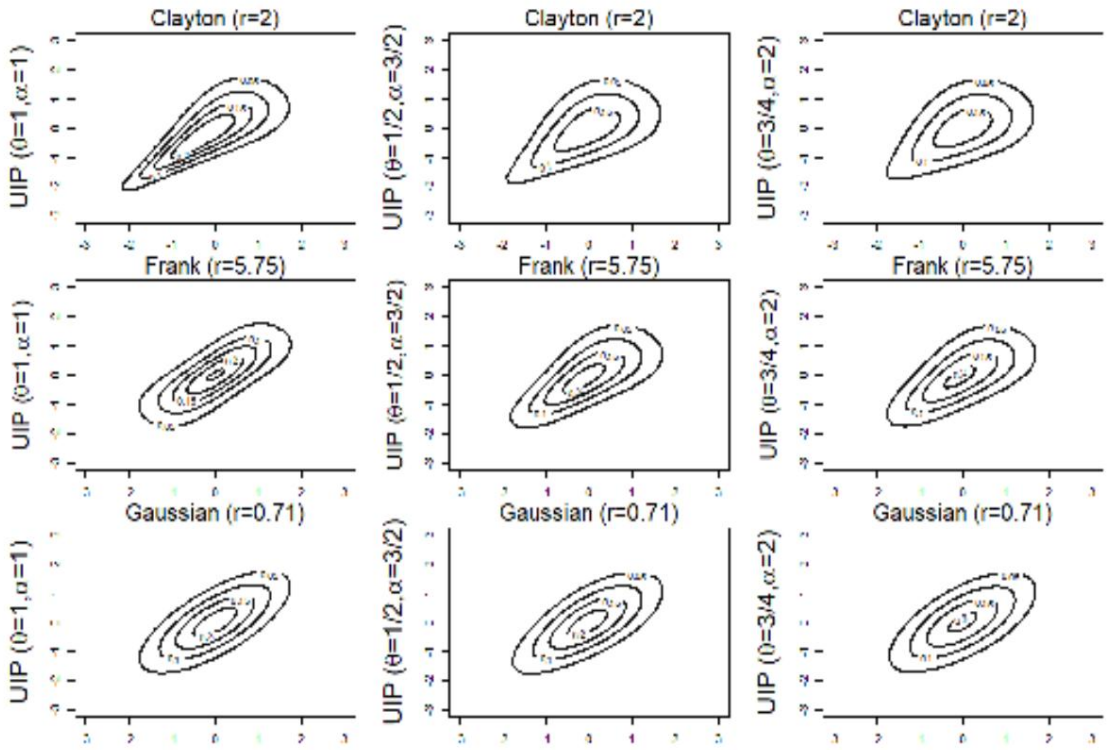


Figura 2. Gráficos de contorno de densidad de las cópulas UIP-Clayton, UIP-Gumbel y UIP-Gaussianas con márgenes normales estándar y parámetros (θ, α) = (1, 1), (1/2, 1/2) y (1/4, 3/4). Para Clayton, Frank y Gaussian, el parámetro r se selecciona de modo que el valor de tau de Kendall sea 1/2.

5. Orden de concordancia y Tau de Kendall

En esta sección examinamos el orden de concordancia y la tau de Kendall. Proporcionamos fórmulas para la tau de Kendall de las cópulas distorsionadas por UL y UIP utilizando la Fórmula (16) y los generadores de Arquímedes proporcionados en la Sección 3. Las fórmulas se pueden adaptar fácilmente para escribir programas para calcular los valores de tau de Kendall en varios valores de parámetros.

5.1. Orden de concordancia

Una familia de funciones cópulas $C_r(u, v) = C(u, v; r)$ con parámetro r está ordenada positivamente, denotada por $C_{r1} \prec C_{r2}$, si $C_{r1}(u, v) \le C_{rr}(u, v)$, y está ordenado negativamente, denotado si $C_{r1}(u, v) \ge C_{rr}(u, v)$ para todo $r1 \le r2$ y $u, v \in I$. Según la definición por $C_{r1} \prec C_{r2}$, de una medida de en el parámetro; ver concordancia, si una familia de las cópulas están ordenadas por un parámetro, su tau de Kendall no es creciente o no decreciente [11].

Se puede calcular la primera derivada con respecto a un parámetro para determinar si la cópula C está ordenada positiva o negativamente por el parámetro, lo que puede ser una tarea desalentadora. Si una cópula es de clase de Arquímedes, además del Teorema A1 del Apéndice C, podemos utilizar el siguiente corolario para examinar los ordenamientos de concordancia en los parámetros; ver [11] o [28].

Teorema 1 ([11]). Sean C1 y C2 cópulas de Arquímedes con generadores ψ_1 y ψ_2, respectivamente. Entonces $C_1 \prec C_2$ se cumple si se cumple una de las siguientes condiciones: (i) $\psi_1 \psi_2^{-1}$ es cóncavo; (ii) ψ_1/ψ_2 $\circ$ ψ no es decreciente en I; y (iii) ψ_1 y ψ_2 son continuamente diferenciables en I y ψ no decrecientes en ψ_1'/ψ_2' es I.

Si la cópula de bases C está ordenada positivamente, entonces, para $r1 \le r2$, $C(T^{-1}(u), t^{-1}(v); r1) \le$ $C(T^{-1}(u), t^{-1}(v); r2)$ para todo $u, v \in I$. Como T es creciente, $T\,C(T^{-1}(u), t^{-1}(v); r1) \le$ $TC(T^{-1}(u), t^{-1}(v); r2)$. Se pueden aplicar argumentos similares a una cópula de base ordenada negativamente. Es decir, una familia de cópulas de distorsión construidas mediante distorsiones admisibles preserva

Machine Translated by Google

el orden de concordancia en el parámetro de la cópula base si la familia de cópulas base está ordenada por el parámetro; ver también [24].

Si la cópula base es de Arquímedes con generador ψ, que está libre de los parámetros θ y α, la nueva familia de cópulas distorsionadas en T se puede escribir como

$$TC\left(T^{-1}(u), t^{-1}(v)\right) = T \circ \psi^{-1}\left(\psi T^{-1}(tu) + \psi T^{-1}(v)\right),$$

con Arquímedes con un generador dado por $\psi T^{-1}(u)$, un continuo, estrictamente decreciente, función convexa tal que ψT con generador ψ. $^{-1}(1) = 0$. Supongamos que debajo de la cópula base está Arquímedes

Sean T1 y T2 la distorsión T evaluada en los valores de los parámetros r1 y r2, donde r1 ≤ r2, respectivamente.

Corolario 1. Sean CT1 y CT2 cópulas de Arquímedes con generadores $\psi \circ T_1^{-1}$ y $\psi \circ T_2^{-1}$, respectivamente. Si se cumple una de las siguientes condiciones: (i) $\psi \circ T_1 \circ \psi^{-1}$ es cóncavo; (ii) $\psi \circ T^{-1}/\psi \circ T^{-1}$ no es decreciente en (0, 1); o (iii) ψ es continuamente diferenciable en I y $(\psi' \circ T_1^{-1}/\psi' \circ T_2^{-1})(T_2' \circ T_2^{-1}/T_1' \circ T_1^{-1})$ no es decreciente en I, entonces CT1 CT2 .

5.2. Tau de Kendall

Para las cópulas inducidas por distorsión T en (1), sustituyendo $T^{-1}(u) = x$ y $T^{-1}(v) = y$ en la definición de tau de Kendall en (A5), entonces su tau de Kendall se puede expresar como

$$\tau = 1 - 4\int_0^1\int_0^1 [t(C(x, y))]2C1|2\,(x|y)C2|1\,(y|x)dxdy,$$

donde $t(v) = dT(v)/dv$, $C2|1\,(u, v) = \partial C(u, v)/\partial u$ y $C1|2\,(u, v) = \partial C(u, v)/\partial v$. Se necesitarán métodos de integración numérica para calcular la tau de Kendall. Defina $\Psi(u) = \psi T$ Por (A6), la tau de Kendall $^{-1}(tú)$. para una cópula distorsionada viene dada por

$$\tau = 1 + 4\int_0^1 \frac{\Psi(tu)}{\Psi'(u)}du = 1 + 4\int_0^1 \frac{\psi(v)}{\psi'(v)}(v)dv. \tag{16}$$

A continuación presentamos las fórmulas explícitas de la tau de Kendall para las distorsiones UL y UIP cuando las cópulas base son cópulas de independencia, Clayton, Gumbel y Frank. Los generadores de Arquímedes para las cópulas base se presentan en la Tabla A1 del Apéndice A.

– Cópula de independencia UL: $1 - 4\int_{10} \frac{\alpha[(1-u)^{-1/\alpha}\log\theta[(1-u)^{-1/\alpha} - 1]^{-1} - 1}{-1/\alpha-1][\theta + (1-u)^{-1/\alpha} - 1]-1\,\alpha\theta u\,(1-u}$ Du.

– Cópula UL-Clayton: $1 - 4\int_{10} \frac{r)\,[1+\theta(-1}{r(1-tu)^{2} + (1-u)-1)]\alpha+1\,\alpha\theta u(-\log u)}$ Du.

– Cópula UL-Gumbel: $1 - 4\int_{10} \frac{[1+\theta(-1+(1}{r(1-tu)^{2} - u)-1)]\alpha+1\,\alpha\theta(1-e\,ru)\log}$ Du.

– Cópula UL-Frank: $1 - 4\int_{10} \frac{[1+\theta(-1+}{r(1-tu)^{2}\,1)]\log 1+\theta(u\,\theta u-1/\alpha-1} \frac{(1-u)-1)]\alpha+1 - \quad \frac{e^{-ru}-1}{e-r-1}}$ Du.

– Cópula de independencia UIP: $1 - 4\int_{10} \frac{\alpha[1+\theta(u^{-1/a} \quad ^{-1/a} - 1)}{}$ Du.

– Cópula UIP-Clayton: $1 - 4\int_{10} \frac{\alpha u(1-u\,r)}{\theta-1(u-1-1)]\alpha+1\,\alpha u(-\log u)\,du.}\,du.\,r\theta u\,2\,[1+$

– Cópula UIP-Gumbel: $1 - 4\int_{10} \frac{r\theta u\,2\,[1+\theta-1(u-1-1)]\alpha+1}{\alpha(1-e\,ru)\log r\theta u\,2\,[1+\theta-1(u-1}$

– Cópula UIP-Frank: $1 - 4\int_{10} \frac{-1)]\alpha+1}{} \quad \frac{e^{-ru}-1}{e\text{-}r-1}$ Du.

Ejemplo 7. Aquí ilustramos solo el orden en el parámetro θ para las distorsiones UL y UIP de la cópula de independencia en el Apéndice C debido al límite de páginas. Si bien no se muestran todos matemáticamente, los gráficos de superficie tau de Kendall en la Figura 3 indican que las distorsiones UL, QUL, UIP y QUP de la

Machine Translated by Google

Las cópulas de independencia dan como resultado nuevas familias que están ordenadas negativamente en los parámetros θ y α. Para la familia de cópulas de independencia UIP en (12), la cópula BB10 está ordenada negativamente por el parámetro θ, ver [12]. La familia de cópulas de independencia QUP en (13) no depende del parámetro α y, por tanto, de las líneas planas a lo largo del eje α .

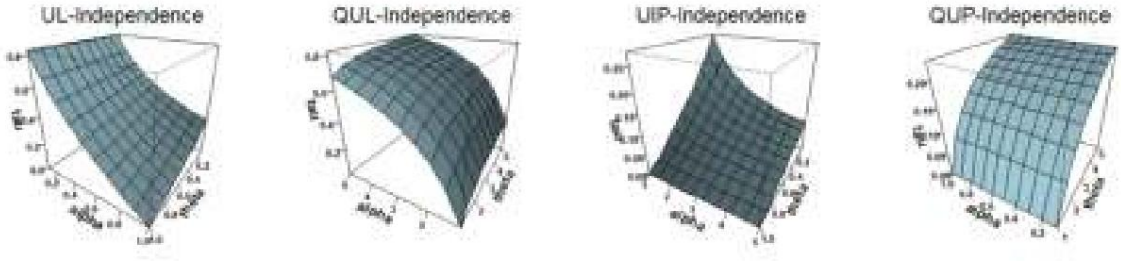


Figura 3. Gráficos de superficie de la tau de Kendall para cópulas de independencia UL, QL, QIP y QUP.

No todas las familias de cópulas generadas por distorsión están ordenadas por concordancia. Por ejemplo, la familia Clayton distorsionada por UL. Cuando una familia de cópulas está ordenada por un parámetro, esperamos que los valores de tau de Kendall aumenten o disminuyan con el parámetro.

La Figura 4 muestra los gráficos de superficie tau de Kendall cuando se aplican distorsiones a la cópula de Clayton con el parámetro r. Estos gráficos indican que las cópulas de distorsión, al igual que la cópula de Clayton base, están ordenadas positivamente por el parámetro r. La cópula UL-Clayton con r = 15 y la cópula QUL-Clayton con r = 20 no están ordenadas por el parámetro α. Respecto a las cópulas UIP-Clayton, el orden de concordancia falla en el parámetro α en r = 25 y en θ = 1/10, así como en el parámetro θ cuando α = 25. Además, dependiendo del valor de α o r , el La cópula QUL-Clayton se puede ordenar positiva o negativamente mediante el parámetro θ cuando r = 2 y α = 3/4.

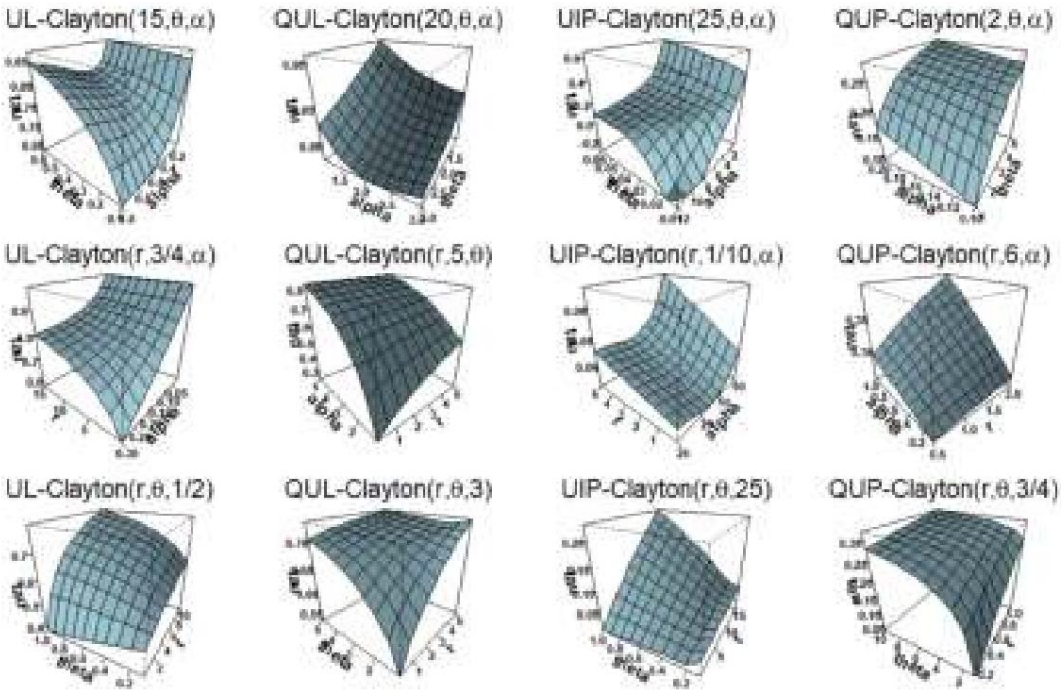


Figura 4. Gráficos de superficie de tau de Kendall construidos en varias combinaciones de valores de parámetros para la cópula UL-Clayton, QL-Clayton, UIP-Clayton y QUP-Clayton.

6. Resultados numéricos

En esta sección, realizamos un estudio de simulación para inspeccionar cómo se comportan las nuevas familias de cópulas de distorsión UL cuando se ajustan a los datos generados por el querido Clayton.

Machine Translated by Google

Cópulas de Gumbel, Gaussiana y Frank, y viceversa. Además, los modelos de cópula se aplican a un conjunto de datos bivariado que consta de las tasas de retorno diarias de Amazon y Acciones de Google.

6.1. Un estudio de simulación

Un algoritmo general para generar sorteos a partir de una cópula bivariada C es el condicional enfoque de distribución, como se describe en [19,24]. Consta de dos pasos: (i) generar dos valores aleatorios uniformes independientes (u1, v) y (ii) resuelven $C(u2|u1) - v = 0$ para u2, donde $C(u|v) = \partial C(u, v)/\partial v$. El par deseado es (u1, u2). Usando este algoritmo, nosotros generó 2000 pseudoobservaciones bivariadas de Clayton, Gumbel, Gaussiano y Cópulas francas. También simulamos el mismo número de pseudoobservaciones del familias de cópulas distorsionadas por UL, donde las cuatro cópulas sirvieron como cópulas base. nosotros no presentar las cifras de distorsiones QIP, QUL y QUP, ya que las conclusiones son similares a los de distorsión UL.

Los valores de los parámetros en las cópulas base se seleccionan para que tengan la tau de Kendall. valor de 0,5. La distorsión UL tiene valores de parámetros de $\alpha = 1/2$ y $\theta = 1/2$. Usamos el método de estimación de pseudoverosimilitud [12] que maximiza (17) para ajustarse a la cópula base y cópulas distorsionadas a los datos. Luego calculamos las probabilidades empíricas usando los modelos de cópula estimados y construyeron las gráficas de probabilidad-probabilidad (PP) de los distribución de probabilidad estimada frente a la teórica. En la primera fila, las cuatro bases. Las cópulas se aproximan a sus contrapartes distorsionadas por UL, UL-Clayton, UL-Gumbel, Cópulas UL-Gaussiana y UL-Frank, y viceversa en la segunda fila. La línea negra sólida es el que resultó del ajuste de los datos al modelo de cópula, a partir del cual las observaciones fueron generados.

Como otra forma de comparación, también calculamos la distancia máxima en la Tabla 3. entre las probabilidades teóricas y empíricas en cada par de datos de (u1, u2). El Me vino a la mente la prueba univariada de Kolmogorov-Smirnov (KS), y para un tamaño de muestra de 2000, la Se utiliza un valor crítico del 95% de 0,03 como umbral ad hoc.

Con base en la Figura 5 y la Tabla 3, la cópula de Clayton, que está representada en rojo en el segunda fila de la Figura 5, muestra mayores desviaciones de la línea de 45 grados cuando se ajusta a la datos generados a partir de cópulas distorsionadas por UL. Según el Lema 1, la cópula distorsionada por UL tiene un coeficiente upd de $2 - (2 - \lambda U)^{\alpha}$, que puede estar en sintonía con cero. Además, tiene límite cero cuando la cópula base tiene límite cero. La cópula de Clayton no tiene parte superior. dependencia de la cola, por lo que uno esperaría que tuviera un mal desempeño cuando se ajusta a los datos generados de la familia distorsionada por UL. Por el contrario, por ejemplo, el UL-Gumbel parece funcionar bien cuando se adapta a los datos generados a partir de las cópulas de Frank y Gauss. En general, el rendimiento de una cópula depende de sus características de dependencia de la cola, y los resultados muestran que la Las cópulas distorsionadas por UL son más flexibles ya que tienen parámetros adicionales.

Tabla 3. Distancia máxima entre las distribuciones de cópula teórica y empírica.

Cópula distorsionada	Aproximando			
	Clayton por	Gumbel por	franco por	gaussiano por
UL-Clayton	0.014	0,021	0,035	0.026
UL-Gumbel	0,028	0,012	0,019	0.008
UL-Frank	0,021	0,012	0,011	0.024
UL-gaussiano	0,031	0,021	0,031	0,016
Cópula base	**Aproximando**			
	UL-Clayton por	UL-Gumbel por	UL-Frank por	UL-Gaussiano por
Clayton	0.035	0,047	0,042	0,042
Gumbel	0.050	0.013	0.034	0.033
Franco	0,046	0.024	0.030	0.032
gaussiano	0.021	0.019	0,017	0.022

Machine Translated by Google

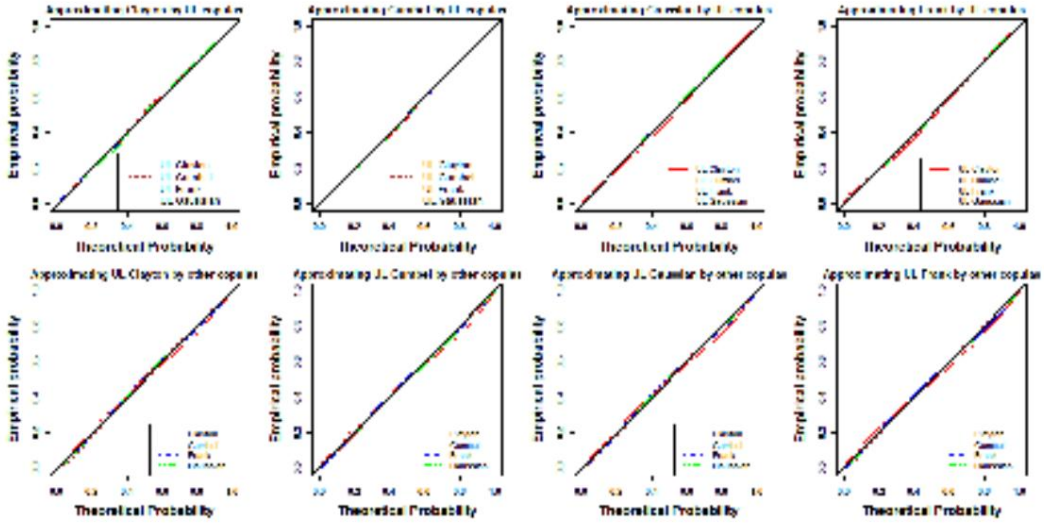


Figura 5. Gráficos PP de UL y cópulas de base.

6.2. Aplicación empírica

Ajustamos las familias propuestas de modelos de cópula a un conjunto de datos bivariados de tasas de retorno diarias en acciones de Amazon y Google. Los datos históricos de los precios diarios de apertura, cierre, máximo y mínimo, y el precio de cierre ajustado de las acciones se pueden descargar de Yahoo Finance. El precio de cierre ajustado tiene en cuenta las divisiones y distribuciones de dividendos. Descargamos los datos de las acciones de Amazon y Google para el período comprendido entre enero de 2014 y diciembre de 2023, lo que equivale a un tamaño de muestra de 2516 puntos de datos diarios. Luego se calcularon las tasas de rendimiento netas diarias en porcentajes en función del precio de cierre ajustado. Para calcular la tasa de retorno de hoy, la diferencia entre el precio de hoy y el precio de ayer se divide por el precio de ayer.

La Tabla 4 muestra las estadísticas resumidas para ambas variables. La correlación muestral de Pearson y la tau de Kendall son 0,71 (valor p < 0,001) y 0,49 (valor p < 0,001), respectivamente, los cuales son significativamente diferentes de cero. En comparación con Google, las medidas de diez dencias centrales para Amazon son más pequeñas, pero no hay una diferencia significativa en las medias. El rendimiento de la tasa diaria de Amazon es significativamente más variable según la prueba F y está más sesgado a juzgar por las medidas de asimetría y los histogramas de la Figura 6.

Tabla 4. Estadísticas descriptivas del conjunto de datos.

	Significar		Mín 1.º Qu Mediana 3.º Qu Máx.					Sesgar
Amazonas	−0,01	1,60	−8,56	−0,86	−0,01	0,91	8.24	−0,09
Google	0,04	1.35	−5,76	−0,67	0,05	0,84	6.65	−0,20

Sea $\{x_i, y_i\}_{i=1}^{n}$ denota las observaciones bivariadas. Las pseudoobservaciones o distribuciones empíricas escaladas $\{u_i, v_i\}_{i=1}^{n}$ se definen como $u_i = \sum_{j=1}^{n} J(x_j \le x_i)/(n+1)$ y $v_i = \sum_{j=1}^{n} J(y_j \le y_i)/(n+1)$, donde $J(\cdot)$ es la función indicadora. La Figura 6 contiene los diagramas de dispersión de y_i versus x_i y v_i versus u_i. Según los histogramas, las tasas de rendimiento de ambas acciones se concentran alrededor de su centro, lo que también se refleja en el diagrama de dispersión resultante. Sin embargo, las pseudoobservaciones calculadas utilizando una distribución empírica escalada están distribuidas uniformemente. Por lo tanto, uno esperaría un diagrama de dispersión más uniformemente disperso.

La estimación de pseudoverosimilitud máxima introducida por [18] se utiliza para estimar los parámetros. Maximiza la función de pseudo-log-verosimilitud, es decir, la log-verosimilitud con las funciones de cópula evaluadas en pseudo-observaciones, dada por

$$L(\gamma) = \log \sum_{i=1}^{n} T(u_i, v_i; \gamma), \tag{17}$$

Machine Translated by Google

donde cT es la función de densidad de probabilidad de la cópula en (A1) y γ es el vector de parámetros.

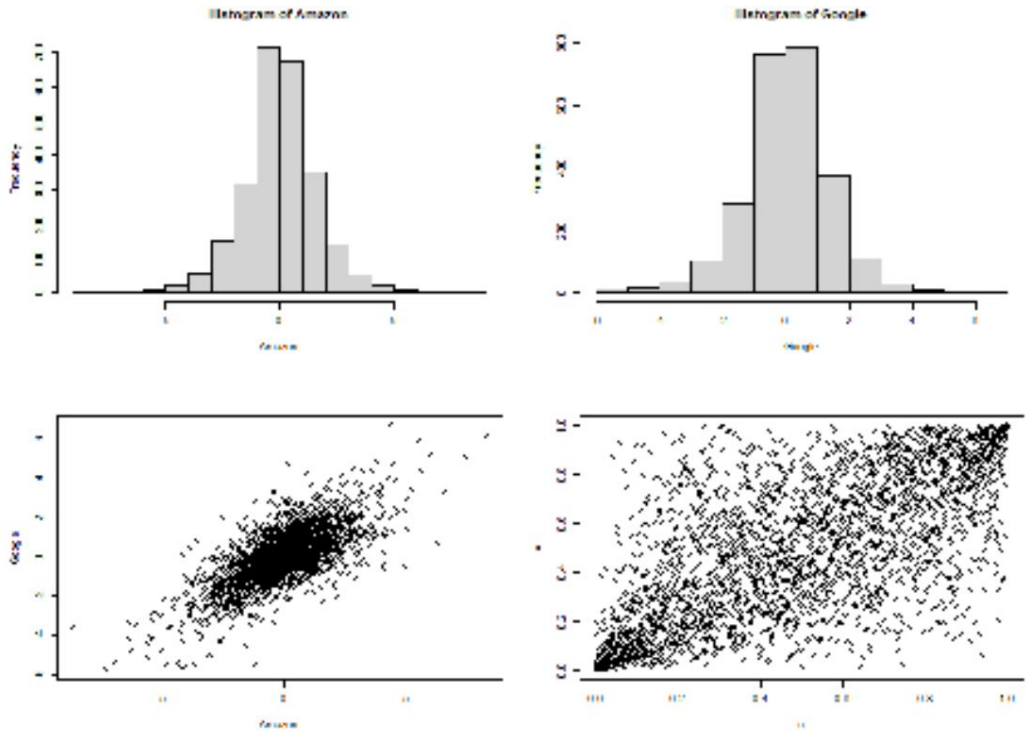


Figura 6. Histogramas de tasas de retorno diarias para las acciones de Amazon y Google, y diagramas de dispersión de Amazon frente a las tasas de retorno de Google y sus correspondientes pseudoobservaciones.

No ajustamos una distribución marginal a cada uno de los rendimientos netos. Nuestro objetivo principal fue comparar las nuevas familias de la cópula distorsionada con las cópulas base. Sea r el parámetro en la cópula base. La Tabla 5 a continuación presenta las estimaciones con el error estándar estimado entre paréntesis, la pseudoverosimilitud máxima (MPL) y los valores AIC. Todas las estimaciones de parámetros se encuentran dentro de los espacios admisibles. Las estimaciones de tau de Kendall τ se calcularon conectando estimaciones de parámetros a la fórmula teórica de tau de Kendall o (16).

Según los gráficos de dispersión, parece que existe una débil dependencia tanto en la cola inferior como en la superior entre los rendimientos netos diarios de las dos acciones. La Tabla 5 muestra que entre las cópulas base consideradas, la cópula gaussiana tiene el mejor rendimiento en términos de MPL y AIC, seguida de Gumbel. La tau de Kendall estimada calculada a partir de la cópula gaussiana produce la coincidencia más cercana a la muestra de tau de Kendall entre Google y Amazon.

La cópula de Frank, que supuestamente es adecuada para datos con una dependencia débil de la cola, tiene el peor rendimiento. Tenga en cuenta que la cópula de Gumbel con un valor de parámetro de 1 representa la cópula independiente donde la tau de Kendall es igual a 0. Funciona mejor que la cópula de Clayton, lo que sugiere que podría haber una upd más fuerte que ltd. Además, una cópula de distorsión que pueda acomodar un rango más amplio de dependencia de la cola superior, por ejemplo, cópulas distorsionadas por UL, puede funcionar bien para ajustar este conjunto de datos de retorno neto.

La Tabla 5 indica que el modelo de cópula distorsionada por UL supera a la cópula base correspondiente. El modelo de cópula UL-Clayton ajustado tiene el valor AIC más grande, con coeficientes de dependencia de cola inferior y superior estimados de 0,44 y 0,50. Es menos satisfactorio que otros modelos de cópula distorsionados por UL, probablemente debido a una menor dependencia de la cola en los datos. Tanto la cópula Frank como la Gaussiana distorsionada por UL tienen un coeficiente de dependencia de la cola superior estimado de 0,39 y funcionan mejor que la cópula Gaussiana en términos de MPL y AIC.

Machine Translated by Google

Las cópulas UIP-Clayton y UIP-Frank no exhiben comportamientos de dependencia de cola y su rendimiento es peor que sus cópulas base. UL-Gumbel y UIP-Gumbel las cópulas son las de mejor rendimiento, siendo UL-Gumbel ligeramente mejor que UIP-Gumbel en términos de MPL y AIC. El modelo UIP-Gumbel produce una tau de Kendall estimada más cerca de la muestra tau de Kendall. Ambos modelos tienen dependencia de cola superior, pero no inferior. dependencia de la cola de la cola.

Según los modelos de cópula empleados en esta aplicación, existe una moderada correlación lineal entre los rendimientos netos diarios de las acciones de Amazon y Google. Además, parece asimétrico en los comovimientos extremos; es decir, los extremos conjuntos son más más probable para valores de retorno neto diario altos que para valores de retorno neto diario bajos.

Tabla 5. MPL, AIC, τ, estimaciones de parámetros y su desviación estándar entre paréntesis r, θ, α para el modelos de cópula base, distorsionada por UL y distorsionada por UIP.

Familia	MPL	τ	AIC	r	θ	α
Clayton	756.0	0,488	−1510,1	1,907(0,055)	–	–
Gumbel	758,7	0.461	−1515,5	1,856(0,033)	–	–
Franco	754.3	0,491	−1506,7	5,582(0,165)	–	–
gaussiano	838,6	0,493	−1675,3	0,699(0,009)	–	–
UL-Clayton	877,3	0,450	−1748,6	4,855(0,192)	0,000(0,174)	0,562(0,029)
UL-Gumbel	906,4	0,461	−1806,9	1,256(0,043)	0,008(0,004)	0,873(0,056)
UL-Frank	900,0	0,476	−1794,0	4,223(0,319)	0,084(0,023)	0,640(0,019)
UL-gaussiano	901,9	0,484	−1797,9	0,527(0,034)	0,028(0,012)	0,686(0,026)
UIP-Clayton	748,3	0,353	−1490,6	1,413(0,097)	0,999(0,030)	1,296(0,108)
UIP-Gumbel	906,4	0,494	−1806,8	1,392(0,036)	0,010(0,006)	1,174(0,097)
UIP-Frank	748,5	0,481	−1494,9	9,281(0,443)	0,999(0,060)	3,484(0,043)
UIP-Gaussiano	887,1	0,472	−1768,3	0,852(0,073)	0,022(0,062)	1,995(0,194)

7. Observaciones finales

El marco propuesto en el artículo se origina en el hecho de que una acumulación La función de distribución con soporte de intervalo unitario es una función de distorsión. Emplea un transformación de una variable aleatoria no negativa en una variable con el apoyo del intervalo unitario. Los parámetros adicionales en la distorsión permiten una mayor flexibilidad de modelado. Como se demuestra en la Sección 3.2, la distorsión de la cópula de independencia crea una nueva familia de cópulas que incluye la cópula base y otras cópulas existentes como sus miembros y Se adapta a una gama más amplia de comportamientos de dependencia de la cola que la cópula de independencia. Nunca soñaría con tener.

El comportamiento de la cola de un modelo de cópula es un factor crucial para determinar si puede ajustarse adecuadamente a los datos. El uso de distorsiones UL y QUL puede transformar una familia de cópulas de base sin dependencia de la cola superior en una nueva familia de cópulas con cola superior dependencia. El coeficiente de dependencia de la cola superior de las cópulas distorsionadas por UL y QUL involucra más parámetros que el de la cópula base. En última instancia, las distorsiones pueden conducir a una mejor acomodación de la dependencia de la cola superior en comparación con la base cópula. Los comportamientos de la cola en las familias de cópulas distorsionadas UIP y QUP son similares a los de la cópula base. Sin embargo, pueden acomodar mejor la cola inferior. dependencia en comparación con las cópulas base.

No estamos seguros de si una CDF o distorsión más complicada, por ejemplo, una con más de dos parámetros, daría como resultado una nueva familia de cópulas con ambas partes superiores y menor dependencia de la cola cuando se aplica a una cópula base sin comportamiento de dependencia de la cola. El marco propuesto en este artículo abre la puerta a un mundo de nuevas distorsiones. Pendiente hasta el límite de longitud de la página, una mayor exploración del orden de concordancia de la nueva familia Se analizará con más detalle el problema de las cópulas distorsionadas. Las distorsiones de las cópulas multivariadas. de dimensiones superiores también pueden ser de interés. A diferencia de las distorsiones de las cópulas bivariadas, las distorsiones de las cópulas multivariadas requieren más atención y serán exploradas en el futuro.

Machine Translated by Google

Contribuciones de los autores: Todos los autores contribuyeron por igual. Todos los autores han leído y aceptado la versión publicada del manuscrito.

Financiamiento: Esta investigación no recibió financiamiento externo.

Declaración de la Junta de Revisión Institucional: El estudio no requirió aprobación ética.

Declaración de Consentimiento Informado: No aplicable.

Declaración de disponibilidad de datos: Los datos están incluidos en el manuscrito.

Conflictos de intereses: Los autores declaran no tener conflictos de intereses.

Apéndice A. Preliminares

Sea F(x, y) la distribución de probabilidad acumulada conjunta (cdf) de variables aleatorias continuas (X,Y), y FX(x) y FY(y) sean las respectivas cdf marginales de X e Y. Según el teorema de Sklar [10], existe una cópula única C, que satisface las condiciones de contorno y monotonicidad , tal que

$$F(x, y) = P(X \le x, Y \le y) = C\,FX(x), FY(y)\,, \quad x, y \quad (-\infty, \infty). \tag{A1}$$

La función de densidad de probabilidad conjunta (pdf), denotada por f(x, y), de (X,Y) es

$$f(x, y) = c\,FX(x), FY(y)\,fX(x)fY(y), \tag{A2}$$

donde fX y fY son las respectivas funciones de probabilidad de X, e Y y c(u, v) es la función de probabilidad de cópula tal que $c(u, v) = \partial\, 2C(u, v)/\partial u\partial v$. La ecuación (A2) implica que una distribución de probabilidad bivariada conjunta se puede separar en marginales univariadas y una estructura de dependencia, donde la estructura de dependencia está representada por una cópula.

Una cópula C por definición tiene las siguientes propiedades: (i) C(u, 0) = C(0, v) = 0, (u, v) I ^{2} donde yo = [0, 1]; (ii) C(u, 1) = u y C(1, v) = v, (u, v) I ^{2}; y (iii) C(u2, v2) − C(u2, v1) − C(u1, v2) + C(u1, v1) ≥ 0, para u1 ≤ u2, v1 ≤ v2, y (u1, u2), (v1, v2) en I^2.

Si C es una cópula de Arquímedes con un generador estricto ψ(·) tal que ψ(0) = ∞, se puede expresar como C(u, v) = ψ es la inversa de ψ(u) + ψ(v), donde ψ^{-1}(t) : [0, 1] → [1, ∞) es convexa, continua y estrictamente decreciente con ψ(1) = 0 y ψ(0) = ∞. Las cópulas de Arquímedes son populares porque admiten fórmulas explícitas y pueden acomodar dimensiones superiores con un solo parámetro. La Tabla A1 destaca algunas cópulas bivariadas de Arquímedes destacadas, sus generadores y sus comportamientos de dependencia de la cola. Sea Π la cópula de independencia.

Tabla A1. Importantes cópulas de Arquímedes y sus generadores.

Nombre	Función de cópula	Generador ψ(t) κL o λL	κU o λU
Π	ultravioleta	− 1)/r kL = 2	kU = 2
Clayton	(tu −r −r + v	− 1) ,r > 0 (e −ru − 1) −1/r λL = 2 −1/r	kU = 2
Franco	$\frac{-1}{r}$ r = 0	$\frac{(e^{-rv} - 1)\log 1 + e^{-r} - 1}{}$, − iniciar sesión $\frac{e^{-rt} - 1}{e\text{-}r - 1}$	kL = 2 kU = 2
Gumbel	mi −((− log(u))r+(− log(v))r) 1/l r > 1	− Iniciar sesión(t) r kL = 2 1/r	λU = 2 − 2 1/r

Los parámetros de dependencia de la cola inferior y superior, λL y λU, están dados por

$$\lambda L = \lim u\to 0+ \; P(FX(X) < u \mid FY(Y) < u) = \lim_{u\to 0+} \frac{C(tú, tú)}{tu}, \tag{A3}$$

$$\lambda U = \lim u\to 0+ \; P(FX(X) > u \mid FY(Y) > u) = \lim_{u\to 1-} \frac{\overline{C}(tu,}{tu)\,(1 - tu)} = 2 - \lim u\to 1- \frac{1 - C(tu, tu)}{1 - tu},$$

Machine Translated by Google

donde $\bar{C}(u, v) = P(U > u, V > v) = 1 - u - v + C(u, v)$. Miden la probabilidad de que una variable aleatoria alcance valores extremos dado que otra variable alcanza extremos.

Se dice que una función no negativa f varía lentamente si fors > 0, limx→0 f(sx)/f(x) = 1. Sean f1 y f2 dos funciones. Si limu→u0 f1(u)/f2(u) = 1, lo denotamos por f1(u) f2(u) como u → u0. Para una cópula bivariada si

$$C(tu, tu) \quad tu \ \kappa_L \ \ell(u), tu \to 0+,$$

donde ℓ(u) varía lentamente en 0+, entonces κL se conoce como el orden de cola inferior de C. La cópula de supervivencia viene dada por

$$C(tu, v) = tu + v - 1 + C(1 - tu, 1 - v). \tag{A4}$$

El orden de cola superior κU cumple la condición de que

$$C(u, u) \quad u \ \kappa_U \ \ell \quad (u), u \to 0 \quad +$$

para alguna función que varía lentamente ℓ (u). Tenga en cuenta que 1 ≤ κL, κU ≤ 2. Si κL > 1 y κU > 1, entonces λL = λU = 0. Si κL = 1, λL = limu→0+ ℓ(u); y si κU = 1, λU = limu→0+ ℓ (u). Cuando κL = 2 y ℓ(u) → q como u → 0 +, para algún q positivo, las variables son casi independientes en la cola inferior. Si 1 < κL < 2, las variables están asociadas positivamente y tienen una dependencia de cola intermedia. Se pueden sacar conclusiones similares para la dependencia de la cola superior; ver [12,29] para más detalles. El valor tau de Kendall de una cópula bivariada se puede expresar como

$$\tau = 1 - 4 \int_0^1 \int_0^1 \frac{\partial C}{\partial u} \frac{\partial C(u, v)}{\partial v} (u, v) \, du \, dv. \tag{A5}$$

Para una cópula de Arquímedes, la tau de Kendall también se puede calcular mediante

$$\tau = 1 + 4 \int_0^1 \frac{\psi(u)}{(tu)} du. \ \psi' \tag{A6}$$

Apéndice B. Gráficos de contorno y derivaciones de órdenes de cola Apéndice

B.1. Gráficos de contorno La

Figura A1 muestra que las colas de dependencia de la cola superior de QUL-Clayton, QUL-Gumbel y QUL-Gaussian parecen más pronunciadas que la cópula base correspondiente para los valores α seleccionados , similares a los de UL-Clayton, UL-Gumbel, y UL-Gaussiano, respectivamente. Las cópulas distorsionadas por UL y QUL no tienen una dependencia de cola inferior cuando las cópulas base no tienen una dependencia de cola inferior. Si bien los parámetros se seleccionan de modo que τ = 0,5, los gráficos de contorno en las Figuras 1 y A1 muestran varias formas y escasez.

La Tabla 2 indica que el coeficiente de dependencia de la cola superior de las cópulas distorsionadas UIP y QUP es el mismo que el de la cópula base. Sin embargo, los coeficientes de dependencia de cola más bajos son (λL) α donde α ≥ 1 y (λL) α donde 0 < α ≤ 1, respectivamente. Es decir, la distorsión UIP con α = 3/2 en la Figura 2 y la distorsión QUP con α = 1/2 en la Figura A2 tienen formas similares, si no iguales, en el lado superior derecho. La asimetría está presente en QUP-Frank y UIP-Frank cuando θ y α no son iguales a 1. Los gráficos de contorno para UIP-Gaussian y QUP-Gaussian muestran pocos cambios.

Machine Translated by Google

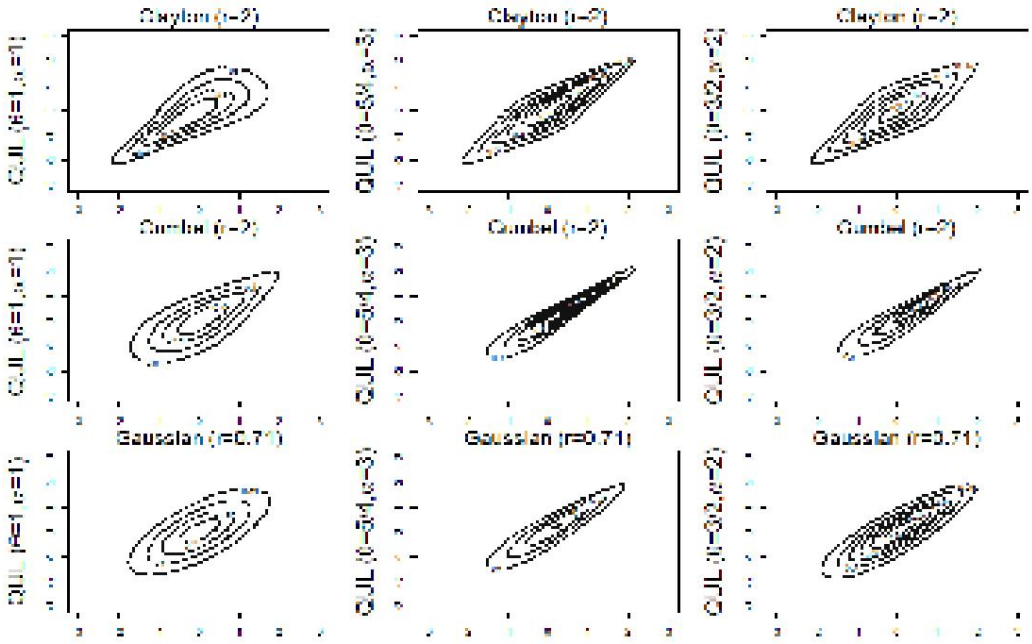


Figura A1. Gráficos de contorno de densidad de las cópulas QUL-Clayton, QUL-Gumbel y QUL-Gaussianas con márgenes normales estándar y parámetros (θ, α) = (1, 1), (3/2, 1/2) y (2, 3/4). Para Clayton, Gumbel y Gauss, el parámetro r se selecciona de modo que el valor de tau de Kendall sea 1/2.

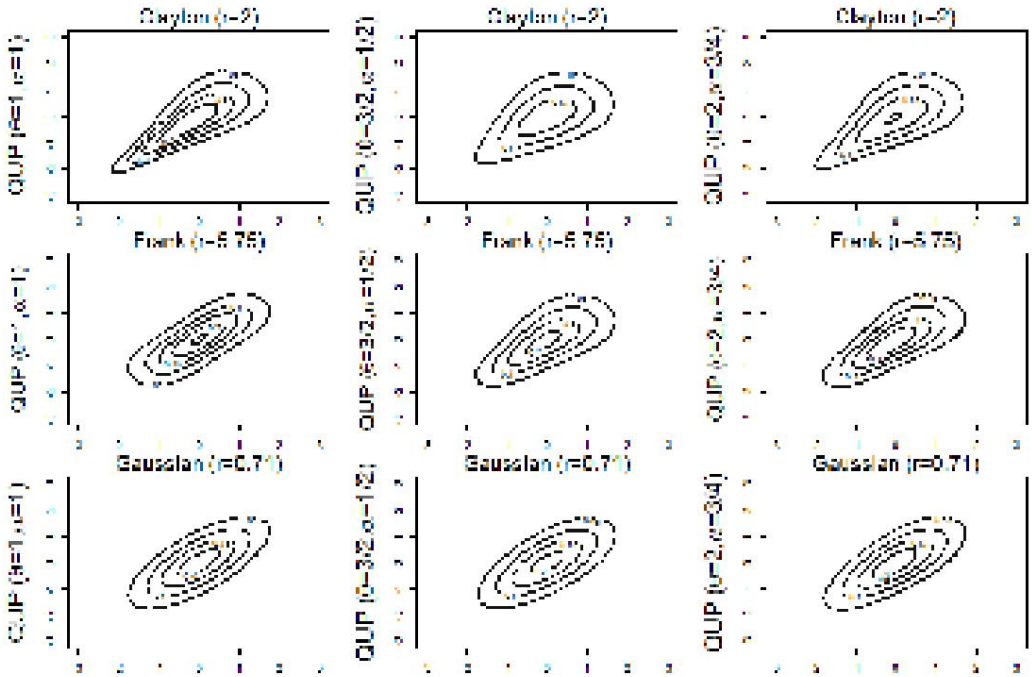


Figura A2. Gráficos de contorno de densidad de las cópulas QUP-Clayton, QUP-Frank y QUP-Gaussianas con márgenes y parámetros normales estándar (θ, α) = (1, 1), (3/2, 1/2) y (2, 3/4). Para Clayton, Frank y Gaussian, el parámetro r se selecciona de modo que el valor de tau de Kendall sea 1/2.

Apéndice B.2. Derivaciones de pedidos de cola

Las siguientes aproximaciones en serie de Talyor son herramientas clave para calcular la órdenes de cola de las distorsiones propuestas. Para dos constantes a y b, tenemos que

$$(1+u)^a \quad 1+au, \quad 1-au,\ 1+au \quad 1+au \quad \frac{1}{1+au} \quad tu, \quad \frac{tu}{b+au} \quad \frac{tu}{b} \text{ como } tu \to 0. \tag{A7}$$

Machine Translated by Google

Supongamos que C(u, u) u κL ℓ(u) como u → 0 + y C(1 − u, 1 − u) u κU ℓ (u) como u → 0 + para algunas funciones que varían lentamente ℓ y ℓ en 0+.

Apéndice B.3. Órdenes de cola inferior

Tenga en cuenta que $G_{|}^{-1}$ = QL y G $_{PAG}^{-1}$ = QP. De (A7) y la Tabla 1, obtenemos que, como u → 0 +,

$$GL(u) = 1 - 1 - 1 - (1 - \frac{\theta u}{\theta)u^{-1/\alpha} - 1}^{\alpha} \quad 1 - (1 - \theta u)^{\alpha} \quad \alpha\theta u;$$

$$QL(u) = \frac{(1 - tu)}{\theta + (1 - u)^{-1/\alpha} - 1} ; \quad \frac{u/\alpha}{\alpha - 1} \quad u/(\alpha\theta); \theta + (1 - u)^{-1/}$$

GP(u) θ αu α QP(u) u 1/α/θ. (A8)

Para cópulas distorsionadas por UL, desde $CG_{|}^{-1}{}_{(u),\,GRAMO|}^{-1}(u) \to 0$ como u → 0 +, aplicando (A8) rendimientos

$$GL\ CG_{|}^{-1}{}_{(u),\,GRAMO|}^{-1}(tu) = 1 - \frac{1 - CG_{|}^{-1}{}_{(u),\,GRAMO|}^{-1}(tú)}{1 - GC_{|}^{-1}{}_{(u),\,GRAMO|}^{-1}(u) + \theta C\,G_{|}^{-1}{}_{(u),\,GRAMO|}^{-1}(tú)}^{\alpha}$$

$$\alpha\theta C\,G_{|}^{-1}{}_{(u),\,GRAMO|}^{-1}(u) \quad \alpha\theta[QL(u)]\kappa L\,\ell\,QL(u)\,\kappa L$$

tú (αθ) 1−κL ℓ QL(u) cuando u → 0 +. (A9)

Por (A8), ℓ QL(u) ℓ u/(αθ) y por definición, (αθ) 1−κL ℓ QL(u) varía lentamente.

Para las cópulas distorsionadas por QUL, QL, según la Tabla 1, (A7) y (A8), tenemos

$$QL\ CQ_{|}^{-1}(u), Q_{|}^{-1}(tu) = \frac{1 - CQ_{|}^{-1}(u), Q_{|}^{-1}(tú)^{-1/\alpha} - 1}{\theta + 1 - CQ_{|}^{-1}(u), Q_{|}^{-1}(tú)^{-1/\alpha} - 1}$$

$$(\alpha\theta)^{-1}CQ_{|}^{-1}(u), Q_{|}^{-1}(u) \quad (\alpha\theta)^{-1}[GL(u)]\,\kappa L\,\ell Q_{|}^{-1}(tú)$$

tú κL (αθ) κL−1 ℓ GL(u) cuando u → 0 +. (A10)

Por (A8), ℓ GL(u) ℓ αθu y por definición, (αθ) κL−1 ℓ GL(u) varía lentamente.

Para las cópulas distorsionadas por UIP, desde $CG_{PAG}^{-1}{}_{(u),\,GRAMO_{PAG}}^{-1}(u) \to 0$ como u → 0 + y por (A8),

$$médico\ de\ cabecera_{PAG}^{-1}{}_{(u),\,GRAMO_{PAG}}^{-1}(tu) = \frac{\theta C\,G_{(u),\,GRAMO_{PAG}}^{-1}{}^{-1}(tú)}{1 - GC_{PAG\ -1}^{-1}{}_{(u),\,GRAMO_{PAG}}^{-1}(u) + \theta C\,G_{PAG}^{-1}{}_{(u),\,GRAMO_{PAG}}^{-1}(tú)}^{\alpha}$$

$$\theta^{\alpha}\,CG_{PAG}^{-1}{}_{(u),\,GRAMO_{PAG}}^{-1}(tú)^{\alpha} \quad \alpha\,[QP(u)]\kappa L\,\ell\,QP(u) \quad \theta^{\alpha}$$

tú κL θ α(1−κL) ℓ QP(u) α cuando u → 0 +. (A11)

Dado que GP(u) θ αu α y ℓ varía lentamente en 0+, para s, α > 0,

$$\lim_{u\to 0+} \frac{\ell(su)}{\ell(u)} = \lim_{u\to 0+} \frac{\ell(su)}{\ell(u)}^{\alpha} = 1;$$

$$\lim_{u\to 0+} \frac{\ell(GP(su))}{\ell(GP(u)} = \lim_{u\to 0+} \frac{\ell(\theta\ a\ s\ \alpha u}{\alpha\ )\,\ell(\theta\ au\ \alpha)} = \lim_{v\to 0+} \frac{\ell(s\ \alpha v)}{\ell(v)} = 1. \quad (A12)$$

Por lo tanto, ℓ QP(u) α está variando lentamente.

Para las cópulas distorsionadas por QUP, $CQ_{PAG}^{-1}(u), Q_{PAG}^{-1}(u) \to 0$ y por (A8), obtenemos

Machine Translated by Google

$$QP\ CQ_{PAG}^{-1}(u), Q_{PAG}^{-1}(tu) = \frac{1}{1+\theta\ CQ_{PAG}^{-1}(u), Q_{PAG}^{-1}(u)\ ]-1/\alpha-1}$$
$$\theta^{-1}\ CQ_{PAG}^{-1}(u), Q_{PAG}^{-1}(tú)^{1/\alpha-1}\ \theta\ [GP(u)]\kappa L\ \ell\ GP(u)^{1/\alpha}$$
$$tú\ \ \kappa L\ \theta\ \kappa L-1\ \ell\ GP(u)^{1/\alpha}\ \text{como tu} \to 0+. \quad \text{(A13)}$$

Por argumentos similares en (A12), $\ell\ GP(u)^{Se}$ puede demostrar que 1/α varía lentamente.

Apéndice B.4. Órdenes de cola

superior Por las aproximaciones en (A7), como u → 0 +,

$$GL(1-u) = 1 - \frac{tu}{\theta+(1-\theta)u^{-1/\alpha}-1\ u}^{\alpha}\ 1-u/\theta^{\alpha},$$
$$QL(1-u) = \frac{}{+\ u-1/\alpha-1\ \theta(1-} = 1-\theta\ \frac{\theta u\ 1/}{(\theta-1)u\ 1/\alpha}\ \frac{\alpha}{1-\theta u\ 1/\alpha\ 1+},$$
$$GP(1-u) = \frac{u)\ u+\theta(1}{-u)}^{\alpha} = 1 - \frac{tu}{\theta+(1-\theta)u}^{\alpha}\ 1-\alpha u/\theta,$$
$$QP(1-u) = \frac{(1-u)\ \theta^{1/\alpha}}{+\ (1-\theta)(1-u)\ 1/\alpha\ \theta u/\alpha}\ \frac{1-u/\alpha}{\theta+(1-\theta)(1-u/\alpha)}$$
$$\frac{1-\theta u/}{\alpha.}\ 1-1-u/\alpha+\theta u/a \quad \text{(A14)}$$

Por (A4), si C(u, u) u κU ℓ (u) como u → 0 +, C(1 − u, 1 − u) 1 − 2u + u κU ℓ (u), como tu → 0 +. Por lo tanto, por (A14),

$$CG_{|}^{-1}(1-u), GRAMO_{|}^{-1}(1-u)\ \ C\ 1-\theta u\ 1/\alpha\ \ 1-, 1-\theta u\ 1/\alpha$$
$$2\theta u\ 1/\alpha+(\theta u^{1/\alpha})\ \kappa U^{1/\alpha}) = 1-kGL(u),$$
$$CQ_{|}^{-1}(1-u), Q_{|}^{-1}(1-u)\ \ C\ 1-(u/\theta)\ \alpha\ \ 1-\ ,\ \ell\ (\theta u\ 1-(u/\theta)\ \alpha$$
$$2(u/\theta)\ \alpha+(u/\theta)\ \alpha)\ \kappa U\ \ell\ ((u/\theta)\ \alpha) = 1-kQL(u), (1-u)$$
$$CG_{PAG}^{-1}(1-u), GRAMO_{PAG}^{-1}\ C\ 1-\theta u/\alpha, 1-\theta u/\alpha\ \ 1-2\theta u/\alpha+(\theta u/$$
$$\alpha)\ \kappa U\ \ell\ (\theta u/\alpha) = 1-kGP(u), (1-u)\ \ C\ 1-\alpha u/\theta, 1-$$
$$CQ_{PAG}^{-1}(1-u), Q_{PAG}^{-1}\ \alpha u/\theta\ \ 1-2\alpha u/\theta+(\alpha u/\theta)\ \kappa U\ \ell\ (\alpha u/\theta)$$
$$= 1-kQP(u). \quad \text{(A15)}$$

Tenga en cuenta que los términos restantes kGL (u), kQL (u), kGP (u) y kQP (u) van a 0 cuando u → 0 +. (1

Para cópulas distorsionadas por UL, por (A14) y luego (A7), $GL_{|}^{-1}$ − u), $GRAM\Phi^{-1}$ (1 − u) es CG dado por

$$1-\frac{1-GC_{|}^{-1}(1-u), GRAMO_{|}^{-1}(1-tu)}{1-GC_{|}^{-1}(1-u), GRAMO_{|}^{-1}(1-u)+\theta C\ G_{|}^{-1}(1-u), GRAMO_{|}^{-1}(1-tu)}^{\alpha}$$
$$1-\frac{kGL}{(u)\ \theta+(1-\theta)kGL\ (u)}^{\alpha}\ 1-kGL\ (u)/\theta^{\alpha}. \quad \text{(A16)}$$

Por lo tanto, por (A4), para 0 < α ≤ 1, y κU ≥ 1, como u → 0 +,

Machine Translated by Google

$$\text{GRAMO}_{\text{l}}\ \text{CG}\ {}^{-1}_{\text{l}}(1-u),\text{GRAMO}_{\text{l}}^{-1}(1-u) = 2u-1+\text{GLCG}\ {}^{-1}_{\text{l}}(1-u),\text{GRAMO}_{\text{l}}^{-1}(1-tu)$$

$$2u-k\text{GL}(u)/\theta^{\ \alpha} = 2u-[2u^{\ 1/\alpha}-\theta\,\kappa\text{U}-1u\,\kappa\text{U}/\alpha\,\ell\ (\theta u^{\ 1/\alpha})]\alpha$$

$$tu\ 2-2^{\ \alpha}[1-2^{\ -1}\,\theta\,\kappa\text{U}-1u\,(\kappa\text{U}-1)/\alpha\,\ell\ (\theta u^{\ 1/\alpha})]\alpha$$

$$u\ 2-2^{\ \ \alpha}[1-\alpha 2^{\ -1}\,\theta\,\kappa\text{U}-1u\,(\kappa\text{U}-1)/\alpha\,\ell\ (\theta u^{1/}{}_{\alpha})]$$

$$\theta\,\kappa\text{U}-1u\,\kappa\text{U}\,\ell\ (\theta u) \text{ si } \alpha = 1. \tag{A17}$$

Para cópulas distorsionadas por QUL, por (A15), $\text{QL}\ \text{CQ}_{\text{l}}^{-1}(1-u),\text{Q}\ {}^{-1}_{\text{l}}(1-u)$ está dada por

$$\frac{1-\text{CQ}\ {}^{-1}_{\text{l}}(1-u),\text{Q}\ {}^{-1}_{\text{l}}(1-tu)^{\ -1/\alpha}-1}{\theta+1-\text{CQ}\ {}^{-1}_{\text{l}}(1-u),\text{Q}\ {}^{-1}_{\text{l}}(1-tu)^{\ -1/\alpha}-1} = \frac{k\text{QL}(u)]-1/\alpha-1}{\theta+k\text{QL}(u)^{\ -1/\alpha}-1}$$

$$= 1-\frac{\theta\,k\text{QL}(u)]1/\alpha}{\theta\,k\text{QL}(u)^{\ 1/\alpha}+1-k\text{QL}(u)^{\ 1/\alpha}}\quad 1-\theta\,k\text{QL}(u)]1/\alpha\,.$$

Por lo tanto, para $\alpha \geq 1$, y $\kappa\text{U} \geq 1$, cuando $u \to 0+$,

$$\text{Q}\hat{}\text{L}\ \text{CG}\ {}^{-1}_{\text{l}}(1-u),\text{GRAMO}_{\text{l}}^{-1}(1-u) = 2u-1+\text{QL}\ \text{CQ}\ {}^{-1}_{\text{l}}(1-u),\text{Q}\ {}^{-1}_{\text{l}}(1-tu)$$

$$2u-\theta\,k\text{QL}(u)]1/\alpha = 2u-\theta\,2(u/\theta)\,\alpha-(u/\theta)\,\alpha)\,\kappa\text{U}\,\ell\ ((u/\theta)\,\alpha)\,(u/\theta)\,\alpha(\kappa\text{U}^{\ 1/\alpha}$$

$$= 2u-2\ 1/\alpha u\ 1-2^{\ \ -1}-1)\,\ell\ ((u/\theta)\,\alpha)^{\ 1/\alpha}$$

$$tu\ 2-2\ 1/\alpha\ 1-2^{\ \ -1-1\ \alpha}\ (u/\theta)\,\alpha(\kappa\text{U}-1)\,\ell\ ((u/\theta)\,\alpha)$$

$$(u/\theta)\,\kappa\text{U}\,\ell\ (u/\theta) \text{ si } \alpha = 1. \tag{A18}$$

Para cópulas distorsionadas por UIP, por (A7) y (A15), $\text{GP}\ \text{CG}_{\text{PAG}}^{-1}(1-u),\text{GRAMO}_{\text{PAG}}^{-1}(1-u)$ es dado por

$$\frac{\theta\text{C G}\ {}^{-1}_{\text{PAG}}(1-u),\text{GRAMO}_{\text{PAG}}^{-1}(1-tu)^{\ \alpha}}{1-\text{GC}\ {}^{-1}_{\text{PAG}}(1-u),\text{GRAMO}_{\text{PAG}}^{\text{PAG}-1}(1-u)+\theta\text{C G}\ {}^{-1}_{\text{PAG}}(1-u),\text{GRAMO}_{\text{PAG}}^{-1}(1-tu)} = \frac{\theta[1-k\text{GP}(u)]}{k\text{GP}(u)]+\theta[1-k\text{GP}(u)]}^{\ \alpha} = 1-\frac{k\text{GP}}{(u)\,k\text{GP}(u)]+\theta[1-k\text{GP}}^{\ \alpha}$$

$$1-k\text{GP}(u)/\theta^{\ \alpha}(u)]\quad 1-\alpha\theta-1\,k\text{GP}(u).$$

Por lo tanto, para $\alpha \geq 1$, y $\kappa\text{U} \geq 1$, cuando $u \to 0+$,

$$\text{GRAMO}_{\text{PAG}}\text{P.G.G.}\ {}^{-1}_{\text{PAG}}(1-u),\text{GRAMO}_{\text{PAG}}^{-1}(1-u) = 2u-1+\text{GP CG}\quad 2u-{}^{-1}_{\text{PAG}}(1-u),\text{GRAMO}_{\text{PAG}}^{-1}(1-tu)$$

$$\alpha\theta-1\,k\text{GP}(u) = 2u-\alpha\theta-1\,[2\theta u/\alpha-(\theta u/\alpha)\,\kappa\text{U}\,\ell\ (\theta u/\alpha)]\quad \alpha\theta-1\,(\theta u/\alpha)$$

$$\kappa\text{U}\,\ell\ (\theta u/\alpha). \tag{A19}$$

Para cópulas distorsionadas por QIP, mediante (A7) y (A15), $\text{QP}\ \text{CQ}_{\text{PAG}}^{-1}(1-u),\text{Q}\ {}^{-1}_{\text{PAG}}(1-u)$ viene dada por

$$\frac{\text{CQ}\ {}^{-1}_{\text{PAG}}(1-u),\text{Q}\ {}^{-1}_{\text{PAG}}(1-tu)\,]^{\ 1/\alpha}}{\text{CQ}\ {}^{-1}_{\text{PAG}}(1-u),\text{Q}\ {}^{-1}_{\text{PAG}}(1-u)\,]\,1/\alpha+\theta\,1-\text{CQ}\ {}^{-1}_{\text{PAG}}(1-u),\text{Q}\ {}^{-1}_{\text{PAG}}(1-u)\,]\,1/\alpha}$$

$$= \frac{[1-k\text{QP}(u)]1/\alpha}{1-k\text{QP}(u)]1/\alpha+\theta\,1-1-k\text{QP}(u)]1/\alpha\,1-}$$

$$= \frac{k\text{QP}(u)/\alpha\,1-}{k\text{QP}(u)/\,\alpha+\theta k\text{QP}(u)/\alpha\,\theta k\text{QP}} = \frac{\alpha-k\text{QP}(u)}{\alpha-k\text{QP}(u)+\theta k\text{QP}(u)}$$

$$= 1-\frac{(u)\quad 1}{\theta k\text{QP}(u)}-\theta k\text{QP}(u)/\alpha.\ \alpha-k\text{QP}(u)+$$

Machine Translated by Google

Por lo tanto,

$$Q\hat{}\ P\ CQ\ {}^{-1}_{PAG}(1-u), Q\ {}^{-1}_{PAG}(1-u) = 2u-1+QPCQ\ {}^{-1}_{PAG}(1-u), Q\ {}^{-1}_{PAG}(1-tu)$$

$$2u-\theta kQP(u)/\alpha = 2u-\theta\alpha-1\,[2\alpha u/\theta-(\alpha u/\theta)\,\kappa U\,\ell\ (\alpha u/\theta)]\,\kappa U(\alpha/)$$

$$^{tú}\,\kappa U-1\,\ell\ (\alpha u/\theta). \tag{A20}$$

Aplicando argumentos similares a los de (A12), podemos demostrar que ℓ en (A17)–(A20) varían lentamente.

Apéndice C. Ordenamiento de concordancia

Una función f es subaditiva si $f(x + y) \le f(x) + f(y)$ para todo xey en su dominio. A la función f es superaditiva si $f(x + y) \ge f(x) + f(y)$ para todo xey en su dominio.

Teorema A1 ([11]). Sean C1 y C2 dos cópulas de Arquímedes generadas por ψ1 y ψ2. Entonces C1 C2 si y sólo si $\psi_1 \circ \psi_2^{-1}$ es subaditivo.

Además, C1 C2 si sólo si $\psi_1 \circ \psi_2^{-1}$ es superaditivo.

Aquí, demostramos usando el Teorema A1 que las familias de independencia UL y UIP las cópulas están ordenadas en el parámetro θ. Para la cópula de independencia de UL, $\Psi_1 \circ \Psi_2^{-1} = \psi \circ T_1^{-1} \circ T2 \circ \psi^{-1}$, donde $\psi(t) = -\log(t)$ y $T(u) = GL(u) = Q\ {}^{-1}_{l}(tú)$. Sea $h(u) = T\ {}^{-1}_{1} \circ T2$. Para $\theta 1 < \theta 2 \le 1$, de la Tabla 1, derivamos que

$$h(u) = QL1 \circ GL2(u) = 1 + \frac{1}{\theta 1(1-u)/(\theta 2u)^{-(x+y)}} = \frac{\theta 2u}{(\theta 2-\theta 1)u+\theta 1}$$

$$\Psi 1 \circ \Psi\ {}^{-1}_{2}(x+y) = -\text{Iniciar sesión}\ \frac{^{\theta\ 2}_{\theta\ mi\ 2}}{\theta 2\,(\theta 2-\theta 1)e^{-(x+y)}+\theta 1} \tag{A21}$$

$$\Psi 1 \circ \Psi\ {}^{-1}_{2}(x) + \Psi 1 \circ \Psi\ {}^{-1}_{2}(y) = -\text{Iniciar sesión}\ \frac{\theta 2e^{-x}}{(\theta 2-\theta 1)e^{-x}+\theta 1}\ \frac{\theta 2e^{-y}}{(\theta 2-\theta 1)e^{-y}+\theta 1} \tag{A22}$$

Deseamos comparar (A21) y (A22) para determinar si $\Psi 1 \circ \Psi^{-(x+y)} + 1 \ge {}^{-1}_{2}$ es subaditivo o desde e e superaditivo. $^{-x} + e^{-y}$ para x, y ≥ 0, obtenemos que

$$\frac{\theta 2}{x+y} + e \ge -e\ \theta 2-\theta 1\ ^{-x} + +\ \theta 1-y\ \frac{}{e\ \theta 2-\theta 1}$$

$$y\ \theta 2e^{-(x+y)} + \ge (\theta 2-\theta 1)e^{-(x+y)} + \theta 1e\ \theta 2-\theta 1\ ^{-x} + \theta 1e^{-y} + \frac{^{\theta\ 2}_{1}}{(\theta 2-\theta 1)}.$$

Por tanto, (A21) es mayor o igual que (A22). Es decir, $\Psi 1 \circ \Psi_2^{-1}$ es superaditivo. Por tanto, la familia de cópulas de independencia de UL está ordenada negativamente por el parámetro θ. Dado que la distorsión UIP da lugar a la misma h(u), la familia de cópulas de independencia UIP también está ordenada negativamente por el parámetro θ.

Referencias

1. Genest, C.; Ghoudi, K.; Rivest, LP Comprensión de las relaciones mediante cópulas por Edward Frees y Emiliano Valdez. N. Am. Actuador. J. 1998, 2, 143-149. [Referencia cruzada]
2. Adés, M.; Preboste, SB; Zang, Y. Cuatro medidas de asociación y sus representaciones en términos de cópulas. Matemáticas Aplicadas 2024, 4, 363–382. [Referencia cruzada]
3. Cherubini, U.; Luciano, E.; Vecchiato, W. Métodos Copula en Finanzas; John Wiley & Sons: Hoboken, Nueva Jersey, EE. UU., 2004. [CrossRef] 4. Libera, EW; Valdez, EA Comprensión de las relaciones mediante cópulas. N. Am. Actuador. J. 1998, 2, 1–25. [Referencia cruzada]
5. Embrechts, P.; McNeil, A.; Straumann, D. Correlación y dependencia en la gestión de riesgos: propiedades y dificultades. En Gestión de riesgos : valor en riesgo y más allá; Dempster, MAH, Ed.; Cambridge University Press: Cambridge, Reino Unido, 2002; págs. 176-223.
6. Ji, L.; Bronceado, KS; Yang, F. Dependencia de la cola y gran cola en riesgos extremos. Asegurar. Matemáticas. Economía. 2021, 99 282–293. [Referencia cruzada]

Machine Translated by Google

7. Alzaid, AA; Alhadlaq, WMA Nueva familia de cópulas de Arquímedes: la familia de cópulas semilogísticas. Matemáticas 2024, 12, 101. [Referencia cruzada]

8. Tootoonchi, F.; Sadegh, M.; Haerter, JO; Raty, O.; Agarra, T.; Teutschbein, C. Cópulas para análisis hidroclimático: una visión general orientada a la práctica. CABLES Agua 2022, 9, e1579. [Referencia cruzada]

9. Chen, L.; Guo, S. Copulas y su aplicación en hidrología y recursos hídricos; Springer: Berlín/Heidelberg, Alemania, 2019; 290p; ISBN 978-981-13-0573-3.

10. Sklar, M. Fonctions de Repartition an Dimensions et Leurs Marges. Ana. ISUP 1959, 8, 229–231.

11. Nelsen, RB Introducción a las cópulas; Springer: Nueva York, NY, Estados Unidos, 2006.

12. Joe, H. Modelado de dependencia con cópulas; Prensa CRC: Boca Ratón, FL, EE. UU., 2015.

13. Yang, Y.; Li, S. Sobre una familia de cópulas de Arquímedes generadas con log-gamma. N. Am. Actuar. J. 2022, 26, 123-142. [Referencia cruzada]

14. Manstavicius, M.; Bagdonas, G. Una clase de transformaciones de cópula de independencia bivariada. Sistema de conjuntos difusos. 2022, 428, 58–79. [Referencia cruzada]

15. Chesneau, C. Una colección de nuevas cópulas de tipo trigonométrica e hiperbólica para la mutilación genital femenina. Matemáticas aplicadas 2023, 3, 147–174. [Referencia cruzada]

16. Chesneau, C. Una nota sobre dos nuevas cópulas logarítmicas y exponenciales de tres parámetros. Asia Pacífico. J. Matemáticas. 2023, 10, 8.

17. Di Bernardino, E.; Rulliere, D. Sobre ciertas transformaciones de las cópulas de Arquímedes: aplicación a lo no paramétrico Estimación de sus Generadores. Depender. Modelo. 2013, 1, 1–36. [Referencia cruzada]

18. Genest, C.; Kilani, G.; Rivest, LP Un procedimiento de estimación semiparamétrica de parámetros de dependencia en familias de distribuciones multivariadas. Biometrika 1995, 82, 543–552. [Referencia cruzada]

19. Morillas, PM Un método para obtener nuevas cópulas a partir de una determinada. Métrica 2005, 61, 169–184. [Referencia cruzada]

20. Durante, F.; Foschi, R.; Sarkoci, P. Cópulas distorsionadas: construcciones y dependencia de la cola. Comunitario. Estadística. Métodos teóricos 2010, 39, 2288–2301. [Referencia cruzada]

21. Valdez, EA; Xiao, Y. Sobre la distorsión de una cópula y sus márgenes. Escanear. Actuador. J. 2011, 4, 292–317. [Referencia cruzada]

22. Samanthi, RGM; Sepanski, J. Una extensión bivariada de la distribución generada beta derivada de cópulas. Comunitario. Estadística. Métodos teóricos 2019, 48, 1043–1059. [Referencia cruzada]

23. Samanthi, RGM; Sepanski, J. Sobre las cópulas bivariadas distorsionadas por Kumaraswamy. Comunitario. Estadística. Métodos teóricos 2020, 51, 2477–2495. [Referencia cruzada]

24. Aldhufairi, FAA; Samanthi, RGM; Sepanski JH Nuevas familias de cópulas bivariadas mediante distorsión unitaria Lomax. Riesgos 2020, 8, 106. [Referencia cruzada]

25. Aldhufairi, FAA; Sepanski, JH Nuevas familias de cópulas bivariadas mediante distorsión unitaria de Weibull. J. estadística. Distribuir. Aplica. 2020, 7, 1–20. [Referencia cruzada]

26. Aldhufairi, FAA Utilizando una cópula unitaria distorsionada de Gompertz para modelar la dependencia en datos antropométricos. Comunitario. Estadística. Aplica. Métodos 2023, 30, 467–483. [Referencia cruzada]

27. Kularatne, TD; Li, J.; Pitt, D. Sobre el uso de cópulas de Arquímedes para el modelado de seguros. Ana. Actuador. Ciencia. 2021, 15, 57–81. [Referencia cruzada]

28. Schweizer, B.; Sklar, A. Espacios métricos probabilísticos; Publicaciones de Dover: Holanda Septentrional, Nueva York, EE. UU., 1983.

29. Hua, L.; Joe, H. Dependencia de la cola intermedia: una revisión y algunos resultados nuevos. En la conferencia sobre órdenes estocásticas en confiabilidad y riesgo Notas en Estadística; Springer: Berlín/Heidelberg, Alemania, 2013; págs. 291–311.

Descargo de responsabilidad/Nota del editor: Las declaraciones, opiniones y datos contenidos en todas las publicaciones son únicamente de los autores y contribuyentes individuales y no de MDPI ni de los editores. MDPI y/o los editores renuncian a toda responsabilidad por cualquier daño a personas o propiedad que resulte de cualquier idea, método, instrucción o producto mencionado en el contenido.

Machine Translated by Google

AppliedMath

MDPI

Artículo

Modelos de epidemia SIRS con retrasos, parciales y temporales Inmunidad y vacunación

Benito Chen Charpentier

Departamento de Matemáticas, Universidad de Texas en Arlington, Arlington, TX 76019, EE. UU.; bmchen@uta.edu

Resumen: La reproducción básica, o número reproductivo, es un índice útil que indica si habrá o no una epidemia. Sin embargo, también es muy importante determinar si una epidemia eventualmente disminuirá y desaparecerá o persistirá como endémica. Diferentes enfermedades infecciosas tienen diferentes comportamientos y los modelos matemáticos utilizados para simularlas deberían capturar los procesos más importantes; sin embargo, los modelos también implican simplificaciones. Las epidemias de influenza suelen ser de corta duración y pueden modelarse con ecuaciones diferenciales ordinarias sin considerar la demografía. Retrasos como el tiempo de infección pueden cambiar el comportamiento de las soluciones. Lo mismo ocurre si existe inmunidad permanente o temporal, o inmunidad total o parcial. La vacunación, el aislamiento y el uso de antivirales también pueden cambiar el resultado. En este artículo, presentamos varios modelos nuevos y los utilizamos para encontrar los efectos de todos los factores anteriores, prestando especial atención a si el modelo puede representar un proceso infeccioso que eventualmente desaparece. Determinamos las soluciones de equilibrio y establecemos la estabilidad del equilibrio libre de enfermedades utilizando v También mostramos que muchos modelos de influenza u otras epidemias de corta duración no tienen soluciones con una epidemia que desaparece. El principal objetivo del artículo es presentar diferentes formas de modelar la inmunidad en modelos epidémicos. Se estudian varios escenarios con diferentes inmunidades ya que una persona puede no volver a infectarse porque tiene inmunidad total o parcial o porque no hubo contactos estrechos. Mostramos que algunos cambios relativamente pequeños, como en la tasa de vacunación, pueden cambiar significativamente la dinámica; por ejemplo, la existencia y el número de equilibrios libres de enfermedad. También ilustramos que si bien la introducción de retrasos hace que los modelos sean más realistas, la dinámica tiene el mismo comportamiento cualitativo.

Palabras clave: influenza; modelos epidémicos; ecuaciones diferenciales; número de reproducción básico; inmunidad parcial

check for updates

Cita: Chen-Charpentier, B. Modelos de epidemia SIRS con retrasos, inmunidad y vacunación parcial y temporal. Matemáticas aplicadas 2024, 4, 666–689. https://doi.org/10.3390/appledmath4020036

Editor académico: Carlo Bianca

Recibido: 7 de mayo de 2024
Revisado: 17 de mayo de 2024
Aceptado: 23 de mayo de 2024
Publicado: 28 de mayo de 2024

Copyright: © 2024 por el autor. Licenciatario MDPI, Basilea, Suiza.
Este artículo es un artículo de acceso abierto. distribuido bajo los términos y condiciones de los Creative Commons Licencia de atribución (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

1. Introducción

La influenza o gripe es una enfermedad contagiosa que se transmite por contacto con una persona infectada. La influenza estacional afecta entre el 3% y el 11% de la población de los Estados Unidos cada año [1]. La gripe alcanza su punto máximo de diciembre a febrero de cada año. Aunque algunos pacientes mueren a causa de la gripe estacional, no es una enfermedad particularmente mortal, con 0,2 muertes por 100.000 habitantes [2]. Otras enfermedades contagiosas que duran poco son el resfriado común, la enfermedad del virus respiratorio sincitial y el estreptococo de garganta. Hay muchos artículos que tratan de modelos matemáticos para epidemias, tanto por contacto directo como por transmisión vectorial. Los modelos matemáticos de epidemias comenzaron a principios del siglo XX. Uno de los primeros esfuerzos es [3] y el artículo fundamental de Kermack y McKendrick [4] es ampliamente reconocido por introducir el primer modelo epidémico compartimental basado en ecuaciones diferenciales ordina El uso de EDO es un buen supuesto de modelización si las diferentes poblaciones están bien mezcladas y hay poca o ninguna dependencia espacial. Kermack y McKendrick introdujeron el modelo SIR (susceptible-infectado-eliminado), que supone que la población puede estar en uno de tres estados: susceptible, infectada o infectiva, y eliminada (recuperada con inmunidad permanente o fallecida). El modelo SIR y sus múltiples variaciones, incluido el SIS (susceptible-infeccioso-susceptible) y SIRS (susceptible-infeccioso-recup

Machine Translated by Google

Los modelos han sido muy influyentes en la comprensión de las epidemias y su control. Entre las modificaciones realizadas al modelo SIR original de Kermack y McKendrick se encuentran la introducción de datos demográficos (no importantes para la gripe estacional), vacunas, diferentes grupos de edad, uso de antivirales y aislamiento. Una modificación importante es la adición de retrasos . Los retrasos aparecen en las epidemias debido al tiempo que tardan en ocurrir ciertos procesos, como el tiempo que tarda un susceptible en infectarse tras un contacto con un individuo infectado, o el tiempo que tardan las vacunas en hacerse efectivas tras ser administradas. Las dos formas principales de introducir retrasos en un modelo basado en ecuaciones diferenciales ordinarias son mediante el uso de ecuaciones diferenciales de retraso (DDE) o mediante la introducción de una población adicional expuesta o latente. Las DDE pueden tener una dinámica diferente a la de las ODE [5,6]. Entre las posibles diferencias se encuentran la introducción de oscilaciones, la introducción de discontinuidades en las derivadas temporales, la no unicidad y la diferente estabilidad. Si bien existe una amplia gama de literatura sobre modelos de epidemias o transmisión de enfermedades que utilizan ecuaciones diferenciales de retardo, la mayoría de los libros sobre EDD tienen una cobertura escasa de la propagación de enfermedades contagiosas, y los libros sobre epidemias tienen poco material sobre EED. D
Los modelos SIR consideran que después de ser infectados, los individuos se recuperan y alcanzan una inmunidad permanente y total frente a la enfermedad. Para la influenza estacional y otras enfermedades contagiosas, la inmunidad es temporal, lo que lleva a modelos SIRS en los que los individuos recuperados pierden la inmunidad por completo después de un cierto período de tiempo. Sin embargo, existen otras posibilidades: los individuos que se recuperan pierden la inmunidad total pero mantienen la inmunidad parcial para siempre; los individuos que se recuperaron pierden la inmunidad total y alcanzan la inmunidad parcial por un período de tiempo; algunos individuos que se recuperaron mantienen una inmunidad total y permanente, mientras que otros solo mantienen una inmunidad permanente parcial, y así sucesivamente. Lo que suceda depende de la infección y la población en particular. Las vacunas también pueden conferir inmunidad total permanente, o inmunidad total que disminuye hasta convertirse en inmunidad parcial, u otros escenarios

Uno de los usos más importantes de los modelos matemáticos de enfermedades contagiosas es predecir si habrá una epidemia o no. El número de reproducción básico R0 , definido como el número de infecciones secundarias producidas por un individuo recién infectado al inicio del proceso infeccioso, es un índice útil para determinar si la epidemia morirá o no. Generalmente se calcula utilizando el método matricial de próxima generación [9,10]. También proporciona la estabilidad del equilibrio libre de enfermedad (DFE) del modelo. En caso de que el número de infectantes crezca inicialmente, también es muy importante determinar si la epidemia eventualmente morirá o no. En este artículo, desarrollaremos modelos matemáticos de enfermedades contagiosas de corta duración como la influenza estacional que incluyen diferentes supuestos sobre la inmunidad total, parcial, permanente y temporal. También agregaremos vacunas con diferentes efectos inmunológicos, así como el uso de antivirales y aislamiento. Los modelos incluirán versiones retrasadas, utilizando DDE y también poblaciones expuestas. Para estos modelos estudiaremos si la epidemia persiste o no. Árbitro. [11] compara algunos modelos básicos para la influenza que usaremos como punto de partida. El resto del documento está organizado de la siguiente manera.
En la Sección 2, presentamos varios modelos nuevos, que son extensiones del modelo SIRS, con diferentes hipótesis sobre la inmunidad total y parcial y la vacunación, determinamos sus equilibrios y analizamos la estabilidad de sus equilibrios libres de enfermedad. En la Sección 3 se presentan simulaciones numéricas y resultados, incluidos los efectos de la introducción de retrasos. También mostramos que algunos cambios relativamente pequeños en los modelos pueden cambiar significativamente la dinámica. Finalmente, en la Sección 4, se presentan las discusiones y conclusiones sobre los modelos y resultados.

2. Materiales y métodos

En esta sección, presentamos varios modelos matemáticos nuevos de epidemias a corto plazo con diferentes tipos de inmunidad. A esto le sumamos la vacunación, el uso de antivirales y el aislamiento o cuarentena. Para determinar si se desarrolla una epidemia según el modelo utilizado, es necesario encontrar los puntos de equilibrio. Si hay un punto de equilibrio sin infectantes y susceptibles igual a la población total, y si el número de infectantes crece con el tiempo, entonces habrá una epidemia. Un equilibrio con cero infecciosos es

Machine Translated by Google

conocido como equilibrio libre de enfermedad (DFE). Existen diferentes métodos para determinar si el número de infecciosos crece: la ecuación diferencial para los infecciosos tiene un lado derecho positivo; el DFE es localmente inestable; utilizando el número de reproducción básico R0 se obtiene R0 > 1; el subsistema de ecuaciones de los compartimentos infecciosos es localmente inestable; el índice R0 calculado utilizando el método matricial de próxima generación es mayor que 1 [10,12]. En caso de que haya una epidemia, es muy importante establecer si eventualmente morirá. Esto se puede calcular determinando si existe un DFE para el tiempo positivo que sea localmente asintóticamente estable. Esto puede suceder durante un tiempo finito o sólo cuando el tiempo tiende al infinito. Esta determinación se llevará a cabo en las siguientes subsecciones para diferentes modelos sin datos demográficos. La primera subsección es una revisión del modelo susceptible-infeccioso-susceptible (SIS), el modelo susceptible-infeccioso-recuperado (SIR) y el modelo susceptible-infeccioso-recuperado-susceptible (SIRS) con diferentes inmunidades.

2.1. Modelos SIS, SIR y SIRS

Un modelo epidémico muy simple, para una población total constante N, que consta de sólo dos compartimentos poblacionales, susceptibles S e infecciosos I, es el modelo SIS. Se supone que un susceptible se infecta con probabilidad β/N > 0 por un contacto con un infeccioso, y luego se recupera sin inmunidad y regresa al compartimento susceptible. El resfriado común es un ejemplo de enfermedad contagiosa que puede modelarse mediante un sistema de ecuaciones SIS.
La población total suele considerarse constante, lo que supone que la enfermedad es rápida en comparación con los efectos demográficos. Una suposición casi equivalente, dado que la tasa de mortalidad es muy pequeña en comparación con el período de la enfermedad, es que el número de nacimientos es igual al número de muertes sin migración. El modelo SIS sin datos demográficos es [13,14]

$$\begin{aligned} \frac{dS}{dt} &= -\frac{b}{}SI + \gamma I \\ \frac{dI}{dt} &= \frac{norte\,\beta}{norte}SI - \gamma I, \end{aligned} \tag{1}$$

donde γ > 0 es la tasa de recuperación y N > 0 es la población total. Como la población es constante, S(t) = N − I(t) y podemos eliminar una ecuación

$$\frac{yo}{dt} = \frac{\beta}{norte}(norte - yo)yo - \gamma yo. \tag{2}$$

El modelo SIS tiene dos estados de equilibrio, I = 0 e I = (1 − físico solo $\frac{\gamma}{b}$)NORTE. El último es para γ ≤ 1. El número de reproducción básico (también llamado número reproductivo básico) R0 es b el número de infecciosos secundarios generados por un único infectante al principio. del proceso de infección para el modelo SIS se puede calcular como el producto de la probabilidad de un nuevo infeccioso, β/N, por el número de susceptibles al inicio del proceso infeccioso, N, por la vida media de una infección . , 1/γ Por lo tanto R0 = β/γ. Si R0 > 1 el número de infecciosos aumenta a partir del DFE, y el otro equilibrio tiene sentido biológico ya que todos los estados no son negativos y si R0 < 1 la infección desaparece . el equilibrio DFE es el único que tiene sentido biológico.

R0 > 1 implica que el DFE es localmente asintóticamente inestable, por lo que al menos uno de los valores propios del sistema linealizado sobre el DFE tiene parte real positiva. Para determinar si la infección está creciendo sólo es necesario observar el compartimento infectado, I(t). El método matricial de próxima generación [9,10] se utiliza ampliamente para calcular el número de reproducción básico en modelos más complejos en epidemiología. Aplicando este método al modelo SIS solo hay un compartimento infeccioso, I, por lo que el J jacobiano de este sistema es una matriz de 1 × 1. Descomponiendo esta matriz como J = T + Σ, con T la transmisión y Σ las matrices de transición, se En T = βS/N y Σ = −γ. La matriz de próxima generación es K = −TΣ el $^{-1}$ produce = βS/(γN). comienzo de la infección S = N − 1, por lo que el radio espectral de K, que es R0, puede aproximarse mediante R0 = β/γ. La estabilidad local del DFE también se puede establecer encontrando un índice que indique cuándo el valor propio del sistema linealizado que consta de

Machine Translated by Google

El compartimento de enfermedad evaluado en el DFE tiene parte real negativa. Una forma sencilla de encontrar si el número de infecciosos crece al inicio de la epidemia es evaluar la ecuación diferencial para I en el estado inicial con S = N:

$$\frac{yo}{dt} = \frac{b}{norte} NI - \gamma I = (\beta - \gamma)I,$$

lo cual es positivo para β/γ > 1. Tenga en cuenta que dado que al inicio de la epidemia I = 1, S = N − 1. Sin embargo, para N grande , S puede considerarse igual a N. Otra forma de encontrar la estabilidad del DFE al inicio de la infección es calculando los valores propios de la matriz jacobiana correspondiente al sistema completo de ecuaciones, pero las fórmulas suelen ser complicados o no pueden determinarse analíticamente. Otra posibilidad es utilizar el polinomio característico y el criterio de Routh- Hurwitz [13,15].

Se pueden agregar dos retrasos al modelo SIS: el tiempo que tarda la enfermedad en desarrollarse en un individuo susceptible después de haber sido infectado por un contacto con un individuo infeccioso, y el tiempo que tarda un individuo recién infectado en recuperarse. Este segundo retraso suele incluirse implícitamente ya que 1/γ es el tiempo medio de recuperación. Hay dos formas de introducir el primer retraso. La primera forma viene dada por SIS Delay Model I:

$$\begin{aligned} \frac{dS}{dt} &= -\frac{b}{norte} S(t-\tau)I(t-\tau) + \gamma I(t) \\ \frac{dI}{dt} &= \frac{b}{norte} S(t-\tau)I(t-\tau) - \gamma I(t). \end{aligned} \tag{3}$$

El supuesto es que un contacto en t − τ produce un nuevo infectivo en el momento t. La segunda forma viene dada por SIS Delay Model II:

$$\begin{aligned} \frac{dS}{dt} &= -\frac{\beta}{norte} S(t)I(t-\tau) + \gamma I(t) \\ \frac{dI}{dt} &= \frac{\beta}{norte} S(t)I(t-\tau) - \gamma I(t). \end{aligned} \tag{4}$$

La suposición ahora es que un individuo susceptible después del contacto con un infeccioso se convierte en infeccioso pero tarda un tiempo τ en infectarse con la enfermedad y poder infectar. Ambos modelos tienen sentido biológico. Otro enfoque para estudiar modelos epidémicos es considerarlos como reacciones químicas, como se muestra, por ejemplo, en [16]. A modo de ilustración, la reacción entre un infeccioso y un susceptible produce dos infecciosos.

Para que se produzca la reacción, ambas especies deben estar al mismo tiempo, lo cual es una advertencia para el Modelo de Retraso II. Sin embargo, dado que en muchos artículos se utilizan ambas formas de introducir el retraso de la infección , incluimos ambas para ver si producen la misma dinámica cualitativamente. La dinámica puede ser diferente ya que sus ecuaciones características son diferentes. Para otros modelos epidémicos que introducen el retraso en la infección como se realizó en el Modelo de Retraso 1, consulte, por ejemplo, [17-19]. Y para otras aplicaciones que utilizan el retardo como en el Modelo de retardo 2, consulte [20,21].

Para el modelo SIS, las soluciones de equilibrio de ambos modelos de retardo son las mismas que las del modelo ODE. Para determinar la estabilidad del DFE para el modelo (3), sea I = i y S = N − s. Entonces la ecuación diferencial linealizada para i es = βi − γi. Tomando $i = e^{rt}$ obtenemos la ecuación característica

$$r = \beta \exp(-r\tau) - \gamma.$$

Esta ecuación es trascendental y no existe una solución exacta, pero r > 0 sólo para β > γ. Entonces tenemos el mismo R0 que el modelo ODE. Un cálculo similar para el modelo (4) arroja el mismo resultado. Este método ha sido aplicado a otros modelos por [20,22].

Machine Translated by Google

Otra forma de introducir el tiempo que tarda en desarrollarse la infección es añadir otro compartimento para la población expuesta o latente (E). Después de un contacto con un infeccioso, el individuo susceptible pasa al compartimento expuesto y pasa allí un tiempo medio igual al tiempo de infección. Los individuos expuestos no pueden infectar. El modelo SIS con población expuesta es

$$\begin{aligned} \frac{dS}{dt} &= -\frac{b}{}SI + \gamma I \\ \frac{dE}{dt} &= \frac{norte\ \beta}{SI} - \frac{1}{mi\ \tau} \\ \frac{dI}{dt} &= \frac{}{\tau} norte\ 1\ mi - \gamma I. \end{aligned} \tag{5}$$

El equilibrio DFE de (5) es el mismo que el de (1). Su estabilidad se puede calcular utilizando el método matricial de próxima generación y da R0 = β/γ. Entonces la estabilidad es la misma que la de (1). Las poblaciones expuestas se han utilizado en muchos modelos epidémicos, como los de [23,24].

El modelo SIR de Kermack y McKendrick [4] considera que los recuperadores infectados tienen inmunidad total y permanente. Añade el compartimento recuperado (R) al modelo SIS:

$$\begin{aligned} \frac{dS}{dt} &= -\frac{b}{}SI \\ \frac{dI}{dt} &= \frac{norte\ \beta}{norte}SI - \gamma I \\ \frac{dR}{dt} &= \gamma I. \end{aligned} \tag{6}$$

Dado que la población total es constante, se puede eliminar una ecuación usando R(t) = N − S(t) − I(t). El nuevo sistema de ecuaciones es

$$\begin{aligned} \frac{dS}{dt} &= -\frac{\beta}{N}SI \\ \frac{dI}{dt} &= \frac{\beta}{norte}SI - \gamma I. \end{aligned}$$

La única solución estable de (6) es I = 0, S + R = N. Al inicio de la infección S(0) = N. La segunda ecuación de (6) establece que en el momento inicial, t = 0, $\frac{yo}{dt} > 0$ para β > γ.
Entonces, para β/γ > 1 el DFE es inestable, y I(t) tiene su valor máximo cuando S = Nγ/β. γNORTE S(∞)
Además, la solución cuando t tiende a infinito es S(∞) = N − R(0) − ln( ) y β S(0)
S(∞) > 0 [13]. Entonces si hay una epidemia esta no desaparece.

El retraso debido al momento de la infección se puede introducir de dos maneras como se realizó para el modelo SIS. La solución de equilibrio de ambos modelos con retraso es el DFE con la misma estabilidad que para el modelo sin retraso. El mismo resultado de estabilidad para el DFE del modelo ODE sin demora es válido para el modelo SIR con la población expuesta.

Un modelo para la propagación de muchas enfermedades diferentes en una población es el modelo SIRS sin datos demográficos [13]. Consta de tres poblaciones: susceptibles (S(t)), infectantes o infectados (I(t)) y recuperados (R(t)). Un susceptible se convierte en infeccioso con una probabilidad dada después de un contacto con un infeccioso. Un infeccioso se recupera con una tasa γ. Un recuperado pierde inmunidad a un ritmo v y se convierte en susceptible. Se supone que la población total es constante, S(t) + I(t) + R(t) = N, y se supone que la enfermedad es lo suficientemente breve como para que no sea necesario considerar los nacimientos y las muertes. El coeficiente de infección se denotará por β/N. Todos los parámetros son positivos. El sistema de ecuaciones diferenciales ordinarias (EDO) que describe el modelo es [13,14]

Machine Translated by Google

$$\begin{aligned}\frac{dS}{dt} &= -\frac{b}{}SI + vR\\ \frac{dI}{dt} &= \frac{norte\ \beta}{norte}SI - \gamma I\\ \frac{dR}{dt} &= \gamma I - vR.\end{aligned} \tag{7}$$

Dado que la población total es constante, se puede eliminar una ecuación usando R(t) = N − S(t) − I(t). El nuevo sistema de ecuaciones es

$$\begin{aligned}\frac{dS}{dt} &= -\frac{\beta}{}SI + v(norte - S - yo)\\ \frac{dI}{dt} &= \frac{norte\ \beta}{norte}SI - \gamma I.\end{aligned} \tag{8}$$

El modelo SIRS tiene dos soluciones de equilibrio: el DFE, S = N, I = 0, R = 0, y (β−γ)Nv el equilibrio endémico, $\beta(\gamma+v)$, el número es R0 = β/γ.
S = γN/β, I = R = N − S − I. La reproducción básica

Existen varios métodos diferentes para introducir retrasos en la infección en el modelo: El primero que llamaremos SIRS Modelo I [17-19] está dado por

$$\begin{aligned}\frac{dS}{dt} &= -\frac{b}{}S(t-\tau)I(t-\tau) + v(N - S(t) - I(t))\\ \frac{dI}{dt} &= \frac{norte\ \beta}{norte}S(t-\tau)I(t-\tau) - \gamma I(t),\end{aligned} \tag{9}$$

donde la hipótesis es que un infeccioso tiene contacto con un susceptible y toma un tiempo τ para que este susceptible se vuelva infectivo. No se cuenta como susceptible hasta ese momento. Un segundo modelo es el SIRS Modelo II [25] descrito por:

$$\begin{aligned}\frac{dS}{dt} &= -\frac{b}{}S(t)I(t-\tau) + v(N - S(t) - I(t))\\ \frac{dI}{dt} &= \frac{norte\ \beta}{norte}S(t)I(t-\tau) - \gamma I(t),\end{aligned} \tag{10}$$

donde se supone que después de un contacto entre un susceptible y un infeccioso, pasa un tiempo τ antes de que el susceptible se convierta en infectivo y luego pueda infectar a otro susceptible.

En lugar de considerar una población recuperada, donde los individuos pasan un tiempo promedio 1/v antes de perder inmunidad, los individuos infecciosos después de dejar de ser infecciosos tienen que esperar un tiempo de retraso 1/v antes de volverse infecciosos. En [26], esto se realiza además de conservar el compartimento recuperado.

Los dos modelos con retrasos presentados anteriormente tienen las mismas soluciones de equilibrio, pero las soluciones reales que dependen del tiempo serán diferentes. También está la cuestión de la influencia de las condiciones iniciales dadas en el intervalo histórico. Para una epidemia, siempre se pueden presentar como soluciones de equilibrio sin enfermedad. Es decir, se puede utilizar el supuesto natural de que la epidemia comienza en t = 0.

Machine Translated by Google

Un modelo que tiene en cuenta la dependencia en el pasado, provocada por el retraso en la transmisión, al incluir una población expuesta, E, es [27,28]

$$\begin{aligned}\frac{dS}{dt} &= -\frac{b}{} SI + vR\\ \frac{dE}{dt} &= \frac{norte\,\beta}{norte} SI - \sigma E\\ \frac{dI}{} &= \sigma E - \gamma I\ dt\\ \frac{dR}{dt} &= \gamma I - vR.\end{aligned} \tag{11}$$

El parámetro adicional es σ = 1/τ. Los modelos de esta forma suelen denominarse modelos SEIRS. Tiene el mismo DFE y R0 que el modelo sin retardo.

2.2. Modelos con diferentes inmunidades

Los tres modelos presentados en la Sección 2.1 tienen diferentes tipos de inmunidad. El modelo SIS no tiene inmunidad, el modelo SIR tiene inmunidad total y permanente y el modelo SIRS tiene inmunidad total pero temporal. Sin embargo, para diferentes infecciones pueden existir diferentes tipos de inmunidad, como parcial y permanente o parcial y temporal. Algunos artículos que consideran la inmunidad a la influenza son [29–31]. Árbitro. [32] incluye una clase reinfectada con inmunidad diferente. Árbitro. [33] modela la inmunidad parcial de la población recuperada y [34] incluye que los recuperados pierden inmunidad pero no todos los infecciosos se recuperan. Árbitro. [35] añade el efecto de la reinfección de la gripe. Árbitro. [36] también tiene en cuenta la disminución de la inmunidad.
Árbitro. [37] modela la pérdida de inmunidad con tasas dependientes del tiempo. Consideraremos varios modelos nuevos con inmunidad diferente basados en el modelo SIRS. El primer modelo supone que todos los individuos infecciosos obtienen inmunidad parcial pero permanente. Modelos similares están dados por [34,38]. Los susceptibles se infectan y luego pasan a ser susceptibles con inmunidad parcial Sp , que pueden reinfectarse con una tasa de infección βp/N. Las ecuaciones del modelo son

$$\begin{aligned}\frac{dS}{dt} &= -\frac{b}{} SI\\ \frac{dI}{dt} &= \frac{norte\,\beta\,\beta_p}{norte} SI + \frac{}{norte} Sp\ I - \gamma I\\ \frac{dSp}{dt} &= -\frac{\beta p}{norte} Sp\ I + \gamma I.\end{aligned} \tag{12}$$

Dado que la población total es constante, se puede eliminar una ecuación usando Sp(t) = N − S(t) − I(t). El sistema tiene dos soluciones de equilibrio: el DFE I = 0, S + Sp = N, γN y S = 0, Sp = I = N − Sp. Al inicio de la infección, Sp = 0, por lo que la ecuación para $\beta\overline{p}$, I es la misma que en el sistema (6), y R0 tiene el mismo valor R0 = β/γ. Los dos modelos de retardo se construyen como antes. El primer modelo es

$$\begin{aligned}\frac{dS}{dt} &= -\frac{b}{} S(t-\tau)I(t-\tau)\\ \frac{dI}{dt} &= \frac{norte\,\beta_p\,\beta}{} S(t-\tau)I(t-\tau) + \frac{}{norte} Sp(t-\tau)I(t-\tau) - \gamma I\\ \frac{dSP}{dt} &= -\frac{norte\,\beta p}{norte} Sp(t-\tau)I(t-\tau) + \gamma I,\end{aligned} \tag{13}$$

Machine Translated by Google

y el segundo es

$$\begin{aligned}\frac{dS}{dt} &= -\frac{b}{N}S(t)I(t-\tau)\\ \frac{dI}{dt} &= \frac{\beta_p \beta}{N}S(t)I(t-\tau) + \frac{\beta_p \beta}{N}Sp(t)I(t-\tau) - \gamma I\\ \frac{dSp}{dt} &= -\frac{\beta_p}{N}Sp(t)I(t-\tau) + \gamma I.\end{aligned} \tag{14}$$

Tienen los mismos puntos de equilibrio que el modelo sin retrasos. El modelo con la población expuesta se construye de manera similar y viene dado por

$$\begin{aligned}\frac{dS}{dt} &= -\frac{b}{N}SI\\ \frac{dE}{dt} &= \frac{\beta_p \beta}{N}SI + \frac{\beta_p \beta}{N}Sp\,I - \sigma E\\ \frac{dI}{dt} &= \sigma E - \gamma I\\ \frac{dSp}{dt} &= -\frac{\beta_p}{N}Sp\,I + \gamma I.\end{aligned} \tag{15}$$

Estos tres modelos también tienen el mismo R0 que (12).

El segundo modelo de inmunidad parcial supone que todos los infecciosos obtienen inmunidad parcial pero que es sólo temporal. Los susceptibles se infectan, luego pasan a los susceptibles con inmunidad parcial y luego regresan a los susceptibles sin inmunidad. El modelo viene dado por el siguiente sistema de EDO

$$\begin{aligned}\frac{dS}{dt} &= -\frac{b}{N}SI + vSp\\ \frac{dI}{dt} &= \frac{\beta_p \beta}{N}SI + \frac{\beta_p \beta}{N}Sp\,I - \gamma I\\ \frac{dSp}{dt} &= -\frac{\beta_p}{N}Sp\,I + \gamma I - vSp.\end{aligned} \tag{16}$$

Nuevamente, dado que la población total es constante, se puede eliminar una ecuación. El sistema tiene dos soluciones de equilibrio: el DFE I = 0, S = N, Sp = 0, y un segundo equilibrio con S > 0, Sp > 0 e I > 0. Al inicio de la infección, Sp = 0, por lo que el La ecuación para I es la misma que la del sistema (6) y R0 = β/γ. Los dos modelos de retardo se construyen de manera similar a antes. Tienen los mismos puntos de equilibrio que el modelo no retrasado. El modelo con la población expuesta tiene esta población agregada de la misma forma. Estos tres modelos también tienen el mismo R0 que (12).

El tercer modelo supone que todos los infecciosos obtienen inmunidad total pero que es sólo temporal, y luego obtienen inmunidad parcial permanente. Los susceptibles se infectan, luego se recuperan con inmunidad temporal total y luego obtienen inmunidad permanente parcial como susceptibles con inmunidad parcial. El modelo viene dado por

$$\begin{aligned}\frac{dS}{dt} &= -\frac{b}{N}SI\\ \frac{dI}{dt} &= \frac{\beta_p \beta}{N}SI + \frac{\beta_p \beta}{N}Sp\,I - \gamma I\\ \frac{dR}{dt} &= \gamma I - vR\\ \frac{dSP}{dt} &= -\frac{\beta_p}{N}Sp\,I + vR.\end{aligned} \tag{17}$$

Machine Translated by Google

Nuevamente, dado que la población total es constante, se puede eliminar una ecuación. El sistema tiene dos soluciones de equilibrio: el DFE I = 0, S + Sp = N, y un segundo equilibrio (βp−γ)Nv
S = 0, Sp = y R = N− $\frac{\gamma N}{\beta p}$, yo = $\frac{\nu(\beta p-\gamma)}{\beta p(\gamma+\nu)}$. Esta segunda solución de equilibrio βp(γ+v) sólo tiene sentido para βp > γ. Al inicio de la infección, Sp = 0, por lo que la ecuación para I es la misma que la del sistema (6) y R0 = β/γ. Sin embargo, si βp < γ, la epidemia eventualmente desaparecerá.

El cuarto modelo supone que todos los infecciosos obtienen inmunidad total pero que es sólo temporal , y luego obtienen inmunidad parcial que también es temporal. Los susceptibles se infectan, luego se recuperan con inmunidad temporal total, luego obtienen inmunidad temporal parcial como susceptibles con inmunidad parcial, que luego se convierten en susceptibles. El modelo viene dado por

$$\begin{aligned}\frac{dS}{dt} &= -\frac{b}{}SI + v1Sp\\ \frac{dI}{dt} &= \frac{\beta}{norte}SI + \frac{\beta p}{norte}Sp\ I - \gamma I\\ \frac{dR}{dt} &= \gamma I - \nu R\\ \frac{dSP}{dt} &= -\frac{\beta p}{norte}Sp\ I + \nu R - v1Sp.\end{aligned} \tag{18}$$

Nuevamente, dado que la población total es constante, se puede eliminar una ecuación. El sistema tiene dos soluciones de equilibrio: el DFE I = 0, S + Sp = N y un segundo equilibrio endémico S > 0, Sp > 0, I > 0 y R = N − Sp − I. Al inicio de la infección, Sp = 0, por lo que la ecuación diferencial para I es la misma que la del sistema (6) y R0 = β/γ.

2.3. Modelos con pérdida de inmunidad y dos grupos

Diferentes individuos pierden inmunidad a diferentes ritmos. Para simplificar, en esta subsección incluiremos dos grupos de individuos, pero los modelos se pueden extender a más grupos.
Árbitro. [34] tiene un modelo con algunos recuperados perdiendo inmunidad. El primer modelo propuesto supone que todos los infectados obtienen inmunidad permanente pero los dos grupos la pierden a ritmos diferentes, v y v1. Un grupo tiene un sistema inmunológico más fuerte debido, entre otras cosas, a la exposición previa a una variación previa de la influenza o a vacunas en temporadas anteriores [29-31].

$$\begin{aligned}\frac{dS}{dt} &= -\frac{b}{}SI + v\,f\,R + v1(1-f)R\\ \frac{dI}{dt} &= \frac{\beta}{norte}SI - \gamma I\\ \frac{dR}{dt} &= \gamma I - v\,f\,R - v1(1-f)R.\end{aligned} \tag{19}$$

Aquí, 0 ≤ f ≤ 1 es la fracción de la población en el primer grupo. Los dos modelos con retraso en la infección vienen dados por

$$\begin{aligned}\frac{dS}{dt} &= -\frac{b}{}S(t-\tau)I(t-\tau) + v\,f\,R + v1(1-f)R\\ \frac{dI}{dt} &= \frac{\beta}{norte}S(t-\tau)I(t-\tau) - \gamma I\\ \frac{dR}{dt} &= \gamma I - v\,f\,R - v1(1-f)R,\end{aligned} \tag{20}$$

Machine Translated by Google

y

$$\begin{aligned}\frac{dS}{dt} &= -\frac{b}{}S(t)I(t-\tau) + v\,f\,R + v1(1-f)R\\ \frac{dI}{dt} &= \frac{norte\,\beta}{norte}S(t)I(t-\tau) - \gamma I\\ \frac{dR}{dt} &= \gamma I - v\,f\,R - v1(1-f)R.\end{aligned} \tag{21}$$

Para ambos modelos de retraso, la ecuación diferencial al inicio del proceso de infección para el compartimento infectivo es la misma que la del modelo SIS (1), por lo que se puede utilizar el mismo procedimiento para determinar que R0 = β/γ. Además, para ambos modelos los puntos de equilibrio son los mismos que los del modelo ODE (19).

El modelo con una población expuesta es

$$\begin{aligned}\frac{dS}{dt} &= -\frac{b}{}SI + v\,f\,R + v1(1-f)R\\ \frac{dE}{dt} &= \frac{norte\,\beta}{norte}SI - \sigma E\\ \frac{dI}{dt} &= \sigma E - \gamma I\\ \frac{dR}{dt} &= \gamma I - v\,f\,R - v1(1-f)R.\end{aligned} \tag{22}$$

Tiene el mismo DFE que (19). La estabilidad del DFE al inicio de la infección todavía está dada por R0 = β/γ.

En el siguiente modelo, una fracción f de los infecciosos obtiene inmunidad total y permanente, y el resto obtiene inmunidad parcial y permanente. Las EDO del modelo son

$$\begin{aligned}\frac{dS}{dt} &= -\frac{b}{}SI\\ \frac{dI}{dt} &= \frac{norte\,\beta}{norte}SI + \frac{\beta p}{norte}Sp\,I - \gamma I\\ \frac{dR}{dt} &= \gamma\,f\,I\\ \frac{dSP}{dt} &= -\frac{\beta p}{norte}Sp\,I + \gamma(1-f)I.\end{aligned} \tag{23}$$

El siguiente modelo supone que una fracción f obtiene inmunidad total y temporal y el resto obtiene inmunidad parcial y temporal. El sistema de ecuaciones es

$$\begin{aligned}\frac{dS}{dt} &= -\frac{b}{}SI + vR + v1Sp\\ \frac{dI}{dt} &= \frac{norte\,\beta p\,\beta}{norte}SI + \frac{}{norte}Sp\,I - \gamma I\\ \frac{dR}{dt} &= \gamma\,f\,I - vR\\ \frac{dSp}{dt} &= -\frac{\beta p}{norte}Sp\,I + \gamma(1-f)I - v1Sp.\end{aligned} \tag{24}$$

En el siguiente modelo, todos los infecciosos obtienen inmunidad total durante un tiempo finito, luego la pierden en diferentes tasas v y v1, y obtienen inmunidad parcial y permanente con diferentes tasas de infección β1 y β2.

Machine Translated by Google

$$\begin{aligned}
\frac{dS}{dt} &= -\frac{b}{norte} SI \\
\frac{dI}{dt} &= \frac{norte\ \beta\ \beta 1\ \beta 2}{norte} SI + \frac{}{norte} T1\ yo + \frac{}{norte} S2\ I - \gamma I \\
\frac{dR}{dt} &= \gamma I - v\ f\ R - v1(1 - f)R \\
\frac{dS1}{dt} &= -\frac{\beta 1}{\beta 2} S1\ I + v\ f\ RN \\
\frac{dS2}{dt} &= -\frac{S2}{norte} I + v1(1 - f)R.
\end{aligned} \tag{25}$$

2.4. Modelos con vacunación y pérdida de inmunidad

La vacunación es una herramienta muy eficaz para lograr inmunidad a la gripe estacional. Generalmente el efecto de la vacuna disminuye con el tiempo. Consideramos que la inmunidad es permanente si se prolonga hasta la próxima temporada de gripe con una nueva variante del virus. La población vacunada se indicará con V. Existen diferentes formas de introducir la tasa de vacunación. La primera forma es que es proporcional al número de susceptibles S [39-41]. Árbitros. [42-44] también consideran la disminución de la vacuna. Árbitro. [45] tiene un modelo de vacunación básico basado en el modelo SIR. Una segunda forma es suponer que sólo un porcentaje de los susceptibles está dispuesto a vacunarse [46], lo cual es una suposición más realista [47]. Esto conduce a una tasa de vacunación de la forma r(prS − V). También puede considerarse proporcional a la población total o simplemente constante [48,49]. Árbitro. [50] considera la vacunación, la inmunidad parcial y la reinfección.

Árbitros. [40,51] también incluyen la cuarentena. Nuestro primer modelo con vacunación supone que los infecciosos se recuperan y obtienen inmunidad total y temporal durante un período de tiempo promedio 1/v. Los individuos vacunados obtienen inmunidad total, pero disminuye con el tiempo a un ritmo ω. El modelo es

$$\begin{aligned}
\frac{dS}{dt} &= -\frac{b}{norte} SI + \omega V + vR - rS \\
\frac{dV}{dt} &= rS - \omega V \\
\frac{dI}{dt} &= \frac{b}{norte} SI - \gamma I \\
\frac{dR}{dt} &= \gamma I - vR.
\end{aligned} \tag{26}$$

En este modelo la tasa de vacunación es proporcional al número de susceptibles. Dado que la ecuación para la población infecciosa es la misma que para (6), el valor de R0 sigue siendo β/γ. Si las vacunas no disminuyen, ω = 0, el sistema tiene un solo punto de equilibrio, S = 0, I = 0, R = 0, V = N. Si ω = 0 hay dos puntos de equilibrio, I = 0, R = 0, S = (ωN/(ω + r), V = rS/ω y S = γN/β, V = rS/ω, R = γI/v e I dada por la solución de I = N − S − V − R. El primer modelo de retardo se obtiene reemplazando S(t) e I(t) en los términos de infección por S(t − τ) e I(t − τ). El segundo reemplazando solo I(). t) término Ambos tienen los mismos puntos de equilibrio y R0 que (26). El sistema con una población expuesta tiene los mismos puntos de equilibrio DFE ya que I = 0 implica E = 0. El R0 calculado por el método matricial de próxima generación sigue siendo R0. = β/γ.

El segundo modelo con vacunación considera que la población infectante obtiene inmunidad total durante un tiempo determinado y luego obtiene inmunidad parcial permanente. La población vacunada obtiene una inmunidad total pero temporal, y luego decae hasta llegar a una inmunidad permanente parcial. Hay un nuevo compartimento, Sp, que representa a los susceptibles con inmunidad parcial.

Machine Translated by Google

$$\begin{aligned}
\frac{dS}{dt} &= -\frac{b}{norte} SI - rS \\
\frac{dSp}{dt} &= -\frac{\beta p}{norte} Sp\, I + \omega V + vR \\
\frac{dV}{dt} &= rS - \omega V \\
\frac{dI}{dt} &= \frac{\beta}{norte} SI + \frac{\beta p}{norte} Sp\, I - \gamma I \\
\frac{dR}{dt} &= \gamma I - vR.
\end{aligned} \tag{27}$$

Este modelo tiene un DFE I = 0, S = 0, R = 0, V = 0, Sp = N. También hay un equilibrio endémico con S = 0, V = 0. Ya que al inicio de la infección, Sp = 0, usando la ecuación diferencial para I, R0 = β/γ. Los dos modelos de retardo se obtienen modificando (27) de forma similar a como se realizó para los modelos anteriores. Los estados de equilibrio y R0 son los mismos que los de (27). La ecuación correspondiente con una población expuesta también tiene el mismo R0.

Si la tasa de vacunación se modifica a r(prS − V) en lugar de rS, para que el modelo concuerde con que no todos quieren vacunarse, el modelo modificado es

$$\begin{aligned}
\frac{dS}{dt} &= -\frac{b}{norte} SI - r(prS - V) \\
\frac{dSp}{dt} &= -\frac{norte\ \beta p}{norte} Sp\, I + \omega V + vR \\
\frac{dV}{dt} &= r(prS - V) - \omega V \\
\frac{dI}{dt} &= \frac{b}{norte} SI + \frac{\beta p}{norte} Sp\, I - \gamma I \\
\frac{dR}{dt} &= \gamma I - vR.
\end{aligned} \tag{28}$$

Este modelo no tiene una solución estable con I = 0 para valores arbitrarios de los parámetros. Esta es una diferencia importante en la dinámica con las de los otros modelos.

El siguiente modelo considera que todos los vacunados obtienen inmunidad total por un tiempo, luego una fracción fv mantiene la inmunidad total, mientras que el resto obtiene inmunidad parcial permanente. Todos los infecciosos obtienen inmunidad temporal total después de recuperarse.

$$\begin{aligned}
\frac{dS}{dt} &= -\frac{b}{norte} SI - r(prS - V) + vR \\
\frac{dSp}{dt} &= -\frac{norte\ \beta p}{norte} Sp\, I + \omega(1 - fv)V \\
\frac{dV}{dt} &= r(prS - V) - \omega(1 - fv)V \\
\frac{dI}{dt} &= \frac{b}{norte} SI + \frac{\beta p}{norte} Sp\, I - \gamma I \\
\frac{dR}{dt} &= \gamma I - vR.
\end{aligned} \tag{29}$$

Si ω = 0 el DFE es I = 0, R = 0, , Sp = 0, V = prS, S = N/(1 + pr). Para ω = 0, el DFE es I = 0, R = 0, V = 0, S = 0, Sp = N. Como Sp(0) = 0, tenemos para ambos casos R0 = β/γ. Los dos modelos de retraso y los modelos de población expuesta se derivan de (29) como se realizó para los modelos anteriores. Los tres modelos tienen el mismo DFE y R0 que (29).

Machine Translated by Google

Como modelo final y más completo añadiremos el uso de antivirales para reducir la tiempo de infección y aislamiento para reducir el número de infecciosos activos. Hay un nuevo compartimento, la población en cuarentena o aislada Q. Los infecciosos se recuperan con total y inmunidad temporal, los vacunados obtienen inmunidad total y temporal y luego inmunidad permanente parcial. Una fracción de los infectados recibe un antiviral y sale del compartimento. a una velocidad más rápida γa. Una fracción fq de los infectados es puesta en cuarentena a una velocidad γq y sale el estado de cuarentena a una tasa γq1. El tiempo promedio que pasa una persona en cuarentena en los compartimentos infectados y en cuarentena es el mismo que el tiempo que pasa un no en cuarentena gasto individual en el compartimento infeccioso. Por lo tanto, 1/γ = 1/γq + 1/γq1. El el modelo es

$$\begin{aligned}
\frac{dS}{dt} &= -\frac{b}{norte} SI - r(prS - V) + vR \\
\frac{dSP}{dt} &= -\frac{\beta p}{norte} Sp\, I + \omega(1 - fv)V \\
\frac{dV}{dt} &= r(prS - V) - \omega(1 - fv)V \\
\frac{dI}{dt} &= \frac{b}{norte} SI + \frac{\beta p}{norte} sp\ yo \\
&\quad - fq\gamma q\, I - fa\gamma a\, I - (1 - fq - fa)\gamma I \\
\frac{dQ}{dt} &= fq\gamma q\, I - \gamma q1Q \\
\frac{dR}{dt} &= fa\gamma a\, I + (1 - fq - fa)\gamma I + \gamma q1Q - vR.
\end{aligned} \tag{30}$$

El DFE es I = 0, R = 0, Q = 0, V = 0, S = 0, Sp = N. Como Sp(0) = 0, para el infeccioso compartimento para crecer necesitamos $R0 = \frac{b}{(1-fq-fa)\gamma+fq\gamma q+fa\gamma a} > 1$. Las versiones retardadas y las El modelo con población expuesta tiene el mismo DFE y R0.

3. Resultados

En esta sección, presentamos resultados numéricos para los modelos dados en la sección anterior. sección. Las simulaciones numéricas, incluido el DDE, se realizaron utilizando el método julia paquete DifferentialEqns.jl [52,53]. Los valores de los parámetros utilizados para la mayoría de las simulaciones son dado en la Tabla 1. Si se utiliza un valor diferente para una simulación particular, se anotará en el título de la figura correspondiente. Los valores de β, γ, v y τ se basan en aquellos dado [54] y CDC [1] correspondiente a la influenza estacional. Todos los demás son estimados. El valor de β tiene una gran variación dependiendo de las condiciones: rural, urbana, suburbana, lugares de trabajo de alta densidad, etc. Los valores de β y γ elegidos para la mayoría de las simulaciones dan R0 = 1,5. Árbitro. [55] también proporciona valores de parámetros similares para la influenza.

Tabla 1. Esta tabla tiene los valores de los parámetros utilizados en las simulaciones.

Símbolo β	Descripción del parámetro	Valor
γ	Tasa de infección	0,5/día
	Tasa de recuperación	1/3/día
v	Tasa de pérdida de inmunidad.	1/30/día
τ	Retraso en el tiempo de infección	1,9 días
σ	1/(tiempo de infección)	1/τ1/d
norte	Población total	1000
βp	Tasa de infección por inmunidad parcial	b/2
v1	Tasa de pérdida o inmunidad parcial	1/60/día
β1	Fracción que pierde inmunidad a una tasa v	0,75
	Tasa de infección para la fracción fr	0,9β/día
β2	Tasa de infección para la fracción 1 − fr	0,8β/día

Machine Translated by Google

Tabla 1. Cont.

Símbolo	Descripción del parámetro	Valor
fp	Fracción que logra inmunidad total	0,5
r	Tasa de vacunación	0,02/día
ω	Tasa de pérdida de inmunidad a la vacunación.	1/60/día
βv	Tasa de infección de vacunados con parcial.	0,8β
pr	Fracción de susceptibles a vacunar	0,2/día
fv	Fracción de vacunados con inmunidad permanente	0,75

Las condiciones iniciales para todos los modelos son S = 900, I = 100 y todas las demás poblaciones igual a cero. El historial de los modelos de retraso es S = 1000, y todas las demás variables retrasadas igual a cero en el intervalo [−τ, 0]. Esto corresponde a una epidemia que comienza en t = 0. Todos Las simulaciones, excepto donde se especifica, tienen R0 = 1,5.

3.1. Modelos SIS, SIR y SIRS

Los modelos SIS, SIR y SIRS son ampliamente utilizados para epidemias con infección por vía directa. contacto. En el modelo SIS, los infecciosos no obtienen inmunidad, en el modelo SIR obtienen inmunidad total. e inmunidad permanente y en el modelo SIRS obtienen inmunidad total pero temporal.

La Figura 1 presenta los gráficos para el modelo SIS, a la izquierda para la población susceptible. y en el derecho de la población contagiada. Como R0 = β/γ > 1 la solución tiende a estado endémico. La gráfica infectiva para el modelo con la población expuesta tiene un significado diferente. límite ya que hay dos clases infectivas, I y E, y solo se traza I. Las soluciones tienen La misma forma, pero la del modelo ODE alcanza el estado de equilibrio más rápido. El Las soluciones de los dos sistemas de retardo son todas idénticas, ya que la única diferencia es cuando se produce un nuevo El infectado deja de contarse como susceptible. El modelo con la población expuesta tiene un tiempo promedio en el compartimiento expuesto igual al retraso pero no el mismo retraso para cada individuo. Se necesitan datos reales para determinar qué modelo es el mejor, ya que todos los modelos, sin demora, con demora y con población expuesta, tienen el mismo dinámica cualitativa.

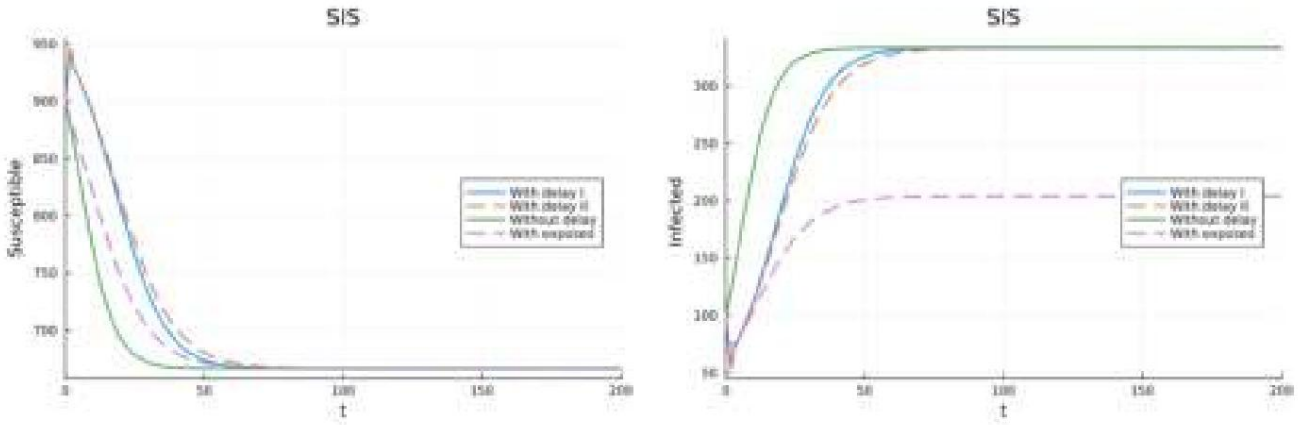


Figura 1. Modelo SIS. (izquierda) Población susceptible. (derecha) Población infecciosa.

La Figura 2 presenta los gráficos para el modelo SIR, a la izquierda para la población susceptible. y para los infecciosos en el derecho. Tenga en cuenta que para I = 0, existen múltiples soluciones S + R = N. Las soluciones simuladas dependen de los valores iniciales. El modelo de retraso que uso la historia. tanto para S como para I, mientras que todos los demás modelos solo usan S(0).

Machine Translated by Google

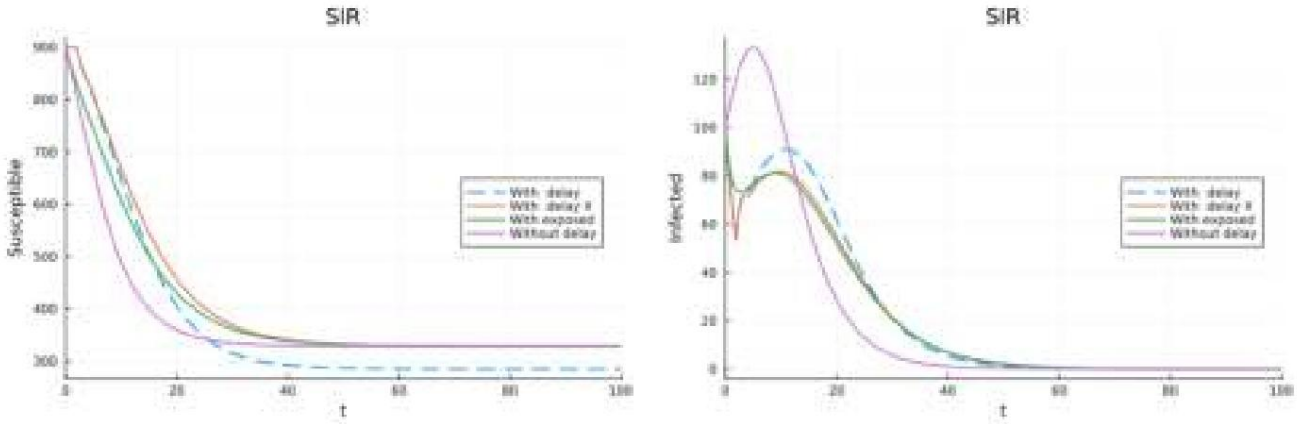


Figura 2. Modelo SIR. (izquierda) Población susceptible. (derecha) Población infecciosa.

La Figura 3 presenta los gráficos de las poblaciones susceptibles (izquierda) e infecciosas (derecha) para el modelo SIR. Este modelo tiene inmunidad total pero temporal. Como era de esperar, las soluciones simuladas tienden al equilibrio endémico. El modelo ODE da el pico de infección más alto, pero todos los modelos tienen la misma dinámica cualitativa.

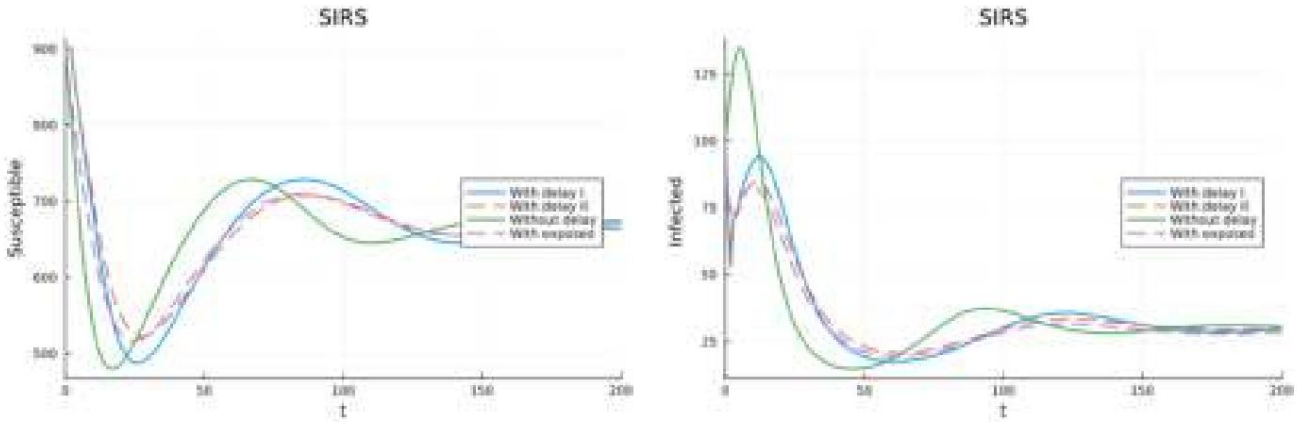


Figura 3. Modelo SIR. (izquierda) Población susceptible. (derecha) Población infecciosa.

3.2. Modelos con diferentes inmunidades

En esta subsección, consideramos modelos con diferentes inmunidades: total, parcial, permanente y temporal en diferentes combinaciones. El primer modelo (12) considera que todos los infecciosos después de recuperarse obtienen inmunidad parcial y permanente. La Figura 4 tiene las soluciones simuladas para la población infecciosa para dos valores diferentes de la tasa de infección para la población con inmunidad parcial, βp. A la izquierda para βp = 0.25 y la solución tiende a I = 0 a pesar de tener R0 > 1, y a la derecha para βp = 0.4 y la solución tiende al estado endémico. Sin embargo, que la epidemia termine o no depende de los valores de Spβp + Sβ − Nγ que aparece en el lado derecho de la ecuación diferencial para la población infectante. Los dos modelos de retardo se comportan de manera diferente al inicio ya que en el modelo con retardo I tanto S como I dependen del histórico, mientras que para el modelo con retardo II solo I depende del histórico.

El siguiente modelo (16) considera que todos los infecciosos se recuperan con inmunidad parcial y temporal. La Figura 5 tiene los gráficos para la población infecciosa, a la izquierda para βp = 0,25, y I tiende a un valor pequeño pero distinto de cero, y a la derecha para βp = 0,4, y I tiende a un valor mucho mayor, lo que muestra el efecto. de β.

Machine Translated by Google

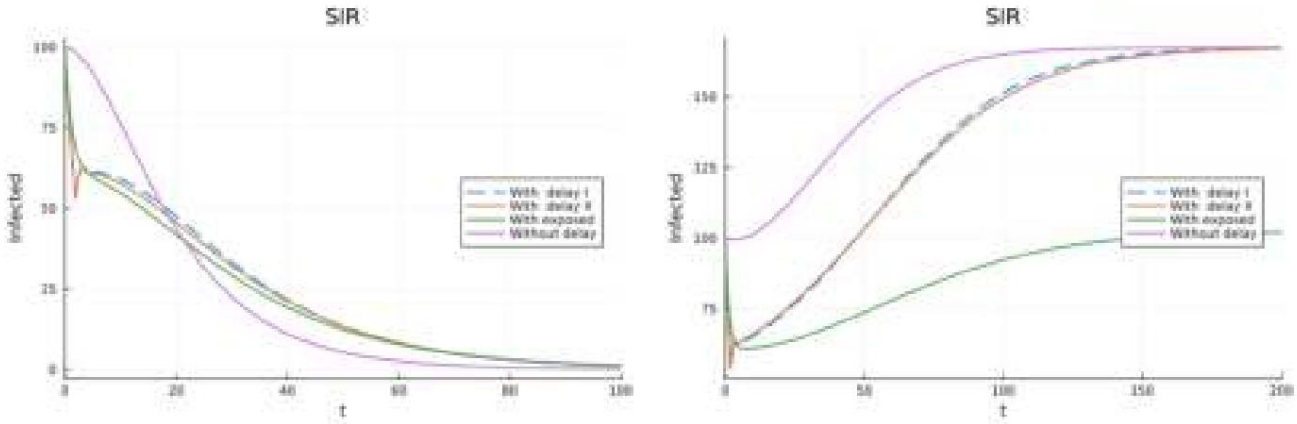


Figura 4. Modelo con inmunidad parcial y permanente (izquierda) βp = 0,25. (derecha) βp = 0,4.

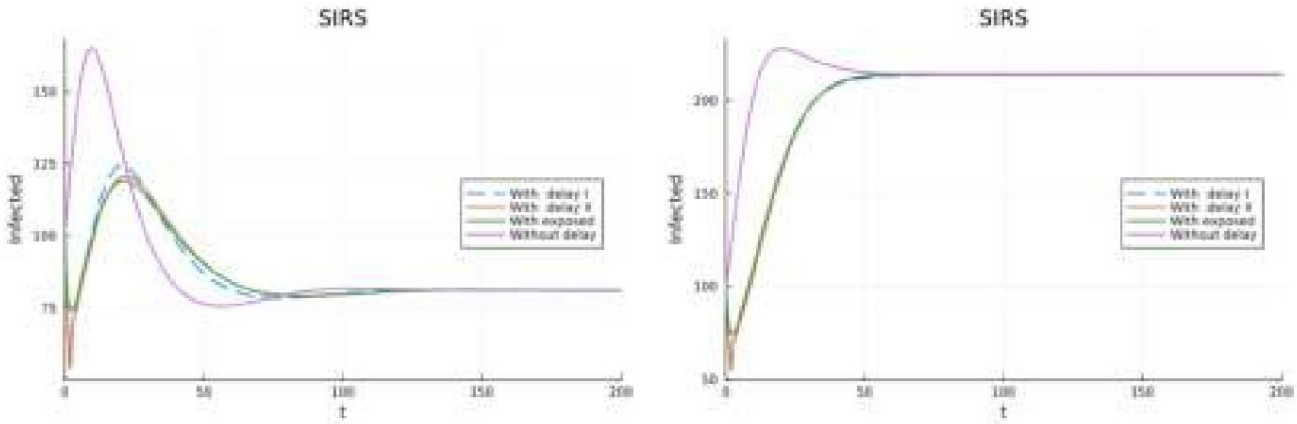


Figura 5. Modelo con inmunidad parcial y temporal (izquierda) βp = 0,25. (derecha) βp = 0,4.

El modelo (17) es para una infección con inmunidad total pero temporal y luego inmunidad parcial y permanente. La Figura 6 tiene los gráficos simulados para la población infecciosa para dos valores de βp. La cifra de la izquierda es para βp = 0,255 e I tiende al DFE; la cifra de la derecha es para βp = 0,4 y I va a un equilibrio endémico. Para ambas simulaciones R0 > 1 , inicialmente el número de infecciosos crece. Sin embargo, que la epidemia termine o no depende de los valores de Spβp + Sβ − Nγ. Todos los modelos tienen los mismos equilibrios.

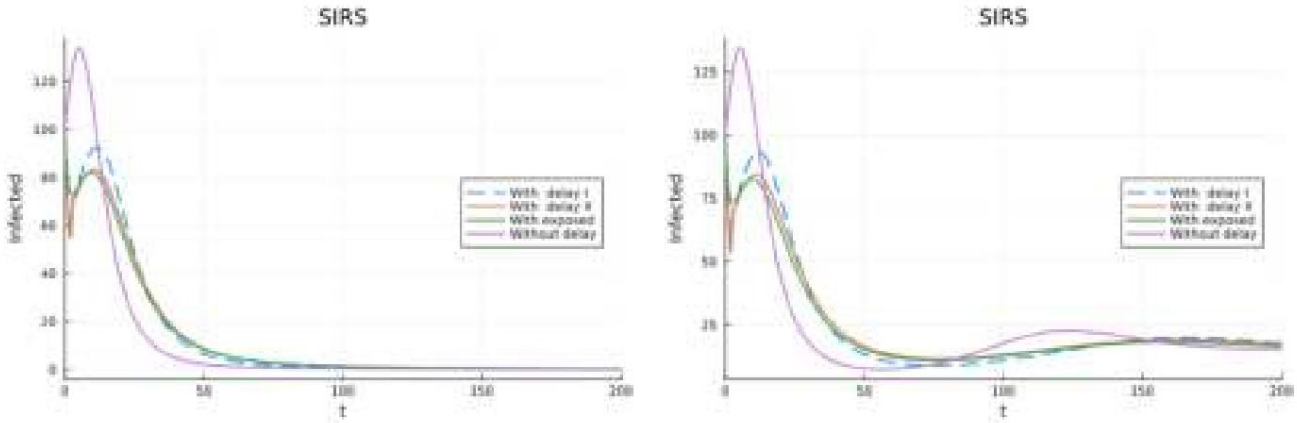


Figura 6. Modelo con inmunidad temporal total y luego parcial permanente (izquierda) βp = 0,25. (derecha) βp = 0,4.

El siguiente modelo (18) supone que después de ser infectado un individuo obtiene inmunidad total que es temporal, y luego inmunidad parcial que también es temporal, perdiendo finalmente todo.

Machine Translated by Google

inmunidad. La Figura 7 muestra la solución temporal de la población infectante. A la izquierda para βp = β/2 y a la derecha para βp = 0,4. En ambos casos la epidemia persiste. El valor límite para βp = 0,4 es ligeramente mayor. Dado que todos los individuos eventualmente pierden completamente la inmunidad, la epidemia no termina.

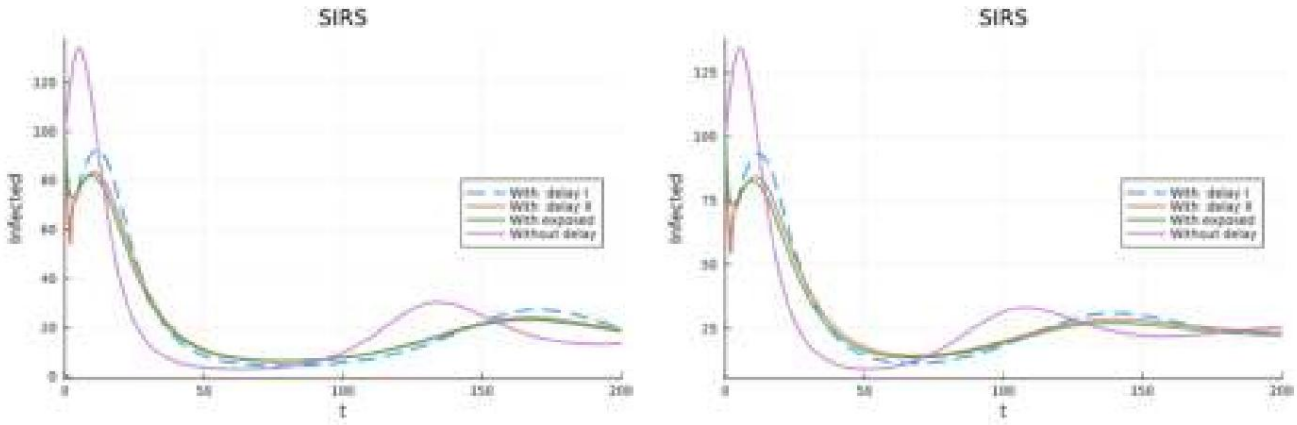


Figura 7. Modelo con inmunidad temporal total y luego parcial (izquierda) βp = 0,25. (derecha) βp = 0,4.

3.3. Modelos con pérdida de inmunidad y dos grupos

Diferentes individuos tienen diferentes reacciones del sistema inmunológico ante una infección. Por ejemplo, todos los individuos pueden obtener inmunidad total a la infección después de recuperarse, pero pueden perderla a diferentes ritmos. Algunas de las razones de la influenza estacional son que los niños han estado expuestos a menos variantes del virus, que las personas mayores tienen sistemas inmunológicos más débiles o que algunas personas pueden tener otros factores de riesgo, como fumar.

Por simplicidad consideramos sólo dos clases de individuos, pero los modelos pueden extenderse a un número arbitrario.

El primer modelo con dos poblaciones distintas (19) considera que los dos grupos obtienen inmunidad total y temporal después de la infección (el supuesto SIR habitual), pero pierden la inmunidad a diferentes tasas v y v1, con fracciones fr y (1 − fr), respectivamente. Las simulaciones se encuentran en la Figura 8. Ambas figuras tienen R0 > 1, la izquierda con β = 0,5 y la derecha con β = 0,666. El valor límite de la figura de la derecha es aproximadamente el doble que el de la izquierda, lo que muestra la gran influencia que tiene la tasa de infección.

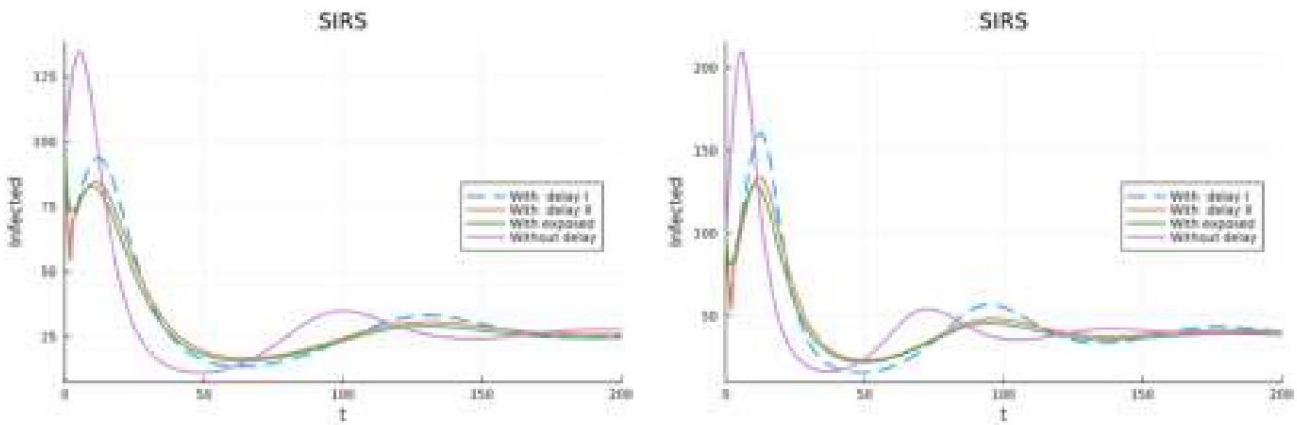


Figura 8. Modelo con dos poblaciones con inmunidad temporal total que salen del compartimento R a diferentes velocidades. (izquierda) β = 0,5. (derecha) β = 0,666

El siguiente modelo (23) supone que una clase obtiene inmunidad total y permanente, mientras que la segunda obtiene inmunidad parcial y también permanente. La fracción de individuos en ambos

Machine Translated by Google

clases se toma como fp = 0,5. Para la clase con inmunidad parcial βp = β/2. Dado que todos los individuos obtienen inmunidad permanente, se espera que la epidemia no persista.
La Figura 9 tiene dos simulaciones para diferentes valores de β. El de la izquierda para β = 0,5 y el de la derecha para β = 1,5. Incluso para el segundo valor bastante grande de β, que corresponde a R0 = 4,5, la epidemia desaparece.

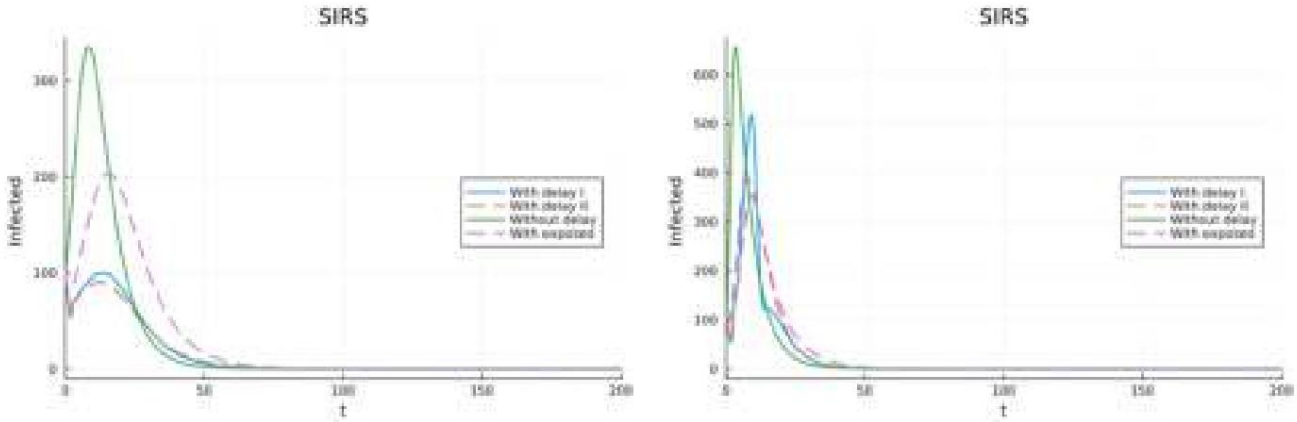


Figura 9. Simulaciones del modelo con dos clases, una con inmunidad total y permanente y otra con inmunidad parcial y permanente. (izquierda) β = 0,5. (derecha) β = 4,5

A continuación consideramos que la mitad de las poblaciones infectantes obtienen inmunidad total y temporal, y la otra mitad inmunidad parcial y también temporal (24). Se realizaron simulaciones para dos valores de βp, la tasa de infección de la población con inmunidad parcial, βp = β/2 (gráfico izquierdo de la Figura 10) y βp = β/8 (gráfico derecho). En ambos casos la epidemia persiste y el valor límite de I tiene un cambio relativamente pequeño. El hecho de que todos los individuos obtengan inmunidad sólo por un tiempo finito es la razón por la que la infección persiste.

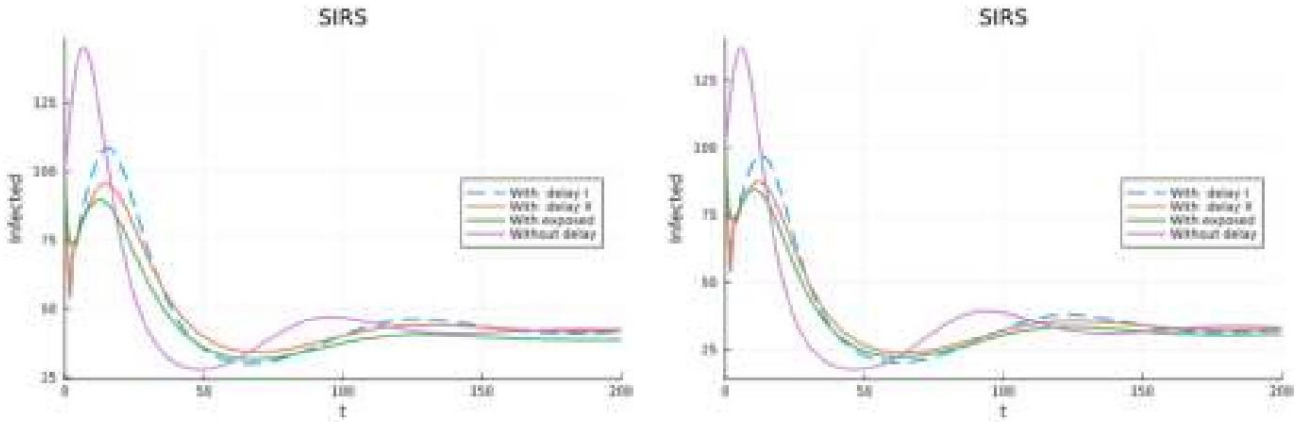


Figura 10. Simulaciones del modelo con dos clases, una con inmunidad total y otra con inmunidad parcial, ambas temporales. (izquierda) βp = β/2. (derecha) βp = β/8.

El modelo (25) supone que todos los infecciosos se recuperan con inmunidad temporal total y luego obtienen inmunidad parcial y permanente. Una fracción fr = 0,75 tiene inmunidad parcial con una tasa de infección β1 = 0,9β y una tasa de pérdida de inmunidad permanente v = 1/30 y el resto tiene una tasa de infección β2 = 0,8β y una tasa de pérdida de inmunidad v1 = 1/60 (gráfico izquierdo de la Figura 11). La solución tiende al equilibrio endémico. Se realizó una segunda simulación cambiando las tasas de infección a β1 = 0,45β y β2 = 0,4β (mitad derecha de la Figura 11). Ahora la solución tiende al DFE. En la segunda simulación, β1/γ < 1 y β2/γ < 1, cuando la mayoría de los individuos recuperados tienen inmunidad permanente parcial, la ecuación diferencial para I muestra que I disminuye.

Machine Translated by Google

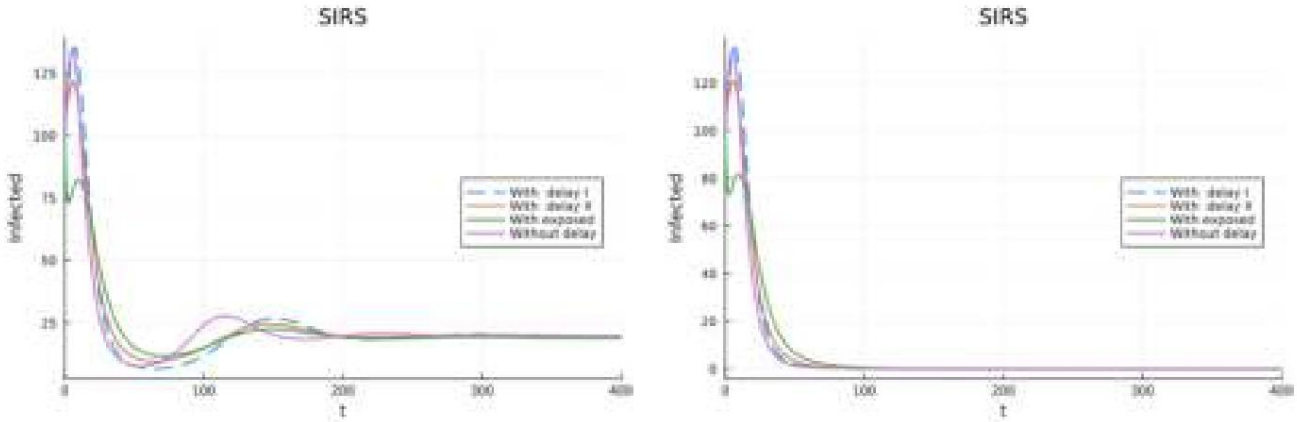


Figura 11. Modelo con dos clases de población. Ambos obtienen inmunidad total (compartimento R) y luego la pierden para obtener inmunidad parcial a diferentes tasas con diferentes tasas de infección. (izquierda) β1 = 0,9β, β2 = 0,8β. (derecha) β1 = 0,45β, β2 = 0,4β.

3.4. Modelos con vacunación y pérdida de inmunidad

Las vacunas pueden ayudar a controlar la influenza estacional. Dado que el virus de la gripe muta de una temporada a otra, es difícil predecir cuál es la eficacia de la vacuna. También es difícil determinar si la vacuna proporciona inmunidad total o parcial, o si el efecto disminuye. Es posible que una persona vacunada no contraiga influenza debido a la vacuna o simplemente porque no estuvo en contacto cercano con una persona infecciosa. De igual forma, que una persona se infecte una sola vez puede deberse a que ha adquirido inmunidad total y permanente, o a no haber tenido contactos estrechos con infecciosos. En esta subsección presentamos simulaciones realizadas para diferentes escenarios, todos con vacunación.

El primer escenario representa individuos que se vacunan a un ritmo rS y alcanzan una inmunidad total. La inmunidad a infectarse con la enfermedad es total y desaparece a un ritmo v, (26). En la parte izquierda de la Figura 12, el efecto de la vacuna disminuye a una tasa ω = 1/60. En la parte derecha, el efecto de las vacunas no disminuye, ω = 0. En
En ambos casos la solución tiende al DFE. En el primer caso también hay una segunda endemia.
equilibrio.

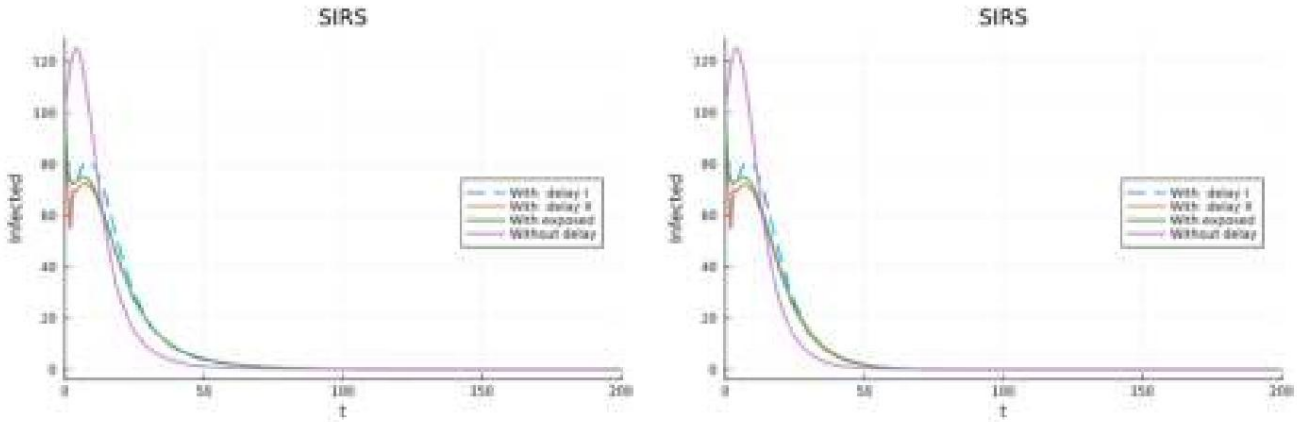


Figura 12. Modelo con vacunación que otorga inmunidad total. (izquierda) La inmunidad no disminuye. (derecha) La inmunidad disminuye.

El modelo del siguiente escenario es similar al modelo dado por (26), excepto que cuando la vacuna disminuye, el individuo aún obtiene inmunidad permanente parcial con βv = 0,8β. La Figura 13 a la izquierda muestra el caso con ω = 1/60 y la solución simulada casi llega al DFE pero luego rebota. A la derecha, la gráfica corresponde a ω = 0 y

Machine Translated by Google

las tendencias de solución al DFE. La figura muestra la importancia que tiene la disminución de la vacuna.

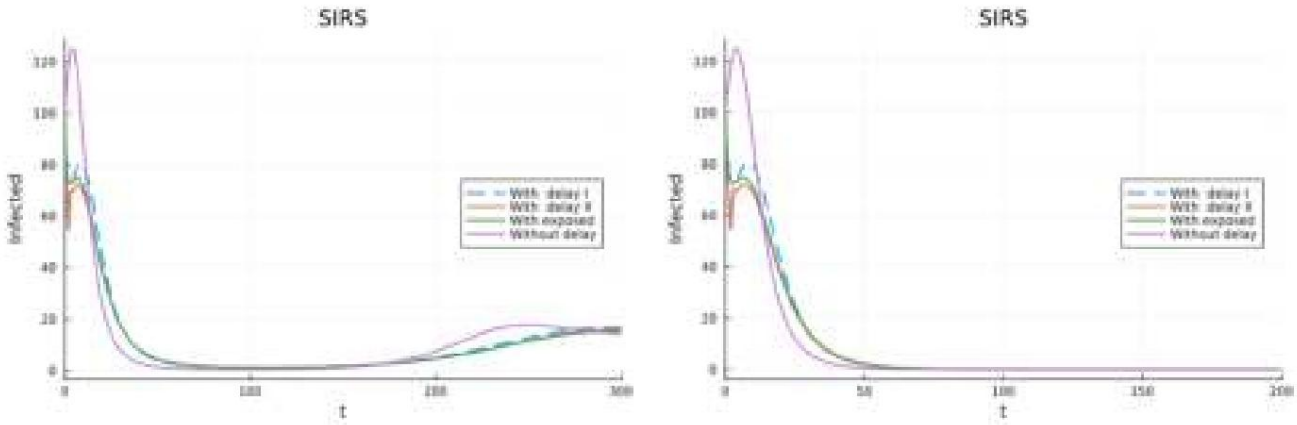


Figura 13. Modelo con vacunación. Los individuos infectados obtienen inmunidad total que disminuye hasta convertirse en inmunidad parcial y permanente. Las personas vacunadas obtienen una inmunidad total que puede disminuir o no. (Izquierda) La inmunidad a la vacunación disminuye hasta convertirse en permanente parcial. (derecha) La inmunidad a la vacunación no disminuye.

Las siguientes simulaciones se realizaron con el modelo (28). La tasa de vacunación es r(prS − V). Para este ritmo de vacunación no existe DFE. Los infectados se recuperan y luego obtienen inmunidad permanente parcial. Los vacunados obtienen una inmunidad total que, en una simulación, disminuye hasta convertirse en parcial permanente, y en la segunda, no disminuye. Usamos pr = 0,4 y la tasa de infección para la población con inmunidad parcial βv = 0,8β. El gráfico de la izquierda de la Figura 14 corresponde a la simulación con una disminución de la vacunación y el de la derecha, sin disminución. En ambos gráficos la solución infectante se acerca a cero y luego rebota.

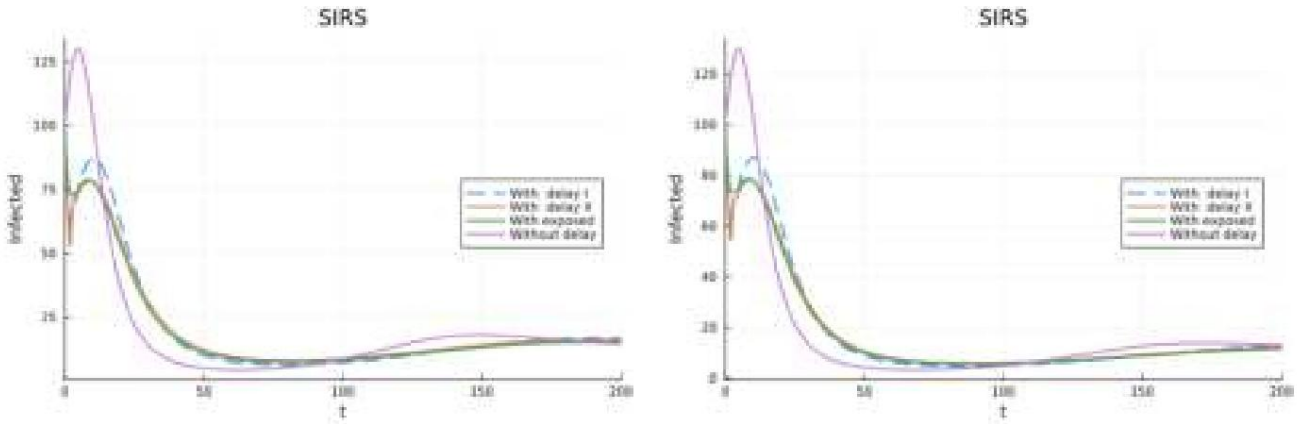


Figura 14. Modelo con recuperados obteniendo inmunidad permanente parcial y vacunados obteniendo inmunidad total que en un caso disminuye y en el otro no. (izquierda) La inmunidad a la vacunación disminuye. (derecha) La inmunidad a la vacunación no disminuye.

Las siguientes simulaciones se realizan con el modelo dado por (29). La tasa de vacunación es r(prS − V) con pr = 0,6 y una fracción fv = 0,75 de la población vacunada obtiene inmunidad total y permanente mientras que una fracción 1 − fr obtiene inmunidad total que disminuye a un ritmo ω hasta inmunidad parcial y permanente con tasa de infección βv = 0,8β. La inmunidad frente a una infección es total y temporal. La solución calculada tiende a un equilibrio endémico. Se realizó una segunda simulación sin disminuir el efecto de la vacuna ω = 0. La solución tiende al DFE. Los gráficos se encuentran en la Figura 15. Este ejemplo muestra la importancia de la disminución o no de la vacuna.

Machine Translated by Google

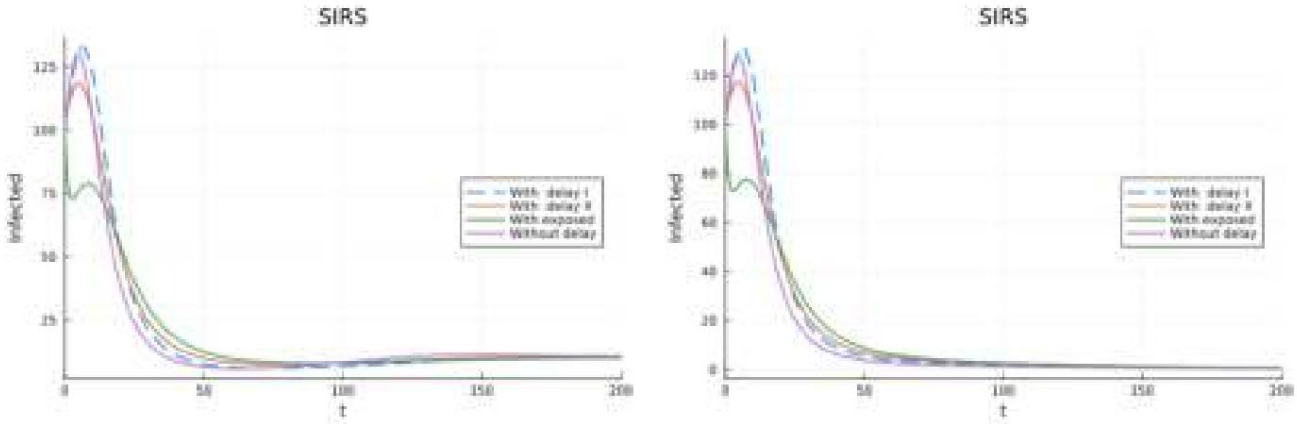


Figura 15. Modelo con dos clases de individuos vacunados, uno con inmunidad total y permanente y otro con inmunidad total que decrece hasta inmunidad parcial y permanente. (izquierda) Menguante inmunidad a la vacunación. (derecha) Inmunidad vacunal permanente.

El modelo final considera la vacunación, el uso de antivirales y el aislamiento. La tasa de vacunación es r(prS − V), una fracción fv = 0,75 de los vacunados mantiene la inmunidad total para siempre y el resto pierde la inmunidad total y luego obtiene inmunidad parcial permanente con una tasa de infección βv = 0,8β. Los infectados se recuperan con inmunidad total y temporal.

A una fracción fa = 0,2 se le administran antivirales y se elimina de la clase infectiva a una tasa γa = 2γ. Se elimina una fracción fq = 0,4 de los infecciosos y se aísla a una velocidad γq = 1.

Se realizaron dos simulaciones. El primero con β = 0,5, que corresponde a R0 = 1,5, se muestra en el gráfico de la izquierda de la Figura 16 y tiene una solución que tiende al DFE. El segundo, con β = 1 que da R0 = 3, se muestra en el gráfico de la derecha y la solución tiende a un equilibrio endémico. La figura muestra que, si bien R0 predice si una epidemia comienza o no, no indica si eventualmente desaparecerá o no.

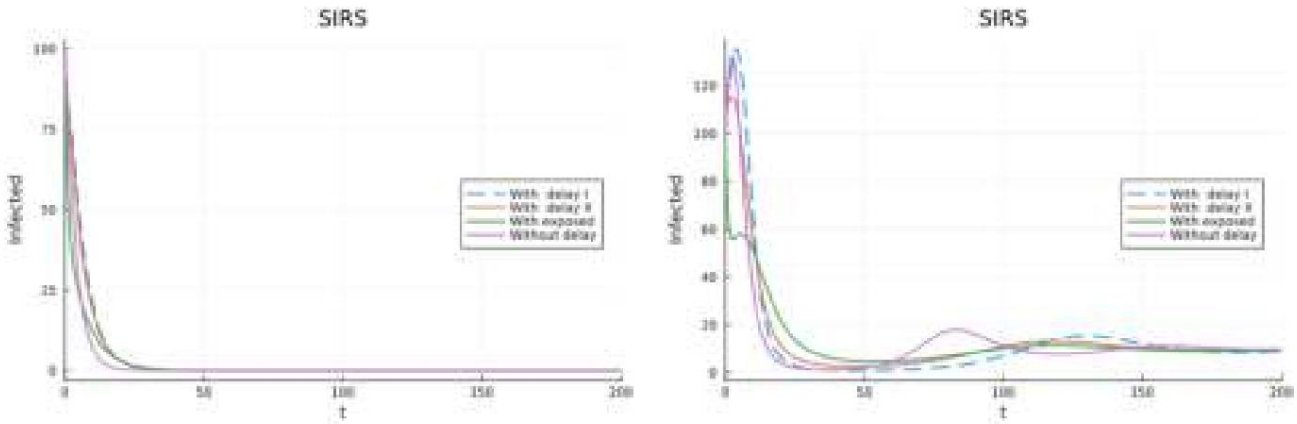


Figura 16. Modelo con vacunación, antivirales y aislamiento (30). (izquierda) β = 0,5. (derecha) β = 1.

4. Discusión

Se han utilizado modelos basados en SIR para muchas enfermedades infecciosas transmitidas por contacto directo. En este artículo, utilizamos modelos sin datos demográficos para infecciones de corta duración como la influenza estacional. La primera razón es que la infección tiene una duración corta y que no es muy mortal. Por tanto, el número de nacimientos y muertes durante el período de infección es pequeño. La segunda razón es que nos permite concentrarnos en los efectos de la pérdida de inmunidad. El número de reproducción básico R0 es un índice muy importante que da un criterio para determinar si habrá una epidemia o no. Por lo general, se calcula como lo hicimos en este artículo asumiendo que al inicio de la infección hay un infectante y N − 1 susceptibles con N la población total. Para la mayoría de nuestros ejemplos

Machine Translated by Google

es R0 = βS(0)/(γN) y simplificado a R0 = β/γ. Si también hay un tiempo inicial del número de infección de individuos vacunados V(0) > 0, entonces R0 = β(N − V(0))/(γN). Y, si por ejemplo sin vacunación R0 = 1,5, vacunando a más de 1/3 de la población total entonces R0 < 1 y no hay epidemia. Se puede realizar una estimación similar si inicialmente también hay un número de individuos Sp > 0 con inmunidad parcial y tasa de infección βp < β.

Los modelos presentados parten de los simples SIS, SIR y SIRS, que no tienen inmunidad, inmunidad total y permanente, e inmunidad total y temporal. Primero se ampliaron para incluir otras situaciones con combinaciones de inmunidad total y parcial, tanto permanente como temporal. Para enfermedades como la influenza estacional es muy difícil determinar si un individuo infectado y recuperado recibió inmunidad total o parcial después de la infección. O si esa inmunidad desaparece o no. Es posible que el individuo no vuelva a infectarse debido a inmunidad total o inmunidad parcial o simplemente porque el individuo no estuvo en contacto cercano con ningún infectante. Los diferentes modelos permiten ejecutar escenarios con diferentes hipótesis y comparar los resultados. A continuación se consideraron dos grupos de población con diferentes reacciones inmunes y se ejecutaron diferentes escenarios. Finalmente, se introdujo la vacunación con los diferentes modelos con diferentes supuestos y el modelo final siendo más completo con antivirales y aislamiento para reducir la infección. Para los distintos modelos que tenían EFD se realizaron simulaciones de dónde murió la epidemia y dónde no. También se estudió el efecto de introducir un retraso en el momento de la infección. El retraso se introdujo de dos maneras diferentes utilizando DDE y también agregando un compartimento expuesto. Aparte de que las soluciones retrasadas tardan más en alcanzar los estados de equilibrio, el comportamiento es muy similar al del modelo ODE correspondiente. Es decir, producen soluciones con los mismos estados de equilibrio y el mismo número de reproducción básico. Sin embargo, agregar el retraso hace que el modelo sea más realista y se espera que los modelos retrasados se ajusten a las condiciones reales.

datos mejor.

Los parámetros relacionados con el tiempo entre el contacto y la infección y con la duración de la infección son bien conocidos para la gripe estacional. Las tasas de infección se pueden estimar a partir del número de infecciosos. Varían de una ciudad a otra y de un grupo de población a otro. Como se mencionó anteriormente, los parámetros relacionados con la inmunidad parcial y los efectos de la vacunación son difíciles de estimar y no se informan. Las simulaciones numéricas que utilizan diferentes modelos y las comparan con datos epidémicos reales pueden ayudar a determinar las hipótesis más realistas.

Dado que las dos versiones de retardo de cada modelo tienen ecuaciones características diferentes, sus comportamientos de solución pueden ser diferentes entre sí y diferentes del comportamiento de las soluciones del modelo no retardado. Sin embargo, para los valores de los parámetros utilizados en las simulaciones eso no sucedió y la dinámica cualitativa fue toda similar. Determinar si un individuo tiene inmunidad total o parcial o no ha entrado en contacto estrecho no es tarea fácil. Propusimos diferentes modelos con diferentes tipos de inmunidad. Se necesitan comparaciones con datos reales para establecer cuál es más realista.

El trabajo futuro considerará diferentes condiciones iniciales, la introducción de aleatoriedad y dependencia espacial. El uso de otros compartimentos como los asintomáticos es otra vía.

Financiamiento: Esta investigación no recibió financiamiento externo.

Declaración de la Junta de Revisión Institucional: No aplicable.

Declaración de Consentimiento Informado: No aplicable.

Declaración de disponibilidad de datos: No se crearon ni analizaron nuevos datos en este estudio. El intercambio de datos no se aplica a este artículo.

Conflictos de intereses: Los autores declaran no tener conflictos de intereses.

Machine Translated by Google

Abreviaturas

En este manuscrito se utilizan las siguientes abreviaturas:

SIS	Susceptible, infeccioso, susceptible
SIR	Susceptible, infeccioso, recuperado
SRIS	Susceptible, infeccioso, recuperado, susceptible
ODE	Ecuación diferencial ordinaria
DDE	Ecuación diferencial de retardo
DFE	Equilibrio libre de enfermedades
R0	Número de reproducción básico

Referencias

1. Centros para el Control de Enfermedades. Influenza (gripe). 2024. Disponible en línea: https://www.cdc.gov/flu/about/keyfacts.htm#:~:text=Gripe%20es%20a%20contagiosa%20respiratoria,a%20gripe%20vacuna%20cada%20año (consultado el 24 de abril de 2024).
2. Centros para el Control de Enfermedades. Influenza. 2024. Disponible en línea: https://www.cdc.gov/nchs/fastats/flu.htm (consultado el 24 abril de 2024).
3. Ross, R.; Hudson, HP Una aplicación de la teoría de las probabilidades al estudio de la patometría a priori—Parte III. Proc. R. Soc. Londres. Ser. Un contener. Papilla. Matemáticas. Física. Carácter 1917, 93, 225–240.
4. Kermack, WO; McKendrick, AG Una contribución a la teoría matemática de las epidemias. Proc. R. Soc. Londres. Ser. Un contener. Papilla. Matemáticas. Física. Carácter 1927, 115, 700–721.
5. Kuang, Y. Ecuaciones diferenciales de retardo: con aplicaciones en dinámica de poblaciones; Prensa académica: Cambridge, MA, EE. UU., 1993.
Bellan, A.; Zennaro, M. Métodos numéricos para ecuaciones diferenciales de retardo; Oxford University Press: Oxford, Reino Unido, 2013. 6.
7. Ma, Z. Modelado dinámico y análisis de epidemias; Científico mundial: Singapur, 2009.
8. Rihan, FA Ecuaciones diferenciales de retardo y aplicaciones a la biología; Springer: Singapur, 2021.
9. Diekmann, O.; Heesterbeek, J.; Roberts, MG La construcción de matrices de próxima generación para modelos epidémicos compartimentales. JR Soc. Interfaz 2010, 7, 873–885. [Referencia cruzada] [PubMed]
10. Van den Driessche, P. Números de reproducción de modelos de enfermedades infecciosas. Infectar. Dis. Modelo. 2017, 2, 288–303. [Referencia cruzada] [PubMed]
11. Edlund, S.; Kaufman, J.; Menosler, J.; Douglas, J.; Bromberg, M.; Kaufman, Z.; Bassal, R.; Chodick, G.; Marom, R.; Shalev, V.; et al. Comparando tres modelos básicos para la influenza estacional. Epidemias 2011, 3, 135-142. [Referencia cruzada]
12. Van den Driessche, P.; Watmough, J. Números de reproducción y equilibrios endémicos subumbrales para modelos compartimentales de transmisión de enfermedades. Matemáticas. Biosci. 2002, 180, 29–48. [Referencia cruzada] [PubMed]
13. Allen, L. Introducción a la biología matemática; Pearson-Prentice Hall: Hoboken, Nueva Jersey, EE. UU., 2007.
14. Edelstein-Keshet, L. Modelos matemáticos en biología; SIAM: Filadelfia, PA, EE. UU., 2005.
15. Murray, JD Biología Matemática I. Introducción; Springer: Berlín/Heidelberg, Alemania, 2002.
16. Fages, F.; Soliman, S. Sobre el cálculo y la optimización de la robustez en BIOCHAM-4. En actas de la Conferencia internacional sobre métodos computacionales en Systems Biolog, Brno, República Checa, 12 a 14 de septiembre de 2018; Springer: Berlín/Heidelberg, Alemania, 2018; págs. 292-299.
17. Cooke, KL; Yorke, JA Algunas ecuaciones que modelan procesos de crecimiento y epidemias de gonorrea. Matemáticas. Biosci. 1973, 16, 75-101. [Referencia cruzada]
18. Jackson, M.; Chen-Charpentier, BM Modelado de la propagación de virus en plantas con retrasos. J. Computación. Aplica. Matemáticas. 2017, 309, 611–621. [Referencia cruzada]
19. Khan, QJA; Krishnan, EV Un modelo epidémico con un retraso en la transmisión. Aplica. Matemáticas. 2003, 48, 193–203. [Referencia cruzada]
20. Al Basir, F.; Takeuchi, Y.; Ray, SA Dinámica de un modelo de enfermedad vegetal retardada con transmisión de la enfermedad de Bedington-DeAngelis. Matemáticas. Biosci. Ing. 2021, 18, 583–599. [Referencia cruzada]
21. Huang, G.; Mamá, W.; Takeuchi, Y. Análisis global del modelo de dinámica de virus de retraso con respuesta funcional de Bedington-DeAngelis. Aplica. Matemáticas. Letón. 2011, 24, 1199–1203. [Referencia cruzada]
22. Wei, HM; Li, XZ; Martcheva, M. Un modelo epidémico de una enfermedad transmitida por vectores con transmisión directa y retraso en el tiempo. J. Matemáticas. Anal. Aplica. 2008, 342, 895–908. [Referencia cruzada]
23. Chan, M.; Johansson, MA Los períodos de incubación de los virus del dengue. PLoS ONE 2012, 7, e50972. [Referencia cruzada] [PubMed]
24. Jeger, M.; Madden, L.; Van Den Bosch, F. Epidemiología de virus vegetales: aplicaciones y perspectivas para el modelado matemático y análisis para mejorar la comprensión y el control de enfermedades. Desinfección de plantas. 2018, 102, 837–854. [Referencia cruzada]
25. Liu, L. Un modelo SIR retrasado con tasa de incidencia general no lineal. Adv. Diferir de. Ecuación. 2015, 2015, 329. [Referencia cruzada]
26. Hethcote, HW Las matemáticas de las enfermedades infecciosas. SIAM Rev. 2000, 42, 599–653. [Referencia cruzada]
27. Wang, N.; Fu, Y.; Zhang, H.; Shi, H. Una evaluación de modelos matemáticos para el brote de COVID-19. Resumen. Clínico. Medicina. 2020, 3, 85–93. [Referencia cruzada] [PubMed]
28. Bjørnstad, ON; Shea, K.; Krzywinski, M.; Altman, N. El modelo SEIRS para la dinámica de enfermedades infecciosas. Nat. Métodos 2020, 17, 557–559. [Referencia cruzada]

Machine Translated by Google

29. Krammer, F. La respuesta de los anticuerpos humanos a la infección y vacunación por el virus de la influenza A. Nat. Rev. Immunol. 2019, 19, 383–397. [Referencia cruzada]
30. Patel, MM; York, IA; Monto, AS; Thompson, MG; Fry, AM Atenuación inmunomediada de la enfermedad gripal después de la infección: Oportunidades y desafíos. Lancet Microbe 2021, 2, e715 – e725. [Referencia cruzada]
31. Han, AX; de Jong, SP; Russell, CA Coevolución de la inmunidad y los virus de la influenza estacional. Nat. Rev. Microbiol. 2023, 21, 805–817. [Referencia cruzada] [PubMed]
32. Rodríguez-Maroto, G.; Atienza-Diez, I.; Ares, S.; Manrubia, S. Estrategias de vacunación en poblaciones estructuradas bajo régimen parcial inmunidad y reinfección. J. Física. Una Matemática. Teor. 2023, 56, 204003. [Referencia cruzada]
33. Katriel, G. Epidemias con inmunidad parcial a la reinfección. Matemáticas. Biosci. 2010, 228, 153-159. [Referencia cruzada] [PubMed]
34. Saha, P.; Mondal, B.; Ghosh, U. Comportamientos dinámicos de un modelo epidémico con inmunidad parcial que tiene incidencia no lineal y tratamiento saturado en entornos deterministas y estocásticos. Fractales de Solitones del Caos 2023, 174, 113775. [CrossRef]
35. Thieme, recursos humanos; Yang, J. Un modelo endémico con tasa de reinfección variable y aplicaciones a la influenza. Matemáticas. Biosci. 2002, 180, 207–235. [Referencia cruzada] [PubMed]
36. Carlsson, RM; Niños, LM; Feng, Z.; Glasser, JW; Heffernan, JM; Li, J.; Röst, G. Modelado de la disminución y el aumento de la inmunidad por infección o vacunación. J. Theor. Biol. 2020, 497, 110265. [Referencia cruzada] [PubMed]
37. Bhattacharya, S.; Adler, FR Un modelo de tiempo transcurrido desde la recuperación con diferentes tasas de pérdida de inmunidad. Toro. Matemáticas. Biol. 2012, 74, 2810–2819. [Referencia cruzada]
38. Viguerie, A.; Carletti, M.; Silvestri, G.; Veneziani, A. Modelado matemático de brotes periódicos con inmunidad menguante: A Posible descripción a largo plazo del COVID-19. Matemáticas 2023, 11, 4918. [CrossRef]
39. Alejandro, YO; Bowman, C.; Moghadas, SM; Veranos, R.; Gumel, AB; Sahai, BM Un modelo de vacunación para la transmisión Dinámica de la gripe. SIAM J. Appl. Din. Sistema. 2004, 3, 503–524. [Referencia cruzada]
40. Sol, C.; Yang, W. Resultados globales de un modelo SIRS con vacunación y aislamiento. Anal no lineal. Aplicaciones del mundo real. 2010, 11, 4223–4237. [Referencia cruzada]
41. Sahu, médico de cabecera; Dhar, J. Análisis de un modelo epidémico SVEIS con inmunidad temporal parcial y tasa de incidencia de saturación. Aplica. Matemáticas. Modelo. 2012, 36, 908–923. [Referencia cruzada]
42. Moghadas, S.; Gumel, A. Un estudio matemático de un modelo para enfermedades infantiles con inmunidad no permanente. J. Computación. Aplica. Matemáticas. 2003, 157, 347–363. [Referencia cruzada]
43. Saad-Roy, CM; Morris, SE; Metcalf, CJE; Mina, MJ; Panadero, RE; Farrar, J.; Holmes, CE; Pybus, OG; Graham, AL; Levin, SA; et al. Consideraciones epidemiológicas y evolutivas de los regímenes de dosificación de la vacuna SARS-CoV-2. Ciencia 2021, 372, 363–370. [Referencia cruzada]
44. Goldberg, Y.; Mandel, M.; Bar-On, YM; Bodenheimer, O.; Freedman, L.; Haas, EJ; Milón, R.; Alroy-Preis, S.; Ceniza, N.; Huppert, A. Inmunidad menguante después de la vacuna BNT162b2 en Israel. N. inglés. J. Med. 2021, 385, e85. [Referencia cruzada] [PubMed]
45. Turkyilmazoglu, M. Un modelo epidémico extendido con vacunación: SIRVI débilmente inmune. Física. Una estadística. Mec. Su aplicación. 2022, 598, 127429. [Referencia cruzada] [PubMed]
46. Ho, SH; Él, D.; Eftimie, R. Modelos matemáticos de dinámica de transmisión y estrategias de vacunas en Hong Kong durante el Temporada de influenza invernal 2017-2018. J. Theor. Biol. 2019, 476, 74–94. [Referencia cruzada] [PubMed]
47. Reluga, TC; Medlock, J. Los mecanismos de resistencia importan en los modelos SIR. Matemáticas. Biosci. Ing. 2007, 4, 553–563. [PubMed]
48. Scherer, A.; McLean, A. Modelos matemáticos de vacunación. Hno. Medicina. Toro. 2002, 62, 187–199. [Referencia cruzada] [PubMed]
49. Safan, M.; Kretzschmar, M.; Hadeler, KP Control de infecciones basado en vacunación en modelos SIRS con reinfección: Especial referencia a la tos ferina. J. Matemáticas. Biol. 2013, 67, 1083–1110. [Referencia cruzada]
50. Krithika, B.; Tamilalagan, P. Revelando la propagación de epidemias que involucran inmunidad parcial y reinfección: ideas de un modelo matemático retardado en el tiempo. euros. Física. J. especificaciones. Arriba. 2023, 232, 2657–2673. [Referencia cruzada]
51. Arif, F.; Majeed, Z.; Rahman, JU; Iqbal, N.; Kafle, J. Modelado matemático y simulación numérica para el brote de COVID-19 que implica pérdida de inmunidad y clase en cuarentena. Computadora. Matemáticas. Métodos Med. 2022, 2022, 3816492. [Referencia cruzada] [PubMed]
52. Rackauckas, C.; Nie, Q. DifferentialEquations.jl: un ecosistema eficaz y rico en funciones para resolver ecuaciones diferenciales en Julia. J. Res. abierta. Software. 2017, 5. Disponible en línea: https://app.dimensions.aion2019/05/05 (consultado el 24 de abril de 2024). [Referencia cruzada]
53. Widmann, D.; Rackauckas, C. DelayDiffEq: Generación de solucionadores de ecuaciones diferenciales de retardo mediante la integración recursiva de solucionadores de ecuaciones diferenciales ordinarios. arXiv 2022, arXiv:2208.12879.
54. Colizza, V.; Barrat, A.; Bartolomé, M.; Vallerón, AJ; Vespignani, A. Modelización de la propagación mundial de la gripe pandémica: casos de referencia e intervenciones de contención. PLoS Med. 2007, 4, e13. [Referencia cruzada]
55. Kanyiri, CW; Marcos, K.; Luboobi, L. Análisis matemático de la dinámica de la influenza A en la aparición de resistencia a los medicamentos. Computadora. Matemáticas. Métodos Med. 2018, 2018, 2434560. [Referencia cruzada]

Descargo de responsabilidad/Nota del editor: Las declaraciones, opiniones y datos contenidos en todas las publicaciones son únicamente de los autores y contribuyentes individuales y no de MDPI ni de los editores. MDPI y/o los editores renuncian a toda responsabilidad por cualquier daño a personas o propiedad que resulte de cualquier idea, método, instrucción o producto mencionado en el contenido.

Machine Translated by Google

AppliedMath

MDPI

Artículo

Un algoritmo eficiente para funciones matriciales elementales básicas con precisión y aplicación específicas

Huizeng Qin [1] y Youmin Lu [2,*]

[1] Escuela de Matemáticas y Estadística, Universidad Tecnológica de Shandong, Zibo 255049, China; qinhz@sdut.edu.cn
[2] Departamento de Ciencias Matemáticas y Digitales, Universidad de Bloomsburg, Bloomsburg, PA 17815, EE. UU. * Correspondencia: ylu@bloomu.edu

Resumen: Si la función matricial f(At) posee las propiedades de f(At) = g $\frac{t}{k}a$, entonces el f(fórmula de recurrencia f i−1 = g(f i), i = N, N − 1, · · · , 1, f (tA) = f0, aquí se puede establecer. $fN = f(AN) = \sum ajA\ j=0^{metro}\ \frac{j}{norte}$, un = $\frac{}{k\ SUST.}$A. Esto proporciona un algoritmo para calcular la función matricial. gordo). Al especificar la precisión del cálculo p , se presenta un método para determinar my N de una manera que minimice el tiempo del algoritmo anterior, proporcionando así un algoritmo rápido para f(At). Es importante señalar que m sólo depende de la precisión del cálculo p y es independiente de las matrices A y t. Por lo tanto, f(AN) tiene un formato de cálculo fijo que se calcula fácilmente. Por otro lado, N depende no sólo de A, sino también de t. Esto proporciona un medio para seleccionar t tal que N sea igual a 0, una propiedad significativa. En resumen, el algoritmo propuesto en este artículo permite a los usuarios establecer un nivel deseado de precisión y luego utilizarlo para seleccionar los valores apropiados para my N para minimizar el tiempo de cálculo. Este enfoque garantiza que tanto la precisión como la eficiencia se aborden simultáneamente. Desarrollamos un algoritmo general, luego lo aplicamos a las funciones matriciales exponenciales, trigonométricas y logarítmicas, y comparamos el rendimiento con el de las funciones internas del sistema de Mathematica y la aproximación de Pade. En la última sección, se proporciona un ejemplo para ilustrar el cálculo rápido de soluciones numéricas para ecuaciones diferenciales lineales.

Palabras clave: función matricial; algoritmos eficientes; alta precisión; ecuaciones diferenciales lineales

MSC: 68W25; 68U99

check for updates

Cita: Qin, H.; Lu, Y. Un eficiente Algoritmo para funciones matriciales elementales básicas con precisión y aplicación especificadas. Matemáticas aplicadas 2024, 4, 690–708. https://doi.org/10.3390/appledmath4020037

Recibido: 26 de enero de 2024
Revisado: 4 de marzo de 2024
Aceptado: 19 de abril de 2024
Publicado: 5 de junio de 2024

Copyright: © 2024 por los autores. Licenciatario MDPI, Basilea, Suiza. Este artículo es un artículo de acceso abierto. distribuido bajo los términos y condiciones de los Creative Commons Licencia de atribución (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

1. Introducción

Las funciones de exponentes matriciales desempeñan un papel crucial en varios campos, incluidas las ecuaciones diferenciales ordinarias y la teoría de control. En particular, son fundamentales para analizar sistemas dinámicos estructurales generales, representados por el problema del valor inicial:

$$Mx'' + Gx' + Kx = r(t), x(0) = x0, x'(0) = x1, \quad (1)$$

donde M, G y K son matrices cuadradas de n por n, y r(t) denota la fuerza externa aplicada. La introducción de variables duales de un sistema hamiltoniano,

$$p = Mx' + Gx/2, \quad (2)$$

da como resultado el siguiente sistema:

$$U' = HU + F(t), U(0) = U0, \quad (3)$$

dónde

Machine Translated by Google

$$tu = \quad , H = \quad \begin{matrix} -M-1G/2 \\ GM\text{-}1G/4 - K\text{-}M\text{-}1G/2 \end{matrix} \quad \begin{matrix} M\text{-}1 \\ , \end{matrix}$$
$$qp \qquad\qquad 0$$
$$F(t) = \quad r(t) , q = x, U0 = \quad \begin{matrix} x0 \\ Mx1 + Gx0/2 \end{matrix} \quad . \tag{4}$$

La solución de (3) se proporciona de la siguiente manera:

$$U(t) = e\ HtU0 + \int_0^t mi\ H(t-s)F(s)ds. \tag{5}$$

Cuando un problema es oscilatorio, la solución formal a menudo involucra tanto el seno como el coseno de una matriz. Ejemplos famosos incluyen la ecuación de Schrödinger en mecánica cuántica :

$$yo\frac{d\psi}{dt} = H(t)\psi,\ \psi(t0) = \psi0, \tag{6}$$

donde H(t) es un operador hermitiano y ψ es una función de onda compleja. Si H(t) = A es una matriz real y constante, la solución de (6) es la siguiente:

$$\psi(t) = e^{-iAt}\psi0 = (\cos(tA) - i\sin(tA))\psi0. \tag{7}$$

Hay otros ejemplos en los que el cálculo del seno y el coseno de una matriz puede resultar interesante. Las ecuaciones de onda proporcionadas por el sistema genérico de segundo orden se escriben de la siguiente manera:

$$x'' + Kx = F(t),\ x(0) = x0,\ x'(0) = x1, \tag{8}$$

donde la solución la proporciona la variación del método constante:

$$x(t) = \cos(tA)x0 + sen1(tA)x1 + \int_0^t pecado1((t-s)A)F(s)ds, \tag{9}$$

donde $A = \sqrt{K}$, $\overline{sen}1(tA) = \int_0^t \cos(sA)ds$. Si A es invertible, entonces (9) se puede escribir como sigue:

$$x(t) = \cos(tA)x0 + sen(tA)A^{-1}x1 + A^{-1}\int_0^t pecado((t-s)A)F(s)ds. \tag{10}$$

Para (5), proporcionar U(kτ)(k = 1, 2, · · ·) con la mayor precisión posible es una preocupación para los ingenieros y matemáticos, donde τ se denomina paso. El método específico para este problema es el siguiente:

$$Reino\ Unido = TUk-1 + \int_0^T e\ HsF(k\tau - s)ds,\ k = 1, 2, \cdots,\ T = mi^{H\tau}. \tag{11}$$

De manera similar, para (10), proporcionar el algoritmo x(kτ)(k = 1, 2, · · ·) con la mayor precisión posible implica introducir vectores $x = y$ y el siguiente algoritmo:

$$\begin{matrix} xk = \cos(\tau A)xk-1 + \sin(\tau A)A^{-1}yk-1 + A \\ yk = -A\int_0^\tau \sin(sA)F(k\tau - s)ds, \\ \sin(\tau A)xk-1 + \cos(\tau A)yk-1 + x0 = x(0), \\ y0 = \int_0^\tau \cos(sA)F(k\tau - s)ds, \\ incógnita \quad (0) = x1. \end{matrix} \tag{12}$$

Para mejorar la precisión de U(kτ) y x(kτ), es esencial desarrollar un algoritmo, cos(τA) y con alta precisión para calcular $T = e^{H\tau}$ manteniendo $\sin(\tau A)$ y la integral con-
estas funciones matriciales. Este artículo mejora los algoritmos desarrollados para el cálculo de e y cos A en la literatura existente, ver [1-15], y también analiza el cálculo de otras matrices trigonométricas, hiperbólicas, logarítmicas y triangulares inversas e hiperbólicas inversas relacionadas.

Machine Translated by Google

2. El teorema general del algoritmo rápido para funciones matriciales

En [16], abordamos el desafío de calcular rápidamente funciones elementales con una precisión especificada. El método de implementación implica una función analítica f(x) que especifica la propiedad,

$$f(x) = gf \quad \frac{}{k} \quad , \text{ donde } k > 1. \tag{13}$$

Si C0 es una aproximación lo más cercana posible a f ejemplo, $\frac{}{}$ con N como un entero positivo, para kN ,

$$C0 = \sum j=0^{\text{metro}} \frac{f(j)(0)(\Delta x)^j}{j!}, \Delta x = \frac{}{kN}, m = 1, 2, \cdots \tag{14}$$

o siendo C0 la aproximación de Padé de f formada: $\frac{}{kN}$, entonces se realizan los siguientes cálculos:

$$\text{para}(i = 0, i < N, i + +)\ Ci+1 = g(Ci). \tag{15}$$

Completando (14) y (15), se obtiene f(x) ≈ CN . Este método también es aplicable al cálculo de funciones matriciales, donde la variable x se reemplaza por la matriz cuadrada A de n por n. Ahora, tenemos lo siguiente:

$$C0 = \sum j=0^{\text{metro}} \frac{f(j)(0)(\Delta A)^j}{j!}, \Delta A = \frac{1}{}A, m = 1, 2, \cdots kN \tag{16}$$

o C0 es la aproximación de Padé de f(ΔA). Nuestro objetivo es determinar my N para minimizar el tiempo de cálculo bajo la condición de que el error de truncamiento y la precisión computacional sean 10-p . Para el cálculo de la función matricial, se puede ignorar el cálculo del coeficiente de (14) , centrándose principalmente en el cálculo de la multiplicación de la matriz en (16) y (15). Si el número de multiplicaciones para completar g(A) es K, entonces el número total de veces para completar (16) y (15) es aproximadamente

$$q(m, N) = (m + KN)Mn, \tag{17}$$

donde Mn es el número de veces que se completa una multiplicación de la matriz n por n. Con base en la discusión anterior, presentamos un algoritmo rápido para la función matricial de la siguiente manera.

Teorema 1. Sea mp un número entero cercano al punto mínimo de h(m), donde

$$h(m) = m + + \frac{pK \ln 10}{(m+1)\ln k} \quad \frac{k}{(m+1)\ln k} \ln \frac{f(m+1)(0)}{(m+1)!} \tag{18}$$

y

$$\text{norte} = \frac{p \ln 10}{(m+1)\ln k} \frac{1 + (m+1)\ln k}{} \ln \frac{f(m+1)(0)}{(m+1)!} + \frac{\ln A}{\ln k} + 1. \tag{19}$$

$Ci+1 = g(Ci)\ y\ (A) \frac{1\ (1)\ C0 (j = 0)(\Delta A)^j kNq(2)\ \text{para}(i = 0, i < N, i + +)}{j!}$, donde $\Delta A = \frac{}{}$A, o C0 es una aproximación de Padé de f(ΔA).
≈ CN.

El algoritmo anterior es un algoritmo rápido para la función matricial general.

Prueba. Con un error de truncamiento de 10-p , para que (16) se cumpla,

$$\frac{f(m+1)(0)(\Delta A)^{m+1}}{(m+1)!} = \frac{f(m+1)(0)}{(m+1)!k^{(m+1)N}} \text{ un } m+1 \leq \frac{f(m+1)(0)}{(m+1)!k^{(m+1)N}} A^{m+1} < 10\text{-}p , \tag{20}$$

donde $\|A\|$ es una norma de A. Entonces,

$$\text{norte} > + \frac{p \text{ en } 10}{(m+1)\ln k} \quad \frac{1}{(m+1)\ln k} \text{En} \frac{f(m+1)(0)}{(m+1)!} \ln \frac{}{\|A\|} + \ln k . \tag{21}$$

Machine Translated by Google

De (17), obtenemos lo siguiente:

$$10\ q(m, N) > g(m) = m + + (m+1)\ln k \frac{pK \ln}{(m+1)\ln k} \frac{k}{} \text{En } f \frac{(m+1)(0)}{(m+1)!} + \frac{K \ln A}{\ln k} \text{ Minnesota.} \quad (22)$$

Tenga en cuenta que el punto mínimo de g(m) no tiene nada que ver con $\frac{K \ln A}{\ln k}$. Al hacer que q(m, N) sea lo más pequeño posible, el algoritmo obtenido por (1) y (2) es un algoritmo rápido, completando así la demostración del teorema. □

Observación 1. En el cálculo específico, A se puede reemplazar por el valor propio aproximado con el módulo máximo en este documento.

$$\frac{f(m+1)(0)}{(m+1)!k(m+1)N} \quad A^{m+1} < 10-p. \quad (23)$$

3. Algoritmo rápido para e A, cos A, sin A, sinh A y cosh A

Supongamos que λ = A es una norma para A. Si dejamos A = λA0, entonces A0 ≤ 1. De (19), podemos obtener lo siguiente:

$$\text{norte} = \frac{p \ln 10 +}{(m+1)\ln k} \frac{1}{(m+1)\ln k} \ln f \frac{(m+1)(0)}{(m+1)!} \lambda + \frac{En}{k \ln k} + 1. \quad (24)$$

Para reducir la extensión del artículo, los algoritmos para varias funciones matriciales se compilan en una tabla.

Teorema 2. Para e A, cos A, sin A, cosh A y sinh A, ideamos el siguiente algoritmo rápido utilizando el teorema 1.

Observación 2. Vale la pena señalar que mp en la Tabla 1 solo depende de la precisión del cálculo p y la función de la matriz, y es independiente de la matriz A.

Tabla 1. El algoritmo rápido para e A, cos A, sin A, cosh A y sinh A.

		MP, Np	Algoritmo
una e	p.f. =	$\frac{9p \text{ en } 10}{8 \text{ en } 2}$, Np = [Ne] + 1	(1) B = $\frac{C}{(2)}$ f o (i = 0, i < N, i + +) C = C $\frac{mp\ B\ j\ A}{}$ 2 (3) e = C A
pecado A sinh A	p.f. =	$\frac{p \ln 10}{2 \ln 3}$, Np = [Nsen] + 1	mp (1) B = C = ∑ 3 2Np , i=1 (2j+1)! (2) f o(i = 0, i < N, i + +) C = 3C $\frac{A}{4C \sin A (3) \sinh A}$ + $\frac{A}{3 \text{ norte}}$ $\frac{A}{3 \text{ norte}}$ $\frac{(B)\ j\ 2\ A}{}$ 3 = C
(I) porque A cosh A	p.f. =	$\frac{pag \text{ en}}{10\ 2 \text{ en } 3}$ + 1, Np = [Npecado]	(1) B = C = I + ∑ 3 2Np , i=1 $\frac{A^2}{(2j)!}$ (2) f o (i = 0, i < N, i + +) C = 4C cos A $\frac{mp\ (B)\ j}{\cosh}$ 3 − 3C A (3) = C
(yo yo) porque A cosh A	p.f. =	$\frac{pag \text{ en}}{10\ 2 \text{ en } 3}$ + 1, Np = [Npecado]	(1) B = $\frac{A^2}{(2j)!}$ C = I + ∑ 3 2Np , i=1 (2) f o (i = 0, i < N, i + +) C = C (3) $\frac{\text{diputado}\ (B)\ j}{}$ 2 − yo porque A cosh A = C
	Ne = + (mp+1)ln 2 $\frac{p \text{ en } 10 \text{ en } \lambda}{\ln 2 \text{ ln } 2}$ − $\frac{En(mp+1)!}{(pf+1)\ln 2}$, Ncos = Nsin = + (2mp+2)ln 3 $\frac{p \text{ en } 10}{\text{en } 3}$	$\frac{p \ln 10 \ln \lambda + (2mp+1)\ln}{\ln 2 \ln(2mp+2)! (2mp+2)en}$ − $\frac{2}{\text{en } \lambda}$ − $\frac{En(2mp+1)!}{(2mp+1) \ln 2}$, $\frac{3}{}$.	

Machine Translated by Google

Ahora, utilizamos el algoritmo rápido general para demostrar específicamente la corrección de los algoritmos de la Tabla 1. Los algoritmos de la Tabla 1 son sencillos; implican colocar las funciones específicas f y g en el algoritmo general, y la parte complicada radica en derivar las fórmulas de mp y Np. Aquí calculamos las fórmulas para mp y Np para la función e para $12 \le p \le 36$. De hecho, la representación gráfica de la función de h(m) respalda nuestros resultados para todos los valores de p. La derivación de mp y Np para las otras funciones es muy similar.

(1) Para la función e A, tenemos $g(u) = u^2$, k = 2, K = 1, y

$$10\,h(m) = m + \frac{p \ln}{(m+1)\ln 2} - \frac{1}{(m+1)\text{en } 2}\,\text{En}(m+1)!. \tag{25}$$

Usando la fórmula de aproximación de Sterling, obtenemos lo siguiente:

$$\ln(2\pi) + \ln(m+1) + (m+1)\ln(m+1) - m - 1 + O\left(\frac{1}{m+1}\right).$$

Ahora, implementándolo en (25), obtenemos lo siguiente:

$$h(m) \approx m + (m \quad \frac{p \ln 10}{+ 1)\ln 2} - \frac{\ln(2\pi)}{2(m+1)\ln 2} - \frac{\ln(m+1) + \quad 1}{2(m+ 1)\ln 2 \ln 2.}$$

Dejar

$$q(x) = x + (x \quad \frac{p \ln 10}{+ 1)\ln 2} - \frac{\ln(2\pi)}{2(x+1)\ln 2} - \frac{\ln(x+1) + \ln 2.\ 2(x+}{\text{Tomando su}\quad 1)\ln 2}$$

derivada con respecto a x, se obtiene

$$q'(x) = 1 - + (x+1)\,2 \ln 2 \quad \frac{p \ln 10}{2 \ln 2} \quad \frac{\ln(2\pi) + \quad}{+ 1)\,2 \ln 2} \quad \frac{2(x+1)}{\ln(x+1)\,2(x} - \frac{1}{2(x+1)\,2 \text{ en } 2.}$$

Para determinar los valores extremos, lo ponemos en 0 y multiplicamos ambos lados por $(m+1)^2$:

$$(x+1)^2 - \frac{1}{\text{en } 2} \quad \ln(x+1)\,(x + 1) + 2 \ln 2 \quad - \frac{2p \ln 10 - \ln(2\pi) + 1\,2 \ln 2}{} = 0.$$

Usando álgebra simple, obtenemos lo siguiente:

$$(x+1)^2 > \frac{(m+1) - \ln(m+1) + (m+1) - \ln(2\pi)\ 2p \ln 10 + 1 + 2 \ln 2\ 2 \ln 2}{} > \frac{2p \text{ en } 10 + 1}{2 \text{ en } 2}.$$

Usando p = 12, obtenemos un límite inferior para x:

$$x > \sqrt{\frac{2p \text{ en } 10 + 1}{2}} - 1 = 6{,}37.\ 2 \text{ en}$$

Por lo tanto, tenemos un límite inferior para m: $m \ge 6$. También podemos resolver la ecuación cuadrática para obtener

$$\text{metro} + 1 = \frac{1}{+ \text{en } 2} \quad \sqrt{\frac{1}{4\ln 2\ 2} + \frac{8p \ln 10 - 4 \ln(2\pi) + 4\ 8 \ln 2}{}}.$$

Ahora, usando el valor más pequeño de m obtenido y p = 36, podemos obtener nuestro primer límite superior:

$$\text{metro} < -1 + \frac{1}{+ \ln 2} \quad \sqrt{\frac{1}{4\ln 2\ 2} - \frac{\text{en } 7}{2 \text{ en } 2} + \frac{276 \ln 10 - 4 \ln(2\pi) + 4\ 8 \ln 2}{}} \approx 11.31.$$

Machine Translated by Google

Podemos usar nuevamente los valores de myp = 12 para actualizar el límite inferior de m de la siguiente manera:

$$\text{metro} > -1 + \frac{1}{\text{en } 2} + \sqrt{\frac{1}{4\ln 2\ 2} + -\frac{\text{en } 12}{2 \text{ en } 2} + \frac{96 \ln 10 - 4 \ln(2\pi) + 4}{8 \text{ en } 2}} \approx 6{,}61.$$

Ahora, nuestro nuevo límite inferior para m es 7. Con este pequeño rango de valores para m, al inspeccionar la fórmula para los valores de m, podemos ver que el valor de m es muy cercano y $\frac{\text{p en } 10}{\text{en } 2}$, podemos afirmar con seguridad que m= $\sqrt{\frac{9\text{p en } 10}{8 \text{ en } 2}}$. Nuestros siguientes soportes de cálculo numérico esta afirmación por completo.

Caso (a): p = 12. Usando la fórmula obtenida por la fórmula cuadrática y la superior límite m = 11, obtenemos un nuevo límite inferior: m ≥ 6,63. Utilizando la fórmula obtenida por el fórmula cuadrática y cota inferior m = 7, también podemos obtener la cota superior: m ≤ 6,63. Usando nuestro reclamo $\sqrt{\frac{9\text{p en } 10}{8 \text{ en } 2}}$, obtenemos m = 6,70. Esto coincide claramente con la parte inferior y superior. encuadernado bastante bien.

Caso (b): p = 18. Usando la fórmula obtenida por la fórmula cuadrática y la superior límite m = 11, obtenemos un nuevo límite inferior: m ≥ 8,05. Utilizando la fórmula obtenida por el fórmula cuadrática y límite inferior m = 7, también podemos obtener el límite superior: m ≤ 8,07. Usando nuestro reclamo $\sqrt{\frac{9\text{p en } 10}{8 \text{ en } 2}}$, obtenemos m = 8,20. Estos valores coinciden muy bien para tomar el entero más cercano.

Caso (c): p = 24. Usando la fórmula obtenida por la fórmula cuadrática y la superior límite m = 11, obtenemos un nuevo límite inferior: m ≥ 9,27. Utilizando la fórmula obtenida por el fórmula cuadrática y cota inferior m = 7, también podemos obtener la cota superior: m ≤ 9,28. Usando nuestro reclamo $\sqrt{\frac{9\text{p en } 10}{8 \text{ en } 2}}$, obtenemos m = 9,47.

Caso (d): p = 36. Usando la fórmula obtenida por la fórmula cuadrática y la superior límite m = 11, obtenemos un nuevo límite inferior: m ≥ 11,29. Usando la fórmula obtenida mediante la fórmula cuadrática y el límite inferior m = 7, también podemos obtener el límite superior: metro ≤ 11,37. Usando nuestro reclamo $\sqrt{\frac{9\text{p en } 10}{8 \text{ en } 2}}$, obtenemos m = 11,60.

Usando mp = $\sqrt{\frac{9\text{p en } 10}{8 \text{ en } 2}}$, podemos obtener

$$\text{norte} = \frac{\text{p en } 10}{(\text{mp}+1)\ln 2} - \frac{\text{En}(\text{mp}+1)!}{(\text{mp}+1)\ln 2 \ln 2} + \frac{\text{en } \lambda}{} + 1. \tag{26}$$

(2) Para $\frac{\text{pecado A}}{\sinh A}$, tenemos $g(u) = 3u \quad 4u^{3}$, k = 3, K = 2 y

$$h(\text{metro}) = \text{metro} + \frac{2\text{p en } 10}{(2\text{m}+2)\text{en } 3} - \frac{2 \text{ en}(2\text{m}+2)!}{(2\text{m}+2)\ln k\ .} \tag{27}$$

En este caso, para p = 12, 16, 24 y 32, los puntos enteros correspondientes para h(m) para lograr su valor mínimo es 3, 3, 5 y 6. Nuevamente, se puede estimar que

$$\text{p.f.} = \sqrt{\frac{3\text{p en } 10}{5 \text{ en } 3}}, \text{ norte} = \frac{\text{p en } 10}{(2\text{mp}+2)\text{en } 3} - \frac{\text{En}(2\text{mp}+2)!}{(2\text{mp}+2)\ln 3 \ln 3} + \frac{\text{en } \lambda}{} + 1. \tag{28}$$

(3) Para $\frac{\text{porque A}}{\cosh A}$, hay dos casos:

Caso (I): $g(u) = 4u^{3} - 3u$, k = 3, K = 2 y (29) se cumple. En comparación con I, los valores en la segunda parte de $C = I + \sum_{j=1}^{\text{diputado}} \frac{(-B)\ j}{(2j)!}$ son muy pequeños, por lo que podemos reemplazar N con N − 1, y p.m. con p.m. + 1.

Caso (II): $g(u) = u^{2} \quad 1$, k = 2, K = 1 y

$$h(\text{metro}) = \text{metro} + \frac{\text{p en } 10}{(2\text{m}+1)\text{en } 2} - \frac{1}{(2\text{m}+1)\text{en } 2}\text{En}(2\text{m} + 1)!. \tag{29}$$

Machine Translated by Google

De manera similar, para p = 12, 16, 24 y 32, los puntos enteros correspondientes para que h(m) alcance su Los valores mínimos son 4, 5, 6 y 7. En consecuencia, podemos determinar lo siguiente:

$$\text{p.f.} = \frac{3p \text{ en } 10}{5 \text{ en } 2}, \text{ norte} = \frac{p \text{ en } 10}{(2mp+1)\text{en } 2} - \frac{\text{En}(2mp+1)!}{(2mp+1)\ln 2 \ln 2} + \frac{\text{en } \lambda}{} + 1. \tag{30}$$

La efectividad de los algoritmos se muestra en la Tabla 1 mediante la implementación de Cálculos en Mathematica. Np y mp se reemplazan con Np,μ = μNp, mp,μ en la Tabla 1, definido correspondientemente por

$$\begin{aligned} \text{pf, } \mu &= \min_{B\ j/j!\ <10-p} \text{ j, B} = \frac{A}{2\ \text{Np, } \mu} \text{ para eA,} \\ \text{mp,}\mu &= \min_{B\ j+1/2/(2j+1)!)\ <10-p} \text{ j, B} = \frac{A^2}{3^{2\text{Np, } \mu}} \text{ para el pecado A,} \\ \text{(I) pf, } \mu &= \min_{B\ j\ /(2j)!\ <10-p} \text{ j, B} = \frac{A^2}{3^{2\text{Np, } \mu}} \text{ para cos A,} \\ \text{(II) pf, } \mu &= \min_{B\ j\ /(2j)!\ <10-p} \text{ j, B} = \frac{A^2}{2^{2\text{Np, } \mu}} \text{ para cos A,} \end{aligned} \tag{31}$$

donde se tiene Np,1 = Np y mp = mp,1.

A modo de comparación, tome A la siguiente forma:

$$A = UDU\text{-}1 \quad , \tag{32}$$

donde U es una matriz n por n generada aleatoriamente, Da = Dial(d1, d2, · · · , dn) es una matriz aleatoria matriz diagonal n por n generada con di (a, b), i = 1, 2, · · · f(D) = , n, entonces f(A) = U f(D)U−1 , Dial(f(d1), f(d2), · · · , f(dn)) . En todos los ejemplos de este artículo coincidimos en error relativo de la siguiente manera:

$$\text{err(g)(A)} = \frac{\| f(A)-U\ f(D)U-1 \|_\infty}{\| f(D)U-1 \|_\infty}. \tag{33}$$

Sean Ai (i = 1, 2, · · · , M) M matrices producidas por (32). De la Figura 1, es evidente que la precisión del cálculo es esencialmente la misma para μ = 1, pero μ = $\frac{2}{3}, \frac{4}{3}$, proporciona la máxima eficiencia. La fórmula del doble ángulo es notablemente más eficiente y precisa que la fórmula del triple ángulo para cos A. De ahora en adelante, excluiremos el caso que utiliza la Fórmula del triple ángulo para cos A.

Ahora yuxtapongamos los algoritmos presentados en la Tabla 1 con las funciones internas y algoritmos comúnmente utilizados que se encuentran en el software matemático. Para facilitar la referencia, dejemos err(g) denota el error de cálculo de la función matricial f(Ai)(i = 1, 2, · · · , M), donde un espacio en blanco g implica la utilización de los algoritmos de las Tablas 1 a 3 de este artículo. Si g = Sym, Pade, · · · significa , cálculo según la función interna, aproximación de Padé, · · · tivamente. Cuando p = 16, mp = 7, 4, , respe- 3 para e A, cos A(cosh A), sen A(senh A), respectivamente. Nosotros denotamos la aproximación de Padé de sen x como r66(x) y la aproximación de Padé de e x y cos x como r88(x). En la Tabla 1, la aproximación de Padé de e A, cos A(cosh A), sen A(sinh A) es como sigue:

$$\begin{aligned} \text{r88(x)} &= \frac{1+\frac{1}{2}+\frac{2}{60}+\frac{3}{60}+\frac{4}{+624\ 9360}+\frac{5}{205920}+\frac{6}{7207200}+\frac{7}{518918400}+\frac{8}{}}{1-\frac{1}{2}+\frac{2}{60}-\frac{3}{60}+\frac{4}{+624\ 9360}-\frac{5}{205920}+\frac{6}{7207200}-\frac{7}{518918400}+\frac{8}{}} \text{ para mi}, \\ 1 \quad \text{r88 (x)} &= \frac{1\pm\frac{260735x^2}{545628}+\frac{4375409x^4}{+1410863280}+\frac{7696415x^6}{1310867072}+\frac{80737373x^8}{2359470072960}}{1\pm\frac{12079x2}{545628}+\frac{34709x4}{1410863280}\pm\frac{109247x6}{65540835360}+\frac{11321x8}{1814976979200}} \text{ para } \frac{\text{porque x}}{\text{cosh x}}, \\ \text{r66(x)} &= \frac{x-\frac{7x}{60\ 2}+\frac{5}{600}}{1\pm\frac{4}{20\ 600}\pm\frac{6}{14400}} \text{ para } \frac{\text{pecado x}}{\text{sinh x}}. \end{aligned} \tag{34}$$

Para los algoritmos de la Tabla 1, el err(g)(A) de una matriz aleatoria de orden 1 a 800 es como sigue.

Machine Translated by Google

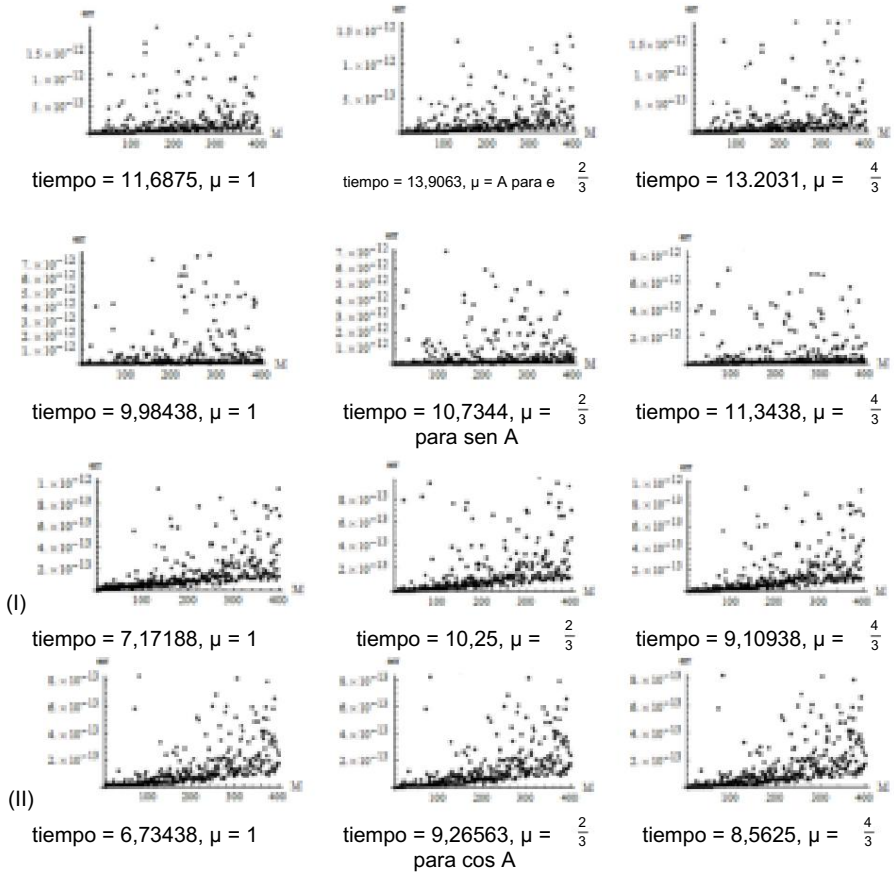


Figura 1. Comparación de errores de los algoritmos de la Tabla 1 para μ = 1, $\frac{2}{3}, \frac{4}{3}$.

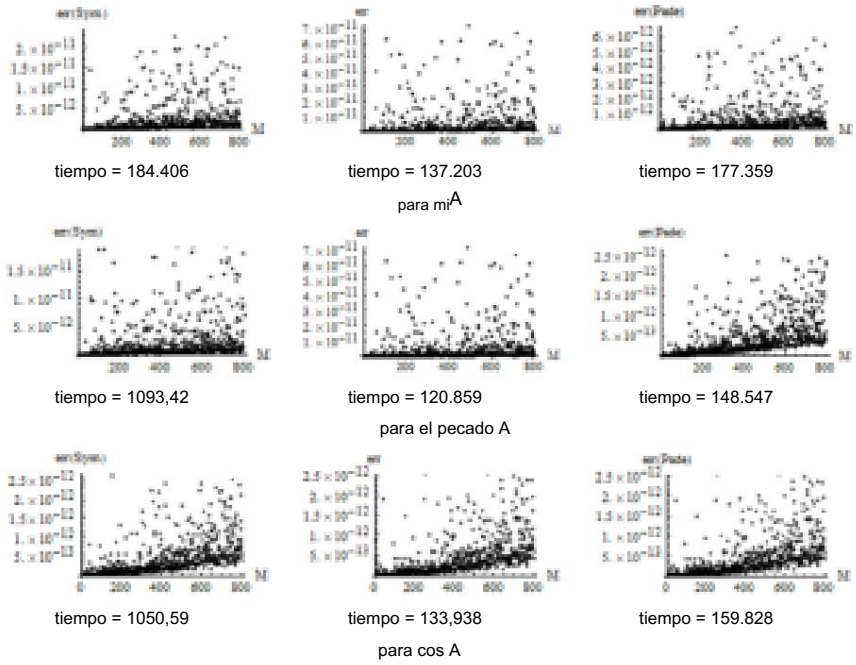
Tabla 2. Los algoritmos rápidos para tan A, sec A, tanh A y sechA.

	diputado, norte	Algoritmo
tan A tanh A	p.f. = $\frac{\overline{\quad}}{\frac{2p \text{ en } 10}{\text{en } 3}}$, norte = norte1 + 1	B = $\frac{A^2}{3\,2N}$, metro c = $\frac{2A\Sigma}{3}$ (±1) jσ(j + 1) 4 j+1 − 1 B j N j=0 f o(i = 0, i < N, i + +) C = C(3I C tan ^{2})(Yo 3C^2)$^{-1}$ A tanh A ≈C
seg A seg hA	p.f. = $\frac{\overline{\quad}}{\frac{2p \text{ en } 10}{\text{en } 3}}$, norte = norte2 + 1	B = $\frac{A^2}{3}$ yo + $\sum$ metro (1) jE jB^j 2N , j=1 f o(i = 0, i < N, i + +) C = C ^{3}(4I - 3C ^{2})$^{-1}$ segundo A segundos ha ≈C
N1 = $\frac{p \ln 10+\ln(2\sigma(mp+2))}{(2mp+3)\ln 3}$	λ + $\frac{En}{\text{en } 3}$, N2 = $\frac{p \ln 10+\ln Emp+1 \ln \lambda +}{(2mp+3)\ln 3 \ln 3}$. ___	

De la Figura 2 se desprende claramente que, para e A, cos A y sen A, la precisión obtenida mediante los tres métodos es esencialmente la misma. Sin embargo, cuando el orden de la matriz es relativamente alto, el método de expansión de Taylor presentado en la Tabla 1 supera a los demás. En particular, para cos y sen A, el tiempo de cálculo de las funciones internas es aproximadamente ocho veces mayor que el de los algoritmos de la Tabla 1. Los cálculos numéricos demuestran que, si bien la velocidad de cálculo de cos A y sen A utilizando las funciones internas Re(e iA) = cos A

Machine Translated by Google

y Im(e iA) = sen A es más rápida que la de las funciones internas cos A y sen A, su precisión de cálculo es relativamente pobre.



tiempo = 184.406 tiempo = 137.203 tiempo = 177.359

para mi A

tiempo = 1093,42 tiempo = 120.859 tiempo = 148.547

para el pecado A

tiempo = 1050,59 tiempo = 133,938 tiempo = 159.828

para cos A

Figura 2. Error de e A, cos A y sen A para tres algoritmos.

Tabla 3. Los algoritmos rápidos para cot A, coth A, csc A y cschA.

	diputado, norte	Algoritmo
cuna A cuna A	p.f. = $\overline{\frac{7p \text{ en } 10}{5 \text{ en } 3}}$, norte = norte1 + 1	B = $\frac{A^2}{3}$ $\frac{C}{2N}$, yo − 2 $\sum^{\text{metro}}$ (±1) jσ(j)B) f o(i = 0, i < N, i + +) (C = 3(C(3B C ^{2})(B 3C^2) $^{-1}$; B = 9B) cuna A cuna A ≈C
csc A cschA	p.f. = $\overline{\frac{5p \text{ en } 10}{2 \text{ en } 3}}$, norte = norte2 + 1	B = $\frac{A^2}{3}$ $\frac{C}{2N}$, yo + 2 $\sum_{j=1}^{\text{metro+1}}$ (±1) jσ(j) 1 − $\frac{}{2}$ $\frac{}{4j}$ Bj f o(i = 0, i < N, i + +) (C = 3C ^{3} (3C ^{2} 4B) $^{-1}$; B = 9B) csc A cschA ≈C
	N1 = $\frac{p \ln 10+\ln(2\sigma(mp+2)) \ln \lambda + (2mp+3)\ln 3}{\text{en } 3}$, N2 = $\frac{p \ln 10+\ln \text{Emp+1} \ln \lambda +}{(2mp+3)\ln 3 \ln 3}$ ___	

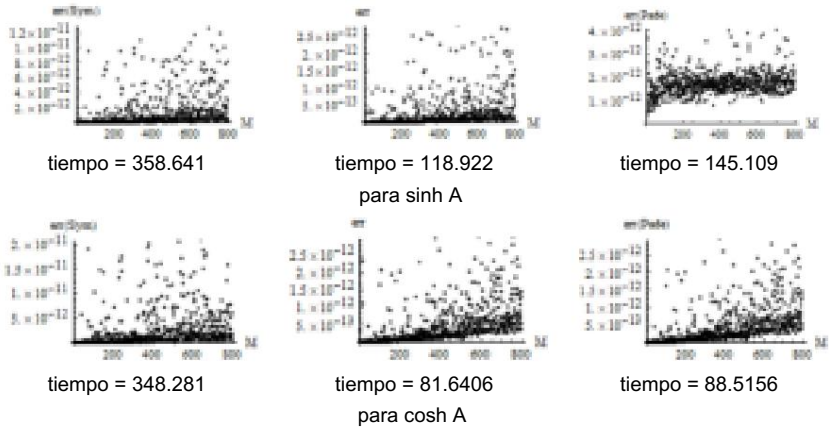
Para cosh A y sinh A, las funciones internas se pueden calcular utilizando la siguiente ecuación:

$$1 \cosh A = 2 \; -(mi^{A} + mi -A), \text{ senh } A = \frac{}{1\,2}(mi^{A} - e -A). \quad (35)$$

Nuestro experimento proporciona los siguientes resultados para err(g)(A) de matrices aleatorias con órdenes de 1 a 800.

De la Figura 3 se desprende claramente que el uso de (35) no sólo duplica el tiempo de cálculo sino que también hace que el error de cálculo sea sustancial. Además, al observar las Figuras 2 y 3, se hace evidente que los algoritmos de la Tabla 1 superan al algoritmo de aproximación de Padé.

Machine Translated by Google



tiempo = 358.641 tiempo = 118.922 tiempo = 145.109

para sinh A

tiempo = 348.281 tiempo = 81.6406 tiempo = 88.5156

para cosh A

Figura 3. Error de sinh A y cosh A para tres algoritmos.

4. Algoritmo rápido para tan A, tanh A, sec A y sechA Ahora contemplemos el cálculo rápido de tan A, tanh A, sec A y sechA.

Aunque estas funciones matriciales se pueden calcular utilizando el siguiente enfoque, nuestro objetivo es calcularlas directamente, como se demostró en la sección anterior.

$$\begin{aligned} &\tan A = \text{sen } A(\cos A)^{-1}, \text{ seg } A = (\cos A)^{-1}, \\ &A = \text{senh } A(\cosh A)^{-1} = mi - mi^{-A}{}_{A} + ee^{-Un-1}, \\ &\text{sechA} = (\cosh A)^{-1} A = 2\,mi + mi^{-Un-1}. \end{aligned} \tag{36}$$

Si bien la aproximación de Padé contribuye significativamente al cálculo de la función matricial, los resultados numéricos de la sección anterior indican que, según el algoritmo presentado en este artículo, ni la precisión ni la velocidad del cálculo son predominantes. Por lo tanto, esta sección no lo adopta.

Teorema 3. Para tan A, sec A, tanh A y sechA, tenemos el siguiente algoritmo rápido basado en el teorema 1.

En la Tabla 2, para σ(j) y E tenemos la siguiente fórmula recursiva [1]: j ,

$$\begin{aligned} s(2n) &= \frac{}{4\,4n+1}\sum_{=1}^{norte-1} \sigma(l)\sigma(2n-l) + \frac{2\,2\,\sigma}{4n+1}(n),\ n = 1, 2, \cdots, \\ \sigma(2norte-1) &= \frac{}{4\,4n-1}\sum_{=1}^{norte-1} \sigma(l)\sigma(2n-1-l),\ n = 2, 3, \cdots \\ En &= -\sum_{=0}^{norte-1} \frac{mi\ k}{(2n-2k)!},\ E0 = 1,\ \sigma(1) = \frac{}{6}. \end{aligned} \tag{37}$$

Utilizando los algoritmos de la Tabla 2, nuestros experimentos generaron los siguientes datos para el error (g) (A) de matrices aleatorias de orden 1 a 200.

Aquí, Sym indica que las funciones matriciales tan A, tanh A, sec A y sechA, expresadas $-A^{-1}\,-A^{-1}\,AA$ por $\text{sin}\,A(\cos A)^{-1}$, $(\cos A)^{-1}$ y $2\,e^{-mi^{-A}}$, respectivamente, son en la Figura 4, la velocidad de cálculo de las funciones internas. Como se observa Tabla 2 supera la de la función interna, particularmente para tan A y sec A, donde la velocidad aumenta aproximadamente 3 veces en comparación con la función interna, consistente. Cuando el orden n de la matriz es relativamente grande, para tanh A y sec hA, la velocidad de cálculo es el doble que la de la función interna, pero la precisión sigue siendo la misma que la de la función interna.

Machine Translated by Google

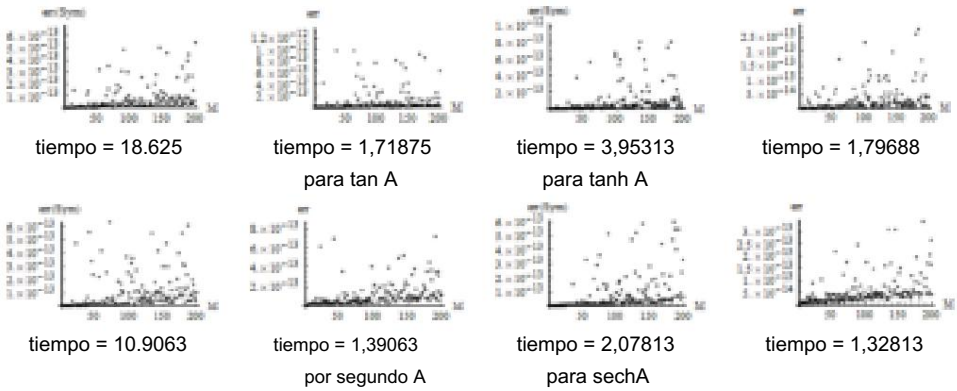


Figura 4. Error de tan A, sec A, tanh A y sechA para dos algoritmos.

Consideremos implementar nuestros algoritmos para calcular las funciones cot A, cot A, csc A y cschA. Por

$$\begin{aligned}
&\text{cuna}(At) \approx A^{-1}\,\sigma(j+1)A^{2j+1};\ (o<|A|<\pi,\ z=0),-2\sum_{j=0}^{metro}\\
&\coth(En) \approx A^{-1}+2\sum_{j=0}^{metro}(-1)^{j}\sigma(j+1)A^{2j+1}\ (0<|A|<\pi),\\
&\operatorname{csc}(En) \approx A^{-1}+2\sum_{j=0}^{metro}\sigma(j+1)\,1-\frac{}{2^{4j+1}}\,A^{2j+1}\ (0<|A|<\pi),\\
&\operatorname{csch}(En) \approx A^{-1}-2\sum_{j=0}^{metro}(-1)^{j}\sigma(j+1)\,1-\frac{}{2^{4j+1}}\,A^{2j+1}\ (0<|A|<\pi),
\end{aligned} \tag{38}$$

vemos que cuando A es pequeño, la norma de la inversa de A será grande, por lo que el cálculo utilizando (38) producirá grandes errores. Por lo tanto, debemos utilizar el siguiente método:

$$\begin{aligned}
&\text{cuna } A = A^{-1}\ \text{cuna1 } A,\ \text{cuna} = A^{-1}\ \coth 1\ un,\\
&\operatorname{csc} A = A^{-1}\ \operatorname{csc}1\ A,\ \operatorname{cschA} = A^{-1}\ \operatorname{csch}1A,
\end{aligned} \tag{39}$$

dónde

$$\begin{aligned}
&\cot 1\ A \approx I-2\sum_{j=1}^{m+1}\sigma(j)\,A^{2j};\ (|z|<\pi,\ z=0),\\
&\coth 1A \approx yo-2\sum_{j=1}^{m+1}(-1)^{j}\sigma(j)A^{2j}\ (|z|<\pi,\ z=0),\\
&\operatorname{csc}1\ A \approx I+2\sum_{j=1}^{m+1}s(j)\,1-\frac{}{2^{4j}}\,A^{2j}\ (|z|<\pi,\ z=0),\\
&\operatorname{csch}1A \approx I+2\sum_{j=1}^{m+1}(-1)^{j}\sigma(j)\,1-\frac{}{2^{4j}}\,A^{2j}\ (|z|<\pi,\ z=0).
\end{aligned} \tag{40}$$

Teorema 4. Usando el Teorema 1, tenemos el siguiente algoritmo para cot A, coth A, csc A y sec hA.

La implementación de los algoritmos de la Tabla 3 proporciona los siguientes resultados para las funciones matriciales en (39).

Aquí, Sym significa que las funciones matriciales cot A, coth A, csc A y sec hA se calculan usando- $-A$ ing las funciones internas con las fórmulas regulares cos A(sin A) (sin A) con $^{-1}$, $^{A}+ee^{-A}$ $^{A}mi-mi^{-1}$, los $^{-1}$, e^{2}, respectivamente. Los resultados de la Figura 5 son esencialmente consistentes : e^{-A-1A} y resultados correspondientes en la Figura 2.

Machine Translated by Google

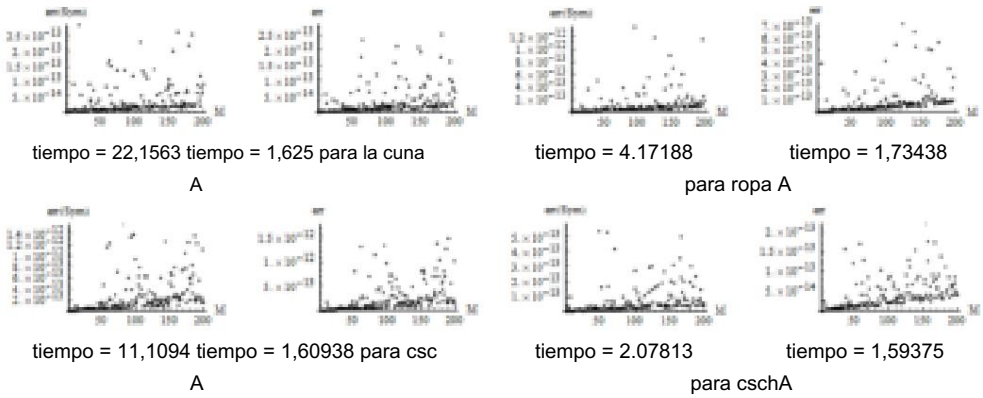


Figura 5. Error de cot A, cot A, csc A y cschA para dos algoritmos.

5. Algoritmo rápido para la función de matriz logarítmica ln A y algunas funciones relacionadas

La gente considera varios algoritmos [17-22] para la función matricial logarítmica ln A. Uno de los métodos más efectivos que notamos fue el siguiente algoritmo. La forma estándar de abordar este problema es utilizar el operador de raíz cuadrada repetidamente para acercar A a la identidad:

$$\text{En A} = 2^{k} \ln Ak \text{ , Ak} = A^{1/2k} \quad . \tag{41}$$

A medida que k aumenta, $A^{1/2k} \to I$, por lo que para k suficientemente grande , podemos aplicar un método directo a $A^{1/2k}$. Este procedimiento para el logaritmo fue introducido por Kenney y Laub [19]. El método específico es elegir k tal que

$$\text{Un}^{1/2k} - \text{Yo} < \alpha < 1, \tag{42}$$

y normalmente, k se selecciona entre 8. Optimizamos ligeramente el método de la siguiente manera:

$$\text{A ln A} = \ln \lambda \cdot I + 2^{k} \ln Ak \text{ , Ak} = \frac{A^{1/2k}}{A} A, \quad A0 = A, \tag{43}$$

y

$$A^{1/2k} - \text{Yo} < \alpha < 1, 0 \tag{44}$$

donde k generalmente se selecciona entre 4.

Para ln x, el algoritmo que converge más rápido es

$$\ln x = \ln 1-y \text{ y} \approx -2 \sum_{j=1}^{\text{metro}} \frac{2j+1}{1+y^{2j+1}}, \tag{45}$$

donde $y = \frac{1-x}{1+x}$. Para Ak , tenemos

$$2j \ln Ak = \ln(I - B) - \ln(I + B) \approx -2 \sum_{j=1}^{\text{metro}} \frac{B^{2j+1}}{2j+1}, \tag{46}$$

donde $B = (I - Ak)(I + Ak)^{-1}$. Aplicando la transformación de Euler-Abel [21], obtenemos la siguiente serie acelerada:

$$1 \ln Ak \approx 3^{-} BF2 \left(yo - 12 \sum_{j=1}^{\text{metro}} \frac{B^{2j}}{(2j+3)(4j^{2}-1)}\right) - F, \sum \tag{47}$$

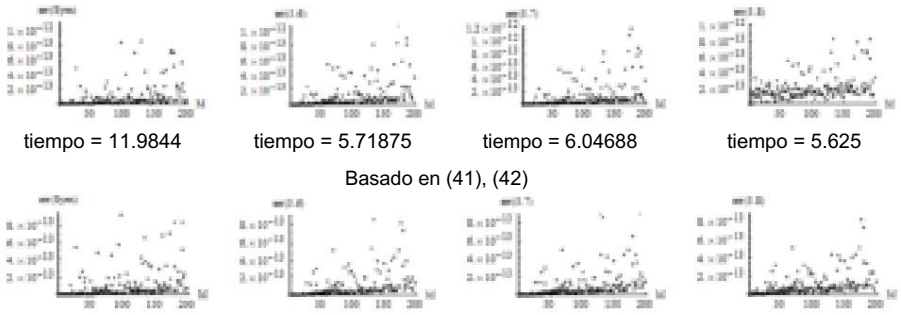
donde $F = 2B(I - B^{2})^{-2}$. Podemos utilizar la siguiente aproximación de Padé de ln Ak [20].

$$\begin{aligned} \text{En Ak} &\approx (B + 2B^{2} + \tfrac{47}{36B}^{3} + \tfrac{11}{36B}^{4} + \tfrac{137}{7560B}^{5})F, \\ F &= (yo + \tfrac{5}{2}B + \tfrac{20}{2}B + 9\tfrac{5}{2}B + 6\tfrac{5}{42B}^{4} + \tfrac{1}{252B}^{5})^{-1}, \end{aligned} \tag{48}$$

Machine Translated by Google

donde B = Ak − I. También proporcionamos algunos resultados numéricos para la función matricial ln A.

Con base en (41)–(44), también proporcionamos algunos resultados numéricos de ln A usando (46)–(48). Se puede ver en la Figura 6 que a través de la optimización (43), la precisión del cálculo de (46) – (48) mejora significativamente y la velocidad de cálculo de (46) – (48) es más del doble que la velocidad interna. función. La ventaja de la velocidad de cálculo de la aproximación de Padé es obvia.



tiempo = 11.9844 tiempo = 5.71875 tiempo = 6.04688 tiempo = 5.625

Basado en (41), (42)

tiempo = 12.1875 tiempo = 5,64063 tiempo = 6,01563 Basado en (43), (44) tiempo = 5.3125

Figura 6. Error de ln A para los distintos algoritmos.

Usamos las siguientes identidades para implementar nuestro algoritmo general para los cálculos de las funciones trigonométricas e hiperbólicas inversas. Para controlar la longitud del artículo, no mostraremos los resultados aquí.

$$\begin{aligned} &\text{arcosen } x = 2\,\text{arcosen}\ \frac{\sqrt{x}}{\sqrt{2(1+\sqrt{1-x^2})}},\\ &\text{arctán } x = 2\,\text{arctán}\ \frac{x}{1+\sqrt{1+x^2}},\\ &\sum_{l=1}^{\infty}\frac{(2l-1)!}{(2l+1)\,l!\,l!}x^{2l+1} = \text{arcsen } x = x + \sum_{l=0}^{\infty}\frac{(2l)!}{(2l+1)\,l!\,l!}\frac{x^{2l+1}}{2^{2l+1}} = \text{arctan } x \end{aligned} \tag{49}$$

$$\begin{aligned} &\text{arcosenhx} = 2\text{arcsenh}\ \sqrt{\frac{x}{2(1+\sqrt{1+x^2})}}, \text{arctanhx} = 2\text{arctanh}\ \frac{x}{1+\sqrt{1-x^2}},\\ &\sum_{l=1}^{\infty}\frac{(-1)^{l-1}(2l-1)!}{(2l+1)\,l!\,l!}x^{2l+1} = \text{arcsinhx} = x + \sum_{l=0}^{\infty}\frac{(-1)^{l}(2l)!}{(2l+1)\,l!\,l!}\frac{x^{2l+1}}{2^{2l+1}} = \text{arctanhx} \end{aligned} \tag{50}$$

Para arcsin A, arctan A, arcsinhA y arctanhA, bajo ciertas condiciones, podemos usar la siguiente identidad.

$$\begin{aligned} &\text{arcosen } A = -i\ln\left(\sqrt{I - A^2} + iA\right), \text{arcsenhA} = \ln\left(\sqrt{I + A^2} + A\right),\\ &\text{arctán } A = \frac{i}{2}\ln(iI + A)(iI - A)^{-1},\\ &\text{arctanhA} = \frac{1}{2}\ln(I + A)(I - A)^{-1}. \end{aligned} \tag{51}$$

6. Fórmula de cálculo mejorada para los problemas de valor inicial (3) y (8)

Ahora, consideremos el cálculo de la integral en (11), es decir,

$$f_k = \int_0^{\tau} e^{H(k\tau - s)} f(s)\,ds. \tag{52}$$

Si se utiliza la integración numérica para calcular esta integral, lleva bastante tiempo y no se puede garantizar la precisión del cálculo. Adoptamos el siguiente método aquí. Porque existe la siguiente relación: nn = 1, 2, · · · ,

$$\begin{aligned} f_k &= \sum_{j=0}^{nn-1}\int_{\frac{j\tau}{nn}}^{\frac{(j+1)\tau}{nn}} e^{H(k\tau - s)} f(s)\,ds = \sum_{j=0}^{nn-1} T_{nn,j} \int_0^{\frac{\tau}{nn}} e^{H(k\tau - \frac{j\tau}{nn} - s)} f(s)\,ds\\ &\approx \sum_{j=0}^{nn-1} T_{nn,j} \sum_{l=0}^{metro} \frac{}{\text{¡Hola !}} \int_0^{\frac{\tau}{nn}} s^{l} F^{H(k\tau - \frac{j\tau}{nn} - s)}\,ds, \end{aligned} \tag{53}$$

Machine Translated by Google

dónde

$$Tnn,0 = I,\ Tnn = e^{h\frac{\tau}{nn}},\ Tnn,j = e^{h\frac{j\tau}{nn}} = TnnTnn,j-1,\ j = 1, 2, \cdots, nn-1. \tag{54}$$

Tnn se puede calcular mediante el siguiente método:

$$\begin{aligned} &Tnn = Yo,\ zz = Yo;\\ &para(i = 1,\ i \le m,\ i++)\ (zz = \frac{\overline{\tau nn}}{I} H \times zz;\ Tnn = Tnn + zz). \end{aligned} \tag{55}$$

Entonces, (53) se puede escribir de la siguiente manera:

$$F\,k \approx \sum_{=0}^{nn-1} j\ Tnn,jHj,\ Hj = \sum_{=0}^{metro} \frac{I}{¡Hola!}\ 0^{\frac{\tau}{nn}}\ s\ IFH(k\tau - \frac{j\tau}{nn} - s)ds, \tag{56}$$

entonces podemos convertir la integral del producto de la función exponencial matricial y la función vectorial en la integral de la función vectorial pura.

La fórmula de cálculo mejorada para (3) es la siguiente:

$$\begin{aligned} &(55), (56),\\ &Reino\ Unido = TUk-1 + \sum_{=0}^{nn-1} j\ Tnn,jHj,\ k = 1, 2, \cdots, \end{aligned} \tag{57}$$

donde $T = e$ (55) y $^{H\tau}$ se calcula mediante el algoritmo de la Tabla 1, y Tnn,j y Hj se calculan mediante (56).

Para

$$0^{\frac{\tau}{nn}}\ s\ IFH(k\tau - \frac{j\tau}{nn} - s)ds, \tag{58}$$

Se pueden utilizar los dos algoritmos siguientes.

(1) Método de la función integrable Si $F(t) = (f1(t), f2(t), \cdots, fv(t), \cdots, fn(t))T$, fv(t) es la suma algebraica de funciones de los siguientes tipos:

$$t^{\alpha},\ I - \omega t\ te,\ I\ t\ pecado\ \omega t,\ t^{"}\ porque\ \omega t, \tag{59}$$

entonces (58) puede expresarse en forma analítica, por lo que Hj también puede expresarse en forma analítica. De esta forma se puede completar el cálculo de (57).

(2) Fórmula integral de tres puntos de Gauss

Tenga en cuenta que, con una precisión de cálculo de 16 bits,

$$\frac{\tau}{nn} < \frac{}{1\ 2\ n16\lambda H} \approx \frac{0.0376434}{|\lambda H|} < 0{,}0376434, \tag{60}$$

y la precisión algebraica de la fórmula integral de tres puntos de Gauss es 5. Usando la fórmula integral de tres puntos de Gauss,

$$\overline{1}\ 2\ \overline{\tau nn}^{\,I+1}\ 3\ \sum i=1\ con(\ 2^{\frac{1+bi}{}}\ IFH(k\tau - \frac{(j+1/2)\tau}{nn}) - \frac{bi\tau}{2nn}), \tag{61}$$

$b2 = 0, b3 = 5$, precisión. Entonces, $\frac{5}{9}w2 = \frac{8}{9}$ (57) puede reemplazarse por el siguiente algoritmo. $\frac{\sqrt{15}}{\sqrt{15}}$ donde $w1 = w3 = 9,\ b1 = -\frac{5}{}$, para aproximar (58) con alta

Algoritmo 1. Algoritmo de solución numérica del problema de valor inicial (3): para nn = 1, 2, · · ·

$$Reino\ Unido = TUk-1 + \frac{}{\overline{\tau 2nn}} \sum_{=0}^{nn-1} j\ Tnn,j \sum^{3} H\ \frac{(1+b\,i)\tau}{2nn}\ FH\ k\tau - \frac{(2i+1+bi)\tau}{2nn}\ i=1\ wi\ e\ \frac{}{2nn},\ k = 1, 2, \cdots, \tag{62}$$

dónde

$$H\ \frac{(1+b\,i)\tau}{2nn} = mi\ \frac{wi\tau}{2nn}\ 1\ yo + \sum^{metro} I!\ I=1\ \frac{(1+bi)\tau}{2nn}\ ^{yo}\ hola\ . \tag{63}$$

Machine Translated by Google

En particular, si dejamos

$$n = \begin{matrix} [2\ np\lambda H\tau] + 1 & 2np\lambda H\tau > 1 \\ 1 & 2np\lambda H\tau \le 1 \end{matrix}, \tag{64}$$

entonces Np = 0, m = mp = $\frac{\overline{9p\ en\ 10}}{8\ en\ 2}$,

$$Wi,H \approx \frac{wi\tau}{2nn} \sum y_{o+i!}^{diputado} {}_{l=1} - \frac{(1+bi)\tau}{2nn}{}^{\sim} hola\ . \tag{65}$$

En este momento, la precisión del cálculo teórico de (62) es p.

Observación 3. Tenga en cuenta que Wi,H(i = 1, 2, 3) en (63) y (65) solo necesita calcularse una vez y es independiente de F(t), lo cual no sólo es eficiente sino también fácil. para programar.

Consideremos ahora la solución numérica del problema de valores iniciales (8). Si se convierte al problema de valor inicial (3), entonces

$$\begin{matrix} x(t) \\ {}^{x}\dot{}(t) \end{matrix} = U, H = \begin{matrix} 0 & I3\times3 \\ -K & 0 \end{matrix}, FH(t) = \begin{matrix} 0 \\ F(t) \end{matrix}, U0 = \begin{matrix} x0 \\ x1 \end{matrix}\ . \tag{66}$$

Por

$$Wi,H = e^{\tau H} \approx \begin{matrix} \cos p(\tau A) & \tau\ sen1,p(\tau A) \\ -K\tau\ sen1,p(\tau A) & \cos p(\tau A) \end{matrix}, \quad A = \sqrt{K}, \tag{67}$$

dónde

$$sin1\ A = \sum_{(2j+1)!}^{\infty} \frac{(-1)}{j=0} {}^{j} A\ 2j = \sum_{j=0}^{\infty} \frac{(-1)}{(2j+1)!}{}^{j} K\ j\ , \tag{68}$$

y (68), obtenemos el siguiente algoritmo.

Algoritmo 2.

$$\begin{matrix} \Delta t = \tau/3N;\ Ta = I;\ z = yo;\ A2 = (A\Delta t) \times (A\Delta t);\ \times zz;\ T\alpha \\ A2\ para(i = 1, i \le m, i + +)\ (zz = -(2i)(2i+1)\ _____ = T\alpha + zz), \end{matrix} \tag{69}$$

y

$$para(i = 0, i < N, i + +)\ 4\ (zz = T\alpha \times T\alpha, \quad zz = zz \times A2,\ A2 = 9A2,\ T\alpha = T\alpha - 3 \quad - T\alpha \times zz)\ , \tag{70}$$

donde mp y Np se determinan usando (31).

Algoritmo 3. Algoritmo de solución numérica del problema de valor inicial (8):

$$\begin{matrix} xk = (\cos(\tau A)xk-1 + \tau\ \sin1(\tau A)yk-1)\ wi\ (2j+1+bi)\tau \\ + \frac{\tau}{2nn} \sum_{=0}^{nn-1} {}^{j} \sum_{yo=1}^{3} \frac{(2j+1+bi)\tau\ (2j+1+bi)\tau}{} _____ A\ F(k\tau -)\ sen1\ 2nn\ 2nn\ 2nn \\ yk = -A\ \sin(\tau A)xk-1 + \cos(\tau A)yk-1\ (2j+1+bi)\tau \\ + \frac{\tau}{2nn} \sum_{=0}^{nn-1} {}^{j} \sum i=1^{3}\ wi\ \cos \frac{(2j+1+bi)\tau}{} A\ F(k\tau -)\ 2nn\ 2nn _____ \end{matrix} \tag{71}$$

Machine Translated by Google

para nn = 1, 2, · · · , donde A = $\sqrt{K}$. En particular, si nn se determina usando (64), el valor teórico la precisión de este algoritmo es p y

$$\begin{aligned} xk &= \cos(\tau A)xk-1 + \tau \sin 1(\tau A)yk-1 + \frac{1}{\tau 2nn}\sum_{j=0}^{nn-1}\sum_{i=1}^{3} wiTWsin\ F(k\tau -), j,yo\ \frac{(2j+1+bi)\tau}{2nn} \\ yk &= \cos(\tau A)yk-1 - A\sin(\tau A)xk-1 + \frac{1}{\tau 2nn}\sum_{j=0}^{nn-1}\sum_{yo=1}^{i=1\ 3} F(k\tau -), wiTWcos\ j,i\ \frac{(2j+1+bi)\tau}{2nn} \\ x0 &= x(0), y0 = x'(0), \end{aligned} \tag{72}$$

dónde

$$\begin{aligned} TWsin\ j,i &= \frac{(1+bi)\tau}{2nn}\ porque\ \frac{j\tau}{nn}\ A\ sen1,mp\ \frac{(1+bi)\tau}{2nn}\ A \\ &\quad + \frac{j\tau}{nn}\ pecado1\ \frac{j\tau}{nn}\ un\ cosmpo\ \frac{(1+bi)\tau}{2nn}\ A\ , \\ TWcos_{j,i} &= porque\ \frac{j\tau}{nn}\ un\ cosmpo\ \frac{(1+bi)\tau}{2nn}\ un \\ &\quad - sen\ j\tau\ \frac{}{nn}\ un\ sinmp\ \frac{(1+bi)\tau}{2nn}\ un\ , \end{aligned} \tag{73}$$

y

$$\begin{aligned} sinmp\ (sA) &= A\sum_{l=0}^{diputado}\frac{(-1)^{\cdot}s\ 2l+\ 1Kl}{(2l+1)!}, \\ sen1,mp\ (sA) &= \sum_{l=0}^{p.f.}\frac{(-1)^{\ \cdot}s\ 2lK\ l}{(2l+1)!}, \\ cosmp\ (sA) &= I + \sum (2l)!\ l=1^{mp+1}\frac{(-1)\quad l\ s\ 2lK\ l}{}. \end{aligned} \tag{74}$$

Si A es reversible, entonces

$$\begin{aligned} xk &= \cos(\tau A)xk-1 + A^{-1}\ pecado(\tau A)yk-1 \\ &\quad + \frac{\tau}{2nn}A^{-1}\sum_{j=0}^{nn-1}\sum i=1^{3}\ wi\ pecado\ \frac{(2j+1+bi)\tau}{}\ \frac{(2j+1+bi)\tau}{A\ F(k\tau -)\ 2nn\ 2nn} \\ yk &= -A\sin(\tau A)xk-1 + \cos(\tau A)yk-1\ (2j+1+bi)\tau\ (2j+1+bi)\tau\ A\ F(k\tau \\ &\quad + \frac{\tau}{2nn}\sum_{j=0}^{nn-1}\sum i=1^{3}\ wi\cos\ \frac{}{-).\ 2nn\ 2nn} \end{aligned} \tag{75}$$

Por lo tanto, (75) se puede escribir en forma de función compleja de la siguiente manera:

$$zk = e^{-i\tau A}zk-1 + \frac{i\tau}{2nn}\ un\ \sum_{j=0}^{nn-1}\ {}_{mi}{}^{-}\ \frac{j\tau i}{nn}\ A\ \sum l=1^{3}\ wl\ e^{-\frac{(1+b\ l)\tau i}{}\ (2j+1+bl)\tau\ 2nn\ AF\ \frac{k\tau -}{2nn}}}, \tag{76}$$

o

$$Zk = e^{-i\tau A}Zk-1 + \frac{\tau}{2nn}\sum_{j=0}^{nn-1}\ {}_{mi}{}^{-}\ \frac{j\tau i}{nn}\ A\ \sum l=1^{3}\ w\ e^{-\frac{(1+b\ l)\tau i}{}\ (2j+1+bl)\tau\ 2nn\ AF\ k\tau - 2nn}}, \tag{77}$$

donde yo = $\sqrt{-1}$,

$$\begin{aligned} zk &= xk + iA-1yk\ ,\ xk = Rezk\ ,\ yk = AImzk\ , \\ Zk &= yk - iAxk\ ,\ yk = ReZk\ ,\ xk = -A^{-1}\ EstoyZk\ . \end{aligned} \tag{78}$$

Primero proporcionamos ejemplos para ilustrar la efectividad del Algoritmo 1. Aunque omitimos ejemplos del Algoritmo 2, luego incluimos un ejemplo de la efectividad del Algoritmo 3.

Ejemplo 1. En (8), sea

$$k = \begin{matrix} 1 & -1 & 5 & 2 & 1 \\ -1 & & & & \frac{4}{} \end{matrix}\qquad F(t) = \begin{matrix} pecado\ t \\ 4\ pecado\ t\ 2 \\ \sqrt{2t}\ \cos\sqrt{3}t + sen\ t\ \cos\ 2 \end{matrix}\ ,$$

$$x0 = \frac{5}{2}\quad 0\text{-}2\quad t\ ,\ x1 = 1\ 1\ 2\quad t\ . \tag{79}$$

Machine Translated by Google

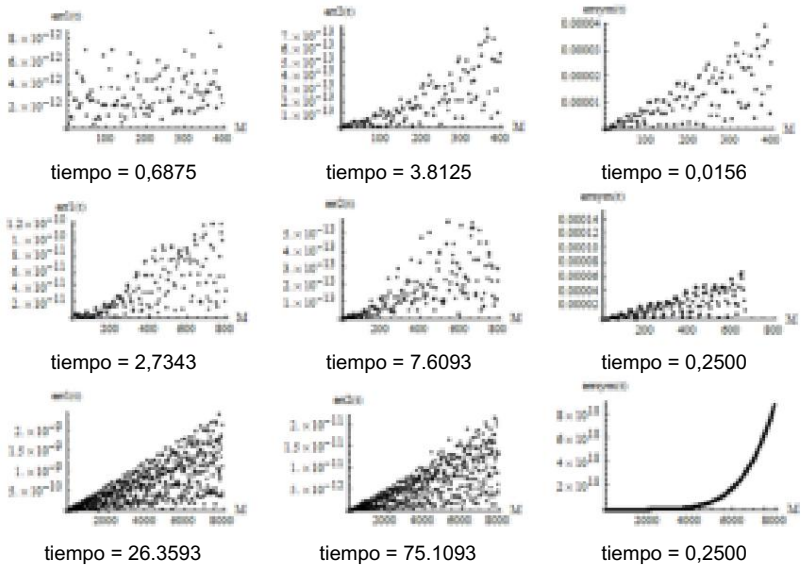
La solución analítica del problema de valor inicial (8) es

$$x(t) = \frac{\sqrt{2}\cos(\sqrt{3}t) + \sqrt{2}\sin(t/2) - \cos(\sqrt{3}t) + \sin(t)}{\cos(2t) - 2\sin(2t)}. \quad (80)$$

Sean $err_{i,k} = x(k\tau) - x_{i,k}$ $i = 1, 2,$ donde $x1,k$, $x2,k$ $(k = 0, 1, 2, \cdots, N)$ son las soluciones numéricas del valor inicial problema (8) determinado por el Algoritmo 1 y el Algoritmo 3, respectivamente. $errsym,k = x(k\tau) - xsym(k\tau)$ donde $xsym(t)$ es la solución numérica obtenida al llamar a la función NDSolve en Mathematica. Los resultados numéricos de $err1,k$, $err2,k$, $errsym,k$ $[0, T]$ $(k = 0, 1, 2, \cdots, N\tau = T)$ son los siguientes.

Se puede ver en la Figura 7 que la precisión del cálculo del Algoritmo 1 es mayor que la del Algoritmo 3, pero la velocidad de cálculo del Algoritmo 3 es tres veces mayor que la del Algoritmo 1. Sin embargo, el error de los resultados numéricos de las funciones internas de Mathematica es mucho mayor. que el error de los algoritmos 1 y 3. En particular, el colapso se calcula en t = 667,787. Después de este punto, el error aumenta exponencialmente con el tiempo t. Sin embargo, los algoritmos 1 y 3 son bastante estables. Por ejemplo, el error de cálculo en [0, 8000] sigue siendo muy pequeño.

En los algoritmos 1 y 3, sean $nn = \mu nnA$, $\mu = 0{,}5, 0{,}25, 0{,}125$, donde nnA se calcula según (64). Los resultados numéricos son los siguientes.



tiempo = 0,6875 tiempo = 3.8125 tiempo = 0,0156

tiempo = 2,7343 tiempo = 7.6093 tiempo = 0,2500

tiempo = 26.3593 tiempo = 75.1093 tiempo = 0,2500

Figura 7. Error de solución numérica (8) para varios algoritmos (nm se calcula según (64)).

En comparación con las Figuras 7 y 8, cuando $\mu = 0{,}5$, la precisión del cálculo no se reduce mucho y la velocidad de cálculo casi se duplica. Para $\mu = 0{,}125$ Algoritmos 1 y 3, la precisión del cálculo sigue siendo alta y la velocidad de cálculo es 10 veces mayor que la de la Figura 8.

Ejemplo 2. En (8), sea $K = (ki,j)n\times n$ una matriz tridiagonal.

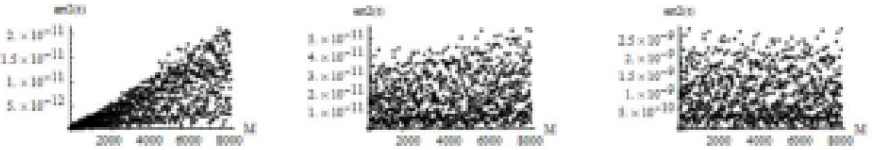
$x0,n)$ y se genera $f(t) = A\cos(\omega t) + B\sin(\omega t) + f1(t)$, $f2(t)$, donde $A = (A1, A2, \cdots, An)$, $B = (B1, B2, \cdots, Bn)$, $x0 = (x0,1, x0,2, \cdots, x1 = (x1,1, x1,2, \cdots, x1,n)$,

$|ki,yo - 9| \le 1$, $ki,i+1 - \frac{3}{2} \le \frac{1}{2}$, $ki+1,i - \frac{3}{2} \le \frac{1}{2}$, $Ai - \le \frac{3}{2} 0 \le \le i \le 2$, $\omega \frac{6}{2} - |x0,i| \le 1$, $|x1,i| \frac{\pi}{2}$, $1, 2, \cdots, norte)$.

Machine Translated by Google

Para n = 50, 100, 200 y 400, los algoritmos 1 y 3 proporcionan los siguientes resultados numéricos .

tiempo = 6,4687, μ = 0,5 tiempo = 3,4062, μ = 0,25 tiempo = 2,0625, μ = 0,125



tiempo = 39,2500, μ = 0,5 tiempo = 20,0312, μ = 0,25 tiempo = 2,9375, μ = 0,125

Figura 8. Error de solución numérica (8) para varios algoritmos.

En la Figura 9, (1) a (4), la izquierda son los dos primeros componentes de la solución numérica y la derecha es la diferencia entre las soluciones numéricas obtenidas por los dos algoritmos. Cabe señalar que a veces ωi (i = 1, 2, · · · , n) está cerca de la frecuencia de la solución y puede producirse aleteo o resonancia. Sin embargo, los valores obtenidos por los dos algoritmos todavía son muy cercanos.

(1) n = 50, tiempo6,1 = 31,6250, tiempo6,4 = 11,2500 (2) n = 100, tiempo6,1 = 60,5625, tiempo6,4 = 27,8125

(3) n = 200, tiempo6,1 = 137,5625, tiempo6,4 = 57,6875 (4) n = 100, tiempo6,1 = 281,3125, tiempo6,4 = 151,2187

Figura 9. Soluciones numéricas y errores de dos algoritmos.

7. Conclusiones

En este artículo, establecimos algoritmos para el cálculo de funciones matriciales exponenciales, funciones matriciales trigonométricas y funciones matriciales hiperbólicas. Nuestros resultados numéricos muestran que estos algoritmos son más rápidos y con mayor precisión en comparación con los algoritmos existentes y los algoritmos del sistema en Mathematica. En la expansión en serie de la función matricial, el número de elementos expandidos mp solo está relacionado con el requisito de precisión del cálculo y no tiene nada que ver con la matriz en sí. Sin embargo, el número de veces en el proceso iterativo Np no sólo está relacionado con la matriz misma sino que también está estrechamente relaciona Cuando |3 npλAt| ≤ 1(|2 npλAt| ≤ 1), Np = 0, el cálculo de la función matricial se convierte en una operación algebraica, por lo que la integral en (9)–(12) se convierte en una integral fácil de calcular. En resumen, nuestro algoritmo permite a los usuarios establecer un nivel requerido de precisión y luego usarlo para elegir valores apropiados para los parámetros my N para minimizar el tiempo de cálculo. De esta manera, se abordan simultáneamente tanto la precisión como la eficiencia.

Contribuciones de los autores: Conceptulación, HQ y YL; metodología, HQ y YL; software, sede; validación, HQ y YL; análisis formal, HQ y YL; investigación, HQ y YL; visualización, HQ; supervisión, sede; administración del proyecto, YL Todos los autores han leído y aceptado la versión de publicación del manuscrito.

Financiamiento: Esta investigación no recibió financiamiento externo.

Declaración de la Junta de Revisión Institucional: No aplicable.

Machine Translated by Google

Declaración de Consentimiento Informado: No aplicable.

Declaración de disponibilidad de datos: los datos están contenidos en el artículo.

Conflictos de intereses: Los autores declaran no tener ningún conflicto de intereses.

Referencias

1. Ruiz, P.; Sastre, J.; Ibáñeza, J.; Defez, E. Computación de alto rendimiento de la matriz exponencial. J. Computación. Matemáticas Aplicadas. 2015, 291, 370–379. [Referencia cruzada]
2. Molert, C.; Préstamos, CV Diecinueve formas dudosas de calcular la exponencial de una matriz. SIAM Rev. Soc. Indiana Matemáticas Aplicadas. 1978, 20, 801–836.
3. Higham, NJ Funciones de matrices: teoría y computación; SIAM: Filadelfia, PA, EE. UU., 2008.
4. Al-Mohy, AH; Higham, Nueva Jersey Cálculo de la acción de la matriz exponencial, con una aplicación a integradores exponenciales. SIAM J. Ciencias. Computadora. 2011, 33, 488–511. [Referencia cruzada]
5. Higham, NJ El método de escalado y cuadratura para la matriz exponencial resistió. SIAM J. Matriz anal. Aplica. 2005, 26, 1179–1193. [Referencia cruzada]
6. Bader, P.; Blanes, S.; Casas, F. Calcular la matriz exponencial con una aproximación optimizada del polinomio de Taylor. Matemáticas 2019, 7, 1174. [CrossRef]
7. Seydaoglu, KM; Bader, P.; Blanes, S.; Casas, F. Calcular la matriz seno y coseno simultáneamente con un número reducido de productos. arXiv 2020, arXiv:2010.00465v1.
8. Hargreaves, GI; Higham, NJ Algoritmos eficientes para la matriz coseno y seno. Número. Algoritmos 2005, 40, 383–400. [Referencia cruzada]
9. Alonso, P.; Ibáñez, J.; Sastre, J.; Peinado, J.; Defez, E. Algoritmos eficientes y precisos para calcular matrices trigonométricas. funciones. J. Computación. Aplica. Matemáticas. 2017, 309, 325–332. [Referencia cruzada]
10. Alonso, P.; Peinado, J.; Ibáñez, J.; Sastre, J.; Defez, E. Computación de funciones trigonométricas matriciales con GPU a través de Matlab. J. Supercomputadora. 2019, 75, 1227–1240. [Referencia cruzada]
11. Al-Mohy, AH; Higham, Nueva Jersey; Relton, SD Nuevos algoritmos para calcular la matriz seno y coseno por separado o simultáneamente . SIAM J. Ciencias. Computadora. 2015, 37, A456–A487. [Referencia cruzada]
12. Serbian, SM; Blalock, SA Un algoritmo para calcular la matriz coseno. SIAM J. Ciencias. Estadístico. Computadora. 1980, 1, 198–204. [Referencia cruzada]
13. Cooper, BL; Bingulac, S. Mejora computacional en el cálculo del logaritmo natural de una matriz cuadrada. Electrón. Letón. 1990, 26, 861–862. [Referencia cruzada]
14. Wu, F.; Zhu, L.; Shi, Q. Método computacional eficiente para la función matricial en problemas dinámicos. Acta Mec. Pecado. 2023, 39, 522451. [Referencia cruzada]
15. Cardoso, JR; Sadeghi, A. Una técnica para mejorar el cálculo de funciones de matrices triangulares. Int. J. Computación. Matemáticas. 2022, 99, 2449–2465. [Referencia cruzada]
16. Jiang, Y.; Él, C.; Qin, H. Un algoritmo acelerado para alta precisión y cálculo rápido de funciones elementales básicas. Matemáticas. Practica. Teoría 2017, 47, 238–246.
17. Fasi, M.; Higham, Nueva Jersey Algoritmo de multiprecisión para calcular logaritmos matriciales. SIAM J. Matrizanal. Aplica. 2018, 39, 472–491. [Referencia cruzada]
18. Ibáñez, J.; Sastre, J.; Ruíz, P.; Alonso, JM; Defez, E. Un algoritmo de Taylor mejorado para calcular el logaritmo matricial. Matemáticas 2021, 9, 2018. [CrossRef]
19. Cheng, SH; Higham, Nueva Jersey; Kenney, CS; Laub, AJ Aproximación del logaritmo de una matriz con una precisión especificada. SIAM J. Matriz anal. Aplica. 2001, 22, 1112–1125. [Referencia cruzada]
20. Kenney, C.; Laub, AJ Estimaciones de condición para funciones matriciales. SIAM J. Matriz anal. Aplica. 1989, 10, 191–209. [Referencia cruzada]
21. Kenney, C.; Laub, AJ Padé estimaciones de error para el logaritmo de una matriz. Internacional. J.Control. 1989, 50, 707–730. [Referencia cruzada]
22. Milovanovic, GV; Ran˘ci´c, SL Algunas mejoras en el cálculo del logaritmo natural de una matriz cuadrada. En Actas de la Tercera Conferencia Internacional sobre Análisis Funcional y Teoría de Aproximación, Aquafredda di Maratea, Potenza, Italia, 23 a 28 de septiembre de 1996; Número 52, Volumen II.

Descargo de responsabilidad/Nota del editor: Las declaraciones, opiniones y datos contenidos en todas las publicaciones son únicamente de los autores y contribuyentes individuales y no de MDPI ni de los editores. MDPI y/o los editores renuncian a toda responsabilidad por cualquier daño a personas o propiedad que resulte de cualquier idea, método, instrucción o producto mencionado en el contenido.

Machine Translated by Google

AppliedMath

MDPI

Artículo

Algunos comentarios sobre el binomio negativo p-generalizado (PNB) Modelo

Daniel Griffith

Facultad de Ciencias Económicas, Políticas y Políticas, Universidad de Texas en Dallas, Richardson, TX 75080, EE. UU.; dagriffith@utdallas.edu

Resumen: Este artículo describe varias propiedades y características seleccionadas de variables aleatorias binomiales negativas (NB), con especial referencia a las especificaciones NB2 (es decir, p = 2) y algunas generalizaciones a NBp (es decir, p ≥ 2). Presenta nuevos resultados (por ejemplo, la función generadora de momentos NBp) con respecto a la relación entre una media muestral y su varianza acompañante, así como ejemplos numéricos y empíricos estadísticos/econométricos espaciales, cuyos parámetros estimadores son la máxima verosimilitud o el método del momento. unos. Finalmente, destaca la metodología de filtrado espacial de vectores propios de Moran dentro del contexto del modelado lineal generalizado, demostrándola en términos de regresión binomial negativa espacial. Su conclusión general refuerza los hallazgos importantes que la literatura ya informa con un ejemplo empírico recientemente reconocido de un fenómeno NB3.

Palabras clave: filtro espacial de vector propio; función generadora de momentos; binomio negativo; densidad de población ; Puerto Rico

1. Introducción

Muchas variables aleatorias (RV) contemporáneas miden recuentos (con un soporte de cero a infinito), lo que implica que son variables de Poisson, pero sin equidispersión (es decir, una violación de $\mu = \sigma^2$), lo que generalmente se denomina sobredispersión. Una propiedad general de esta descripción distributiva es una variable de Poisson con una media distribuida gamma (es decir, una mezcla gamma-Poisson), que se traduce muy bien en un binomio negativo (NB)
RV. Sin embargo, un NB estándar no siempre explica con éxito esta discrepancia entre media y varianza en tales circunstancias. Cameron y Trivedi [1] introducen una versión generalizada del modelo de probabilidad NB ahora conocido como modelo NBp, donde p es un entero positivo, cuyo objetivo es capturar mejor la sub/sobredispersión; esta variable difiere del modelo de proceso binomial negativo (PNB) (por ejemplo, [2]). La literatura contiene poca discusión sobre este antiguo modelo generalizado, con la excepción del tratamiento inicial que le dieron Cameron y Trivedi y los artículos de Greene [3] y Di et al. [4]. Es de destacar que existe un paquete R (versión ≥ 3.00) para implementar este modelo [4]. El propósito de este artículo es abordar esta brecha en la literatura (1) aclarando/corrigiendo varios detalles técnicos engañosos en parte del artículo de Greene; (2) presentar nuevos resultados; (3) iluminar propiedades de muestra pequeñas y grandes seleccionadas en modelos NB; y (4) presentar nuevos casos empíricos de estadística espacial/econometría que impliquen autocorrelación espacial (SA).

El contexto de esta discusión es una conceptualización de mezcla paramétrica, análoga a una formulación anterior conjugada bayesiana: la especificación NB describe una variable de conteo Y que tiene una distribución de Poisson, con su media μ distribuida gamma en lugar de constante. En consecuencia, el RV gamma tiene parámetros $\alpha\mu^{2-p}$ y μ^{p-1}/α, donde μ denota su media y $1/\alpha$ denota su parámetro de dispersión, lo que produce un modelo NBp global con parámetros $r = \alpha\mu^{2-p}$ y probabilidad $p = 1/(1 + \mu^{p-1}/\alpha)$ para algún entero no negativo $p > 0$; p = 1, 2 son especificaciones estándar de NB, siendo la primera la Poisson RV y la segunda la popular variante NB2.

check for updates

Cita: Griffith, DA Algunos Comentarios sobre el modelo p-binomial negativo generalizado (NBp). Matemáticas aplicadas 2024, 4, 731–742. https://doi.org/10.3390/appliedmath4020039

Editor académico: Tommi Sottinen

Recibido: 24 de marzo de 2024
Revisado: 3 de junio de 2024
Aceptado: 6 de junio de 2024
Publicado: 11 de junio de 2024

Copyright: © 2024 por el autor. Licenciatario MDPI, Basilea, Suiza. Este artículo es un artículo de acceso abierto. distribuido bajo los términos y condiciones de los Creative Commons Licencia de atribución (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

Machine Translated by Google

2. Materiales y métodos: Aclaración de la presentación de Greene

Las ecuaciones confusas que aparecen en Greene [1] incluyen (2-13) y (2-14), particularmente cuando se las contrasta con la ecuación del apéndice (A-2) con respecto al exponente p Además, a esta última ecuación le falta un subíndice i en la variable λ que aparece en el numerador de su tercer término en su lado derecho. La discusión relativa a las ecuaciones (2-4) y (2-5) sería más clara si la distribución gamma propuesta se expresara en su simbología estándar como Γ(θ, 1/θ). Las ecuaciones de Greene (2-13), en las que establece incorrectamente q = 1/1(1 + θ) en lugar de q = 1/(1 + 1/θ), y (2-14), en las que establece s = λ/(λ + θλ2–p) en lugar de 1/(1 + μ p–2/α): produce, para NBps independientes e idénticamente distribuidos (iid), la siguiente función de masa de probabilidad:

$$\frac{\Gamma\left(y+\alpha\mu^{3-p}\right)}{\Gamma(y+1)\,\Gamma\left(\alpha\mu^{3-p}\right)}\left(\frac{1}{1+\mu p^{-2/\alpha}}\right)^{\alpha\mu^{3-p}}\left(\frac{\mu p^{-2/\alpha}}{1+\mu p^{-2/\alpha}}\right)^{y},\ y=0,1,2,\ldots, \tag{1}$$

donde, adoptando en este artículo una notación más típica que aparece en la literatura (particularmente para las especificaciones gamma-Poisson), μ (que denota la media) es equivalente a λ y α es lo mismo que θ en el artículo de Greene, los subíndices i no se utilizan en este artículo. versión porque representa una ecuación global, y la relación del término de funciones gamma también se puede reescribir empleando coeficientes binomiales combinatorios. La diferencia es que la ecuación de Greene (2-14) tiene un exponente de (2–p) en lugar de (p–2). La ecuación (1) presenta los resultados revisados

$$E(Y)=\mu \text{ y } E[(Y-\mu)^{2}]=\mu(1+\mu p^{-1/\alpha}). \tag{2}$$

La especificación de la ecuación de Greene (2-14) representa

$$\frac{\Gamma\left(y+\alpha\mu^{2-p}\right)}{\Gamma(y+1)\,\Gamma\left(\alpha\mu^{2-p}\right)}\left(\frac{1}{1+\alpha\mu^{1-p}}\right)^{\alpha\mu^{2-p}}\left(\frac{\alpha\mu^{1-p}}{1+\alpha\mu^{1-p}}\right)^{y},\ y=0,1,2,\ldots,$$

que no es su ecuación (2-15), incluso para el caso estándar específico de p = 2. Su función de masa de probabilidad debería ser la Ecuación (1), en lugar de

$$\frac{\Gamma\left(y+\alpha\mu^{2-p}\right)}{\Gamma(y+1)\,\Gamma\left(\alpha\mu^{2-p}\right)}\left(\frac{1}{1+\alpha\mu^{1-p}}\right)^{\alpha\mu^{2-p}}\left(\frac{\alpha\mu^{1-p}}{1+\alpha\mu^{1-p}}\right)^{y},\ y=0,1,2,\ldots,$$

que no es ni el resultado de una probabilidad de mezcla paramétrica de Poisson definida por Prob(y) = e−μμ y/(y!) con μ ~ Γ(αμ2–p , μ p–1/α), ni el resultado del reemplazo de Greene α con αμ2–p en esta mezcla, lo que da como resultado E(Y) = μ y $E[(Y-\mu)^{2}]=\mu(1+\mu p^{-3/\alpha})$.

3. Resultados: Articulación de variaciones de NBp

Tanto los resultados analíticos de las estadísticas matemáticas como los resultados experimentales de la simulación ayudan a establecer una comprensión de una generalización de NB1 y NB2 a NBp. Esta sección proporciona algunos de cada uno.

3.1. La función generadora de momento NBp y la normalidad asintótica

Particularmente porque el método de los momentos es una técnica de estimación útil para el parámetro α [5], es deseable conocer los momentos de la especificación NBp. La ausencia de estimadores de máxima verosimilitud de forma cerrada (MLE, por ejemplo, ver [6-10]) hace que esta invención resultante sea aún más atractiva.

3.1.1. La función generadora de momento NBp (mgf)

El NBp mgf es

$$\left(\frac{1}{1+\mu p^{-1/\alpha}\left(1-e^{t}\right)}\right)^{\alpha\mu^{2-p}},$$

Machine Translated by Google

siendo el primer momento sobre el origen

$$\frac{\partial}{\partial t}\left(\frac{\alpha\mu^{2-p}}{\alpha\mu^{2-p}+\mu}\right)^{-\alpha^{-1}}\left(1-e^{t}\right)^{-\alpha^{-1}}\Bigg|_{t=0} = m,$$

y, por tanto, los momentos en los que el mgf medio es

$$\left(\frac{\alpha\mu^{2-p}}{\alpha\mu^{2-p}+\mu}\right)^{-\alpha^{-1}}\left(1-e^{t}\right)^{-\alpha^{-1}} = \left(1+\alpha\mu^{p-1}\right)^{-\alpha^{-1}}\left(1-e^{t}\right)^{-\alpha^{-1}} = \left(\frac{1}{1+\alpha\mu^{p-1}}\right)^{\alpha^{-1}}\left(1-e^{t}\right)^{-\alpha^{-1}} \quad (3)$$

El segundo momento respecto a la media es $\mu(1 + \mu p-1/\alpha)$.

, es visible en esta primera expresión anterior. Aquí 3.1.2.

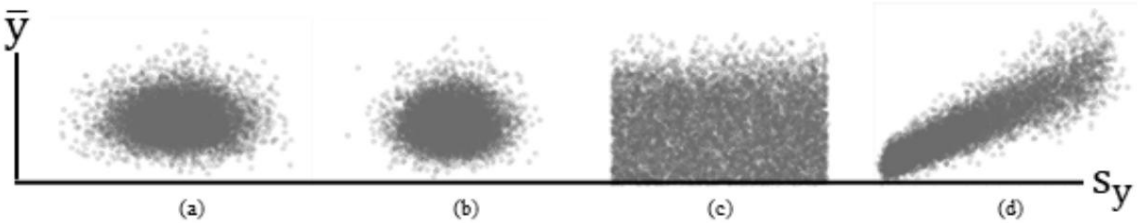
Normalidad asintótica y NBp RV

Hacer coincidir momentos es una forma de establecer equivalencia entre distribuciones [11].

asimetría = $\frac{1+2\mu p^{-1/\alpha}}{\sqrt{\mu}\,\sqrt{1+\mu p^{-1/\alpha}}}$, y curtosis = $3 + \frac{6\mu^{p-2}}{\alpha} + \frac{1}{\mu\left(1+\mu p^{-1/\alpha}\right)}$.

); Figura 1a]. La Figura 1 muestra diagramas de dispersión de El término de asimetría de Poisson, $1/\sqrt{\mu}$, es visible en esta primera expresión anterior. Aquí, la asimetría no necesariamente converge en cero, y la curtosis no necesariamente converge en 3, a menos que el parámetro de sobredispersión, 1/α, así como p exhiban ciertos comportamientos y/o propiedades. En otras palabras, la normalidad asintótica de la distribución NBp no es obvia. Bagui y Mehra [12] ofrecen formas de demostrar esta característica asintótica, una de las cuales implica su mgf.

Mientras tanto, un teorema clásico establece que si se extraen muestras aleatorias de la misma distribución normal (es decir, iid), entonces sus estadísticas de media y varianza muestral son independientes [porque s2 es una función de (yi − _ y); Figura 1a]. La Figura 1 muestra diagramas de dispersión de las relaciones entre la muestra s2 y s2 de los resultados del experimento de simulación para 10.000 replicaciones de muestras de tamaño n = 100 (un tamaño apropiado bien conocido para garantizar que el clásico Teorema del Límite Central esté en funcionamiento). Para este experimento, la correlación lineal y y s2 para un VD normal estándar (es decir, iid) es decir, iid) es –0,014 (Figura 1a), mientras que entre _ un VD NBp con p = 2, μ = 1000 y α = 1000 (es 0,027 (Figura 1b). Una extensión de este teorema convencional y bien conocido es la siguiente:



(a): Figura 1. Relaciones simuladas seleccionadas (10.000 replicaciones, n = 100) entre _ iid y y s2 . Izquierda RV normales. Centro izquierdo (b): iid NB2 RV. Centro derecho (c): vehículos recreativos normales. Derecha (d): i~id NBp RVs.

Teorema 1. La correlación lineal entre los RV y y s2 para una mezcla de independientes y no idénticas normales distribuidos (i~id) es cero.

Prueba del teorema 1. Para un RV normal, μ puede ser cualquier valor real y σ^2 puede ser cualquier número real positivo. Debido a que estos dos conjuntos de valores se pueden emparejar de cualquiera de las infinitas formas posibles que existen, cada valor de μ se empareja con valores relativamente grandes, intermedios y

Machine Translated by Google

valores pequeños de σ^2, y cada valor de σ^2 se combina con valores relativamente grandes, intermedios y pequeños de μ. En consecuencia, la correlación lineal entre estas dos variables es cero. □

Este resultado (ver Figura 1c; r = 0,001) contrasta con el de un NBp RV para el cual μ aparece en su fórmula de varianza, introduciendo una correlación lineal entre los dos (Figura 1d; r = 0,905). La extensión formal de la postulación aquí es la siguiente:

Teorema 2. La correlación lineal entre $\bar{y}$ y s^2 para una mezcla de i~id NBp RVs es > 0, $\frac{\sqrt{3(2p+1)}}{p+2}$ $p \geq 1$.

Prueba del teorema 2. El coeficiente de correlación del momento del producto lineal es

$$\rho_c = \frac{\sum \sigma_0 (\mu - \bar{A}_c)_0 \mu \, d\mu \, \ddot{a}/c \, \acute{o} \times (\mu_1 + \mu \, p-1/\alpha - \bar{A}_c)_0 \mu_1 + \mu \, p-1/\alpha \, d\mu \, \ddot{a}/c \, \acute{o} \, d\mu \, \copyright \ldots c_0 \mu}{d\mu \, \ddot{a}/c \, \acute{o}2 \, d\mu \times \ldots c_0 \mu_1 + \mu \, p-1/\alpha \, d\mu \, \ddot{a}/c \, \acute{o}2 \, d\mu_0 (\mu - \bar{A}_c)_0 (\mu_1 + \mu \, p-1/\alpha - \bar{A}_c)}.$$

$\lim_{c\to\infty} \rho_c = \frac{\sqrt{3(2p+1)}}{p+2}$ porque $p > 0$. □

Este nuevo teorema indica que si p = 1 (es decir, un RV de Poisson), entonces ρ = 1, y para el caso estándar NB2 de p = 2, ρ ≈ 0,968. El parámetro p debe ser 20 antes de que esta correlación lineal sea aproximadamente 0,5. El parámetro de dispersión α no afecta esta correlación lineal. Este nuevo teorema indica que para el NBp antes mencionado, la media y la varianza de la muestra, $\bar{y}$ y s^2 están correlacionadas (Figura 1d). En otras palabras, para una mezcla de distribuciones NBp, la media y la varianza de la muestra están correlacionadas, aunque esta correlación lineal permanece cero cuando se agrupan las distribuciones normales, puede llegar a 1 para las distribuciones NBp agrupadas.

Es de destacar que el Teorema 1 es válido incluso para una distribución normal restringida en tal forma en que $\sigma^2 = \mu(1 + \mu \, p-1/\alpha)$, con restricciones adecuadas tanto en p como en μ (por ejemplo, p debe ser par y el intervalo de dominio μ [–1, 0] deben eliminarse del rango de correlación σ^2 lineal). La correlación lineal ocurre para este caso particular de distribución normal, por ejemplo, imponiendo una restricción como μ > 0, que luego también permite p ser impares y lineales que se describirán mediante el teorema 2.

3.2. MLE y Método de Estimación de Momentos (MME) del Parámetro de Dispersión α

Existen varias técnicas de estimación, incluidas MLE, MME (p. ej., ver [6,13]) y mínimos cuadrados (ver [14-16]). Con frecuencia, estas técnicas producen exactamente los mismos estimadores. La estimación del parámetro de dispersión 1/α de un NBp a menudo se realiza mediante MLE y/o MME. Un μ más cercano a cero y un tamaño de muestra pequeño pueden crear una situación en la que un MLE y/o MME sea inexacto y/o inestable. MME es la más simple de las dos técnicas para estimar 1/α [5]. Un MLE no puede producir una solución que sea un estimador en forma cerrada para este parámetro NBp (ver Apéndice A).

3.2.1. Estimación del exponente del parámetro p

En lugar de establecer p = 2, la ecuación (1) alude a p como parámetro, aunque restringido a los números enteros positivos (es decir, discreto, no continuo). Como tercer parámetro, su estimación MME requiere, por ejemplo, agregar el tercer momento con respecto a la media, m3, al conjunto de ecuaciones de estimación. Respectivamente,

$$m = \sum_{yo=1}^{norte} yi/norte = \bar{y},$$

Machine Translated by Google

$$\alpha = \mu \sum_{i=1}^{n} \frac{\mu^{p}}{(y_i - \mu)^2/(n-1) - \mu}, \text{ y}$$

$$\frac{1 + 2\mu p^{-1}/\alpha}{\mu \left(1 + \mu p^{-1}/\alpha\right)} = \frac{\sum_{i=1}^{n} (y_i - \mu)^3}{n - 1}.$$

Resolver este triplete de ecuaciones para p produce una ecuación que no contiene p, ni tampoco la que incorpora el cuarto momento. Por tanto, el MME no puede dar una estimación de p.

Mientras tanto, un MLE, que aquí implica el cálculo de diferencias porque p es un número entero, no logra producir una solución de forma cerrada.

La implicación de este resultado es que generalizar el término de varianza estándar de NB2 desde $\mu(1 + \mu/\alpha)$ al término generalizado de NBp $\mu(1 + \mu p^{-1}/\alpha)$ es similar a la estimación de cuasi verosimilitud con respecto al parámetro p. . Denote parece demasiado grande, entonces $\frac{1}{\alpha}$ y μ^{-1} , es un k ésimo múltiplo de μ (es μ^{k}), donde k es un número entero positivo, se puede factorizar a partir de él, efectivamente decir, μ estableciendo p = 2 + k. Esta factorización no es constante entre los valores de μ_i , con el subíndice i denotando una observación individual, cuando la media varía de una observación a otra (esta es la situación que presenta Greene [1]).

3.2.2. Un ejemplo numérico simple n = 3

Considere una muestra de tamaño n = 3, siendo el número de recuentos por observación, Y como sigue: $y_i = \{2, 5, 8\}$. Entonces

$$\bar{y} = (2 + 5 + 8)/3 = 5 \text{ y } s^{2} = 18/(3 - 1) = 9 = 5\ 1 + 5/\ 25/4\ .$$

Los MLE son

$$\mu = 5 \text{ y } \alpha^{-1} \approx 0{,}0476.$$

Las MME son

$$\mu = 5 \text{ y } \alpha^{-1} = 0{,}1600.$$

La Tabla 1 resume los resultados de un experimento de simulación de Monte Carlo con 10.000 repeticiones (para invocar la Ley de los Grandes Números), para NB2[5, 5(1 + 5/0,0476)] y NB2[5, 5(1 + 5/0,1600)]

La Figura 2 muestra las estimaciones colectivas de MLE y MME para estos resultados de simulación.

Como era de esperar, ambas técnicas de estimación arrojan una estimación sólida de la media poblacional, μ. Las técnicas proporcionan diferentes estimaciones del parámetro de dispersión, aunque ambas se encuentran dentro del intervalo de confianza del 95% de la estimación de la otra. Los MLE son cero con más frecuencia que los MME. La mayoría de estos MME cero en realidad son estimaciones negativas que han sido reemplazadas por cero mediante truncamiento (estos valores cero se emplearon al calcular las estadísticas resumidas que aparecen en la Tabla 1). Estos valores deben considerarse inviables porque α es el parámetro de forma de un VD gamma y, como tal, debe ser positivo.

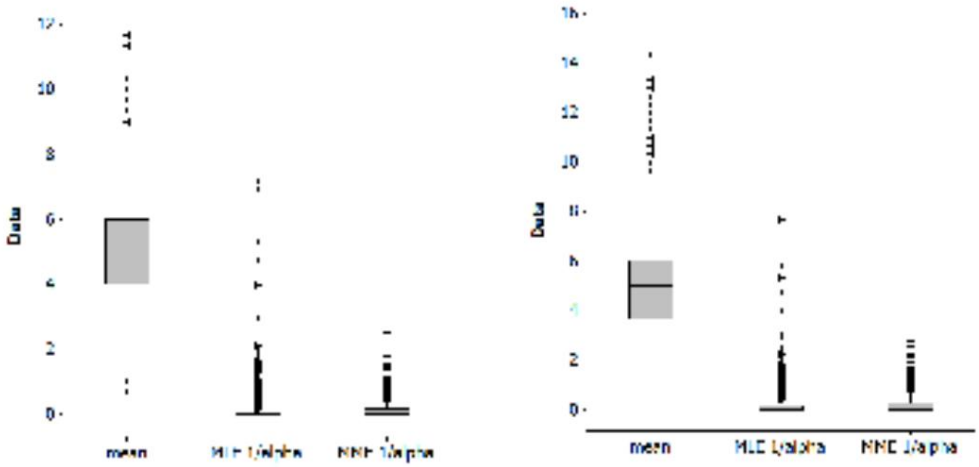
La Figura 2 proporciona un diagrama visual de caja y bigotes: los bordes superior e inferior de un cuadro indican el primer y tercer cuartil, la línea interna del cuadro indica la mediana y las puntas de los bigotes indican 1,5 veces el rango intercuartil (es decir, el tercero menos los valores del primer cuartil): comparación entre los dos estimadores (es decir, los MLE y MME), destacando que los MLE tienen más valores atípicos (indicados por asteriscos). El exceso de ceros y la frecuencia de estimaciones atípicas para las MLE y las estimaciones negativas para las MME se encuentran entre las razones por las que estos estimadores se consideran inexactos y/o inestables. Una implicación general de estos resultados es que ninguno de los estimadores es particularmente bueno para muestras pequeñas, pero que el MME bien puede ser preferible. Este pequeño caso de muestra ilustra que un MLE y un MME pueden producir estimaciones diferentes (ver también [17]).

Machine Translated by Google

Tabla 1. Resumen de los resultados del experimento de simulación Monte Carlo para NB2[5, 5(1 + 5/0.0476)] y NB2[5, 5(1 + 5/0,1600)].

Parámetro	Estimar	Estándar Desviación	Rango	% 0 s	% < 0
		utilizando MLE como parámetros de población			
μ	4.984	1,410	0,7–11,7		
1/αMLE 1/	0,073	0.231	0,0–7,2	71.1	
αMME	0.113	0.213	0,0–2,7	5.4	52.2
		utilizando MME como parámetros de población			
μ	5,003	1,714	0,3–14,3		
1/αMLE	0,137	0,315	0,0–7,7	0,295 0,0–2,8	58,4
1/αMME	0,190			1.3	40,6

, PARA REVISIÓN POR PARES



(a) (b)

Figura 2. Diagramas de caja del experimento de simulación de Monte Carlo (como es habitual, los asteriscos indican valores atípicos). Izquierda (a): NB2[5, 5(1 + 5/0,0476)]. Derecha (b): NB2[5, 5(1 + 5/0,1600)].

3.2.3. Filtrado espacial de vectores propios de Moran: una breve descripción general

La siguiente sección presenta un ejemplo sencillo de interés, especialmente para los criminólogos, ecologistas, empresarios y comercializadores, científicos ambientales, epidemiólogos, analistas geoespaciales, sociólogos, planificadores urbanos y urbanistas, y economistas urbanos y regionales, entre otros, a saber, la densidad de población, quizás uno de los factores más utilizados covariables en las ciencias sociales y del comportamiento; Sirve como prefacio a la Sección 4. Los datos georreferenciados requieren una metodología de estadística espacial/econometría para poder procesarlos adecuadamente. dar cuenta de SA latente en esos datos. SA es la correlación entre observaciones que surge. debido a su relativa proximidad en el espacio geográfico, porque o bien se ven afectados por algún factor común o interactúan directamente entre sí a través del espacio geográfico. Es factor mon o interactúan directamente entre sí en el espacio geográfico. Es análogo a, pero más complejo, la correlación serial en datos de series de tiempo (por ejemplo, ver [18]).

El filtrado espacial de vectores propios de Moran (MESF) es una metodología relativamente novedosa para El filtrado espacial de vector propio de Moran (MESF) es una metodología relativamente novedosa para analizar el SA contenido en los datos [19-21]. MESF utiliza un conjunto de variables proxy sintéticas, que son extraídos como vectores propios de una matriz de pesos espaciales (SWM) n por n doblemente centrada, Matriz de pesos espaciales (SWM), digamos C, que a menudo es binaria (0-1) y que une n objetos geográficos en el espacio (indicando cuáles están directamente correlacionados por pares mediante una entrada de uno), y luego los suma. vectores como variables de control para una especificación de modelo de regresión; Los SWM son análogos a vectores como variables de control para una especificación de modelo de regresión; SWM Matrices de adyacencia de gráficos planos no dirigidos. Estas variables de control identifican y aíslan dependencias espaciales estocásticas entre un conjunto de observaciones georreferenciadas, filtrando estas dependencias espaciales estocásticas entre un conjunto de observaciones georreferenciadas, filtrando estas dependencias de los residuos de un modelo y agregándolas a la respuesta media del modelo, permitiendo así que la construcción del modelo de regresión proceda con observaciones georreferenciadas que imitar ser independiente.

SWM centrado explotado por MESF. Este índice se puede escribir en forma matricial de la siguiente manera:

$$\underline{\quad} \quad \underline{\quad} \quad \underline{\quad}$$

Machine Translated by Google

El índice SA del coeficiente de Moran (MC; [22]) , uno de varios índices de este tipo que existen (por ejemplo, el índice Geary (GR; [23]), proporciona la base para el MESF; su numerador contiene el SWM doblemente centrado explotado por el MESF. Este índice se puede escribir en forma matricial de la siguiente manera:

$$(n/1^{t} C1)Y^{t} (I - 11T/n)C(I - 11T/n)Y/Y T \qquad (yo - 11T/n)Y, \tag{4}$$

donde n es el número de, por ejemplo, polígonos en un archivo de forma de sistema de información geográfica; Y es un vector n por 1 de valores de atributos; I es una matriz de identidad de n por n; 1 es un vector de unos de n por 1; El superíndice T denota transposición de matriz y, aquí, C es un SWM binario 0–1 para el cual 1 denota que las unidades de área de fila y columna (es decir, ubicaciones, por ejemplo, polígonos de archivos de forma) comparten un límite común de longitud distinta de cero, y 0 denota que no lo hacen (es decir, la definición de adyacencia de la torre, fraseología basada en las piezas del juego de ajedrez y sus movimientos). Las funciones propias (es decir, los valores propios y los vectores propios emparejados) se extraen de la siguiente matriz, que está doblemente centrada porque la matriz C se multiplica antes y después por la matriz de proyección (I − 11T/n), que aparece en el numerador de la expresión. (4) de la siguiente manera:

$$(I - 11T/n)C(I - 11T/n). \tag{5}$$

Cuando se multiplica por (n/1 TC1), un valor propio de la matriz (5) se convierte en el MC que mide el SA en su patrón de mapa de vector propio asociado de elementos de números reales [24,25]. El signo de un valor propio indica la naturaleza de SA representada por su correspondiente vector propio, mientras que su magnitud indica el grado de SA. Los valores propios extremos de la expresión matricial (4) determinan los límites de MC, que no son necesariamente ±1 [26].

Extraer las funciones propias de los SWM ajustados constituye una descomposición espectral de esas matrices. Los vectores propios extraídos con valores propios asociados relativamente lejos de cero y, por lo tanto, que representan un SA nada despreciable, pueden considerarse como una representación de componentes globales, regionales o locales de SA debido a los patrones de mapa particulares que representan cuando se visualizan [27,28] . En otras palabras, SA se manifiesta en términos de valores similares (SA positivo; PSA) o diferentes (SA negativo; NSA) de la agrupación de la variable Y en un mapa: puntos calientes (es decir, agrupaciones similares de valores relativamente altos), puntos fríos (es decir, , agrupación de valores relativamente bajos similares) o contrastes (es decir, agrupación de valores relativamente dife

3.2.4. Una ilustración empírica de juguetes sobre la densidad de población de Puerto Rico en 2010

Puerto Rico está dividido en cinco regiones administrativas agrícolas, según un esquema que se le ha aplicado durante un siglo o más, cada una de las cuales se centra en una de las principales áreas urbanas de la isla (Figura 3). La distribución geográfica de la densidad de población de 1899 para esta resolución espacial muestra una variación extra-Poisson excesiva (desviación = 1723, que es mucho mayor que su valor esperado de 1) y un PSA moderado. La tabla de la Figura 3b informa el vector propio de SWM ajustado que representa este PSA [MC = 0,221, siendo su máximo (MCmax) 1,094 y GR = 0,474]. Este escenario reconoce una fuente bien reconocida de sobredispersión, a saber, la correlación entre observaciones (es decir, datos dependientes, autocorrelación), que es frecuente en las estadísticas espaciales; por ejemplo, SA representa aproximadamente la mitad de cualquier sobredispersión detectada (por ejemplo, variación extra-Poisson).

Machine Translated by Google
Vector propio SWM que representa este PSA [MC = 0,221, con su máxima

(a)

region	SJ	A	M	P	C	eigenvector
SJ: San Juan	0	1	0	1	1	0.37175
A: Arecibo	1	0	1	1	0	−0.37175
M: Mayagüez	0	1	0	1	0	−0.60150
P: Ponce	1	1	1	0	1	0.00000
C: Caguas	1	0	0	1	0	0.60150

(b)

Figura 3. Las regiones administrativas agrícolas (AAR) de Puerto Rico de 1899; indicadas por bordes rojos. Las líneas negras denotan los límites de los municipios. Mapa de coropletas terciles: negro, gris y blanco, respectivamente, denotan una densidad de población relativamente alta, intermedia y relativamente baja. Izquierda (a): mapa de coropletas del municipio-AAR. Derecha (b): la matriz de pesos espaciales AAR y su vector propio relevante.

Una descripción NB2 de esta distribución geográfica de la densidad de población produce la MLE global de μ = e5,6361 ≈ 280,4 personas por milla cuadrada, con un parámetro de dispersión MLE de α^{-1} ≈ 0,0010 (es decir, casi un Poisson RV). La especificación del filtro espacial de vector propio (ESF), ≈ 280,4 personas por milla cuadrada, con un parámetro de dispersión MLE

-1 ≈ 0,0010 (es decir, casi un Poisson RV). El filtro espacial de vector propio (ESF) describe una combinación lineal de un subconjunto de vectores propios extraídos del SWM ajustado, específicamente una densidad de población media que varía espacialmente por región [es decir, 283,6 (SJ), 278,6 (A), 295,7 (M), 271,6 (P) y 272,4 (C)] alrededor de una media global de 280,3 y reduce el exceso de SWM, describe una densidad de población media que varía espacialmente por región [es decir, 283,6 (S parámetro de dispersión a 0,0006 (aún más parecido a un Poisson RV).

El MME global de μ también es e5.6361 ≈ 280,4. Sin embargo, el parámetro de dispersión MME de α^{-1} es –0.0001, lo que resultaría en que se redondearía a 0. Finalmente, la variación espacial ≈ 280,4. Sin embargo, el parámetro de dispersión MME

-1 de la densidad de población media por región es el siguiente: 280,1 (SJ), 280,6 (A), 280,8 (M), 280,4 (P), y 280,0 (C). Estos resultados parecen inferiores a sus homólogos MLE anteriores.

4. Discusión: dos ejemplos empíricos del MESF

Los ejemplos anteriores en las Secciones 3.2.2 y 3.2.4 son sólo para fines de ejemplificación. Esta sección presenta dos ejemplos empíricos cuyos resultados proporcionan información implicaciones sustantivas.

4.1. Un estudio de caso empírico sobre la densidad de población urbana de Puerto Rico en 2010

La Figura 4a muestra la distribución geográfica de la densidad de población urbana (personas por milla cuadrada) en 2010 en la isla principal de Puerto Rico. El patrón del mapa que caracteriza esta distribución geográfica tiene un MC de 0,487 y un GR de 0,541, lo que implica moderado PSA. A partir de los MLE para una especificación del modelo NB2, la densidad global es e7.0549 ≈ 1158.5, con un parámetro de dispersión de 0,6707 (es decir, una variación considerable extra-Poisson). Históricamente, la población se ha concentrado en las tierras bajas costeras de la isla, la zona preferida ubicaciones de los colonos españoles iniciales debido a su conveniencia y accesibilidad, proporcionando una justificación para incluir la elevación media como covariable en una regresión NB2. Ante todo debido a la forma en que se expanden las áreas urbanas (por ejemplo, el establecimiento de infraestructura en franja rural-urbana), esta variable debe contener PSA. En este caso, ejecutando NB2 paso a paso La estimación de regresión da como resultado la selección de cinco vectores propios de PSA y ningún vector propio de NSA de el conjunto candidato extraído del SWM ajustado para el cual MCj/MCmax ≥ 0,25, con subíndice j que denota el jésimo valor propio. La Tabla 2 resume los resultados de MLE. revela que SA representa aproximadamente el 60% de la sobredispersión en estos datos, y el único La covariable junto con el FSE construido representa más de las tres cuartas partes de la variación espacial en la densidad de población urbana en la isla principal de Puerto Rico. Con a final α^{-1} < 0,2, una reducción drástica desde casi 0,7, p = 2 parece adecuado para el NB especificación del modelo.

Machine Translated by Google

subíndice j que denota el jésimo valor propio. La Tabla 2 resume los resultados de MLE. revela

variación inicial en la densidad de población urbana en la isla principal de Puerto Rico. Con una especificación de α^{-1} modelo final.

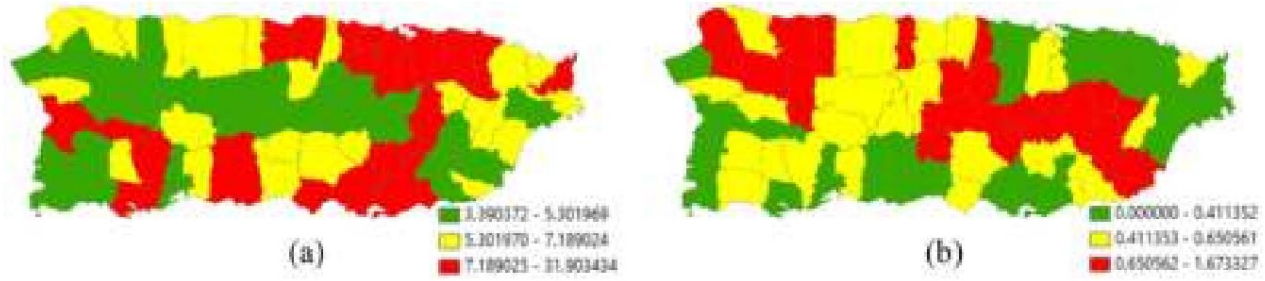


Figura 4. Densidad de población de resolución geográfica de municipio 2010 (es decir, personas por milla cuadrada) a través de la isla principal de Puerto Rico; verde, amarillo y rojo, respectivamente, denotan niveles relativamente bajos, valores de densidad intermedios y altos. Izquierda (a): población urbana (la variación extra-Poisson es 41.183 >> 1). Derecha (b): población rural (la variación extra-Poisson es 1427 >> 1).

Tabla 2. Resumen de los resultados de la estimación del modelo.

Modelo	Densidad de población urbana		Densidad de población rural			
Especificación	NB2		NB2		NB3	
	α^{-1}	Pseudo–R2	α^{-1}	Pseudo–R2	α^{-1}	Pseudo-R2
Sólo interceptar	0.6707	0	1.8760	0	1.8285	0
covariable única	0.4848	0,18	1.3964	0,48	1.1847	0,48
Covariable + FSE	0.2021	0,79	1.0399	0,52	1.0065	0,65

NOTA: pseudo—R^2 (análogo al de Efron) denota el R2 de una regresión lineal bivariada de lo observado en el Densidad de población prevista por NBp (p. ej., [29–32]).

4.2. Una ilustración empírica de la densidad de población rural de Puerto Rico en 2010

La Figura 4b muestra la distribución geográfica de la densidad de población rural en 2010 en la isla principal de Puerto Rico. El patrón del mapa que caracteriza esta distribución geográfica. tiene un MC de 0,496 y un GR de 0,511, lo que implica un PSA moderado. De los MLE para un NB2 especificación del modelo, la densidad global es e4.1298 ≈ 62.2, con un parámetro de dispersión de 1.8760. Los suelos rurales y urbanos son mutuamente excluyentes, lo que justifica la inclusión del densidad de suelo rural como covariable. Porque el suelo rural tiende a concentrarse fuera del urbano. áreas, lo que sugiere que contiene PSA, y principalmente porque la expansión del suelo urbano significa la contracción del suelo rural, particularmente a lo largo de la franja rural-urbana, con esta competencia que sugiere que contiene NSA, esta variable debe contener una mezcla de PSA-NSA. En este caso, la estimación gradual de NB2 da como resultado la selección de cinco vectores propios, cuatro con PSA y uno con NSA (es decir, una mezcla). La Tabla 2 resume los resultados de MLE. Él revela que SA representa aproximadamente el 26% de la sobredispersión en estos datos, y el Una sola covariable junto con el ESF construido representa más de la mitad de la variación en la densidad de población rural en la isla principal de Puerto Rico. Con un α^{-1} final > 1, p = 2 puede parecer inadecuado para la especificación del modelo NB. Estableciendo p = 3—con estimación ejecutado a través de SAS PROC NLMIXED (ver Apéndice B), en el que la densidad de probabilidad NB3 La función de probabilidad logarítmica se escribió expresando el NB modificado como una distribución general. [ver https://stats.oarc.ucla.edu/sas/faq/how-can-i-compute-negative-binomial-models-with-random-intercepts-and-slopes-using-nlmixed/ (consultado por última vez el 20 de marzo de 2024)], altera α^{-1}, que permanece mayor que uno, mientras aumenta los valores pseudo-R2 cuando hay un FSE presente; estos últimos valores disminuyen cuando p = 4, lo que implica que p = 3 es el valor adecuado.

4.3. Una guarida empírica de la población rural de Puerto Rico en 2010

Las dos distribuciones geográficas de densidad de población tienen promedios a nivel insular marcadamente diferentes y variaciones fuera de Poisson. Aunque tienen un índice SA muy similar valores, sus ESF indican que tienen patrones de mapas latentes muy diferentes (por ejemplo, uno se caracteriza por PSA, mientras que el otro se caracteriza por una mezcla de PSA y

Machine Translated by Google

Agencia Nacional de Seguridad). Una población clasificada como urbana no es rural, y viceversa, lo que implica una relación negativa entre estas dos distribuciones geográficas. La correlación bivariada de Poisson RV entre ellos es –0,62 (ver [33]). Finalmente, la densidad de población urbana parece estar bien descrita por un modelo NB2, mientras que la densidad de población rural puede beneficiarse de ser descrita por un modelo NB3.

5. Comentarios finales e implicaciones

Este artículo contribuye a la literatura académica de diversas maneras. Primero, aclara/corrige varios detalles técnicos inexactos en el artículo anterior de Greene. En segundo lugar, presenta nuevos resultados que incluyen el mgf para el modelo NBp, dos nuevos teoremas y aplicaciones empíricas espaciales/econométricas originales. En tercer lugar, ejemplifica propiedades de muestras grandes y pequeñas seleccionadas en modelos NBp. Finalmente, una vez más, demuestra el poder del MESF.

Una implicación destacada es que los analistas cuantitativos, y especialmente los estadísticos espaciales aplicados, que estudian variables de Poisson y/o NB necesitan ampliar su caja de herramientas para incluir especificaciones NBp para las cuales p > 2. Otra es la necesidad de una apreciación y adopción más amplias del MESF. Metodología de tipo para datos correlacionados, incluidos conjuntos de datos dependientes tradicionales, series temporales y proyectos de redes sociales contemporáneos. Una tercera repercusión de este artículo es el reconocimiento de que la variable aleatoria NBp necesita una página de Wikipedia paralela a la existente para NB (https://en.wikipedia.org/wiki/Negative_binomial_distribution; consultado por última vez el 15 de mayo de 2024).

Quizás la limitación más grave de NBp es la falta de MLE de forma cerrada. Aunque este inconveniente de las variables no es nuevo en las estadísticas, constituye un impedimento para el éxito de la utilización de una variable. Este es ciertamente un tema que debería motivar investigaciones posteriores sobre la PNB. Otra de sus limitaciones es su novedad, lo que significa que aún no se han determinado las fortalezas y debilidades de su empleo frente a vehículos recreativos alternativos .

Financiamiento: Esta investigación no recibió financiamiento externo.

Declaración de la Junta de Revisión Institucional: No aplicable.

Declaración de Consentimiento Informado: No aplicable.

Declaración de disponibilidad de datos: Los datos empleados en este documento están disponibles en la Oficina del Censo de los Estados Unidos (es decir, los datos de Puerto Rico) o se pueden obtener (con errores de muestreo) ejecutando un generador de números pseudoaleatorios adecuado (este documento utilizó implementaciones de SAS). .

Agradecimientos: El autor es profesor titular Ashbel Smith de ciencias de la información geoespacial.
enencias.

Conflictos de intereses: El autor declara no tener conflictos de intereses.

Apéndice A. Estimación de los parámetros iid NBp

MLE

$$L(\alpha, \mu|p) = \prod_{yo=1}^{norte} \text{Å yi ãαμ2−p 1 − ãÅ} \frac{1}{1 + \mu p-1/\alpha} \text{ãαμ2−pÅ 1 −} \frac{1}{1 + \mu p-1/\alpha}^{\text{ãyi}}$$

$$LN\ L(\alpha, \mu|p) = \sum_{yo=1}^{norte} LN\hat{\imath}\ \Gamma\ \text{Ä yi + αμ2−p äó −} \sum_{yo=1}^{norte} LN\ \Gamma\ yi + 1 - nLN\hat{\imath}\ \Gamma\ \text{Ä αμ2−p äó+}$$

$$\text{2−p /α ä LNÅ 1 Ä nμ−} \overline{1 + \mu p-1/\alpha} \quad \sum_{i=1}^{norte} yiLN\text{Ç μ+pμp−1/α å ã +}$$

Tanto ∂LN L(α, μ|p) /∂μ como ∂LN L(α, μ|p) /∂α producen la función poligamma (p. ej., ver [34]), que involucra la derivada del logaritmo de la gamma función que impide

Machine Translated by Google

formular algebraicamente soluciones en forma cerrada. Además, la relación entre α, μ y p determina si existe o no un MLE.

MME (usando el estimador de varianza insesgado)

$$\mu = \frac{\sum_{i=1}^{n} y_i}{n}$$

$$\mu(1 + \mu/\alpha) = \frac{\sum_{i=1}^{n} y_i^2 - n\bar{y}^2}{n-1} \qquad \alpha = \frac{n\bar{y}^2}{\sum_{i=1}^{n} y_i^2/(n-1) - \sum_{i=1}^{n} y_i/n}$$

Apéndice B. Código SAS PROC NLMIXED

El siguiente código SAS engaña su procedimiento para estimar una especificación BN3 al no incluir un término de efectos aleatorios. El siguiente conjunto de datos de entrada (PASO 3) contiene los vectores propios de SWM ajustados necesarios, las variables de respuesta (es decir, recuentos de población) y compensación (es decir, área logarítmica), y la covariable (es decir, el logaritmo del porcentaje de área de tierra rural; para evitar datos indefinidos). términos para tierra rural cero, cero fue reemplazado por uno, lo cual es trivial porque la medida más pequeña distinta de cero es 9703): PROC NLMIXED DATA = STEP3; PARÁMETROS B0 = 0 B1 = 0 B2 = 0 B3 = 0 B4 = 0 B5 = 0 B6 = 0 A = 1,4; P = 3; *XB = B0 + COMPENSACIÓN; *XB = B0 + B1*PCTRL +

COMPENSACIÓN; XB = B0 + B1*PCTRL + B2*COL10

+ B3*COL12 + B4*COL15 + B5*COL20 +B6*COL44 + COMPENSACIÓN; MU = EXP(XB); M = 1/A;

LL =

LGAMMA(Y + M) − LGAMMA(Y

+ 1) - LGAMMA(M)

+Y*LOG(A*MU**(P - 1)) - (Y + M)*LOG(1 + A*MU**(P - 1)); MODELO Y ~

GENERAL(LL); PREDECIR MU SALIDA

= MU; CORRER;

XB diferencia entre una media constante, una regresión bivariada con la covariable del porcentaje del área rural y una regresión MESF que incluye esta covariable anterior. P = 3 permite la estimación de NB3; establecer P en 2 estima NB2, y establecerlo en 1 estima NB1 (es decir, un Poisson). Finalmente, el archivo SAS temporal WORK.MU almacena los recuentos previstos para el posprocesamiento. Las opciones NB1 y NB2 admiten comprobaciones de salida comparativas con otros módulos de software.

Referencias

1. Cameron, C.; Trivedi, P. Análisis de regresión de datos de recuento; Cambridge University Press: Nueva York, NY, Estados Unidos, 1998.
2. Barreto-Souza, W.; Ombao, H. El proceso binomial negativo: un modelo manejable con inferencia compuesta basada en probabilidad. Escanear. J. estadística. 2021, 49, 568–592. [Referencia cruzada]
3. Greene, W. Formas funcionales del modelo binomial negativo para datos de recuento. Economía. Letón. 2008, 99, 585–590. [Referencia cruzada] 4. Di, Y.; Schäfer, DW; Cumbie, JS; Chang, JH El modelo binomial negativo de NBP para evaluar la expresión genética diferencial de RNA-Seq. Estadística. Aplica. Gineta. Mol. Biol. 2011, 10, 1–28. [Referencia cruzada]
5. Zhang, Y.; Vosotros, Z.; Lord, D. Estimación del parámetro de dispersión de la distribución binomial negativa para el análisis de datos de accidentes: método de máxima verosimilitud con arranque. Transp. Res. Rec. J. Transp. Res. Junta 2019, 2019, 15-21. [Referencia cruzada]
6. Pearson, K. Método de Momentos y Método de Máxima Verosimilitud. Biometrika 1936, 28, 34–59. [Referencia cruzada]
7. Hendry, D.; Nielsen, B. Modelización econométrica: un enfoque de probabilidad; Princeton University Press: Princeton, Nueva Jersey, EE. UU., 2007.
8. Cámaras, RL; Acero, Dirección General; Wang, S.; Welsh, A. Estimación de máxima verosimilitud para encuestas por muestreo; Prensa CRC: Boca Ratón, FL, Estados Unidos,
2012. 9. Rossi, R. Estadística matemática: una introducción a la inferencia basada en la probabilidad; Wiley: Nueva York, NY, EE. UU., 2018.
10. Ward, M.; Ahlquist, J. Máxima verosimilitud para las ciencias sociales: estrategias de análisis; Cambridge University Press: Nueva York, Nueva York, Estados Unidos, 2018.

Machine Translated by Google

11. Hoeffding, W. El poder de las pruebas para muestras grandes basadas en permutaciones de observaciones. Ana. Matemáticas. Estadística. 1952, 23, 169–192. [Referencia cruzada]
12. Bagui, S.; Mehra, K. Sobre la convergencia de la distribución binomial negativa. Soy. J. Matemáticas. Estadística. 2019, 9, 44–50.
13. Bowman, K.; Shenton, L. Estimador: Método de momentos, Enciclopedia de Ciencias Estadísticas, vol. 3 (Prueba de normalidad de la estimación ocular de D'Agostino), 2ª ed.; Kotz, S., Read, C., Balakrishnan, N., Vidakovic, B., Johnson, N., Eds.; Wiley: Nueva York, Nueva York, Estados Unidos, 2006; págs. 2092-2098.
14. van de Geer, S. Un nuevo enfoque para la estimación de mínimos cuadrados, con aplicaciones. Ana. Estadística. 1987, 15, 587–602. [Referencia cruzada] 15. van de Geer, S. Estimación de mínimos cuadrados. En Enciclopedia de Estadística en Ciencias del Comportamiento; Everitt, B., Howell, D., Eds.; Wiley: Chichester, Reino Unido, 2005; Volumen 2, págs. 1041-1045. Disponible en línea: https://people.math.ethz.ch/~geer/bsa199_o.pdf (consultado el 1 de marzo de 2024).
16. Wolberg, J. Análisis de datos mediante el método de mínimos cuadrados: extracción de la mayor cantidad de información de los experimentos; Springer: Berlín, Alemania, 2005.
17. Delicado, P.; Goria, MN Una pequeña comparación de muestra de los métodos de máxima verosimilitud, momentos y momentos L para Distribución de potencia exponencial asimétrica. Computadora. Estadística. Análisis de datos. 2008, 52, 1661–1673. [Referencia cruzada]
18. Griffith, DA Una familia de observaciones correlacionadas: desde las independientes hasta las fuertemente interrelacionadas. Estadísticas 2020, 3, 166–184. [Referencia cruzada]
19. Griffith, D. Filtrado espacial. En Enciclopedia de Ciencias de la Información Geográfica; Kemp, K., Ed.; SABIO: Thousand Oaks, CA, EE. UU., 2008; págs. 413–415.
20. Griffith, D.; Getis, A. Filtrado espacial. En Enciclopedia de SIG, 2ª ed.; Shekhar, S., Xiong, H., Zhou, X., Eds.; Springer: Cham, Suiza, 2016; págs. 1-14. [Referencia cruzada]
21. Griffith, D.; Chun, Y. Autocorrelación espacial y filtrado espacial. En Manual de ciencia regional, 2ª ed.; Fischer, M., Nijkamp, P., Eds.; Springer-Verlag: Berlín, Alemania, 2022; págs. 1863–1892 (revisado, actualizado a la versión 2014).
22. Moran, P. La interpretación de mapas estadísticos. Estadística JR. Soc. B 1948, 10, 243–251. [Referencia cruzada]
23. Geary, RC La relación de contigüidad y el mapeo estadístico. Incorporación. Estadística. 1954, 5, 115–127+129–146. [Referencia cruzada]
24. Tiefelsdorf, M.; Boots, B. La distribución exacta de I. Environ de Moran . Plan. A 1995, 27, 985–999. [Referencia cruzada]
25. Griffith, DA Autocorrelación espacial y funciones propias de la matriz de pesos geográficos que acompaña a los datos georreferenciados. Poder. Geogr. Poder. 1996, 40, 351–367. [Referencia cruzada]
26. de Jong, P.; Sprenger, C.; van Veen, F. Sobre los valores extremos de I de Moran y c de Geary. Geogr. Anal. 1984, 16, 17–24. [Referencia cruzada]
27. Borcard, D.; Legendre, P. Análisis espacial a toda escala de datos ecológicos mediante coordenadas principales de matrices vecinas. Ecológico. Modelo. 2002, 153, 51–68. [Referencia cruzada]
28. Borcard, D.; Legendre, P.; Avois-Jacquet, C.; Tuomisto, H. Diseccionando la estructura espacial de datos ecológicos a múltiples escalas. Ecología 2004, 85, 1826–1832. [Referencia cruzada]
29. Cameron, AC; Windmeijer, FAG Una medida R cuadrado de bondad de ajuste para algunos modelos de regresión no lineal comunes. J. Economía. 1997, 77, 329–342. [Referencia cruzada]
30. Christensen, R. Teoría general de la predicción y el papel de R2, manuscrito no publicado; Departamento de Matemáticas y Estadística, Universidad de Nuevo México: Albuquerque, NM, EE.UU., 2007; Disponible en línea: https://www.math.unm.edu/~fletcher/JPG/rsq.pdf (consultado el 1 de marzo de 2024).
31. Hoetker, G. El uso de modelos logit y probit en la investigación de gestión estratégica: cuestiones críticas. Estrategia. Gestionar. J. 2007, 28, 331–343. [Referencia cruzada]
32. Cakmakyapan, S.; Demirhan, H. Un pseudocoeficiente de determinación basado en Monte Carlo para modelos lineales generalizados con resultado binario. J. Aplica. Estadística. 2016, 44, 2458–2482. [Referencia cruzada]
33. Kawamura, K. La estructura de la distribución bivariada de Poisson. Matemáticas Kodai. J. 1973, 25, 246–256. [Referencia cruzada]
34. Mortici, C. Estimaciones muy precisas de las funciones poligamma. Asíntota. Anal. 2010, 68, 125-134. [Referencia cruzada]

Descargo de responsabilidad/Nota del editor: Las declaraciones, opiniones y datos contenidos en todas las publicaciones son únicamente de los autores y contribuyentes individuales y no de MDPI ni de los editores. MDPI y/o los editores renuncian a toda responsabilidad por cualquier daño a personas o propiedad que resulte de cualquier idea, método, instrucción o producto mencionado en el contenido.

Machine Translated by Google

AppliedMath

MDPI

Artículo

Relación entre la función de Langevin inversa y la función r0-r1-Lambert W

Roy Howard

Escuela de Ingeniería Eléctrica, Computación y Ciencias Matemáticas, Universidad Curtin, Perth, WA 6845, Australia; r.howard@curtin.edu.au

Resumen: Se define la relación entre la función inversa de Langevin y la función r0 -r1 -Lambert W propuesta. La relación derivada conduce a nuevas aproximaciones para la función de Langevin inversa con límites de error relativo más bajos que las aproximaciones publicadas comparables. Se detallan aproximaciones de alta precisión , basadas en las aproximaciones de raíz de Schröder del primer tipo. Se detallan varias aplicaciones.

Palabras clave: función de Langevin inversa; función de Lambert W; Aproximación de Schröder

1. Introducción

La función de Langevin surge en diversos contextos, incluido el modelo clásico de paramagnetismo [1] y el modelo ideal de cadena libremente unida, por ejemplo, [2]. Se define, por ejemplo, [3], según:

$$y = L(x) = 0, \quad \coth(x) - \frac{1}{x}, \quad x \quad (0, \infty) \quad x = 0 \quad L(-x) = -L(x) \tag{1}$$

Si bien la función de Langevin está bien definida, su forma analítica inversa, $^{-1}$, no tiene un denotada L, y los importantes esfuerzos de investigación han llevado a muchas aproximaciones con los artículos, por ejemplo, y en orden cronológico, de Itskov (2012) [4], Nguessong (2014) [5], Darabi (2015) [6], Jedynak (2015) [7], Kröger (2015) [8], Marchi (2015) [9], Rickaby (2015) [10], Petrosyan (2017) [11], Jedynak (2018) [12], Marchi (2019) [13] y Howard (2020) [14], que proporcionan aproximaciones útiles. En la Figura 1 se muestran gráficas de la función de Langevin y la función de Langevin inversa.

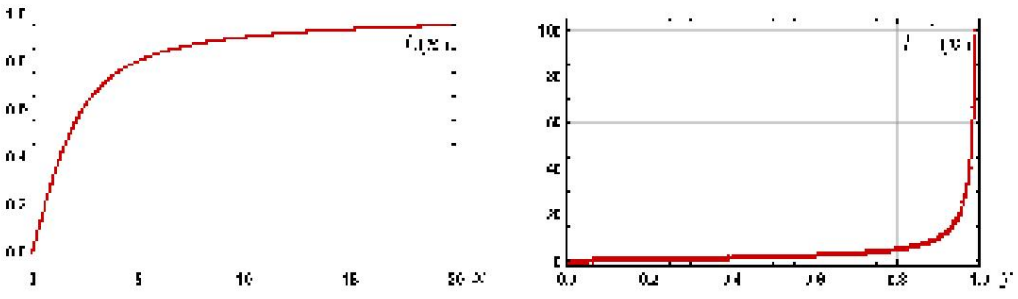


Figura 1. Gráficas de las funciones de Langevin y de Langevin inversa.

Mező y Keady [15] detallan la relación entre la función inversa de Langevin ő y una generalización de la función Lambert W. Si bien esta relación es importante, no facilita directamente la búsqueda de aproximaciones para la función de Langevin inversa, ya que la función W de Lambert generalizada no tiene una forma analítica explícita. La contribución de este artículo es detallar una relación entre la función de Langevin inversa y la función r0-r1-Lambert W, que es una generalización de la función r-Lambert W.

check for updates

Cita: Howard, RM Relación entre la función de Langevin inversa y función r0-r1-Lambert W. Matemáticas aplicadas 2024, 4, 743–762. https://doi.org/10.3390/appliedmath4020040

Editor académico: Valery Karachik

Recibido: 25 de marzo de 2024
Revisado: 18 de mayo de 2024
Aceptado: 21 de mayo de 2024
Publicado: 14 de junio de 2024

Copyright: © 2024 por el autor. Licenciatario MDPI, Basilea, Suiza. Este artículo es un artículo de acceso abierto. distribuido bajo los términos y condiciones de los Creative Commons Licencia de atribución (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

Machine Translated by Google

eso ha sido considerado por Mez˝o [16], Mez˝o y Baricz [17], y Jamilla [18]. La relación derivada conduce a nuevas aproximaciones para la función de Langevin inversa con límites de error relativo potencialmente más bajos que aproximaciones comparables. Se detallan aproximaciones mejoradas, basadas en las aproximaciones de raíz de Schröder del primer tipo, por ejemplo, [19-21] . También se detallan varias aplicaciones de los resultados.

La sección 2 detalla las definiciones que sustentan el artículo. La sección 3 detalla la relación entre la función inversa de Langevin y la función r0-r1-Lambert W propuesta. La sección 4 detalla aproximaciones para la función r0-r1-Lambert W y, por tanto, la función inversa de Langevin. La sección 5 detalla el establecimiento de aproximaciones de mayor precisión mediante el uso de aproximaciones de raíz de Schröder del primer tipo. En la Sección 6 se detallan varias aplicaciones de los resultados. En la Sección 7 se detalla un resumen de los resultados y las conclusiones .

2. Definiciones

Utilizando la definición basada en funciones exponenciales para la cotangente hiperbólica función, la siguiente definición alternativa para la función de Langevin es válida:

$$y = L(x) = \frac{x-1+(1+x)e^{-2x}}{x(1-e-2x)}, \quad x > 0, \; y \quad (0, 1). \tag{2}$$

Luego se deduce que la relación entre xey = L (x) está definida de manera equivalente por cualquiera de las ecuaciones:

$$mi^{-2x} = \frac{1 - x(1 - y)}{1 + x(1 - y)} \tag{3}$$

$$f(x, y) = 1 - (1 - y)x - [1 + (1 + y)x]e^{-2x}, \; f(x, y) = 0, \tag{4}$$

Para x > 0, y (0, 1). La gráfica de la función f se muestra en la Figura 2. Esta función sustenta un vínculo entre la función inversa de Langevin y una generalización de la función Lambert W. Las siguientes definiciones son fundamentales para establecer el vínculo requerido.

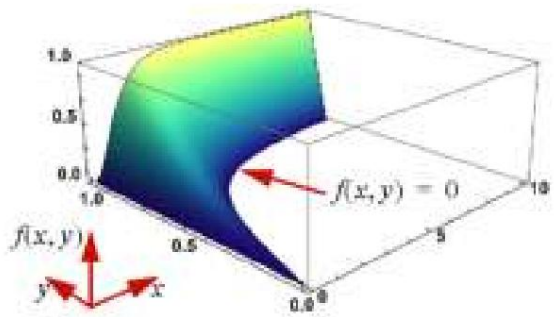


Figura 2. Gráfica de f(x, y).

Definición 1. Función de Lambert W.

La función Lambert W, denotada como W, es la inversa de h, es decir, h −1 (z) = W(z), donde h ( ) = ( ) se define según:

$$z = h(w) = wew . \tag{5}$$

= ( ) = .

Para el caso general, w C, la función Lambert W es multivaluada, por ejemplo, [22]. Para el caso real, w (−∞, ∞), la rama principal de la función Lambert W, denotada W0, está asociada con el intervalo w [−1, ∞) mientras que la rama negativa, denotada W−1, está asociado con el intervalo w (−∞, −1). (−∞, ∞) [−1, ∞) (−∞, −1)

ő

, ,
, ,

Machine Translated by Google

Definición 2. Función W de Lambert generalizada.

La función generalizada de Lambert W, por ejemplo, [15], se denota

$$W. \begin{matrix} \alpha 1, \cdots, \alpha n \\ \cdots, \beta n \end{matrix} \beta 1, \quad z \tag{6}$$

y es la inversa de la función h que se define según

$$z = h(w) = mi \quad {}^{w} \cdot \frac{(w - \alpha 1) \cdot \cdot (w - \alpha n)}{(w - \beta 1) \cdot \cdot (w - \beta n)} \tag{7}$$

Para el caso general w, α1, · · · , αn, β1, · · · , βn C pero a menudo el caso real es de interés.

Definición 3. Función r-Lambert W.

La función r-Lambert W, por ejemplo, [16-18], denotada h −1 o Wr , es la inversa de hr definida según

$$hora(w) = rw + wew . \tag{8}$$

Para el caso general w,r C. Una generalización de esta función es la función r0-r1-Lambert W.

Definición 4. Función r0-r1-Lambert W.

La función r0-r1-Lambert W, denotada h −1 definida $_{r0,r1}$ o Wr0,r1 ,es el inverso de hr0,r1 que es según:

$$hr0,r1 (w) = r0 + r1w + wew . \tag{9}$$

Para el caso general w,r0,r1 C. En este artículo se considera el caso real.

2.1. Enlace: Función de Langevin inversa y función W de Lambert generalizada

La manipulación de (3) (ver [15] Sección 2.2.1) conduce a la relación

$$l^{-1}(y) = -2W \quad \frac{\frac{2}{y+1}}{\frac{2}{\text{años} - 1}} \frac{y - 1}{y + 1} \tag{10}$$

que vincula directamente la función inversa de Langevin con la función general de Lambert W.
Si bien este vínculo es importante, no facilita directamente la búsqueda de aproximaciones para la función inversa de Langevin, ya que se desconoce una aproximación analítica explícita para la función Lambert W y se están realizando investigaciones para desarrollar aproximaciones, por ejemplo, [23,24].
La relación entre la función de Langevin inversa y la función r0-r1-Lambert W, que produce nuevas aproximaciones, se detalla a continuación después de la siguiente subsección, que enumera las aproximaciones establecidas para la función de Langevin inversa.

2.2. Aproximaciones publicadas

Las aproximaciones representativas de la función de Langevin inversa incluyen:

$$L^{-1}_{0,1} y \, 25 \, 75 \quad \frac{3y \, 24y \, 22y}{} (y) = \cdot 1 - + 1 - {}^{2} \tag{11}$$

$$L^{-1}_{0,2} (y) = 1 \quad \frac{3y}{- y} \frac{\text{años} \cdot 1}{- 2 \, 5} \quad \frac{2}{} \quad {}^{2} - \frac{4}{\text{años} + 15 \, 15} \quad \frac{2}{} \quad {}^{6} \tag{12}$$

$$L^{-1}_{0,3} y (y) = 3y + sen \, 5 \quad \frac{2}{} \quad \frac{7}{\text{años } 2} \quad y \quad \frac{3}{+ 1 - y} \tag{13}$$

Machine Translated by Google

$$L^{-1}_{0,4} = \frac{2y^3 - y(y)}{1 - y^2} - \frac{años^{10/3}}{2} + 3\,años^5\, y - \frac{76}{100}(y - 1) \quad (14)$$

$$L^{-1}_{0,5}(y) = \frac{3y + a2y^2 + a3y^3 + a4y^4}{1 - y + b2(y - 2) + b3(y^2 - y^3) + b4(y^3 - y^4)}, \text{ [4/4] aprox.} \quad (15)$$

a2 = −6,98408968, a3 = 5,69026957, a4 = −1,35415696, b2 = −1,33411915, b3 = 0,03919556, b4 = 0,64694651,

que, respectivamente, han sido publicados en [5,8,11,13,14]. Los límites de error relativos (consulte la definición a continuación) que están asociados con estas aproximaciones en el intervalo [0, 1), respectivamente, son: $9{,}69 \times 10^{-3}$, $2{,}75 \times 10^{-3}$, $7{,}22 \times 10^{-4}$ y $1{,}80 \times 10^{-4}$. $1{,}79 \times 10^{-3}$

2.2.1. Error relativo vinculado

Para una función arbitraria f , definida en el intervalo [α, β], una función aproximada fA tiene un error relativo, en un punto x1, definido según re(x1) = 1 − fA(x1)/ f(x1). El límite de error relativo para la función de aproximación, en el intervalo [α, β], se define según:

$$reB = \max\{|re(x1)| : x1 \quad [\alpha, \beta]\} \quad (16)$$

En este artículo se ha utilizado Mathematica para facilitar el análisis y obtener resultados numéricos. En general, los resultados de error relativo asociados con las aproximaciones se han obtenido muestreando uniformemente el intervalo [0, 1] con 1000 puntos.

3. Relación entre la función Langevin inversa y r0-r1-Lambert W

Con la función f definida según (4), se da el caso de que f(x, y) = 0 cuando $x = L^{-1}$ (y). Entonces se cumple el siguiente resultado:

Teorema 1. Relación: función r0-r1-Lambert W y función de Langevin inversa

Con hr0,r1 yf definidos, respectivamente, por (9) y (4), el resultado general

$$hr0(y),r1(y)(w) = r0(y) + r1(y)w + wew = k\, f(y) f\left(\frac{-1}{1+y} - \frac{w}{2}, y, y\right) = -1, \quad (17)$$

es válido para el caso de x, y R cuando

$$-2/(1+y)\; r0(y) = 2, \; r1(y) = 1+y\; \frac{(1+y)\,4e}{(1+y)}, \; \frac{(1-y)e^{-2/}}{(1+y)\,kf(y)} = \frac{2e^{-2/}}{1+y}, y = -1. \quad (18)$$

Luego se deduce que

$$hr0(y),r1(y)(w) = \frac{4e^{-2/(1+y)}}{(1+año)^2} + \frac{w(1-y)e^{-2/(1+y)}}{1+y} + wew, \; y = -1, \quad (19)$$

y la función de Langevin inversa se puede definir según

$$I^{-1}(y) = \frac{-1}{1+y} - \frac{h^{-1}_{(0)}\, r0(y),r1(y)}{2}, \quad 0 \le y < 1. \quad (20)$$

Prueba del teorema 1. La prueba se detalla en el Apéndice A. □

3.1. Determinación de la raíz

de h De acuerdo con (20), la función de Langevin inversa se puede definir una vez que se conoce la raíz apropiada de hr0(y),r1 (y) . Se puede demostrar fácilmente que

$$hr0(y),r1(y)\left(\frac{-2}{1+año}\right) = 0, \; y = -1, \quad (21)$$

Machine Translated by Google

implica h (0) = ( ) = r0(y) < 1 < r1(y), lo cual es así. Sin embargo, este resultado está asociado con una rama de h −1 que especificada por (20). La rama de h −1 que se requiere se puede ver fácilmente considerando la naturaleza de h para el caso de 0< <1 0 < y < 1.

3.1.1. Graficos

Para el caso de 0 ≤ y < 1, se da el caso de que 1/e < r0(y) ≤ 4/e ( ), para el caso de 0 ≤ y < 1 se muestran en la Figura 3. La gráfica de hr0(y),r1 (y) (w) se muestra en la Figura 4 para dos valores establecidos de y para ilustrar la naturaleza de la variación de hr0(y),r1 (y) (w) para valores establecidos de y. La variación, con y, la Figura 5. está a la izquierda de −2/(1 + y), se muestra en

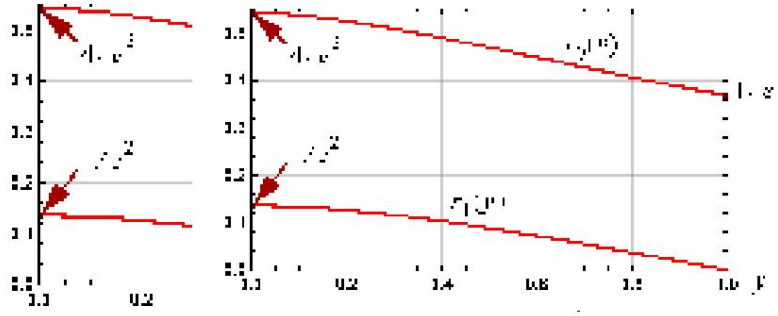


Figura 3. Gráficas de r0(y) y r1(y) para el caso de y (0,1). 1 ( )

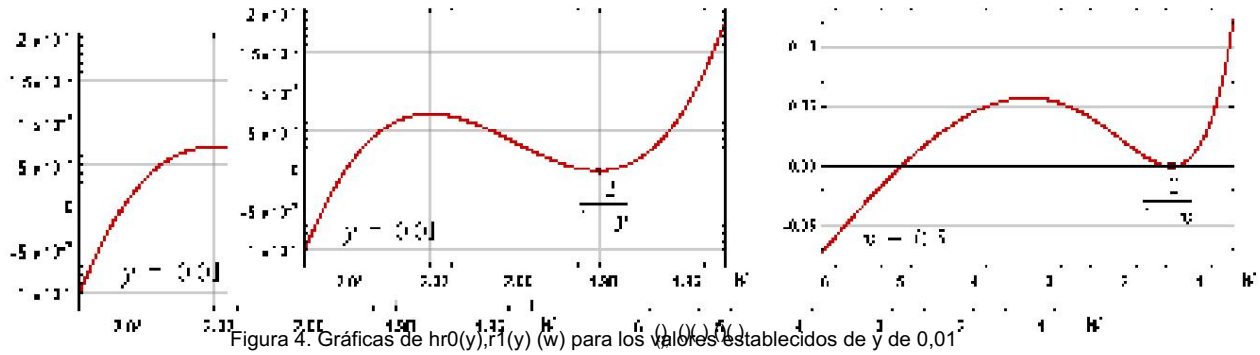
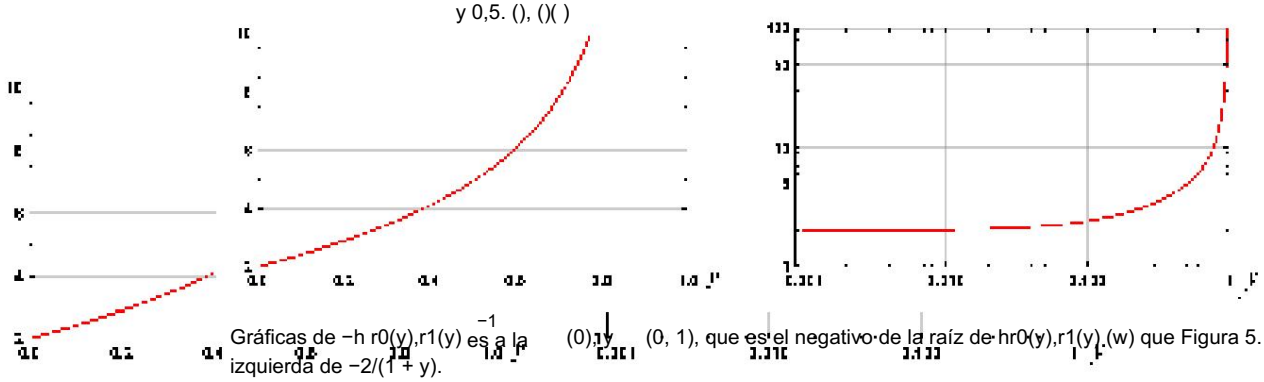


Figura 4. Gráficas de hr0(y),r1(y) (w) para los valores establecidos de y de 0,01 y 0,5. (), ()( )



Gráficas de −h r0(y),r1(y) es a la (0,1), que es el negativo de la raíz de hr0(y),r1(y) (w) que Figura 5. izquierda de −2/(1 + y).

3.1.2. Rama requerida para la raíz de

h La variación de hr0(y),r1 (y) (w), con w, se ilustra en la Figura 4 junto con la primera raíz negativa de w = −2/(1 + y) como se define por (21). La raíz requerida está a la izquierda del máximo local de hr0(y),r1 (y) (w) que se encuentra a la izquierda de la primera raíz negativa. Para establecer dónde ocurre este máximo local, considere la derivada de hr0(y),r1 (y) (w):

$$(1) \quad (w) = r1(y) + (1 + w)e \; h\, r0(y),r1(y) \qquad w \,. \tag{22}$$

Machine Translated by Google

La raíz de esta función, denotada como wm, ocurre cuando (1 + wm)e wm = − ( ) 0< ( ) ≤ después de un cambio 1 / =1+ = −er1(y). Como 0 < r1(y) ≤ 1/e 0 ≤ 2 para ( ) caso que = 0, ( ) = 1/ −1 / , −1/e, y = de wm(y) = we 0<1<+ w0<1<+ w0, donde w1e <1 0 ≤ y < 1 (ver Figura 3), se da el 1/ − ( ), −er1(y), 0 < y < 1, 0 < r1 (y) < 1/e

$$= \quad = \quad 2 \quad 2 \tag{23}$$

con soluciones, consistentes con la naturaleza de la función Lambert W, por ejemplo, [22], definida de acuerdo con

$$= \quad \begin{matrix} -1, & 0\ 0 \quad y = 0 = \\ W0[-er1(y)] \text{ o } W-1[-er1(y)], [-( )] [-( )], \text{ donde } W0 & < y < 1 \end{matrix} \tag{24}$$

es la solución de la rama principal y W−1 0 < 1

es la rama negativa. Es la rama negativa la que proporciona la solución requerida y, por tanto:

$$= \quad \begin{matrix} -2, & y = 0 = 0 \\ -1 + W-1[-er1(y)], & 0 < y < 1 \end{matrix} \quad -1 + [-( )], 0 < \tag{25}$$

Por tanto, el objetivo es establecer la raíz de hr0(y),r1 (y) ( ) ( ) para el caso de w ≤ wm. La naturaleza (z), se ilustra en la Figura hr0(y),r1 (y) (w), y su inversa w = h ≤ caso de w ≤ wm y un ( ) ( ) ( ) r0(y),r16 para el valor fijo de y.

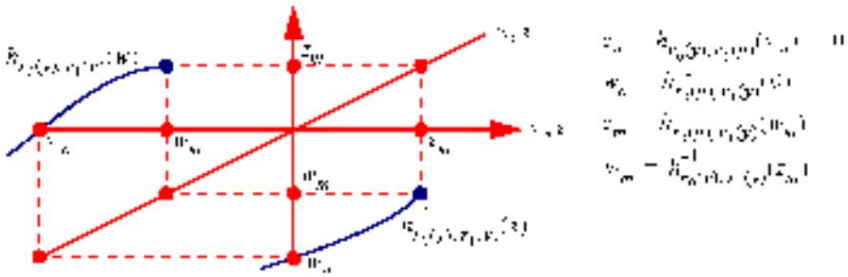


Figura 6. Ilustración de las funciones z = hr0(y),r1(y) (w) y w = h (), ()( ) w ≤ wm. ≤ = (), () (0)(y),r1(y)(z) para y fijo y

4. Aproximaciones para la función de Langevin inversa

con el teorema 1, una aproximación a h r0(y),r1 (y) ()() (0) 0≤ <1 a una aproximación a la inversa de (0), para 0 ≤ y < 1, conduce de acuerdo Langevin función. Para establecer aproximaciones para −1 (0), primero, considere las siguientes dos aproximaciones de dominio restringido que h r0(y),r1 (y) surgen del análisis detallado en el Apéndice B:

$$(0) \approx -2 - 4y, \ 0 < y \quad 1, \ h \ r0(y),r1 \ (y) \ y = 1 - \Delta, \ 0 < \Delta \quad 1.$$

$$h \ (0) \approx r0(y),r1 \ (y) \quad \frac{-2}{1-y}, \tag{26}$$

Estas aproximaciones son consistentes con la aproximación de primer orden:

$$(0) = h \ 1,r0(y),r1 \ (y) \quad \frac{-2}{} = 2y = 1-y \quad \frac{-2}{} 1 + y - y \ 1-y \quad 2, \quad 0 \le y < 1, \tag{27}$$

La coherencia surge, primero, al notar −2/(1 − y) ≈ −2(1 + y), 0 < y 1 y, por lo tanto, la aproximación para y → 0 es −2 − 4y según sea necesario. En segundo lugar, para el caso de y = 1 − Δ, 0 < Δ 1, la aproximación tiende a −2/(1 − y) según sea necesario. En la Figura 7 se muestra una medida del error en esta aproximación y el límite de error relativo asociado con el intervalo (0, 1) es 4,45 × 10−2 .

Machine Translated by Google

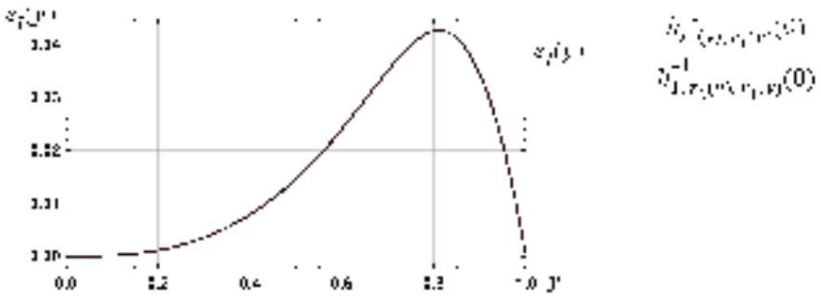


Figura 7. Gráfica, en el intervalo y (0, 1), de una medida del error relativo, el , en la aproximación h (0) tal como se define en (27), a h (0).

La aproximación definida por (27) y la medida de error relativo el mostrada en la Figura 7 son consistentes con

$$h^{-1}(0) = h^{-1}(0)\cdot[1 - el(y)],\ r0(y),r1(y)\ 1,r0(y),r1(y) \tag{28}$$

y esta relación proporciona una base útil para establecer mejores aproximaciones. Las aproximaciones mejoradas incluyen:

$$\cdot[1+ -],\quad 0 \le < 1,\ \cdot 1 + c21\sin(\pi y) + c22\sin \tag{29}$$

$$= \cdot 1 + y - y \tag{30}$$

$$\cdot 1 + y - y + y 3,r0(y),r1(y)\ 1 - y \cdot 1 + \tag{31}$$

$$h^{-1}(0) = 5,r0(y),r1(y) \qquad (1 - y)(y - yo)e^{-3(y-yo)} \tag{32}$$

$$+ \alpha 3y + \alpha 4y + \alpha 5y + \alpha 7y + \alpha 8y = \tag{33}$$

En la Figura 8 se muestran gráficos de la magnitud de los errores relativos en estas aproximaciones. Los respectivos límites de error relativo asociados con estas aproximaciones y 4,58 × para el intervalo (0, 1), son: 4,66 × 10−3 2,33 × 10− 4 y 4,58 × 10−5 .

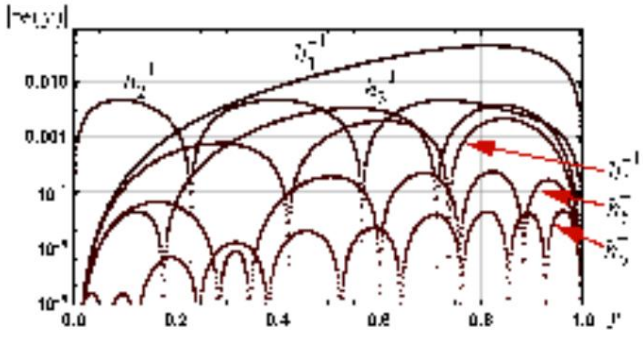


Figura 8. Gráficas de la magnitud del error relativo en aproximaciones a h (0) tal como se define r0(y),r1(y) por (27) y (29) a (33).

Machine Translated by Google

4.1. Aproximaciones de Langevin asociadas

Dada la relación $L^{-1}(y) = \frac{-1}{1+y} - h^{-1}_{(0)\,r0\,(y),r1\,(y)\,2}$, como se establece en el Teorema 1, las aproximaciones anteriores para $h^{-1}(0)$, como se especifica en (27) y (28) a (33), conducen, respectivamente, a r0(y),r1 (y) las siguientes aproximaciones para la inversa Función de Langevin:

$$L^{-1}_{1}\ y(y) = \cdot 1 - 2\,3\,1 - y^{\frac{3\ años}{}}{}^{2} \tag{34}$$

$$L^{-1}_{2} + 1-y\ 1-y\ d21 = \frac{=}{1000}, \frac{d21\sin(\pi y)+d22\sin(\pi y\ 2)\ 2y\ (y)}{d22 = \frac{301}{}\ \frac{-92}{1000}.} \tag{35}$$

$$L^{-1}_{3}(y) = /1\ \frac{3y}{-y}\ años \cdot 1\ \frac{2}{-2/3} - \frac{y^{5/2}}{6} + \frac{y^{9/2}}{6} \tag{36}$$

$$( ) = 1--1\ \cdot 1- \quad - \quad + \quad - \quad (1-)(-) \cdot ()$$

$$L_{4}\ y\ 3\ \frac{3y\ y\ (y) = \cdot 1 - 2\ 1 - - \frac{y^{3}}{6} - \frac{y^{6}}{6} + \frac{y^{2}}{3} + \frac{y^{5}}{6}\ años - \frac{y^{7}}{6}}{} \tag{37}$$

$$L^{-1}_{5}(y) = 1\ \frac{3y}{-y\ ( ) = 1-}\ \cdot 1 - 2\ \frac{2\ y\ 1 - \frac{5/2\ 9/2\ yy + 3}{6\ 6\ 6\ (1-)}}{3}\ + \frac{5\ c51y}{3} - \frac{(1-y)(y-yo)\ 6\ 2\ 3\cdot 1 - + - 4}{7\ yy\ 6\ 6} \cdot mi^{-3(y-yo)}{}^{2} \tag{38}$$

$$L^{-1}_{6}(y) = 1\ \frac{3y}{-y}\ \cdot \frac{1-\ \frac{2}{y\ 3} - \frac{3}{y}\ - \frac{5}{años +}\ \frac{6}{años + 6} - \frac{2}{+}\ 3 + 2\ 2}{3\ \frac{\{1+\alpha 0y(1-y\ 2)\ }{}\ \cdot 1 - y^{2}\ 6\ 2 + 3\ + 2\ 3\ - \frac{+3\ 4}{años\ 2}\ años + 2\ + 5\ - \frac{6}{y}\ +}{1 + \alpha 2y\ 3 + \alpha 3y\ 4 + \alpha 4y\ 5 + \alpha 5y\ 7 + \alpha 7y\ 8 + \alpha 8y} \tag{39}$$

En la Figura 9 se muestran gráficos de la magnitud de los errores relativos en estas aproximaciones. $(0,1)$ $4{,}94 \times 10$ $1{,}83 \times 10$ $4{,}37 \times 10$ $2{,}44 \times 10$ $2{,}58 \times 10$ $4{,}85 \times 10$ Los límites de error relativos asociados con estas aproximaciones y para el intervalo (0, 1), respectivamente, son: $4{,}94 \times 10^{-2}$ $1{,}83 \times 10^{-2}$ $2{,}44 \times 10^{-3}$ $2{,}58 \times 10^{-4}$ y $4{,}85 \times 10^{-5}$.

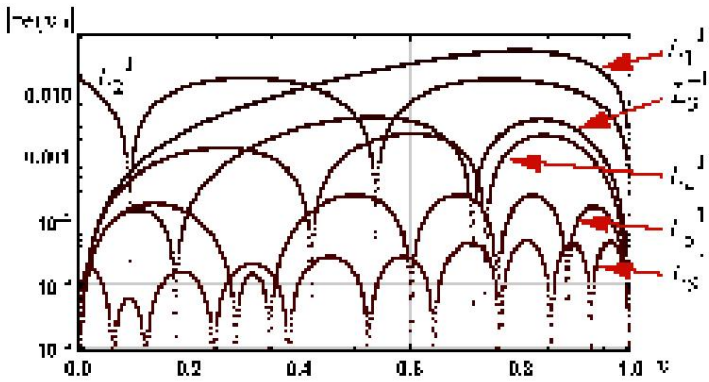


Figura 9. Gráficas de la magnitud del error relativo en aproximaciones de L (34) a (39). $^{-1}$ po(y)() según lo definido

Notas

$$( ) = \cdot 1 - 3\ ^{3}_{1--1}$$

Los coeficientes utilizados en las aproximaciones h y L son diferentes y 2,r0(y),r1 (y) han sido elegidos para minimizar sus respectivos límites de error relativos durante el intervalo [0, 1).

La cuarta aproximación, dada por (37), tiene un límite de error relativo de $2{,}44 \times 10^{-3}$ $\sin(\ ) + \sin(\ )\ (\ ) =$ en el intervalo [0, propuesto por Kröger [8] y expuesto en 1- (12). La aproximación propuesta incluye un polinomio 301 de orden siete (en lugar de ^{2} 1). y es similar en forma al $^{+}$ y logra un límite de error relativo ligeramente inferior. 1000 Las aproximaciones quinta (38) y sexta (39) logran límites de error relativo más bajos, respectivamente , que las aproximaciones publicadas proporcionadas por Nguessong [5] y Marchi [13] y que se establecen en (14) y (15). $^{-}$ $^{1-92}$ ^{1000} .

$$( ) = 1-\ \cdot 1-\ \frac{}{3}\ \frac{-}{6}\ \frac{/}{+}\ 6$$

$$( ) = 1-\ \cdot 1-\ \frac{3}{3}\ \frac{-}{6}\ \frac{-}{6}\ \frac{+}{3}\ \frac{+}{6}\ \frac{-}{6}$$

Machine Translated by Google

Las aproximaciones propuestas complementan la lista de aproximaciones publicadas para la función de Langevin inversa y con límites de error relativos consistentes o inferiores a aproximaciones publicadas comparables. Aún quedan por proponer aproximaciones adicionales.

5. Aproximaciones de orden superior basadas en

Schröder Las siguientes aproximaciones generales para una función inversa surgen de la aproximación de raíz de Schröder de primer tipo, por ejemplo, [25] (Ecuación (21)), [19] (Ecuación (17)), [20] (Sección 3) y [21]:

Teorema 2. Aproximaciones basadas en Schröder para una función inversa Considere una función real f que es diferenciable hasta al menos orden n y también es estrictamente monótona en un intervalo alrededor de un punto x0 que incluye la raíz xo de g(x) = f (x) − yo que está cerca de x0. Una serie de Taylor de enésimo orden para g basada en el punto (y0, x0), y0 = f(x0) − yo, produce la aproximación de Schröder de enésimo orden para $^{-1}$ (yo), denotado xn, y xo = f definida según

$$xn = x0 + \sum_{k=1}^{n} \frac{(-1)^k [f(x0) - yo]^k}{k!} \cdot D(k) f^{-1}[f(x0)], \quad n \in \{1, 2, \cdots\}, \tag{40}$$

donde D(k) es el operador que denota la diferenciación de orden k. Una forma explícita es

$$\begin{aligned} xn = x0 &- f(1)\frac{f(x0)-yo}{(x0)} - \frac{[f(x0)-yo]^2 f(2)(x0)}{2[f(1)(x0)]^3} - \frac{[f(x0)-yo]^3 f(3)(x0)}{6[f(1)(x0)]^4} \cdot -1 + f(1)\frac{3[f(2)(x0)]^2}{(x0)f(3)(x0)} - \\ &\frac{[f(x0)-yo]^4 f(4)(x0)}{24[f(1)(x0)]^5} 1 - f(1)\frac{10 f(2)(x0)f(3)(x0)\cdot}{(x0)f(4)(x0)\,15 f(2)} + \frac{15[f(2)(x0)]^2}{[f(1)(x0)]^2 f(4)(x0)} - \\ &\frac{[f(x0)-yo]^5 f(5)(x0)}{120[f(1)(x0)]^6} \cdot \frac{(1)(x0)f\,(x0)\,f(4)(x0) - 1 + f}{(5)(x0)\,105[f(2)(x0)]^2} + \frac{10[f(3)(x0)]}{f(1)(x0)f(5)(x0)^4} - \\ &\frac{f(3)(x0)}{[f(1)(x0)]^2 f(5)(x0)} + \frac{105[f(2)(x0)]^3}{[f(1)(x0)]^3 f(5)(x0)} \\ &\frac{[f(x0)-yo]^6 f(6)(x0)}{720[f(1)(x0)]^7} \cdot - f\frac{21 f(2)(x0)f(5)(x0)\,1}{(1)(x0)f(6)(x0)\,280 f(2)^2} - \frac{35 f(3)(x0)f(4)(x0)\,f}{(1)(x0)f(6)(x0)} + \frac{210[f(2)(x0)]^2 f(4)(x0)}{[f(1)(x0)]^2 f(6)(x0)^5} + \\ &\frac{(x0)[f(3)(x0)]\,[f(1)(x0)]^2}{f(6)(x0)} - \frac{1260[f(2)(x0)]^3 f(3)(x0)}{[f(1)(x0)]^3 f(6)(x0)} + \frac{945[f(2)(x0)]^4}{[f(1)(x0)]^4 f(6)(x0)} - \\ &\cdots + \frac{(-1)^{n}[f(x0)-yo]^{n}}{n!} \cdot D(n) f^{-1}[f(x0)] \end{aligned} \tag{41}$$

Prueba del teorema 2. Este resultado surge de una aproximación en serie de Taylor para $^{-1}$ en el el punto g (y0, x0) y el teorema de la función inversa. Véase, por ejemplo, [21]. □

5.1. Aproximaciones de primer, segundo y tercer orden para h−1 (0)

Considere hr0(y),r1 (y) (w) = r0(y) + r1(y)w + wew, con r0 y r1 como se definen en (18), y (0) donde el argumento de y es una aproximación inicial para h (0) se denota como h $^{-1}$ r0(y),r1 (y) $^{-1}$ 0 implícita y se supone que está en el intervalo [0, 1). Luego se deduce que el primero, el segundo y (0), aproximaciones basadas en Schröder de tercer orden para h $^{-1}$ r0(y),r1 (y) (0), denotado, respectivamente, h $^{-1}$ 1 h $^{-1}$ 2 (0), yh $^{-1}$ 3 (0) son:

$$h_{1}^{-1}(0) = 2\,1 - y \frac{-4 + (1+y)\,2w\,2\exp w + 1+y}{2 + (1+años)\,2(1+w)\exp w + 1+y\,w=h} \quad {}^{-1}_{0}(0) \tag{42}$$

Machine Translated by Google

$$h_2^{-1}(0) = \frac{-4+(1+y)\,2w\,2\exp w+\sqrt{1+y}\,(1+w)^{\frac{2}{2}}}{1-y\,2+(1+y)\exp w+\sqrt{\frac{2}{1+año}}} - \frac{4+(1-y\,2)w+(1+y)\,2w\exp w+\sqrt{1+y}^{\frac{2}{2}}}{2(1+y)\,1-y+(1+y)(1+w)\exp w+\sqrt{\frac{2}{1+año}}}\,(2+w)\exp w+\sqrt{1+y}^{\frac{2}{3}}\,h_{ancho=alto_0}^{-1}(0) \quad (43)$$

$$h_3^{-1}(0) = h_2^{-1}(0) - \frac{4+1-y^{2}\,w+(1+y)\,2w\exp w+\sqrt{1+y}^{\frac{2}{2}}\exp w+\sqrt{1+y}^{\frac{2}{3}}\cdot}{6(1+año)^{2}\,1-y+(1+y)(1+w)\exp w+\sqrt{1+y}^{\frac{2}{2}}} \quad (44)$$

$$\frac{-(1-y)(3+w)+(1+y)(9+8w+2w\,2)\exp w+\sqrt{\frac{2}{1+y\,5}}}{h_{ancho=alto_0}^{-1}(0)}$$

Las pruebas de estos resultados se detallan en el Apéndice C.

Resultados

Los límites de error relativos asociados con las aproximaciones anteriores, y basados en $h_0^{-1}(0)$ como se especifica en (27) y (29) a (33), se detallan en la Tabla 1. Los casos específicos se detallan a continuación.

5.2. Aproximaciones de primer orden: casos explícitos

La forma general de aproximación de primer orden especificada por (42) conduce a las aproximaciones detalladas a continuación que son válidas para $y \in [0, 1)$. Las aproximaciones asociadas para la función de Langevin inversa se pueden definir de acuerdo con el Teorema 1.

5.2.1. Caso 1

Con la aproximación inicial $h_{1,r0(y),r1(y)}^{-1}(0)$ especificado por (27), se deduce que

$$h_{hora\,1}^{-1}(0) = \frac{-4(1-y)^{2}+4\,1+2y-y\exp 3\,2^{\frac{-6\,años}{1-año}}\,años\cdot 1-\frac{2}{3}}{(1+y)(1-y)^{3}-(1-y\,2)\,1+4y+y^{2}-2\,años^{3}\exp.\frac{-6\,años}{1-año}\,años\cdot 1-\frac{2}{3}} \quad (45)$$

que tiene un límite de error relativo de 2,69 × 10−4 . El límite de error relativo en la aproximación asociada para la función de Langevin inversa es 3,36 × 10−4 .

5.2.2. Caso 2

Con la aproximación $h_1^{-1}(0)$ definida por (30), se deduce que $h_{3,r0(y),r1(y)}^{-1}$

$$h_1^{-1}(0) = \frac{-4(1-y)\,2+(1+y)\,2\,[-2-2y+2y\,2+y\,7/2-y\,9/2]\,[1+3y-2y\,2-y\,7/2+y^{2}\exp[pE(y)]}{(1-año)^{3}\,(1+y)-(1-y)(1+y)^{2}\,9/2]\exp[pE(y)]}$$

$$pE(y) = \frac{-6y_2}{1-año}\cdot 1-\frac{2}{y\,3}-\frac{y^{5/2}}{6}\,años+6^{\frac{9/2}{}} \quad (46)$$

que tiene un límite de error relativo de 8,52 × 10−6 . El límite de error relativo en la aproximación asociada para la función de Langevin inversa es 1,17 × 10−5 .

5.2.3. Caso 3

Con la aproximación $h_{4,r0(y),r1(y)}^{-1}(0)$ definido por (31), se deduce que

$$h_{hora\,1}^{-1}(0) = \frac{-4(1-y)\,2+4(1+y)\,2\,1+y-y^{+y\,2\,\frac{4\,7\,2-y\,6-y\,2}{}}\exp[pE(y)]}{(1-año)^{3}\,(1+y)-(1-y)(1+y)^{2}\,[1+3y-2y\,2-y\,4+2y\,6-y\,7\,]\exp[pE(y)]}$$

$$pE(y) = \frac{-6y_2}{1-año}\cdot 1-\frac{2}{y\,3}-\frac{3}{y\,6}-\frac{4}{y\,6}\,años+3^{\frac{5}{}}\,años+6^{\frac{6}{}}-\frac{7}{y\,6} \quad (47)$$

que tiene un límite de error relativo de 1,92 × 10−6 . El límite de error relativo en la aproximación asociada para la función de Langevin inversa es 2,66 × 10−6 .

Machine Translated by Google

5.2.4. Caso 4

Con la aproximación h 6,r0(y),r1 (y) $^{-1}$ (0) definido por (33), se deduce que

$$h_1^{-1}(0) = \frac{2+4(1+y)\,2\,2-y\,6-y\,1+y-y\,2+y\,2 \quad \frac{4\,7-4(1-y)}{} \quad \frac{2}{} \quad 2p\,(y)\exp[pE(y)]}{(1-y)^{3}(1+y)+(1-y)(1+y)^{2}\,1-y-2\,1+y-y^{+y\,2\,\frac{4\,7\,2-y\,6-y\,2}{}}\,p(y)\exp[pE(y)]} \tag{48}$$

dónde

$$p(y) = 1 + \alpha 0y^{2}(1-y)^{2}\,1 + \alpha 2y^{2} + \alpha 3y^{3} + \alpha 4y^{4} + \alpha 5y^{5} + \alpha 7y^{7} + \alpha 8y^{8} \tag{49}$$

$$pE(y) = \frac{-6\ años}{1-y^{2}} \cdot \frac{y\,1-3\,\frac{2}{}-\frac{3}{y}-\frac{4}{y\,6}\,\frac{5}{años+3}\,\frac{6}{años+6}-\frac{7}{y\,6}+}{\frac{\alpha 0y(1-y\,2)}{3}\cdot 1-y^{\,2}\,\frac{3}{+\ años}-\frac{3y}{2}\,\frac{4}{\frac{3}{años+2}}\,\frac{5}{}-\frac{6}{y}\cdot} \tag{50}$$
$$1+\alpha 2y^{\,2}+\alpha 3y^{\,3}+\alpha 4y^{\,4}+\alpha 5y^{\,5}+\alpha 7y^{\,27}+\alpha 8y^{\,8}$$

Esta aproximación tiene un límite de error relativo de 4,07 × 10−10. El límite de error relativo en la aproximación asociada para la función de Langevin inversa es 6,27 × 10−10 .

5.3. Aproximaciones de segundo orden

Con la aproximación h 4,r0(y),r1 (y) $^{-1}$ (0) definido por (31), se deduce, de (43), que:

$$h_{\text{hora}\,2}^{-1}(0) = h_1^{-1}(0) + \frac{-(1-y)^{3}\,1-y^{\,2\,3-y}-\frac{4}{y}\,\frac{5}{años+2\,5}\,y+\frac{6}{2\,6}+}{(1+años)^{3}\,1-y^{\,2}+y^{\,3}-\frac{2\,4}{\frac{3\,años}{2}\,años+\frac{3}{2}}-\frac{y}{2}\exp[pE(y)]}^{2}\cdot \tag{51}$$
$$\frac{{}^{8\,años}\,1-años^{\frac{2}{}-\frac{y}{4}\,\frac{3}{años+2}\,\frac{5}{}-\frac{y}{4}\,\frac{6}{}}\exp[pE(y)]}{(1+y)[(1-y)\,2-(1+y)[1+3y-2y\,2-y\,4+2y\,6-y\,7\,]\exp[pE(y)]]^{3}}$$

donde h 1 $^{-1}$ (0) y pE(y) están definidos por (47). El límite de error relativo para esta aproximación es 9,12 × 10−9 . El límite de error relativo en la aproximación asociada para la función de Langevin inversa es 1,16 × 10−8 .

5.4. Aproximación directa de Schröder de primer orden a la función de Langevin inversa

Considere la aproximación de Schröder de primer orden a la función de Langevin inversa, basado en una aproximación inicial denotada L $_{0,}^{-1}$ y definido por

$$L_1^{-1}(y) = x - \frac{L(x)-y}{L\,(1)\,(x)}\,{}_{=L\,0}{}^{-1x}(y) \tag{52}$$

La naturaleza de L.$_1^{-1}$ Depende de la forma asumida para la función de Langevin. Para el caso de L definida según L(x) = coth(x) − 1/x, se deduce que:

$$L_1^{-1}(y) = \frac{3x - x\sinh(x)\quad 2\sinh(x\,) - x\cosh(x) + xy\sinh(x)\quad - \sinh(x)}{2x\qquad\qquad 2}\,{}_{=L\,0}{}^{-1x}(y) \tag{53}$$

Para el caso de L definido según (3), se sigue que:

$$L_1^{-1} = 1 - \frac{x[2-x+xy]-2x\,2+2x\,(y)^{\,2}+xy\,e^{\,-2x}+x[2+x+xy]e^{\,-4x}}{2[1+2x\,2\,]e-2x+e-4x}\,{}_{=L\,0}{}^{-1x}(y) \tag{54}$$

Machine Translated by Google

Comparación

Las aproximaciones definidas por (53) y (54) se pueden contrastar con la aproximación la función de Langevin inversa de acuerdo con la relación establecida en el Teorema 1, es decir,

$$L^{-1}(y) \approx \frac{-1}{1+y} - \frac{h_1^{-1}(0)}{2} \tag{55}$$

donde $h_1^{-1}(0)$ es la aproximación de primer orden definida según (42). para comparar el aproximaciones definidas por (54) y (55), considere las aproximaciones iniciales

$$L_0^{-1}(y) = \frac{3y}{1-y} \cdot \frac{1}{2} - \frac{2}{15} + \frac{2}{15} \tag{56}$$

$$h_0^{-1}(0) = \frac{-2}{1+y} \cdot \frac{y}{2} - \frac{y^4}{2} + \frac{y^6}{2} - \frac{y^7}{2} \tag{57}$$

El primero fue propuesto por Kröger [8] (indicado anteriormente en (12)) y tiene un límite de error relativo de 2,75 × 10−3 . La segunda aproximación se especifica en (31) y tiene un límite de error relativo de 2,13 × 10−3 . El uso de (56) en (54) produce la aproximación

$$L_1^{-1}(y) = -pE(y) \frac{(1-y)^2 \left(1 + \frac{y}{2} - \frac{2}{y} - \frac{3}{5}\right) + \left(\frac{3}{5} + \frac{5}{y} - \frac{8}{y}\right)}{(1+y)^2 \left(1 - \frac{y}{2} - \frac{2}{2} - \frac{3y}{5} - \frac{4}{10} - \frac{5}{8}\right) - 2pE(y) - \left(\frac{2}{2} + \frac{2}{17y} - \frac{83y}{10} + \frac{y}{50} + \frac{26}{37y} - \frac{159}{10} - \frac{y}{10} - \frac{5}{4y} + \frac{14}{25}\right) - pe(y)}$$

$$(1-y^2)^2 \left(1 + 2pE(y) - 2\left(1 + 16y^2 - \frac{67}{5} + \frac{28y}{25} + \frac{6}{25} + \frac{4}{25} + \frac{12y}{25} + \frac{14}{25}\right) - 2pE(y)\right)$$

$$pE(y) = \frac{-6}{1-y^2} \cdot \frac{2y}{3} \cdot \left(1 - \frac{2}{5} - \frac{3}{6} - \frac{4}{6} + \frac{2}{3} - \frac{6}{5} - \frac{7}{6}\right) \tag{58}$$

con un límite de error relativo de 7,34 × 10−6 . En contraste, el uso de (57) en (42), y luego (55), conduce a la aproximación especificada en (47), es decir,

$$h_1^{-1}(0) = \frac{-4(1-y)^2 + 4(1+y)^2 \left(1 + y - \frac{y^2}{2}\right) - \frac{7y^2}{2} + 6y - 2}{(1-y)^3 (1+y) - (1-y)(1+y)^2 [1 + 3y - 2y^2 - y^4 + 2y^5 - 6y^6 + 7y^6] \exp[pE(y)]} \exp[pE(y)] \tag{59}$$

$$pE(y) = \frac{-6y}{1-y^2} \cdot \left(1 - \frac{2}{3y} - \frac{3}{6y} - \frac{4}{6y} + \frac{5}{3y} - \frac{6}{6y} - \frac{7}{6y}\right)$$

que tiene un límite de error relativo de 1,92 × 10−6 . La aproximación asociada para $L^{-1}(y)$ es

$$L_1^{-1}(y) = \frac{-1}{1+y} - \frac{h_1^{-1}(0)}{2} \tag{60}$$

y tiene un límite de error relativo de 2,66 × 10−6 .

Claramente, (59) y (60) tienen una forma más simple que (58) y producen una aproximación para (y) con un límite de error relativo inferior. Esta es, en parte, la razón fundamental para el uso de la Enlace −1 L entre la función de Langevin inversa y la función r0-r1-Lambert W.

6. Aplicaciones

6.1. Aproximaciones de

límites Las aproximaciones detalladas en (26) conducen a las siguientes aproximaciones para el Función de Langevin inversa después de usar la relación definida en el Teorema 1:

$$\begin{aligned} L^{-1}(y) &\approx \frac{3y}{1+y} + \frac{2y}{1-y}, && 0 < y < 1, \\ L^{-1}(y) &\approx \frac{2}{\Delta}, && y = 1 - \Delta,\ 0 < \Delta < 1. \end{aligned} \tag{61}$$

Machine Translated by Google

9 ( ) ≈ 3 + 297 + 1539 + =1−Δ 0 < Δ 1 Estas ( ) ≈
5 175 875

(1−) para aproximaciones se pueden contrastar con aproximaciones establecidas. Primero, una aproximación en serie de Taylor⁄ la función de Langevin inversa, por ejemplo, [4], es

$$L^{-1}(y) \approx 3y + \frac{9y^3}{5} + \frac{297a^5}{175} + \frac{1539a^7}{875} + \cdots \quad (62)$$

( ) ≈ 1+ 3 , 0 < 1,

En segundo lugar, una aproximación de primer orden para el caso de y = 1 − Δ, 0 < Δ 1, es L −1 (y) ≈ 1/(1 − y) (ver Apéndice B de [8]). Las gráficas de los errores relativos asociados con estas = 1 − Δ, 0 < Δ 1. aproximaciones se muestran en la Figura 10. De estas gráficas se desprende claramente que las aproximaciones definidas en (61) son inferiores a las aproximaciones establecidas para argumentos cercanos a cero o cercano a uno. Sin embargo, sirven como base útil para establecer aproximaciones con límites de error relativos bajos en el intervalo (0, 1) para h −1 −1 y, por tanto, L (0,1) , como se resume en la Tabla 2.

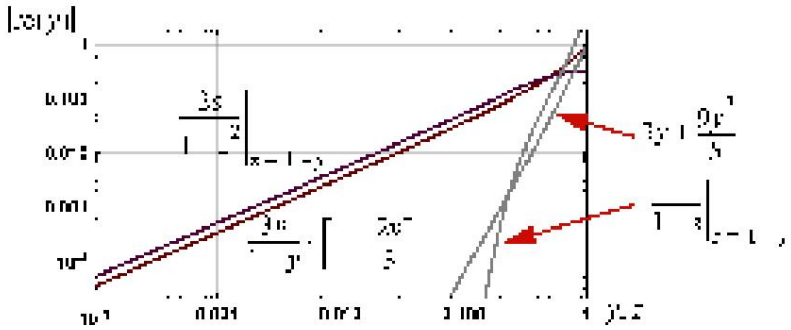


Figura 10. Gráfica de los errores relativos en aproximaciones a la función de Langevin inversa para los casos de 0 < y 1 e y = 1 − Δ, 0 < Δ 1. 0< 1 = 1 − Δ, 0 < Δ 1.

6.2. Resultados Integrales

Como $\int_0^{\cdot} L(\lambda)d\lambda = \ln[\sinh(x)] - \ln(x)$, el resultado integral general

( ) = ln[sinh( )] − ln( )

$$\int_{y_1}^{y} f^{-1}(\lambda)d\lambda = yf^{-1}(y) - y_1 f^{-1}(y_1) - \int_{f^{-1}(y_1)}^{f^{-1}(y)} f(\lambda)d\lambda \quad (63)$$

asumiendo f^{-1} y f están bien definidos, produce

$$\int_0^y L^{-1}(\lambda)d\lambda = yL^{-1}(y) + \ln L(L^{-1}(y)) - \ln\sinh(L^{-1}(y)), \quad y \in (0,1). \quad (64)$$

Las aproximaciones propuestas se pueden utilizar para establecer aproximaciones para la integral de la función de Langevin inversa. Por ejemplo, el uso de las aproximaciones especificadas por (36) y (37), en (64), produce aproximaciones para la función de Langevin inversa con los límites de error relativos respectivos de 1,40 × 10−5 y 3,91 × 10−6 . La última aproximació

$$\int_0^y L^{-1}(\lambda)d\lambda \approx 1-y\ln\frac{3y^2}{2\text{ años}\cdot 1-\frac{2}{3}-\frac{3}{y6}-\frac{4}{y6}\text{ años}+3\frac{5}{\text{años}+6}-\frac{6}{y6}-\frac{7}{y6}+} \quad (65)$$
$$\frac{}{1-y^2\text{ años}\cdot 1-\frac{2}{3}-\frac{3}{y}-\frac{4}{y}y+3\frac{4}{5}\text{ años}+6\frac{6}{}-\frac{7}{y}-}$$
$$\text{en sinh } 3y\frac{}{1-\text{año}^2\,2y\cdot 1-3\frac{6}{}-\frac{6}{3y6}-\frac{6}{}\text{años}+3\frac{y}{5}6y+6\frac{6}{}-\frac{7}{y6}}, \quad y \in (0,1).$$

6.3. Raíces de α + βw +

wew Consistente con (21) y el Teorema 1, las raíces de hr0(y),r1 (y) (w) = r0(y) + r1(y)w +
0 < y < 1 son y 1+y 1+y especificación 2L −1 wew, (y). Una extensión deseable de estos resultados sería una de las raíces de

$$h_{\alpha,\beta}(w) = \alpha + \beta w + we^w \quad (66)$$

Machine Translated by Google

para valores arbitrarios de α y β. Dadas las definiciones de r0 y r1, esto, en general, no es posible. Sin embargo, se pueden establecer raíces para casos específicos. Por ejemplo, las raíces de

$$\frac{(1+k)^2}{e^{1+k}} + \frac{kw}{kd^{+}} + \text{bueno}, \quad 0 < k < 1, \tag{67}$$

son $-1 - k$ y $-1 - k - 2L\ 1+k$, $^{-1}\ \frac{1-k}{}$ y surgen del caso de y = $\frac{1-k}{}$ 0 < k < 1. Como 1+k,

ejemplo, las raíces de

$$\frac{16}{9e^{4/3}} + \frac{w}{3e^{4/3}} + \text{wew} \tag{68}$$

son $-4/3$ y $-4/3 - 2L$ $^{-1}(1/2) \approx -4.927$, y surgen del caso de y = 1/2, k = 1/3.

Respuesta al impulso del sistema

lineal Tenga en cuenta que $\alpha + \beta$ consistente $^{-vo} =$ vse-re siempre que $\alpha + \beta v - ve - v = 0$ y ^{wo} wo, es decir, las raíces de $\alpha + \beta w + wew = 0$ son las negativas de las raíces de $\alpha + \beta v - ve - v = 0$.

De acuerdo con (67), considere un sistema lineal con una respuesta al impulso definida por

$$h(t) = 1 - k1t - k2te^{-t}, \quad t \geq 0, \tag{69}$$

donde $k1 = 2$, $k2 = (1+k)\ (1+k)$ según $\frac{e^{1+k}}{2}$, transferencia de dicho sistema está definida $(1 +) 0 < k < 1$. La función de

$$H(s) = \frac{1}{s} - \frac{k1}{2\,s} - \frac{}{k2\ (1 + s)^2} \tag{70}$$

y es consistente con un sistema que comprende un integrador, una cascada de dos integradores y un sistema de segundo orden. Las raíces de la respuesta al impulso son $1 + k$ y $1 + k + 2L\ 1\ 10/,1\ 4/,1\ 2/$ La $^{-1}\ \frac{1-k}{1+k}$. respuesta al impulso se muestra en la Figura 11 para los casos de k {1/10, 1/4, 1/2}.

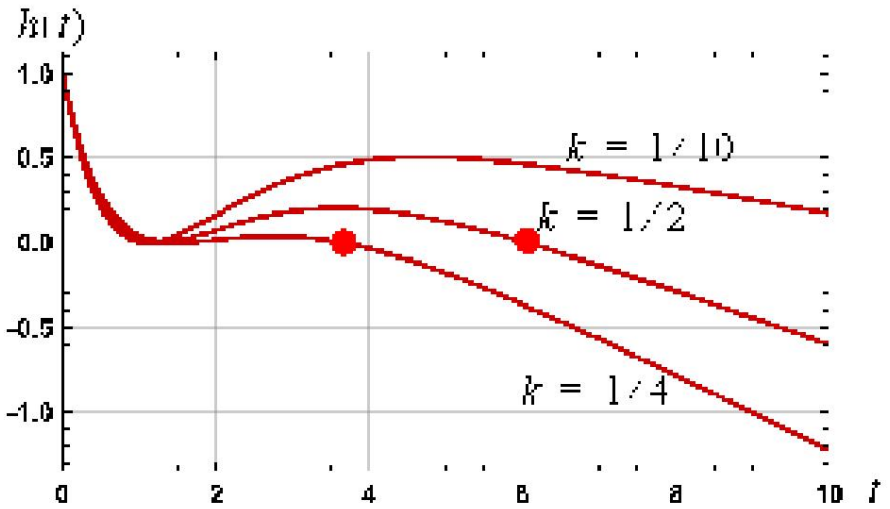


Figura 11. Gráfica de la respuesta al impulso de la función de transferencia definida por (69) para valores de k {1/10, 1/4, 1/2}.

7. Resumen de resultados y conclusiones

7.1. Resumen de resultados

Las aproximaciones a h (0), junto con sus límites de error relativos asociados, r0(y),r1 (y) se detallan en la Tabla 1. Las aproximaciones asociadas para la función de Langevin inversa $^{-1}(y) = 1 - - ,$ son detallado en la Tabla 2. $1\ 10/,1\ 4/,1\ 2/$

$(),() (0)$

Machine Translated by Google

Tabla 1. Resumen de aproximaciones para $h^{-1}_{(0)}(y)$, tal como se define en (27) y (29) a (33), y sus límites de error relativos $r0(y),r1(y)$ para el intervalo [0, 1).

Aproximación	Definición	Error relativo vinculado	Error relativo Con destino: 1º Orden—(42)	Error relativo Con destino: 2do Orden—(43)	Error relativo Con destino: 3.º Orden—(44)
$h^{-1}_{(0)\,h\,1,r0\,(y),r1\,(y)}$	$\frac{-2}{1-y}\cdot 1 + y - y^{2}$	$4,45 \times 10{-}2$	$2,69 \times 10{-}4$	$1,87 \times 10{-}5$	$1,32 \times 10{-}6$
$h^{-1}_{2,r0\,(y),r1\,(y)}(0)$	$\frac{-2}{1-año}\cdot 1 + c21\text{sen}\,(\pi y) + c22\text{sen}\,\pi y^{2}$	$4,66 \times 10{-}3$	$2,20 \times 10{-}4$	$2,42 \times 10{-}5$	$3,68 \times 10{-}6$
$h^{-1}_{(0)\,3,r0\,(y),r1\,(y)}$	$\frac{-2}{}\cdot 1 + y - y\,1{-}y^{2} - \frac{y^{7/2}}{2}\frac{9/2}{años+2}$	$3,71 \times 10{-}3$	$8,52 \times 10{-}6$	$7,58 \times 10{-}8$	$7,74 \times 10{-}10$
$h^{-1}_{(0)\,4,r0\,(y),r1\,(y)}$	$\frac{-2}{}\cdot 1 + y - y\,1{-}y^{2} - \frac{y^{4}}{2} + y^{6} - \frac{y^{7}}{2}$	$2,13 \times 10{-}3$	$1,92 \times 10{-}6$	$9,12 \times 10{-}9$	$4,59 \times 10{-}11$
$h^{-1}_{(0)\,5,r0\,(y),r1\,(y)}$	$\frac{-2}{}\cdot 1 + y - y\,1{-}y^{2} - \frac{y^{7/2}}{2}\frac{9/2}{años+2} + c51y^{5}(1-y)(y-yo)e^{-3(y-yo)^{2}}$	$2,33 \times 10{-}4$	$3,15 \times 10{-}8$	$1,68 \times 10{-}11$	$1,04 \times 10{-}14$
$h^{-1}_{(0)\,h\,6,r0\,(y),r1\,(y)}$	$\frac{-2}{1-y}\cdot 1 + y - y^{2} - \frac{y^{4}}{2} + y^{6} - \frac{y^{7}}{2}\cdot \frac{1+\alpha 0y^{2}}{1+\alpha 2y^{2}+\alpha 3y^{5}+\alpha 4y^{7}+\alpha 5y^{8}+\alpha 7y+\alpha 8y}(1-y)^{2}$	$4,58 \times 10{-}5$	$4,07\times 10{-}10$	$2,43 \times 10{-}14$	$1,92 \times 10{-}18$

Tabla 2. Resumen de aproximaciones, basado en las aproximaciones a $h^{-1}_{(0)}$ detalladas en la Tabla 1, $r0(y),r1(y)$ para la función de Langevin inversa junto con sus límites relativos asociados para el intervalo [0, 1). Los límites de error relativo para las aproximaciones basadas en Schröder de primer, segundo y tercer orden surgen del uso directo de las aproximaciones en la forma de aproximación general definida en el Teorema 2.

Aprox.	Definición	Error relativo vinculado	Error relativo Encuadernado: primer orden	Error relativo vinculado: Segundo Orden	Error relativo vinculado: Tercer orden
$L^{-1}_{1}(y)$	$\frac{2\,años}{1-año^{2}}\,años\cdot 1\,\frac{2}{-3}$	$4,94 \times 10{-}2$	$2,40 \times 10{-}3$	$1,20 \times 10{-}4$	$5,87 \times 10{-}6$
$L^{-1}_{2}(y)$	$\frac{2\,años}{1-año^{2}}\,\frac{d21\text{sin}\,(\pi y)+d22\text{sin}(\pi y+1-y^{2})}{}$	$1,83 \times 10{-}2$	$2,94 \times 10{-}4$	$5,29 \times 10{-}6$	$9,23 \times 10{-}8$
$L^{-1}_{3}(y)$	$\frac{-3y}{}\cdot 1 - 2\,1{-}y^{\frac{y^{2}}{3}} - \frac{5/2}{\frac{años}{6}} + \frac{9/2}{\frac{años}{6}}$	$4,37 \times 10{-}3$	$1,62 \times 10{-}5$	$6,59 \times 10{-}8$	$2,64 \times 10{-}10$
$L^{-1}_{4}(y)$	$\frac{2\,años}{1-año^{2}}\,años\cdot 1\,\frac{2}{-3} - \frac{y^{3}}{6} - \frac{y^{4}}{6}\frac{5}{años+3}\frac{6}{años+6} - \frac{y^{7}}{6}$	$2,44 \times 10{-}3$	$5,32 \times 10{-}6$	$1,24 \times 10{-}8$	$2,84 \times 10{-}11$
$L^{-1}_{5}(y)$	$\frac{3\,años}{1-y^{2}}\,años\cdot 1\,\frac{2}{-3} - \frac{y^{5/2}}{6} + \frac{y^{9/2}}{6} - \frac{c51y\,5(1-y)(y-yo)}{2}\cdot m^{2-3(y-yo)^{2}}\cdot\frac{3}{\frac{años}{2\,1-años}}$	$2,58 \times 10{-}4$	$6,47 \times 10{-}8$	$1,67 \times 10{-}11$	$4,22 \times 10{-}15$
$L^{-1}_{6}(y)$	$y\,1-\frac{2}{3} - \frac{y^{3}}{6} - \frac{y^{4}}{6}\frac{5\,6\,aa+}{3\,6} - \frac{7}{años+6}\cdot\frac{\alpha 0\,y(1-y^{\frac{2}{3}})^{2}}{1+\alpha 2y\,\alpha 5y^{2}+\alpha 3y+\alpha 4y\,4+\alpha 8y}\cdot\frac{1-y^{2}+años^{3}-\frac{4}{a2os+}}{\frac{3y}{2}-\frac{5}{2}\frac{6}{2}}$	$4,85 \times 10{-}5$	$2,32 \times 10{-}9$	$1,14 \times 10{-}13$	$5,48 \times 10{-}18$

7.2. Conclusiones

Se definió la relación entre la función inversa de Langevin y la función r0-r1-Lambert W. La relación derivada conduce a nuevas aproximaciones para la función de Langevin inversa con límites de error relativos potencialmente más bajos que las aproximaciones publicadas comparables.

Se demostró que se pueden definir aproximaciones mejoradas para la función r0-r1-Lambert W, con límites de error relativo significativamente más bajos, utilizando las aproximaciones de raíz de Schröder del primer tipo. Luego se pueden definir aproximaciones para la función de Langevin inversa utilizando la relación establecida entre las dos funciones. Un enfoque alternativo es utilizar directamente una aproximación establecida para la función de Langevin inversa en la forma de aproximación general de Schröder. Se dieron ejemplos para ilustr

Machine Translated by Google

que el primer enfoque probablemente conduzca a una expresión analítica más simple para un límite de error relativo establecido.

Se mostró cómo las aproximaciones definidas para la función de Langevin inversa se pueden utilizar para establecer aproximaciones para la integral de la función de Langevin inversa, con, en general, límites de error relativos bajos.

El vínculo entre la función inversa de Langevin y la función r0-r1-Lambert W se utilizó para establecer las raíces de la respuesta al impulso de un sistema lineal personalizado que comprende un integrador, una cascada de dos integradores y un sistema de segundo orden.

En términos de investigaciones futuras, se observa que se pueden utilizar enfoques iterativos para establecer aproximaciones con límites de error relativos más bajos y se han detallado varios resultados para la función inversa de Langevin, por ejemplo, [14,21]. Berinda [26] y Gdawiec [27], por ejemplo, proporcionan una visión general de los enfoques iterativos que pueden utilizarse para problemas de punto fijo. En general, los enfoques iterativos conducen a aproximaciones analíticas con un alto grado de complejidad y se basan en una aproximación inicial. Para lograr resultados con una complejidad modesta, se requiere un orden de iteración bajo junto con una aproximación inicial que tenga un límite de error relativo bajo y con baja complejidad. Por lo tanto, las nuevas aproximaciones son de interés y justifican una investigación continua.

Financiamiento: Esta investigación no recibió financiamiento externo.

Declaración de la Junta de Revisión Institucional: No aplicable.

Declaración de Consentimiento Informado: No aplicable.

Declaración de disponibilidad de datos: los datos están contenidos en el artículo.

Agradecimientos: El autor se complace en agradecer el apoyo de A. Zoubir, SPG, Technische Universität Darmstadt, Darmstadt, Alemania, quien organizó una visita donde se establecieron los resultados fundamentales que sustentan este artículo.

Conflictos de intereses: El autor declara no tener ningún conflicto de intereses.

Apéndice A. Prueba del teorema 1

Considere la función $f(x, y) = 1 - (1 - y)x - [1 + (1 + y)x]e^{-2x}$ para el caso general de $x, y \in \mathbb{R}$, $y \neq -1$ y las transformaciones

$$z = 1 + (1 + y)x, \; x = \frac{z-1}{1+y}, \tag{A1}$$

$$w = \frac{-2z}{1+y}, \quad z = \frac{-(1+y)w}{2}. \tag{A2}$$

Las definiciones de f2 y f1, según $f2(w, y) = f1(z, y) = f(x, y)$, conducen entonces a

$$f1(z, y) = f\left(\frac{z-1}{1+y}, y\right) = 1 + \frac{(1-y)z}{y} - z\exp\left(\frac{-2z}{1+y}\right)\cdot\exp\left(\frac{2}{1+y}\right) \tag{A3}$$

$$f2(w, y) = f1\left(\frac{-(1+y)w}{2}, y\right) = \frac{2}{1+y}\left(\frac{(1-y)w}{2} + \frac{(1+y)w}{2}\cdot\exp(w)\cdot\exp\left(\frac{2}{1+y}\right)\right) \tag{A4}$$

Luego se deduce que

$$\frac{2e^{-2/(1+y)}}{1+y}\cdot f2(w, y) = \frac{4e^{-2/(1+y)}}{(1+y)^2} + we^{w+1+y}\frac{(1-y)e^{-2/(1+y)}w}{1+y} = r0(y) + r1(y)w + we^{w} = hr0(y),r1(y)(w) \tag{A5}$$

cuando

$$(1+y)\,r0(y) = \frac{4e^{-2/}}{y}, \; r1(y) = (1+y)\left(1 + \frac{(1-y)e^{-2/(1+y)}}{y}\right). \tag{A6}$$

Machine Translated by Google

De (A1) y (A2), se da el caso de que x = f2(w, y) = f(x, $\frac{-1}{1+y} - \frac{w}{2}$ y luego se sigue, de la relación y), que

$$k_{f(y)f2(w,\,y) = k\,f(y)f\left(\frac{-1}{1+años} - \frac{w}{2},\, y = hr0(y),r1(y)(w)\right.} \tag{A7}$$

cuando $k_{(1+y)\,f(y) = \frac{2e^{-2/}}{1+y},}$ que es el primer resultado requerido.

Para establecer el último resultado indicado, considere el caso específico de f(x, y) = 0, 0 < y < 1, que define el camino

$$(x, y) : x = L^{-1}(y),\ y \quad (0, 1)\ . \tag{A8}$$

A lo largo del camino definido por f(x, y) = 0, 0 < y < 1, y con x = $\frac{-1}{1+y} - \frac{}{w2,}$ es el caso que

$$kf(y)f\left(\frac{-1}{1+años} - \frac{w}{2},\, y = r0(y) + r1(y)w + wew = hr0(y),r1(y)(w) = 0\right. \tag{A9}$$

$^{-1}_{(0).\ Por\ lo\ tanto,\ con\ w\ = - 2L\ r0(y),r1(y)\ 1+y\ -1\ h\ (0)\ r0(y),r1(y)\ 2}$ $\frac{-2\ -1}{}$ es decir, w = h (w), se deduce que $L^{-1}(y) = \frac{-1}{1+y} -$ $\frac{}{}$, que es el último resultado indicado. Este resultado también es válido para el caso de y = 0.

Apéndice B. Aproximaciones de dominio restringido para hr0(y),r1 (y) (w)

Apéndice B.1. Caso 1: 0 < y << 1

Para el caso de 0 < y 1, y de acuerdo con el gráfico de la izquierda que se muestra en la Figura 4, la raíz de hr0(y),r1 (y) que es de interés está cerca, pero a la izquierda, de la raíz conocida de − 2/(1 + y) (ver (21)). Considere hr0(y),r1 (y) (w), definido por (19), y el caso de hr0(y),r1 (y) (w) = 0, es decir,

$$\frac{4e^{-2/(1+y)}}{(1+años)^2} + \frac{w(1-y)e^{-2/(1+y)} + wew}{} = 0.\ 1+y \tag{A10}$$

Con la forma para w de w = −2/(1 + y) + ε(y), se deduce que:

$$\frac{4}{(1+años)} \quad \frac{1-y+2}{1+años}\cdot\frac{-2}{y} + \varepsilon(y) + 1 + \frac{-2}{} + \varepsilon(y)\ e = 0,^{\varepsilon(y)}{}_{1+y} \tag{A11}$$

La simplificación y una expansión en serie de Taylor para la función exponencial alrededor del punto cero conducen a:

$$2 + (1-y)\varepsilon(y) + [-2 + (1+y)\varepsilon(y)]\ 1 + \varepsilon(y) + \frac{2\,\varepsilon(y)}{} + \cdots = 0,\ 2 \tag{A12}$$

Las expansiones de primer, segundo y tercer orden de la función exponencial conducen a las aproximaciones:

$$\begin{array}{l} \varepsilon^2{}_{(y)[1+y] = 0,\ primer\ orden\ (1+y)\varepsilon(y)\ (y)\ 1+2} \\ 2\,\varepsilon \quad \frac{}{} = 0,\ segundo\ orden\ (1+3y)\varepsilon(y) \\ 2\,\varepsilon \quad \frac{(1+y)\varepsilon\ (y)\ y + +\ 6\ 6}{} \ \frac{}{^2(y)} = 0,\ tercer\ orden \end{array} \tag{A13}$$

con soluciones:

$$\begin{array}{ll} \varepsilon(y) = 0, & primer\ orden \\ \varepsilon(y) = 0, \varepsilon(y) = & \frac{-2y}{1+y},\ segundo\ orden \\ = 0, \varepsilon(y) = 2(1+y) & \frac{-1-3y \pm \sqrt{1-18y-15y\ \varepsilon(y)^2}}{},\ tercer\ orden \end{array} \tag{A14}$$

Machine Translated by Google

Considere la primera solución distinta de cero para ε(y), que es la solución de segundo orden de ε(y) = −2y/(1 + y). Para el caso de y → 0, esta solución implica w = −2, lo cual es consistente con la raíz −2/(1 + y).

La solución de tercer orden, que es consistente con ε(y) → 0, ya que y → 0, es $-1-3y+\sqrt{1-18y-15y^2}$ ε(y) = 2(1+y) y conduce a la aproximación:

$$\text{hr0(y),r1 (y)} \quad \frac{-2-6y-24y^2}{1+y} \approx 0,\ 0 < y \quad 1. \tag{A15}$$

Por eso:

$$\text{r0(y),r1 (y)}\ 1 + y \quad \frac{-2 - 6y - 24y^2}{-1\ (0) \approx h} \approx -2 - 4y - 20y^2 \tag{A16}$$

después de usar la aproximación $(1 + y)^{-1} \approx 1 - y + y^2$.

Apéndice B.2. Caso 2: y -> 1

Para y ≈ 1, las expansiones en serie de Taylor alrededor del punto uno producen las aproximaciones:

$$\begin{aligned} r0(y) &= \frac{4e^{-2/(1+y)}}{(1+y)^2} \approx \frac{1}{mi} \frac{1-y}{2e}\left(1-(1-y)+\frac{(1-y)^2}{8e}+\cdots\right) \\ (1+y)\,r1(y) &= \frac{(1-y)e^{-2/(1+y)}}{1+y} \approx \frac{1-año}{2e} - \frac{(1-año)^3}{16e} \cdots \end{aligned} \tag{A17}$$

Sustitución de estas aproximaciones en

$$\text{hr0(y),r1 (y) (w)} = r_0(y) + r1(y)w + wew = 0 \tag{A18}$$

rendimientos

$$\frac{1}{mi}\left(1 - y + \frac{(1-y)^2}{8e} + \cdots\right) + \left(\frac{1-y}{2e} - \frac{(1-y)^3}{16e} - \cdots\right) w + wew = 0. \tag{A19}$$

A medida que y → 1 desde abajo, la raíz requerida de (A18) se vuelve cada vez más negativa (esto es evidente al comparar los gráficos de la izquierda y la derecha que se muestran en la Figura 4). Para el caso de |w| 1, y w negativo, el término wew se vuelve cada vez más pequeño, lo que lleva a

$$w = -\frac{\frac{1}{mi}\left(1-(1-y)+\frac{(1-y)^2}{8e}+\frac{3(1-y)^3}{16e}+\cdots\right)}{\frac{1-y}{2e} - \cdots} \approx \frac{-2}{1-y},\ y = 1 - \Delta,\ 0 < \Delta \quad 1, \tag{A20}$$

y el resultado requerido

$$\Delta,\ 0 < \Delta \quad 1.\ \text{r0(y),r1 (y)}^{-1}(0) \approx h\ 1-y,\ y = 1 - \frac{-2}{} \tag{A21}$$

Apéndice C. Prueba de las aproximaciones de Schröder de primer, segundo y tercer orden

Con z = hr0(y),r1 (y) (w) = r h(w) = $r_0(y) + r1(y)w + wew$, se deduce, usando la notación hr0(y),r1 (y) (w), eso

$$\begin{aligned} &h^{(1)}(w) = r1(y) + (1 + w)e^w,\ h^{(3)}(w) = (3 + w)e^w, \\ &h^{(2)}(w) = (2 + w)e^w,\ \cdots \\ &h^{(k)}(w) = (k + w)e^w. \end{aligned} \tag{A22}$$

Luego se deduce, del Teorema 2, que la aproximación de Schröder de primer orden es

$$\text{hora 1}^{-1}(z) = w - h\frac{h(w) - z}{(1)(w)} \quad \text{ancho=alto}_0^{-1}(z) = \frac{-r0(y) + w^{\text{2 nosotros}} + z}{r1(y) + (1 + w)e^w} \quad \text{ancho=alto}_0^{-1}(z) \tag{A23}$$

Machine Translated by Google

asumiendo una función de aproximación inicial hz = 0 es h_0^{-1} para h -1 . La aproximación para el caso de

$$h_1^{-1} = r1(y) \frac{-r0(y) + w\,(0)^{2}\, e^{w}}{+ (1 + w)e^{w}} \quad \text{ancho=alto}_0^{-1}\ (0) \tag{A24}$$

La sustitución de r0 y r1, como se especifica en (18), produce:

$$h_1^{-1}(0) = 2\,1 - y \frac{-4 + (1 + y)\,2w\,2\,2\exp w + \frac{1+y}{2}}{+ (1 + años)\ 2\,(1 + w)\exp w + \frac{1+y}{2}\, w=h} \quad {}_0^{-1}\ (0) \tag{A25}$$

Apéndice C.1. Segundo Orden

La aproximación de Schröder de segundo orden a h, tal como se define en el teorema 2, es

$$\begin{aligned} h_2^{-1}(z) &= h_1^{-1}(z) - \frac{[h(w)-z]^{2}\, h\,(2)\,(w)}{2[h\,(1)\,(w)]^{3}} \quad \text{ancho=alto}_0^{-1}\ (z) \\ &= \frac{-r0(y)+w^{2}\, e^{w+z}}{r1(y)+(1+w)e^{w}} - \frac{[r0(y)+r1(y)w+wew-z]^{2}\,(2+w)e^{w}}{2[r1(y)+(1+w)e^{w}]^{3}} \quad \text{ancho=alto}_0^{-1}\ (z) \end{aligned} \tag{A26}$$

y la aproximación para el caso de z = 0 es

$$h_2^{-1}(0) = \frac{-r0(y)+w^{\text{2 nosotros}}}{r1(y)+(1+w)e^{w}} - \frac{[r0(y)+r1(y)w+wew]^{2}\,(2+w)e^{w}}{2[r1(y)+(1+w)e^{w}]^{3}} \quad \text{ancho=alto}_0^{-1}\ (0) \tag{A27}$$

La sustitución de r0 y r1 produce:

$$h_2^{-1}(0) = \frac{-4+(1+y)\,2w\,2\quad 2\exp w+\frac{1+y}{2}}{1-y\,2+(1+y)^{2}(1+w)\exp w+\frac{2}{1+año}} - \frac{4+(1-y\,2)w+(1+y)\,2w\quad \exp w+\frac{2}{1+y}}{2(1+y)\,1-y+(1+y)(1+w)\exp w+\frac{2}{1+año}} \frac{{}^{2}\,(2+w)\exp w+\frac{1+y}{2}}{3} \quad \text{ancho=alto}_0^{-1}\ (0) \tag{A28}$$

Apéndice C.2. Tercer orden

La aproximación de Schröder de tercer orden a h, tal como se define en el teorema 2, es

$$h_3^{-1}(z) = h_2^{-1}(z) - \frac{[h(w) - z]^{3}\, h\,(3)\,(w)}{6\,h\,(1)\,(s)} \cdot \frac{3\,h\,(2)\,(w)^{2}}{-1 + h\,(1)\,(w)h\,(3)\,(w)} \quad \text{ancho=alto}_0^{-1}\ (z) \tag{A29}$$

y la aproximación para el caso de z = 0 es

$$h_3^{-1}(0) = h_2^{-1} - \frac{[r0(y)+r1(y)w+wew]\,(0)^{2}\,(3+w)e^{w}}{6[r1(y)+(1+w)e^{w}]^{4}} \cdot \frac{}{-1 + \frac{3(2+w)^{2\ \text{2 nosotros}}}{[r1(y)+(1+w)e^{w}](3+w)e^{w}}} \quad \text{ancho=alto}_0^{-1}\ (0) \tag{A30}$$

La sustitución de r0 y r1 produce el resultado requerido:

$$\begin{aligned} h_3^{-1}(0) = h_2^{-1}(0) &- \\ &\frac{4 + 1 - y^{2}\, w + (1 + y)\,2w\exp w + \frac{1+y}{2}\quad {}^{3}\ \exp w + \frac{1+y}{2}}{-(1-y)(3+w)+(1+y)(9+8w+2w\,2)\exp w+\frac{2\,1+y}{5}} \cdot \\ &\frac{}{6(1+año)^{2}\ 2\,1-y+(1+y)(1+w)\exp w+\frac{1+y}{2}} \quad \text{ancho=alto}_0^{-1}\ (0) \end{aligned} \tag{A31}$$

Machine Translated by Google

Referencias

1. Langevin, P. Sur la théorie du magnétisme. J. Física. Teor. Aplica. 1905, 4, 678–693. [Referencia cruzada]
2. Fiasconaro, A.; Falo, F. Resultados analíticos del modelo de cadena extensible de articulación libre. Física. Una estadística. Mec. Aplica. 2019, 532, 121929. [Referencia cruzada]
3. Iliafar, S.; Vézenov, D.; Jagota, A. Estiramiento de una cadena libremente articulada en dos dimensiones. arXiv 2013, arXiv:1305.5951.
4. Itskov, M.; Dargazany, R.; Hornes, K. Taylor expansión de la función inversa con aplicación a la función de Langevin. Matemáticas. Mec. Sólidos 2012, 17, 693–701. [Referencia cruzada]
5. Nguessong, AN; Beda, T.; Peyraut, F. Un nuevo enfoque de error basado para aproximar la función de Langevin inversa. Reol. Acta 2014, 53, 585–591. [Referencia cruzada]
6. Darabi, E.; Itskov, M. Una aproximación simple y precisa de la función de Langevin inversa. Reol. Acta 2015, 54, 455–459. [Referencia cruzada]
7. Jedynak, R. Revisión de la aproximación de la función de Langevin inversa. Reol. Acta 2015, 54, 29–39. [Referencia cruzada]
8. Kröger, M. Aproximantes simples, admisibles y precisas de las funciones inversas de Langevin y Brillouin, relevantes para fuertes deformaciones y flujos de polímeros. J. No Newton. Mecánica de fluidos. 2015, 223, 77–87. [Referencia cruzada]
9. Marchi, antes de Cristo; Arruda, EM Un enfoque de minimización de errores para aproximaciones inversas de Langevin. Reol. Acta 2015, 54, 887–902. [Referencia cruzada]
10. Rickaby, SR; Scott, NH Una comparación de modelos de elasticidad del caucho de estiramiento limitado. Int. J. Mec. no lineal. 2015, 68, 71–86. [Referencia cruzada]
11. Petrosyan, R. Aproximaciones mejoradas para algunos modelos de extensión de polímeros. Reol. Acta 2017, 56, 21-26. [Referencia cruzada]
12. Jedynak, R. Un estudio exhaustivo de los métodos matemáticos utilizados para aproximar la función de Langevin inversa. Matemáticas. Mec. Sólidos 2018, 24, 1992–2016. [Referencia cruzada]
13. Marchi, antes de Cristo; Arruda, EM Aproximaciones de Langevin inversas racionales y minimizadoras de errores generalizadas. Matemáticas. Mec. Sólidos 2019, 24, 1630-1647. [Referencia cruzada]
14. Howard, RM Aproximaciones analíticas para la función de Langevin inversa mediante linealización, aproximación de error e iteración. Reol. Acta 2020, 59, 521–544. [Referencia cruzada]
15. Mezo, I.; Keady, G. Algunas aplicaciones físicas de funciones de Lambert generalizadas. euros. J. Física. 2016, 37, 065802. [Referencia cruzada]
16. Mez˝o, I. Sobre la estructura del conjunto solución de una ecuación generalizada de Euler-Lambert. J. Matemáticas. Anal. Aplica. 2017, 455, 538–553. [Referencia cruzada]
17. Mezo, I.; Baricz, Á. Sobre la generalización de la función de Lambert W. Trans. Soy. Matemáticas. Soc. 2017, 369, 7917–7934. [Referencia cruzada]
18. Jamila, C.; Mendoza, R.; Mez˝o, I. Soluciones de ecuaciones diferenciales de retardo neutro utilizando una función W de Lambert generalizada. Aplica. Matemáticas. Computadora. 2020, 382, 125334. [Referencia cruzada]
19. Petkovi´c, MS; Petkovi´c, LD; Herceg, Đ. Sobre las familias de métodos de búsqueda de raíces de Schröder. J. Computación. Aplica. Matemáticas. 2010, 233, 1755–1762. [Referencia cruzada]
20. Dubeau, F. Aproximaciones polinómicas y racionales y el vínculo entre los procesos de Schröder de primer y segundo tipo. Abstr. Aplica. Anal. 2014, 2014, 719846. [Referencia cruzada]
21. Howard, RM Schröder basado en la aproximación de función inversa. Axiomas 2023, 12, 1042. [CrossRef]
22. Corless, RM; Gonnet, GH; Liebre, DEG; Jeffrey, DJ; Knuth, DE Sobre la función Lambert W. Adv. Computadora. Matemáticas. 1996, 5, 329–359. [Referencia cruzada]
23. Howard, RM Aproximaciones analíticas para la rama principal de la función Lambert W. euros. J. Matemáticas. Anal. 2022, 2, 14. [Referencia cruzada]
24. Lóczi, L. Evaluación garantizada y de alta precisión de la función Lambert W. Aplica. Matemáticas. Computadora. 2022, 433, 127406.
25. Schröder, E. Über unendlich viele Algorithmen zur Auflösung der Gleichungen. Matemáticas. Ana. 1870, 2, 317–365. [Referencia cruzada]
26. Berindé, V.; Takens, F. Aproximación iterativa de puntos fijos; Springer: Berlín/Heidelberg, Alemania, 2007.
27. Gdawiec, K.; Kotarski, W.; Lisowska, A. Polinomiografía basada en métodos no estándar de búsqueda de raíces similares a los de Newton. Abstr. Aplica. Anal. 2015, 2015, 797594. [Referencia cruzada]

Descargo de responsabilidad/Nota del editor: Las declaraciones, opiniones y datos contenidos en todas las publicaciones son únicamente de los autores y contribuyentes individuales y no de MDPI ni de los editores. MDPI y/o los editores renuncian a toda responsabilidad por cualquier daño a personas o propiedad que resulte de cualquier idea, método, instrucción o producto mencionado en el contenido.

Machine Translated by Google

AppliedMath

MDPI

Artículo

Aplicaciones de la Geometría Diferencial Vinculación Topológica Bifurcaciones a campos de flujo caóticos

Peter D. Neilson 1,* y Megan D. Neilson 2

1 Escuela de Ingeniería Eléctrica y Telecomunicaciones, Universidad de Nueva Gales del Sur, Sydney, NSW 2052, Australia Investigador

2 independiente, Escuela Superior de Ingeniería Eléctrica y Telecomunicaciones, Universidad de Nueva Gales del Sur, Sydney, NSW 2052, Australia; megan.neilson@gmail.com * Correspondencia: p.neilson@unsw.edu.au

Resumen: En cada punto p de una n-variedad M suave existen (n + 1) espacios tensoriales simétricos sesgados que abarcan formas r diferenciales ω con r = 0, 1, · ·,· n. Debido a que d ◦ d es siempre cero donde d es el diferencial exterior, se deduce que cada forma r exacta (es decir, ω = dλ donde λ es una forma (r−1)) es cerrada (es decir, dω = 0) pero No todas las formas r cerradas son exactas. Esto implica la existencia de un tercer tipo de forma r diferencial que es cerrada pero no exacta. Estas formas se denominan formas armónicas . Cada n-variedad suave tiene una estructura topológica subyacente. Existen muchas estructuras topológicas posibles diferentes . Lo que distingue una estructura topológica de otra es el número de huecos de diversas dimensiones que posee. La teoría de formas diferenciales de De Rham relaciona la presencia de agujeros de r dimensiones en la topología subyacente de una variedad n suave M con la presencia de campos armónicos de forma r en la variedad suave. Se requiere una gran cantidad de teoría para comprender el teorema de De Rham. En este artículo resumimos la geometría diferencial que vincula los agujeros en la topología subyacente de una variedad suave con campos armónicos en la variedad. Exploramos la aplicación de la teoría de De Rham a (i) sistemas visuales, (ii) mecánicos, (iii) eléctricos y (iv) de flujo de fluidos. En particular, consideramos los campos de flujo armónico en la solución acuosa intracelular de células biológicas y proponemos , sobre bases matemáticas, un posible papel de los campos de flujo armónico en el plegamiento de cadenas polipeptídicas de proteínas.

Palabras clave: complejos simpliciales; grupos de homología singulares; grupos de cohomología singulares; complejos de cadenas singulares ; álgebra de Grassmann; Formas laplacianas y armónicas; grupos de cohomología de de Rham; teorema de de Rham; plegamiento de proteínas

MSC: 53Z10; 53Z30

check for updates

Cita: Neilson, PD; Neilson, MD Aplicaciones de geometría diferencial que vinculan bifurcaciones topológicas con campos de flujo caóticos. Matemáticas aplicadas 2024, 4, 763–790. https://doi.org/10.3390/appliedmath4020041

Editor académico: Massimo Ferri

Recibido: 23 de abril de 2024
Revisado: 11 de junio de 2024
Aceptado: 11 de junio de 2024
Publicado: 15 de junio de 2024

Copyright: © 2024 por los autores. Licenciatario MDPI, Basilea, Suiza. Este artículo es un artículo de acceso abierto. distribuido bajo los términos y condiciones de los Creative Commons Licencia de atribución (CC BY) (https://creativecommons.org/licenses/by/4.0/).

1. Introducción

Una conocida cita de Einstein dice que en la medida en que las leyes de las matemáticas se refieren a la realidad, no son ciertas, y en la medida en que son ciertas, no se refieren a la realidad. Si bien esto puede parecer controvertido, lo cierto es que las teorías matemáticas de los procesos físicos requieren interpretación en términos de observaciones experimentales. La geometría diferencial es una estructura lógica de proposiciones, teoremas y lemas demostrables cuya verdad se deriva de axiomas, definiciones y los conectivos elementales de la lógica formal. Aplicar la geometría diferencial a los sistemas físicos requiere la construcción de puentes entre eventos y variables en el mundo físico (preferiblemente, pero no necesariamente, mensurables) y nociones bien definidas en la estructura deductivamente lógica de la geometría diferenc Estos puentes son hipótesis intuitivas que vinculan la realidad con las matemáticas y su valor deriva de la calidad de la intuición del practicante. Si la hipótesis es cierta, entonces se pueden aplicar ciertas leyes de las matemáticas al sistema del mundo real. Pero incluso si se demuestra experimentalmente que las leyes matemáticas describen el comportamiento en el mundo real, esto no prueba la hipótesis. Otras hipótesis pueden funcionar igual de bien.

Machine Translated by Google

De hecho, no es posible probar las conjeturas si-entonces, aunque se puede demostrar fácilmente que son falsas. En las matemáticas aplicadas, los "puentes si-entonces" entre la realidad y las matemáticas siguen siendo siempre intuitivos. Con respecto a la certeza, lo primario son las matemáticas.

Dicho y entendido esto, es valioso demostrar que las mismas matemáticas pueden aplicarse de manera útil a una variedad de sistemas del mundo real.

El teorema de cohomología de las formas diferenciales de De Rham es una pieza matemática notable que vincula áreas de las matemáticas que antes se pensaba que no estaban relacionadas. Creemos que los hechos matemáticos proporcionados por el teorema de De Rham tienen gran relevancia para el comportamiento no lineal y la presencia de puntos críticos y singularidades en los espacios de estados de muchos sistemas diferentes del mundo real. En este artículo, consideramos una interpretación del teorema en términos de observaciones experimentales de la dinámica del flujo de fluidos. En esencia, la teoría de la cohomología de De Rham relaciona la existencia de agujeros de varias dimensiones en la topología subyacente de variedades suaves y riemannianas con un tercer tipo de campo de flujo de forma diferencial en la variedad, uno que es cerrado pero no exacto. Un campo de flujo de este tipo se denomina campo armónico . En otras palabras, nuestro objetivo es demostrar los vínculos entre los campos armónicos de las matemáticas y los flujos físicos de sistemas no lineales del mundo real con singularidades y puntos críticos observados en la naturaleza.

Para ilustrar el tipo de procesos dinámicos no lineales en el mundo que pueden entenderse a través de esta vinculación matemática, consideremos las burbujas y el flujo de fluido asociados con el vertido de una copa de champán, o la generación y el flujo de espuma asociados con una ola del océano rompiendo en el océano. una playa, o el caótico y turbulento flujo de un arroyo de montaña entre rocas. Incluso la metamorfosis de una oruga en polilla puede estar dentro del alcance de la teoría. Ciertas singularidades en la dinámica no lineal de los sistemas visuales, mecánicos, eléctricos y de flujo de fluidos pueden relacionarse con campos armónicos. En particular, examinaremos el posible papel de los vórtices y remolinos en los campos de flujo armónico inducidos por la presencia de orgánulos y otras moléculas inmersas en la solución acuosa intracelular de células biológicas en la cinética del plegamiento de proteínas.

Nuestra búsqueda de este vínculo se inspiró en dos acontecimientos aparentemente no relacionados. En primer lugar, en nuestro análisis de la geometría del espacio visual binocular 3D [1–3], nos dimos cuenta de que se deben tener en cuenta tres geometrías 3D diferentes. Por supuesto, existe la geometría euclidiana 3D del mundo exterior. Pero la geometría del espacio visual tridimensional codificado por los receptores sensoriales está deformada (o curvada) por el hecho de que el tamaño de las imágenes en las retinas de los ojos varía en proporción inversa a la distancia euclidiana entre el objeto en el entorno y el punto nodal. del ojo [4]. Este espacio 3D intrínseco corresponde a una variedad de Riemann curva con una métrica proporcional a la inversa de la distancia euclidiana al cuadrado [1]. Esto da lugar a una singularidad en el origen donde la métrica llega al infinito. El espacio visual 3D percibido conscientemente vuelve a ser diferente porque entran en juego muchos mecanismos cognitivos diferentes de percepción de profundidad.

La singularidad en el origen se soluciona mediante la existencia de un agujero alrededor del egocentro en el espacio visual. La existencia de tal agujero se confirma por el hecho de que no es posible ver tu propia cabeza, y mucho menos tu propio egocentro. Esto plantea la pregunta: ¿qué efecto tiene tal agujero en el espacio visual sobre el flujo percibido de imágenes visuales?

El segundo evento se refería a un grupo inusual de diagnósticos de cáncer de mama que ocurrieron en mujeres que trabajaban juntas en un estudio de televisión australiano [5,6]. Se especuló que los cánceres podrían ser causados por fugas de radiación electromagnética del transmisor de televisión. Se trajo a expertos para probar esta hipótesis. No encontraron evidencia de niveles de radiación problemáticos en el lugar de trabajo de las mujeres. En las teorías físicas del electromagnetismo, existen al menos dos tipos de campos eléctricos. Uno está dado por gradientes de voltaje y la curvatura del otro está dada por campos magnéticos cambiantes. Estos dos tipos de campo son independientes porque la curvatura de un gradiente siempre es cero. Surge la pregunta: ¿existe un misterioso tercer tipo de campo eléctrico que no habría sido detectado por las técnicas de medición estándar? La geometría diferencial sugiere que sí. Según el teorema de de Rham, la curvatura de un campo eléctrico puede ser cero sin que sea el gradiente

Machine Translated by Google

f
de un voltaje. La pregunta es: ¿existe ese tercer tipo de campo en la naturaleza y, de ser así, qué propiedades tiene?

Aplicar la geometría diferencial a sistemas físicos requiere un conocimiento profundo de la teoría matemática, porque la secuencia en la que se aplica no sigue necesariamente la hermosa secuencia lógica en la que se desarrollaron y demostraron los teoremas. La aplicación a menudo requiere saltar de un punto de la teoría matemática a un punto distante con una gran cantidad de matemáticas uniendo los dos puntos. Además, la comprensión del teorema de De Rham no es trivial. Requiere familiaridad con una gran cantidad de geometría diferencial. En consecuencia, en el Apéndice A proporcionamos detalles de todos los pasos en nuestra propuesta de vinculación matemática entre la aparición y desaparición de agujeros en la topología subyacente de variedades suaves y riemannianas y la aparición y desaparición de campos de flujo armónico en la variedad. Mientras tanto, en la siguiente sección brindamos al lector una introducción más conceptual a la geometría diferencial utilizada para relacionar bifurcaciones topológicas con campos de flujo caóticos en la naturaleza.

2. Resumen de conceptos relevantes de geometría diferencial 2.1. Estructura topológica Una propiedad

importante de los espacios topológicos es que no están equipados con una métrica. No tienen medida de tamaño. El tamaño y la topología pueden considerarse propiedades independientes. Un espacio topológico, como el conjunto de todos los puntos de una 2-esfera, puede expandirse al tamaño de un planeta o contraerse al tamaño de un átomo sin cambiar su estructura topológica. Siempre que no se corte, rasgue, perfore o se peguen partes del mismo, se puede deformar de cualquier manera en cualquier cantidad sin cambiar su estructura topológica. Por ejemplo, las 2 esferas se pueden estirar hasta formar una cuerda larga y delgada y hacer un nudo sin cambiar su topología.

Todas las diferentes formas en las que se puede deformar un espacio topológico sin cortar, rasgar, perforar o pegar se pueden mapear de manera continua, uno a uno, sobre y continuamente invertible entre sí. Este es un tipo especial de mapa llamado homeomorfismo . Dos espacios homeomórficos cualesquiera tienen las mismas propiedades topológicas. De hecho, una propiedad topológica se puede definir como aquella que se conserva mediante mapas homeo

Las transformaciones que dejan la estructura topológica sin cambios se denominan homotopías, siendo un ejemplo muy conocido la transformación de una taza de café con asa en un donut y viceversa [7]. El mapeo matemático ([8], p. 348]) se muestra en la Figura 1.

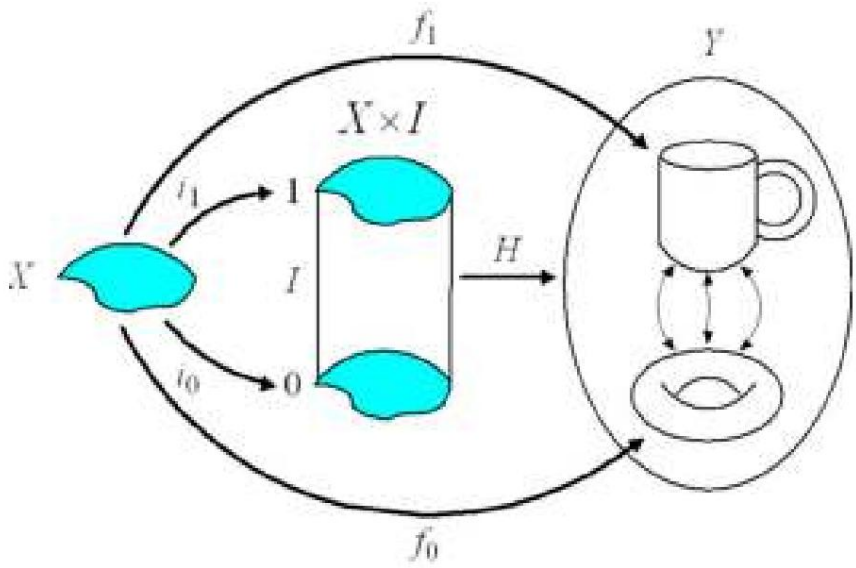


Figura 1. Un diagrama que ilustra la homotopía. El 2-toro (superficie de una dona) en el espacio topológico Y y la taza de café con asa en el espacio topológico Y son estructuras topológicas equivalentes.

Estos no sólo están relacionados homeomórficamente (es decir, continuos, uno a uno, sobre e invertibles), sino que

Machine Translated by Google

hay aplicaciones homotópicas continuas f0, f1 : X → Y que asignan el espacio topológico X tanto al 2-toro como a la taza. El hecho de que el donut pueda transformarse continuamente en taza implica la existencia de todas las formas intermedias en la transformación. El agujero en el donut se transforma en el agujero entre la taza y su asa. Esto se ilustra en el diagrama mediante el caso especial en el que Y = X × I, f0 = i0, f1 = i1 , donde el intervalo I varía de 0 a 1. Para cada punto x X, el mapa f0(x) = i0(x) = (x, 0) corresponde al toro 2, el mapa f1(x) = i1(x) = (x, 1) corresponde a la taza y todos los demás puntos a lo largo del intervalo I corresponden al formas intermedias. Esto se representa en el diagrama mediante la homotopía H. De manera más general, si bien los espacios topológicos X e Y pueden incluir huecos de varias dimensiones, deben tener la misma estructura topológica para que exista la homotopía H. En otras palabras, para que f0, f1 : X → Y sean mapas homotópicos, los homomorfismos inducidos (f0) ,(f1) : Hr(X) → Hr(Y) para todo r ≥ 0 entre los grupos de homología para X y Y debe ser igual, es decir, (f0) = (f1) .

Sin embargo, no hay manera de que una 2 esferas pueda transformarse en un 2 toro sin perforar agujeros en la esfera. Las 2 esferas y los 2 toros no tienen la misma estructura topológica. AA Markov demostró en 1958 (ver [8]) que no existe ningún algoritmo que pueda determinar todas las diferentes estructuras topológicas que puede tener un espacio topológico con dimensión 4 o mayor. Lo mejor que podemos decir es que dichos espacios tienen una gran cantidad de posibles estructuras topológicas diferentes. Básicamente, lo que hace topológico
estructuras diferentes es el diferente número de agujeros de varias dimensiones que existen en el
espacio topológico (por ejemplo, la diferencia entre la topología de una esfera y la topología de un toro es atribuible a los agujeros en el toro).

2.2. Bifurcaciones topológicas

Basado en Smale [9], Abrahams y Marsden [10] proporcionan una descripción de las bifurcaciones en la topología subyacente de una variedad suave. Un mapa suave f : M → N entre variedades suaves M y N es localmente trivial en y0 N si hay una vecindad U de (y) es una subvariedad suave y0 N tal que f M $^{-1}$ apropiadamente incrustada (es decir, un conjunto de niveles) en (y0) para para cada y U en N y existen aplicaciones difeomorfas h| f −1(y) cada y : f $^{-1}$ (y) → f $^{-1}$ U (es decir, h proyecta conjuntos de niveles $^{-1}$ (y) en M difeomorficamente en el nivel $^{-1}$ (y0) en M).

cercanos f set f El conjunto de bifurcación de f : M → N es Σf := y0 N f no es localmente trivial en En otras palabras, cuando y pasa por Σf , f $^{-1}$ y0 . (y) cambia su estructura topológica. denota el conjunto cerrado de puntos críticos x de f en M, es decir,

$$\sigma(f) = x \quad \text{METRO| } dx\, f : TxM \to Tf(x)N \text{ no es sobreyectivo.}$$

Según el teorema de Sard, los valores singulares f(σ(f)) en N tienen medida cero (es decir, son puntos aislados con interior vacío y su unión no es densa en ninguna parte en N). Los valores singulares de f(σ(f)) son los valores de bifurcación de f en N [10]. En otras palabras, los conjuntos de niveles en M de valores de bifurcación singulares en N tienen una estructura topológica que difiere de las estructuras topológicas de los conjuntos de niveles vecinos en M. El número de agujeros en el nivel establecido en M de un valor de bifurcación en N difiere de el número de agujeros en los conjuntos de niveles vecinos. Esto puede estar relacionado con los subgrupos de isotropía, el teorema de la rebanada, el teorema del tubo, la subvariedad de tipo órbita (H), la subvariedad de tipo isotropía (H) y la subvariedad de tipo punto fijo (H) descrita por Ortega y Ratiu [11] que amplían la noción de puntos críticos en subvariedades.

2.3. Teoría de De Rham

Ofrecemos aquí sólo una breve descripción. Un resumen más ampliado de diferencial
La geometría que vincula las bifurcaciones con los campos de flujo armónico se proporciona en el Apéndice A.

Machine Translated by Google

Cada diferencial exacto de la forma r ω con r = 0, 1, · · , n en un suave n-dimensional · variedad M o una variedad de Riemann n-dimensional (M, g) es cerrada (es decir, dω = 0) porque d ◦ d es siempre cero. Pero una forma r diferencial cerrada no es necesariamente exacta. Las formas r diferenciales que son cerradas pero no exactas dan lugar a un tercer tipo de forma r diferencial. Por lo tanto, existen tres tipos de formas r diferenciales, (i) formas r ω cuya derivada exterior no es cero (es decir, dω = 0), (ii) formas r cerradas que son diferenciales exactas (es decir, ω = dλ donde λ es una (r − 1)-forma) y (iii) r-formas ω que son cerradas pero no exactas. Este tercer tipo de forma r se llama forma r armónica y está contenida en el espacio armónico r-ésimo Hr (M) de la variedad con r = 0, 1, · · · n. También es un elemento en el r-ésimo grupo de cohomología de Rham Hr dR(M) con r = 0, 1, · · · n. Hr dR(M) es isomorfo con Hr (M). Tanto Hr (M) como Hr dR(M) son subespacios vectoriales además de ser grupos.

Si ω es una forma r cerrada (es decir, dω = 0) en una variedad suave M, y zr es un ciclo r-dimensional en la topología subyacente X de la variedad, entonces la integral de ω sobre zr se llama período de ω (el término período deriva de los períodos de integrales elípticas). Según el teorema de Stokes, si el ciclo r zr resulta ser un límite br de una cadena (r + 1) cr+1 (es decir, br = ∂cr+1), entonces el período desaparece ya que

$$\omega = \omega = br \quad \omega = d\omega = 0 \text{ (ya que } d\omega = 0\text{).}$$
$$z_r \qquad \partial c_{r+1} \qquad c_{r+1}$$

(Ver Apéndice A.1 para definiciones de cadenas, ciclos y límites).

De esto obtenemos el primer teorema de De Rham: una forma cerrada es globalmente exacta si y sólo si todos sus períodos desaparecen [12]. Saber qué formas son globalmente exactas y cuáles no tiene consecuencias importantes. Por ejemplo, (i) una forma 1 suave es conservadora (es decir, su integral de línea es independiente de la trayectoria) si y sólo si es exacta, (ii) la integral de cada forma globalmente exacta sobre una variedad suave compacta orientada sin límite es cero, (iii) la integral de formas globalmente cerradas sobre límites es cero y (iv) si la integral de una forma r cerrada sobre una subvariedad r-dimensional sin límite no es cero, entonces la forma r no es exacta y la subvariedad no es el límite de una subvariedad orientada compacta con límite.

Una propiedad notable de los espacios armónicos y los grupos de cohomología de De Rham es que son homotópica y topológicamente invariantes. Son isomorfos a los grupos de cohomología singulares Hr (M), y estos a su vez son isomorfos a los grupos de homología singulares Hr(M) en la estructura topológica subyacente de la variedad suave M (ver Apéndice A.12). Esto se representa en la Figura 2.

Esta es de hecho una propiedad notable, porque no es inmediatamente obvio que la cercanía y exactitud de las formas diferenciales en una variedad suave deban tener algo que ver con la estructura topológica subyacente de la variedad. ¡Pero la teoría de De Rham muestra matemáticamente que sí! El r-ésimo grupo de homología Hr(M) es isomorfo al r-ésimo grupo de cohomología de Rham Hr dR(M) con r = 0, 1, · · · n. Hr(M) detecta la presencia de agujeros de r dimensiones en la topología subyacente de la variedad suave M. La dimensión de Hr(M) es igual al número de agujeros de r dimensiones en la topología, y esto se conoce como r-ésimo Betti número br del colector. Por lo tanto, la teoría de formas diferenciales de De Rham proporciona el vínculo que buscamos entre los agujeros en la estructura topológica de una variedad suave o riemanniana y los campos de flujo armónicos en la variedad.

Machine Translated by Google

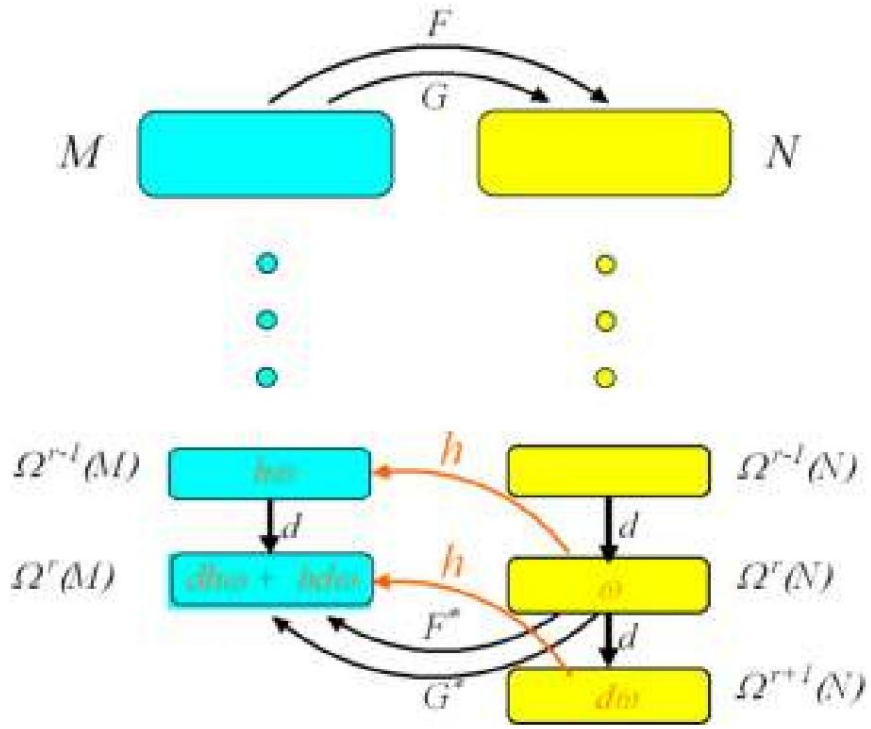


Figura 2. Un diagrama que ilustra el hecho de que si F, G : M → N son aplicaciones difeomorfas entre : Hr dR(N) → Hr dR(M) entre las variedades F de Rham Los grupos de cohomología Hr dR(N) y Hr dR(M) para todos los valores de r , suaves M y N, entonces las aplicaciones de retroceso son iguales, es decir, F. En otras palabras, si los mapas F y G son difeomorfos, tienen la misma estructura topológica homeomórfica y, = . En por lo tanto, son invariantes de homotopía. . Esto significa que las variedades equivalentes a homotopía tienen grupos de Rham isomórficos. Las etiquetas Ωr−1 (M), Ωr (M) representan los espacios tensoriales simétricos sesgados de grado (r − 1) y grado (r) para formas diferenciales en cada punto de la variedad suave M. Las etiquetas Ωr−1 (N), Ωr (N), Ωr+1 (N) representan los espacios tensoriales simétricos sesgados de grado (r − 1), grado (r) y grado (r + 1) para formas diferenciales en cada punto de la variedad suave. N. Ωr+1 (N) → Ωr (M) para todos los valores de r se denominan operadores de homotopía. Examen de los mapas h : el diagrama muestra que si la forma r ω es cerrada, es decir, dω = 0, entonces d(h ω) = G ω − F ω. En otras palabras, G ω y F ω difieren por un diferencial exacto d(h ω). Se dice que los mapas F son cohomólogos. Así es exactamente como de Rham. se definen los grupos de cohomología iguales. Por lo tanto, los grupos cohomólogos tienen grupos de cohomología isomórficos

: Hr dR(N) → Hr dR(M)

2.4. El operador laplaciano de Hodge

Por definición, el operador laplaciano de Hodge en una variedad de Riemann es Δ : dd + d d, donde d es la derivada exterior d : Ωr → Ωr+1 d , donde, re es el operador estrella de Hodge : Ωr → Ωn −r y Ωr es el espacio tensor simétrico sesgado de r-formas para r = 0, 1, · · · n en cada punto p en una variedad suave n-dimensional M o una variedad de Riemann (M, g). El operador laplaciano Δ juega un papel importante en la teoría de formas diferenciales de De Rham porque es el único operador que puede usarse para calcular formas armónicas en una variedad de Riemann. El espacio armónico Hr (M) = ker(ΔΩr) para r = 0, 1, · · · n. (Tenga en cuenta que las formas r armónicas para todos los valores de r en una variedad suave o una variedad de Riemann están asociadas con la presencia de agujeros de dimensión r en la topología subyacente de la variedad).

2.5. Formas 0 y 1 armónicas

Las formas 0 armónicas están dadas por la ecuación Δ φ = d d φ = div grad φ, donde φ es una función suave de valor real en la variedad suave M o variedad de Riemann (M, g).
Los campos potenciales en un fluido son formas 0 armónicas irrotacionales e incompresibles (es decir, libres de divergencia) . Nos centramos aquí en calcular formas 1 armónicas en variedades de Riemann porque

Machine Translated by Google

de su importancia en la dinámica no lineal de los sistemas físicos que se analizarán más adelante.

Sea λ una forma diferencial 1 en Ω1 = T p M en cada punto p (M, g). Primero, calculamos el Laplaciano de Hodge de una λ de 1 forma en una variedad de Riemann (M, g), es decir, ∆λ := dd λ + d dλ, y luego calculamos la λ de 1 forma armónica estableciendo la Laplaciano ∆λ = 0.

El tensor de curvatura de Ricci se puede describir geométricamente de la siguiente manera. Sea (M, g) una variedad de Riemann y p un punto perteneciente a (M, g) (es decir, p (M, g)). Para cada vector unitario ξ TpM, Ric(ξ, ξ) es la suma de las curvaturas seccionales de los 2 planos en TpM abarcadas por (ξ, ξ2), · · · ,(ξ, ξn), donde (ξ1, · · · , ξn) es cualquier base g-ortonormal para TpM con ξ1 = ξ. Dado cualquier vector unitario ξ TpM, elija una base g-ortonormal (ξ1, · · · , ξn) para TpM tal que ξ1 = ξ. Entonces Ric(ξ, ξ) viene dada por

$$R11 = Ri11^{i} = \sum_{yo=1}^{norte} Rm(\xi_i, \xi_1, \xi_1, \xi_i) = \sum_{yo=1}^{norte} (\xi_i \xi - \xi \xi_i)\xi, \xi_i .$$

Siguiendo a Lang [13] (págs. 423–424), el producto interno del tensor de curvatura de Ricci con el vector ξ (es decir, Ric(λ), ξ) se define como la forma escalar tal que, con con respecto a un marco ortonormal ξ1, · · · , ξn, y cualquier campo vectorial unitario ξ con ξ1 = ξ, tenemos

$$\mathrm{Ric}(\lambda), \xi = \sum_{yo=1}^{norte} (\xi \xi_i - \xi_i \xi)\lambda, \xi_i$$

donde denotamos por λ, ξ el valor de la forma 1 λ en el campo vectorial ξ. Como operador 2 en formas 1, el laplaciano ∆ :) − Ric). Escrito en términos ξi de variables, esto significa $\Delta\lambda, \xi = -\sum_{yo=1}^{norte}$ ξi ξi λ, ξ − ∑i=1 (ξ ξi − ξi ξ)λ, ξi . Ω1 → Ω1 está dado por (− ∑

Para calcular el producto interno de Hodge Laplaciano ∆λ, ξ = (dd + d d)λ, ξ , calculamos (i) dd λ, ξ y (ii) d dλ, ξ por separado y luego sumamos: (i) Tenemos d λ = −divλ = −tr(λ) = − ∑ el campo de 1 forma λ y ξi para i $\sum_{yo=1}^{norte}$ ξi λ, ξi , donde λ es el dual de = 1, · · · n es una g-ortonormal marco de campos vectoriales.

Estos vectores ortonormales deben calcularse utilizando el algoritmo ortogonalizador de Gram-Schmidt de la métrica de Riemann en cada punto. Vemos que d λ es una función de valor real f = − ∑ ξi λ, ξi en (M, g). Para cada punto $\sum_{yo=1}^{norte}$ (M, g), tenemos ξ f = ξ f = df , ξ . Así podemos escribir

$$\xi(-\sum_{1}^{norte} \xi_i \lambda, \xi_i) = d(-\sum_{yo=1}^{norte} \xi_i \lambda, \xi_i), \xi = dd \lambda, \xi = -\sum_{yo=1}^{norte} \xi \xi_i \lambda, \xi_i . \quad (1)$$

(ii) Para obtener d dλ, primero observamos que si ω es una forma r, podemos usar la proposición dω(ξ0, ξ1 , · · · , ξr) = $\sum_{yo=0}^{r}$ (−1) i ξiω(ξ1 , · · · , ξi−1 , ξ0, ξi+1 , · · · , ξr) para obtener

$$d\lambda(\xi, \eta) = \xi\lambda, \eta - \eta\lambda, \xi \quad (2)$$

donde λ es una forma 1 y ξ y η son vectores. Usando el hecho de que para cualquier forma r ω

$$d\omega = \sum_{yo=0}^{r} (\lambda_i \xi_i\omega) \text{ y } d(\omega) = \sum_{yo=1}^{norte} (\lambda_i \xi_i(\omega))$$

donde ξi para i = 1, · · · , n es un marco ortonormal y λi para i = 1, · · · , n es un coframe ortonormal, se deduce que

$$d\omega = d\omega = \sum_{yo=1}^{norte} (\xi_i\omega \circ \xi_i) \quad (3)$$

donde (ξiω ◦ ξi) es la notación para ξiω contraída por ξi [13].

De las Ecuaciones (3) y (2) con η = ξi , obtenemos [13] (p. 424)

$$d\, d\lambda, \xi = \sum_{yo=1}^{norte} \xi_i \xi_i\lambda, \xi_i - \sum_{yo=1}^{norte} \xi_i \xi_i \lambda, \xi \quad (4)$$

Finalmente, sumamos las ecuaciones (1) y (4) para obtener

$$\Delta\lambda, \xi = (dd + d\, d)\lambda, \xi = -\sum_{yo=1}^{norte} \xi_i \xi_i \lambda, \xi - \sum_{yo=1}^{norte} (\xi \xi_i - \xi_i \xi)\lambda, \xi_i . \quad (5)$$

Machine Translated by Google

Por lo tanto, para una forma 1 en una variedad de Riemann n-dimensional (M, g), Hodge 2 ξi
El operador laplaciano tiene la forma Δ = (− ∑) − Ric, donde Ric es el operador de curvatura de Ricci. La forma
yo=1
armónica 1 se obtiene estableciendo el Laplaciano Δ λ igual a cero.

3. Ejemplos físicos Existe

una infinita variedad de comportamientos en la naturaleza de sistemas dinámicos no lineales asociados a diferentes singularidades, puntos críticos y degeneraciones en sus ecuaciones diferenciales. La mayoría de estos comportamientos aún no han sido descritos, y mucho menos resueltos. Una familia de tales comportamientos considerada en este artículo está asociada con la existencia de agujeros de varias dimensiones en la topología subyacente del espacio de estados del sistema. Semejante
Los agujeros crean formas diferenciales y campos vectoriales que son cerrados pero no exactos y crean espacios 'prohibidos' en el estado dinámico del sistema. Estos agujeros pueden cambiar drásticamente el comportamiento dinámico del sistema. Los subgrupos de isotropía en la acción de un grupo de Lie sobre un sistema dinámico, por ejemplo, crean estados dinámicos en los que un campo vectorial (o un subconjunto de campos vectoriales) puede entrar pero no salir; un tipo de trampa de estado dinámico de punto de equilibrio . Pero los sistemas dinámicos no lineales pueden pasar a campos vectoriales vecinos y fluir alrededor de la trampa, creando comportamientos no lineales novedosos [11]. A continuación, consideramos las consecuencias sobre el comportamiento dinámico de las singularidades en las ecuaciones diferenciales que dan lugar a formas diferenciales que son cerradas pero no exactas en sistemas físicos visuales, mecánicos, eléctricos y de flujo de fluidos.

3.1. El sistema visual humano

El espacio visual binocular 3D humano codificado por receptores sensoriales visuales, propioceptivos y vestibulares está deformado (o curvo) porque el tamaño de las imágenes 2D en la retina de cada ojo varía en proporción inversa a la distancia euclidiana entre el punto nodal del ojo y el objeto. en el medio ambiente. Geométricamente hablando, el espacio visual 3D corresponde a una variedad riemanniana egocéntrica curvada negativamente con una bola 3D de puntos que faltan aproximadamente en el origen (es decir, el egocentro) [1,2]. Este agujero intrínseco corresponde al hecho de que no es posible ver la propia cabeza. ¿Qué influencia tiene tal agujero sobre el origen en el espacio visual deformado en 3D en el flujo percibido de imágenes visuales?

La geometría diferencial y el teorema de De Rham proporcionan una posible respuesta a esta pregunta. (Tenga en cuenta que es posible que existan otros agujeros en el espacio visual 3D debido a una patología neural. Conocidos como escotomas, no se consideran aquí).

Dado un agujero 2D alrededor del origen (es decir, una bola 3D faltante de puntos alrededor del origen encerrada dentro de un ciclo 2), el teorema de de Rham implica que las formas r diferenciales cerradas para r = 0, 1, 2, 3 no son globalmente exacto. La forma de volumen unitario Ω = dx1 dx2 dx3 es una forma 3 cerrada en el espacio visual 3D deformado y, por lo tanto, el agujero alrededor del origen impide que esta forma 3 de volumen sea globalmente exacta. Esto es mejor, porque si no hubiera un agujero alrededor del origen y la forma de volumen cerrado Ω fuera globalmente exacta, entonces todos los períodos Ω iguales a la forma de volumen integrados en todas las 3 cadenas (es decir, volúmenes) encerrados por 2- los ciclos desaparecerían. La topología subyacente
c3
del espacio visual 3D deformado sería tal que el volumen percibido Ω de cualquier espacio 3D cerrado sería igual a cero. En otras palabras, no sería posible detectar el tamaño (es decir, el volumen) de espacios 3D cerrados.
c3

Sabemos por experiencia cotidiana que podemos sentir visualmente el tamaño de la habitación en la que nos encontramos. Esto es lo que predice el teorema de De Rham.

3.2. Sistemas mecánicos En

ausencia de todas las fuerzas externas, la mecánica hamiltoniana se desarrolla en el paquete cotangente de 2n dimensiones T Q = P (o espacio de fase de momento) de una variedad de configuración de n dimensiones Q. Usando la posición (configuración) q y el momento p como coordenadas canónicas, una forma canónica 1 θ = pidqi y una forma canónica cerrada ω = dqi dpi existen en cada punto (q, p) en T Q = P (es decir, en el espacio de fase de momento del hamiltoniano sistema). La estructura simpléctica de la forma bidimensional ω = dqi dpi define una geometría simpléctica en el sistema. Contracción del campo simpléctico de 2 formas ω por un campo vectorial ξ

Machine Translated by Google

da un campo de 1 forma λ = Cξω en T Q = P. Sin embargo, hay un descubrimiento maravilloso que se aplica a todos los sistemas mecánicos hamiltonianos (y a todos los sistemas cuya dinámica puede describirse mediante ecuaciones diferenciales de segundo orden). Existe un campo vectorial especial ξh para el cual el campo de 1 forma λ = Cξh ω es un campo de 1 forma exacto igual a dh, donde h es una función suave de valor real en el espacio de fase de momento T Q = P. La función h de valor real se llama función hamiltoniana, y el campo vectorial ξh se llama campo vectorial hamiltoniano del sistema mecánico hamiltoniano. Si la función de valor real h es igual a la energía total del sistema mecánico (es decir, igual a la suma de la energía cinética y la energía potencial del sistema expresadas como una función de la posición y el momento), entonces el flujo integral de la El campo vectorial hamiltoniano ξh en el espacio de fase de momento T Q = P es igual al movimiento libre del sistema mecánico. Dado que, por definición, no existe una fuerza externa que inyecte o disipe energía, el flujo integral del campo vectorial hamiltoniano ξh conserva la energía total h del sistema. El flujo integral está confinado a una capa de contorno de energía constante (2n − 1) dimensional (es decir, h constante) en T Q = P. Diferentes niveles de energía corresponden a flujos en diferentes capas de contorno de energía constante en el espacio de fase de momento de la sistema.

Sin embargo, si existe un agujero unidimensional en el espacio de fase de momento, entonces, según el teorema de De Rham, la forma unidimensional λ = Cξω no puede ser globalmente exacta y, por lo tanto, ni la función hamiltoniana h ni el campo vectorial hamiltoniano ξh pueden ser definido globalmente! El campo exacto de 1 forma λ = dh se reemplaza por un campo armónico de 1 forma

$$\lambda \quad H^1(P) = HdR^1(P)$$

eso es cerrado pero no exacto.

3.3. Sistemas electrodinámicos

En electrodinámica en el espacio-tiempo 4D de Minkowski, se pueden escribir las ecuaciones de Maxwell en términos de formas diferenciales como

$$F^{\mu\nu}_{,\nu} = F\mu\nu,\rho + F\nu\rho,\mu + F\rho\mu,\nu = 0 \text{ (fuente gratuita)}$$
$$F^{\mu\nu}_{,\nu} = \frac{4\pi}{do} J\,\mu \text{ (fuente acoplada),}$$

donde F es una forma 2 electromagnética sesgada en cada punto del espacio-tiempo 4D de Minkowski, J es un 4-vector actual en cada punto, c es la velocidad de la luz en el vacío y es el operador estelar de Hodge (todo en gaussiano). unidades).

Las ecuaciones sin fuente son verdaderas si y sólo si existe un campo de 4 covectores A (es decir, un campo de 1 forma) llamado campo de potencial electromagnético tal que F = dA (es decir, F es un campo de 2 formas cerrado y globalmente exacto). campo de forma en el espacio-tiempo plano de Minkowski 4D).

Tal campo A de 1 forma puede alterarse mediante la adición de un diferencial exterior dψ de una función suave de valor real ψ en la variedad de Minkowski 4D (es decir, A = A + dψ) sin cambiar el campo electromagnético cerrado de 2 formas F. F no cambia porque d ◦ d = dd es siempre cero y F se puede escribir como.

$$F = dA = dA + ddψ = dA$$

Cualquier transformación del potencial electromagnético A que no tenga efecto sobre el campo electromagnético F se denomina transformación de calibre. Cualquier elección de ψ que satisfaga la ecuación de onda no homogénea se denomina calibre de Lorentz; es decir,

$$\Box\psi = -A\,\mu\,{}_{,\mu} \text{ de modo que}$$
$$A_{\mu,\mu} = A\,\mu,\mu + \Box\psi = A\,\mu\,{}_{,\mu} - A\,\mu,_{\mu} = 0$$

donde el operador □ es el operador de onda o el d'alembertiano

$$\Box\psi = 2\psi - \frac{1}{\text{tazas}^2}\frac{\partial^2\psi}{\partial t^2}$$

Machine Translated by Google

Esto muestra que un potencial electromagnético A en el calibre de Lorentz puede transportar una onda electromagnética □ψ propagándose a la velocidad de la luz c en el vacío sin cambiar el campo electromagnético F. El campo electromagnético F es el medio en el que la onda se propaga sin necesidad de un éter. Es importante notar que

$$\Box\psi = -A\,\mu\ {}_{,\mu} = -\text{div} \circ \text{graduado}\ \psi = -\quad 2\psi = -\Delta\psi,$$

es decir, el operador de onda d'Alembertiano □ en un espacio-tiempo plano de Minkowski 4D es igual al laplaciano ∆ en un espacio-tiempo plano 4D.

Sin embargo, si existe un agujero unidimensional en la variedad espacio-temporal 4D, entonces, según el teorema de De Rham, el potencial electromagnético A no puede ser globalmente exacto y, en consecuencia, no se puede definir el calibre de Lorentz. El campo A de 1 forma de potencial electromagnético global se reemplaza por un campo armónico de 1 forma. Esto deja incierta la cuestión de la existencia de un éter.

3.4. Flujo de fluido

Los principios de la mecánica de fluidos han sido durante mucho tiempo un estándar en la enseñanza de la física [14-16]. Considere una corriente de agua que fluye lentamente. Las partículas de agua cercanas al límite se adhieren al límite y, por lo tanto, tienen velocidad cero aguas abajo. Las moléculas de agua en esta capa límite forman enlaces de hidrógeno entre sí que aumentan la tensión superficial y alteran la distribución dipolar de la carga eléctrica. Los puntos alejados del límite hacia el centro de la corriente tienen velocidades promedio de flujo aguas abajo más altas. Este patrón de flujo se conoce como flujo laminar. El movimiento aleatorio de las partículas de agua debido a la energía térmica y los campos eléctricos provoca un intercambio de partículas de agua entre las distintas láminas, donde las partículas de mayor velocidad aceleran a las más lentas y viceversa. Esto da lugar a tensiones de corte y transferencia de energía cinética y momento entre las distintas capas, produciendo una propiedad del agua conocida como viscosidad dinámica μ. La viscosidad varía entre diferentes fluidos y proporciona una medida de la facilidad de flujo. La miel, por ejemplo, tiene una viscosidad mayor que el agua y fluye con menos facilidad.

Consideremos ahora un arroyo de montaña que fluye rápidamente entre rocas. Las partículas de agua no pueden fluir a través de las rocas, por lo que tienen que fluir alrededor de ellas. Las rocas forman agujeros 2D de varios tamaños en el espacio del agua (es decir, la superficie 2D de cada roca es un ciclo de 2 que forma un límite alrededor de un agujero en el espacio del flujo de agua). Si la corriente fluye lo suficientemente rápido (debido a un amplio suministro de energía cinética proporcionada por la gravedad), el flujo alrededor de las rocas induce grandes remolinos y vórtices en el flujo. Estos grandes remolinos se propagan hacia afuera, transfiriendo energía cinética a remolinos y vórtices más pequeños. Estos, a su vez, generan remolinos aún más pequeños. Esta cascada de energía cinética de remolinos más grandes a más pequeños depende de la velocidad del flujo (es decir, la energía cinética disponible) y la tasa de disipación de energía en el agua debido a la viscosidad. Los remolinos de varios tamaños aparecen y desaparecen de forma impredecible, lo que provoca que las velocidades y presiones tangenciales locales varíen de un punto a otro de forma caótica e impredecible. Esto aumenta la cantidad de mezcla. El aumento de las velocidades del flujo local reduce la presión local, lo que permite que los gases del agua se vaporicen y produzcan burbujas. La formación y el estallido de burbujas disipan la energía y producen un sonido de "gorgoteo". El patrón caótico de flujo asociado con tal cascada de energía desde remolinos grandes a pequeños con una mínima disipación de energía se llama flujo turbulento. El flujo turbulento se observa en muchos flujos de fluidos en la naturaleza, como las olas del océano, las nubes de tormenta, el humo, las burbujas de champán y más.

3.4.1. Ecuaciones de Navier-Stokes

Las ecuaciones de Navier-Stokes son ecuaciones diferenciales parciales no lineales que describen el movimiento de fluidos viscosos en el espacio euclidiano 3D. Fueron desarrollados durante muchos años durante la primera mitad del siglo XIX por Navier y Stokes. En esencia, son una expresión de la segunda ley del movimiento de Newton; es decir, la fuerza es igual a la masa por la aceleración. El

Machine Translated by Google

Las ecuaciones incompresibles de Navier-Stokes se pueden derivar de la forma convectiva de la ecuación de impulso de Cauchy y tiene la forma

$$\cdot u = 0 \tag{6}$$

$$\partial u/\partial t + (u\cdot\ )u = -\rho \frac{1}{-} p + \mu\ 2u + F. \tag{7}$$

La ecuación (6) es una ecuación de continuidad que expresa la conservación del momento y se aplica a fluidos incompresibles. La solución de las ecuaciones da la velocidad del flujo u como un campo vectorial sobre el espacio 3D del fluido. El vector u da la dirección y magnitud de la velocidad del fluido en cada punto del espacio 3D del fluido en cada momento. $\partial u/\partial t$ es la aceleración del tiempo. $(u\cdot\ )u$ o $u\cdot(\ u)$ es aceleración espacial o convectiva (por ejemplo, remolinos y vórtices). La aceleración convectiva no es lineal y describe el cambio de velocidad u con la posición. Hace que las ecuaciones sean difíciles de resolver y es la causa predominante de turbulencia. u es el gradiente de velocidad. Es una matriz jacobiana o un tensor de segundo orden. Su componente simétrica $1/2(\ u + (\ u)^t)$ describe el estiramiento y el corte, mientras que su componente simétrico sesgado $1/2(\ u - (\ u)^t)$ describe la velocidad de rotación. ρ es la densidad del fluido, mientras que p es el gradiente de presión. μ 2u es la divergencia de difusión del esfuerzo cortante debido a la viscosidad, y μ es la viscosidad dinámica del fluido El tensor de esfuerzo cortante simétrico τ es igual a $\mu(\ u + (\ u)^t)$, y la divergencia de τ está dada por

$$2\mu\ \cdot\tau = 2\mu\ \cdot(1/2(\ u + (\ u)^t)) = \mu\ 2u = \mu\Delta u \tag{8}$$

donde $\Delta u = 2u$ es el laplaciano de u en el espacio euclidiano. F representa todas las fuerzas externas que actúan sobre el fluido, incluidas la gravedad y las fuerzas electromagnéticas.

Las ecuaciones de Navier-Stokes son difíciles, si no imposibles, de resolver analíticamente. Sin embargo , los estudios de simulación por computadora que utilizan supuestos simplificadores como el promedio temporal y espacial o el filtrado de paso de banda hacen un buen trabajo al hacer coincidir los flujos de fluidos simulados con los medidos experimentalmente [16].

3.4.2. Número de Reynolds El

número de Reynolds es un número adimensional definido por Osborne Reynolds (1842-1917) como

$$\mathrm{Re} = \frac{\text{fuerzas de inercia}}{\text{fuerzas viscosas}} = \frac{\rho u L}{\mu}$$

donde ρ = densidad, u = velocidad, L = longitud característica (por ejemplo, tamaño de las rocas en un arroyo de montaña) y μ = viscosidad dinámica.

El número de Reynolds juega un papel importante en la interpretación de las ecuaciones de Navier-Stokes y el teorema de De Rham porque a medida que Re aumenta experimentalmente de un número pequeño a un número grande, se observa que el flujo de fluido sufre una transición de flujo laminar a flujo turbulento. La transición se puede dividir en etapas según el tamaño del número de Reynolds. Cuando las fuerzas viscosas debidas a tensiones cortantes en el fluido dominan las fuerzas de inercia, el número de Reynolds es pequeño (por ejemplo, Re para miel ≈ 0,001). Cuando Re es pequeño, el flujo de fluido está dominado por la disipación de energía debido a las tensiones de corte viscosas y, en consecuencia, los remolinos y vórtices se amortiguan fuertemente. El resultado es un flujo laminar estable. A medida que aumenta Re , se observan ondas de presión 3D que fluyen hacia afuera entre las capas de la lámina. Esto aumenta la mezcla y aumenta la transferencia de energía-momento entre capas.
Para números de Reynolds grandes, dominan las fuerzas de inercia y la amortiguación viscosa es mínima.
El flujo turbulento resulta del nivel relativamente bajo de disipación de energía y amortiguación.
Cuando la inyección de energía cinética excede la disipación de energía, las inestabilidades en las interacciones convectivas no lineales crean una cascada de transferencia de energía de remolinos grandes a pequeños y vórtices.

Para números de Reynolds muy grandes, la divergencia de difusión del esfuerzo cortante debida a la viscosidad μ 2u en la ecuación de Navier-Stokes puede ignorarse de manera efectiva. Extremadamente turbulento

Machine Translated by Google

Los resultados son flujo, y la cascada de energía produce flujos rotacionales caóticos a pequeña escala de moléculas individuales. Ignorar μ 2u equivale a establecer el laplaciano del campo vectorial velocidad en el espacio euclidiano igual a cero (es decir, Δu = 2u = 0), donde 2u es igual al laplaciano Δu en el espacio euclidiano donde la curvatura de Ricci es cero . Según la teoría de cohomología de De Rham, dicho campo vectorial de velocidad es un campo de flujo armónico global asociado con la existencia de agujeros en la topología subyacente del espacio fluido euclidiano 3D.
Dado un amplio suministro de energía cinética, la existencia de agujeros (como partículas, rocas y burbujas) en el espacio de flujo del fluido se asocia con turbulencias extremas, debido a la falta de disipación de energía en el fluido y a la presencia de cinética rotacional caótica. energía a niveles moleculares.

4. Una propuesta teórica 4.1.

Células biológicas

Los procesos biológicos regulan la temperatura del cuerpo humano hasta 37 ◦C. Por tanto, existe un amplio suministro de energía cinética térmica para el movimiento de la solución acuosa dentro de las células biológicas. Las células biológicas están densamente repletas de orgánulos y otras moléculas. El líquido intracelular no puede fluir a través de estos orgánulos, por lo que tiene que circular alrededor de ellos. En otras palabras, el espacio 3D de la solución acuosa intracelular contiene una gran cantidad de orificios de diversas formas y tamaños. Como se describió anteriormente, el teorema de De Rham muestra que la presencia de tales agujeros en la topología subyacente del espacio 3D de un fluido tiene una influencia dramática en los campos de flujo vectorial y uniforme. Basándonos en la teoría matemática, proponemos que el movimiento térmico del fluido intracelular alrededor de los orgánulos crea una transferencia de energía cinética con una mínima disipación de energía desde remolinos y vórtices grandes a muy pequeños a nivel molecular, y que estas rotaciones se escalan adecuadamente para desempeñar un papel. en

4.2. Plegamiento de

proteínas El estudio del plegamiento de proteínas comenzó hace unos 60 años. Se sabe mucho, pero aún queda mucho por aprender [17-19]. Las proteínas son cadenas polipeptídicas de aminoácidos ensambladas por ribosomas dentro de las células biológicas. Cientos de miles de átomos pueden estar involucrados en largas cadenas peptídicas. El orden en fila única de las cadenas laterales de aminoácidos en cada cadena peptídica se especifica mediante codones dentro de las moléculas de ARNm. La secuencia de aminoácidos determina la estructura tridimensional en la que la cadena peptídica se pliega espontáneamente. El plegamiento ocurre en milisegundos y tiene lugar incluso mientras un ribosoma ensambla la cad
La conformación 3D nativa de la cadena peptídica es de vital importancia. Determina la función de la proteína. Las diferentes funciones de miles de proteínas diferentes proporcionan la maquinaria molecular de los sistemas vivos. Si las proteínas no se pliegan correctamente, entonces no funcionan correctamente, provocando muchos trastornos como la esclerosis lateral amiotrófica, la anemia falciforme, la enfermedad de Alzheimer y el síndrome de las vacas locas, por nombrar algunos. Pero, ¿cuáles son los mecanismos que controlan el plegamiento de las proteínas y qué causa que funcione mal?

Dado el número extremadamente grande de posibles ángulos de enlace entre los átomos y moléculas dentro de la cadena peptídica, el hecho de que el plegamiento correcto ocurra muy rápidamente indica que no se trata de un proceso aleatorio de prueba y error. Debe haber un proceso físico que dirija la vía de plegado. También hay una paradoja. La gran cantidad de configuraciones posibles de la cadena peptídica desplegada se pliega en una estructura tridimensional única de la proteína funcional. Esto viola la segunda ley de la termodinámica. ¡La entropía conformacional de la proteína disminuye a medida que su estructura plegada se vuelve más orden

La paradoja puede resolverse considerando la solución acuosa intracelular en la que están sumergidas las cadenas peptídicas. Se sabe que esto juega un papel crucial en el plegamiento de proteínas. Si una proteína plegada se sumerge en urea (que actúa como detergente), se despliega y se vuelve a plegar correctamente cuando se devuelve al líquido intracelular. Las concentraciones de iones, el pH y la temperatura del líquido intracelular deben regularse cuidadosamente a niveles apropiados mediante procesos biológicos. Si la regulación falla, las cadenas peptídicas ensambladas en ribosomas no se pliegan correctamente. Una teoría termodinámica predominante sobre el plegamiento de proteínas invoca el papel de la solución acuosa intracelular de la siguiente manera. Propone que la falta de polarización de

Machine Translated by Google

Las cadenas laterales de aminoácidos hidrófobos distribuidas a lo largo de la cadena peptídica hacen que las moléculas de agua que rodean estas cadenas laterales se estructuren, impulsadas por sus dipolos eléctricos y los enlaces de hidrógeno entre las moléculas de agua. Los gradientes termodinámicos de energía libre hacen que la cadena peptídica se pliegue para encerrar sus cadenas laterales hidrofóbicas dentro de su núcleo, ocultándolas así de las moléculas de agua circundantes. Esto libera las moléculas de agua de la estructura ordenada impuesta por las cadenas laterales hidrófobas, aumentando así la entropía del agua. El aumento de la entropía del agua excede la disminución de la entropía configuracional asociada con el plegamiento de proteínas, por lo que el plegamiento espontáneo es termodinámicamente factible.

Además de los gradientes termodinámicos de energía libre dentro del líquido intracelular, existen muchos otros campos de energía potencial químicos y eléctricos que rodean la cadena peptídica y que pueden influir en su plegamiento, incluidos los enlaces covalentes, iónicos y de hidrógeno, los gradientes de voltaje y las fuerzas de Van der Waals. entre los átomos de la proteína. Diferentes configuraciones de la cadena peptídica alinean sitios atractivos a lo largo de la cadena. Estos alineamientos permiten la formación de enlaces de hidrógeno, enlaces disulfuro, puentes salinos y otros. Los campos de energía potencial asociados impulsan el plegamiento de la proteína hacia la configuración energética mínima especificada por la secuencia de aminoácidos en la cadena peptídica. Sin embargo, no hay garantía de que un conjunto particular de enlaces moleculares y puentes salinos corresponda a la configuración funcional de energía mínima. Si se trata sólo de una configuración local de energía mínima, entonces se debe realizar trabajo sobre la molécula para sacar su configuración de la trampa de energía local y llevarla a un valle de energía más profundo.

4.3. La propuesta

No parece descabellado sugerir que el movimiento (movimiento browniano) de átomos y moléculas en cadenas peptídicas causado por el movimiento de la solución acuosa intracelular juega un papel importante en la cinemática del plegamiento de proteínas. El teorema de De Rham muestra que los agujeros en la topología subyacente del fluido intracelular están asociados con campos de flujo armónico global en el espacio 3D del fluido. El estudio experimental del flujo de fluidos junto con el número de Reynolds, las ecuaciones de Navier-Stokes y el teorema de De Rham proporcionan una conexión (consistente con la observación experimental) entre agujeros, flujo turbulento extremo y campos de flujo armónico. El término μ 2u = μ∆u en la ecuación de Navier-Stokes describe la divergencia de difusión del esfuerzo cortante debido a la viscosidad dinámica μ del fluido. Los estudios experimentales muestran que grandes proporciones de fuerzas inerciales a fuerzas viscosas (es decir, un número de Reynolds grande) están asociadas con un flujo turbulento extremo, y el término μ 2u = μ∆u en la ecuación de Navier-Stokes puede ignorarse efectivamente. Esto equivale a establecer el Laplaciano del campo de velocidad ∆u igual a cero. En geometría diferencial, así es exactamente como se definen los campos vectoriales armónicos y de forma 1 (es decir, en el espacio euclidiano ∆u = 2u = 0). En otras palabras, las observaciones experimentales del flujo de fluidos implican que los campos de flujo armónicos asociados con el flujo alrededor de orgánulos y otras moléculas (incluidas las moléculas de proteínas) corresponden a un flujo turbulento extremo. La existencia de agujeros en el espacio fluido 3D evita que los campos de flujo cerrados de forma unidimensional y vectorial sean globalmente exactos y los reemplaza con campos de flujo armónicos de una forma. En otras palabras, dado un amplio suministro de energía cinética, la teoría de De Rham proporciona una descripción formal del flujo de líquido intracelular responsable del movimiento browniano de moléculas en proteínas y otras estructuras dentro de las células biológicas.

Como se describe en la Sección 3.4, matemáticamente hablando, los campos de flujo armónicos asociados con huecos en la topología subyacente del espacio fluido se pueden calcular igualando el Laplaciano del campo de flujo a cero (es decir, ∆λ := dd λ + d dλ = 0). Pero en el espacio euclidiano, el laplaciano de un campo de flujo vectorial viene dado por ∆u = 2u, y de acuerdo con las ecuaciones de Navier-Stokes (6) y (7), establecer ∆u = 2u = 0 es equivalente a la ausencia de esfuerzo cortante viscoso y ausencia de disipación de energía en el fluido. Resultados de flujo extremadamente turbulento . Un amplio suministro de energía cinética térmica con una disipación de energía insignificante en el fluido provoca una cascada de transferencia de energía cinética desde grandes remolinos y vórtices a otros cada vez más pequeños hasta que finalmente una gran cantidad de energía cinética está disponible en el fluido.

Machine Translated by Google

la forma de caóticas rotaciones 3D de moléculas de agua individuales. Recordando que las moléculas de agua forman enlaces de hidrógeno entre sí en las proximidades de las cadenas laterales de aminoácidos hidrófobos y se adhieren a cadenas laterales de aminoácidos hidrófilos, parece razonable proponer que esta energía cinética está disponible para realizar trabajo (es decir, inyectar energía) en la superficie. átomos y moléculas de cadenas polipeptídicas de proteínas.

Los pares de rotación aplicados a las moléculas de proteínas tienen el nivel de energía y la escala adecuados para operar en los ángulos de enlace entre átomos y moléculas dentro de la cadena peptídica. Además, operan simultáneamente en todas las partes de la cadena peptídica y, cuando se combinan con gradientes de energía potencial termodinámicos, químicos y eléctricos que rodean la cadena peptídica, pueden desempeñar un papel importante en el plegamiento de proteínas. No sólo suministran la energía necesaria para romper con las trampas energéticas locales, sino que también pueden acelerar el plegamiento hacia configuraciones de energía mínima. Al igual que el "recocido" en los algoritmos de búsqueda de varianza mínima para redes neuronales, los caóticos campos de flujo armónico dentro de los fluidos intracelulares permiten el plegamiento para encontrar la configuración única de energía mínima absoluta de la proteína. Por supuesto, esto requiere el nivel adecuado de energía disponible regulada biológicamente. Si la temperatura corporal es demasiado baja, el flujo de líquido intracelular no tiene energía suficiente para romper las cadenas peptídicas y separarlas de las trampas de energía locales, y si la temperatura corporal es demasiado alta, puede romper enlaces y desnaturalizar las proteínas.

Una buena analogía para el movimiento browniano de las moléculas de proteínas y las cadenas laterales causado por el flujo turbulento de líquido intracelular alrededor de los orgánulos densamente empaquetados dentro de la célula viene dada por el movimiento caótico de las hojas y ramas de un árbol generado por el flujo del viento a través del árbol.

5. Discusión

En resumen, el teorema de De Rham muestra que los agujeros en la topología subyacente de una variedad suave están asociados con campos de forma diferencial cerrados en la variedad que no son exactos. Conocidos como campos de forma armónica, estos se pueden calcular en una variedad de Riemann estableciendo su Δ laplaciano de Hodge := dd + d d igual a cero, y en un espacio euclidiano estableciendo el Δ laplaciano = 2 igual a cero. La pregunta es: ¿existen en la naturaleza tales campos de formas armónicas y, de ser así, qué propiedades poseen?

Las matemáticas nos dicen que los campos de flujo integrales en una esfera, por ejemplo, serán muy diferentes de los campos de flujo en un toro, siendo la diferencia atribuible al agujero en el toro. Asimismo, para los sistemas físicos descritos en la Sección 3, las mismas matemáticas predicen con certeza que el comportamiento de los campos de forma vectorial y diferencial en sus variedades suaves debería verse profundamente alterado por la aparición de cualquier agujero o agujeros en la topología subyacente de la variedad. ¿Tenemos evidencia experimental de que este es el caso? ¿Y cuáles son las implicaciones para modelar estos sistemas?

5.1. Sistemas mecánicos En

la sección 3.2, demostramos matemáticamente que, según el teorema de De Rham, la presencia de huecos en el espacio de fase de momento de un sistema mecánico impide que la función hamiltoniana h y el campo vectorial hamiltoniano ξh se definan globalmente en el espacio de fase. Sin embargo, a pesar de que se trata de una predicción matemática definitiva, hasta que no se disponga de datos experimentales apropiados, sigue siendo incierto cómo podría comportarse exactamente un sistema físico-mecánico con agujeros en su espacio de fase de momento. Aún no se ha descubierto qué tipo de campo de flujo reemplazará al campo globalmente exacto de 1 forma $\lambda = C\xi h\ \omega$ en un sistema hamiltoniano. En otras palabras, el teorema de De Rham implica que los cambios en la topología del espacio de estados de un sistema mecánico requieren un modelo matemático diferente al de la mecánica hamiltoniana.

Este problema se aplica de manera similar a la mecánica lagrangiana. Mientras que la mecánica hamiltoniana se desarrolla en el fibrado cotangente (T Q) sobre la variedad de configuración (Q), la mecánica lagrangiana se desarrolla en el fibrado tangente (TQ) sobre la variedad de configuración (Q). La transformación de Legendre FL = TQ → T Q asigna puntos en TQ a puntos en T Q. Las mecánicas lagrangiana y hamiltoniana a menudo se describen como duales entre sí.

Machine Translated by Google

otra porque, si las ecuaciones son regulares en todas partes, ambas derivan la misma trayectoria de movimiento en la configuración múltiple Q. Sin embargo, la dualidad entre los sistemas lagrangianos y hamiltonianos requiere que la transformación de Legendre FL(. → pi sea una q suave uno a uno difeomorfismo, donde L(q, . q) es el lagrangiano igual a la diferencia i entre la energía cinética y la energía potencial, $q^{i}) = \frac{\partial L}{\partial \dot{q}^{i}}$

. es el i − ésimo componente de la velocidad vector q y pi es el i − ésimo componente del vector de momento p. La dualidad también requiere que tanto la segunda forma diferencial lagrangiana simpléctica en TQ como la segunda forma diferencial simpléctica hamiltoniana en T Q sean no degeneradas. Sin embargo, muchos sistemas mecánicos del mundo real no satisfacen estos requisitos. De hecho, es este hecho el que subyace a la importancia del teorema de Emmy Ntre; es decir, si un lagrangiano L(q, . q) es degenerado y no varía a lo largo de una coordenada q, entonces la componente asociada del momento pi se conserva. La eliminación de degeneraciones y singularidades de las variedades introduce agujeros que, como lo muestra la teoría de De Rham, cambian el comportamiento del sistema mecánico alejándolo del predicho por las mecánicas clásicas lagrangiana y hamiltoniana. Hasta donde sabemos, el modelado de sistemas mecánicos del mundo real para tener en cuenta los agujeros en el colector aún está por llegar. Es de esperar que esta exposición de los antecedentes teóricos pueda ayudar a estimular la recopilación de los datos experimentales necesarios.

5.2. Sistemas electrodinámicos

En la Sección 3.3, demostramos matemáticamente que la presencia de agujeros en el espacio-tiempo 4D de Minkowski impide que las ecuaciones sin fuente de Maxwell sean verdaderas. Se impide que el campo electromagnético cerrado F de 2 formas sea globalmente exacto, y el campo potencial electromagnético de 1 forma no se puede definir globalmente. En consecuencia, no pueden existir transformaciones de calibre globales del campo potencial, incluido el calibre de Lorentz. ¿Se reemplazará el campo electromagnético de 1 forma por un campo armónico de 1 forma y, de ser así, qué propiedades poseerá? Nuevamente, se requieren observaciones experimentales apropiadas de sistemas físicos electrodinámicos para responder estas preguntas.

De hecho, las implicaciones afectan a la naturaleza misma de la radiación electromagnética. El movimiento giratorio de las partículas cargadas en el sol crea corrientes eléctricas cambiantes J y densidades de carga cambiantes ρ. Pero, como lo muestran las ecuaciones de electromagnetismo de Maxwell, J y ρ son exactamente las fuentes de energía para los campos magnéticos B y los campos eléctricos E. En el vacío del espacio entre el Sol y la Tierra, J y ρ son ambos cero. Sin embargo, utilizando la curvatura de la relación curvatura y las relaciones div y curvatura entre E y B en las ecuaciones de Maxwell, incluso con J y ρ iguales a cero, los campos E y B acoplados satisfacen las ecuaciones de onda 2E − 1/c 2 ∂ 2E/∂t 2 y 2B − 1/c 2∂ 2B/∂t 2 En otras palabras, E y B forman ondas electromagnéticas acopladas que pueden viajar a la velocidad de la luz c a través del vacío a pesar de que tanto J como ρ son cero .

Pero en física quedan cuestiones sin resolver. Por analogía con las ondas en el agua y las ondas sonoras en el aire, se puede ver que las ondas requieren un medio en el que viajar. Pero ¿cuál es el medio en el que viaja una onda electromagnética en el vacío? A finales del siglo XIX, Lorentz propuso que algo llamado éter llenaba todo el espacio y que las ondas electromagnéticas viajaban a la velocidad de la luz en el éter. Sin embargo, el experimento del interferómetro de 1887 de Michelson y Morley demostró que no había diferencia en la velocidad de la luz proveniente de una estrella cuando la Tierra se movía hacia la estrella en comparación con cuando la Tierra se alejaba de la estrella. Esto desafió la idea de que tanto la luz como la Tierra se mueven en el espacio en relación con un éter fijo. Más recientemente, el enfoque matemático de la física de partículas conocido como teoría de calibre ofrece una explicación alternativa , a saber, que los propios campos eléctricos y magnéticos proporcionan el medio por el que viaja la luz. Sin embargo, como se describe en la sección 3.3, la teoría de las formas diferenciales de De Rham muestra que la propia teoría de calibre no se cumple en las proximidades de los agujeros de la variedad. Y esto es justo lo que la presencia de estrellas y planetas crea en el tejido del espacio-tiem

La vida en la Tierra depende fundamentalmente de la capacidad de las ondas electromagnéticas para transportar energía a través del espacio vacío. Maxwell se apresuró a afirmar que la luz es un fenómeno electromagnético.

Machine Translated by Google

ola. A través de la fotosíntesis, las plantas absorben energía de la luz que incide sobre ellas, transformándola en energía química almacenada en moléculas de azúcar. Esta energía química almacenada, a su vez, proporciona la mayor parte de la energía biológica necesaria para sustentar la vida en la Tierra. Por lo tanto, la comprensión de los campos electromagnéticos y las consecuencias de la aparición de agujeros en ellos es una tarea importante.

5.3. El sistema visual humano

Tanto para los sistemas mecánicos como para los electrodinámicos, las implicaciones del teorema de De Rham aún no se han abordado mediante experimentos. Sin embargo, para el sistema visual humano examinado en la Sección 3.1, la situación es algo diferente. En este caso, tenemos información extra relevante derivada de nuestra propia conciencia. Sabemos, por ejemplo, que hay un agujero en el egocentro de nuestro espacio visual binocular tridimensional percibido porque no podemos ver nuestra propia cabeza. También sabemos que el tamaño de una imagen visual percibida aumenta a medida que disminuye la distancia euclidiana entre el egocentro y el objeto del entorno . Esto es consistente con datos experimentales que muestran que el tamaño de una imagen en la retina cambia en proporción inversa a la distancia euclidiana entre el punto nodal del ojo y el objeto en el entorno [4]. Esto crea una singularidad en el egocentro del espacio visual 3D donde el tamaño de la imagen en la retina se vuelve infinito. Pero esta singularidad no es un problema, porque está contenida dentro del agujero alrededor del egocentro. El teorema de De Rham muestra que tal agujero debería evitar que los campos de flujo de imágenes percibidos sean globalmente exactos, reemplazándolos con campos de flujo armónicos.

La pregunta de si tales campos de flujo armónico existen o no en el sistema visual humano y qué propiedades poseen se puede responder hasta cierto punto a partir de la experiencia personal. A medida que un objeto del mundo exterior se acerca cada vez más al egocentro, podemos mantenerlo en el campo visual binocular haciendo converger los ojos. El tamaño percibido de la imagen se hace cada vez más grande y comienza a exceder el tamaño del espacio visual percibido. Con el tiempo, llegamos al límite de la visión bizca. La imagen visual percibida se desenfoca y se desintegra de forma caótica. Ya no se puede percibir como una imagen única coherente. Se vuelve físicamente imposible acercar el objeto al egocentro. Ha llegado al borde del agujero en el espacio visual binoc
Así, la experiencia personal muestra que los campos armónicos de flujo de imágenes visuales en el espacio visual tridimensional se desintegran de una manera caótica e impredecible en una región cercana al límite del agujero alrededor del egocentro; en otras palabras, se trata de un fenómeno de capa límite.

En este punto, es importante señalar que la experimentación en la ciencia visual humana nos dice que el espacio visual puede tener múltiples asignaciones. Argumentamos que el mapeo fundamental es el revelado por el sorprendente hecho de que hay personas que, como resultado de un daño cerebral, son visualmente ciegas pero, sin embargo, pueden moverse en su entorno 3D y realizar movimientos guiados visualmente con precisión, tal como si pudieran hacerlo. ver. A este fenómeno se le ha denominado visión ciega [20]. Demuestra que el cerebro tiene un sistema capaz de formar un espacio visual 3D egocéntrico subconsciente que tiene un mapeo uno a uno con el mundo exterior euclidiano. Este sistema visual subconsciente tiene que tener en cuenta la deformación de este espacio visual 3D, una deformación que se produce porque la detección de imágenes por parte de las retinas se produce en una superficie 2D. Más precisamente, la deformación se debe a que el tamaño de una imagen en las retinas 2D varía en proporción al ángulo subtendido en el punto nodal del ojo por el objeto correspondiente en el mundo exterior euclidiano 3D [4]. Como resultado de esta deformación, los cambios en el lugar de la cabeza en el entorno y los cambios en la postura del cuerpo cambian no sólo el tamaño calculado y la posición de las imágenes visuales codificadas de los objetos en el entorno y del cuerpo en ese entorno, pero también cambian el contorno, las oclusiones, la forma, las curvaturas, las velocidades relativas y las aceleraciones relativas de lo que se ve.

En otras palabras, existe un mapeo conforme entre el mundo exterior euclidiano y el espacio visual 3D subconsciente deformado. Un mapeo conforme preserva la forma infinitesimal de los objetos y los ángulos entre las coordenadas, pero no se preservan los tamaños de las imágenes codificadas. Las imágenes de objetos se reducen de tamaño a medida que aumenta la distancia relativa entre el egocentro y el objeto en el espacio euclidiano. Sin embargo, a pesar de esto

Machine Translated by Google

Al deformarse, el mapeo conforme con el mundo exterior es uno a uno, uniforme y uniforme, y la misma deformación se aplica a las imágenes visuales del entorno y a las imágenes visuales del cuerpo o partes del cuerpo en ese entorno. Por lo tanto, este espacio visual 3D deformado, subconsciente y egocéntrico se puede utilizar para la planificación y ejecución de movimientos guiados visual

Dicho esto, hay mucha evidencia de que las percepciones conscientes y subconscientes del mundo visual no son las mismas. Esto último puede estar mediado por el procesamiento cognitivo de arriba hacia abajo de imágenes visuales, que se considera basado en hipótesis estadísticas bayesianas derivadas de la experiencia a lo largo del tiempo. Esto introduce todo tipo de ilusiones en el mundo visual percibido conscientemente. Una de las favoritas es la ilusión de máscara cóncava. Debido a toda una vida de mirar caras de forma convexa, cuando una persona mira la superficie interior cóncava de una mascarilla, la ve conscientemente como convexa. Además, la ilusoria cara convexa parece girar y seguirlos mientras se mueven. Un espacio visual conscientemente percibido que contenga ilusiones no es lo que se necesita para planificar y controlar movimientos precisos guiados visualmente. El espacio visual 3D subconsciente, derivado directamente de señales aferentes y libre de intervención cognitiva estadística, es el más apropiado para el movimiento guiado visualmente [1–3].

El apoyo experimental a esta descripción del sistema visual humano proviene de un experimento realizado por primera vez por Taylor en 1941, pero repetido muchas veces desde entonces (ver [3]). En la oscuridad, una persona se lleva la mano a la cara. Luego, un breve destello de luz brillante quema (o blanquea) una imagen de la mano en las retinas. Esto da lugar a la percepción de una imagen residual de la mano que puede persistir durante varias decenas de segundos. Si, en la oscuridad, la persona acerca la mano a la cara o acerca la cara a la mano, sucede algo extraordinario. La imagen residual de la mano parece disminuir de tamaño a pesar de que la imagen blanqueada en las retinas permanece sin cambios. Esto es lo contrario de lo que suele ocurrir porque, a medida que la mano se acerca a la cara bajo la luz, parece aumentar de tamaño. Nuestra explicación es que el sistema nervioso puede transformar los movimientos del cuerpo sentidos propioceptivamente en cambios anticipados en su imagen visual [3]. Si la imagen visual anticipada se sustrae de la imagen visual real, la imagen corporal percibida no parece cambiar de tamaño sino sólo de posición en el espacio (la llamada constancia de tamaño). Durante el experimento de imagen residual, el movimiento de la mano genera el cambio anticipado en la imagen retiniana de la mano, pero la imagen retiniana real se graba en las retinas y, por lo tanto, no cambia. Restar el tamaño anticipado de la imagen de la imagen retiniana real genera un aumento ilusorio en el tamaño de la imagen residual.

Hemos construido una descripción de la geometría de Riemann de los cálculos necesarios para transformar los movimientos del cuerpo sentidos propioceptivamente en cambios anticipados en imágenes visuales tridimensionales del cuerpo [1–3]. Este no es un cálculo simple, porque la relación es multivariable y no lineal. Requiere 116 variables para especificar la configuración del cuerpo que se mueve en un entorno 3D local. Para cada lugar y postura del cuerpo existe un espacio visual deformado en 3D, y en cada punto de ese espacio visual existe una imagen visual codificada del entorno. Estas imágenes visuales cambian de un lugar y postura a otro. Geométricamente hablando, cada espacio visual 3D corresponde a una variedad riemanniana egocéntrica y curvada negativamente.

Esta relación entre movimiento y visión se complica por el hecho de que al espacio visual le falta la bola de puntos 3D alrededor del egocentro [1,2]. Este agujero intrínseco es lo que nos inspiró en primer lugar a considerar cómo la geometría diferencial y el teorema de De Rham en particular podrían iluminar la influencia de este agujero en el flujo percibido de imágenes visuales. La caótica descomposición de las imágenes visuales cerca del límite del agujero se describió anteriormente.

5.4. Flujo de fluido

Pasando a la Sección 3.4, nuestra mejor interpretación hasta la fecha de las propiedades de los campos de flujo armónico del mundo real proviene de los estudios experimentales del flujo de fluidos. Como se establece en esa sección, las ecuaciones de Navier-Stokes describen el flujo de fluido en el espacio euclidiano 3D donde el laplaciano del campo de flujo vectorial u está dado por $\Delta u = \quad 2u$. El vector armónic

Machine Translated by Google

El campo asociado con la presencia de agujeros (p. ej., partículas, rocas, burbujas, etc.) en el espacio euclidiano de un fluido se obtiene poniendo el Laplaciano a cero (es decir, ∆u = 2u = 0). Cuando esto se sustituye en la ecuación de Navier-Stokes, se obtiene una turbulencia extrema consistente con observaciones experimentales de turbulencia extrema asociada con números de Reynolds grandes. Las fuerzas de inercia ρuL, las velocidades, las presiones, las temperaturas y la energía cinética cambian de forma caótica en el tiempo y de un punto a otro en el espacio. A medida que el número de Reynolds de un fluido aumenta experimentalmente, las fuerzas viscosas que disipan energía dentro del fluido se vuelven insignificantes y el comportamiento del flujo del fluido se aproxima al de un campo de flujo armónico teórico. Debido a la falta de amortiguación viscosa dentro del fluido, se produce una cascada de transferencia de energía cinética de remolinos grandes a pequeños y se produce un campo de flujo caótico. Los vórtices rotacionales y la energía rotacional quedan disponibles a pequeña escala de moléculas individuales en el fluido. La presencia de tales remolinos de pequeña escala en un flujo de fluido turbulento extremo creado por la presencia de agujeros en el espacio del fluido sugiere un papel en el plegamiento de proteínas para la turbulencia extrema en el fluido intracelular de las células biológicas debido a los agujeros en el fluido creados por la Organelos densamente empaquetados dentro del líquido.

Las fuerzas de inercia que cambian de un punto a otro en el fluido dan lugar a una matriz simétrica de masa-inercia correspondiente a un tensor métrico de Riemann de energía cinética [21]. Esto brinda la oportunidad de mejorar las ecuaciones de Navier-Stokes. Una métrica de energía cinética de Riemann permite que el flujo de fluido se describa mediante campos vectoriales en una variedad de Riemann curva 3D en lugar de en el espacio euclidiano 3D plano de las ecuaciones de Navier-Stokes. Como se muestra en la Sección 3.4, el campo de flujo armónico de 1 forma λ en una variedad de Riemann se obtiene estableciendo el Laplaciano de la 1 forma λ en la variedad de Riemann 2 a cero (es decir, ∆λ = (− ∑)λ − Rico (λ) = 0). La curvatura de Ricci en cada punto de ξi la variedad de Riemann tiene en cuenta la dinámica no lineal atribuible a fuerzas de inercia que cambian de un punto a otro en el fluido. Estas son las fuerzas no lineales dentro del fluido que son predominantemente responsables del flujo turbulento. Por lo tanto, describir el flujo de fluido en una variedad de Riemann dotada con una métrica de Riemann apropiada dará una descripción más detallada del flujo de fluido que la proporcionada por las ecuaciones de Navier-Stokes.

Sin embargo, debido a la naturaleza caótica del flujo turbulento, se requieren métodos estocásticos para obtener estimaciones de una métrica de Riemann de energía cinética apropiada en función de la posición en la variedad. Parece necesario un modelo estocástico de series temporales de la cambiante métrica de Riemann .

Se han utilizado redes neuronales de aprendizaje profundo para modelar el flujo de fluidos turbulentos [16]. Sin embargo, a diferencia de una solución de “caja negra”, la geometría diferencial ofrece una explicación matemática más general del flujo turbulento que puede proporcionar una comprensión de los fenómenos físicos subyacentes.

5.5. Plegamiento de

proteínas Existe amplia evidencia experimental que muestra que la solución acuosa intracelular juega un papel importante en el plegamiento de proteínas. Sin embargo, la disponibilidad de datos que pueden usarse para verificar o refutar hipótesis de trabajo sobre la naturaleza precisa de las interacciones agua-proteína está limitada por las dificultades experimentales involucradas. Dicho esto, los resultados disponibles no respaldan la popular teoría hidrofóbica del plegamiento de proteínas. En cambio, los datos respaldan que el agua forma enlaces de hidrógeno entre puntos de la cadena peptídica y luego atrae esos puntos entre sí. El trabajo para medir el papel del agua en el plegamiento de proteínas mediante la medición de patrones de difracción de neutrones es prometedor, pero la técnica no es fácil y requiere un análisis estocástico de los datos medidos y el uso de simulaciones por computadora para encontrar el modelo que mejor se ajuste a los datos. Mientras tanto, la conclusión general de que el movimiento del agua desempeña un papel en el plegamiento de proteínas parece fuera de toda duda [19,22,23].

Nuestra propuesta sobre el movimiento browniano de los ángulos de enlace dentro de las cadenas peptídicas asociadas con la rotación turbulenta de moléculas de agua que fluyen alrededor de orgánulos dentro de la célula es una hipótesis novedosa. Se basa únicamente en las matemáticas del flujo de fluidos (es decir, las ecuaciones de Navier-Stokes) junto con el teorema de De Rham sobre la influencia de los agujeros en los campos de flujo. Está a la espera de pruebas experimentales que están más allá de los recursos del

Machine Translated by Google

autores presentes. Mientras tanto estamos convencidos de lo siguiente. El formalismo lógico y la certeza dentro de la formulación matemática expuesta aquí hacen que sea muy probable que el movimiento browniano observado experimentalmente en la solución acuosa intracelular durante el plegamiento de proteínas esté contribuyendo fundamentalmente al proceso de plegamiento. Esperamos que esto pueda proporcionar una guía para futuras investigaciones.

6. Conclusiones

Un objetivo de este artículo ha sido demostrar el valor de aplicar la geometría diferencial para comprender cualquier sistema físico donde pueden ocurrir agujeros en un campo de flujo circundante. Según la teoría de De Rham, los agujeros alteran el carácter de los campos circundantes, transformándolos en campos armónicos cerrados pero no exactos. Como resultado, el consiguiente flujo de un fluido se vuelve caótico. Hemos discutido este fenómeno en relación con la visión binocular 3D, que falla debido a un agujero en el campo de flujo visual alrededor del egocentro. Ahora señalamos que la presencia de un apantallamiento electromagnético alrededor de una fuente de radiación electromagnética, como un transmisor de televisión, crea un agujero en el espacio 3D circundante de campos electromagnéticos. Sugerimos que fue la creación de campos de flujo caóticos que subyacían al problema de los estudios de televisión [5,6] lo que en parte nos llevó a examinar esta teoría en primer lugar. Durante las investigaciones realizadas no se descubrió ninguna causa del cáncer , pero el estudio finalmente fue abandonado [24].

Para los sistemas mecánicos y electrodinámicos donde aparentemente faltan datos de flujo, esperamos que las ideas presentadas aquí sean de interés para quienes trabajan directamente en esos campos. En particular, somos optimistas de que nuestra propuesta sobre el plegamiento de proteínas pueda estimular el deseo de saber más sobre la geometría diferencial y De Rham utilizando la información del tutorial escrita aquí. En resumen, confiamos en que nuestra demostración del alcance del teorema de cohomología de De Rham en diversos sistemas del mundo real pueda ayudar a centrar la atención futura en general en esta extraordinaria pieza de las matemáticas. Después de todo, la noción de que las mismas matemáticas pueden aplicarse a una variedad de sistemas es lo que subyace a una gran cantidad de modelos y simulaciones, ya sea con variables mecánicas, eléctricas, químicas, económicas, fisiológicas, sociológicas o cosmológicas, etc. encontró un lugar clave en la física para quienes buscan comprender la naturaleza. Para aquellos en otros lugares que tienen inclinaciones matemáticas además de un punto de apoyo en un campo relevante, el trabajo involucrado para comprender y aplicar la geometría diferencial promete ser, en última instancia, mu

Contribuciones de los autores: Conceptualización, PDN; metodología, PDN; redacción: preparación del borrador original, PDN y MDN; redacción: revisión y edición, MDN y PDN Todos los autores han leído y aceptado la versión publicada del manuscrito.

Financiamiento: Esta investigación no recibió financiamiento externo.

Declaración de la Junta de Revisión Institucional: No aplicable.

Declaración de Consentimiento Informado: No aplicable.

Declaración de disponibilidad de datos: No se crearon ni analizaron nuevos datos en este estudio. El intercambio de datos no se aplica a este artículo.

Conflictos de intereses: Los autores declaran no tener conflictos de intereses.

Apéndice A

Como se indicó en la Introducción, se requiere un conocimiento sustancial de la teoría geométrica para apreciar exactamente cómo el teorema de De Rham vincula los agujeros en la topología de variedades suaves con el comportamiento caótico de los campos de flujo integrales en una varie Una comprensión completa del vínculo depende de la familiaridad con (i) complejos simpliciales, (ii) grupos de homología singulares, (iii) complejos de cadenas singulares, (iv) mapas de cadenas, (v) cálculo de grupos de homología singulares, (vi) el Teorema de Mayer-Vietoris, (vii) invariancia de homotopía, (viii) cohomología singular y el operador colímite, (ix) tensores simétricos sesgados, (x) álgebra de Grassmann y (xi) grupos de cohomología de Rham y la operación de la estrella de Hodge.

Machine Translated by Google

ator, así como conocimiento del operador laplaciano y campos de flujo armónico. Algunos de ellos han sido abordados en el texto. En este Apéndice A, proporcionamos una descripción más detallada de la geometría diferencial involucrada. Las pruebas están disponibles en una variedad de textos [7,11,12,25–30].

Apéndice A.1. Complejos simple

pegados a lo largo de sus caras límite para formar subespacios compactos conectados orientados de 0, 1, 2 y 3 dimensiones en R3. Estas estructuras pegadas se denominan complejos simpliciales, en los que se pueden formar formas diferenciales de 0, 1, 2 y 3er grado, respectivamente. estar integrado. Un estándar 3-simplex ∆3 = (e0,e1,e2,e3) en R3 se define de manera que e0 = (0, 0, 0), e1 = (1, 0, 0), e2 = (0, 1, 0) y e3 = (0, 0, 1).

Este ∆3 estándar 3-símplex define un sistema de coordenadas cartesianas en R3 donde e0 define el origen y e1,e2,e3 definen vectores de longitud unitaria a lo largo de cartesianos mutuamente ortogonales, respectivamente. incluye todos los $^{1}, ^{2x}, ^{3x}$ Un 2-simplex (P0, P1, P2) en R3 define un triángulo en coordenadas 3D x espacio que puntos a lo largo de sus caras límite orientadas (P1, P2), (P0, P2), (P0, P3) también. como todos los puntos internos. Dados los vectores (v0, v1, v2) desde el origen e0 hasta cada uno de los puntos P0, P1 y P2, respectivamente, si los vectores v1 − v0 y v2 − v0 son linealmente independientes, decimos que los puntos P0, P1 y P2 son afínmente independientes (a diferencia de un espacio vectorial, un espacio afín no abarca el origen). A medida que el vector v0 se mueve en R3, el 2-símplex afínmente independiente (P0, P1, P2) se puede mover a cualquier lugar en R3 . Al pegar una colección de 2-símplices a lo largo de caras límite comunes, se puede construir una superficie orientada compacta bidimensional conocida como complejo simplicial en R3 como una cadena de 2, denotada por c2. Se pueden construir muchas 2 cadenas . Esta superficie se denomina diferentes en R3 . En R3 se pueden construir muchas cadenas 0 (puntos discretos), cadenas 1 (modelos de estructura de alambre), cadenas 2 (superficies bidimensionales) y cadenas 3 (subespacios tridimensionales). En general, los complejos simpliciales son poliedros cerrados y acotados.

Dada una variedad suave 3D M, un 2-simplex (P0, P1, P2) en R3 se puede mapear mediante un mapa suave σ2 en M. Tal mapa suave es un mapa no lineal que da lugar a un 2-simplex singular en M. Debido al mapeo no lineal, es posible que un 2-símplex singular ya no parezca un triángulo. Puede perder rango e incluso puede aparecer como un solo punto. Por esta razón se le conoce como simplex singular . Lo mismo se aplica a uno, dos y tres simples simples en variedades 3D.

De todas las posibles 2 cadenas c2 que se pueden construir en R3 , hay algunas 2 cadenas que se cierran sobre sí mismas (como 2 esferas) y tienen límites vacíos (es decir, ∂c2 = 0). Estas 2 cadenas con límites vacíos se denominan 2 ciclos y se representan mediante la notación z2. Una propiedad similar se aplica a las cadenas 0, 1, 2 y 3 en R3 , dando 0, 1, 2 y 3 ciclos con límites cero.

Apéndice A.2. Grupos de homología singulares

Se pueden utilizar 2 ciclos singulares en la topología subyacente X de M para detectar la existencia de agujeros bidimensionales en M. El grupo singular de 2 homologías en la topología subyacente de M se define como todos los 2 ciclos (representados por el grupo abeliano libre Z2(M)) en la topología subyacente de M que no son límites de 3 cadenas c3. Estos 2 ciclos encierran agujeros bidimensionales en la topología subyacente. Defina B2(M) como todos los 2 ciclos en Z2(M) que son límites de 3 cadenas c3. El grupo singular de 2 homologías se define entonces por el cociente H2(M) = Z2(M)/B2(M) (es decir, el grupo de todos los 2 ciclos que no son límites de 3 cadenas). La dimensión del grupo abeliano libre H2(M) es igual al número de huecos bidimensionales en la topología subyacente X de M. Esto puede generalizarse a grupos de homología de r-ésimo grado Hr(M) = Zr(M)/Br(M) con r = 0, 1, · · · n en la topología subyacente X de una variedad uniforme de n dimensiones M. Esta descripción intuitiva de grupos de homología singulares se puede formalizar definiendo una secuencia de grupos de homología singulares adjuntos a la topología subyacente. de colectores lisos.

Machine Translated by Google

Apéndice A.3. Complejos de cadenas singulares

Una larga secuencia de grupos abelianos conectados por homomorfismos

$$\alpha_{r+1} \cdots \to G_{r+1} \to G_r \xrightarrow{\alpha_r} G_{r-1} \to \cdots$$

se dice que es exacto si im $\alpha_{r+1} = \ker \alpha_r$ para todo r.

El primer teorema de isomorfismo para grupos implica que una secuencia corta de 5 términos

$$0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$$

es exacta si y sólo si el homomorfismo $\alpha : A \to B$ es inyectivo y el homomorfismo $\beta : B \to C$ es sobreyectivo. Esto implica que para una secuencia corta de 4 términos

$$0 \to A \xrightarrow{\alpha} B \to 0$$

para ser exactos, el homomorfismo $\alpha : A \to B$ tiene que ser tanto inyectivo como sobreyectivo. Este hecho es útil cuando se calculan grupos de homología singulares de espacios topológicos de dimensiones superiores (por ejemplo, esferas de altas dimensiones que se describen a continuación).

Una larga secuencia de grupos (cadenas) abelianos libres junto con homomorfismos

$$\cdot \to C_{r+1} \to C_r \xrightarrow{\partial_{r+1}\ \partial_r} C_{r-1} \to \cdots$$

se llama complejo de cadena singular, denotado por C , si la composición de dos homomorfismos consecutivos cualesquiera $\partial_r \circ \partial_{r+1}$ es cero. Esto es lo mismo que requerir im $\partial_{r+1} = \ker \partial_r$.

El r-ésimo grupo de homología singular del complejo de cadena singular C es

$$H_r(C) = \ker \partial_r / \text{im}\, \partial_{r+1} = Z_r(X)/B_r(X)$$

donde $Z_r(X)$ es el espacio de todos los r-ciclos en X, y $B_r(X)$ es el espacio de todos los r-ciclos en X que son límites de (r + 1)-cadenas C_{r+1} en X. Se sigue que el complejo de cadena singular C es exacto si y sólo si $H_r(C)$ es cero para todos los valores de r. En otras palabras, los grupos de homología singulares $H_r(C)$ proporcionan una medida precisa de la falla del complejo de cadena C para ser exacto. Esto ocurre cuando hay un agujero de dimensión r en cualquiera de las cadenas r.

Apéndice A.4. Mapas en cadena F#

Un mapa de cadenas $F_\# : C(X) \to D(Y)$ entre complejos de cadenas singulares asociados con espacios topológicos X e Y, respectivamente, es una colección de homomorfismos $F_\# : C_r \to D_r$ para todo r tal que $F_\# \circ \partial_r = \partial_r \circ F_\#$, como se ilustra a continuación:

$$\begin{array}{ccccccc} X \to C(X) & \cdots & \xrightarrow{\partial_{r+1}} & C_r(X) & \xrightarrow{\partial_r} & C_{r-1}(X) & \to \cdots \\ \downarrow F_\# & & & \downarrow F_\# & & \downarrow F_\# & \\ Y \to D(Y) & \cdots & \to & D_r(Y) & \to & D_{r-1}(Y) & \to \cdots \end{array}.$$

Un mapa continuo $F : X \to Y$ entre espacios topológicos X e Y induce un mapa de cadena singular $F_\# : C(X) \to D(Y)$ entre el complejo de cadena singular $C(X)$ y el complejo de cadena singular $D(Y)$. Este mapa está definido por los homomorfismos $F_\# : C_r(X) \to D_r(Y)$ para todos los valores de r, como se ilustra. Dado que $F_\#$ asigna $\ker \partial_r$ a $\ker \partial_r$ y im ∂_{r+1} a im ∂_{r+1} , se deduce que el mapeo continuo $F : X \to Y$ entre espacios topológicos X e Y induce un homomorfismo de grupo de homología singular $F_* : H_r(C(X)) \to H_r(D(Y))$ para todos los valores de r.

Debido a que la imagen de cualquier simplex singular es conexa, tiene que estar completamente en un componente conexo por trayectoria del espacio topológico X. Por lo tanto, se deduce que el r-ésimo grupo de homología singular $H_r(X)$ para X es igual a la suma directa de los r-ésimos grupos de homología singulares $H_r(X_\alpha)$ para cada uno de los componentes X_α conectados por ruta (es decir, $\alpha_* : H_r(X_\alpha) \to H_r(X)$).

Machine Translated by Google

Apéndice A.5. Computación de grupos de homología singulares

La definición dada anteriormente de grupos de homología singulares implica la noción de complejos de cadenas r que consisten en todas las cadenas r posibles que pueden construirse en un espacio topológico. Se trata de un número muy grande de cadenas r, por lo que no ayuda mucho a decirnos cómo podemos calcular realmente grupos de homología singulares. Sin embargo, se han desarrollado poderosas herramientas basadas en la llamada secuencia exacta de Mayer-Vietoris y el lema en zigzag para calcular los grupos de homología para la mayoría, si no para todos, los espacios topológicos. Algunos grupos de homología singulares se pueden deducir directamente de la definición, en particular (i) grupos de homología singulares de grado cero H0(X) para cualquier espacio topológico X y (ii) todos los grupos de homología singulares Hr(X) para todos los valores de r en un espacio topológico discreto X.

Apéndice A.5.1. Grupos de homología singular de grado cero

Para cada espacio topológico X, el límite ∂C0 del complejo de cadena de dimensión cero C0 es cero porque no existe una cadena C−1 . De ello se deduce que cada cadena 0 singular es un ciclo 0 (es decir, C0(X) = Z0(X)). Por otro lado, una cadena 0 singular es una combinación lineal formal de puntos xi en X con coeficientes enteros (es decir, c0 = ∑i aixi). Suponiendo que X está conexo por caminos, podemos definir un mapa sobreyectivo : C0(X) → Z por : ∑i aixi → ∑i ai Z, donde Z es el grupo de todos los números enteros. El núcleo de este mapa corresponde a todos los puntos B0(X) en X que son puntos límite de una cadena 1 y, en consecuencia, según el primer teorema de isomorfismo para grupos,

$$Z0(X)/\ker = Z0(X)/B0(X) = H0(X).$$

Así, para cualquier espacio topológico X, el espacio de homología singular cero H0(X) es isomorfo al grupo de números enteros Z, y la dimensión de H0(X) es igual al número de componentes conectados por caminos en X. Esto La dimensión es el número de Betti b0 .

Apéndice A.5.2. Grupos de homología para espacios topológicos discretos

Para un espacio discreto X, cada punto aislado corresponde a una componente conexa de X, por lo que H0(X) es un grupo abeliano libre isomorfo al grupo de números enteros Z con dimensión b0 igual al número de puntos discretos en X. Aunque cada componente conectado de X es un único punto aislado, el límite ∂σr para un r-símplex singular σr está dado por (−1) rσr−1, donde σr ◦ Fi,r suma alterna ∂σr = ∑ singular $i_{yo=0}^{r}$ (−1) rσr ◦ Fi,r = ∑ $_{yo=0}^{r}$ representa el ésima cara límite. Para espacios de un solo punto, las caras singulares σr ◦ Fi,r son todas iguales a σr−1. Por tanto, ∂ : Cr → Cr−1 es un isomorfismo cuando r es par y cero cuando r es impar.
En consecuencia, obtenemos un complejo de cadena singular.

$$\cdots \to C3 \to C2 \to C1 \to C0 \to 0$$

donde los mapas de límites ∂ : Cr → Cr−1 son mapas 0 para valores impares de r y son isomorfismos para valores pares de r. De ello se deduce que la secuencia es exacta en cada grupo abeliano Cr excepto C0. Pero ya hemos tratado con H0(X) para C0. Dado que los grupos de homología singular Hr(X) proporcionan una medida de la desviación de la exactitud, se deduce que el grupo de homología singular Hr(X) para cada punto aislado de un espacio topológico discreto X es igual a cero para todo r > 0 .

Apéndice A.6. Teorema de Mayer-Vietoris

El teorema de Mayer-Vietoris es una herramienta importante para calcular grupos de homología singulares para espacios topológicos. Si X es un espacio topológico y U, V son subconjuntos abiertos de X cuya unión es X, entonces para cada r existe un homomorfismo llamado homomorfismo de conexión ∂ : Hr(X) → Hr−1(U ∩ V) , de modo que la siguiente secuencia larga de grupos de homología sea exacta:

$$\cdots \to \text{Hora}(U \cap V) \xrightarrow{yo} {}^{j}\text{Hr}(U) \quad \text{Hr}(V)\ k \quad \to -I \quad \text{Hr}(X) \xrightarrow{\partial} \quad \text{Hr}-1(U \cap V) \to \cdots$$

Machine Translated by Google

donde i j : Hr(U ∩ V) → Hr(U) Hr(V) con (i k − l : Hr(U) j)[c] = (i [c], j [c]) y] donde Hr(V) → Hr(X) con (k − l)([c], [c [c ′]) = k [c] − l [c ′[c] y ′] representan clases de homología singulares [8].

Para ilustrar cómo se puede combinar el teorema de Mayer-Vietoris con otras propiedades de los espacios topológicos para calcular grupos de homología singulares, considere calcular los grupos de homología singulares para todos los valores de r y para todas las dimensiones n de una esfera unitaria. Considere una unidad n-esfera S n con un subconjunto abierto U igual a S n sin el polo sur y un subconjunto abierto V igual a S n sin el polo norte. Una parte de la secuencia exacta de Mayer-Vietoris es

$$\text{Hora(U)} \quad \text{Hr(V)} \to \text{Hr(S n)} \to \overset{\partial}{} \quad \text{Hr−1(U} \cap \text{V)} \to \text{Hr−1(U)} \quad \text{Hr−1(V)}$$

donde el homomorfismo conector es ∂ : Hr(S n) → Hr−1(U ∩ V). Todos los puntos en U se pueden contraer hacia el polo norte y todos los puntos en V se pueden contraer hacia el polo sur. Esto significa que tanto U como V son contráctiles y, por lo tanto, simplemente están conectados (es decir, no contienen agujeros). Resulta que

$$\text{Hora(U)} \quad \text{Hr(V)} = 0 \text{ y Hr−1(U)} \quad \text{Hr−1(V)} = 0$$

En consecuencia, cuando r > 1 y n ≥ 1, la secuencia exacta se puede reescribir como

$$0 \to \text{Hr(S n)} \to \overset{\partial}{} \text{Hr−1(U} \cap \text{V)} \to 0.$$

Esta es una secuencia exacta de 4 términos como se describió anteriormente y, en consecuencia, el homomorfismo de conexión ∂ : Hr(S n) → Hr−1(U ∩ V) es un mapa biyectivo (es decir, un isomorfismo). Así como el ecuador de una unidad de 2 esferas S, la esfera ecuatorial (U ∩ V) de una unidad de n-esfera S n esfera S^2 es topológicamente equivalente a un círculo unitario S^1 S n−1 . Esto significa Hr(S n) = Hr−1(U ∩ V) = Hr−1(S es topológicamente equivalente a una unidad (n − 1)-) Hr(S n) = 0 si 0 < r < n. Ya sabemos que H0(S n) = Z y Hr(S n) = 0 norte-1 para r > 1 y n ≥ 1 y si r > n, entonces tenemos todos los grupos de homología singulares para esferas de dimensiones superiores.

Apéndice A.7. Invariancia de homotopía

Una T de 2 toros^2 tiene dos agujeros unidimensionales y un agujero bidimensional. En otros) y uno-
En palabras, tiene un grupo de homología singular de grado uno bidimensional H1(T grupo ^{2}
de homología singular de grado dos H2(T del toro 2 permite que se ^{2}). La geometría de la lámina de caucho.
deforme en muchos tamaños y formas diferentes. Siempre que no sea Cortados, rasgados, perforados o
pegados, todos los diferentes tamaños y formas de T se mapean homeomórficamente entre sí y la ^{2} poder
estructura topológica permanece sin cambios (es decir, el número de agujeros y por lo tanto los grupos de
homología singulares H1(T y H2(T número de los huecos de varias dimensiones, y por lo tanto sus ^{2})
grupos de ^{2}) permanecen sin cambios). Una propiedad similar se cumple para todos los espacios topológicos. El
homología singulares, son homotópica y topológicamente invariantes. Sean f0, f1 : X → Y aplicaciones
homotópicas continuas (es decir, f0(X) en Y y f1(X) en Y pueden transformarse en. entre sí sin cambiar su
estructura topológica), entonces f0(X) y f1(X) (junto con todas las formas intermedias a lo largo de la ruta
de transformación de homotopía entre f0(X) y f1(X)) tienen los mismos grupos de homología singulares.
En otras palabras, los mapas de homotopía continua f0, f1 : X → Y inducen homomorfismos de grupos de
homología singulares (f0) ,(f1) : Hr(X) → Hr(Y) para todos los valores de r. Se dice que un mapa
continuo F: X → Y entre espacios topológicos X e Y es una equivalencia de homotopía si hay un mapa de
retroceso continuo G: Y → X tal que la composición F ◦ G es homotópica al mapa de identidad IY en Y y la
composición G ◦ F es homotópica al mapa de identidad IX en X. Si F: X → Y es una equivalencia de
homotopía entre los espacios topológicos X e Y, entonces para cada r ≥ 0, el mapa F : Hr(X) → Hr(Y) es
un isomorfismo.

Machine Translated by Google

Apéndice A.8. Cohomología singular

Los grupos de homología singular son functores covariantes de la categoría de espacios topológicos a la categoría de grupos abelianos libres. Para muchas aplicaciones, incluida la teoría de formas diferenciales de De Rham , es más útil tener functores contravariantes. La cohomología singular es una versión retraída de la homología singular. No proporciona ninguna información nueva , simplemente presenta la misma información de forma diferente.

La transformación de funtor covariante a functor contravariante se basa en el siguiente hecho. Si A y B son grupos abelianos bajo suma puntual, entonces el conjunto Hom(A, B) de homomorfismos de grupo de A a B es en sí mismo un grupo abeliano bajo suma puntual. Para cualquier grupo abeliano fijo G, cada homomorfismo f# : A → B da Hom(B, G)) = φ A. Este mapa de retroceso se puede utilizar para relacionar los grupos libres un homomorfismo de retroceso inducido f y φ grupo$^{\#}$: Hom(B, G) → Hom(A, G). Si φ $_{B}$ abeliano^{A} Hom(A, G), entonces f$^{\#}$ (φ ^{B} de todas las r-cadenas singulares Cr(X) en un espacio topológico X al grupo abeliano libre de todas las r-cocadenas singulares C r (X; G) = Hom(Cr(X), G) en el mismo espacio topológico X. Para grupos abelianos libres, el grupo G es igual al grupo de números enteros Z. Un mapa continuo f : X → Y entre espacios topológicos X e Y induce un homomorfismo f# : Cr(X) → Cr(Y) entre los grupos abelianos libres de todas las posibles cadenas r en X e Y (es decir, c ^{Y} $_{r}$ = f#c $_{r}$ donde nosotros Y usa la notación c $_{r}$ para representar una única cadena r en Y). También induce un homomorfismo de retroceso f : C r (Y; G$^{\#}$) → C r (X; G), o equivalentemente, f : Hom(Cr(Y), G) → Hom(Cr(X), G) : φ Y(f#c X r) → (f #φ Y)(c X r) (es decir, f$^{\#}$ donde φ Hom(Cr(Y), G)).

Operador colímite

El operador de frontera ∂ : Cr+1(X) → Cr(X) entre grupos abelianos libres de todas las cadenas (r + 1) posibles y todas las cadenas r posibles en X induce un operador colímite de retroceso δ : C r (X; GRAMO) → C r+1(X; G) entre todas las r-cocadenas posibles y todas las posibles (r + 1)-cocadenas en X (observe los valores crecientes de r). Para el colímite de una sola cadena r c # δφY)(c X r), donde $_{r}$ simplemente, esto $^{\#}$: (δφY)(f#c X (δφX)(c X r) = (φ X)(∂c X r) → (f en X, obtenemos f r). Más . muestra que f (δδφX)(c X r) = (φ $^{\#}$ ∘ δ = δ ∘ f r) y $^{\#}$ También se deduce que δ ∘ δ = δδ = 0, porque X)(∂∂c X complejo y mapa de ∂∂ = 0. Como δδ = 0, existe una cocadena exactamente larga cocadenas f $^{\#}$ entre los espacios topológicos X e Y (observe la disminución de r).

$$\begin{array}{ccccccc} r-1 & & \overset{\delta}{\longrightarrow} & & \overset{\delta}{\longrightarrow} & & \\ \cdots \to C^{r+1}(X;G) & \to & C^{r}(X;G) & \to & C^{r-1}(X;G) & \to & \cdots \\ \downarrow f^{\#} & & \downarrow f^{\#} & & \downarrow f^{\#} & & \\ \cdots \to D^{r+1}(Y;G) & \xrightarrow{\delta} & D^{r}(Y;G) & \to & D^{r-1} & \to & \cdots \end{array}$$

Una r-cocadena φ $_{C}$ r (X; G) es un r-cociclo φ $_{X}$ $_{Z}$ r (X; G) si δφX = 0 (es decir, tiene un r (X; colímite vacío), y es un cociclo acotado φ que G) si existe un cociclo B (X; G) tal $_{C}$ $^{r-1}$ δψX = 0 (es decir, es cofrontera de un (r − 1)-cocadena). Tanto los ψ r-cociclos Z r (X; G) como los r-cociclos delimitadores B r (X; G) se conservan mediante mapas de cocadenas f, por lo que podemos definir el r-ésimo$^{\#}$ grupo de cohomología singular. de X con coeficientes en G sea el cociente Hr (X; G) = Zr (X; G)/B r (X; G).

Se puede concluir que un mapa continuo f : X → Y entre espacios topológicos X e Y induce un homomorfismo de retroceso f r (Y; G) → C r (X; G) entre r-ésimas cocadenas$^{\#}$ singulares : C en Y y r-ésimas cocadenas singulares X. Esto a su vez induce un retroceso homomor- : Hr (Y; G) → Hr (X; G) entre el r-ésimo grupo de cohomología los grupos singular Hr (Y; G) fismo f en Y y el r-ésimo grupo de cohomología singular Hr (X; G) en X. Así como de homología singulares son homotópica y topológicamente invariantes, también lo son los grupos de cohomología singulares.

En otras palabras, los grupos de cohomología singulares son isomórficos a sus correspondientes grupos de homología singulares.

Apéndice A.9. Tensores simétricos sesgados en una variedad n suave M

En cada punto p en una variedad n suave M, existe un espacio tangente de n dimensiones TpM y un espacio cotangente de n dimensiones T $_{pag}$ M. Un covector o λ de 1 forma en T p $_{METRO}$

Machine Translated by Google

actúa sobre un vector v en TpM y lo transforma en un número real (es decir, λ(v) = λ, v → R). Un tensor asimétrico de tipo (0, 2) T(X, Y) que opera sobre dos vectores X e Y en TpM para dar un número real cambia de signo si el orden de los dos vectores se invierte (es decir, T(X, Y) = −T(Y, X)). Si (e1, · · · ,en) son vectores de base unitaria que abarcan el vector tangente n) son espacio TpM en p M y (ε el ^{1}, · · · ,ε covectores de base unitaria duales a (e1, · · · ,en) que espacio vectorial cotangente T abarcan M, entonces podemos escribir el sesgo de tipo (0, 2) -tensor simét T(X, Y) en la forma T = Tj1 j2 $\varepsilon^{j1\ j2}$, dónde

$$\varepsilon^{j1\ j2} = (\varepsilon j1 \quad \varepsilon j2 - \varepsilon j2 \quad \varepsilon j1) = \varepsilon\ j1 \quad \varepsilon\ j2 \text{ con } j1 < j2$$

Esto define el producto de cuña ε j1 ε j2 como base para el tipo (0, 2) sesgado-simétrico espacio tensor Ω2 en p M, y Tj1 j2 es el componente del tensor (vector) j1 ε j2 de tipo (0, 2) simétrico proyectado sobre el vector base ε simetrización (o alterno) sesgado . Esto se puede generalizar a un sesgo -n operador A para un tipo (0, s), s = 0, 1, · · · tensor T por la fórmula , sesgo-simétrico

$$AT(X1, \cdots, Xs) = (-1)^{\kappa} \frac{1}{s!} \sum T(Xa1, \cdots, Xas)$$

donde esta! es el número de permutaciones P(n,r) = es la de n objetos tomados r a la vez (n−r)! cuando el orden importa y con n = s y r = s en este caso, suma de todos los s! permutaciones de los s vectores (X1, · · · , Xs) , y κ es la paridad de cada permutación (es decir, κ = 0 si el número de intercambios de dos vectores es par y κ = 1 si el número de intercambios es impar).

Si Φ es una forma r1 y Π es una forma r2, entonces Φ Π = A(Φ Π) es una forma (r1 + r2).

El producto de cuña es (i)

asociativo Φ (Π Ψ) = (Φ Π) Ψ, (ii) (anti)-
conmutativo (Φ Π = (−1) r1 r2Π Φ), y (iii) distributivo (Φ + Ψ)
Π = Φ Π + Ψ Π donde Φ y Ψ son ambas formas r1 aquí.

Los vectores base (tensores simétricos sesgados) que abarcan el r-ésimo tensor simétrico sesgado j1 ···jr espacio Ωr (M) son ε = ε
j1 · · · ε jr : j1 < · · · < jr .

Apéndice A.10. Álgebra de Grassmann En cada p

M en una n-variedad suave M, existe una familia de espacios tensoriales simétricos sesgados Ω0 Ωn . Los tensores simétricos sesgados en estos espacios tensoriales Ω1 Ω2 simétricos sesgados se denominan formas diferenciales. Una forma r en Ωr actúa sobre r vectores v1, · · · , vr en TpM y los transforma en un número real para todos los valores de r = 0, · · · n. La dimensión de cada espacio tensor simétrico sesgado Ωr viene dada por r!(n−r)! . Según esta fórmula, los espacios tensoriales asimétricos Ω0 y Ωn son ambos unidimensionales. La dimensión aumenta linealmente con r desde Ω0 hasta Ω. La dimensión de Ωr es igual a la dimensión de Ωn−r . De hecho, el espacio tensor simétrico sesgado Ωr es isomorfo al espacio tensor simétrico sesgado Ωn−r. Un operador isomorfo conocido como operador estrella y luego disminuye linealmente a Ωn . Esto implica que el de Hodge transforma las formas r en Ωr en formas (n − r) en Ωn−r para todos los valores de r. Para r < 0 y r > n, los espacios tensoriales simétricos sesgados Ωr son espacios vacíos y las formas diferenciales se establecen en cero. Un operador d, conocido como derivada exterior, transforma una forma r en una forma (r + 1). Es importante destacar que la composición d ◦ d = dd es siempre cero. De ello se deduce que existe una secuencia de espacios tensoriales simétricos sesgados conectados por el operador diferencial exterior

$$0 \xrightarrow{d} \Omega 0 \quad d \to \cdots \xrightarrow{d} \Omega n \to 0$$

donde cada flecha corresponde a un diferencial exterior d. Un operador d transforma una forma r en una forma (r − 1). Más generalmente, d = (-1) d , donde es el Hodge

Machine Translated by Google

operador estrella, es decir, un isomorfismo entre Ωr y Ωn−r . Los operadores d y d no son son inversos entre sí, son adjuntos. La suma directa de espacios tensoriales asimétricos forma un álgebra graduada Ω = Ω0 · · · Ωn conocida como álgebra exterior o álgebra de Grassmann. Cualquier forma r ω para la cual dω = 0 se dice que es una forma r cerrada, mientras que cualquier forma r ω igual a dλ donde λ es una forma (r − 1) se dice que es una forma r exacta.

Apéndice A.11. Grupos de cohomología de De Rham

Del álgebra de Grassmann y del hecho de que d ◦ d = dd = 0, se deduce que toda forma diferencial exacta es cerrada. Sin embargo, no se sigue de ello que toda forma diferencial cerrada sea exacta. La teoría de las formas diferenciales de De Rham deja claro que el hecho de que una forma diferencial cerrada sea exacta o no depende únicamente de la estructura topológica subyacente de la variedad. Específicamente, depende de la presencia o ausencia de agujeros de varias dimensiones en la topología subyacente. Como lo ilustra la cantidad de teoría resumida en este apéndice, se necesita una cantidad significativa de geometría diferencial para apreciar este hecho notable. Debido a que el diferencial exterior d : Ωr → Ωr+1 es lineal, su núcleo ker(dr) = Z r (M) = formas r cerradas y su imagen im(dr−1) = B r (M) = r- exacto las formas son espacios vectoriales lineales. Dado que toda forma exacta es cerrada, se deduce que las formas r exactas de Rham es un subespacio de las son r (M) Z r (M)). Por lo tanto, la cohomología r-ésima formas r cerradas (el grupo B se puede definir como el espacio vectorial cociente

$$h_{r\,dR}(M) = {}^{Z_r}(M)/B_r(M)$$

(es decir, el espacio de todas las formas cerradas que no son exactas).

Si bien Hr dR(M) es un espacio vectorial real, también es un grupo de Lie bajo suma. Dado que Ωr (M) = 0 para r < 0 y r > n, se deduce que Hr dR(M) es igual a cero para estos valores de r. También se deduce que Hr dR(M) = 0 para 0 ≤ r ≤ n si y sólo si toda forma r globalmente cerrada en M es globalmente exacta.

El espacio cociente vectorial $^{Z_r}(M)/B_r(M)$ divide el espacio de formas r cerradas en clases de equivalencia. Para cualquier forma r cerrada ω en M que no sea exacta, denotamos [ω] la clase de equivalencia de ω. Si ω y ω′ pertenecen a la misma clase de equivalencia, [ω] entonces difieren en una forma r exacta (es decir, ω − ω′ = dη), y decimos que ω y ω′ son cohomólogos. Esto es fácil de visualizar en términos de grupos de homología, donde simplemente significa que pueden existir múltiples ciclos r que contengan el mismo agujero de dimensión r. Así como un mapa continuo f : X → Y entre espacios topológicos X e Y induce un homomorfismo de retroceso f : Hr (Y; G) → Hr (X; G) entre el r-ésimo grupo de cohomología singular Hr (Y; G) en Y y el r-ésimo grupo de cohomología singular Hr (X; G) en X, un mapa suave F : M → N entre variedades suaves M y N induce un retroceso F : Hr dR(N) → Hr dR (M) entre los grupos de cohomología r-ésimo de Rham de N y M, respectivamente. Esto corresponde a un functor contravariante entre la categoría de mapas suaves entre variedades suaves y la categoría de grupos de cohomología de De Rham.

Podemos formarnos una imagen interesante de una relación entre los grupos de cohomología de Rham de grado uno H1 dR(M) en una variedad suave conectada y la topología subyacente X de la variedad medida por el grupo fundamental Π1(X, q). Todos los ciclos 1 en todas las 2 cadenas posibles en la topología subyacente de la variedad suave M que contienen un agujero unidimensional corresponden no solo a H1 dR(M) sino también al grupo fundamental Π1(X, q). El grupo fundamental consta de todos los bucles asociados con el punto q en la topología subyacente X de M que no se pueden reducir a un punto. El conjunto de homomorfismos de grupo Hom (Π1(X, q), R) desde el grupo fundamental Π1(X, q) hasta el grupo aditivo de números reales R es un espacio vectorial bajo suma y multiplicación por constantes. Podemos definir una aplicación lineal Φ : H1 dR(M) → Hom(Π1(X, q), R) de la siguiente manera: Dada una clase de cohomología [ω] H1 dR(M), defina Φ[ω] : Π1(X, q) → R por Φ[ω] = ω = un número real, donde [γ] es cualquier clase de homotopía de γ trayectoria en Π1(X, q) y γ es cualquier curva suave por tramos que represente la misma clase de trayectoria de bucle [γ]. El mapa Φ : H1 dR(M) → Hom(Π1(X, q), R) es un homomorfismo bien definido entre grupos. Se deduce que H1 dR(M) = 0 cuando el topol-

Machine Translated by Google

La tecnología X subyacente a la variedad suave M está simplemente conectada (es decir, no tiene agujeros unidimensionales ni bucles que no puedan contraerse hasta un punto).

Apéndice A.12. Teorema de De Rham

El teorema de De Rham depende del hecho de que, al igual que la homología singular y los grupos de cohomología singulares, los grupos de cohomología de De Rham Hr dR(M) para todos los valores de r son homotópica y topológicamente invariantes. Supongamos que los mapas suaves F : M → N y G : M → N entre variedades suaves M y N son homotópicos (es decir, F(M) en N se puede transformar en G(M) en N y viceversa). Defina un operador de homotopía h : Ωr (N) → Ωr−1 (M) para todos los valores de r. Supongamos que ω Ωr (N) y η = hω Ωr−1 (M) de modo que d(hω) + h(dω) donde G y f = G ω − F ω, son mapas de retroceso suaves desde Ωr (N) a Ωr (M) para todos los valores de r. Si dejamos que ω sea una forma r cerrada en N (es decir, dω = 0), encontramos d(hω) = G ω − F ω. En consecuencia, G ω y F ω difieren en una y f son cohomol- que G : Hr dR(N) → Hr dR(M) para todos los ' GRAMO forma exacta d(hω) en M, por lo que se deduce valores de r son ogos, y la atracción inducida -back mapas F isomórficos. Los grupos de cohomología de De Rham para todos los valores de r son homotópica y topológicamente invariantes. Se puede concluir que si M y N son variedades suaves equivalentes a homotopía con o sin límite, entonces Hr dR(N) = Hr dR(M) para cada valor de r. La equivalencia de homotopía suave F : M → N entre variedades suaves induce isomorfismos entre grupos de cohomología de De Rham.

Del teorema de Stokes y la definición de período zr ω donde zr es un ciclo r y ω es una forma r, podemos definir un paréntesis , :Cr(M) × C r (M) → R, donde Cr(M) es el grupo abeliano libre de todos posibles r-cadenas cr Cr(M) en M y C r (M) es el grupo abeliano libre r (M) en M, estableciendo cr , ω := de r- formas) ω C ω. Para cada forma r ω, el mapa cr −→ cr , ω es lineal en Cr(M), ya que toda cr es una combinación lineal de r-cadenas, es decir, C r(M), el mapa ω −→ cr , ω es lineal en C r (M). Esto induce un paréntesis en el grupo de homología r singular Hr(M) y el grupo de cohomología r-ésimo de Rham Hr dR(M) estableciendo

$$[zr], [\omega] = zr, \omega = \int_{zr} \omega$$

donde [zr] Hr(M) es una clase de homología r, [ω] Hr dR(M) es una clase de cohomología r-ésima de Rham , zr es cualquier representante de [zr] y ω es cualquier representante de [ω]. El período ωzr solo es igual a cero si la forma r cerrada ω es exacta; de lo contrario, es un número real distinto de cero. Para cualquier clase de cohomología fija de Rham r-ésima [ω], el mapa f [ω] : Hr(M) → R dado por f [ω] ([zr]) = [zr], [ω] = ω es un funcional lineal bien definido en Hr(M). El teorema de De Rham afirma que esta correspondencia entre funcionales lineales en Hr(M) y las clases de cohomología de De Rham [ω] Hr dR(M) es biyectiva. El mapa bilineal , : Hr(M) × Hr dR(M) → R no es degenerado en ambos argumentos [27]. Esencialmente, el teorema de de Rham dice que el grupo de cohomología r-ésimo de Rham es isomorfo con el espacio dual del grupo de homología r-ésimo (es decir, con el espacio del grupo de cohomología singular). Esto implica que una forma r ω es globalmente exacta si y sólo si todos sus períodos desaparecen. Esto sólo puede suceder si no hay agujeros de dimensión r en la topología subyacente de la variedad suave.

Referencias

1. Neilson, PD; Neilson, Doctor en Medicina; Adiós, RT Una teoría de la geometría riemanniana de la percepción visual binocular tridimensional. Visión 2018, 2, 43. [Referencia cruzada] [PubMed]
2. Neilson, PD; Neilson, Doctor en Medicina; Adiós, RT Una teoría geométrica de Riemann de selección de sinergia para movimientos guiados visualmente. Visión 2021, 5, 26. [CrossRef] [PubMed]
3. Neilson, PD; Neilson, Doctor en Medicina; Adiós, RT La teoría de la geometría de Riemann del movimiento guiado visualmente explica la imagen residual ilusiones y constancia de tamaño. Visión 2022, 6, 37. [CrossRef] [PubMed]
4. Katz, M.; Kruger, PB El ojo humano como sistema óptico, Capítulo 33. En Oftalmología clínica de Duane; Tasman, W., Jaeger, EA, Eds.; Lippincott, Williams y Wilkins: Filadelfia, PA, EE.UU., 2006; Volumen 1.

Machine Translated by Google

5. Armstrong, B.; Aitken, J.; Sim, M.; Swan, N. Informe final del Panel Independiente de Revisión e Investigación Científica; Corporación Australiana de Radiodifusión: Sydney, Nueva Gales del Sur, Australia, 2007.
6. Sitas, F.; O'Connell, DL; van Kemenade, CH; Corto, PM; Zhao, K. Riesgo de cáncer de mama entre empleadas de la Australian Broadcasting Corporation en Australia. Medicina. J. Aust. 2010, 192, 651–654. [Referencia cruzada] [PubMed]
7. Vieira, L. Torus a Mug Morphing Animation. 2007. Disponible en línea: https://commons.wikimedia.org/wiki/File:Mug_and_ Torus_morph.gif (consultado el 7 de junio de 2024).
8. Lee, J. Introducción a las variedades topológicas, 2ª ed.; Textos de Posgrado en Matemáticas; Springer: Nueva York, NY, EE.UU., 2011; ISBN 978-1-4419-7939-1.
9. Smale, S. Topología y mecánica. I. Invenciones Matemáticas. 1970, 10, 305–331. [Referencia cruzada]
10. Abraham, R.; Marsden, JE Fundamentos de Mecánica, 2ª ed.; AMS Chelsea Pub./Sociedad Americana de Matemáticas: Providence, RI, EE.UU., 2008; ISBN 978-0-8218-4438-0.
11. Ortega, J.-P.; Ratiu, T. Mapas de impulso y reducción hamiltoniana; Progreso en Matemáticas; Birkhäuser: Basilea, Suiza, 2004; ISBN 978-0-8176-4307-2.
12. Flandes, H. Formas diferenciales con aplicaciones a las ciencias físicas; Edición revisada; Publicaciones de Dover: Mineola, Nueva York, EE. UU., 1989; ISBN 978-0-486-66169-8.
13. Lang, S. Fundamentos de Geometría Diferencial; Textos de Posgrado en Matemáticas; Springer: Nueva York, NY, EE.UU., 1999; Volumen 191, ISBN 978-1-4612-6810-9.
14. Feynman, RP; Leighton, RB; Sands, M. Las conferencias Feynman sobre física, vol. II Cap. 41: El flujo de agua húmeda. 1963/2013. Disponible en línea: https://www.feynmanlectures.caltech.edu/II_41.html (consultado el 21 de abril de 2024).
15. Smits, AJ Conferencias sobre mecánica de fluidos: flujos viscosos y turbulencias; Smits, AJ, Ed.; Departamento de Ingeniería Mecánica, Universidad de Princeton: Princeton Nueva Jersey, EE. UU., 2009.
16. Brunton, SL; Noack, BR; Koumoutsakos, P. Aprendizaje automático para mecánica de fluidos. Año. Rev. Mec. de fluidos. 2020, 52, 477–508. [Referencia cruzada]
17. Dobson, CM Plegado y plegamiento incorrecto de proteínas. Naturaleza 2003, 426, 884–890. [Referencia cruzada] [PubMed]
18. Eneldo, KA; MacCallum, JL El problema del plegamiento de proteínas, 50 años después. Ciencia 2012, 338, 1042–1046. [Referencia cruzada] [PubMed]
19. La Real Institución. La física de la vida: cómo el agua pliega las proteínas, con Sylvia McLain; Vídeo de YouTube, 45:07. 2017. Disponible en línea: https://www.youtube.com/watch?v=SaSntOkK-Vk (consultado el 7 de junio de 2024).
20. Goodale, MA; Milner, D. Vista invisible: una exploración de la visión consciente e inconsciente; Oxford University Press: Oxford, Reino Unido, 2013; ISBN 978-0-19-175300-8.
21. Bullo, F.; Lewis, AD Control geométrico de sistemas mecánicos: modelado, análisis y diseño de sistemas de control mecánico simples; Textos de Matemática Aplicada; Springer: Nueva York, Nueva York, Estados Unidos, 2005; ISBN 978-0-387-22195-3.
22. La Real Institución. Aceleradores de partículas y plegamiento de proteínas: una nueva solución, con Sylvia McLain y Alan Soper; Vídeo de YouTube , 7:21. 2016. Disponible en línea: https://www.youtube.com/watch?v=uSErKi_HWg8 (consultado el 7 de junio de 2024).
23. Rhys, NH; Soper, AK; Dougan, L. La capacidad de formar enlaces de hidrógeno del aminoácido glutamina revelada por difracción de neutrones. experimentos. J. Física. Química. B 2012, 116, 13308–13319. [Referencia cruzada] [PubMed]
24. Stewart, BW "¡No habrá más!": El legado del grupo de cáncer de mama de Toowong. Medicina. J. Aust. 2007, 187, 178–180. [Referencia cruzada] [PubMed]
25. Lee, JM Colectores de Riemann: una introducción a la curvatura; Textos de Posgrado en Matemáticas; Springer: Nueva York, NY, EE.UU., 1997; ISBN 978-0-387-98271-7.
26. Morita, S. Geometría de Formas Diferenciales; Sociedad Estadounidense de Matemáticas: Providence, RI, EE. UU., 2001; ISBN 978-0-8218-1045-3.
27. Szekeres, P. Un curso de física matemática moderna: grupos, espacio de Hilbert y geometría diferencial; Prensa de la Universidad de Cambridge: Cambridge, Reino Unido, 2004; ISBN 978-0-511-26329-3.
28. Lee, JM Introducción a las variedades suaves, 2ª ed.; Textos de Posgrado en Matemáticas; Springer: Nueva York, NY, EE.UU., 2013; ISBN 978-1-4419-9981-8.
29. Tu, LW Geometría diferencial: conexiones, curvatura y clases características; Textos de Posgrado en Matemáticas; Saltador Publicaciones internacionales: Cham, Suiza, 2017; ISBN 978-3-319-55082-4.
30. Lee, JM Introducción a las variedades de Riemann, 2ª ed.; Textos de Posgrado en Matemáticas; Publicaciones internacionales Springer: Cham, Suiza, 2018; ISBN 978-3-319-91754-2.

Descargo de responsabilidad/Nota del editor: Las declaraciones, opiniones y datos contenidos en todas las publicaciones son únicamente de los autores y contribuyentes individuales y no de MDPI ni de los editores. MDPI y/o los editores renuncian a toda responsabilidad por cualquier daño a personas o propiedad que resulte de cualquier idea, método, instrucción o producto mencionado en el contenido.

Machine Translated by Google

AppliedMath

MDPI

Artículo

Formas innovadoras de desarrollar y utilizar un propósito específico Alternativas para resolver problemas de optimización ordenada y enrutamiento de redes combinatorias difíciles

Santosh Kumar 1,* y Elías Munapo 2

1 Departamento de Ciencias Matemáticas y Geoespaciales, Facultad de Ciencias, Universidad RMIT, Melbourne, VIC 3000, Facultad de
2 Economía y Ciencias de la Decisión de Australia, Universidad North West, Campus Mafikeng, Mafikeng 2735, Sudáfrica; emunao@gmail.com
* Correspondencia: santosh.kumarau@gmail.com

Resumen: Este artículo revisa algunas contribuciones recientes de los autores y sus asociados y destaca algunas ideas innovadoras que los llevaron a abordar algunos problemas difíciles de optimización ordenada y enrutamiento de redes combinatorias. El viajante, que se encuentra en la categoría NP difícil, ha sido revisado y resuelto como un gráfico conectado más corto restringido por índice y, por lo tanto, abre una pregunta sobre su categoría 'NP Difícil'. El problema de enrutamiento a través de nodos especificados 'K' y soluciones óptimas ordenadas son exigentes desde el punto de vista computacional, pero se han hecho computacionalmente factibles. Todos estos enfoques se basan en la creación estratégica y el uso de una solución alternativa en esa situación. La eficacia de estos métodos requiere más investigación.

Palabras clave: red conectada más corta; gira mínima de vendedor ambulante; NP Problemas difíciles; ruta a través de K nodos especificados; teorema de equilibrio de índices; red conectada con índice restringido; creación de una solución alternativa; soluciones óptimas ordenadas

check for updates

Cita: Kumar, S.; Munapo, E. Formas innovadoras de desarrollar y utilizar alternativas de propósito específico para resolver problemas de optimización ordenada y enrutamiento de redes combinatorias difíciles. Matemáticas aplicadas 2024, 4, 791–805. https://doi.org/10.3390/appliedmath4020042

Editor académico: Carlo Bianca

Recibido: 18 de abril de 2024
Revisado: 6 de junio de 2024
Aceptado: 6 de junio de 2024
Publicado: 18 de junio de 2024

Copyright: © 2024 por los autores. Licenciatario MDPI, Basilea, Suiza. Este artículo es un artículo de acceso abierto. distribuido bajo los términos y condiciones de los Creative Commons Licencia de atribución (CC BY) (https://creativecommons.org/licenses/by/4.0/).

1. Introducción

Las matemáticas de OR han brindado apoyo en muchos procesos de toma de decisiones industriales, comerciales y gubernamentales de la vida real . La tecnología cambiante siempre ha sido una fuente de nuevos desafíos y, en consecuencia, se han desarrollado nuevos métodos y herramientas matemáticos para brindar el apoyo necesario a la toma de decisiones para tomar decisiones mejores e informadas. Los problemas de enrutamiento y sus variantes en la teoría de grafos han atraído mucha atención por parte de los investigadores debido a diversas aplicaciones industriales. La ruta más corta, el gráfico conectado más corto y el viajante de comercio son sólo algunos de ellos. De vez en cuando se han identificado variantes de estos problemas. La optimización ordenada es otro campo en el que obtener un óptimo ordenado, es decir, la k -ésima mejor solución, k ≥ 2, puede ser una tarea muy desafiante. En este artículo, revisamos algunas ideas innovadoras específicas para desarrollar una alternativa en una situación determinada y utilizamos esa solución alternativa para resolver algunas situaciones difíciles de optimización combinatoria; por ejemplo, nos encontramos con lo siguiente: (1) Se soluciona el recorrido mínimo del viajante de comercio, que está en la categoría 'NP Difícil' como un gráfico conectado mínimo restringido por índice;
(2) La ruta que pasa a través de un conjunto de nodos específicos, que es computacionalmente exigente debido a que la ruta está restringida a pasar a través de un conjunto de nodos específicos, también se resuelve como un gráfico conectado mínimo restringido por índice;
(3) La segunda mejor solución óptima para la ruta más corta y el problema de asignación.

Todos estos enfoques desarrollan una solución alternativa y utilizan esa solución de manera estratégica para resolver el problema bajo consideración. Estos nuevos enfoques necesitan más investigación para establecer su eficiencia, si la hay. Ese trabajo se está realizando actualmente. Entre los problemas anteriores, se ha clasificado el recorrido mínimo del viajante de comercio (MTST)

Machine Translated by Google

como 'NP Duro'; ver [1]. De manera similar, la determinación de la ruta más corta a través de nodos especificados 'K' y la búsqueda de una solución óptima ordenada se han clasificado como computacionalmente exigentes; ver [2,3]. El enfoque basado en el gráfico conectado más corto restringido por índice para el problema del viajante plantea una pregunta: "¿Es el TSP realmente un problema 'NP-Difícil' ?" Una pregunta similar surge nuevamente para el camino a través de K nodos especificados. La pregunta es: "¿Es el camino a través de nodos específicos exigente desde el punto de vista computacional?" El desarrollo del árbol de expansión mínimo restringido se basa en la idea de desarrollar una alternativa y hacer uso de esa alternativa para encontrar el árbol de expansión mínimo restringido

Ampliamos la idea de desarrollar una solución alternativa para la ruta más corta y un problema de asignación y encontramos la segunda mejor solución óptima para estos problemas, sin aumentar excesivamente los cálculos. Este enfoque se encuentra en su etapa inicial y es necesario más trabajo computacional para establecer su eficiencia computacional para diferentes problema

Muchas situaciones de la vida real se pueden representar en forma de red, que está formada por nodos y enlaces. Estos enlaces se forman uniendo los nodos, y los pesos asociados con estos enlaces proporcionan una especie de medida de la relación entre los nodos. La red que representa una situación física puede ser una red dirigida o no dirigida. El análisis de una red no dirigida suele ser más exigente en comparación con un gráfico dirigido, simplemente porque las direcciones en un gráfico dirigido proporcionan información adicional, que falta en una red no dirigida. Muchas situaciones industriales se han modelado como redes dirigidas y se han analizado mediante técnicas de programación dinámica; véanse [4–6]. Estos problemas, una vez transformados en un modelo de red, se analizaron utilizando la teoría y las herramientas relacionadas con las redes. El resultado del análisis se puede interpretar en términos de la situación de la vida real que se modeló como una red. Para conocer los desarrollos recientes, consulte [7,8

En general, una red se puede representar mediante G(N, L), donde N representa un conjunto de algunas entidades representadas por los nodos N = {1, 2,. . . , n} y L es el conjunto de enlaces que unen varios nodos que representan propiedades asociadas con pares de estos nodos. Estos enlaces se indican con (i, j), donde i y j son los nodos 1, 2 ,. . . , n y i = j. En las redes ponderadas, asociamos un número con cada enlace, lo que proporciona una medida cuantitativa de la relación entre esos dos nodos. Por ejemplo, puede ser simplemente la distancia entre los dos nodos, o puede representar el tiempo de viaje entre el par de nodos, i y j.
Este peso asociado al vínculo (i, j) se denota por cij, que puede ser una cantidad positiva o negativa, según la situación.

Algunas ideas utilizadas en este artículo se revisaron en la Sección 2, y sus aplicaciones, en la forma del vendedor ambulante y los problemas de rutas, se presentaron en la Sección 3. El problema del óptimo ordenado se analizó en la Sección 4. Finalmente, Las observaciones finales se presentan en la Sección 5.

2. Modificación de un problema clásico de red conectada más corta y sus aplicaciones La red conectada más corta

(SCN) de cualquier red conectada no dirigida G(N, L) es un problema de red clásico. Comprende (n − 1) enlaces, donde n representa el número de nodos en la red dada G(N, L). Dado que la red dada está conectada, debe haber una ruta única desde cada nodo a todos los demás nodos, y esa ruta estará libre de bucles. En un gráfico tan conectado, si se elimina algún vínculo, el gráfico se desconectará. Este SCN se puede obtener mediante un enfoque codicioso según [9-11]. En [12] se analiza una modificación del problema del gráfico conectado más corto , que se presenta brevemente a continuación.

2.1. Modificación de la Red de Conexión Más Corta (SCN)

2.1.1. Declaración de la red conectada más corta modificada

Para una red G(N, L), el SCN puede obtenerse mediante cualquier método conocido. Desde el SCN, se puede encontrar un índice de nodo para cada nodo. El índice de nodo es el número de enlaces que emanan de ese nodo en el SCN. Sea ni el índice del nodo i, donde 1 ≤ ni ≤ (n − 1). En la modificación discutida en [12] en la que se restringe el SCN, el

Machine Translated by Google

El índice de nodo ni satisface 1 ≤ ni ≤ 2 para cada nodo. Esto significa que si el SCN original no satisface la restricción sobre el valor del índice en cualquier nodo, el SCN obtenido utilizando el enfoque clásico debe modificarse para satisfacer la condición 1 ≤ ni ≤ 2, i. En la red de índice restringido dada, dos nodos tendrán un valor de índice de 1 y los otros nodos tendrán un valor de índice de 2; por tanto, formarán un camino que unirá dos nodos y pasará por todos los demás nodos restantes de la red. Podemos suponer que en la red conectada dada, todos los nodos son tales que al menos dos enlaces emanan de cada nodo.

2.1.2. Algunos términos esenciales y un teorema para modificar el SCN obtenido mediante el enfoque codicioso de un SCN restringido por índice (IRSCN)

Los términos necesarios son los siguientes:

T1. Índice de un nodo: Dado que cada enlace seleccionado genera dos índices, el número total de índices de los enlaces seleccionados (n − 1) que forman el SCN será 2(n − 1). Para una ruta que pasa por todos los nodos, dos de estos nodos tendrán solo un valor de índice (desde donde comienza la ruta hasta donde termina), y los nodos restantes (n − 2) tendrán un índice de 2. La ruta restringida por índice SCN o IRSCN formarán una ruta que unirá dos nodos, que pasará por todos los demás nodos restantes de la red.

T2. Enlaces básicos y no básicos: un enlace que une dos nodos i y j se dice que es básico si xij = 1, es decir, ese enlace pertenece al SCN y, si xij = 1, entonces es un enlace no básico.

T3. Nodos de índice alto y bajo: una vez que se ha identificado un SCN utilizando cualquier método conocido, conocemos los valores de los índices de los nodos para todos los nodos. Estos valores vienen dados por 1 ≤ ni ≤ (n − 1). Deben cambiarse para satisfacer la condición 1 ≤ ni ≤ 2 i, es decir, para los nodos de origen y de destino, el índice ni = 1, y para todos los demás nodos restantes (n − 2) , ni = 2. El alto - y los nodos de índice bajo se pueden identificar fácilmente por su valor actual frente al valor de índice de nodo deseado para ese nodo.

T4. Nodos y enlaces vecinos: se dice que un nodo i es un nodo vecino del nodo j, si los nodos i y j están conectados por un solo enlace. En una red completamente conectada, todos los nodos son nodos vecinos ya que todos los nodos están conectados mediante un enlace directo. Un enlace vecino se define como un enlace que emana de un nodo vecino.

Teorema 1. Teorema de equilibrio de índices: Agregar la misma constante a todos los enlaces que emanan del mismo nodo no cambia el mérito relativo de cualquier enlace seleccionado en el SCN obtenido mediante el enfoque codicioso. Sin embargo, el teorema del equilibrio de índices puede crear alternativas. A partir del SCN obtenido mediante el método codicioso, conocemos el índice de cada nodo; por lo tanto, los nodos de índice alto y bajo se pueden identificar fácilmente. El teorema de equilibrio de índices comienza su trabajo agregando constantes a todos los enlaces que emanan del mismo nodo, que ha sido clasificado como un nodo de índice alto. La idea es crear alternativas para la selección de un enlace, y así reducir el índice en un nodo de alto índice y aumentar el índice de un nodo de bajo índice; cuando se elimina un enlace de un nodo de índice alto y se coloca en un nodo con un valor de índice bajo, el desequilibrio se reduce. Este proceso se repite hasta que todos los nodos alcanzan el índice requerido y satisfacen todas las restricciones de índice

Teorema 2. Para cualquier solución SCN dada, un enlace básico puede modificarse agregando una constante a todos los enlaces que emanan de ese nodo, es decir, cambiando el estado de un enlace de básico a no básico y el de algún otro enlace de no básico. básico a básico. Las aplicaciones repetidas pueden generar una ruta SCN que une dos nodos y pasa por todos los demás nodos de una red, es decir, satisfará la condición de índice en cada nodo.

2.1.3. Ilustración numérica

Considere un ejemplo trivial simple pero interesante: los valores cij dados en la Tabla 1.

Machine Translated by Google

Tabla 1. Una red de 6 nodos completamente conectada con enlaces ponderados.

i\j	1	2	3	4	5	6
1	-	1	1	1	1	1
2	1	-	2	2	2	2
3	1	2	-	3	3	3
4	1	2	3	-	4	4
5	1	2	3	4	-	5
6	1	2	3	4	5	-

El SCN se puede obtener utilizando el enfoque codicioso, que se muestra en la Figura 1. Se puede comenzar desde cualquier nodo, digamos el nodo 6, y seleccionar el enlace (6,1) como primer enlace; el próximo Los enlaces seleccionados pueden ser (1,2), (1,3), etc. El SCN será el que se muestra en la Figura 1.

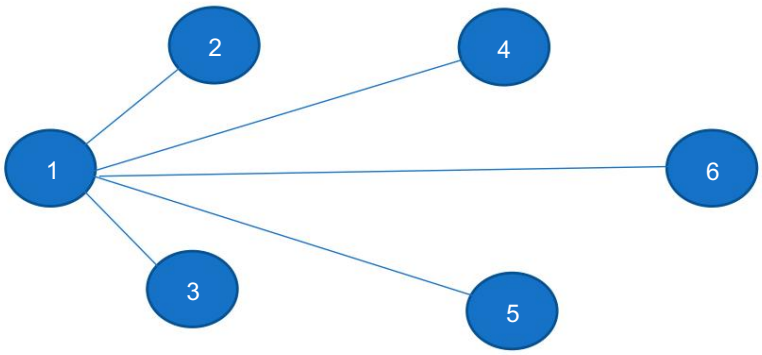


Figura 1. Gráfico conectado más corto de la red de la Tabla 1.

La matriz de incidencia del SCN en la Figura 1 se muestra en la Tabla 2.

Tabla 2. Matriz de incidencia del SCN en la Figura 1.

Nodos	1	2	3	4	5	6
Incidencia	5	1	1	1	1	1

Esto significa que si queremos un camino, digamos, entre los nodos 5 y 6, los nodos 5 y 6 son equilibrado con un valor de índice de nodo de 1 y todos los demás nodos están desequilibrados. El nodo 1 es un nodo de índice alto, y los nodos 2, 3 y 4 son nodos de índice bajo. Como aplicación del teorema de equilibrio del índice, agregue un valor constante de 1 a todos los enlaces que emanan del índice alto. nodo, que es el nodo 1; los méritos relativos de los enlaces (1,2), (1,3), (1,4), (1,5) y (1,6) permanecer sin cambios. Sin embargo, se ha desarrollado una alternativa agregando 1 a todos los enlaces. que emana del nodo 1. La alternativa es un enlace que comienza en el nodo 2, en lugar de un enlace comenzando desde el nodo 1. Por lo tanto, el enlace (1,4) se puede eliminar y reemplazar por el enlace (2,4). Este cambio es posible ya que se ha creado una alternativa para comenzar desde el nodo 2. Tenga en cuenta que reducirá el índice del nodo de índice alto 1 en 1 y aumentará el valor del índice de el nodo de índice bajo 2 por 1. Esto se muestra en la Figura 2. De manera similar, nuevamente, si se suma 1 a todos enlaces que emanan de los nodos 1 y 2, el enlace (1,5) puede eliminarse y reemplazarse por el enlace (3,5). Una aplicación más del teorema de equilibrio de índices dará como resultado el SCN como se muestra en la Figura 3.

Sin embargo, los nodos 3 y 4 siguen siendo nodos de índice bajo y el nodo 1 es un nodo de índice alto. Otra aplicación proporciona el SCN restringido por índice, como se muestra en la Figura 3.

Machine Translated by Google

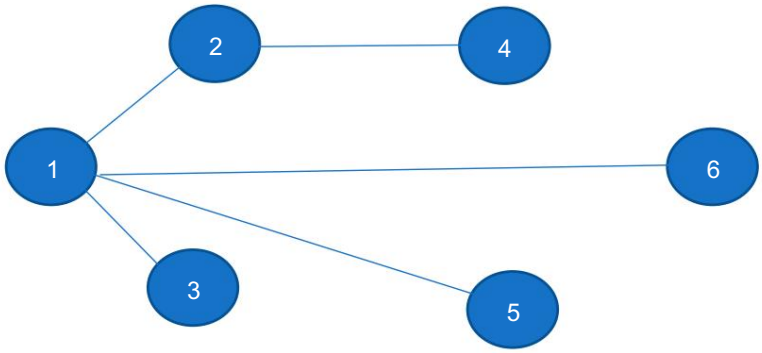


Figura 2. SCN mejorado.

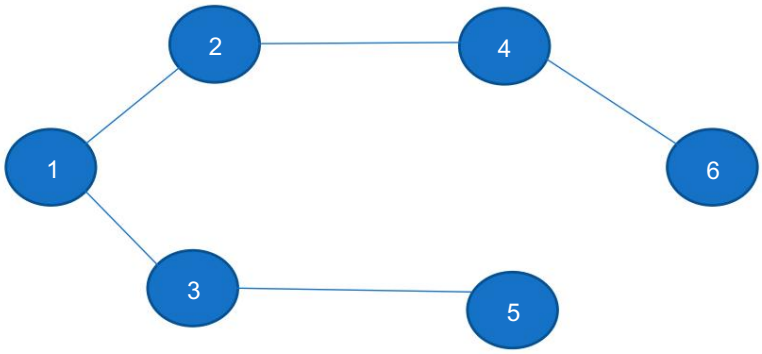


Figura 3. SCN restringido por índice de la red en la Tabla 1.

Supongamos que nuestro interés fuera encontrar el gráfico conectado más corto con una restricción de índice de 2 en todos los nodos excepto en los nodos inicial y final, donde el índice está restringido a 1. La incidencia total del SCN seleccionado será 2(n − 1) = 10, como en este ejemplo, n = 6. En En otras palabras, el IRSCN será un camino entre dos nodos y pasará por (n − 2) número de nodos. La red conectada en la Figura 3 es un IRSCN requerido y es posible que no ser único.

3. El SCN restringido por índices y sus aplicaciones

3.1. El problema del viajante

El recorrido mínimo del viajante de comercio es un problema bien conocido en la teoría de grafos. Se trata de la determinación de un recorrido de costo mínimo que pase por todos los nodos y regresando a la ciudad de origen. Este clásico problema del viajante tiene muchas implicaciones industriales. aplicaciones, y ha sido clasificado como un problema NP Difícil; véase [1,13-15]. Otro El problema que se ocupa de todos los nodos en la teoría de grafos es el grafo conectado más corto, que ha sido obtenido por el enfoque codicioso; ver [11]. En un artículo anterior [12], los autores desarrolló un enfoque para convertir la red conectada más corta (SCN) en un gráfico conectado más corto restringido por índice, como se analiza en la Sección 2. El SCN restringido por índice se puede utilizar para encontrar el recorrido mínimo del viajante de comercio. El SCN restringido por índice es desarrollado utilizando el teorema de equilibrio de índices [12]. Por lo tanto, surge una pregunta: ¿es la ¿El problema del viajante es realmente NP Difícil? Se ha utilizado el SCN en una forma modificada. para encontrar el recorrido del viajante de diferentes maneras, [16,17], y finalmente, el mínimo El árbol de expansión bajo la restricción del índice se ha establecido como igual al árbol de viaje. gira de vendedores [18]. Explicamos brevemente estos enfoques en las Secciones 3.1.1–3.1.3.

3.1.1. Un enfoque basado en árboles de expansión mínima para encontrar el recorrido mínimo Tour del Vendedor (MTST) de la Red G(N, L): Enfoque 1 (Complejidad rC2)

Proponemos iniciar el recorrido mínimo del viajante desde el nodo con un valor mínimo del índice de nodo, es decir, el nodo desde donde se encuentra un número mínimo de enlaces emanando. También es cierto que cada nodo tendrá dos enlaces, que formarán parte

Machine Translated by Google

del recorrido mínimo de viajante de comercio. Si el nodo de índice mínimo se denota por
el nodo p y su valor de índice es r, es decir, r enlaces emanan del p, dos de estos
Los enlaces formarán parte del recorrido mínimo del vendedor. Examinamos todas esas posibilidades.
Denotemos estos r enlaces que emanan del nodo p por pr . El mínimo de viaje
El recorrido del vendedor se puede identificar evaluando un número rC2 de recorridos más cortos restringidos por índice.
árboles conectados de la red modificada G((N − 1),(L − pr)). Este símbolo indica
que la red modificada se obtiene después de eliminar el nodo p y todos los enlaces que están
que emana del nodo p. Por lo tanto, la red modificada tendrá (N − 1) nodos y la red modificada tendrá (N − 1) nodos y la
El número de enlaces vendrá dado por (L − pr). Cada árbol con índice restringido se convierte en un árbol itinerante.
recorrido del vendedor cuando los dos bordes iniciales seleccionados arbitrariamente se unen a ese camino o ese
IRSC. El tour de vendedor mínimo requerido pasa a ser el de valor mínimo.
de estos viajes de vendedores ambulantes de rC2 .

En lo anterior, sugerimos iniciar el recorrido del vendedor ambulante desde un nodo con un
valor mínimo del índice de nodo. Para explicar este punto, se puede hacer referencia a la red de la Figura 4.
y tenga en cuenta que el valor del índice del nodo 1 es 3; los nodos 2, 3 y 4 tienen un valor de índice de 5; y
Los ánodos 5 y 6 tienen un valor de índice de 4. Estos valores de índice son el número de emanaciones
enlaces desde ese nodo. Para minimizar el trabajo computacional, es posible que prefiramos iniciar el
recorrido del vendedor desde el nodo 1, donde solo se requerirá un número de evaluaciones de 3c2 ; más
Sería necesario si comenzamos desde cualquier otro nodo.

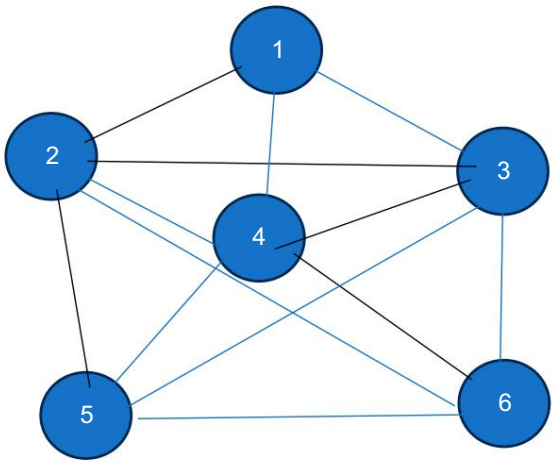


Figura 4. Dada la red G(6, 13) .

Este enfoque se discutió en [17].

3.1.2. Un enfoque basado en árboles de expansión mínima para encontrar el recorrido mínimo
Tour de vendedores (MTST) de G(N, L): Enfoque 2 (Complejidad pr)

Como se discutió anteriormente, una vez más podemos usar el enlace más corto restringido por índice.
gráfica para encontrar el recorrido mínimo del vendedor ambulante. Considere una modificación ligeramente
ción de la red dada. Esta vez, consideramos una modificación y denotamos esta modificación.
red por G(N,(L − 1)). Significa que no hay cambios en el número de nodos, y sólo
se ha eliminado un enlace que conecta el nodo p con algún otro nodo, digamos con el nodo q. Dejar
enlaces se denotarán por {(p, q1), (p, q2), . . ., (p, qr)}. En este enfoque, se propone , , estos r
encuentre r número de árboles conectados más cortos que unen los nodos (p, q 1), (p, q 2), , , (p, q r), después
eliminando solo un enlace a la vez. Por ejemplo, eliminemos el enlace (p, q 1 ,), el) y encontrar
IRSCN uniendo los nodos p y q1. Después de unirse al enlace (p, q 1), el índice restringido
El camino se convertirá en un recorrido de vendedor ambulante. De manera similar, encontramos el más corto restringido por índice.
Redes conectadas que unen los nodos p y qr , r = 1, 2 ,. . ., r. Después de unirse al enlace (p, qr)
y el gráfico conectado más corto mínimo, formará un recorrido de vendedor y el mínimo
de estos r recorridos será el recorrido mínimo requerido de vendedor ambulante. Por lo tanto, la complejidad se reduce
a r árboles conectados más cortos con índice restringido en comparación con la complejidad
rC2 , como fue el caso en el enfoque anterior.

Machine Translated by Google

Este enfoque se discutió en [16].

3.1.3. Un enfoque basado en un árbol de expansión mínimo para establecer que el recorrido mínimo del vendedor es equivalente a un árbol de expansión mínimo restringido por índice: enfoque 3 (complejidad 1)

Una vez más, deje que el nodo seleccionado sea p. La red dada G(N, L) puede Agregar el nodo p ′, modificarse y los enlaces que emanan del nodo p también se agregan del nodo p ′como enlaces que emanan ; la red dada se transforma en G{(N + 1),(L + pr)}. debería observó que esta vez, en lugar de eliminar algunos nodos o enlaces, estamos agregando el nodo p y todos los enlaces que emanan de este nodo a la red modificada. Para mantener una diferencia en el nodo seleccionado, su duplicado y, de igual forma, los enlaces que están emanando del nodo duplicado, hemos utilizado la notación p y p han sido duplicados, su uso es posible sólo una vez según la ′. Dado que el nodo p y todos los enlaces que emanan de este existencia del duplicado. Los enlaces son virtuales. Nodo p sobre p, estamos hablando de un nodo p, que es real. A continuación, explicamos ′también estas entidades virtuales, es decir, el nodo p. Dado que los nodos p y p

′ y los enlaces virtuales asociados desde el nodo p son ′.

′ esencialmente los mismos, la ruta restringida por índice desde el nodo p formará Pensando ′ un recorrido de vendedor. Así, en este caso, con la ayuda de algún creativo y en el nodo p , encontramos que el árbol conectado más corto restringido por índice que une los nodos p y p . nos da el recorrido mínimo requerido por el vendedor.

En la modificación anterior de la red dada, el enfoque codicioso habitual no es directamente aplicable para encontrar el árbol de expansión mínimo. El número pr de enlaces adicionales no es de expansión mínimo ′ real; solo uno de ellos existe cuando se selecciona para ser parte del árbol agregado al nodo p , y el otro enlace debe desaparecer. Esto significa que tan pronto como uno de estos vínculos se vuelve básico, el otro debe desaparecer. En otras palabras, si el enlace p, qj es básico, el enlace ′, qj tendrá una longitud infinita. De manera similar, si el enlace duplicado es básico, el enlace original se cambia a un valor infinito. Sin embargo, si estos vínculos no son básicos, pueden coexistir como vínculos no básicos, ya que no tienen una existencia significativa en el viaje del viajante de comercio.

Red modificada y pasos para obtener el árbol de expansión mínimo de G{(N + 1),(L + pr)}

Paso 1: Para la red dada G(N, L), desarrolle la red G{(N + 1),(L + pr)} y organice los enlaces en orden creciente. Inicialmente, todos los enlaces no son básicos, por lo que todos los enlaces duplicados serán los mismos que en G(N, L) . Los enlaces que han sido duplicados se denominan enlaces de Tipo 2 y los que no han sido duplicados se denominan enlaces de Tipo 1. El número de enlaces de Tipo 1 es (L − pr) y el de Tipo 2 es 2 pr . Los enlaces de tipo 2 son tales que al menos dos enlaces tienen el mismo valor de longitud, inicialmente.

Paso 2: Establezca un contador K = 1. Seleccione el enlace de longitud mínima, inclúyalo en el árbol de expansión mínima y vaya al Paso 3.
Paso 3: Si el enlace seleccionado es Tipo 1 y K < n, establezca K = K + 1 y seleccione la siguiente longitud mínima. Si K = n, vaya al Paso 7.
Paso 4: Si el borde seleccionado era Tipo 2 y K < n, primero, establezca la longitud del enlace duplicado igual a ∞ y luego reorganice las longitudes en orden creciente. Establezca K = K + 1 y seleccione el siguiente mínimo. Si K = n, vaya al Paso 7.
Paso 5: Si el enlace seleccionado forma un ciclo con el árbol de expansión formado hasta el momento, deséchelo o inclúyalo. Si este último enlace fue Tipo 1, regrese al Paso 3. Si el enlace seleccionado fue Tipo 2, vaya al Paso 6.

Paso 6: Para el borde seleccionado, verifique lo siguiente:

6.1: Si la inclusión del borde no conduce a un nodo aislado en la red, deséchelo.

6.2: Si existen alternativas, opte por la arista que desarrolle un mejor equilibrio entre el índice de nodos que forman el árbol de expansión mínimo.

Regrese al Paso 4 ya que el enlace seleccionado es del Tipo 2.

Machine Translated by Google

Paso 7: Deténgase cuando se haya obtenido el árbol de expansión mínimo.

Ilustración numérica

Considere la siguiente red G(6, 13) que se muestra en la Figura 4, se dan los pesos de los enlaces en la Tabla 3. (6,13)

Tabla 3. Distancias de enlace para la red en la Figura 4.

PIE	1	2	3	4	5	6
1	-	12	10	10	-	-
2	12	-	15	11	11	16
3	10	15	-	7	14	12
4	10	11	7	-	10	11
5	-	11	14	10	-	9
6	-	16	12	11	9	-

Ahora, considere la red modificada que será G((6+1),(13+3)) o G(7, 16) como (7,16) como se muestra en la Figura 5, donde el nodo 1 se ha duplicado como 1'.

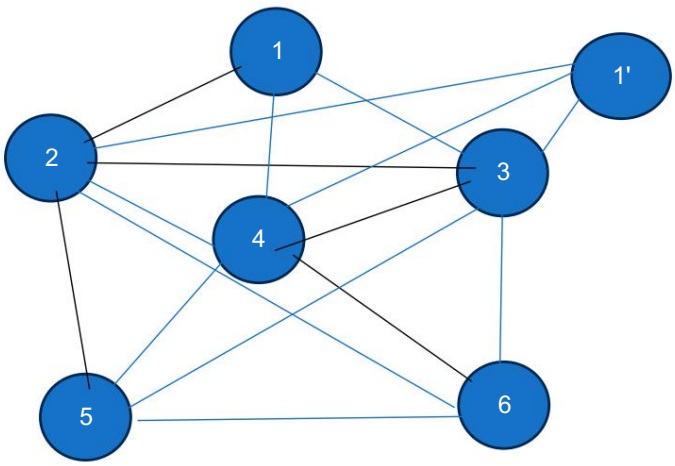


Figura 5. Red modificada G(7, 16). (7,16)

Los enlaces de tipo 1 son (2,3), (2,4), (2,5), (2,6), (3,4), (3,5), (3,6), (4,5), (4,6) y (5,6).

Los enlaces de tipo 2 son (1,2), (1,3), (1,4), (1',2), (1',3) y (1',4).

Disponiendo los enlaces en orden creciente, tenemos 7 (3,4)T1, 9(5,6)T1, 10(4,5)T1, 10(1,3)T2, 10(1,4)T2, 10(1',3)T2, 10(1',4)T2, 11(2,4)T1, 11(2,5)T1, 11(4,6)T1, 11(4,6)T1, 12 (3,6)T1, 12(1,2)T2, 12(1',2)T2, 15 (2,3)T1 y 16(2,6)T1.

Determinación de MST: Enlace 7(3,4) T1 incluido, enlace 9(5,6) T1 incluido, 10(4,5) T1 incluidos, 10(1,3) T2 incluidos; sin embargo, ahora el enlace (1',3) se cambia a infinito, por lo que la lista de orden creciente se cambia como se indica a continuación:

10(1,4) T2, 10(1',4) T2, 11(2,4) T1, 11(2,5) T1, 11(4,6) T1, 11(4,6) T1, 12 (3,6) T1, 12(1,2) T2, 12(1',2) T2, 15 (2,3) T1 y 16(2,6) T1, ∞(1',3) T2.

Tenga en cuenta que el enfoque codicioso identificará el siguiente enlace (1,4) de longitud 10, pero hace un circuito, y si vamos a favor de (1',4), también forma un circuito 1-3-4-1', ya que los nodos 1 y 1' son lo mismo. También está excluido. Se excluye el siguiente enlace (2,4) y se incluye el enlace (2,5). El enlace (4,6) está excluido. Finalmente, tenemos la Figura 6.

Machine Translated by Google

∞

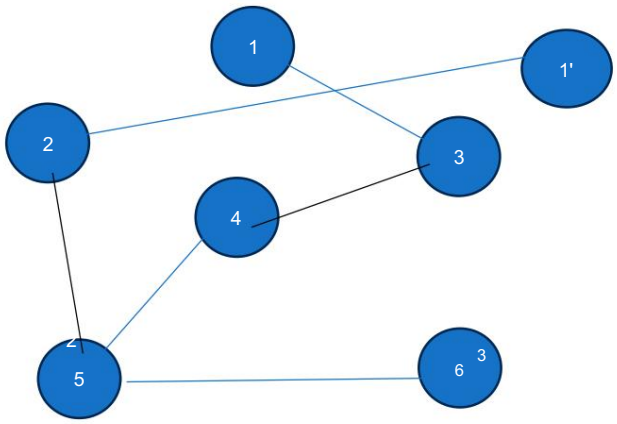


Figura 6. Árboles de expansión por enfoque codicioso.

Mirando la Figura 6, encontramos que los nodos 5 y 6 están desequilibrados como índice de nodo en El nodo 5 es 3 y el índice de nodo en el nodo 6 es 1. Si se agrega 1 a todos los nodos que emanan de nodo 5, el peso asociado con el enlace (4,5) será igual al peso del enlace (4,6); por lo tanto, resultará en una alternativa. Crea la posibilidad de eliminar el enlace. (4,5) y reemplazándolo por el eslabón (4,6). Esto se muestra en la Figura 7.

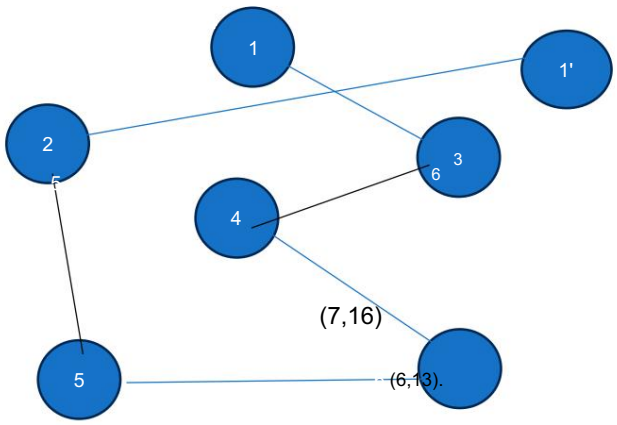


Figura 7. El camino 1 -> 3 -> 4 -> 6 -> 5 -> 2 -> 1'.

Este camino en la red G(7, 16) tiene una interpretación de un mínimo restringido por índice. árbol de expansión que une los nodos 1 y 1' bajo la restricción del índice de nodos y un recorrido por el vendedor interpretación para la red G(6, 13). Se muestra en la Figura 8. El costo de este recorrido será 12(1,2) + 11(2,5) +9(5,6) + 11(6,4) + 7(4,3) + 10((3,1) = 60 (1-> 2 - > 5 -> 6 -> 4 -> 3 -> 1).

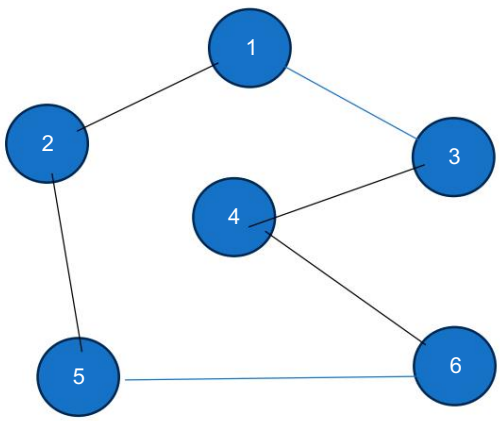


Figura 8. El árbol generador y el recorrido mínimo del viajante de comercio.

tt

Machine Translated by Google

3.2. La ruta más corta en una red no dirigida bajo la condición de pasar a través de K nodos específicos

Kalaba [19] sugirió un problema de enrutamiento restringido que encuentra una ruta desde el nodo de origen al nodo de destino, pasando por un número K de nodos específicos. Utilizando la técnica de programación dinámica, este problema se intentó en [20]. Cuando el conjunto especificado está vacío, el problema se reduce a un problema de enrutamiento ordinario, y cuando el conjunto de nodos especificados contiene todos los nodos y regresa al nodo desde el que comenzamos, se convierte en el problema del recorrido del vendedor. Debido a la naturaleza combinatoria del problema, es difícil resolverlo mediante la técnica de programación dinámica; sin embargo, cuando se intenta solucionar este problema utilizando el árbol de expansión más corto bajo la restricción del índice de nodo, la carga computacional se vuelve manejable. Todos los nodos se incluyen en el SCN y se cambia el aspecto combinatorio, que hace que el problema sea computacionalmente exigente para el enfoque de programación dinámica. Para los nodos especificados, simplemente aplicamos una restricción de índice asegurándonos de que el nodo especificado sea visitado en la ruta. Se logra imponiendo una restricción de índice de 2 o más y asegurándose de que se cumpla la restricción.

Véase, por ejemplo, [21], donde los autores consideraron una aplicación industrial de llamadas de servicio y el problema se desarrolló en forma matemática como una ruta a través de K enlaces específicos.

Todos estos problemas de enrutamiento, donde la ruta está restringida para pasar a través de un conjunto de nodos o enlaces específicos, son computacionalmente exigentes, pero cuando estos problemas se intentan como un SCN restringido por índice, la complejidad computacional se reduce. De hecho, es fácil encontrar el SCN restringido por índice bajo cualquier condición, como

$1 \le n_i \le k_i$, donde k_i es un número dado entre $1 \le k_i \le (n-1)$.

4. Óptimo ordenado

La mejor solución de K th ($K \ge 2$) en general es un problema de optimización ordenada. La k th mejor, ($K \ge 2$), se define como la mejor solución después de excluir ($K-1$) las soluciones óptimas ordenadas . No existe una teoría bien definida para la determinación de soluciones óptimas ordenadas para problemas de optimización, excepto en [22], donde los autores desarrollaron una ecuación característica que identifica soluciones óptimas ordenadas para un problema de programación entera lineal pura . Aquí, consideraremos las matemáticas de los problemas OR para encontrar la segunda mejor solución, utilizando el concepto de solución óptima alternativa. Se han propuesto muchas sugerencias en el pasado, pero pueden ser exigentes desde el punto de vista computacional. Estos problemas surgen en diversas situaciones de la vida real y, en general, se puede identificar la solución mejor u óptima, ya que todos los métodos existentes tienen condiciones definidas para la optimización de la solución y, cuando esas condiciones se satisfacen, se puede concluir fácilmente que la solución óptima. se ha obtenido la solución. Sin embargo, la existencia de th, no ha sido definido; más bien, la K th óptimo se define como el mejor, excluyendo una condiciones de optimización para la K mejor solución es th mejor solución surge en muchos diferentes que $K \ge 2$ ($K-1$) número de mejores soluciones. Una necesidad de las K situaciones (consulte [2] para ver un ejemplo), donde se requiere la mejor K para que un programa la k th lineal analice el problema de programación matemática del punto extremo. Se requiere la mejor solución por diferentes razones en diferentes situaciones. En esta sección, abordamos el problema del óptimo ordenado, que puede ser computacionalmente exigente, nuevamente, desarrollando una solución alternativa y usando esa solución para identificar la segunda mejor solución. Se han utilizado dos problemas para ilustrar la determinación de la segunda mejor solución: (1) la segunda ruta más corta para un problema de ruta más corta y (2) el problema de asignación. Para el problema de asignación, Murthy [3] propuso un método para clasificar todas las asignaciones en orden de costos crecientes. En un problema de asignación nXn , Murthy [3] obtuvo la segunda mejor solución resolviendo n problemas de asignación asignando un alto costo a una de las asignaciones óptimas a la vez, y la mejor de todas estas n soluciones fue la segunda mejor solución para ese problema. Su idea es que el segundo mejor se diferenciará del mejor en al menos una tarea, y evalúa

En este artículo, la idea es aumentar la longitud de la solución óptima actual en una cierta cantidad para establecer dos o más soluciones óptimas alternativas. uno de estos

Machine Translated by Google

f

Las soluciones alternativas son la solución óptima existente y la otra es una solución recién creada.
solución óptima alternativa. Dado que el valor de la solución óptima aumenta en un
De manera controlada, estas dos soluciones óptimas alternativas tienen una interpretación interesante.
de las mejores y segundas mejores soluciones con respecto a la información de costos original. Nosotros
ilustrar la idea tomando el problema de la ruta más corta en una red y una asignación
problema para encontrar la mejor y la segunda mejor solución.

4.1. La segunda mejor ruta para el problema de enrutamiento convencional en una red dirigida

Ilustración numérica para encontrar la segunda mejor ruta más corta

Apliquemos la idea de desarrollar soluciones alternativas al enrutamiento de la red.
problema considerado en la Figura 9, el cual se transformó en la Figura 10 para encontrar la solución óptima.

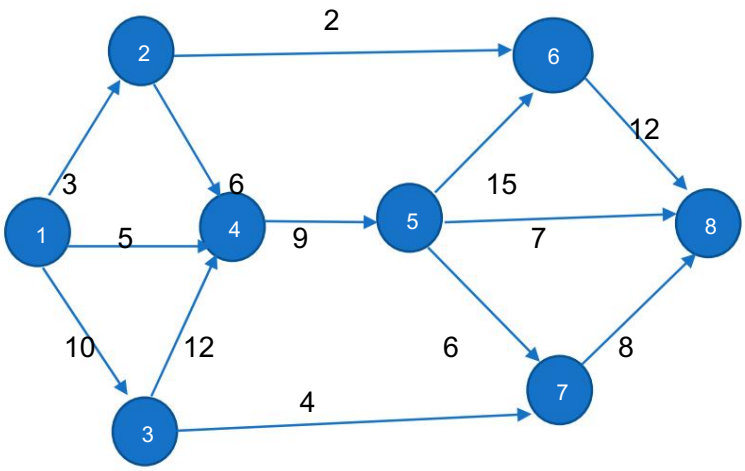


Figura 9. Red dirigida.

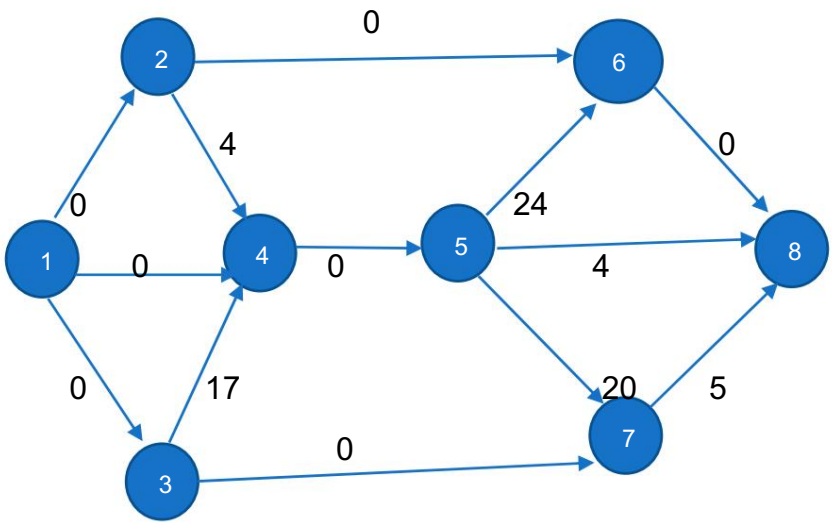


Figura 10. La red final reconstruida después de 7 iteraciones.

La solución óptima para la red en la Figura 9 se identificó como 17 unidades, y la
la ruta era 1 -> 2 -> 6 -> 8.

Utilizamos el enfoque de etiquetado de peso mínimo desarrollado en [23]. el transformado
La red después de siete iteraciones del enfoque de etiqueta de peso mínimo se muestra en la Figura 10.

Las etiquetas generadas al aplicar el enfoque de etiquetado de peso mínimo son
resumido en la Tabla 4.

Machine Translated by Google

Tabla 4. Etiquetas obtenidas por el método de etiqueta de peso mínimo.

Nodo	1	2	3	4	5	6	7	8
Etiqueta	(1,0)	(1,3)	(1,10)	(1,5)	(4,14)	(2,5)	(3,14)	(6,17)

Llegamos al nodo 8 desde el nodo 6, llegamos al nodo 6 desde el nodo 2 y llegamos al nodo 2 desde el nodo 1, lo que da como resultado que la ruta óptima sea 1 -> 2 -> 6 -> 8 y el costo mínimo siendo 17.

El costo mínimo distinto de cero es 4 en la Figura 10, para dos enlaces. Cambiando el peso asociado al enlace (2,4) no genera un segundo camino, pero sí cambia el peso asociado al enlace (5,8) generará un segundo camino. Podemos generar dos alternativas. caminos cuando el enlace (6,8) se incrementa en 4 unidades. Cuando el problema se resuelva nuevamente, Tenemos dos caminos, el que da lo mejor y el segundo lo mejor.

El mejor camino es, como se indicó anteriormente, un costo de 17, y la ruta es 1 a 2 a 6 a 8.

La segunda mejor ruta tendrá un costo de 21 y la ruta será de 1 a 4 a 5 a 8.

Aumentar el peso del enlace de (2,6) o (6,8) en 4 unidades dará lugar a dos alternativas soluciones.

La idea principal es aumentar la longitud de la solución óptima actual en 4 unidades, y si podemos establecer dos caminos óptimos alternativos con una longitud total de 21 unidades después de aumentar la longitud de la ruta óptima actual en 4 unidades, la ruta anterior es óptima y la otra es un nuevo camino que tiene una longitud igual a 21. El camino anterior será la mejor solución y el otro camino será la segunda mejor solución. Por tanto, a partir de estos dos caminos alternativos, Establecemos las soluciones óptimas y la segunda mejor para el problema de la ruta más corta.

Esta idea se puede extender fácilmente a la red no dirigida utilizando el etiquetado enfoque discutido en [24].

4.2. Ilustración numérica para la segunda mejor solución a un problema de tarea

Reconsideramos el problema de asignación considerado por Murthy [3]. la tarea El problema se presenta en la Tabla 5.

Tabla 5. Valores de costos de asignación para la Ilustración de Murthy.

R\C	1	2	3	4	5	6	7	8	9V ↓1	10
1 →	7	51	52	87	38	60	74	66	0	20
2	50	12	0	64	8	53	0	46	76	42
3	27	77	0	18	22	48	44	13	0	57
4	62	0	3	8	5	6	14	0	26	39
5	0	97	0	5	13	0	41	31	62	48
6	79	68	0	0	15	12	17	47	35	43
7	76	99	48	27	34	0	0	0	28	0
8	0	20	9	27	46	15	84	19	3	24
9	56	10	45	39	0	93	67	79	19	24
10	27	0	39	53	46	24	69	46	23	1

→ y ↓ indican una línea horizontal y una vertical.

La solución óptima con un costo total de 0 viene dada por las siguientes asignaciones, que se denota por A (1), como se indica a continuación:

$$A(1) = \{(1,9), (2,7), (3,3), (4,8), (5,6), (6,4), (7,10), (8,1), (9,5), (10,2)\}.$$

Estas asignaciones en la Tabla 5 se muestran en rojo. Se dibujan líneas verticales para cubrir el cero. entradas en las columnas 9, 3, 4, 1, 5, 2 y 7. Las filas están cubiertas por tres líneas horizontales 4, 5 y

Machine Translated by Google

7. Para la segunda mejor tarea, Murthy resolvió 9 problemas estableciendo una tarea de la solución óptima en un tiempo igual a ∞. Por ejemplo, si establecemos el vínculo (1,9) = ∞, el problema se puede resolver nuevamente. De manera similar, Murthy resolvió 9 problemas de asignación, que Murthy resumió en [3] de la siguiente manera:

$$M1 = \{\overline{(1, 9)}\} = 10,$$
$$M2 = \{(1, 9), \overline{(2, 7)}\} = 14$$
$$M3 = \{(1, 9), (2, 7), (3, 3)\} = \text{Costo } 14$$
$$M4 = \{(1, 9), (2, 7), (3, 3), (4, 8)\} = \text{Costo } 1$$
$$M5 = \{(1, 9), (2, 7), (3, 3), (4, 8), (5, 6)\} = \text{Costo } 12$$
$$M6 = \{(1, 9), (2, 7), (3, 3), (4, 8), (5, 6), (6, 4)\} = \text{Costo } 53$$
$$M7 = \{(1, 9), (2, 7), (3, 3), (4, 8), (5, 6), (6, 4), (7, 10)\} = \text{Costo } 45$$
$$M8 = \{(1, 9), (2, 7), (3, 3), (4, 8), (5, 6), (6, 4), (7, 10), (8, 1)\} = \text{Costo } 47$$
$$M9 = \{(1, 9), (2, 7), (3, 3), (4, 8), (5, 6), (6, 4), (7, 10), (8, 1) (9, 5)\} = \text{Costo } 56$$

De lo anterior, concluyó que la segunda mejor solución corresponde a la cadena M4. Esta cadena de asignaciones está dada por

A (2) = {(1,9), (2,7), (3,3), (4,2), (5,6), (6,4), (7,8), (8 ,1), (9,5), (10,10)} con costo total 1.

Ahora aplicamos el concepto de crear una alternativa. De la Tabla 5, el elemento mínimo distinto de cero es 1 en la celda (10,10). En la columna 10 y la fila 10, tenemos una asignación en la solución óptima indicada por la cadena A (1). Estas asignaciones están en las celdas (7,10) y (10,2). Para crear una alternativa, el costo asociado con el elemento (7,10) se incrementa en una unidad o el costo asociado con el elemento (10,2) se incrementa en una unidad. Por lo tanto, la columna 10 o la fila 10 no tendrán un número cero y, al restar el mínimo, crearemos dos elementos cero en las celdas (10,2) y (10,10) de la fila 10, lo que dará lugar a dos soluciones alternativas. Un caso similar se dará con un aumento del coste en la celda (7,10).

Las dos soluciones óptimas alternativas conducirán a la mejor y la segunda mejor solución.

Primera solución: {(1,9), (2,7), (3,3), (4,2), (5,6), (6,4), (7,8), (8,1), (9.5), (10,10)}.

Segunda solución: {(1,9), (2,7), (3,3), (4,8), (5,6), (6,4), (7,10), (8,1), (9,5), (10,2)}.

Con respecto a los datos de la Tabla 5, la primera solución es la segunda mejor con un costo total de 1. La segunda solución es la solución óptima con un costo total de 0, como se obtiene en [3].

El desafío sigue siendo encontrar la K° mejor cuando K > 2.

5. Comentarios finales

En este artículo, hemos considerado algunos problemas de optimización combinatoria de redes, que se transformaron en una forma que se volvió computacionalmente manejable. La idea central detrás de estas transformaciones es el desarrollo de una alternativa y, al utilizar esa alternativa, obtuvimos la solución deseada. Los problemas abordados en este artículo son los siguientes: (1) El problema del viajante; (2) La ruta a través de K nodos especificados; (3) La solución óptima ordenada, es decir, la segunda mejor solución para la ruta más corta y el problema de asignación.

Todos los enfoques utilizan la solución alternativa, que se ha creado de forma diferente en cada situación y se utiliza para reducir la carga computacional. El aspecto combinatorio del viajante y el camino a través de nodos específicos se redujo a una forma manejable mediante el desarrollo del árbol de expansión mínimo restringido por índice.

La segunda mejor solución para la ruta más corta y el problema de asignación se obtuvo aumentando la solución óptima a una posición controlada y estratégica para obtener dos soluciones óptimas, una de las cuales era la óptima y la otra era la segunda mejor solución.

Nuestro enfoque también planteó algunas preguntas que merecen una mayor investigación.

1. ¿El problema mínimo del recorrido del vendedor ambulante es realmente NP Difícil?
2. ¿El problema del enrutamiento restringido es computacionalmente exigente?

Machine Translated by Google

3. Las optimizaciones ordenadas requieren más investigación para encontrar la K th mejor cuando K ≥ 3 para los dos casos analizados en este artículo y para otras situaciones.

Este artículo se ha limitado a algunos problemas clásicos debido a la larga asociación de los autores con ellos; sin embargo, sería deseable investigar la aplicación de estas ideas a los problemas de enrutamiento considerados por [25-27].

Los enfoques revisados en este artículo necesitan más investigación para establecer la eficiencia computacional de diversas situaciones discutidas en este artículo. Esto está bajo investigación y se informará cuando esté listo.

Contribuciones de los autores: Conceptualización, SK y EM; Software, SK y EM; Análisis formal, EM; Investigación, SK; Curación de datos, EM; Redacción: borrador original, SK y EM. Todos los autores han leído y aceptado la versión publicada del manuscrito.

Financiamiento: Esta investigación no recibió financiamiento externo.

Declaración de la Junta de Revisión Institucional: No aplicable.

Declaración de Consentimiento Informado: No aplicable.

Declaración de disponibilidad de datos: Las contribuciones originales presentadas en el estudio están incluidas en el artículo.

Agradecimientos: Los autores expresan su más sincero agradecimiento a todos los revisores anónimos por sus sugerencias y comentarios constructivos.

Conflictos de intereses: Los autores declaran no tener ningún conflicto de intereses.

Referencias

1. Nemhäuser, G.; Wolsey, LA Optimización entera y combinatoria; John Wiley & Sons: Hoboken, Nueva Jersey, EE. UU., 1988. [CrossRef]
2. Kumar, S.; Wagner, D. Algunos algoritmos para resolver problemas de programación matemática de puntos extremos. Ópera de Nueva Zelanda. Res. J. 1979, 7, 128–149.
3. Murty, KG Un algoritmo para clasificar todas las tareas en orden de costo creciente. Ópera. Res. 1968, 16, 682–687. [Referencia cruzada]
4. Bellman, R.; Dreyfus, S. Programación dinámica aplicada; Princeton University Press: Princeton, Nueva Jersey, EE. UU., 1962.
5. Hastings, Nueva Jersey Programación dinámica con aplicaciones de gestión; O Serie; Sociedad Estadounidense de Matemáticas: Butterworth, Malasia; Londres, Reino Unido, 1973.
6. Saksena, JP Programación dinámica; Publicaciones Siddarth: Delhi, India, 2006.
7. Dorado, B.; Raghavan, S.; Wasil, E. El problema de las rutas de vehículos: últimos avances y nuevos desafíos; Springer Ciencia + Negocios Medios: Berlín/Heidelberg, Alemania, 2008; ISBN 978-0-087-77777-1. [Referencia cruzada]
8. Marrekchi, E.; Besbes, W.; Dhouib, D.; Demir, E. Una revisión de los avances recientes en la literatura de investigación de operaciones sobre los problemas de enrutamiento ecológico y sus variaciones. Ana. Ópera. Res. 2021, 304, 529–574. [Referencia cruzada]
9. Anupam, G. MST determinista, algoritmos avanzados, CMU. Disponible en línea: http://www.cs.cmu.edu/~anupamg/advalgos15/conferencias/conferencia01.pdf (consultado el 1 de abril de 2024).
10. Kruskal, JB Sobre el subárbol de expansión más corto de un gráfico y el problema del viajante. Proc. Soy. Matemáticas. Soc. 1956, 7, 48–50. [Referencia cruzada]
11. Prim, R. Red conectada más corta y algunas generalizaciones. Sistema de campana. Tecnología. J. 1957, 36, 1389–1401.
12. Munapo, E.; Kumar, S.; Lesaoana, M.; Nyamugure, P. Un árbol de expansión mínimo con índice <=2. Agosto. Soc. Ópera. Res. Toro. 2016, 34, 1–14.
13. Hoffman, KL; Pedberg, M.; Rinaldi, G. El problema del viajante; Editorial académica Kluer: Amsterdam, The Nerther-tierras, 2001. [CrossRef]
14. Korté, B.; Vygen, J. Optimización combinatoria: teoría y aplicaciones; Capítulo 21; Springer: Berlín/Heidelberg, Alemania, 2006; págs. 501–535.
15. Laporte, G. El problema del viajante: una descripción general de los algoritmos exactos y aproximados. euros. J. Ópera. Res. 1992, 59, 231–247. [Referencia cruzada]
16. Kumar, S.; Munapo, E.; Lesaoana, M.; Nyamugure, P. ¿Es realmente difícil el problema del viajante? Capítulo 3 en Ingeniería y Tecnología: Innovaciones e investigaciones recientes; Mathani, A., Ed.; Casa de Publicaciones de Investigación Internacional: Delhi, India, 2017; págs. 37–58, ISBN 978-93-86138-06-4.
17. Kumar, S.; Munapo, E.; Lesaoana, M.; Nyamugure, P. Un enfoque basado en un árbol de expansión mínima para el problema del viajante . Opsearch 2018, 55, 150–164.

Machine Translated by Google

18. Kumar, S.; Munapo, E.; Siguake, C.; Al-Rabeeah, M. El árbol de expansión mínimo con índice de nodo <= 2 es equivalente al recorrido mínimo del viajante de comercio. En Matemáticas en Ciencias de la Ingeniería: Teorías, Tecnologías y Aplicaciones Novedosas; Capítulo 8; Serie de Ingeniería Matemática, Fabricación y Ciencias de la Gestión; Ram, M., Ed.; CRC Press: Boca Raton, FL, EE. UU., 2020. Disponible en línea: https://www.crcpress.com/Mathematical-Engineering-Manufacturing-and-Management-Sciences/ serie-de-libros/CRCMEMMS (consultado el 1 de abril de 2024).

19. Kalaba, R. Sobre algunos análisis combinatorios de redes de comunicación. Proc. Síntoma. Aplica. Matemáticas. 1960, 10. [Referencia cruzada]
20. Saksena, JP; Kumar, S. El problema de enrutamiento con K nodos especificados. Ópera. Res. 1966, 14, 909–913. [Referencia cruzada]
21. Kumar, S.; Munapo, E.; Nyamugure, P.; Tawanda, T. Ruta a través de K enlaces especificados: una heurística que utiliza la amplitud mínima enfoque de árbol. Gráfico. Era Univ. J. 2023, 11, 221–238.
22. Kumar, S.; Munapo, E.; Jones, BC Una ecuación entera controlaba el camino descendente hacia un programa entero puro. Indio J. Matemáticas. 2007, 49, 211–237.
23. Munapo, E.; Jones, antes de Cristo; Kumar, S. Un método de etiqueta de peso mínimo y su aplicación a la red CPM. ORION S. Afr. J. Ópera. Res. 2008, 24, 37–48. [Referencia cruzada]
24. Kumar, S.; Munapo, E.; Ncube, O.; Sigauke, C. Un método de etiquetado de peso mínimo para determinar el camino más corto en un Red no dirigida. Int. J. Sistema. Asegurar. Ing. Gestionar. 2013, 4, 13-18.
25. Karimi-Managham, M.; Mohammadi, M.; Meyer, AM; Karimi-Managham, E.; Talbi, G. El aprendizaje automático al servicio de la metaheurística para la resolución de problemas de optimización combinatoria: un estado del arte. euros. J. Ópera. Res. 2022, 296, 393–422. [Referencia cruzada]
26. Labidi, H.; Azzouna, N.; Hassine, K.; Goudir, MS Un algoritmo genético mejorado para resolver rutas de vehículos multiobjetivo problema con consideraciones medioambientales. Computación de procedimiento. Ciencia. 2023, 225, 3866–3875. [Referencia cruzada]
27. Sabeeniano, RS; Kumar, V.; Ravi, G.; Pudasaini, A. Una aplicación para compartir carga en camiones. Int. J. Tecnología reciente. Ing. (IJRTE) 2019, 8, 3203–3207. [Referencia cruzada]

Descargo de responsabilidad/Nota del editor: Las declaraciones, opiniones y datos contenidos en todas las publicaciones son únicamente de los autores y contribuyentes individuales y no de MDPI ni de los editores. MDPI y/o los editores renuncian a toda responsabilidad por cualquier daño a personas o propiedad que resulte de cualquier idea, método, instrucción o producto mencionado en el contenido.

Machine Translated by Google

AppliedMath

MDPI

Artículo

La mecánica que sustenta la computación no determinista en las redes neuronales corticales

Elizabeth A. Stoll

Instituto Occidental de Estudios Avanzados, Denver, CO 80202, EE. UU.; stoll@westerninstitute.org

Resumen: Las neuronas corticales integran señales ascendentes y ruido eléctrico aleatorio para generar resultados de señalización, lo que conduce a patrones de actividad estadísticamente aleatorios. Sin embargo, clásicamente, la neurona se modela como una unidad computacional binaria, que codifica la entropía de Shannon. Aquí, el potencial de membrana neuronal se modela como una función del comportamiento iónico inherentemente probabilístico. En este nuevo modelo, cada neurona calcula la probabilidad de pasar de un estado apagado a un estado encendido, codificando así la entropía de von Neumann. Los estados puros de los componentes se integran en una cantidad física de información, y la derivada de esta distribución de probabilidad de alta dimensión produce valores propios en todo el sistema cuántico de múltiples escalas. De acuerdo con el teorema de Hellman-Feynman, la resolución del estado del sistema va acompañada de un cambio espontáneo en la distribución de carga, por lo que este estado definido del sistema se convierte instantáneamente en el pasado a medida que emerge una nueva distribución de probabilidad. Este modelo mecanicista produce predicciones comprobables sobre la longitud de onda de la energía libre liberada tras la compresión de información y la relación temporal de estos eventos con los resultados fisiológicos. En general, este modelo demuestra cómo las neuronas corticales podrían lograr resultados de señalización no deterministas mediante un proceso computacional de detección de coincidencias ruidosas.

Palabras clave: neurona cortical; computación neuronal; cálculo termodinámico; codificación probabilística; sistemas cuánticos de dos estados; información cuántica; entropía de von Neumann

check for updates

Cita: Stoll, EA La mecánica Respaldo de la computación no determinista en redes neuronales corticales. Matemáticas aplicadas 2024, 4, 806–827. https://doi.org/10.3390/appliedmath4030043

Editores académicos: Thomas Woolley y Gaige Wang

Recibido: 25 de noviembre de 2023

Revisado: 25 de mayo de 2024

Aceptado: 30 de mayo de 2024

Publicado: 26 de junio de 2024

Copyright: © 2024 por el autor. Licenciatario MDPI, Basilea, Suiza. Este artículo es un artículo de acceso abierto. distribuido bajo los términos y condiciones de los Creative Commons Licencia de atribución (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

1. Introducción

Las neuronas corticales tienen resultados de señalización altamente impredecibles [1], en contraste con las neuronas espinales, cuyos resultados de señalización se predicen fácilmente analizando las entradas ascendentes [2]. Las neuronas corticales permiten que el ruido eléctrico aleatorio controle los resultados de la señalización, y las fluctuaciones espontáneas por debajo del umbral en el potencial de membrana contribuyen significativamente a la probabilidad de un potencial de acción [3,4]. De hecho, estas células mantienen activamente un " estado cortical superior", rondando el umbral del potencial de acción y permitiendo que los eventos estocásticos impulsen los resultados de la señalización [5]. Este proceso de detección de coincidencias ruidosas da como resultado patrones de disparo probabilísticos, que pueden modelarse con estadísticas bayesianas [6], modelos de conectividad aleatoria [7], análisis fanofactor de la variabilidad entre picos [8] o modificando las ecuaciones de Hodgkin-Huxley para tener en cuenta el ruido eléctrico [9-11]. En particular, los modelos más confiables de neuronas corticales se componen de ecuaciones diferenciales estocásticas, que proporcionan un potencial de membrana analógico basado en la probabilidad de apertura del canal iónico y conducen a un comportamiento de pico tras la activación del umbral [12]. Este método modela el "ruido" en la sinapsis como un proceso de Ornstein-Uhlenbeck [13] o aplicando las ecuaciones de Fo

Dado el éxito de estos modelos computacionales inherentemente probabilísticos, vale la pena revisar si las transiciones de estado en las neuronas corticales deberían modelarse según supuestos clásicos. De hecho, el movimiento aleatorio de unos pocos iones puede empujar a una neurona cortical por encima del umbral del potencial de acción [15,16], los intervalos entre picos en el nivel de una sola unidad son estadísticamente aleatorios [3,4], y la codificación de la población en el nivel de la red neuronal es estadísticamente aleatorio [6-8]. Los eventos de escala cuántica afectan materialmente los resultados de escala macro a nivel de la unidad computacional [9-11], y los modelos probabilísticos replican efectivamente la actividad de las redes neuronales corticales [12-14]. Además, los enfoques clásicos no pueden explicar la

Machine Translated by Google

extraordinaria eficiencia energética del cerebro, mientras que un modelo de computación cuántica puede hacerlo [17]. Para comprender la mecánica subyacente a un proceso computacional en el que la incertidumbre cuántica se preserva in situ, será útil desarrollar un marco matemático para un sistema cuántico de múltiples escalas.

Clásicamente, una neurona que se activa se considera una unidad binaria, siempre en un estado activado o desactivado. Aquí, la neurona se modela como un sistema cuántico de dos estados, con cierta probabilidad de cambiar de un estado apagado a uno encendido. Este estado neuronal depende de la posición inherentemente probabilística y del impulso de cada ion en las proximidades. En este modelo cuántico de múltiples escalas, el voltaje de una neurona cortical es una función de todos los estados puros de los componentes, y un estado óptimo del sistema en el presente contexto se selecciona a partir de una distribución de probabilidad, durante un proceso de generación y compresión de información cuántica. Este proceso computacional produce valores propios para todos los vectores de estado e inmediatamente restaura la incertidumbre en todo el sistema, lo que provoca el siguiente ciclo computacional. Este modelo teórico ofrece un mecanismo mediante el cual las neuronas corticales podrían lograr resultados de señalización no deterministas a través de un proceso de computación cuántica a temperatura ambiente y produce predicciones específicas que pueden probarse en el laboratorio.

2. Métodos

2.1. Modelado de la neurona cortical como un sistema cuántico de dos estados

Durante el estado activo, las neuronas corticales permanecen en su umbral de potencial de acción, lo que permite que tanto las señales ascendentes como el ruido eléctrico aleatorio provoquen un resultado de señalización. Entonces, si bien una neurona se interpreta clásicamente como una puerta lógica binaria en un estado 'encendido' o 'apagado', codificada como 1 o 0, también podría describirse como si tuviera cierta probabilidad de convertirse a un estado 'encendido' o permanecer en ese estado. un estado "apagado".

En este nuevo enfoque, una neurona cortical integra señales ascendentes con ruido eléctrico aleatorio, definiendo su estado de voltaje en función del tiempo, a medida que el sistema se perturba. La neurona comienza en un estado apagado φ, sin disparar un potencial de acción, y con el tiempo t, alcanza otro estado χ. Y así, durante un período de tiempo, de t0 a t, el estado de la neurona evoluciona de φ a χ. La trayectoria temporal tomada de un estado a otro está dada por:

$$\chi|U(t, t0)|\varphi \quad . \tag{1}$$

La probabilidad de un cambio de estado se puede representar de alguna manera:

$$\sum \quad \chi|k \quad k|U(t, t0)|j \quad j|\varphi \quad , \tag{2}$$

tal que U está completamente descrito por los estados base k y j, que representan el estado apagado inicial de la neurona y el estado de la neurona después del paso de tiempo, respectivamente:

$$k|U(t, t0)|j \quad , \tag{3}$$

El intervalo de tiempo puede entenderse como t = t0 + Δt, por lo que identificar el estado de Se puede entender que la neurona χ en el momento t toma un camino de un estado a otro:

$$\chi|\psi(t0 + \Delta t) \quad = \quad \chi|U(t0 + \Delta t, t0)|\psi(t0) \quad . \tag{4}$$

Si Δt = 0, no puede haber cambio de estado. En este caso:

$$|\psi(t0 + \Delta t) \quad = |\psi(t0) \quad . \tag{5}$$

En cualquier otro caso, el estado de la neurona en el instante t viene dado por la base ortonormal estados k y j, con amplitudes de probabilidad:

$$Ck\ (t0 + \Delta t) = \quad k|\psi(t0) \tag{6}$$

Machine Translated by Google

Y:

$$Cj(t0 + \Delta t) = \quad j|\psi(t0) \quad . \tag{7}$$

Por lo tanto , el estado neuronal $|\psi$ en el momento t puede describirse como un vector de estado normado ψ, en una superposición de dos estados base ortonormales k y j, con amplitudes de probabilidad Ck y Cj . La suma de los módulos al cuadrado de todas las amplitudes de probabilidad es igual a 1:

$$|ck|^2 + |Cj|^2 = 1. \tag{8}$$

Dado que la neurona inicia la evolución temporal en el estado $|\psi(t0) = j$, su estado probable en el tiempo t viene dado por:

$$k|\psi(t0 + \Delta t) \quad = \quad k|U(t0 + \Delta t, t0)|\psi(t0) \quad . \tag{9}$$

Esta ecuación también se puede escribir en forma expandida como la suma de todas las probabilidades de transición:

$$k|\psi(t) \quad = \sum j \quad k|U(t0 + \Delta t, t0)|j \quad j|\psi(t0) \quad . \tag{10}$$

Para el vector de estado $\psi(t)$, la probabilidad de un cambio de estado en el tiempo t se describe mediante la Matriz U, Ukj(t):

$$Ukj(t) = \quad k|U(t0 + \Delta t, t0)|j \quad . \tag{11}$$

Y así, todas las amplitudes de probabilidad dependen de la cantidad de tiempo que ha pasado, Δt:

$$Ck\ (t0 + \Delta t) = \sum j\ Ukj(t)Cj(t0)\ . \tag{12}$$

La neurona sufre perturbaciones a lo largo del tiempo Δt. Si $\Delta t = 0$, no puede haber cambio de estado y k = j. Si $\Delta t > 0$, existe cierta probabilidad de un cambio de estado, donde k = j. Como tal, el sistema cuántico de dos estados se describe mediante el delta de Kronecker δkj:

$$\delta kj = \quad k|j \quad = \quad \begin{matrix} 1, \text{ si } k = j \\ 0, \text{ si } k \quad = j \end{matrix} \tag{13}$$

Cada uno de los coeficientes de la matriz U Ukj difiere de δkj en alguna cantidad proporcional a Δt, tal que:

$$Ukj(t) = \delta kj + Wkj\Delta t\ . \tag{14}$$

donde $Wkj = -(i/\bar{h})\hat{H}\ kj(t)$, donde el término del operador hamiltoniano $\hat{H}\ kj(t)$ representa las derivadas temporales de cada uno de los coeficientes de Ukj:

$$Ukj(t) = \delta kj - \frac{i}{\bar{h}}\hat{H}\ kj(t)\Delta t\ . \tag{15}$$

Por tanto, la amplitud de probabilidad Ck (t) en el instante t viene dada por:

$$Ck\ (t0 + \Delta t) = \sum_j [\delta kj - \frac{i}{\bar{h}}\hat{H}\ kj(t)\Delta t]Cj(t0)\ . \tag{16}$$

Dado que la suma de $\delta kjCj(t0) = \psi(t0)$, la última ecuación se simplifica a:

$$Ck\ (t0 + \Delta t) - Ck\ (t0) = -\frac{i}{\bar{h}}\sum_j [\hat{H}\ kj(t)\Delta t]\psi(t0)\ . \tag{17}$$

Al dividir ambos lados de la ecuación por Δt, resulta evidente que cualquier cambio de estado es una suma de todas las posibles perturbaciones que afectan al sistema bajo investigación, desde su condición inicial en el tiempo t0, incluidos los eventos aleatorios. Nuevamente, Ck (t) es la probabilidad

Machine Translated by Google

amplitud i|ψ(t0) de encontrar el vector de estado ψ en uno de los estados base k = j o k = j en el tiempo t. Y así, la derivada temporal de esta función de probabilidad produce el camino tomado:

$$\frac{Ck\,(t0 + \Delta t) - Ck\,(t0)}{\Delta t} = -\frac{i}{h^{-}} \sum_{j} H\hat{}\,kj(t)\psi(t0)\,. \tag{18}$$

Por lo tanto, el sistema se describe mediante la ecuación de Schrödinger dependiente del tiempo, con cualquier cambio de estado relacionado con un cambio en la distribución de energía:

$$dt\,\frac{d\psi}{} = H\hat{}\,\psi\,.\,ih^{-} \tag{19}$$

Donde i es la unidad imaginaria, $h^{-} = h/2\pi$ es la constante de Planck reducida, $H\hat{}$ es el operador hamiltoniano, que corresponde a la energía total del sistema, y ψ es el vector propio que describe el sistema cuántico después del tiempo t. Esta ecuación describe el estado propio de un sistema en función de la cantidad de tiempo t que ha pasado y la cantidad de energía disponible para su redistribución a través del sistema, dada por el hamiltoniano $H\hat{}$. Esta ecuación permite definir el estado de un sistema como una distribución de probabilidades que representan los posibles caminos tomados desde la última vez que se definió el estado del sistema. Combinando las ecuaciones (18) y (19), encontramos que la probabilidad de un cambio de estado es igual a la suma de todas las derivadas temporales del vector de estado normado:

$$\frac{Ck\,(t0 + \Delta t) - Ck\,(t0)\,\Delta t}{} = \sum dt\,\frac{d\psi}{}\,. \tag{20}$$

2.2. Modelado del potencial de membrana de la neurona cortical en función de los estados puros de los componentes

El potencial de membrana dinámico de una neurona cortical ha sido modelado previamente con la ecuación de Fokker-Planck [14], con funciones propias φi y coeficientes αi formando un conjunto completo en el que se puede expandir la distribución del potencial de membrana P(V, t) :

$$P(V, t) = \sum_{i} \alpha i\varphi i(V)e^{-\lambda i\, t/\tau m}\,. \tag{21}$$

Fundamentalmente, todos los valores propios son reales y estrictamente negativos, y las funciones propias obedecen a L(φi) = −λiφi y λi > 0. Un componente clave de este modelo es τm, la constante de tiempo de membrana de la neurona, dada por:

$$\tau m = \frac{gL}{\text{centímetro}} + veae + viai\,, \tag{22}$$

donde ve y vi son las entradas sinápticas excitadoras e inhibidoras, respectivamente; ae y ai son las fuerzas sinápticas excitadoras e inhibidoras, respectivamente; cm es la capacitancia de la membrana; y gL es la conductancia de la membrana en el contexto de una corriente de fuga, dada por:

$$gL = \frac{\text{Illinois}}{V - VL}\,. \tag{23}$$

La corriente total entonces viene dada clásicamente por la suma:

$$In = gL(t)(V - VL) + ge(t)(V - Ee) + gi(t)(V - Ei)\,, \tag{24}$$

donde Ee y Ei son los potenciales de inversión de las sinapsis excitadoras e inhibidoras, respectivamente. En las neuronas espinales, las entradas sinápticas excitadoras e inhibidoras dominan cualquier cambio en el potencial de membrana, y la corriente de fuga no contribuye materialmente a los resultados de la señalización. Mientras tanto, en las neuronas corticales, las entradas excitadoras e inhibidoras están equilibradas,

Machine Translated by Google

tal que predominen las corrientes de fuga. Para modelar la contribución de iones individuales a la corriente de fuga, podemos modelar la densidad de corriente de forma probabilística:

$$IL = \frac{}{q\ 2m}(\psi \quad) , s\psi - \psi s\psi \tag{25}$$

donde m es la masa de cada carga puntual q. En este modelo, el estado neuronal $|\psi n$ evoluciona con el tiempo, y el resultado de la señalización en el tiempo t es función de todos los estados iónicos. En particular, el estado de cada ion en el sistema $|\psi i$ también evoluciona con el tiempo. La ecuación de Schrödinger dependiente del tiempo para la posición de cada ion es:

$$yo^{-}\frac{d}{dt}|\psi(r, t) = H\hat{} |\psi(r, t) . \tag{26}$$

Y la ecuación de Schrödinger dependiente del tiempo para el momento de cada ion es:

$$yo^{-}\frac{d}{dt}|\psi(s, t) = H\hat{} |\psi(s, t) . \tag{27}$$

La posición de cada ion viene dada por una distribución de amplitudes de probabilidad a lo largo de los ejes x, y y z . Y así, el impulso de cada ion s lo proporciona la derivada de la función de onda con respecto a su posición r:

$$ih^{-}\frac{d\psi}{dr} + z\,\partial x\,\partial y\,\frac{\partial\,\partial\,\partial}{\partial z} = ih^{-}\frac{}{}x + \frac{}{}y = ih^{-} . \tag{28}$$

Como la masa, la energía y la carga eléctrica se conservan, podemos emplear la ecuación de continuidad:

$$\frac{\partial\rho}{} = - \quad r \cdot IL , \partial t \tag{29}$$

donde $\rho = |\psi(r, t)|^2$. Por lo tanto:

$$\frac{\partial|\psi(r, t)|^2}{\partial t} = - \quad r \cdot [\frac{}{q\ 2m}(\psi \quad s\psi - \psi s\psi \quad)] . \tag{30}$$

Dado que $r = -s/ih^{-}$:

$$\frac{\partial|\psi(r, t)|^2}{\partial t} = \psi \quad \frac{qs2}{2mih^{-}}\psi - \psi\,\psi\,\frac{qs2}{2mih^{-}} . \tag{31}$$

La ecuación de Schrödinger dependiente del tiempo para describir el hamiltoniano del sistema, con respecto al impulso, es:

$$ih^{-}\frac{\partial\psi}{\partial t} = (\frac{}{2m})\psi . \tag{32}$$

Y su conjugado complejo es:

$$-ih^{-}\frac{\partial\psi}{\partial t} = (\frac{s^2}{2m})\psi . \tag{33}$$

Por lo tanto:

$$\frac{\partial\rho}{\partial t} = \frac{\partial|\psi(r, t)|^2}{\psi\,\partial t\,\partial t} \frac{q\partial\psi}{} \frac{q\partial\psi}{} = \psi + . \tag{34}$$

Y así, la suma de todas las corrientes de fuga sinápticas para una neurona determinada en el tiempo t es:

$$IL = \frac{q\partial\rho/\partial t\ \partial\psi/}{\partial r} . \tag{35}$$

Machine Translated by Google

Cuando las entradas excitadoras e inhibidoras están equilibradas, durante un estado cortical elevado, la densidad de corriente de una neurona se describe efectivamente mediante la suma de todas las corrientes de fuga:

$$En = \frac{q\partial\rho/\partial t}{\partial\psi/\partial r}. \quad (36)$$

Por lo tanto, la probabilidad de que la neurona cruce el umbral del potencial de acción (Vth) está dada por la relación entre la corriente y el voltaje, y cualquier valor propio negativo para ψ(r, t) incita a la neurona a regresar a su potencial de reposo:

$$V - Vth = \frac{1}{gL} \frac{q\partial\rho/\partial t}{\partial\psi/\partial r}, \quad (37)$$

2.3. Modelado de la información codificada por el sistema multiestado, en términos de la entropía de von

Neumann Durante los estados corticales elevados, las neuronas mantienen activamente un potencial de reposo justo cerca del umbral de voltaje para disparar un potencial de acción [5], lo que permite que la fuga de iones estocásticos bloquee un cambio de estado en la unidad computacional [3,4]. En el momento t, la neurona ha alcanzado el umbral de voltaje o no; existe cierta probabilidad de que se produzca cualquiera de los dos resultados. La neurona es un sistema cuántico de dos estados, que depende de la posición inherentemente probabilística y del impulso de cada ion en la vecindad. Para modelar el sistema a nivel de neurona o de ion, ρx se define como la probabilidad de que un objeto pase de un estado puro a otro:

$$\rho x := |\psi x \quad \psi x|. \quad (38)$$

La matriz de densidad ρ está compuesta por un conjunto de estados puros mutuamente ortogonales ρx, cada uno de los cuales tiene cierta probabilidad de ocurrir px:

$$\rho = \sum px\rho x = \sum px|\psi x \quad \psi x|, \quad (39)$$

Debido a que la neurona cortical es sensible a eventos de nivel cuántico, permanece suspendida en un estado de incertidumbre, codificando físicamente "información" (la entropía de von Neumann del sistema, o la suma mixta de todos los estados puros que la componen). La cantidad de información codificada por la neurona se calcula rastreando el volumen de amplitudes de probabilidad a través de una matriz de densidad de alta dimensión:

$$S(\rho) = -Tr(\rho \ln \rho). \quad (40)$$

Aquí, el macroestado del sistema es la suma mixta de todos los microestados puros que lo componen, o la suma mixta de todos los productos externos multiplicada por sus probabilidades de transición. El producto interno ay ψ proporciona la probabilidad que el vector de estado |ψx asigna al vector propio ay . Como tal, la probabilidad de medir un determinado valor propio ay es igual a:

$$p(ay) = ay \rho ay. \quad (41)$$

El objeto está en algún estado en el momento t0 y tiene cierta probabilidad de ser encontrado en otro estado en el momento t, después de haber sido definido transitoriamente como una distribución de estados posibles. La probabilidad de encontrar un sistema en cualquier estado propio se puede calcular aplicando la regla de Born, que iguala el producto interno del vector de estado y su valor esperado con la probabilidad de transición a un estado real particular [18]. Esta regla establece que cualquier medida del observable tiene una probabilidad p(ψa) de ser igual a un valor exacto a, en el tiempo t para el vector de estado ψ:

$$p(\psi a(t)) = |\quad a|\psi(t)\quad|^2. \quad (42)$$

El cuadrado del valor absoluto de esta función de onda es un número real, pero la función de onda en sí es una amplitud de probabilidad de valor complejo, que existe a lo largo de un eje.

Machine Translated by Google

ortogonal a todos los valores propios reales. Cada valor propio definido a lo largo de este eje depende de la cantidad de tiempo que ha pasado. Si hay suficientes restricciones, el estado puro de cada componente se puede definir resolviendo el hamiltoniano en el tiempo t. Si no hay suficientes restricciones , algunos estados puros de los componentes no se resolverán y el sistema producirá una cantidad neta de entropía durante ese ciclo computacional.

2.4. Generación de una distribución de posibles estados del sistema a partir de la incertidumbre cuántica

Para comprender la mecánica subyacente a este proceso, se puede conceptualizar un cálculo de posibles estados del sistema a través de la red neuronal modelando la distribución de Poisson de la densidad de corriente para cada ion en el sistema, en relación con cada región de la membrana neuronal, generando así una distribución de caminos posibles. Pero debido a que estos eventos no ocurren con la misma probabilidad o independientemente del estado anterior del sistema, es más apropiado modelar la distribución de probabilidad de los vectores de posición y momento como r(x) y s(x), respectivamente, cada uno definido en tres espacios espaciales. ejes, con una función de peso w(x) > 0:

$$(r,s) = \int_{-\infty}^{\infty} r'(x)s(x)w(x)dx \, . \tag{43}$$

En este sistema Sturm-Liouville, la función de peso w(x) representa matemáticamente un oscilador armónico cuántico. w(x) encarna la incertidumbre cuántica del estado del ion y genera naturalmente un espacio de Hilbert para cada posición y vector propio de momento. Si esta operación es simétrica o hermitiana, tal que:

$$\begin{aligned}(r, Ls) &= \int_{-\infty}^{\infty} L(r)'(x))s(x)w(x)dx \\ &= \int_{-\infty}^{\infty} r'(x)L(s(x))w(x)dx \\ &= (Lr,s) \, ,\end{aligned} \tag{44}$$

entonces todos los polinomios formarán conjuntos completos en el espacio de Hilbert, con valores propios reales y funciones propias ortonormales, proporcionando soluciones a la ecuación de diferenciación lineal de segundo orden dada por:

$$Lu = \lambda u \, , \tag{45}$$

En el cual, λ es una constante y L es un operador hermitiano definido por las funciones reales de x, α, β y γ:

$$L = \alpha(x) \frac{d^2}{dx^2} + \beta(x) \frac{d}{dx} + \gamma(x) \, . \tag{46}$$

La distribución de resultados para r(x) y s(x) viene dada por ρ, la suma mixta de estados mutuamente ortogonales ρx, cada uno de los cuales ocurre con cierta probabilidad px:

$$\rho(r,s) = \sum(p_{r(x)} \rho_{r(x)})(p_{s(x)} \rho_{s(x)}) \, . \tag{47}$$

La evolución del estado del sistema ρ a lo largo del tiempo t viene dada por el método de Liouville-von Ecuación de Neumann:

$$ih \frac{d}{dt} \rho = ih \sum p_x(|\psi_x\rangle\langle\psi_x| + |\psi_x\rangle\langle\psi_x|) \, , \tag{48}$$

con estados propios proporcionados por el operador hamiltoniano:

$$ih \frac{d}{dt}|\psi_x\rangle = \hat{H}|\psi_x\rangle \, . \tag{49}$$

y su conjugado hermitiano:

$$-ih \frac{d}{dt}\langle\psi_x| = \langle\psi_x|\hat{H} \, . \tag{50}$$

Machine Translated by Google

Sustituyendo el operador y el conjugado en la ecuación de Liouville-von Neumann se obtiene:

$$ih^{-}\frac{d}{dt}\rho = ih^{-}\sum(H^{\wedge}\rho x - \rho x H^{\wedge}) = [H^{\wedge}, \rho x]\ . \quad (51)$$

Debido a que el tiempo en sí es intrínsecamente incierto, la evolución temporal del estado del sistema la proporciona un operador de desplazamiento de tiempo U:

$$U = mi^{-iH^{\wedge}t}\ . \quad (52)$$

Una transformación unitaria, impulsada por este operador de cambio de tiempo, permite que la matriz de densidad para evolucionar de ρx(t0) a ρx(t):

$$\rho x(t0) \rightarrow \rho x(t) = U\rho x(t0)U^{\dagger}\ . \quad (53)$$

Someterse a esta transformación unitaria produce valores propios para cada componente en estado puro, resolviendo la posición y el impulso de cada ion y el estado de voltaje de cada neurona. Esos valores propios solo se resuelven cuando se resuelve el hamiltoniano, cuando se reduce la dimensionalidad de la matriz de densidad, cuando la función de onda colapsa, cuando se extraen correlaciones lineales o cuando se selecciona un solo estado del sistema de una distribución de posibles estados del sistema. Todos estos son procesos equivalentes, que representan la compresión de la "información" o entropía de von Neumann. Y lo que es más importante, este sistema de múltiples escalas tiene un único punto de tiempo t, en el que se resuelven tanto los estados iónicos como los estados neuronales (aunque puede permanecer cierta incertidumbre o entropía al finalizar cada ciclo computaci

2.5. Reducción de la distribución de probabilidad a un único estado de sistema observable

Debido a que la posición y el momento de cualquier ion en el tiempo t dependen de si ha interactuado con otros iones o campos eléctricos durante la evolución temporal, se debe tener en cuenta la interdependencia de los vectores propios. Al calcular las posiciones y los momentos de todos los iones dentro de un sistema de temperatura definida en relación entre sí, este enfoque garantiza que los iones no ocupen la misma posición, giro y estado de energía, lo que los haría idénticos. Este resultado haría que el universo perdiera masa y, por lo tanto, está prohibido por el principio de exclusión de Pauli [19]. Este principio de conservación se puede aplicar a cualquier sistema termodinámico, por lo que la cantidad total de energía disponible para el sistema no es mayor que la cantidad de energía almacenada en el sistema más la cantidad neta de energía que ha ingresado al sistema a lo largo del tiempo t. En el sistema nervioso central, que atrapa energía térmica para impulsar el trabajo computacional, la distribución de los estados del sistema está representada por el operador hamiltoniano:

$$yo^{-}\frac{d}{dt}|\rho(t) = H^{\wedge}|\rho(t)\ . \quad (54)$$

En un sistema lejos del equilibrio que está acoplado termodinámicamente a su entorno circundante, no se puede suponer que el movimiento browniano de los iones ocurre en el límite clásico, por lo que se debe aplicar la formulación de la mecánica cuántica de las ecuaciones de Fokker-Planck, conocidas como Modelo de Caldeira-Leggett [20-22]. Aquí, el operador hamiltoniano representa la posición r y el momento s de cada ion con masa m y carga q, moviéndose en un potencial V(r), donde ω es un oscilador armónico cuántico y c es una constante de ac

$$H^{\wedge}(r,s) = \frac{\cdot}{} + V(r) + 2m2^{-1}\sum\frac{(st - crt)^{2}}{\text{metro}} + m\omega rt^{2\,2}\ . \quad (55)$$

Dado que el estado de cada ion depende (en parte) del estado de todos los demás iones, todo el sistema multiescala debe considerarse en conjunto, con las posiciones de los iones dependientes de los voltajes de la membrana y viceversa. Esta función de onda combinada relaciona cada vector de estado ψ con el paso del tiempo t y el operador hamiltoniano $H^{\wedge}$. El espectro del operador hamiltoniano es el conjunto de posibles resultados en el tiempo t. El operador hamiltoniano $H^{\wedge}$ está relacionado

Machine Translated by Google

la función lagrangiana de la posición r, su derivada temporal r˙ y el tiempo t. Se calcula tomando la transformada de Legendre, para minimizar la acción necesaria para efectuar el cambio:

$$H\hat{}\,(r_i, s_i, t) = \sum_{i=1}^{n} s_i \dot{r}_i - L(r_i, \dot{r}_i, t)\,. \tag{56}$$

Los espacios vectoriales representados por r˙ y s˙ se generan mediante la suma de todas las perturbaciones y oscilaciones cuánticas, que se representan matemáticamente mediante la función de peso dentro de la ecuación de Sturm-Liouville. Como resultado, la derivada del operador hamiltoniano está relacionada con cualquier cambio en la posición, el impulso y el tiempo y, por lo tanto, se puede calcular tomando las derivadas parciales de estos vectores propios:

$$dt\,.\,\frac{\partial H\hat{}}{\partial r_i}\,dr_i + \frac{\partial H\hat{}}{\partial s_i}\,ds_i + \frac{\partial H\hat{}}{\partial t}\,dt = dH\hat{} \tag{57}$$

La distribución del espacio de fases ρ(r,s) describe la probabilidad de que se seleccione un estado particular del sistema del volumen total del espacio de fases d n r dn s. Para describir este volumen del espacio de fases, la ecuación de Liouville produce la evolución de ρ(r,s) en el tiempo t:

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^{n} \frac{\partial \rho}{\partial r_i}\frac{\partial \rho}{\partial s_i}\,\dot{r}_i + \dot{s}_i = \frac{d\rho}{dt} = 0\,. \tag{58}$$

Como resultado de esta restricción geométrica, los vectores de estado (ρ, ρr˙ , ρs˙) se conservan en todo el sistema y el campo vectorial (r˙ ,s˙) tiene divergencia cero. Esto también permite conservar la temperatura y la energía en todo el sistema. La ecuación de continuidad viene dada por:

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^{n} \frac{\partial(\rho \dot{r}_i)}{\partial r_i} + \frac{\partial(\rho \dot{s}_i)}{\partial s_i} = 0\,. \tag{59}$$

Debido a que ∂ρ / ∂t es igual a cero, la continuidad en la densidad de probabilidad se puede denominar como:

$$\rho \sum_{i=1}^{n} \frac{\partial \dot{r}_i}{\partial r_i} + \frac{\partial \dot{s}_i}{\partial s_i} = 0\,. \tag{60}$$

Y porque los espacios vectoriales r˙ y s˙ se definen como:

$$\dot{r}_i = \frac{\partial H\hat{}}{\partial s_i}\,, \tag{61}$$

y:

$$\dot{s}_i = -\frac{\partial H\hat{}}{\partial r_i}\,, \tag{62}$$

las leyes de continuidad persisten al tomar la segunda derivada del operador hamiltoniano. Este proceso identifica el límite observable del volumen total de posibles estados del sistema:

$$\rho \sum_{i=1}^{n} \frac{\partial}{\partial r_i}\frac{\partial H\hat{}}{\partial s_i} - \frac{\partial}{\partial s_i}\frac{\partial H\hat{}}{\partial r_i} = 0\,. \tag{63}$$

La relación entre la matriz de densidad ρ y el operador hamiltoniano que describe la distribución de energía a través del sistema se vuelve evidente en esta ecuación, al igual que la relación entre los vectores de posición y momento y el operador hamiltoniano.

Una vez que se traza un rastro a través de las matrices de densidad que representan r(x) y s(x), la función de ponderación w(x) que sustenta la distribución de probabilidad se puede resolver en relación con los otros parámetros. Y así, durante un cálculo de todo el sistema, se seleccionan valores propios y los estados propios ψr(x) y ψs(x) se resuelven transitoriamente. Cabe señalar que, en esta formulación de eventos que ocurren temporalmente, los operadores impulsan el ciclo computaci

Machine Translated by Google

en lugar de los estados vectoriales. Es decir, el operador hamiltoniano puede evolucionar con el tiempo (permitiendo cambios en la posición y el impulso) mientras que los estados del vector permanecen independientes del tiempo. Por lo tanto, cualquier evento que ocurra dentro de una evolución temporal puede contribuir al resultado de un cálculo. Esto incluye cualquier oscilación cuántica y cualquier corriente de entrada clásica que ocurra durante esa evolución temporal.

2.6. Restaurar la incertidumbre después de que el estado del sistema se defina transitoriamente

Al relacionar cualquier cambio probabilístico en la posición y el impulso de cada ion con la distribución de energía en todo el sistema, el operador hamiltoniano representa de manera efectiva todos los resultados observables. El cambio unitario de base que guía este proceso computacional está definido por el factor de fase exp(iHt /¯h). Por lo tanto, cualquier inversión de la dirección del tiempo hace que las energías positivas se vuelvan negativas, un resultado que no está permitido por la primera ley de la termodinámica, a menos que el espín también se invierta con la introducción de un momento dipolar eléctrico [23]. Un momento dipolar eléctrico es un cambio espontáneo en el estado energético de un electrón en presencia de un campo eléctrico externo; Se espera que este evento ocurra aquí, ya que la distribución de probabilidad se reduce y se calculan los valores propios.

Inicialmente, cuando el sistema fue perturbado por su entorno externo, todos los estados puros componentes (y el operador hamiltoniano) se integraron para crear un volumen de probabilidad. Ahora, podemos tomar la derivada de ese volumen de probabilidad (y el operador hamiltoniano) para definir valores propios en la región límite de esa distribución de probabilidad de alta dimensión . Al hacerlo, se reduce la distribución de los estados del sistema a un único estado del sistema actualizado, con un resultado no determinista:

$$\frac{\text{Delaware}}{d\lambda} = \psi\lambda| \frac{dH\hat{}}{|\psi\lambda} . d\lambda \quad (64)$$

Esta ecuación es el teorema de Hellman-Feynman y se encuentra en el núcleo de la electrodinámica cuántica [24,25]. Aquí, cualquier cambio en la posición o momento de un ion es proporcional al cambio en el hamiltoniano, porque el hamiltoniano corresponde a todas las energías potenciales y cinéticas del sistema. Es útil señalar que lo que es incierto no es la cantidad total de energía en el sistema, sino cómo se distribuye esta energía.

La medida de posibles estados del sistema, o configuraciones energéticas, está directamente relacionada con la entropía total codificada por el sistema al inicio del ciclo computacional [26]. El estado del sistema, una vez completado el ciclo computacional, está relacionado con cualquier parámetro λ que contribuya al operador hamiltoniano, como cualquier cambio en la posición atómica, el momento, el orbital o la intensidad del campo eléctrico. Y así, la derivada de la energía total en el sistema E está relacionada con el producto interno de los vectores de estado ψ y la derivada del hamiltoniano Hˆ, ambos con respecto al parámetro λ.

Al desarrollar esta solución a la ecuación de Schrödinger, Richard Feynman descubrió que cualquier colapso de la función de onda tiene un efecto discreto sobre los átomos que la componen. Este evento va acompañado de un cambio en la distribución de carga en cada átomo del sistema, en relación con su distancia recién definida entre sí [24,25]. Para dos átomos que interactúan en una separación D, el momento dipolar inducido para cada átomo es 1/D7 . Cuanto menor es la distancia entre los átomos, mayor es el momento dipolar y mayor es el impulso del momento angular. Como resultado, cualquier perturbación del sistema (por ejemplo, la apertura de un canal iónico o un cambio en el campo eléctrico local) cambia el estado del sistema, de una manera relacionada con el estado original no perturbado y la derivada del hamiltoniano. La asignación de ubicaciones atómicas, en relación con otros átomos del sistema, va acompañada de una alteración en la organización de los electrones alrededor del núcleo. Este evento provoca fuerzas de van der Waals entre átomos vecinos [27,28].

En resumen: al integrar amplitudes de probabilidad de valores complejos, el sistema genera una cantidad de información cuántica. Al tomar la derivada de este volumen, el sistema identifica valores propios en la región límite observable de ese espacio vectorial de alta dimensión . A medida que se seleccionan los valores propios, se eliminan todos los demás estados propios. En ese punto, la función de onda colapsa y la información contenida en el sistema se comprime abruptamente.

Machine Translated by Google

De acuerdo con el principio de Landauer, cualquier compresión de información va acompañada de una liberación de energía libre [29-32]. Esta energía libre se puede utilizar para realizar trabajo en el sistema, por ejemplo, restaurar el potencial de membrana en reposo a medida que se reduce la incertidumbre dentro del campo receptivo de la neurona . Fundamentalmente, este intercambio bidireccional de energía libre y entropía garantiza que la energía total del sistema se conserve durante el transcurso del cálculo, respetando al mismo tiempo la primera y segunda leyes de la termodinámica [26].

Al pasar de un estado anterior a un estado posterior, el sistema actualiza una solución a un problema computacional: ¿Cuál es el estado óptimo del sistema para codificar el entorno circundante? Sin embargo, este estado definido del sistema es transitorio, porque va acompañado de una restauración inmediata de la incertidumbre, a través de una alteración en la distribución de carga alrededor de cada átomo. El momento dipolar inducido por el colapso de la función de onda provoca entonces nuevas interacciones atómicas, y para mediciones sucesivas con resultados discretos, que no destruyen el entrelazamiento del sistema de partículas, cada medición con valor a establece la base para un nuevo estado, que luego pasa por el tiempo subsiguiente. evolución, de acuerdo con el postulado de la proyección de von Neumann [33]. Y así, inmediatamente después de que se resuelve una función de onda, el sistema comienza nuevamente a evolucionar con el tiempo, formando un nuevo estado de sistema probabilístico y pasando por otro ciclo computacional.

2.7. Conversión de estados probabilísticos del sistema en resultados de señalización temporalmente irreversibles

Tomar la derivada del hamiltoniano implica tomar la derivada de todos los estados puros de los componentes, con respecto a todas las perturbaciones desde el último evento de detección. Este proceso comprime la información, generando valores propios en la región límite de la distribución de probabilidad de alta dimensión. Durante este proceso computacional, se reduce la distribución de valores complejos de los posibles estados del sistema y se definen transitoriamente los estados puros de los componentes en todo el sistema (por ejemplo, la posición de los iones y el momento). Pero este estado de sistema definido de múltiples escalas se convierte inmediatamente en el pasado, a medida que se altera la distribución de carga en cada átomo y emerge un nuevo estado de sistema probabilístico.

Dado que el valor esperado para la energía de un átomo dado ai |ρ|ai es proporcional al valor esperado para su espín ψ|k|ψ , el cambio de energía debido al momento dipolar inducido provoca un cambio de signo en ambos valores y la simetría temporal se rompe. Esencialmente, a medida que los estados fundamentales pierden degeneración, los momentos dipolares resultantes deberían alterar la atracción entre los iones de sodio fuera de la neurona y los átomos que componen la bicapa lipídica de la membrana neural. Se espera que las fuerzas de van der Waals resultantes causen una violación permitida de la simetría de inversión del tiempo, afectando así la causalidad dentro del sistema a medida que se comprime la información, se identifican los valores propios y se observan las corrientes de fuga.

Sin embargo, un sistema sólo demostrará violaciones de la simetría de inversión del tiempo si la incertidumbre cuántica contribuye a la dinámica de fluctuación-disipación térmica. Sólo si los eventos de señalización ascendentes coincidentes y las oscilaciones aleatorias desencadenan un cambio en la resistencia de la membrana, dentro de los parámetros temporales de la disipación de iones y la cinética de rectificación de la bomba de iones, los eventos inherentemente probabilísticos contribuirán a activar un cambio de estado neuronal.

2.8. Condiciones bajo las cuales las fluctuaciones cuánticas contribuyen a la dinámica de disipación

Una estructura física que genera activamente un potencial de reposo electroquímico generará entropía y calor. La membrana neural proporciona resistencia y, por tanto, cierta cantidad de energía libre se disipará en entropía a medida que se produzcan interacciones eléctricas.

Sin embargo, en un sistema que atrapa calor, esta energía no necesariamente se pierde de manera irreversible; se puede utilizar para realizar un trabajo [34]. El teorema de fluctuación-disipación de Callen-Welton afirma que las fluctuaciones térmicas impulsan una función de respuesta que afecta la impedancia de la estructura, dada por χ (t):

$$\chi(t) = -\beta \, dt \, \frac{A(t)\theta(t)}{} \, . \quad (65)$$

donde β = kBT, θ(t) es la función escalón de Heaviside y A(t) es el valor esperado de Ok (t), un observable que está sujeto a fluctuaciones térmicas en un sistema dinámico. Si las escalas de tiempo de decoherencia dentro del sistema son más largas que las escalas de tiempo de disipación de iones

Machine Translated by Google

y la dinámica de ionización, la incertidumbre cuántica puede contribuir materialmente al comportamiento de los iones en la membrana neuronal [35]. Es decir, si un cambio dependiente del tiempo en el voltaje de una neurona depende de cualquier fluctuación cuántica en la posición o el momento de los iones cercanos, entonces el promedio conjunto de cada observable (una medida de fluctuación, dada por el operador hermitiano [Ok (t), Oj(t0)]) estará relacionada con la función de respuesta (una medida de disipación, dada por χ(t − t0), en función del tiempo. Esta relación está dada por la fórmula de fluctuación-disipación de Kubo [36 ,37]:

$$\chi(t - t0) = i\theta(t - t0) \quad [\, Ok\,(t), Oj(t0)\,] \quad . \tag{66}$$

Esta función de respuesta, χ(t), se puede escribir como una función de eventos oscilatorios:

$$\chi(t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi}\, mi\, i\omega t \chi(\omega)\ . \tag{67}$$

Si t < 0, entonces iωt será negativa y e χ(t) iωt será cero, por lo que toda la función de respuesta. será cero. Como tal, los estados cuánticos que sustentan esta función espectral solo pueden contribuir causalmente a la dinámica de disipación a medida que el tiempo avanza.
La transformada de Fourier de la función de respuesta proporciona una dinámica de disipación y fluctuación en el dominio de la frecuencia, dada por χ(ω):

$$\chi(\omega) = -yo \int_{0}^{\infty} tr\,(\, mi\beta \hat{H}\, [\, Ok\,(t), Oj(t0)\,]\,)\, e^{i\omega t}\, dt\ . \tag{68}$$

Esta función dependiente del tiempo proporciona la densidad de posibles estados para un sistema de partículas, que surgen en presencia de una perturbación o un campo eléctrico cambiante. Y entonces [Ok (t), Oj(t0)] es simplemente una descripción de cómo la matriz de densidad, la función de onda o el operador hamiltoniano cambia a lo largo de una evolución temporal, en comparación con la probabilidad de que todos los estados de energía se reordenen de la misma manera. bajo inversión del tiempo. Como tal, la función de respuesta χ(ω) se puede escribir en forma expandida:

$$\begin{aligned} \chi(\omega) = -yo \int_{0}^{\infty} e^{i\omega t}\, dt \sum_{mn} e^{-Em\beta}\, [\; & m|Ok|n\; n|Oj|m e^{i(Em-En)t} \\ & - m|Oj|n\; n|Ok|m e^{i(En-Em)t}\,] \end{aligned} \tag{69}$$

Si la incertidumbre en la posición y el impulso de un ion se mantiene en presencia de un campo eléctrico en constante cambio, entonces las fluctuaciones cuánticas pueden contribuir al comportamiento de los iones, afectando el estado de los iones que interactúan con el potencial electroquímico de la membrana neural. En este caso, las fluctuaciones cuánticas pueden contribuir a la probabilidad de un cambio de estado en la unidad computacional a macroescala, de un estado apagado a un estado encendido.
Sin embargo, el estado de cada neurona en el momento en que se resuelve el operador hamiltoniano regirá su respuesta. Las neuronas en un estado cortical elevado, que permiten que los eventos estocásticos activen un resultado de señalización, pueden ser empujadas hacia el umbral del potencial de acción o alejadas de él a medida que la incertidumbre se reduce abruptamente. De hecho, sólo se estimularán las neuronas en estado cortical elevado, lo que permite que un ruido aleatorio active un resultado de señalización. Mientras tanto, las neuronas que reciben estimulación por encima del umbral desde un estado de reposo exhibirán patrones de activación deter

2.9. Supuestos del Modelo 2.9.1.
Las neuronas están funcionalmente aisladas pero siguen siendo sensibles a las perturbaciones externas

Este modelo supone que las neuronas están funcionalmente aisladas, pero siguen siendo sensibles a las perturbaciones externas. La arquitectura de una red neuronal de capas profundas permite perturbaciones desde fuera del propio sistema, mientras que la fisiología de las unidades computacionales individuales dentro de las capas más profundas permite que se produzcan cálculos aislados.

Con respecto a la arquitectura de la red: las neuronas en el tálamo, que reciben señales de la periferia, exhiben patrones de activación deterministas que representan fielmente mapeos dentro de sus campos receptivos [38,39]. Estas neuronas hacen sinapsis con capas más profundas.

Machine Translated by Google

de la red neuronal, con señales en cascada sobre neuronas sensoriales primarias y neuronas corticales de orden superior, que integran cada vez más múltiples entradas con ruido para producir resultados probabilísticos [40,41]. Las neuronas de la corteza motora y los ganglios basales integran diversas entradas para iniciar el movimiento voluntario e inhibir otros movimientos [42,43]. Estas neuronas hacen sinapsis con regiones motoras posteriores en el tálamo, la formación reticular y la médula espinal, y estas señales convergen en neuronas periféricas con resultados deterministas que instruyen la extensión y flexión de grupos de músculos específicos [44,45].

Esencialmente, las neuronas de orden superior dentro de las capas más profundas de la corteza cerebral reciben mensajes desde el interior del propio sistema, y estas perturbaciones representan el mundo exterior.

Con respecto a la fisiología neuronal: en las neuronas corticales, las entradas excitadoras e inhibidoras están activamente equilibradas, de modo que las fluctuaciones térmicas dominan cualquier salida del potencial de equilibrio [5]. Mientras que un estímulo por encima del umbral ciertamente puede conducir a un comportamiento de picos en las neuronas de la corteza sensorial primaria [46,47], el ruido eléctrico aleatorio impulsa un comportamiento de picos en las neuronas corticales de orden superior [3,4]. En estas áreas, los resultados de la señalización son más probabilísticos [48,49]. Por esta razón, modelar la contribución del ruido a los resultados de la señalización es un enfoque válido. En particular, en las neuronas periféricas y las capas superficiales de los circuitos corticales, el ruido aleatorio es insuficiente para afectar los resultados de la señalización, y las propias unidades computacionales son muy resistentes al ruido, por lo que las fluctuaciones térmicas no contribuyen materialmente a los resultados neuronales, y estos términos desaparecen. de la ecuación. Sin embargo, una neurona en estado cortical elevado, que se sitúa cerca del umbral del potencial de acción y permite que un ruido eléctrico aleatorio controle los resultados de la señalización, es muy sensible al ruido. En este caso, el ruido debe tenerse en cuenta mecánicamente para modelar eficazmente la probabilidad de que la neurona pase del estado apagado al estado encendido.

2.9.2. La incertidumbre en el estado de los iones individuales afecta el estado de voltaje de la neurona

Este modelo supone que un cambio en el estado de muy pocos iones puede afectar la probabilidad de que una neurona active un potencial de acción. Cuando las neuronas se encuentran en un estado cortical elevado, un aumento de 1 mV en el potencial de membrana puede fácilmente empujar a una neurona por encima del umbral [15]. Este cambio de voltaje corresponde a 1 pA de corriente, o sólo 10 iones de sodio que se mueven a través de un canal en el transcurso de un microsegundo [16]. Dado que el movimiento aleatorio de unos pocos iones de sodio puede empujar a una neurona cortical por encima del umbral del potencial de acción, y estas células tienen corrientes de fuga significativas [50], los eventos de escala cuántica parecen afectar los resultados de escala macro en el nivel de la unidad computacional.

Se espera que las neuronas en estado cortical elevado codifiquen cantidades máximas de información, con amplitudes de alta probabilidad para cada resultado k = j y k = j. Sin embargo, para los iones, las amplitudes de probabilidad tanto en el espacio de posición como en el de momento serán cero en casi todas partes. Para los iones ubicados fuera del cerebro, la probabilidad de que un ion se ubique dentro de una neurona cortical en el siguiente instante es insignificante. Y para los iones que ya están ubicados dentro del cerebro, las amplitudes de probabilidad se centran en gran medida alrededor de los valores de posición y momento más recientes, excepto en el contexto de la apertura de un canal iónico cercano o un cambio en el campo eléctrico local. Un resultado poco probable para un ion podría ser un cambio de posición correspondiente a una fuga a través de una membrana neural, y este evento afectaría el estado de voltaje de la neurona.

Aquí, cada ion se describe mediante una función de onda o una distribución de estados propios. Durante la transformación unitaria, cualquier estado no distinguible (por ejemplo, correlaciones lineales entre estados puros) se elimina, comprimiendo así la entropía de von Neumann del sis
A medida que se asignan valores propios para cada ion, se define el voltaje de la membrana neuronal y la célula alcanza o no el umbral del potencial de acción. Si se reduce la incertidumbre durante el cálculo y se asignan valores propios, entonces la información se comprime, se libera energía libre y la neurona restaura su potencial de reposo. Si la información no se puede comprimir y no se pueden asignar valores propios, entonces no se libera energía libre y la neurona dispara un potencial de acción, codificando así la "incertidumbre" en su campo receptivo.

Machine Translated by Google

2.9.3. Las escalas de tiempo estimadas de decoherencia cumplen con los criterios para un sistema cuántico

El modelo de decoherencia en sistemas neuronales, centrado en las interacciones entre los iones de sodio y el potencial de membrana neuronal, fue proporcionado originalmente por Max Tegmark [35]. Este modelo se recalculó utilizando los datos más recientes de dispersión de Coulomb, que se presentan en un artículo hermano, junto con la formulación de álgebra matricial de generación y compresión de información cuántica [51]. Estas estimaciones demuestran que las escalas de tiempo de decoherencia en la membrana neuronal son de hecho más largas que las tasas de disipación de iones y, por lo tanto, cumplen con los criterios para un sistema cuántico, pero solo para las neuronas corticales en estado activo.

Una neurona cortical en estado activo está efectivamente aislada, porque las entradas excitadoras e inhibidoras están equilibradas, por lo que las fluctuaciones térmicas aleatorias impulsan los resultados de la señalización. Si bien hay entradas ascendentes, estas no impulsan los resultados de la señalización en las neuronas corticales; las fluctuaciones térmicas sí lo hacen. Por esta razón, cada unidad computacional se modela como un sistema cuántico de dos estados, con cierta probabilidad de pasar del estado apagado al estado encend Por supuesto, la neurona no está verdaderamente aislada; existe dentro de un circuito, por lo que un estímulo por encima del umbral aún puede impulsarlo por encima del umbral del potencial de acción y provocar que se active. El concepto de modelar la contribución del ruido a nivel cuántico al estado de la unidad computacional siempre es válido; sin embargo, en muchos circuitos, las escalas de tiempo de decoherencia son demasiado cortas, o el ruido aleatorio es insuficiente para afectar los resultados de la señalización, o las unidades computacionales en sí mismas son resistentes al ruido, por lo que las fluctuaciones térmicas no contribuyen materialmente a los resultados neuronales, y estos términos naturalmente caen Como las neuronas obedecen a la dinámic

Sólo en un estado cortical elevado se espera que la incertidumbre tanto en el potencial de la membrana neuronal como en la posición de los iones extracelulares se mantenga durante escalas de tiempo suficientemente largas para que se reduzcan las correlaciones lineales: para que la función de onda colapse, para que la información se comprima, para que la energía libre se libere. ser liberado, para que se altere el momento atómico y para que se observen cambios de voltaje en la unidad computacional, al finalizar el cálculo.

3. Resultados

3.1. La longitud de onda esperada de la liberación espontánea de energía libre durante la compresión de información Si

las redes neuronales corticales participan en cálculos cuánticos a temperatura ambiente, entonces estos sistemas termodinámicos alejados del equilibrio deben intercambiar bidireccionalmente energía libre por información, y toda la energía libre se gasta en la generación de información durante la compresión de información. La etapa inicial del ciclo de computación termodinámica se recupera parcialmente durante la etapa de compresión de la información. Como tal, se deben liberar cantidades discretas de energía libre, locales para cualquier reducción de la incertidumbre, a medida que se selecciona un estado óptimo del sistema a partir de alguna distribución de probabilidad. Cualquier fluctuación térmica debería corresponder a un cambio en el momento dipolar atómico, aumentando el momento angular de los iones individuales. Por lo tanto, estos eventos deberían ser observables. Para evaluar esta hipótesis, podemos calcular los efectos esperados de la dinámica de fluctuación-disipación cuántica en el sistema nervioso central de los mamíferos.

Si la perturbación del estado de un ion durante algún tiempo de evolución depende de cualquier cambio aleatorio en el espín o el estado energético de un electrón componente, entonces el promedio conjunto de cada observable (una medida de fluctuación, dada por $\hat{x}(t), \hat{x}(0)$) estará relacionada con la función de respuesta (una medida de disipación, dada por $\chi(t)$) en el dominio de la frecuencia [36].
Esta relación esencialmente modela cómo el comportamiento oscilatorio de un sistema cuántico, a lo largo de un eje imaginario, afecta la dinámica térmica del sistema observable:

$$\chi(t) = i\theta(t - t0) \quad \hat{x}(t), \hat{x}(0) \quad , \tag{70}$$

Machine Translated by Google

donde θ(t) es la función escalón de Heaviside y la respuesta χ(t) proporciona un valor esperado para x(t), que es un 'observable' dependiente del tiempo sujeto a fluctuaciones térmicas en un sistema dinámico. La ecuación dependiente del tiempo viene dada por:

$$f(t) = \frac{d\omega}{2\pi} \, i\omega t \, mi \, f(\omega), \tag{71}$$

y su transformada de Fourier viene dada por:

$$f(\omega) = f(t) \, mi \, -i\omega t dt \, . \tag{72}$$

Si los valores propios han de ser reales, la suma de la parte real y la parte imaginaria de la función de respuesta durante algún tiempo de evolución χ(t − t0) también debe ser real. La función de respuesta completa viene dada por:

$$\chi(\omega) = Re\chi(\omega) + Im\chi(\omega) \, . \tag{73}$$

La parte real de la función de respuesta viene dada por:

$$\begin{aligned} Re\chi(\omega) &= \frac{1}{2} \left[\chi(\omega) - \chi^{\dagger}(\omega) \right] \\ &= \frac{1}{2} \int_{-\infty}^{\infty} dt \, ei\omega t \left[\chi(t) - \chi(-t) \right]. \end{aligned} \tag{74}$$

Y la parte imaginaria de la función de respuesta viene dada por:

$$\begin{aligned} Im\chi(\omega) &= -\frac{i}{2} \left[\chi(\omega) - \chi^{\dagger}(\omega) \right] \\ &= -\frac{yo}{2} \int_{-\infty}^{\infty} dt \, ei\omega t \left[\chi(t) - \chi(-t) \right] . \end{aligned} \tag{75}$$

La energía de las fluctuaciones cuánticas E(ω) está relacionada con la frecuencia ω:

$$mi(\omega) = \frac{1}{-} \frac{1}{\beta h^{-} \omega, 2} \frac{h^{-} \omega \coth}{2} - \tag{76}$$

donde β = 1/kBT. El espectro de potencia clásico está relacionado con la densidad espectral cuántica de valores complejos de tal manera que el ruido cuántico puede contribuir a la densidad térmica local bajo ciertas condiciones:

$$\int_{-\infty}^{\infty} x\hat{}(t) - x\hat{}(t0) \quad e^{i\omega t} dt = \frac{1}{\beta E(\omega)} \int_{-\infty}^{\infty} x\hat{}(t) - x\hat{}(t0) \quad e^{i\omega t} dt. \tag{77}$$

Si E > kBT, con una temperatura alta, una amplia distribución de electrones en los estados de energía y una baja ocupación de los estados de energía, entonces la contribución cuántica al comportamiento de los iones es insignificante y el comportamiento de los iones se reducirá a las estadísticas de Boltzmann-Maxwell. Si E < kBT, con electrones a bajas frecuencias obedeciendo la ley de Rayleigh-Jeans, entonces la contribución cuántica al comportamiento de los iones también es insignificante y nuevamente el comportamiento de los iones se reducirá a la estadística clásica de Boltzmann-Maxwell. Sólo en casos de alta densidad de partículas, cuando la energía retenida por un ion es mayor que su potencial químico, y E > kBT, las fluctuaciones cuánticas contribuirán al comportamiento del ion. Se predice que este será el caso en el cerebro de los mamíferos.

Si las fluctuaciones cuánticas realmente contribuyen al comportamiento de los iones en los sistemas biológicos, entonces E debe ser mayor que kBT. Desde:

$$kBT = (1{,}38 \times 10{-}23 \text{ J/K})(310 \text{ K}) = 4{,}28 \times 10{-}21 \text{ J}, \tag{78}$$

Machine Translated by Google

y:

$$f = \frac{mi}{h} = \frac{4{,}28 \times 10{-}21 \text{ kg m2/s2}}{6{,}63 \times 10{-}34 \text{ kg m2/s}} = 6{,}46 \times 1012 \text{ Hz}, \quad (79)$$

se espera que se observen partículas de alta energía de E > kBT en el sistema nervioso central a 37 ◦ C (310 K). Específicamente, estas partículas de alta energía deben tener una frecuencia de f > 6,46 x 1012 Hz, una longitud de onda de λ < 46 micrones o una energía de E > 0,0267 eV, dentro del espectro de luz infrarroja. De hecho, se han observado emisiones espontáneas de fotones en este rango en tejido cerebral de mamíferos [52-55] y se ha demostrado que la estimulación infrarroja del cerebro tiene un efecto funcional sobre la actividad neuronal [56-58]. Se necesitan más estudios para medir las longitudes de onda exactas de estas emisiones de fotones y correlacionar temporalmente estos eventos con los resultados de la señalización neuronal.

En resumen, se predice que los fotones, específicamente en el rango infrarrojo del espectro electromagnético, deberían liberarse tras la compresión de información en el cerebro de los mamíferos. Dado que la primera ley de la termodinámica establece que la energía no se puede crear ni destruir, cualquier sistema capaz de reducir la entropía para lograr un cálculo no determinista debe liberar energía libre al comprimir la información. Si la computación cuántica ocurre en las redes neuronales corticales, entonces se deberían observar fluctuaciones térmicas espontáneas, localmente para reducir la incertidumbre. Se espera que estas fluctuaciones térmicas impulsen la activación sincrónica de un conjunto estadísticamente aleatorio de neuronas en toda

Por lo tanto, en este enfoque, la reducción de la entropía termodinámica se combina con la selección de un estado óptimo del sistema a partir de una gran distribución de probabilidad (uno que esté termodinámicamente favorecido para correlacionarse con el entorno circundante) y la liberación de energía térmica libre (que es utilizado para instanciar físicamente la solución a ese problema computacional).

3.2. Predicciones específicas de este

modelo Si las redes neuronales corticales son de hecho sistemas de computación cuántica en lugar de sistemas de computación clásicos , entonces se deberían observar pruebas de generación y compresión de información cuántica en la neocorteza. Como tal, esta teoría hace predicciones específicas con respecto a la eficiencia energética del cerebro [17]; con respecto a la dispersión de culombio y las escalas de tiempo de decoherencia en la membrana neuronal [51]; y con respecto a los efectos esperados de la estimulación electromagnética y la intervención farmacológica en las redes neuronales corticales [59]. Algunas predicciones específicas adicionales del marco teórico, impulsadas por el presente modelo, incluyen las siguientes.

3.2.1. La energía térmica libre se libera espontáneamente durante la computación a medida que se comprime la

información. Las partículas infrarrojas con longitudes de onda de λ < 46 micrones of > 6,46 × 1012 Hz deberían aparecer espontáneamente en la membrana neural durante el procesamiento de información cortical. Esta predicción debe probarse con dispositivos sensibles de detección de infrarrojos en lugar de electrodos o sistemas de imágenes clásicos; La liberación espontánea de partículas de longitud de onda infrarroja debe observarse en el cerebro a medida que la incertidumbre se resuelve en resultados de señalización. Se debería observar un aumento cuantitativo de estas partículas, por ejemplo, tras el reconocimiento perceptivo de un estímulo visual o auditivo muy incierto, con una fuerte correspondencia temporal con los potenciales relacionados con eventos P300 en la corteza cerebral. Por el contrario, esta liberación espontánea de energía térmica libre no debería ocurrir en el caso de un ataque epiléptico, cuando la activación constitutiva del canal iónico, en lugar del procesamiento de la información, conduce a una actividad neuronal altamente sincronizada en toda la corteza cerebral. Por supuesto, se han observado emisiones espontáneas de fotones en este rango en el tejido cerebral de los mamíferos [52-55], y se ha demostrado que la estimulación infrarroja del cerebro tiene un efecto funcional sobre la actividad neuronal [56-58], pero se necesitan más estudios. necesario para medir las longitudes de onda exactas de estas emisiones de fotones y evaluar si estos eventos están correlacionados temporalmente con los res

Si el cerebro genera y comprime cíclicamente entropía, el sistema debería demostrar una eficiencia energética mucho mayor de lo esperado en condiciones clásicas (Tabla 1).

Machine Translated by Google

Tabla 1. Comparación del poder explicativo: la eficiencia energética de la red neuronal.

Modelos clásicos		
Mecanismo propuesto	Observación prevista	Evidencia a favor y en contra
La eficiencia energética del sistema es el resultado de una ponderación sináptica óptima, una distribución óptima de los canales iónicos y otros mecanismos moleculares.	La producción neta de entropía física en el cerebro humano es compatible con los supuestos clásicos, según los cuales el recambio de ATP produce cierta cantidad de entropía.	El costo computacional de cada pico es de la asombrosa cifra de 0,1 W: en el contexto de la ingesta calórica conocida, este requerimiento de energía está "desviado en un factor de 108 ". [60]
modelo actual		
Mecanismo propuesto La	Observación prevista	Evidencia a favor/en contra
eficiencia energética del sistema es el resultado de la compresión de la información, con resultados neuronales provocados por la extracción de correlaciones, consistencias o "valor predictivo".	La producción neta de entropía física en el cerebro humano es demasiado baja para mantener los supuestos de un sistema clásico, y la cantidad de trabajo realizado por caloría muestra un uso casi perfecto.	"La eficiencia energética del cerebro humano es consistente con este modelo de computación no determinista". [17] "Este proceso computacional maximiza la disponibilidad de energía gratuita". [26]

3.2.2. La liberación espontánea de energía térmica libre durante la compresión de información provoca una activación sincronizada en toda la red neuronal

Este modelo describe tanto un proceso computacional como un proceso termodinámico, ya que la compresión de información es tanto la selección de un estado óptimo del sistema a partir de una gran distribución de probabilidad como la reducción de la entropía. Se prevé que este evento termocomputacional que afecta a todo el sistema, que resuelve la incertidumbre en los estados puros de los componentes, conducirá a una liberación espontánea de energía libre. Por lo tanto , las neuronas que reducen la incertidumbre durante el cálculo deberían recuperar energía libre y restaurar su potencial de reposo; las neuronas que retienen incertidumbre en su campo receptivo deberían distribuir energía libre para generar entropía y d Por lo tanto, este cálculo de todo el sistema debería dar como resultado la activación sincrónica de un conjunto estadísticamente aleatorio de neuronas a través de la red. De hecho, la actividad neuronal sincrónica se observa en un rango de frecuencias en las redes neuronales corticales y se considera un correlato de procesos cognitivos de orden superior [61,62]. Fundamentalmente, las oscilaciones de alta frecuencia que se observan durante las tareas de percepción no pueden modelarse mediante acoplamiento y reclutamiento bajo supuestos y escalas de tiempo clásicos [63,64]. Aquí, se predice que los eventos de compresión de información provocarán una actividad sincronizada espontánea en neuronas escasamente distribuida Se predice que esta actividad coordinada ocurrirá en células que permiten que el ruido aleatorio controle los resultados de la señalización (p. ej., circuitos neuronales corticales), pero no en células que actúan de manera totalmente determinista (p. ej., circuitos reflejos espinales). Si bien esta actividad oscilatoria, que ocurre en un rango de frecuencias, se ha observado en el cerebro de los mamíferos, estudios adicionales podrían explorar la posible correlación entre la codificación probabilística y la actividad a nivel de red en especies de aves y cefalópodos. Si la actividad sincrónica es causada por métodos clásicos de propagación de señales, en lugar de ser el resultado de un cálculo no determinista de todo el sistema, entonces tanto las simulaciones clásicas de redes neuronales corticales como de circuitos reflejos espinales deberían demostrar fácilmente oscilaciones rápidas y lentas. Si, por el contrario, la actividad sincrónica a través de la red es causada por la compresión de información combinada con la liberación de energía libre, entonces el disparo sincrónico debería eliminarse mediante la absorción de las longitudes de onda predichas y debería ser provocado por la introducción de estas longitudes de onda. En resumen, si un modelo clásico es correcto, entonces los mecanismos clásicos deberían generar fácilmente oscilaciones en un rango de frecuencias anidadas, y si el modelo actual es correcto, entonces la generación y compresión cíclica de información debería generar fácilmente estas oscilaciones (Tabla 2).

Machine Translated by Google

Tabla 2. Una comparación del poder explicativo: la activación sincrónica de conjuntos neuronales dispersos.

Modelos clásicos		
Mecanismo propuesto El	Observación prevista	Evidencias a favor y en
disparo sincrónico observado en un rango de frecuencias anidadas es el resultado de la codificación de información, con resultados de señalización neuronal provocados por un estímulo común.	Una combinación de uniones comunicantes, sinapsis químicas, acoplamiento efáptico, cambios en la concentración de iones y optimización de la conectividad neuronal a lo largo del tiempo provoca una activación sincrónica.	contra Estos eventos no se simulan fácilmente: "Sin embargo, es difícil identificar la contribución exacta de cada mecanismo a un tipo específico de oscilación". [65] El problema "no es trivial". [66]
modelo actual		
Mecanismo propuesto El	Observación prevista	Evidencias a favor y en
disparo sincrónico observado en un rango de frecuencias anidadas es el resultado de la compresión de información, con resultados de señalización neuronal provocados por la extracción de correlaciones.	Un pulso de fotones infrarrojos reduce el potencial de membrana de algunas neuronas, mientras que otras neuronas en estado cortical elevado se activan, lo que resulta en una activación sincrónica pero no en una actividad ictal en toda la red.	contra Las predicciones de este modelo deben probarse: específicamente, se espera que la liberación espontánea de fotones infrarrojos esté temporalmente correlacionada con las oscilaciones neuronales, pero no con la actividad ictal.

4. Discusión

La neurona se considera clásicamente como un transistor, siempre en estado encendido o apagado . Aquí, la neurona cortical se modela como un qubit, con cierta probabilidad de pasar de un estado apagado a un estado encendido durante algún tiempo de evolución. En este nuevo enfoque, un cambio de estado en la unidad computacional se basa en eventos inherentemente probabilísticos a nivel iónico.

El ruido desempeña un papel fundamental en las redes neuronales corticales y otros sistemas dinámicos complejos, facilitando la transmisión de información a través de un mecanismo de resonancia estocástica [67,68]. El ruido eléctrico aleatorio provoca transiciones de fase en unidades computacionales individuales, lo que lleva a intervalos entre picos altamente variables que se observan en las neuronas corticales [69]. El ruido eléctrico aleatorio también conduce a una actividad sincronizada y oscilaciones a gran escala a nivel de sistemas [70]. La activación resonante inducida por el ruido y la estabilidad mejorada por el ruido se observan tanto en redes neuronales [71] como en arquitecturas memrísticas [72].

El resultado de un proceso computacional ruidoso, a nivel de red, es la "codificación poblacional dispersa": la activación sincrónica de un conjunto estadísticamente aleatorio de neuronas escasamente distribuidas a través de la red [73,74]. Esta actividad neuronal se corresponde con la realización de una percepción multisensorial [62,63,75,76]. En particular, se observa escasa codificación poblacional y realización de percepciones en redes neuronales que conservan la sensibilidad al ruido eléctrico aleatorio (como la neocorteza de los mamíferos), pero no en redes neuronales que son robustas al ruido eléctrico aleatorio (como los circuitos reflejos espinales).

Muchos modelos anteriores han utilizado métodos clásicos para describir el intervalo entre picos o la activación sincrónica de conjuntos estadísticamente aleatorios de neuronas a nivel de red [9-14,67-72], pero ninguno de estos métodos proporciona una visión mecanicista de cómo funcionan las neuronas corticales individuales. las neuronas controlan los resultados de la señalización al permitir que el ruido termoeléctrico afecte los resultados de la señalización neuronal. El presente enfoque demuestra cómo la codificación ruidosa en la sinapsis está vinculada mecánicamente a resultados de señalización probabilística. Este método realiza predicciones específicas que pueden probarse en el laboratorio para evaluar si el cerebro es un sistema clásico (como se suele suponer) o un sistema cuántico de múltiples escalas (como se presenta

Hameroff y Penrose también han propuesto que el cerebro es un sistema cuántico, con una reducción objetiva orquestada del estado del sistema correspondiente a la experiencia consciente [77]. De hecho, el colapso de estados propios alternativos es una propiedad clave de las redes neuronales corticales tanto en la teoría OrchOR como en el modelo actual. Sin embargo, la teoría OrchOR se centra en el papel de la decoherencia cuántica en la dinámica de los microtúbulos, mientras que el presente informe modela la contribución del ruido electrónico aleatorio al potencial de membrana de las neuronas corticales . El primer enfoque no se conecta con los mecanismos establecidos de procesamiento de información neuronal, mientras que el segundo enfoque es altamente compatible con las contribuciones conocidas del movimiento browniano a los resultados de la señalización de las neuronas corticales.

Machine Translated by Google

Si bien la idea del cerebro como un sistema cuántico a menudo se ha considerado muy controvertida, los investigadores en neurociencia computacional han adoptado recientemente el concepto para impulsar avances significativos en el modelado de redes neuronales y el aprendizaje automático [78-80], basándose en esfuerzos anteriores para incorporar ruido en el modelo de Hodgkin y Huxley | Sin embargo, estos enfoques, cuando se implementan en la arquitectura informática clásica, no permiten la compresión de información física ni la liberación de energía gratuita. Se tendría que construir una red de unidades computacionales con comportamiento de activación probabilística y escalas de tiempo de decoherencia similares a las del cerebro de los mamíferos para lograr este proceso computacional en una arquitectura de red neuronal diseñada.

En este modelo, los estados de los componentes inherentemente probabilísticos se integran para poblar un operador hamiltoniano. Luego se diferencia el operador hamiltoniano con respecto a todas las perturbaciones del sistema. La redistribución de energía que resulta de este proceso computacional asigna valores propios para la posición y el momento de cada ion en el sistema en un único momento. Este proceso define el estado de voltaje de cada neurona, ya que los eventos de nivel cuántico afectan los resultados a nivel de la unidad computacional. Este estudio demuestra cómo las neuronas en un estado cortical elevado retienen un estado de incertidumbre, generando y comprimiendo físicamente información para lograr resultados de señalización no deterministas.

Este modelo es explícitamente una forma de cálculo cuántico de temperatura ambiente. Hay tres formas adicionales de describir este proceso computacional, proporcionadas en informes hermanos: De

acuerdo con las leyes de la termodinámica, se debe gastar energía libre para crear información y esta energía libre se recupera parcialmente al comprimir la información.

En este modelo, los estados puros de los componentes probabilísticos se pueden representar algebraicamente mediante una matriz de densidad [51]. La matriz de densidad sufre un cambio unitario de base, a medida que el estado del sistema se ve perturbado por el entorno que lo rodea a lo largo de un tiempo de evolución. La diagonalización de la matriz de densidad produce un determinante cero, lo que lleva a observables en la región límite de esa distribución de probabilidad de alta dimensión.

De acuerdo con las leyes de la holografía, estos estados puros componentes probabilísticos también pueden representarse geométricamente como ondas o funciones de onda de valores complejos [59]. Estas amplitudes de probabilidad de valores complejos interfieren de manera constructiva y destructiva en la sensible superficie polimérica de detección de carga de la membrana neural. Como resultado de esta interferencia física entre amplitudes de probabilidad, la función de onda colapsa y los valores propios se actualizan en la región límite de la distribución de probabilidad de alta dimensión. Este proceso de codificar físicamente información en la superficie de la membrana neural del polímero genera una proyección holográfica del contenido codificado.

De acuerdo con las leyes de la termodinámica, la entropía siempre debe aumentar y la energía total del sistema debe conservarse. Según los supuestos clásicos, la mayor parte de la energía libre en la corteza cerebral debería disiparse en entropía, ya que se selecciona un estado óptimo del sistema a partir de una gran distribución de probabilidad y todos los demás estados posibles del sistema se pierden. Sin embargo, la eficiencia energética observada en el cerebro sugiere que este no es el caso. Al modelar el sistema generando y comprimiendo cíclicamente información cuántica, se puede explicar la extraordinaria eficiencia energética del cerebro sin violar la primera o segunda ley de la termodinámica [17]. Este novedoso enfoque también forja una profunda conexión entre el procesamiento predictivo y los límites termodinámicos de la computación cuántica [26]. Lograr una codificación poblacional dispersa, a través de un proceso de cálculo inherentemente probabilístico, puede conducir a redes neuronales más inteligentes y energéticamente eficientes .

En resumen, el modelo actual de mecánica hamiltoniana se complementa con estos modelos de mecánica matricial, mecánica ondulatoria y computación termodinámica instanciada por software húmedo, que demuestran el mismo proceso de generación y compresión de información . De hecho, las neuronas corticales pueden describirse mejor como qubits, que codifican la entropía de von Neumann, en lugar de bits clásicos, que codifican la entropía de Shannon. Este marco teórico para la computación cuántica a temperatura ambiente puede no sólo proporcionar información útil sobre el funcionamiento de los sistemas biológicos, sino también impulsar avances en la ingeniería inteligente.

Machine Translated by Google

Financiamiento: El autor recibió apoyo para este trabajo del Western Institute for Advanced Study, con generosas donaciones de Jason Palmer, Bala Parthasarathy y Dave Parker.

Declaración de disponibilidad de datos: todos los métodos, materiales y datos necesarios para replicar este estudio están incluidos en el manuscrito.

Conflictos de intereses: El autor declara no tener conflictos de intereses. Los financiadores no tuvieron ningún papel en el diseño del estudio; en la recopilación, análisis o interpretación de datos; en la redacción del manuscrito; o en la decisión de publicar los resultados.

Referencias

1. Softky, WR; Koch, C. La activación altamente irregular de las células corticales es inconsistente con la integración temporal de EPSP aleatorios. J. Neurosci. 1993, 13, 334–350. [Referencia cruzada] [PubMed]
2. Poderes, RK; Binder, MD Modulación efectiva de la velocidad de activación de motoneuronas y corriente sináptica. J. Neurofisiol. 1995, 74, 793–801. [Referencia cruzada] [PubMed]
3. popa, EA; Kincaid, AE; Wilson, CJ Fluctuaciones espontáneas del potencial de membrana por debajo del umbral y variabilidad del potencial de acción de las neuronas corticoestriatales y estriatales de rata in vivo. J. Neurofisiol. 1997, 77, 1697–1715. [Referencia cruzada] [PubMed]
4. Dorval, AD; El ruido blanco del canal JA es esencial para las oscilaciones periumbrales en las neuronas estrelladas entorrinales. J. Neurosci. 2005, 25, 10025–10028. [Referencia cruzada] [PubMed]
5. Haider, B.; Duque, A.; Hasenstaub, AR; McCormick, DA La actividad de la red neocortical in vivo se genera a través de una dinámica equilibrio entre excitación e inhibición. J. Neurosci. 2006, 26, 4535–4545. [Referencia cruzada] [PubMed]
6. Beck, JM; Mamá, WJ; Kiani, R.; Hanks, T.; Churchland, Alaska; Roitman, J.; Shadlen, Minnesota; Latham, PE; Pouget, A. Probabilístico Códigos de población para la toma de decisiones bayesiana. Neurona 2008, 60, 1142-1152. [Referencia cruzada] [PubMed]
7. Maoz, O.; Tka˘cik, G.; Esteki, MS; Kiani, R.; Schneidman, E. Aprendizaje de representaciones neuronales probabilísticas aleatoriamente circuitos conectados. Proc. Nacional. Acad. Ciencia. Estados Unidos 2020, 117, 25066–25073. [Referencia cruzada] [PubMed]
8. Fayaz, S.; Fakharian, MA; Ghazizadeh, A. La presentación de estímulo puede mejorar la irregularidad de los picos a través de subcorticales y corticales. regiones. Computación PLoS. Biol. 2022, 18, e1010256. [Referencia cruzada] [PubMed]
9. Rinzel, J.; Miller, RN Cálculo numérico de soluciones periódicas estables e inestables de las ecuaciones de Hodgkin-Huxley. Matemáticas. Biosci. 1980, 49, 27–59. [Referencia cruzada]
10. Rowat, P. Estadísticas de intervalo entre picos en el modelo estocástico de Hodgkin-Huxley: coexistencia de ráfagas de frecuencia gamma y Cocción muy irregular. Computación neuronal. 2007, 19, 1215–1250. [Referencia cruzada]
11. Austin, TD El surgimiento de las ecuaciones deterministas de Hodgkin-Huxley como límite del mecanismo subyacente del canal iónico estocástico. Ana. Aplica. Prob. 2008, 18, 1279–1325. [Referencia cruzada]
12. Montbrió, E.; Pazó, D. La teoría exacta del campo medio explica el doble papel de las sinapsis eléctricas en la sincronización colectiva. Física. Rev. Lett. 2020, 125, 248101. [Referencia cruzada] [PubMed]
13. Goldman, JS; Kusch, L.; Aquilué, D.; Yalcinkaya, BH; Depannemaecker, D.; Ancourt, K.; Nghiem, T.-AE; Jirsa, V.; Destexhe, A. Una simulación neuronal integral del sueño de ondas lentas y una dinámica de vigilia altamente receptiva. Frente. Computadora. Neurociencias. 2021, 16, 1058957. [Referencia cruzada] [PubMed]
14. Ostojic, SJ Distribuciones de intervalos entre picos de neuronas con picos impulsadas por entradas fluctuantes. J. Neurofisiol. 2011, 106, 361–373. [Referencia cruzada] [PubMed]
15. Armstrong, CM; Hille, B. Canales iónicos dependientes de voltaje y excitabilidad eléctrica. Neurona 1998, 20, 371–380. [Referencia cruzada] [PubMed]
16. Chung, SH-H.; Hoyles, M.; Allen, T.; Kuyucak, S. Estudio de corrientes iónicas a través de un canal de membrana modelo utilizando browniano dinámica. Biofísica. J. 1998, 75, 793–809. [Referencia cruzada]
17. Stoll, EA Un modelo termodinámico de computación no determinista en redes neuronales corticales. Física. Biol. 2024, 21, 016003. [Referencia cruzada]
18. Born, M. Interpretación estadística de la mecánica cuántica. Ciencia 1955, 122, 675–679. [Referencia cruzada] [PubMed]
19. Pauli, W. Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren. Z. Phys. 1925, 31, 765–783. [Referencia cruzada]
20. Caldeira, AO; Leggett, AJ Path enfoque integral al movimiento browniano cuántico. Física. A 1983, 121, 587–616. [Referencia cruzada]
21. Caldeira, AO; Leggett, AJ Túneles cuánticos en un sistema disipativo. Ana. Física. 1983, 149, 374–456. [Referencia cruzada]
22. Chang, L.-D.; Waxman, D. Ecuación cuántica de Fokker-Planck. J. Física. C 1985, 18, 5873–5879. [Referencia cruzada]
23. Khriplovich, IB; Lamoreaux, SK Violación de CP sin extrañeza: momentos dipolares eléctricos de partículas, átomos y moléculas; Springer: Berlín, Alemania, 2012 24.

Feynman, RP Forces in Molecules. Física. Rev. 1939, 56, 340. [CrossRef]

25. Esteve, JG; Falceto, F.; García Canal, C. Generalización del teorema de Hellmann-Feynman. Física. Letón. A 2010, 374, 819–822. [Referencia cruzada]
26. Stoll, EA Un proceso energéticamente eficiente de computación no determinista impulsa la aparición de modelos predictivos y comportamientos exploratorios. Frente. Cognir. 2024, 2, 1171273 [Referencia cruzada]
27. Bethe, HA El cambio electromagnético de los niveles de energía. Física. Rev. 1947, 72, 339. [CrossRef]
28. Holstein, BA La interacción de van der waals. Soy. J. Física. 2001, 69, 441–449. [Referencia cruzada]

Machine Translated by Google

29. Landauer, R. Irreversibilidad y generación de calor en el proceso informático. IBM J. Res. Desarrollo. 1961, 5, 183–191. [Referencia cruzada]
30. Berut, A.; Arakelyan, A.; Petrosyan, A.; Ciliberto, S.; Dillenschneider, R.; Lutz, E. Verificación experimental de Landauer. Principio que vincula la información y la termodinámica. Naturaleza 2012, 483, 187–189. [Referencia cruzada]
31. Jun, Y.; Gavrilov, M.; Beckhoefer, J. Prueba de alta precisión del principio de Landauer en una trampa de retroalimentación. Física. Rev. Lett. 2014, 113, 190601. [Referencia cruzada]
32. Yan, LL; Xiong, TP; Rehan, K.; Zhou, F.; Liang, DF; Chen, L.; Zhang, JQ; Yang, WL; Mamá, ZH; Feng, M. Demostración de un solo átomo del principio cuántico de Landauer. Física. Rev. Lett. 2018, 120, 210601. [Referencia cruzada] [PubMed]
33. Von Neumann, J. Fundamentos matemáticos de la mecánica cuántica; Springer: Berlín, Alemania, 1932; Volumen 205, pág. 116277.
34. Callen, HB; Welton, TA Irreversibilidad y ruido generalizado. Física. Rev. 1951, 83, 34. [CrossRef]
35. Tegmark, M. Por qué el cerebro probablemente no sea una computadora cuántica. inf. Ciencia. 2000, 128, 155-179. [Referencia cruzada]
36. Kubo, R. El teorema de fluctuación-disipación. Prog. Rep. Física. 1966, 29, 255–284. [Referencia cruzada]
37. Martyushev, LM; Nazarova, AS; Seleznev, VD Sobre el problema de la producción mínima de entropía en el estado estacionario de no equilibrio. J. Física. A 1966, 40, 371–380. [Referencia cruzada]
38. Taube, JS Los análisis de intervalos entre picos revelan patrones de disparo irregulares en intervalos cortos, pero no largos, en las células de dirección de la cabeza de rata. J. Neurofisiol. 2010, 104, 1635–1648. [Referencia cruzada] [PubMed]
39. Naito, T.; Sadakane, O.; Okamoto, M.; Sato, H. Ajuste de orientación de la supresión envolvente en el núcleo geniculado lateral y la corteza visual primaria del gato. Neurociencia 2007, 149, 962–975. [Referencia cruzada] [PubMed]
40. Averbeck, BB; Latham, PE; Pouget, A. Correlaciones neuronales, codificación y computación de poblaciones. Nat. Rev. Neurociencias. 2006, 7, 358–366. [Referencia cruzada]
41. Shadlen, MN; Newsome, WT La descarga variable de las neuronas corticales: implicaciones para la conectividad, la computación y codificación de información. J. Neurosci. 1998, 18, 3870–3896. [Referencia cruzada]
42. Lee, D.; Puerto, Países Bajos; Kruse, W.; Georgopoulos, AP Variabilidad y ruido correlacionado en la descarga de neuronas en áreas motoras y parietales de la corteza de primates. J. Neurosci. 1998, 18, 1161–1170. [Referencia cruzada]
43. Cui, H.; Anderson, RA Diferentes representaciones de planes motores potenciales y seleccionados por distintas áreas parietales. J. Neurosci. 2011, 31, 18130–18136. [Referencia cruzada]
44. Jackson, A.; Mavoori, J.; Fetz, EE Correlaciones entre las mismas células de la corteza motora y los músculos del brazo durante una tarea entrenada, gratis comportamiento y sueño natural en el mono macaco. J. Neurofisiol. 2007, 97, 360–374. [Referencia cruzada]
45. Russo, AA; Bittner, SR; Perkins, SM; Seely, JS; Londres, BM; Lara, AH; Miri, A.; Marshall, Nueva Jersey; Kohn, A.; Jessell, TM; et al. La corteza motora incorpora órdenes similares a las de los músculos en una respuesta poblacional desenredada. Neurona 2018, 97, 953–966. [Referencia cruzada]
46. Hubel, DH; Wiesel, TN Campos receptivos de neuronas individuales en la corteza estriada del gato. J. Physiol. 1959, 148, 574–591. [Referencia cruzada]
47. Mendonça, PR; Vargas-Caballero, M.; Erdélyi, F.; Szabó, G.; Paulsen, O.; Robinson, HP Dinámica estocástica y determinista de disparos intrínsecamente irregulares en interneuronas inhibidoras corticales. eLife 2016, 5, e16475. [Referencia cruzada]
48. Wimmer, K.; Compté, A.; Roxin, A.; Peixoto, D.; Renart, A.; de la Rocha, J. Dinámica de integración sensorial en una red jerárquica explica las probabilidades de elección en el área cortical MT. Nat. Comunitario. 2015, 6, 6177 [Referencia cruzada] [PubMed]
49. Melloni, L.; Schwiedrzik, CM; Müller, N.; Rodríguez, E.; Singer, W. Las expectativas cambian las firmas y el momento de los correlatos electrofisiológicos de la conciencia perceptiva. J. Neurosci. 2011, 31, 1386–1396. [Referencia cruzada]
50. Huang, S.; Hong, S.; De Schutter, E. Las corrientes de fuga no lineales afectan la fisiología de las neuronas de los mamíferos. Frente. Celúla. Neurociencias. 2015, 9, 432.
51. Stoll, EA El ruido eléctrico aleatorio impulsa el cálculo no determinista en redes neuronales corticales. bioRxiv 2022. [CrossRef]
52. Isojima, Y.; Isoshima, T.; Nagai, K.; Kikuchi, K.; Nakagawa, H. Bioquímica-iluminiscencia ultradébil detectada en el hipocampo de rata rebanadas. Neurorreporte 1995, 6, 658–660. [Referencia cruzada]
53. Kobayashi, M.; Takeda, M.; Sato, T.; Yamazaki, Y.; Kaneko, K.; Ito, K.; Kato, H.; Inaba, H. Imágenes in vivo de la emisión espontánea de fotones ultradébiles del cerebro de una rata se correlacionan con el metabolismo energético cerebral. Neurociencias. Res. 1999, 34, 103–113. [Referencia cruzada] [PubMed]
54. Kataoka, Y.; Cui, Y.; Yamagata, A.; Niigaki, M.; Hirohata, T.; Oishi, N.; Watanabe, Y. La oxidación del tejido neural dependiente de la actividad emite fotones ultradébiles intrínsecos. Bioquímica. Biofísica. Res. Comunitario. 2001, 285, 1007–1011. [Referencia cruzada] [PubMed]
55. Tang, R.; Dai, J. Imágenes espaciotemporales de actividades biofotónicas inducidas por glutamato y transmisión en circuitos neuronales. MÁS UNO 2014, 9, e85643. [Referencia cruzada] [PubMed]
56. Amaroli, A.; Marcoli, M.; Venturini, A.; Passalacqua, M.; Agnati, LF; Signore, A.; Rafetto, M.; Maura, G.; Benedicenti, S.; Cervetto, C. Los fotones láser del infrarrojo cercano inducen la liberación de glutamato desde las terminales nerviosas cerebrocorticales. J. Biofotónica 2018, 11, e201800102. [Referencia cruzada] [PubMed]
57. Naeser, MA; Ho, MD; Martín, PI; Hamblin, señor; Koo, BB Aumento de la conectividad funcional dentro de las redes neuronales intrínsecas en el accidente cerebrovascular crónico después del tratamiento con fotobiomodulación transcraneal roja/infrarroja cercana. Fotobiomódulo. Fotomed. Cirugía Láser . 2020, 38, 115-131. [PubMed]
58. Tan, X.; Rajguru, S.; Joven, H.; Xia, N.; Valores, SR; Xiao, X.; Richter, CP Energía radiante necesaria para la estimulación neuronal infrarroja. Ciencia. Rep. 2015, 5, 13273. [CrossRef] [PubMed]
59. Stoll, EA El modelado de la interferencia de electrones en la membrana neuronal produce una proyección holográfica de información representativa. contenido. bioRxiv 2022. [CrossRef]

Machine Translated by Google

60. Levy, WB; Calvert, VG La comunicación consume 35 veces más energía que la computación en la corteza humana, pero ambos costos son necesario para predecir el número de sinapsis. Proc. Nacional. Acad. Ciencia. EE.UU. 2021, 118, e2008173118. [Referencia cruzada]

61. Buzsaki, G.; Draguhn, A. Oscilaciones neuronales en redes corticales. Ciencia 2004, 304, 1926-1929. [Referencia cruzada]

62. Engel, AK; Singer, W. Unión temporal y correlatos neuronales de la conciencia sensorial. Tendencias Cogn. Ciencia. 2001, 5, 16-25. [Referencia cruzada]

63. Stacey, WC; Krieger, A.; Litt, B. Reclutamiento de redes para oscilaciones coherentes en un modelo informático del hipocampo. J. Neurofisiol. 2011, 105, 1464–1481. [Referencia cruzada]

64. Whittington, MA; Cunningham, Missouri; LeBeau, FEN; Racca, C.; Traub, RD Múltiples orígenes del ritmo gamma cortical. Desarrollo. Neurobiol. 2010, 71, 92-106. [Referencia cruzada]

65. Timofeev, I.; Bazhenov, M.; Señor, J.; Sejnowski, T. Sincronización neuronal y ritmos talamocorticales durante el sueño, la vigilia y la epilepsia. En Mecanismos básicos de las epilepsias de Jasper, 4ª ed.; Oxford University Press: Oxford, Reino Unido, 2012.

66. Gansel, KS Sincronía neuronal en redes corticales: mecanismos e implicaciones para el procesamiento y codificación de información neuronal. Frente. Integral Neurociencias. 2022, 16, 900715. [Referencia cruzada]

67. Astumián, RD; Moss, F. El papel constructivo del ruido en el transporte impulsado por fluctuaciones y la resonancia estocástica. Caos 1998, 8, 533–538. [Referencia cruzada] [PubMed]

68. Lucarini, V. Resonancia estocástica en sistemas de desequilibrio. Física. Rev.E 2019, 100, 062124. [CrossRef]

69. Lindner, B.; Schimansky-Geier, L. Enfoque analítico del sistema estocástico Fitz-Hugh-Nagumo y resonancia de coherencia. Física. Rev.E 1999, 60, 7270–7276. [Referencia cruzada]

70. Lindner, B.; García-Ojalvo, J.; Neiman, A.; Schimansky-Geier, L. Efectos del ruido en sistemas excitables. Física. República de 2004, 392, 321–424. [Referencia cruzada]

71. Valenti, D.; Augello, G.; Spagnolo, B. Dinámica de un sistema FitzHugh-Nagumo sometido a ruido autocorrelacionado. euros. Física. JB 2008, 65, 443–451. [Referencia cruzada]

72. Surazhevsky, IA; Demim, Virginia; Ilyasov, AI; Emelyanov, AV; Nikiruy, KE; Rylkov, VV; Shchanikov, SA; Bordanov, IA; Gerasimova, SA; Guseinov, DV; et al. Persistencia asistida por ruido y recuperación del estado de memoria en una red neuromórfica de picos memristivos. Caos 2021, 146, 110890.

73. Geisler, C.; Brunel, N.; Wang, XJ Contribuciones de la dinámica de membrana intrínseca a oscilaciones rápidas de la red con irregularidades Descargas neuronales. J. Physiol. 2005, 94, 4344–4361.

74. Brunel, N. Dinámica de redes escasamente conectadas de neuronas excitadoras e inhibidoras. J. Computación. Neurociencias. 2000, 8, 183–208. [Referencia cruzada]

75. Csibra, G.; Davis, G.; Spratling, MW; Johnson, MH Oscilaciones gamma y procesamiento de objetos en el cerebro infantil. Ciencia 2000, 290, 1582–1585. [Referencia cruzada] [PubMed]

76. Herrmann, CS; Knight, RT Mecanismos de atención humana: potenciales y oscilaciones relacionados con eventos. Neurociencias. Biocomportamiento. Rdo. 2001, 25, 465–476. [Referencia cruzada] [PubMed]

77. Hameroff, SR; Penrose, R. Conciencia en el universo: una revisión de la teoría 'OrchOR'. Física. Vida Rev. 2014, 11, 39–78. [Referencia cruzada] [PubMed]

78. Khalid, M.; Wu, J.; Ali, TM; Ameen, T.; Altaher, AS; Moustafa, AA; Zhu, Q.; Xiong, R. Modelado computacional cortico-hipocampal utilizando redes neuronales de inspiración cuántica. Frente. Computadora. Neurociencias. 2020, 14, 80. [Referencia cruzada] [PubMed]

79. Wang, Z.; Xu, M.; Zhang, Y. Red neuronal acoplada por pulsos cuánticos. Red neuronal. 2022, 152, 105-117. [Referencia cruzada] [PubMed]

80. Jeswal, SK; Chakraverty, S. Desarrollos y aplicaciones recientes en redes neuronales cuánticas: una revisión. Arco. Computadora. Métodos Ing. 2019, 26, 793–807. [Referencia cruzada]

81. Adair, RK Ruido y resonancia estocástica en canales iónicos dependientes de voltaje. Proc. Nacional. Acad. Ciencia. Estados Unidos 2003, 100, 12099–12104. [Referencia cruzada]

Descargo de responsabilidad/Nota del editor: Las declaraciones, opiniones y datos contenidos en todas las publicaciones son únicamente de los autores y contribuyentes individuales y no de MDPI ni de los editores. MDPI y/o los editores renuncian a toda responsabilidad por cualquier daño a personas o propiedad que resulte de cualquier idea, método, instrucción o producto mencionado en el contenido.

Machine Translated by Google

AppliedMath

MDPI

Artículo

Proposiciones de cálculo de bondad de red

Marina Bershadsky 1,†, Božidar Ivankovi´c 2,* ,† y Marko Puši´c 2,†

1 Departamento de Ciencias de la Computación, Shamoon College of Engineering (SCE), 56 Bialik, Be'er Sheva 8410802, Israel; marinab@ac.sce.ac.il Universidad

2 de Ciencias Aplicadas Hrvatsko Zagorje Krapina, Šetalište Hrvatskog Narodnog Preporoda 6, 49000 Krapina, Croacia; marko.pusic@vhzk.hr

* Correspondencia: bozidar.ivankovic@vhzk.hr; Tel.: +385-98-407-189

† Estos autores contribuyeron igualmente a este trabajo.

Resumen: Acuñamos el término “bondad de la red” para un valor que definimos para una red integrada en un entorno determinado como una métrica que describe la idoneidad de esa red para satisfacer una demanda. Se proponen tres fórmulas para calcular la métrica a partir de tres valores de variables. La primera variable considera partes del entorno gravitadas por la red. Para estas partes del entorno, definimos un valor que mide los costos del usuario al negarle el uso de la red. Por último, pero no menos importante, se consideran los costes de mantenimiento de la red. Los resultados se obtienen tras centrarse en infraestructuras y redes de transporte, pero también pueden utilizarse para otros tipos de redes.

Palabras clave: evaluación de redes; bondad de la red; centro de demanda; inercia de la red

1. Introducción

Una red está diseñada para satisfacer las demandas de un entorno específico. hay tres Factores esenciales a considerar al evaluar una red en cualquier entorno: - El rango de cobertura de la red dentro del entorno; - La disponibilidad de la red para los usuarios, reflejada en los costos incurridos por el usuario; - Los costes de mantenimiento asociados a la red.

Si la demanda excede la capacidad de la red, es posible que algunas áreas no reciban servicio, que el servicio se encarezca o que aumenten los gastos de mantenimiento. Además, se pueden integrar varias redes en el mismo entorno para satisfacer la demanda de un entorno.

Para evaluar la función de una red, en este artículo definimos la función de “bondad de la red” y proponemos fórmulas para calcular los valores de bondad y comparar dos o más redes de infraestructura integradas en el mismo entorno. También nosotros

Demuestre aproximadamente todos los teoremas, proposiciones, lemas y corolarios, ya que hasta la fecha no se ha realizado ningún enfoque similar en la literatura disponible.

Actualmente, el término predominante utilizado para evaluar el funcionamiento de la red es "Calidad de servicio" (QoS), y a lo largo de los años se han publicado muchos millones de artículos sobre el tema . Un trabajo publicado recientemente por Kanellopoulos et al. en [1] ofrece un estudio extenso de arquitecturas y protocolos de red. Sin embargo, los autores sólo presentan una comparación cualitativa de los requisitos de red para algunas aplicaciones de ciudades inteligentes (SCA) y no atribuyen una puntuación general que permita clasificar entre las alternativas. Karakus y Durresi escribieron un estudio exhaustivo [2] para examinar exhaustivamente la QoS en redes definidas por software (SDN)/OpenFlow. Los autores proporcionan una descripción detallada de la arquitectura SDN y dividen 43 métodos diferentes según dos enfoques para mantener la capacidad de una red para proporcionar los servicios necesarios para un escenario de tráfico seleccionado. Si bien está claro que el Marco de Calidad y Resultados (QoF, por sus siglas en inglés) suave supera al enfoque de QoF estricto, no se define ninguna métrica cuantitativa.

check for updates

Cita: Bershadsky, M.; Ivankovic, B.; Puši´c, M. Bondad de la red Proposiciones de cálculo. Matemáticas aplicadas 2024, 4, 828–842. https://doi.org/10.3390/appliedmath4030044

Editor académico: Takayuki Hibi

Recibido: 12 de enero de 2024
Revisado: 10 de junio de 2024
Aceptado: 20 de junio de 2024
Publicado: 2 de julio de 2024

Copyright: © 2024 por los autores. Licenciatario MDPI, Basilea, Suiza.

Este artículo es un artículo de acceso abierto. distribuido bajo los términos y condiciones de los Creative Commons Licencia de atribución (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

Machine Translated by Google

En otros trabajos, las incorporaciones de redes se analizan en términos de resolución de problemas complejos una vez que se define una red. En [3], el Análisis de Redes Sociales (SNA) se presenta como una poderosa técnica para construir y mantener relaciones sociales dentro de la red. En las redes sociales, los usuarios interactúan y se comunican entre sí y también publican contenido personalizado. La incrustación de redes con atributos dinámicos novedosos (DANE) se estudia experimentalmente en [4] para avanzar en tareas de aprendizaje como la clasificación de nodos, la agrupación de redes y la predicción de enlaces. La integración de redes aprovecha la proximidad de los nodos para aprender una representación vectorial de nodos de baja dimensión para cada nodo dentro de la red social existente y, por lo tanto, estudia cómo se establecen y evolucionan las relaciones sociales. Sin embargo, los valores del SNA también se dan sólo de forma cualitativa, ya que se evalúa el DANE. En [5], Salkhi et al. Considere las redes como un medio para evaluar el aislamiento estructural y la integración de nodos, evaluar su accesibilidad y navegabilidad dentro del gráfico y predecir posibles cambios estructurales que puedan afectar la evolución de la red. Sin embargo, nuevamente, no se proporciona ningún valor cuantitativo para medir el aislamiento o la integración en el escenario dado.

El área de aplicación que elegimos para mostrar nuestro enfoque son las redes de transporte público. Normalmente, las redes de infraestructura de tráfico surgen de forma natural a medida que aumenta la demanda en un entorno. Sin embargo, una red existente no puede satisfacer las necesidades del medio ambiente a tiempo. En la última sección, presentamos varias alternativas para rutas de transporte público que podrían acomodar un modo de transporte con suficiente volumen y velocidad para hacer que el transporte público sea lo más atractivo posible. A pesar de buscar en la literatura disponible, no pudimos encontrar un instrumento para seleccionar cuantitativamente variantes dibujadas heurísticamente.

Bogdanova et al. investigue las instalaciones de transporte de un entorno urbano confortable en [6], donde la “comodidad de la red” se mide por el número de usuarios con menos de 500 (u 800) metros para viajar entre la población total. La calificación de seguridad de la red de transporte público se calcula dividiendo el número de ciudadanos heridos en el transporte público por el número total de pasajeros que lo utilizan. Los sistemas de transporte se han vuelto cada vez más vitales para las economías modernas, como señalan Colón et al. describir en [7]. En [8]. Guliot propone una herramienta de decisión que ofrece varios escenarios alternativos para reconfigurar los servicios de tránsito. Cada escenario incluye un conjunto de líneas existentes que se cerrarán y el correspondiente perfil de retraso causado por el cierre de cada línea. Otro estudio, de Wang et al. [9], aborda el desarrollo urbano desequilibrado mejorando la eficiencia del transporte, reduciendo la congestión del tráfico y los accidentes, y coordinando el transporte. Kuo et al. en [10] enfatizan la importancia de un transporte público fácilmente disponible que proporcione demanda en tiempo real , ofrezca un servicio conveniente y de alta calidad y opere de manera energéticamente eficiente.
En [11], Nes et al. presentar un modelo de optimización para seleccionar rutas, asignar frecuencias y determinar el número de pasajeros en la red existente generando un conjunto de rutas posibles.

El resto de este documento está organizado de la siguiente manera. En la Sección 2, presentamos nuestra terminología y describimos la metodología, incluidos lemas y teoremas y sus demostraciones que subrayan nuestro enfoque. En la Sección 3, presentamos los resultados y los discutimos. En la siguiente sección, ofrecemos un ejemplo de nuestro enfoque propuesto (Sección 4), y en la sección final, ofrecemos conclusiones (Sección 5).

2. Terminología y teoría básica

En esta sección, definimos los términos que luego usamos en teoremas y lemas. Extraemos nuestra terminología de los campos de la teoría de conjuntos, la teoría de grafos y el cálculo geométrico analítico. Primero, asumimos que cualquier entorno se encuentra dentro de algún “terreno”: un estado, un continente o la Tierra en general. El suelo es un plano bidimensional.

Definición 1 (Entorno y tamaño de la red). El entorno de red U es parte del terreno definido únicamente por sus fronteras administrativas o de cualquier otra forma única que determine qué ubicaciones están dentro y cuáles fuera del terreno. El área de U se denota μ(U) y se mide en cuadrados.

Machine Translated by Google

El entorno de la red se divide en zonas que comprenden puntos de demanda donde se concentra la demanda. Diferentes zonas no comparten los mismos puntos de demanda: matemáticamente , están mutuamente inconexas.

Definición 2 (Zona de demanda y punto de demanda). Zona de demanda Ti , i = 1, . . . ,n es parte del entorno de red: Ti U. Además, para todo i, j, U = $\bigcup_{i=1}^{n}$ Ti y Ti ∩ Tj = para i = j.

Punto de demanda Pij Ti , j = 1, . . . , m(i) es un punto con valor asociado positivo Q(Pij), que es la demanda concentrada en Pij.

El número de puntos de demanda es m(i) y varía de una zona a otra. La definición formal se da a continuación.

Definición 3. Supongamos que para cada Ti existe m(i) N, y un conjunto finito $\bar{N}$ = {1, 2, . . . , m(i)} con función inyectiva pi : $\bar{N}$ → Ti , denotada pi(j) = Pij, y pi(j) = pi(k) cuando j = k. Entonces, Qi : pi($\bar{N}$) → R+ es la cantidad de demanda dada en el punto de demanda Pij denotado Qi(Pij) = Qij.

Para calcular la bondad de la red, necesitamos coordenadas, un par de números que Determina de forma única la posición de un punto en el plano.

Definición 4 (Entorno de red coordinado). Un entorno está coordinado si cualquier punto puede definirse de forma única mediante un par de números reales (x, y).

Las coordenadas son necesarias para los cálculos algebraicos y lo garantizamos mediante demostrando el siguiente Lema.

Lema 1. Se puede coordinar cualquier entorno de red.

Prueba del Lema 1. Establecemos una ubicación conocida en el entorno como un par (0, 0). Establecemos la dirección oeste-este a través de (0, 0) como abscisa y la dirección sur-norte como eje de ordenadas. Entonces, cualquier punto P U está definido únicamente por el par P ≡ (x, y). Generalmente es Pij ≡ (xij, yij).

Nuestro objetivo es evaluar una red, y para ello necesitamos definir algebraicamente una “segmento” entre dos puntos.

Definición 5. Un segmento entre dos puntos, A y B, con coordenadas (xA, yA) y (xB, yB), respectivamente, es un conjunto formado por todos los puntos de la recta que conecta los dos puntos.

$$\overline{AB} = \{P : P = (1 - t)A + tB,\ t \in [0, 1]\}. \quad (1)$$

En terminología gráfica [12], un gráfico es un par G = (N, E) definido por un conjunto dado de nodos N y un conjunto de aristas E dadas como una relación, E N × N. Esto significa que E es un conjunto de pares tomados del conjunto de todos los pares posibles de todos los nodos dados en N. En nuestro artículo, una red es un gráfico conexo G = (N, E), donde E es un conjunto de segmentos y N es un conjunto de sus puntos finales. Debido a que E N × N, una red está definida correctamente y lo demostramos formalmente en el siguiente teorema, donde E = N × N.

Teorema 1. Cualquier conjunto de segmentos define una gráfica.

Prueba del teorema 4. Para un conjunto dado de segmentos $\{\overline{AiBi}$, i = 1, 2, . . .s}, denotamos cada segmento $\overline{AiBi}$ = (Ai , Bi) como un elemento en el conjunto de aristas E = {(Ai , Bi) i = 1, 2, . . .s}}, con | E| = s. Además, cada segmento $\overline{AiBi}$ da dos puntos en el conjunto de nodos N = {Ai , Bi i = 1, 2 ,. . .s}} con |N| ≤ 2s elementos, ya que diferentes segmentos pueden tener un punto final en

Machine Translated by Google

común. Entonces, el conjunto de segmentos dado genera G = (N, E) con E = N × N N × N, en el conocido sentido de definición de gráfico. □

Un camino es una secuencia de segmentos encadenados. Segmentos de ruta $\overline{\{AiBi , i = 1, \ldots , p\}}$ se consideran encadenados si el punto final de un segmento coincide con el punto inicial del siguiente segmento o con el punto final del camino: Bi = Ai+1 , i = 1, . . . , p − 1. En la Figura 3, se ilustra una ruta. Necesitamos el término de ruta para la siguiente definición.

Definición 6 (Gráfico conectado). Un grafo es conexo si cualquier par de nodos tiene un camino que comienza en uno y termina en el otro nodo.

El concepto de gráfico conexo es necesario para definir una red en nuestro artículo, ya que no todo gráfico es una red. La definición 6 nos lleva a definir la red que evaluamos en nuestro artículo, que se detalla a continuación.

Definición 7 (Red). Una red es un gráfico conectado que consta de un conjunto de segmentos.

En el siguiente teorema, demostramos que cada segmento definido por la Ecuación (1) corresponde a una ecuación única derivada de los puntos límite del segmento.

Teorema 2. Sean A = (xA, yA) y B = (xB, yB) puntos dados en un entorno cubierto por una cuadrícula de coordenadas. Entonces, el segmento AB se define como el conjunto

$$\overline{AB} = \{(x, y) \quad [xA, xB] \times [yA, yB] : (yB - yA)x + (xA - xB)y + xByA - xAyB = 0\}. \quad (2)$$

Prueba del teorema 2. Del Lema 1 y la Ecuación (1), obtenemos x − xA = t(xB − xA), y − yA y − yA = t(yB − yA), $\frac{x - xA}{xB - xA} = \frac{y}{\quad}$. Como t [0, 1], tenemos (x, y) yB − yA [xA, xB] × [yA, yB]. Ahora, (yB − yA)(x − xA) − (xB − xA)(y − yA) = 0 produce la ecuación (2). y − yA = t. Si (x, y)

Además, de la ecuación (2), se desprende que $\frac{x - xA}{xB - xA} = \frac{\quad}{yB - yA}$

[xA, xB] × [yA, yB], entonces t [0, 1], y la Ecuación (1) resulta inmediatamente. □

La figura 1 ilustra las posiciones de diferentes puntos fuera del segmento que le concierne. Habiendo probada la ecuación (2), el teorema 2 nos permite calcular la distancia al cuadrado entre un punto y un segmento. En el siguiente teorema, derivamos las fórmulas de distancia al cuadrado para la posición de cada punto representado en la Figura 1.

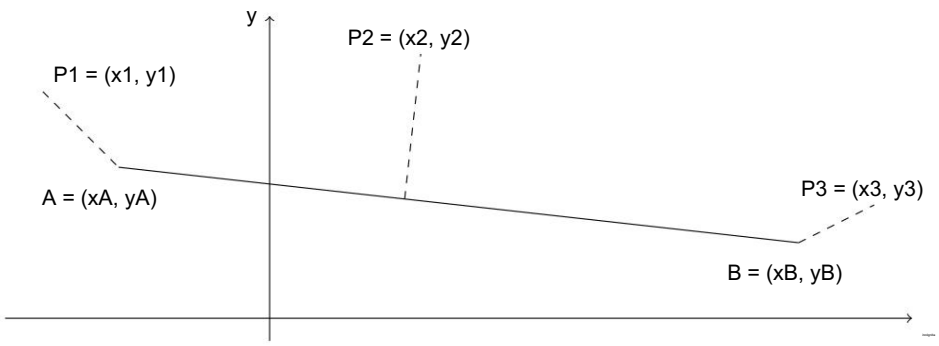


Figura 1. Distancias del segmento relacionadas con posiciones de zonas de demanda.

Teorema 3. La distancia al cuadrado entre el punto P = (x, y) {P1, P2, P3} y el segmento $\overline{AB}$ está dada por

$$d^2_{\text{días}}(P, \overline{AB}) = (x - xA)^2 + (y - yA)^2, \text{ si } (x - xA)(xB - xA) + (y - yA)(yB - yA) < 0 \quad (3)$$

$$d^2_{\text{días}}(P, \overline{AB}) = (x - xB)^2 + (y - yB)^2, \text{ si } (x - xB)(xA - xB) + (y - yB)(yA - yB) < 0 \quad (4)$$

Machine Translated by Google

o

$$\text{días}^2(P, \overline{AB}) = + \frac{[(yB - yA)x + (xA - xB)y + xByA - xAyB]^2}{(yB - yA)^2 \quad (xA - xB)^2} \tag{5}$$

en cualquier otro caso.

La ecuación (5) se obtiene después de definir $-\rightarrow AB$ como la combinación lineal de vectores $i = (1, 0) - (0, 0)$ y $j = (0, 1) - (0, 0)$, que representan los pasos unitarios en las direcciones este y norte, respectivamente. Entonces, el vector $-\rightarrow AB$ está dado por

$$-\rightarrow AB = (xB - xA) \quad i + (yB - yA) \quad j, \tag{6}$$

donde $A = (xA, yA)$ y $B = (xB, yB)$ ya están definidos. El coseno del ángulo (PAB), con el vértice en A, viene dado por la fórmula

$$\text{porque}(\quad PAB) = \frac{-\rightarrow AP \cdot -\rightarrow AB}{-\rightarrow AP2 \cdot -\rightarrow AB2\ .} \tag{7}$$

Aquí, observamos que $i \cdot j = 0$, y $i\ 2 = j\ 2 = 1$. Además, $\cos(\quad PAB) < 0$ para $90^\circ \leq$ y $\cos(\quad PAB) > 0$, para $0^\circ \leq \quad PAB \leq 90^\circ$. $\quad PAB \leq 180^\circ$

Además, tomamos el producto cruzado $i \times j$ con $i \times j = - j \times i$ y área $|i \times j| = 1$. Es decir, $i \times j$ es un vector perpendicular a i y j, por lo que no pertenece al entorno terrestre.
Luego, sigue la demostración del Teorema 3 . Se pueden encontrar más detalles sobre los productos vectoriales en [13].

Prueba del teorema 3. En el caso en que P = P1, tenemos $-\rightarrow AP \cdot -\rightarrow AB < 0$, y entonces la ecuación (3) es válida. Si P = P3, entonces $-\rightarrow BP \cdot -\rightarrow BA < 0$, y la ecuación (4) es válida.

En otros casos, igualamos las áreas del triángulo $\Delta(PAB)$ calculadas utilizando dos enfoques,

$$\frac{1}{2} \cdot d(P, \overline{AB}) \cdot |\overline{AB}| = \frac{1}{2} |-\rightarrow AP \times -\rightarrow BP|, \tag{8}$$

donde el producto vectorial se calcula en un espacio tridimensional después de que el entorno U esté integrado en el sentido P = (x, y, 0). El valor absoluto en el lado derecho de la ecuación (8) es igual

$$\begin{aligned} &|[(x - xA) \quad i + (y - yA) \quad j] \times [(x - xB) \quad i + (y - yB) \quad j]| = |x(yA \\ &- yB) + y(xB - xA) + xAyB - xByA| \cdot |i \times \quad j| = |(yB - yA)x + (xA - \\ &xB)y + xByA - xAyB|, \end{aligned} \tag{9}$$

porque área $|i \times \quad j| = 1$. En el lado izquierdo de la ecuación (8), tenemos

$$d(P, \overline{AB}) \cdot \sqrt{(xA - xB)^2 + (yA - yB)^2}. \tag{10}$$

Sustituyendo las ecuaciones (9) y (10) en la ecuación (8), obtenemos (5) después de ordenar. □

En la Definición 2, definimos un punto de demanda como un punto al que se le ha asignado un número positivo que representa la demanda en ese punto. Es importante señalar que no todos los puntos tienen un valor de demanda. Según la Definición 2, una zona de demanda contiene sólo un número finito de puntos de demanda. Si un punto en la zona de demanda no tiene demanda, no es un punto de demanda.

3. Resultados y Discusión

En esta sección, explicamos los conceptos básicos de nuestra teoría, comenzando con el centro de demanda de la zona. Cada zona de demanda tiene un punto central de demanda único. Este punto se calcula de manera similar al cálculo del centro de gravedad en física. Aquí, el cálculo se da como un teorema que sigue.

Machine Translated by Google

Teorema 4 (Zona centro de demanda). Cada zona de demanda Ti tiene un punto central de demanda único Pi con coordenadas Pi = (xTi , yTi), calculado mediante la fórmula

$$xTi = \frac{Qi1 \cdot xi1 + Qi2 \cdot xi2 + \cdots + Qim(i) \cdot xim(i)}{Q(Ti)}$$
$$yTi = \frac{Qi1 \cdot yi1 + Qi2 \cdot yi2 + \cdots + Qim(i) \cdot yim(i)}{Q(Ti)}, \quad (11)$$

donde Q(Ti) = Qi1 + Qi2 + · · · + Qim(i) es la demanda total de la zona concentrada en Pi .

El número de puntos de demanda dentro de Ti depende de Ti y, por tanto, se denota como m(i).

Prueba del teorema 4. Probamos el cálculo de la coordenada x a partir de la ecuación (11) tomando los puntos P1 := Pi1 y P3 := Pi3 de la Figura 1. Si suponemos que la masa Qi1 > Qi3 , entonces xi1 y su centro de gravedad xi13. se colocan en el eje de abscisas como se muestra en la Figura 2. xi3 ,

Figura 2. Cálculo del centro de demanda de dos puntos de demanda.

El centro xi13 del "balanceo" en la Figura 2 debe satisfacer la siguiente ley de "equilibrio":

$$Q(Pi1) \cdot (xi13 - xi1) = Q(Pi3) \cdot (xi3 - xi13). \quad (12)$$

Según la ecuación (12), la posición de la abscisa del centro de gravedad es

$$xi13 = \frac{Qi1xi1 + Qi1xi1}{Qi1 + Qi1}. \quad (13)$$

A continuación, ampliamos el cálculo para incluir P2 := Pi2 . Suponiendo Q(Pi2) ≥ Qi1 + Qi3 , obtenga el nosotros centro de gravedad entre (13) con Qi1 + Qi3 y P2 := Pi2 con Q(Pi2):

$$xi132 = \frac{Qi1xi1 + Qi3xi3 + Qi2xi2}{Qi1 + Qi3 + Qi2}. \quad (14)$$

Continuando, obtenemos un penúltimo centro de gravedad:

$$xi1..[m(i)-1] = \frac{\sum_{j=1}^{m(i)-1} Qijxij}{\sum_{j=1}^{m(i)-1} Qij}. \quad (15)$$

Junto con el último punto de demanda, llegamos a la Ecuación (11) (este procedimiento comúnmente se denomina "promedio ponderado"). □

La distancia cuadrada entre la zona de demanda y la red es la distancia cuadrada mínima entre el centro de la zona de demanda y un segmento de red. El siguiente teorema garantiza existencia y unicidad.

Teorema 5. Para una red dada $\overline{AjBj}^{s}_{j=1}$, y para cada Pi , donde i = 1, . . . , n, existe

$$yo^{2\,2\,re} = re(Pi) = \min_j \{d^2(Pi , \overline{AjBj})\}, \text{ para cada } i = 1, \ldots , \text{norte.} \quad (16)$$

Machine Translated by Google

En la prueba a continuación, mostramos la inducción como una sugerencia para aplicaciones de programas. En la sección Ejemplo, utilizamos únicamente la herramienta Excel para calcular la distancia mínima, aunque el número de segmentos y zonas es grande.

Prueba del teorema 5. Como la red es finita, la distancia mínima siempre existe.
A continuación, realizamos una inducción por índice j, 1 ≤ j ≤ s.

Para inducción básica j = 1, tenemos $d_i^2 = d^2(P_i, \overline{A_1B_1}) = \min_{j\le 1}\{d^2(P_i, \overline{A_jB_j})\}$. Para j = 2, $d_i^2 = \min\{d^2(P_i, \overline{A_1B_1}), d^2(P_i, \overline{A_2B_2})\}$, $\min_{j\le 1}\{d^2(P_i, \overline{A_jB_j})\} = \min_{j\le 2}\{d^2(P_i, \overline{A_jB_j})\}$.

Para el paso j = s − 1, suponemos que para la gráfica $\bigcup_{j=1}^{s-1}\overline{A_jB_j}$ existe $d_i^2 = \min_{j\le s-1}\{d^2(P_i, \overline{A_jB_j})\}$.

Si j = s, separamos $\bigcup_{j=1}^{s}\overline{A_jB_j} = \bigcup_{j=1}^{s-1}\overline{A_jB_j} \cup \overline{A_sB_s}$, y luego tenemos

$$d_i^2 = \text{mínimo}\left\{\min_{j\le s-1}\{d^2(P_i, \overline{A_jB_j})\}, d_i^2(P_i, \overline{A_sB})\right\}$$
$$= \min_{j\le s}\{d^2(P_i, \overline{A_jB_j})\},$$

para cada i = 1, . . . , n. Se ha demostrado que existe una distancia cuadrada para cada zona Fi con centro de demanda en el punto Pi . Es único porque si existe el mínimo, siempre es un solo valor.
□

A continuación, introducimos la "inercia" de la zona de demanda. Recordamos que en el concepto general de bondad de la red, una puntuación más alta indica un mejor servicio de red dentro del entorno.

Para determinar la bondad de la red, calculamos la inercia de la zona de demanda que "gravita" hacia la red. La inercia de una zona de demanda se determina multiplicando la demanda de la zona por la distancia cuadrada mínima entre el centro de la zona de demanda y un segmento de red. Formulamos la inercia de una zona de demanda utilizando la siguiente definición.

Definición 8 (Inercia de la zona de demanda). La inercia de la i-ésima zona de demanda es producto

$$M_i = Q(T_i) \cdot d_i^2, \quad i = 1, \ldots, n, \tag{17}$$

donde d_i^2 se calcula a partir de la Ecuación (16) y Q(Ti) se calcula en el Teorema 4.

Cabe señalar que no se consideran todas las zonas donde hay demanda de servicio de red para el cálculo de la inercia del sistema. Para utilizar el servicio de red, sus usuarios deben conectarse a la red de forma independiente.

En una red de infraestructura, dos tipos de distancia influyen en la elección de red por parte del usuario para un servicio. Una distancia corta atrae al usuario a elegir la red propuesta sobre otras opciones. Por ejemplo, si una parada de autobús está cerca del apartamento de alguien, esa persona solo usaría el autobús en lugar de su vehículo. Por el contrario, una distancia mayor es un límite más allá del cual el usuario utilizará indirectamente la red determinada. En la siguiente definición, definimos formalmente ambos tipos de distancias: distancias y zonas gravitantes y atractoras.

Definición 9 (Zonas y distancias gravitantes y atractoras). Zona de demanda Ti , donde i = 1, . . . , n, gravita hacia la red si y solo $d_i < D$ para la distancia gravitante D, que es un valor determinado de antemano.

Zona de demanda Tj , con j = 1, . . . ,n, es atraído por la red si ff $d_j < d$ para una distancia de atracción d, que también es un valor determinado de antemano.

Machine Translated by Google

En [14] se puede encontrar más información sobre d y D relacionados con las redes de transporte público. El valor de D representa una distancia aceptable de hasta dos kilómetros que se puede caminar para llegar a la red. Al mismo tiempo, d es la distancia de unos cientos de metros que un usuario está dispuesto a caminar en lugar de utilizar cualquier medio de transporte diferente.

La inercia total de una red es la suma de las inercias de todas las zonas de demanda gravitantes. La calidad de una red está determinada por qué tan baja es la inercia de su red. Un ejemplo sencillo se ilustra en la Figura 3.

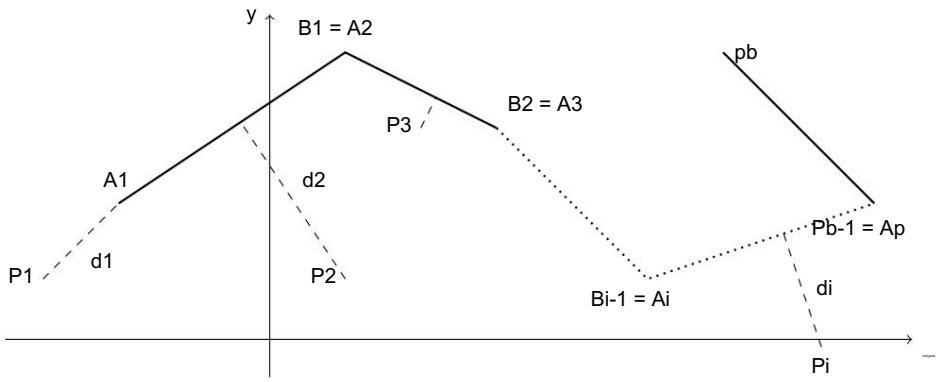


Figura 3. Cálculo de la inercia de la zona de momento a lo largo del camino en una red.

Definición 10 (Inercia de la red y número de zonas gravitantes). La inercia de la red se calcula sumando la inercia de todas las zonas gravitantes de la red.

$$m = \sum_{yo=1}^{k} Mi = \sum_{yo=1}^{k} Q(Ti)d_{i,}^{2} \quad (18)$$

Hay k zonas de demanda que gravitan hacia la red, cada una representada por Ti .

La longitud total de una red es la suma de las longitudes de sus segmentos.

Definición 11 (Longitud de la red). Para una red dada la longitud es $\overline{AjBj}$, $j=1^{s}$

$$l = \sum_{j=1}^{s} |AjBj| = \sum_{j=1}^{s} \sqrt{(xBj - xAj)^{2} + (yBj - yAj)^{2}}. \quad (19)$$

A la hora de calcular la bondad de una red tenemos en cuenta tres factores esenciales: el número de zonas gravitantes (k), la inercia (M) y la longitud (l). Una red se considera buena si tiene tantas zonas conectadas y su inercia y longitud son las menores posibles.

A continuación, proporcionamos una definición débil de una fórmula de función de bondad de red necesaria para probar las proposiciones como nuestros resultados principales.

Definición 12. Una función de bondad de red de tres variables se define adecuadamente mediante una fórmula g(k, M, l) si aumenta a medida que k aumenta y disminuye a medida que l y M disminuyen.

Cuando se define adecuadamente de acuerdo con la Definición 12, el valor de la función de bondad de la red proporciona una evaluación cuantitativa de la red en el sentido de que dos o más redes integradas en el mismo entorno se clasifican de manera que una red con un valor de bondad más alto sea la "mejor".

Proponemos tres fórmulas que aumentan con el número de zonas gravitantes y disminuyen con la longitud e inercia de la red. Se presentan en el orden en que

Machine Translated by Google

fueron desarrollados y tienen la misma prueba porque el aumento y la disminución se rigen por derivadas parciales positivas y negativas, respectivamente.

Proposición 1. Si k zonas son atraídas hacia la red y la red tiene una inercia de M con una longitud de l, entonces la función de bondad de la red viene dada por la fórmula

$$3\ g1(k, M, l) = l\ \frac{k\ 2 + k}{2\ + M}. \tag{20}$$

La Proposición 1 describe cómo la Ecuación (20) cuantifica la bondad de la red mediante la Definición 12. A continuación, analizamos las potencias atribuidas a l y k en (20). Observamos que la dimensión de l 2 en el denominador es la misma que la de M, es decir, km2 . El nominador resalta el número de zonas gravitantes, comparándolo con los grados del valor del denominador. Es decir, l 2 está en grado dos, pero M de (18) está en total grado tres ya que se obtiene multiplicando la demanda por la distancia al cuadrado. De acuerdo con la Definición 12, la Ecuación (20) define una función de bondad correcta en tres variables: k, M y l.

Las derivadas parciales de la ecuación (20) vienen dadas por

$$\frac{\partial g1(k, M, l)}{\partial k} = \frac{2k + 3k\ l\ ^2}{2\ + M\ k} = \frac{2 + 3k}{k\ 3}\ g1(k, M, l), k + \tag{21}$$

$$\frac{\partial g1(k, M, l)}{\partial M} = - \frac{2 + k\ 3}{2\ + M)\ ^2} = \frac{-1}{+ M}\ g1(k, M, l), l\ 2 \tag{22}$$

$$\frac{\partial g1(k, M, l)}{\partial l} = -2l\ (l\ \frac{(lk\ 2 + k\ 3}{2\ + M)\ ^2} = \frac{^{-2\ litros}}{2 + M}\ g1(k, M, l).\ l \tag{23}$$

Debido a que la longitud de la red suele ser superior a medio kilómetro, de las ecuaciones (21) a (23), se cumple la siguiente desigualdad:

$$\frac{\partial g1(k, M, l)}{\partial k}\ 0 > \partial M\ \frac{\partial g1(k, M, l) >}{} > \frac{\partial g1(k, M, l)}{\partial l}. \tag{24}$$

Proposición 2. Dados k, M y l de la Proposición 1, la función de bondad de la red se puede expresar mediante la siguiente ecuación:

$$g2(k, M, l) = + l\ 2\ \frac{k\ 2\ k\ 3}{}\ \frac{}{_{METRO}}. \tag{25}$$

A continuación, ofrecemos las derivadas parciales de la función en (25):

$$\frac{\partial g2(k, M, l)}{\partial k} = \frac{2\ 2k\ 3k +}{_{yo\ 2}}\ \frac{}{M}$$

$$\frac{\partial g2(k, M, l)}{\partial M} = - \frac{k}{3}$$

$$\frac{\partial g2(k, M, l)}{\partial l} = - \frac{M2}{2k\ 2\ l\ 3}.$$

Esto conduce a las siguientes desigualdades:

$$\frac{\partial g}{\partial l}\ \partial g \geq 0 \geq \partial k, \frac{\partial g}{\partial M}. \tag{26}$$

Proposición 3. Sean k, M y l como en las proposiciones anteriores. Sean d y D valores positivos de la Definición 9 y se calcule Q(Ti) como en el Teorema 4.

En el entorno dado, n representa el número total de zonas de demanda que la red pretende atender. Entonces, la función de bondad de la red se define mediante la ecuación

Machine Translated by Google

$$\text{gramo}(k, M, l) = 2\,2 \cdot \text{norte}^{\frac{k}{2 \cdot \mu(U) + \cdot D \cdot l}} \quad \frac{k\,3 \cdot d^{\,2} \cdot \sum_{yo=1}^{k} Q(Ti)}{2 \cdot \text{norte}^{3} \cdot M}. \tag{27}$$

La ecuación (27) no tiene dimensiones. Esto es particularmente valioso cuando todas las zonas de demanda gravitan atractivamente hacia la red en el sentido de la Definición 9. Teniendo en cuenta los valores de n, d, D y el valor de la demanda total $\sum_{yo=1}^{\text{norte}} Q(Ti)$, el siguiente resultado valida el uso de la ecuación (25) en la función de bondad de la red.

Corolario 1. Si n = k, μ(U) = D · l, y di = d para cada Ti U, entonces el valor calculado por la Ecuación (27) es igual a uno, es decir, 1 = 100%.

Prueba del Corolario 1. Si las distancias de todas las zonas a la red son di = d, y si todas zonas gravitan hacia la red con k = n, entonces (18) ofrece $M = d^{\,2} \cdot \sum_{yo=1}^{\text{norte}} Q(Ti)$. Si di ≤ d, La ecuación (27) produce un resultado superior al 100% y la red es demasiado buena o está sobrecapacitada. □

Si se cumplen las condiciones del Corolario 1 , entonces decimos que la red está disponible con distancia D y atractivo con distancia d. A continuación demostramos todas las proposiciones.

Las pruebas de las proposiciones 1-3. En el sentido de la Definición 12, las pruebas de todas las proposiciones son las mismas ya que se satisfacen las solicitudes de base. Siguiendo (26), todos los valores g1, g2 y g aumentan con k y disminuyen con l y M, según sea necesario. □

En la Ecuación (27), la potencia en n cuando n^{2} se utiliza porque en nuestra sección de ejemplo $n^{2} \geq \mu(U) \leq \leq \mu(U)$. cubo sustantivo, masculino— ^{3} se utiliza porque en nuestro ejemplo, $n^{3} \geq \sum i=1^{\text{norte}} Q(Ti)$ cuando $n \sum_{yo=1}^{\text{norte}} Q(Ti)$.

Para evitar la confusión que pueda surgir por la aparición de exponentes cuadrados y cúbicos en las ecuaciones (20) a (27), agregamos el siguiente teorema que usa la función techo x de un número real x para determinar el entero más bajo mayor que o igual ax.

Teorema 6. Sean los valores k, l, M como en la Proposición 1 y sean n, d, D y Q(Ti) como en la Proposición 3. Definimos α y β de la siguiente manera:

$$\alpha = \log n\, \mu(U) \quad \text{y } \beta = \log n\, Q(Ti) \quad . \sum_{yo=1}^{\text{norte}} \tag{28}$$

Si α, β > 0, entonces el valor de la función

$$G(k, M, l) = \frac{k\,\alpha \cdot \mu(U)\,2 \cdot}{\text{norte}^{\,\alpha} \cdot D \cdot l} + \frac{k\,\beta \cdot d\,2 \quad \cdot \sum_{yo=1}^{\text{norte}} Q(Ti)}{\cdot \text{norte}\beta \cdot M\,2} \tag{29}$$

depende cada vez más de k y decrecientemente de ly M. Dependiendo de valores arbitrarios para d y D, G puede ser menor o mayor que 100%.

Prueba ∂G . De nuevo, derivadas parciales $\frac{}{\partial k} > 0 > \frac{\partial G}{\partial l}, \frac{\partial G}{\partial M}$. □

Observación 1. La fórmula (27) es un caso especial de (29) cuando α = 2 y β = 3.

Machine Translated by Google

En la siguiente sección, proponemos una red de corredores de Transporte Público Urbano (PCT). El método combina física ambiental y observaciones sociales, estadísticas de comportamiento y heurística, y aplica la teoría propuesta para evaluar varias variantes de red.

4. Ejemplo El

transporte urbano es un pilar crucial que apoya la movilidad ciudadana en ciudades que experimentan crecimiento demográfico y expansión territorial. El transporte privado tiende a congestionar el tráfico urbano hasta el punto de no poder conectar localidades dentro del área urbana.

Es necesario invertir en nuevos modos de transporte público para atraer a los habitantes a utilizarlo gracias a su disponibilidad, velocidad y capacidad. Los nuevos modos necesitan nuevos corredores. A continuación se muestra un ejemplo del procedimiento propuesto para el diseño heurístico y la evaluación de propuestas de redes de corredores de transporte público urbano. A continuación, demostramos el proceso paso a paso.

El objetivo inicial era proponer una red para el modo de transporte público con gran capacidad y velocidad. Por lo general, las redes se integran mediante el árbol de expansión o el método de mínimos cuadrados, pero propusimos una red después de varias variantes dibujadas heurísticamente y elegimos la mejor. En la sección anterior, propusimos varias funciones para calcular la bondad de una red integrada en un entorno para atender las demandas de tráfico de los usuarios. En esta sección, los aplicamos para comparar varias variantes de red obtenidas heurísticamente luego de exhaustivas investigaciones estadísticas.

El principal requisito previo era tener en cuenta los obstáculos naturales del entorno. En la Figura 4, se representan cuatro barreras rojas en la ciudad de Zagreb, de norte a sur: la montaña, el ferrocarril, el río y la carretera.

Figura 4. Entorno natural del condado de la ciudad de Zagreb.

La división de zonas se adoptó de la antigua división administrativa de comunidades locales y se muestra en la Figura 5. Los puntos de demanda son edificios, escuelas, supermercados y otros lugares donde pueden aparecer personas. La demanda es el número de habitantes o usuarios extraído de varios informes. Los centros de demanda se calculan mediante (11) después de establecer el sistema de coordenadas , según el Lema 1. También es necesario calcular el momento de inercia.

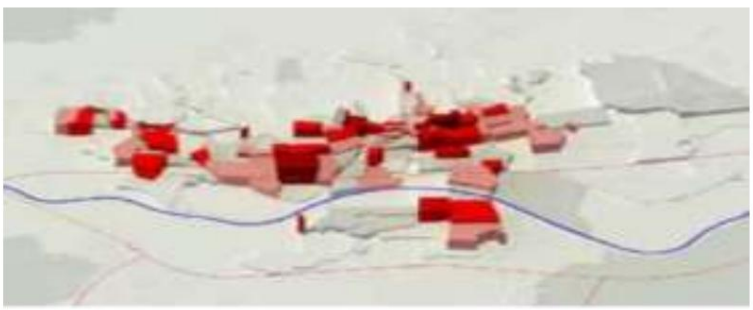


Figura 5. Centros de población locales de la ciudad de Zagreb. Cuanto más oscuro sea el color rojo, mayor será la población en ese lugar. Fuente: Plan de Zagreb [15].

Machine Translated by Google

Una vez que se dividió el entorno (ver Figura 5), y teniendo en cuenta las barreras que se muestran en la Figura 4, dibujamos corredores compuestos por segmentos (ver Ecuación (1)). Los segmentos se integraron de acuerdo con las principales direcciones del flujo de tráfico. Las direcciones del flujo se obtuvieron después del monitoreo estadístico y se presentan en la Figura 6.

Figura 6. Flujos de pasajeros de la ciudad de Zagreb.

El proceso de diseño heurístico se guió a través de las nueve variantes de la red de corredores presentadas en las Figuras 7 a 9. La inercia de la red para cada uno de ellos se calculó según las Ecuaciones (17) y (18).

Figura 7. Las primeras variantes de red de corredores propuestas.

Figura 8. Segunda variante de red de corredores propuesta.

Machine Translated by Google

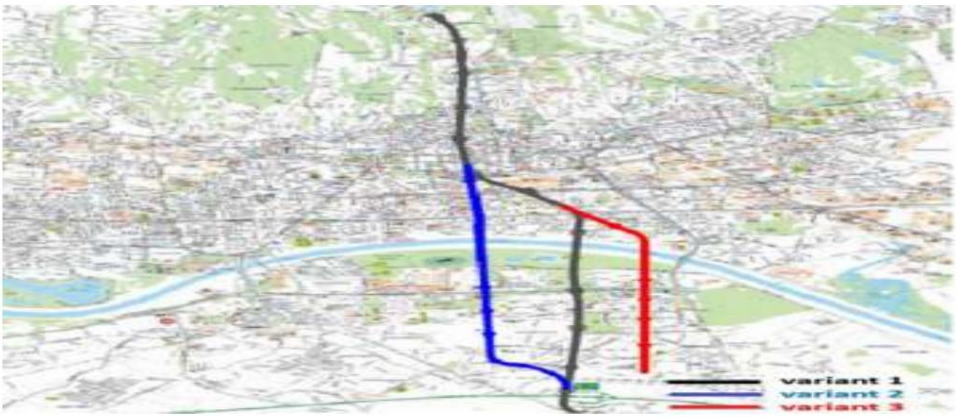


Figura 9. La tercera red de corredores propuesta.

Las variantes de la red de corredores en la Figura 7 están dibujadas en un círculo, basándose en la fluye de la Figura 6. Solo las zonas hasta D = 4 km se observan como zonas gravitantes para cálculo de inercia.

Las variantes de la Figura 9 están dibujadas según las direcciones norte-sur de la Figura 6. Los corredores nuevos y modificados implican la consideración de la Figura 5.

Los corredores en la Figura 8 se extienden en dirección oeste-este siguiendo los flujos de Figura 6. Luego, estas direcciones se conectan con líneas en dirección norte-sur. El El proceso continúa considerando la dirección oeste-sur en la Figura 6. Un buen tráfico público La red atrae a los ciudadanos a utilizar la red de tráfico público en lugar del transporte privado. dispositivos si se crea para reducir la cantidad de caminata necesaria, según lo expresado por la red inercia (ver Ecuación (18)).

Las fórmulas propuestas en las Proposiciones 1 a 3 se utilizaron para comparar cuantitativamente diferentes redes de corredores, considerando la disponibilidad de la red a través del número de corredores gravitantes. zonas de demanda y costos de red aproximadamente proporcionales a la longitud de la red.

En la ciudad de Zagreb, el tamaño del área ambiental es μ(U) = 250 km2, la demanda de tráfico es $\sum_{yo=1}^{norte} Q(Ti) = 680.000$ y el número total de piezas del entorno es n = 160. Valores d = 200 m y D = 4 km se tomaron siguiendo las recomendaciones de [14].

En las Figuras 7 a 9 se proponen tres tipos de variantes, según el flujo de pasajeros. instrucciones de la Figura 6. En la Tabla 1, se enumeran los valores de bondad para cada variante, y los mejores Se consideran los valores de bondad de la variante. Luego, recopilamos la red final en la Figura 10. tomando las primeras variantes de las Figuras 7 y 8, y la tercera variante de la Figura 9.

Tabla 1. Bondad de la red de las redes utilizando las Ecuaciones (20), (25) y (27).

Cifra	Variante	M (km2)	l (km)	k	g1/km2	g2/km2	gramo/%
Figura 7	1	27.000	22.3	47	3.86	8.29	13,33%
	2	17.000	27,5	36	2.70	4.46	6,65%
	3	9.000	16.4	21	1.05	2.67	3,61%
Figura 8	1	17.000	19.6	36	1.05	6.12	8,95%
	2	13.000	19.2	33	2.76	5.72	7,82%
	3	11.000	18.7	28	2.00	4.24	5,76%
Figura 9	1	13.000	11.8	21	2.77	3,88	4,78%
	2	4000	13.7	18	2.00	3.18	3,36%
	3	10.000	9.8	20	0,83	4.96	5,23%
Figura 10	final	58.000	53,7	213	159	182	81,52%
Figura 11	tranvía	25.000	116	93	21	107	19,80%

A modo de comparación, calculamos los valores para el tranvía temporal para electricistas de Zagreb. red que se muestra en la Figura 11.

Machine Translated by Google

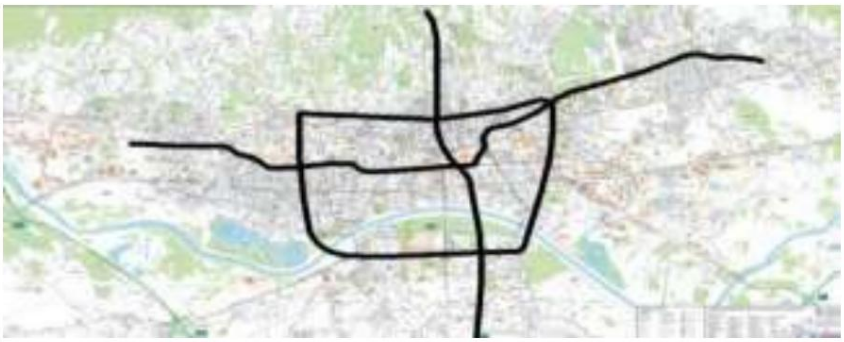


Figura 10. Variante final de red de corredores propuesta.

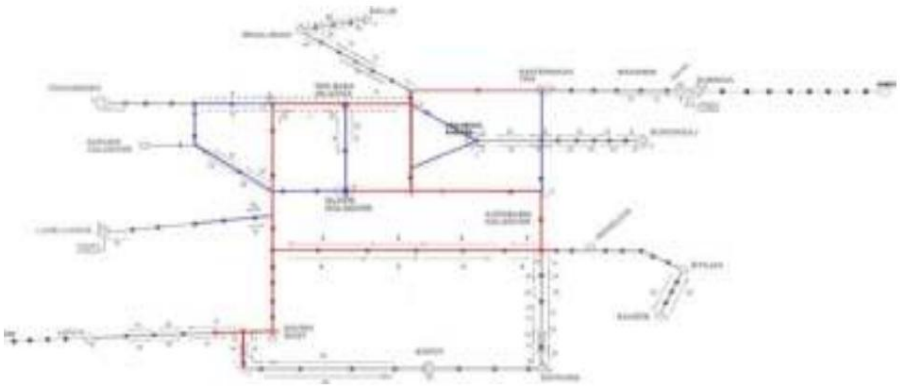


Figura 11. Red ZET existente. Fuente: Horario de la sucursal ZET (2020).

5. Conclusiones

Ofrecemos expresiones matemáticas que cuantifican la bondad de diferentes redes en función de (i) cobertura, (ii) disponibilidad y (iii) gastos de mantenimiento. Aunque el modelo se desarrolla heurísticamente, demostramos que puede ser una herramienta confiable para evaluar varias redes sugeridas para integrar en un entorno.

El debate en curso sobre el uso de heurísticas para crear fórmulas matemáticas implica el uso de reglas generales, la intuición y el sentido común para encontrar una solución que sea lo suficientemente buena y rápida. Las heurísticas basadas en la experiencia y la intuición son a veces mejores que los métodos exactos o los procedimientos de simulación exhaustivos. Además, si no hay nada similar en la literatura disponible para evaluar cuantitativamente una red con un valor único, entonces el enfoque heurístico ofrece una buena solución.

Sin embargo, es crucial evaluar los resultados de la heurística. La primera cuestión que abordamos es seleccionar la mejor variante de red entre las variantes obtenidas heurísticamente. Proponemos varias variantes para resolver el problema en la Sección 4 y justificamos nuestro enfoque para elegir la mejor opción.

Finalmente, planeamos desarrollar este trabajo aún más en el futuro.

Contribuciones de los autores: Conceptualización y validación, MB; metodología y análisis formal, BI; investigación, recursos y preparación del borrador original, MP Todos los autores han leído y aceptado la versión publicada del manuscrito.

Financiamiento: Esta investigación no recibió financiamiento externo.

Declaración de la Junta de Revisión Institucional: No aplicable.

Declaración de disponibilidad de datos: No se aplica el intercambio de datos.

Agradecimientos: Agradecemos a la Universidad de Ciencias Aplicadas de Krapina y a la Facultad de Ingeniería Sami Shamoon por su apoyo moral.

Conflictos de intereses: Los autores declaran no tener conflictos de intereses.

Machine Translated by Google

Referencias

1. Kanellopoulos, D.; Sharma, VK; Panagiotakopoulos, T.; Kameas, A. Arquitecturas y protocolos de red para aplicaciones de IoT en ciudades inteligentes: desarrollos recientes y perspectivas. Electrónica 2022, 12, 2490. [CrossRef]
2. Karakus, M.; Durresi, A. Calidad de servicio (QoS) en redes definidas por software (SDN). Una supervivencia. J. Net. Computadora. Aplica. 2017, 80, 200–218. [Referencia cruzada]
3. Smit, LC; Dikken, J.; Schuurmans, MJ; de Wit, Nueva Jersey; Bleijenberg, N. Valor del análisis de redes sociales para desarrollar y evaluar intervenciones sanitarias complejas: una revisión del alcance. BMJ Abierto 2020, 10, e039681. [Referencia cruzada] [PubMed] [PubMed Central]
4. Li, J.; Duro, D.; Hu, X.; Tang, J.; Yi, C.; Liu, H. Integración de redes atribuidas para el aprendizaje en un entorno dinámico. En Actas del CIKM'17, Singapur, 6 a 10 de noviembre de 2017; Sesión 2D: Integración de red 2. [CrossRef]
5. Salkhi Khasraghi, G.; Volchenkov, D.; Nejat, A.; Hernández, R. El campus universitario como una red dinámica peatonal compleja: Un estudio de caso sobre patrones de transitabilidad para peatones en la Universidad Tecnológica de Texas. Matemáticas2023, 12, 140. [CrossRef]
6. Bogdanova, TV; Evdokimov, KA Componente de transporte de un entorno urbano confortable. Inov. Inversiones 2021, 2021, 163–170.
7. Colón, C.; Hallegatte, S.; Rozenberg, J. Análisis de criticidad de la red de transporte de un país a través de una cadena de suministro basada en agentes. modelo. Nat. Sostener. 2021, 4, 209–215. [Referencia cruzada]
8. Guillot, M.; Furno, A.; Aghezzaf, E.; El Faouzi, N. Reducción de la red de transporte basada en una subred óptima. Comunitario. Transp. Res. 2022, 2, 100079. [Referencia cruzada]
9. Wang, S.; Wang, Z. Estrategia de regulación del volumen de transporte y desarrollo colaborativo para una aglomeración urbana. Sostenibilidad 2022, 15, 14742. [CrossRef]
10. Kuo, Y.; Leung, JM; Yan, Y. Transporte público para ciudades inteligentes: innovaciones recientes y desafíos futuros. euros. J. Ópera. Res. 2023, 306, 1001–1026. [Referencia cruzada]
11. Nes, VR; Hamerslag, R.; Immers, R. Diseño de Redes de Transporte Público. Transp. Res. Rec. 1988, 1202, 74–83.
12. Wilson, RJ Introducción a la teoría de grafos, 4ª ed.; Longman Group Ltd.: Harlow, Reino Unido, 1996. Disponible en línea: https://www.maths.ed.ac.uk/~v1ranick/papers/wilsongraph.pdf (consultado el 20 de febrero de 2023).
13. Lang, S. Introducción al álgebra lineal. En Springer Books sobre matemáticas elementales; Springer: Berlín/Heidelberg, Alemania, 1986; ISBN 978–1-4612-7002-7. Disponible en línea: https://www.math.nagoya-u.ac.jp/~richard/teaching/f2014/Lin_alg_Lang.pdf (consultado el 15 de abril de 2024). 14. Šimunovi´c, LJ.; Cosi´c, M. Osobitosti Pješa˘ckoga Prometa. En Nemotorizirani Promet; Hrvoje Oro; Facultad de Ciencias del Tráfico y el Transporte : Zagreb, Croacia, 2015; págs. 3–9.
15. Disponible en línea: https://www.zagreb.hr/UserDocsImages/arhiva/zagrebplan-ciljeviiprioritetirazvojado2020.pdf (accedido en 12 de diciembre de 2020).

Descargo de responsabilidad/Nota del editor: Las declaraciones, opiniones y datos contenidos en todas las publicaciones son únicamente de los autores y contribuyentes individuales y no de MDPI ni de los editores. MDPI y/o los editores renuncian a toda responsabilidad por cualquier daño a personas o propiedad que resulte de cualquier idea, método, instrucción o producto mencionado en el contenido.

Machine Translated by Google

AppliedMath

MDPI

Artículo

Estrategias financieras a largo y mediano plazo sobre acciones utilizando Redes bayesianas dinámicas

carlos lewis †, Marco Antonio Caruana †, y David Paul Suda *,†

Departamento de Estadística e Investigación Operativa, Facultad de Ciencias, Universidad de Malta, 2080 Msida, Malta; karl.lewis.17@um.edu.mt (KL); mark.caruana@um.edu.mt (MAC)

* Correspondencia: david.suda@um.edu.mt; Tel.: +356-9986-9884 Estos

† autores contribuyeron igualmente a este trabajo.

Resumen: Diseñar una estrategia comercial financiera que permita ganancias a largo plazo es un problema muy común en las finanzas. Este artículo tiene como objetivo formular un marco matemáticamente riguroso para el problema y comparar y contrastar los resultados obtenidos. El principal enfoque considerado se basa en Redes Bayesianas Dinámicas (DBN). Dentro del entorno DBN, se consideran y aplican una estrategia comercial tanto a largo como a corto plazo en doce acciones obtenidas de mercados desarrollados y en desarrollo. Se concluye que tanto las estrategias a largo plazo como las de mediano plazo propuestas en este documento superan la estrategia comercial de referencia de comprar y mantener (B&H). A pesar de las claras ventajas de las estrategias comerciales anteriores, se discuten las limitaciones de este modelo junto con sus posibles mejoras.

Palabras clave: finanzas; redes bayesianas dinámicas; estrategias comerciales; acciones

1. Introducción

En este artículo, se utilizan redes bayesianas dinámicas (DBN) para estudiar el problema de obtener y probar una estrategia financiera cuyo rendimiento sea mayor que la estrategia de comprar y mantener para una acción determinada. Una red bayesiana (BN) es una representación gráfica y compacta de una función de densidad de probabilidad conjunta (PDF) que hace uso de la independencia condicional y se puede utilizar para modelar un sistema en una instancia de tiempo. Las DBN extienden la BN a más de un período de tiempo, es decir, las DBN son una extensión de las BN impulsada temporalmente.

Algunos beneficios de utilizar DBN sobre otros modelos, como los modelos básicos de series de tiempo, son los siguientes. En primer lugar, las DBN tienen la capacidad de incorporar lo que se conoce como variables ocultas (o latentes), que gobiernan, o se cree que gobiernan, las observaciones de la variable aleatoria observable. Estas variables ocultas son variables aleatorias cuyo verdadero valor no se puede medir directamente, a veces debido a insuficiencias instrumentales y otras simplemente porque se supone que existen hipótesis sobre la existencia de estas variables. Ejemplos de tales variables ocultas son la inteligencia de un individuo (que no puede medirse con precisión utilizando ningún instrumento) o el estado (alcista o bajista) de un mercado financiero (cuya existencia en sí es hipotética). En segundo lugar, dado que son una rama de los BN, los DBN permiten representaciones compactas de una distribución conjunta que de otro modo sería posiblemente engorrosa. Russell y Norvig [1] sostienen que, debido a la independencia condicional, representar una distribución conjunta como un DBN puede eliminar cualquier término redundante, lo que, a su vez, también simplifica la estimación de parámetros. Murphy [2] afirma que otro beneficio de los DBN es la facilidad con la que se pueden introducir variaciones en el modelo en cuestión. Los DBN se pueden utilizar para representar alternativamente tanto modelos ocultos de Markov (HMM) como HMM con variaciones, como HMM con una mezcla de salidas gaussianas, HMM autorregresivos, HMM de entrada-salida y HMM jerárquicos. Los DBN también se pueden utilizar para representar modelos de filtro de Kalman donde, a diferencia de los HMM, el estado oculto es continuo en lugar de discreto. Por último, las DBN también tienen la capacidad de responder preguntas sobre diferentes tipos de razonamiento o inferencia, como

check for updates

Cita: Lewis, K.; Caruana, MA; Suda D. Estrategias financieras a largo y medio plazo sobre acciones utilizando redes bayesianas dinámicas.

Matemáticas aplicadas 2024, 4, 843–855. https://doi.org/10.3390/appledmath4030045

Editor académico: Tommi Sottinen

Recibido: 28 de abril de 2024
Revisado: 21 de junio de 2024
Aceptado: 24 de junio de 2024
Publicado: 3 de julio de 2024

Copyright: © 2024 por los autores. Licenciatario MDPI, Basilea, Suiza.

Este artículo es un artículo de acceso abierto. distribuido bajo los términos y condiciones de los Creative Commons Licencia de atribución (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

Machine Translated by Google

como inferencia diagnóstica (de efectos a causas), inferencia causal (de causas a efectos), inferencia intercausal (entre causas del mismo efecto) e inferencia mixta (una mezcla de dos de los tipos de inferencia anteriores).

En este trabajo de investigación utilizamos la relación precio-beneficio (PE) en los modelos analizados. En la literatura se pueden encontrar varios artículos que aplican dicho ratio PE para examinar las estrategias de inversión. En particular, se puede mencionar a Basu [3]. Esta relación también se ha utilizado en una amplia gama de literatura, incluida la de Lleó [4] y Angelini [5]. Chang y Tian [6] también utilizan BN para modelar las relaciones cualitativas y cuantitativas entre varias variables que afectan la dinámica del índice bursátil S&P 500, con el objetivo de optimizar las decisiones comerciales, es decir, cuándo abrir una posición corta y cuándo invertir en una posición larga. Damián et al. [7] basan su trabajo en un estudio anterior, el de Tayal [8], que emplea HHMM para analizar datos financieros. Los resultados obtenidos por Tayal se reproducen en el artículo de Damiano et al. y se brindan más información. Damián et al. [7] concluyen que la inferencia probabilística permite la identificación de dos estados distintos en datos de alta frecuencia que están marcados principalmente por presión de compra y venta.

Históricamente, se cree que Paul Dagum impulsó el desarrollo de las DBN a principios de la década de 1990, concretamente en el trabajo de Dagum et al. [9] en 1992 y por Dagum et al. [10] en 1995. Desde entonces, los DBN se han utilizado en diversas áreas, a saber, en finanzas y economía. Sin embargo, otras áreas incluyen (pero no se limitan a) reconocimiento de voz (Murphy [2], Zweig, Russell [11] y Nefian et al. [12]), biología (Yao et al. [13], Raval et al. [14], Murphy y Mian [15]), robótica (Prembida et al. [16], Patel et al. [17]), procesamiento de imágenes (Delage et al. [18]) y detección de fallas (Cozar et. otros [19]). Benhamou et al. analizan las técnicas de estimación de parámetros de DBN . [20], mientras que Murphy [21] trabaja en los fundamentos teóricos de las DBN.

El resto de este artículo está estructurado de la siguiente manera. La sección 2 presenta los fundamentos teóricos de las DBN. En la Sección 3, los DBN se aplican a doce acciones de mercados tanto desarrollados como en desarrollo, y se presenta el retorno de la inversión. La razón para considerar tanto los mercados desarrollados como los en desarrollo es probar la solidez del modelo bajo diferentes condiciones económicas y patrones de crecimiento. Finalmente, en el apartado de conclusiones se destacan los puntos clave y los resultados obtenidos.

2. Marco teórico

Este modelo se basa en el trabajo de Wang [22] y explota dos fenómenos financieros conductuales: la volatilidad conductual y la reversión a la media. La volatilidad del comportamiento se refiere al fenómeno por el cual eventos relacionados con el mercado, como operaciones irracionales ejecutadas por operadores sin experiencia, hacen que el precio de negociación de una acción se desvíe de su valor "verdadero". Eventos como estos afectan continuamente el precio de mercado, ya sea hasta que los efectos causados se anulan entre sí , o hasta que los inversores racionales equilibran estos efectos con sus operaciones racionales. Este fenómeno se conoce como reversión a la media. La teoría de la reversión a la media es aplicable no sólo al precio de una acción, sino a cualquier métrica relacionada con el precio, como la relación precio-beneficio (PE) (precio de la acción dividido por el beneficio por acción). Como resultado, uno puede decidir comprar (o vender) una acción sólo cuando alguna métrica que sigue la teoría de reversión a la media está muy lejos de la media, sabiendo que el valor "verdadero" de la métrica volverá en algún momento del año. el futuro.

Un modelo gráfico dirigido (DGM) es una representación de un modelo probabilístico que utiliza supuestos de independencia condicional a través de un gráfico dirigido, donde los vértices (tenga en cuenta que los términos “estado”, “variable aleatoria”, “nodo” y “vértice” pueden ser usados indistintamente) del gráfico representan variables aleatorias y los arcos representan la dependencia condicional entre pares de variables aleatorias. Un arco desde el vértice A hasta el vértice B representa la afirmación “A causa B” (Murphy [2]). Los DGM son convenientes debido a su compacidad al exhibir alguna distribución de probabilidad conjunta: para N variables aleatorias binarias, la forma general cerrada de la distribución de probabilidad conjunta de estas variables aleatorias puede necesitar O(2 N) parámetros, mientras que el modelo gráfico puede dar los mismos información utilizando menos parámetros debido a la omisión de términos mediante declaraciones de indepen

Machine Translated by Google

Una Red Bayesiana es un DGM que representa un conjunto de variables aleatorias X1, ..., Xn y sus dependencias condicionales a través de un gráfico acíclico dirigido (DAG), denotado por G = (V, E). Cada vértice en V representa una variable aleatoria Xi , y cada arista dirigida en E representa la dependencia condicional que tiene una variable aleatoria de otra. Se asocia una distribución de probabilidad condicional a cada nodo Xi , y la distribución de probabilidad conjunta en el conjunto de vértices V viene dada por P[V] = ∏ P[Xi | padres (Xi)]. La generalización de una BN a múltiples intervalos de tiempo da lugar a la definición de una DBN. Una Red Bayesiana Dinámica se define como un par (B0, B→), donde B0 es el BN que representa la probabilidad previa (o la probabilidad del estado inicial), es decir, la distribución de probabilidad de las variables aleatorias en el tiempo 0, y B → es un BN temporal de dos sectores que describe las probabilidades de transición desde el tiempo t − 1 hasta el tiempo t para cualquier nodo X en el conjunto de vértices V del DAG G = (V, E), denotado por P[Xt |Xt−1] .

La distribución de probabilidad conjunta del conjunto de vértices V en todos los intervalos de tiempo viene dada por:

$$P[V] = P[V0]\, P[V1{:}T] = \prod_{Xi \in V} P[Xi0 \mid \text{padres}(Xi0)] \prod_{t=1}^{t} \prod_{Xi \in V} P[\text{Salir} \mid \text{padres}(\text{salir})],$$

donde Vt1 :t2 se refiere al conjunto de todos los vértices indexados desde el momento t1 hasta el momento t2. Un DBN puede parametrizarse mediante su matriz de transición At = P[Xt = j | Xt−1 = i], y su distribución a priori π(i) = P[X1 = i]. Si se supone que las probabilidades de transición son constantes para todos los intervalos de tiempo (At = A), entonces se dice que son homogéneas y tienen una función de distribución conjunta mucho más compacta . Cuando se representa gráficamente una DBN, normalmente se muestran dos o tres intervalos de tiempo (el intervalo de tiempo inicial y uno o dos posteriores), ya que se supone que su estructura se replica a lo largo del tiempo.

Una propiedad importante de los DAG es que “los nodos se pueden ordenar de manera que los padres aparezcan antes que los hijos. Esto se llama ordenamiento topológico y se puede construir a partir de cualquier DAG”. (Murphy [23]). Dado tal ordenamiento, la propiedad de Markov ordenada (o propiedad de Markov local) se define como la “suposición de que un nodo sólo depende de sus padres, no de todos sus predecesores en el ordenamiento” (Murphy [23]). Esta suposición es de hecho una generalización de la propiedad de Markov para cadenas de Markov.

2.1. Inferencia para

DBN El objetivo de la inferencia es inferir el valor de los estados latentes dadas las observaciones de los estados observables, es decir, inferir los marginales P[Xt = i | y1:τ]. Si τ = t, el proceso se conoce como filtrado (o 'now-casting'); si τ > t, entonces esto es suavizamiento; y si τ < t, entonces se estaría realizando una predicción (o previsión). Un algoritmo de inferencia comúnmente utilizado para DBN es el algoritmo hacia adelante-hacia atrás. En este algoritmo, la programación dinámica se implementa a través de dos pasos, conocidos como pases, que se ejecutan en sentido contrario : uno avanza en el tiempo y el otro retrocede. Tenga en cuenta que se supone que se conocen la matriz de probabilidad de transición, la matriz de probabilidad de emisión y las probabilidades anteriores . En el pase hacia adelante, el valor de αt(i) := P[Xt = i | y1:t] se encuentra de forma recursiva. En el pase hacia atrás, el valor de βt(i) := P[yt+1:T | Xt = i] también se encuentra de forma recursiva pero moviéndose en orden contracronológico (del tiempo T al tiempo 2). Una vez completados los pases hacia adelante y hacia atrás, el valor de γt(i) := P[Xt = i | y1:T] finalmente se puede obtener:

$$P[Xt = i \mid y1{:}T] = \frac{P[yt+1{:}T \mid Xt = i]P[Xt = i \mid y1{:}t\,]}{P[y1{:}T]} \quad \gamma t(i) =$$

donde $P[y1{:}T] = \prod_{T\ t=1\ ct}$.

2.2. Aprendizaje de

DBN En este contexto, el aprendizaje se refiere al proceso de estimación de parámetros. En la estimación de parámetros, el aprendizaje se puede abordar mediante un enfoque de Máxima Verosimilitud (ML) o mediante el enfoque de Máximo A Posteriori (MAP). Si se utiliza ML, los datos se utilizan para obtener estimaciones de parámetros mediante la solución del problema de optimización:

Machine Translated by Google

θ_{ml} = arg máxθ P[Y | θ] = arg maxθ log P[Y | θ] , donde θ es el conjunto de parámetros a estimar. Normalmente, θ contiene la matriz de transición y los parámetros relacionados con la distribución de probabilidad utilizada en la matriz de emisión. Por otro lado, si se utiliza el método MAP, el problema de optimización anterior se ajusta ligeramente para convertirse en θ arg maxθ log P[Y | θ] + log P[θ] , donde P[θ] es la distribución previa del parámetro. El $= _{MAPA}$ enfoque para resolver los problemas de optimización anteriores es a través de una adaptación del algoritmo de Expectativa-Maximización (EM) conocido como algoritmo de Baum-Welch , que ha demostrado brindar un óptimo local a los problemas de optimización anteriores (Baum et al. [24], Dempster y otros [25]). Utiliza el algoritmo de avance-retroceso como subrutina. Por lo tanto, en los casos en que se desconocen los parámetros del modelo, primero se utiliza el algoritmo de Baum-Welch para estimar (o aprender) los parámetros del modelo. Luego, se utiliza el algoritmo de avance-retroceso para inferir los marginales posteriores.

2.3. Aplicación de la teoría La

principal hipótesis del modelo es que el precio de las acciones de una empresa no siempre es igual al "verdadero" valor intrínseco de la empresa. Utilizando los fenómenos de volatilidad conductual y reversión de la media, los efectos temporales que hacen que las métricas de las acciones se desvíen de su valor "real" se clasifican en dos: efectos a corto plazo (duración de unos pocos días) y efectos a mediano plazo (duración de varias semanas).).

Se plantea la hipótesis de que el valor fundamental de una empresa i en el momento t, denotado por $P_{él}$, es directamente proporcional a sus ganancias anuales en el momento t, denotado por Ei,t , con el factor fundamental Relación PE, denotada por $PE_{i,t}$, actuando como constante de proporcionalidad:

$$Fosa = (PE_{i,t})(Ei,t).$$

Por otro lado, el ratio PE observado (disponible abiertamente en el dominio público), Pi,t donde Pi,t viene dado por PEi,t = Ei,t empresa i en el denota el precio de negociación real de a denotado por PEi,t , momento t.

Tenga en cuenta que a partir de este momento, el índice i se elimina, ya que el análisis de las acciones es realizado de forma univariada. La ecuación modelo que resulta de lo anterior viene dada por:

$$Pto = P\ t\ (1 + Zt)(1 + \varepsilon t) = \quad yt = \ln PE \quad (1 + Zt) + \ln 1 + \varepsilon t$$

$$= \quad yt = \ln PE \quad (1 + Zt) + \varepsilon t \tag{1}$$

donde yt = ln Pt efectos , {Zt}t N es una cadena de Markov de tiempo discreto que modela el mediano plazo) del ruido, y εt N (0, σ Debido a la ^{2} es una variable aleatoria que modela los efectos del ruido a corto plazo. ecuación del modelo anterior, el modelo utilizado sólo es aplicable para empresas que tienen ganancias positivas durante todo el período en estudio. Además, se debe tener en cuenta que yt es un cantidad observable (ya que tanto Pt como Et lo son); sin embargo, PE y Zt no lo son. Se supone que tanto Zt como PE tienen valores discretos; por lo tanto, Zt {a1, . {b1, . . , bN} donde M y N representan el número de posibles estados latentes de Zt y PE , respectivamente.

Los supuestos de independencia condicional utilizados en este modelo se presentan a continuación:

$$t \geq 1, \quad yo, j \geq 1,$$

$$Zt \quad PE \quad | \quad . \tag{2}$$

$$Zi \quad Zj \mid Zk \quad k \quad \{i + 1, \ldots, j - 1\}\ yi \quad Zj \tag{3}$$

$$| Zk , PE \quad k \quad \{i, \ldots, j - 1\}\ o\ k \quad \{j + 1, \ldots, i\}.\ yi \quad yj \mid \tag{4}$$

$$Zk , PE \quad k \quad \{i, \ldots, j\}. \tag{5}$$

Machine Translated by Google

Para t ≥ 2 y r,s {1, . . . , M}, la matriz W := [wrs]M×M se define como la matriz de probabilidad de transición, donde

$$wrs := P[Zt = as \mid Zt-1 = ar].$$

Tenga en cuenta que wrs [0, 1] y $\sum_{s=1}^{\text{METRO}}$ wrs = 1. Para t ≥ 1, m {1, . . . , M} y n {1, . . . , N}, la matriz Dt := [dmn(yt)]M×N se define como la matriz de probabilidad de emisión en el tiempo t, donde

$$dmn(yt) := P[yt \mid Zt = am, PE = bn].$$

donde dmn(yt) N (ln[bn(1 + am)], σ vectores ^{2}). Para metro {1, . . . , M} y n {1, . . . , N}, el u := (um)m N y v := (vn)n N se definen como los vectores de probabilidad iniciales que contienen los distribuciones de probabilidad iniciales, donde

$$um := P[Z1 = soy]; \qquad vn := P[PE = bn].$$

Tenga en cuenta que um, vn [0, 1], $\sum_{m=1}^{\text{METRO}}$ um = 1 y $\sum_{n=1}^{\text{norte}}$ vn = 1. Estas distribuciones previas sirven para incorporar cualquier conocimiento experto que el investigador pueda tener disponible. Una representación gráfica del modelo descrito anteriormente se puede encontrar en la Figura 1.

Habiendo establecido los principios necesarios para la inferencia, el objetivo del análisis ahora queda más claro: inferir el valor de PE para estimar el precio fundamental de una acción y formular una estrategia comercial basada en este conocimiento. Además, inferir el valor de Zt también es útil, ya que puede utilizarse para probar una estrategia comercial alternativa. En este contexto, los datos utilizados se dividirán en un conjunto de entrenamiento y uno de prueba. El conjunto de parámetros del modelo viene dado por θ = {W, u, v, σ 2}. El espacio de posibles parámetros viene dado por el conjunto:

$$\Theta = \theta : um, vn, wrs \quad [0, 1], \quad \sum_{\text{metro}=1}^{\text{METRO}} um = 1, \quad \sum_{\text{norte}=1}^{\text{norte}} vn = 1, \quad \sum_{s=1}^{\text{METRO}} 2\, wrs = 1, \sigma > 0 . \tag{6}$$

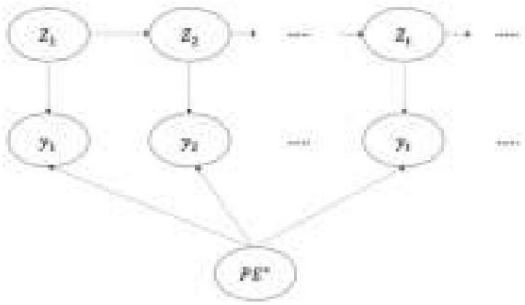


Figura 1. Representación del modelo descrito anteriormente utilizando un DBN, donde yt es observable y PE y Zt son variables latentes y discretas.

Luego siguen el aprendizaje y la inferencia. Dado que los parámetros en θ son desconocidos, primero será necesario estimarlos (procedimiento de aprendizaje). Como se mencionó, el algoritmo utilizado para obtener las estimaciones del MAP es el algoritmo de Baum-Welch. Después de estimar los parámetros desconocidos del modelo , se utiliza el algoritmo hacia adelante y hacia atrás en las estimaciones de los parámetros para inferir el valor constante de PE para el conjunto de entrenamiento y el conjunto de prueba, y para inferir el valor de Zt mediante el suavizado para el conjunto de entrenamiento, y mediante filtrado para el conjunto de prueba:

- Probabilidades de filtrado: P[ZT, PE | y1:T, θ];
- Probabilidades de suavizado: P[Zt , PE | y1:T, θ], donde t {1, . . . , T - 1}.

2.4. Inferencia con parámetros conocidos

Según el paso hacia adelante del algoritmo hacia adelante-hacia atrás, se realizan las siguientes definiciones y se derivan ecuaciones para las probabilidades de filtrado:

$$t \quad \{1, \ldots, T\}, \text{ metro } \{1, \ldots, M\}, \text{ norte } \{1, \ldots, N\},$$

Machine Translated by Google

$$\alpha tmn := P[Zt = am, PE = bn \mid y1:t, \theta];$$

$$\alpha 1mn = \sum_{m'=1}^{METRO} \frac{dmn(y1)umvn}{\sum_{norte'=1}^{norte} dm'n'(y1)um'vn'};$$

$$\alpha tmn = \frac{dmn(yt)\sum_{yo=1}^{METRO} P[Zt-1 = ai, PE = bn \mid y1:t-1]wmi}{\text{Connecticut}};$$

$$ct := \sum_{metro=1}^{METRO}\sum_{norte=1}^{norte} dmn(yt)\sum_{yo=1}^{METRO} P[Zt-1 = ai, PE = bn \mid y1:t-1]wmi.$$

Después de definir las probabilidades de filtrado, las probabilidades de suavizado se dan mediante la estimación denotada por γtmn:

$$t \in \{1, \ldots, T-1\}, metro \in \{1, \ldots, M\}, norte \in \{1, \ldots, N\},$$

$$\gamma tmn := P[Zt = am, PE = bn \mid y1:T].$$

A continuación, según el paso hacia atrás del algoritmo hacia adelante-hacia atrás, se debe dar la definición de βtmn para poder obtener los valores de γtmn:

$$\beta tmn := \frac{P[yt+1:T \mid Zt = am, PE = bn]}{P[yt+1:T \mid y1:t]} = \frac{P[yt+1:T \mid Zt = soy, PE = bn]}{\prod_{t'=t+1}^{t} ct'}$$

Por lo tanto, según el algoritmo hacia adelante-hacia atrás, se tiene:

$$\gamma tmn = \alpha tmn\beta tmn.$$

Lo que queda por derivar son las expresiones para βtmn $t \in \{1, \ldots, T-1\}$:

$$\beta(T-1)mn = \frac{b'_{(T-1)mn}}{\text{Connecticut}} = \frac{\sum_{metro=1}^{METRO} idin(yT)wim}{\sum_{metro=1}^{METRO}\sum_{n=1}^{norte} cT}.$$

Para βtmn $t \in \{1, \ldots, T-2\}$, $metro \in \{1, \ldots, M\}$, $norte \in \{1, \ldots, norte\}$:

$$b'_{tmn} = \sum_{yo=1}^{METRO} \beta(t+1)indin(yt+1)wim; \qquad \beta tmn = \frac{\sum_{i=1}^{M} \beta(t+1)indin(yt+1)wim}{ct+1}.$$

Con expresiones encontradas tanto para las probabilidades de suavizado como para las de filtrado, la más Los valores probables de las variables latentes PE y Zt se encuentran mediante marginación:

$$PE = \arg\max_{bn} \sum_{metro=1}^{METRO} P[Zt = am, PE = bn \mid y1:T] = \arg\max_{mn} \sum_{metro=1}^{METRO} \gamma tmn$$

Para encontrar la estimación Zt para el estado latente Zt , se utiliza suavizado en el conjunto de entrenamiento mientras se aplica el filtrado en el conjunto de prueba:

$$Zt = \arg\max_{soy} \sum_{norte=1}^{norte} \gamma tmn \quad t \in \{t : yt \in Xtrain\};$$

$$Zt = \arg\max_{soy} \sum_{norte=1}^{norte} \alpha tmn \quad t \in \{t : yt \in Xtest\}.$$

donde Xtrain y Xtest denotan el conjunto de entrenamiento y el conjunto de prueba, respectivamente.

Machine Translated by Google

2.5. Aprendizaje de parámetros desconocidos

El problema de optimización en el aprendizaje viene dado por $\theta_{MAP} = \arg\max_{\theta \in \Theta} P(\theta \mid y_1^t)$ y es resuelto usando el algoritmo de Baum-Welch (tenga en cuenta que en las siguientes expresiones, θ técnicamente debe escribirse con un superíndice de sombrero (ˆ) ya que es un conjunto de estimaciones de parámetros): 1. Establezca j = 1.

2. Establecer $\theta(1) - \theta(0) > \delta$

3. Mientras que $\theta(j) - \theta(j-1) > \delta$ hacer

(i) Calcular probabilidades $P(Z_t = a_m, E = b_n \mid y_{1:T}, \theta(j))$ t, m, n

(ii) Resuelva el problema de maximización restringida.

$$\arg\max_{\theta \in \Theta} \{Q(\theta; \theta(j)) + \ln P[\theta]\} \quad \theta(j+1) =$$

donde $Q(\theta; \theta(j)) = E_{Z_{1:T}, E \mid y_{1:T}, \theta(j)} [\ln P[y_{1:T}, Z_{1:T}, E \mid \theta]]$.

(iii) Incremento j

Se puede obtener una expresión en forma cerrada para $Q(\theta; \theta(j))$ utilizando las probabilidades de suavizado:

$$\begin{aligned} Q(\theta; \theta(j)) = {} & \sum_{t=1}^{T-1} \sum_{m=1}^{M} \sum_{n=1}^{N} \alpha_{tmn}^{(j-1)} \beta_{tmn}^{(j-1)} \ln[d_{mn}(y_t)] + \sum_{metro=1}^{M} \sum_{norte=1}^{N} \alpha_{tmn}^{(j-1)} \ln[d_{mn}(y_T)] \\ & + \sum_{t'=2}^{T-1} \sum_{yo,m=1}^{M} \sum_{n=1}^{N} \beta_{tmn}^{(j-1)} \frac{d_{mn}^{(j-1)}(y_t)}{c_t^{(j-1)}} w_{mi}^{(j-1)} \alpha_{t-1,i}^{(j-1)} \ln[w_{mi}] \\ & + \sum_{yo,m=1}^{M} \sum_{n=1}^{N} \frac{d_{mn}^{(j-1)}(y_T)}{do_t^{(j-1)}} w_{mi}^{(j-1)} \alpha_{T-1,pulg.}^{(j-1)} \ln[w_{mi}] \\ & + \sum_{metro=1}^{M} \sum_{norte=1}^{N} \alpha_{1\,minuto}^{(j-1)} \beta_{1\,minuto}^{(j-1)} (\ln[u_m] + \ln[v_n]). \end{aligned}$$

En el argumento del problema de maximización anterior, ln P[θ] es el logaritmo de] distribución previa P[θ], donde se supone $P[\theta] = P[u]P[v]P[W]P[\sigma^2]$ ors. Dentro desde la independencia de pride esta distribución previa, se pueden incluir dos tipos de conocimiento experto: conocimiento previo del valor aproximado de PE y conocimiento previo de la persistencia de los efectos del ruido a mediano plazo, que se codifican a través de los valores previos P[v] y el P[W] anterior , respectivamente.

El prior para el vector v está representado por la distribución de Dirichlet:

$$f(v) = \frac{\Gamma(k1 + k2 + \ldots + kN)}{\Gamma(k1)\Gamma(k2)\ldots\Gamma(kN)} \prod_{n=1}^{N} v_n^{kn-1} \quad (7)$$

Los valores kn para n {1, . . . , N} corresponden intuitivamente al grado de creencia que tiene un experto sobre el evento de que bn sea el valor "verdadero" de PE . Para este análisis, kn = 1 n. El a priori para la matriz W también se deriva de la distribución de Dirichlet $f(W) = \prod_{metro=1}^{M} f(w_m)$, donde $w_m = (w_{im})_{i=1,\ldots,M}$ y

$$= \frac{\Gamma(k1m + k2m + \ldots + kMm)}{\Gamma(k1m)\Gamma(k2m)\ldots\Gamma(kMn)} \prod_{metro=1}^{M} w_{soy}^{kim-1} .$$

Dado que W es la matriz de transición para la cadena de Markov oculta {Zt}t N, entonces las entradas diagonales de W representan la probabilidad de que un estado particular persista (permanezca como está en el siguiente punto temporal). Cuanto mayor sea el valor de Zt (o, correspondientemente, am), mayor será

Machine Translated by Google

la persistencia del Estado. En este análisis, las entradas fuera de la diagonal (se sugiere al lector consultar Wang [22] para obtener más información sobre los valores de las fuera de la diagonal) se establecen en 0.

Teniendo todo lo anterior en consideración, el logaritmo del P[θ] anterior se convierte en

$$\ln P[\theta] = \ln P[W] + \ln P[v] = \sum_{metro=1}^{METRO} \sum_{yo=1}^{METRO} (kim - 1)\ln[wim] + \sum_{norte=1}^{norte} (kn - 1)\ln[vn] + s1,$$

donde $s1 = \ln \frac{\Gamma(k1+k2+\ldots+kN)}{\Gamma(k1)\Gamma(k2)\ldots\Gamma(kN)} + \ln \frac{\Gamma(k1m+k2m+\ldots+kMm)}{\Gamma(k1m)\Gamma(k2m)\ldots\Gamma(kMn)}$.

La constante s1 sólo se incluye por razones de integridad; se vuelve irrelevante cuando maximizando en el algoritmo de Baum-Welch.

Con los antecedentes establecidos, el problema de optimización restringida ahora está completamente definido. Tenga en cuenta que sólo se considerarán las restricciones de igualdad en (6) , ya que las restricciones de desigualdad terminarán siendo satisfechas aún. Por tanto, el método de los multiplicadores de Lagrange se puede utilizar para resolver este problema de optimización. Las expresiones para los estimadores de las cuatro variables en cuestión se dan a continuación:

$$tu(j)_{mi} = \sum_{norte=1}^{norte} {}_{1mn\,;\,1\,minuto}^{(j-1)\,(j-1)\,\beta\,\alpha} \tag{8}$$

$$v(j)_{mi} = \frac{\sum_{m=1}^{M} {}_{1mn}^{(j)}(Z1, PE) + (kn - 1)\sum_{m=1}^{q}}{1 + \sum_{norte'=1}^{norte} (kn' - 1)} = \frac{\sum_{m=1}^{M} \alpha\, {}_{1mn}^{(j-1)\,(j-1)}}{1 + \sum_{norte'=1}^{norte} (kn' - 1)} + \frac{(kn - 1)\,\beta}{1 + \sum_{norte'=1}^{norte} (kn' - 1)}; \tag{9}$$

$$w(j)_{mi} = \frac{\sum_{t'=2}^{t} \sum_{n=1}^{norte} q\, {}_{t',m,i,n}^{(j)}(Zt', Zt'-1, PE) + (kmi - 1)}{\sum_{t'=2}^{t} \sum_{metro=1}^{METRO} \sum_{n=1}^{norte} q\, {}_{t',m,yo,n}^{(j)}(Zt', Zt'-1, PE) + \sum_{metro'=1}^{METRO} (km'yo - 1)} = \frac{\omega\, {}_{t',m,i,n}^{(j)}}{\sum_{m'=1}^{M} \omega\, {}_{t',m,i,n}^{(j)}}, \tag{10}$$

dónde

$$\omega t\, {}_{',m,i,n}^{(j)} = \sum_{t'=2}^{T-1} \sum_{n=1}^{norte} \beta\, {}_{'mn}^{(j-1)\,t} \frac{d\, {}_{mn}^{(j-1)}(yt')}{c\, {}_{t'}^{(j-1)}}\, w\, {}_{mi}^{(j-1)}\, {}_{t',(j-1)\,\alpha}{}_{-1,i,n} + \sum_{norte=1}^{norte} \frac{d\, {}_{mn}^{(j-1)}(yT)}{do_t^{(j-1)}}\, w\, {}_{mi}^{(j-1)}\, \alpha\, {}_{(T-1)en}^{(j-1)} + (kmi - 1)$$

$$\sigma 2\,(j-1) = \frac{\sum_{t=1}^{T-1} \sum_{metro=1}^{METRO} \sum_{n=1}^{norte} \alpha\, {}_{tmn}^{(j-1)\,(j-1)\,\beta}\,(yt - \ln[bn(1 + am)])2}{\sum_{t=1}^{T-1} \sum_{metro=1}^{METRO} \sum_{n=1}^{N} \alpha\, {}_{1tmn}^{(j-1)\,(j-1)\,\beta_{tmn}} \sum_{m=1}^{M} \alpha\, {}_{tmn}^{(j-1)}} + \frac{\sum_{metro=1}^{METRO} \sum_{n=1}^{norte} \alpha n = 1\, {}_{Tmn}^{(j-1)}\,(yT - \ln[bn(1 + am)])2}{\sum_{t=1}^{T-1} \sum_{metro=1}^{METRO} \sum_{n=1}^{norte} \alpha\, {}_{tmn}^{(j-1)\,(j-1)\,\beta_{tmn}} + \sum_{metro=1}^{METRO} \sum_{n=1}^{norte} \alpha n = 1\, {}_{tmn}^{(j-1)}}. \tag{11}$$

3. Metodología de Análisis y Resultados

Se eligieron doce acciones para ser analizadas en este documento (ver Tabla 1): nueve de un mercado desarrollado (Estados Unidos) y tres de mercados emergentes (Brasil y China). El conjunto de entrenamiento contiene datos del 1 de enero de 2011 al 31 de diciembre de 2019, mientras que el conjunto de prueba contiene datos del 1 de enero de 2020 al 30 de septiembre de 2020. Utilizando un índice de Sharpe móvil de 12 meses como punto de referencia, todas las acciones mostraron rentabilidades medias o superiores a la media (con respecto a un índice de referencia de S&P) durante el período de prueba, con la excepción de las dos acciones brasileñas que tuvieron un rendimiento inferior [26]. Se recopilaron dos conjuntos de datos para cada acción durante el período mencionado anteriormente: datos de precios diarios y datos de ganancias trimestrales por acción.

Machine Translated by Google

Tabla 1. Tabla de resultados, que muestra el retorno de la inversión como porcentaje del monto invertido.

Los valores entre paréntesis son negativos.

Existencias	Umbrales para la estrategia LT 0,10				Umbrales para la estrategia MT 0,05				Beneficio de B&H
	0,05	0,20	0,15		0,03	0,10	0,07		
AAPL (Estados Unidos)	55,30%	55,30%	55,30%	55,30%	69,51%	57,82%	48,06%	76,02%	55,30%
ADBE (Estados Unidos)	25,46%	37,86%	44,45% 0%		46,65%	46,65%	46,65%	46,65%	46,65%
NEGRO (EE. UU.)	12,98%	12,98%	12,98%	12,98%	33,92%	34,66%	20,85%	20,85%	12,98%
COSTO (EE. UU.)	22,59%	22,59%	22,59%	22,59%	50,12%	42,85%	28,21% 0%		22,59%
HD (EE. UU.)	28,71%	28,71%	28,71%	28,71%	57,76%	51,31%	37,99%	47,35%	28,71%
ITUB (BZ)	(39,86%)	(29,36%)	(29,36%)	(27,57%)	9,10%	27,17%	29,34%	61,09%	7,07% 7,07% (9,59%) (55,85%)
MA (Estados Unidos)	60,41%	47,70%							32,57% 24,66% 34,39% 11,93%
MCD (Estados Unidos)	11,38%	11,38%	11,38%	11,38%	18,94%	18,94%	30,63%	30,63%	11,38%
NTES (CHN)	47,89%		50,84%	57,35%	39,54%	39,54%	39,54%	39,54%	39,54%
NVDA (EE.UU.)	125,90%	125,90%	125,90%	125,90%	149,00%	156,79%	125,90%	125,90%	125,90%
SAN (BZ)	(14,35%)	(17,78%)	(17,78%)	(57,17%)	(57,17%)	(57,17%)	(57,17%)	(57,17%)	
UNH (EE.UU.)	7,95%	7,95%	19,88%	15,90%	72,41%	21,83%	38,58%	2,87%	7,95%
Promedio	26,12%	25,07%	32,38%	27,61%	44,79%	37,74%	32,60%	29,79%	20,83%
Totales Lavadora/Lavadora/Lavadora	4/7/1	4/7/1 5/6/1 17/27/4		4/7/1	9/3/0	9/3/0 7/4/1 31/14/3		6/4/2	

Después de la fase preliminar de limpieza y preparación de datos, los valores iniciales de los vectores a = (am) y b = (bn) se establecen a continuación. Los antecedentes del vector v y la matriz W se establecen como se explicó más temprano; el valor anterior para u se establece como la distribución uniforme discreta y el valor inicial de σ^2 se establece en 5. Luego se realiza el aprendizaje de parámetros antes de la inferencia de los estados latentes. PE y Zt . A continuación se muestra la simulación de las estrategias comerciales. Las estrategias comerciales propuestas. por Wang [22] y utilizado en este artículo se comparará con la estrategia de referencia de B&H.

Para ambas estrategias comerciales propuestas, indique la cantidad de efectivo disponible en ese momento. t; Sea Nt las unidades de un valor mantenido en el momento t; dejemos que Ttrain represente el tamaño del conjunto de entrenamiento; y deje que T represente el tamaño del conjunto de datos (suma de los tamaños de entrenamiento y prueba conjuntos). Tanto las estrategias comerciales a largo plazo como las de mediano plazo pueden describirse mediante tres posibles cursos de acción (etiquetados (i), (ii) y (iii) a continuación) en el momento t; es decir, cursos de acción (i) a (ii) son comunes a ambas estrategias comerciales:

(i) Si PEt ≤ At(1 − Tr) e It > 0, entonces compre el valor utilizando todo el efectivo disponible It . Por tanto, $N_{t+1} = \frac{El}{PT}$ y It+1 = 0.

(ii) Si PEt ≥ At(1 + Tr) e It = 0, entonces vender todas las unidades poseídas Nt . Por tanto, Nt+1 = 0 y It+1 = PtNt .

(iii) Si no se cumplen ni (i) ni (ii), no ejecute ninguna operación. Por tanto, Nt+1 = Nt y Eso+1 = Eso .

Tenga en cuenta que At se considera una línea de base y depende de la estrategia comercial. El El valor umbral Tr (0, 1) actúa como un indicador de sensibilidad que define en qué medida el inversor quiere permitir PEt = $\frac{punto}{y}$ desviarse de la línea de base At antes de desencadenar un determinado curso de acción en las estrategias comerciales. Esto se desprende claramente de cómo son los cursos de acción. definido. El beneficio total al final del período de negociación viene dado por IT + PTNT − ITtrain+1.

La primera estrategia comercial es la llamada "estrategia a largo plazo", donde la negociación se realiza con respecto al valor constante de PE , por lo que At = PE . La estrategia alternativa es la "estrategia a mediano plazo", donde el comercio se realiza con respecto a la dinámica valores de PE (1 + Zt), donde cada Zt se estima dinámicamente mediante filtrado. Por lo tanto, para la estrategia de mediano plazo, At = PE (1 + Zt).

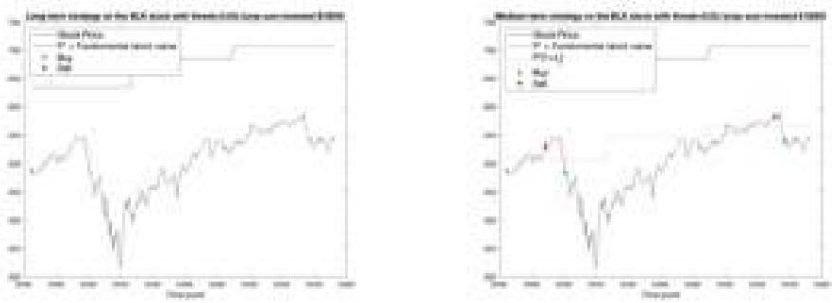
Los resultados de las acciones BLK e ITUB se presentan en detalle gráfico en este documento. En primer lugar, la estrategia a largo plazo basada en los datos de las acciones de BLK sugiere que el inversor compra el stock en el punto temporal 2265 = Ttrain + 1 y mantiene el stock durante el resto del período. Claramente, esta estrategia coincide con la estrategia de B&H y, como resultado, el beneficio de la estrategia a largo plazo sería igual al beneficio de la estrategia de B&H, que equivale a 1298,04 USD.

lo que equivale a un retorno del 12,98% sobre la inversión inicial de USD 10.000. Por otra parte, el

Machine Translated by Google

La estrategia a mediano plazo sugiere que el inversor compre la acción en los momentos 2265, 2301 y 2437, la venda en los momentos 2289 y 2431 y la mantenga durante el resto de los momentos. Esto arrojaría una ganancia de USD 3391,85; lo que equivale a un retorno del 33,92% sobre la inversión inicial de USD 10.000 en el período de nueve meses que cubre el conjunto de pruebas. Esto significa que la estrategia a medio plazo supera a la estrategia de B&H en un 20,94%. La Figura 2 ilustra la estrategia a largo plazo y la estrategia a mediano plazo.

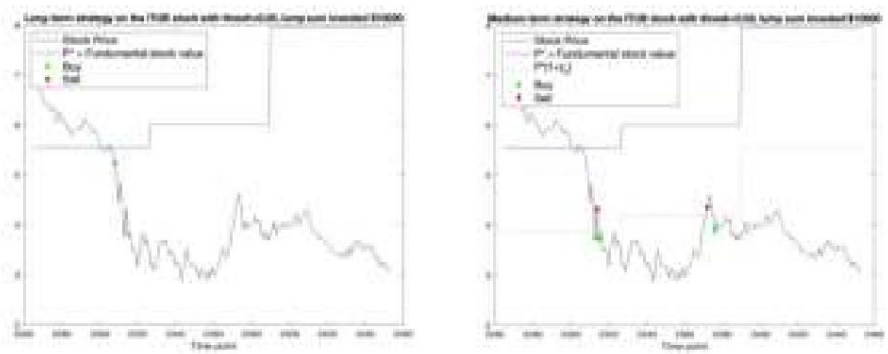
A continuación, analizamos los datos bursátiles de ITUB. La estrategia a largo plazo sugiere que el inversor compre la acción en el momento 2308 y la mantenga durante el resto del período. La implementación de esta estrategia generaría una pérdida de USD 3986,16, lo que equivale a una pérdida del 39,86% sobre la inversión inicial de USD 10.000. La estrategia de comprar y mantener, sin embargo, produciría una pérdida mayor de 5.584,67 dólares. Por el contrario, la estrategia a medio plazo sugiere que el inversor compra la acción en los momentos 2313, 2315 y 2376, la vende en los momentos 2314 y 2372 y la mantiene durante el resto de los momentos. En realidad, esto produciría un beneficio de 909,81 dólares. Aunque esto supone sólo un rendimiento del 9,1% sobre la inversión inicial de 10.000 dólares, la estrategia a medio plazo proporciona al inversor una estrategia mediante la cual puede obtener beneficios en un período en el que las acciones realmente están colapsando. La Figura 3 ilustra la estrategia a largo plazo y la estrategia a mediano plazo, respectivamente.



(a) Estrategia a largo plazo b) Estrategia a mediano plazo

Figura 2. Estrategias simuladas para la acción BLK.

De manera más general, vemos en la Tabla 1 que la estrategia media ha sido consistentemente superior (para varios umbrales) para BLK, COST, HD, ITUB, MA, MCD, NVDA y UNH. Para ITUB, para todos los umbrales excepto uno, la estrategia genera una ligera ganancia aunque se registran pérdidas para otras estrategias. La estrategia a largo plazo ha sido consistentemente superior para NTES y SAN. Ninguna estrategia ha ofrecido ventajas a ADBE, mientras que para AAPL, el éxito de la estrategia a medio plazo depende de la elección del umbral. En la conclusión se ofrecerá un análisis más detallado de estos resultados.



(a) Estrategia a largo plazo b) Estrategia a mediano plazo

Figura 3. Estrategias simuladas para el stock IUTB.

Machine Translated by Google

4. Conclusiones

Mediante el uso de DBN, se construye el modelo de movimiento de acciones de Wang [22] para nuestro conjunto de datos de acciones existente. Este modelo incluye dos estados latentes: uno que modela los efectos del ruido a mediano plazo y el otro que modela el verdadero ratio PER fundamental de una empresa, este último se supone constante durante todo el período estudiado. El algoritmo hacia adelante-hacia atrás y el algoritmo de Baum-Welch (variante del algoritmo EM) se utilizan para realizar el aprendizaje y la inferencia de parámetros. Con base en este modelo ajustado, las estrategias de largo y mediano plazo se aplican a las doce acciones aquí estudiadas, nueve de estas acciones cotizan en un mercado desarrollado y el resto en uno emergente. En general, tanto la estrategia a largo plazo como la estrategia a mediano plazo superan a la estrategia comercial de referencia de B&H 17 y 31 veces, respectivamente, de un total de 48 ejecuciones experimentales para cada estrategia (cuatro para cada una de las doce acciones). Las estrategias propuestas por Wang sólo pierden ante B&H cuatro y tres veces, respectivamente. Además, el rendimiento superior de estas estrategias comerciales es sustancial. Mientras que el beneficio medio de todas las acciones cuando se utiliza el B&H es del 20,83%, el beneficio medio de todas las acciones y de todos los umbrales para la estrategia a largo plazo es del 27,78% y el de la estrategia a medio plazo es del 36,23%. Por último, resulta que estas estrategias comerciales ofrecen al inversor sugerencias comerciales que, en determinados casos, pueden incluso convertir una pérdida en la estrategia B&H en una ganancia, como se demostró en el caso de las acciones de ITUB. Esta acción estaba en caída, pero la estrategia a medio plazo todavía produjo beneficios sobre la suma invertida. Todos estos resultados se muestran en la Tab

Como ocurre con cualquier modelo estadístico, el modelo implementado en este trabajo tiene margen de mejora. La primera y más importante limitación es el hecho de que el modelo sólo puede modelar las acciones que tuvieron un EPS positivo durante el período estudiado. Esto se debe al lado izquierdo de la ecuación del modelo, es decir, la ecuación (1). En determinados momentos, como durante pandemias o recesiones, puede resultar considerablemente difícil encontrar una empresa que no haya registrado pérdidas en al menos un trimestre del período estudiado. Una segunda limitación es la falta de uso de conocimientos expertos. En las inversiones del mundo real, la experiencia en el campo se considera muy valiosa y, por lo general, solo se recomienda a los inversores experimentados que operen activamente. Dado que el modelo de este artículo permite la incorporación de conocimiento experto, es de hecho una limitación que dicho conocimiento no se utilice. Otra limitación relacionada con el conocimiento experto es la limitación de que en este análisis no se tienen en cuenta tarifas o comisiones comerciales . Aunque el modelo propuesto permite que se consideren las comisiones en las estrategias comerciales, al momento de escribir este artículo no se disponía de conocimiento sobre los valores reales de las tarifas o comisiones cobradas y, por lo tanto, dichos gastos no se pudieron tener en cue Es bien sabido que ciertas tarifas comerciales a veces pueden acumularse al ejecutar numerosas operaciones de tal manera que una ganancia puede convertirse en una pérdida cuando estas tarifas pasan a primer plano, lo que haría que la estrategia de comprar y mantener sea más rentable.

Las mejoras en las limitaciones mencionadas anteriormente pueden agregar valor al modelo y deben considerarse en trabajos futuros. Además, otra mejora que hace que el modelo aquí sea más aplicable en escenarios de la vida real es la incorporación de alguna variable que mida la liquidez de una acción. Esto se debe a que las sugerencias proporcionadas por las estrategias comerciales sobre cuándo comprar y vender una acción se vuelven inútiles si la acción en sí no es líquida. Los precios pueden cambiar cuando las acciones se vuelven lo suficientemente líquidas como para que surja una oportunidad comercial real, y si el precio cambia significativamente, eso obviamente hace que la sugerencia en sí sea inútil. En este estudio, no se previó que la liquidez fuera un problema, ya que las 12 acciones consideradas son acciones de gran capitalización.

La elección de los datos históricos utilizados es siempre una decisión importante debido a que la duración de la serie temporal juega un papel importante. En principio, se prefieren las series temporales más largas a las más cortas, pero si los datos históricos contienen cambios de régimen, esto puede inhibir el desempeño del modelo en sus pronósticos. Además, los futuros cambios de régimen que puedan ocurrir durante el período de prueba también pueden afectar el desempeño de pronóstico del modelo.

Finalmente, otras mejoras que se pueden implementar en futuros estudios son una experimentación y pruebas más profundas y la posibilidad de vender acciones en corto. Para empezar , se puede implementar la experimentación a nivel de cartera con las acciones aquí estudiadas.

Machine Translated by Google

para comprender cómo cambian las ganancias cuando se considera un grupo de acciones en conjunto; por ejemplo, se pueden tomar muestras de nueve acciones de las doce estudiadas aquí para formar carteras, y se puede identificar la cartera "óptima". Aparte de esto, se puede dedicar más tiempo a la validación cruzada en estudios futuros para mejorar aún más las sugerencias proporcionadas por las estrategias comerciales en este trabajo. Además, las propiedades de todos los estimadores utilizados podrían derivarse en trabajos futuros, para comprender mejor su comportamiento. Finalmente, esta estrategia no permite la venta al descubierto de acciones; Se podrían proponer extensiones al modelo que permitan esto .

Contribuciones de los autores: Los tres autores trabajaron igualmente duro para escribir este artículo. KL se encargó de la codificación en MATLAB y de la sección de resultados, MAC y DPS se encargaron de la metodología y el contenido teórico del artículo. Todos los autores han leído y aceptado la versión publicada del manuscrito.

Financiamiento: Esta investigación no recibió financiamiento externo.

Declaración de disponibilidad de datos: todos los datos están disponibles abiertamente en Yahoo Finance en https://finanzas.yahoo.com/.

Conflictos de intereses: Los autores declaran no tener conflictos de intereses.

Abreviaturas

En este manuscrito se utilizan las siguientes abreviaturas:

BN	Red bayesiana
DAG	Gráfico acíclico dirigido por
DBN	Red bayesiana dinámica
DGM	Modelo gráfico dirigido por
EM	Expectativa de EM: Maximización
HNN	Modelo de Markov oculto de
PDF	Función de densidad de probabilidad
MAP	MAP Máximo A Posteriori
ML	Probabilidad máxima de

Referencias

1. Russell, S.; Norvig, P. Inteligencia artificial: un enfoque moderno, 2ª ed.; Prentice Hall: Upper Saddle River, Nueva Jersey, EE. UU., 1995.
2. Murphy, K. Redes bayesianas dinámicas: representación, inferencia y aprendizaje. Doctor en Filosofía. Disertación, Universidad de California, Berkeley, California, Estados Unidos, 2002.
3. Basu, S. Rendimiento de las inversiones de las acciones ordinarias en relación con sus relaciones precio-beneficio: una prueba del mercado eficiente hipótesis. J. Finanzas. 1977, 23, 3. 663–673. 4.
Lleó, S.; Ziemba, WT Historia de dos índices: Predicción de caídas del mercado de valores en China. 2020. [Referencia cruzada]
5. Angelini, N.; Bormetti, G.; Marmi, S.; Nardini, F. El valor importa: previsibilidad de la rentabilidad de los índices bursátiles. SSRN 2012. [Referencia cruzada]
6. Chang, Kansas; Tian, Z. Análisis de mercado y estrategias comerciales con redes bayesianas. En Actas de la XVIII Internacional Conferencia sobre fusión de información, FUSION 2015, Washington, DC, EE. UU., 6 a 9 de julio de 2015; págs. 1922-1929.
7. Damián, L.; Peterson, B.; Weylandt, M. Cambio de régimen y comercio técnico con redes bayesianas dinámicas en alta Frecuencia Bolsas de Valores, 2017.
8. Tayal, A. Cambio de régimen y negociación técnica con redes bayesianas dinámicas en mercados bursátiles de alta frecuencia; Universidad de Waterloo: Waterloo, ON, Canadá, 2009.
9. Dagum, P.; Galper, A.; Horvitz, E. Modelos de redes dinámicas para pronósticos. En Actas de la Octava Conferencia sobre Incertidumbre en Inteligencia Artificial, Stanford, CA, EE.UU., 17 a 19 de julio de 1992; págs. 41–48.
10. Dagum, P.; Galper, A.; Horvitz E. Razonamiento y pronóstico inciertos. Int. J. Pronóstico. 1995, 11, 73–87. [Referencia cruzada]
11. Zweig, Z.; Russell S. Reconocimiento de voz con redes bayesianas dinámicas. 1998. Disponible en línea: http://www.cse.msu.edu/ ~cse841/papers/zweig.pdf (consultado el 28 de abril de 2024).
12. Nefián, A.; Liang, L.; Pi, X.; Liu, X.; Murphy, K. Redes bayesianas dinámicas para el reconocimiento de voz audiovisual. EURASIP J. Adv. Proceso de señal. 2002, 11, 783042. [Referencia cruzada]
13. Yao, XQ; Zhu, H.; Ella, Z. Un enfoque dinámico de red bayesiana para la predicción de la estructura secundaria de proteínas. Bioinformación de BMC. 2008, 9, 49. [Referencia cruzada] [PubMed]
14. Raval, A.; Ghahramani, Z.; Wild, DL Un modelo de red bayesiana para el pliegue de proteínas y el reconocimiento remoto de homólogos. Bioinformáticas 2002, 18, 788–801. [Referencia cruzada] [PubMed]

Machine Translated by Google

15. Murphy, K.; Mian, S. Modelado de datos de expresión genética mediante redes bayesianas dinámicas. 1996. Disponible en línea: https://www.cs.ubc.ca/~murphyk/Papers/ismb99.pdf (consultado el 28 de abril de 2024).
16. Prembida, C.; Souza, FAA; Faria, DR Red bayesiana dinámica para clasificación semántica de lugares en robótica móvil. Auton. Robot. 2016, 5, 1161–1172. [Referencia cruzada]
17. Patel, M.; Miró, JV; Dissanayake, G. Redes bayesianas dinámicas para el aprendizaje de interacciones entre caminantes robóticos de asistencia y usuarios humanos. 2010. Disponible en línea: https://opus.lib.uts.edu.au/bitstream/10453/16294/1/2010000382.pdf (consultado el 28 de abril de 2024).
18. Delage, E.; Lee, HAY Un modelo de red bayesiana dinámica para la reconstrucción 3D autónoma a partir de una única imagen interior. En Actas de la Conferencia de la IEEE Computer Society de 2006 sobre visión por computadora y reconocimiento de patrones (CVPR'06), Nueva York, NY, EE. UU., 17 a 22 de junio de 2006; Volumen 2, págs. 2418–2428. [Referencia cruzada]
19. Cózar, J.; Puerta J. ;Gamez, JA Una aplicación de redes bayesianas dinámicas al monitoreo de condiciones y predicción de fallas en Un sistema sensorizado: un estudio de caso. Int. J. Computación. Intel. Sistema. 2017, 10 , 176. [Referencia cruzada]
20. Benhamou, E.; Atif, J.; Laraki, R. Un nuevo enfoque para el aprendizaje en redes bayesianas dinámicas (DBN). Electrón SSRN. J. 2018. [Referencia cruzada]
21. Murphy, K. Introducción a los modelos gráficos. 2001. Disponible en línea: https://www.cs.ubc.ca/~murphyk/Bayes/bnintro.HTML (consultado el 28 de abril de 2024).
22. Wang, H. Inferencia probabilística bayesiana sobre la dinámica de los precios de las acciones a nivel empresarial. Doctor en Filosofía. Disertación, Universidad Suranaree de Tecnología, Nakhon Ratchasima, Tailandia, 2016.
23. Murphy, K. Aprendizaje automático: una perspectiva probabilística; Prensa del MIT: Cambridge, MA, EE. UU., 2012.
24. Baum, LE; Petries, T.; Soulés, G.; Weiss, N. Una técnica de maximización que ocurre en el análisis estadístico de métodos probabilísticos. Funciones de las Cadenas de Markov. Ana. Matemáticas. Estadística. 1970, 41, 1, 167–171. [Referencia cruzada]
25. Dempster, AP; Laird, Nuevo México; Rubin, DB Máxima probabilidad de datos incompletos mediante el algoritmo EM. Estadística JR. Soc. Ser. B 1977, 39, 1–22. [Referencia cruzada]
26. Laboratorio de carteras. Disponible en línea: https://portfolioslab.com/ (consultado el 28 de abril de 2024).

Descargo de responsabilidad/Nota del editor: Las declaraciones, opiniones y datos contenidos en todas las publicaciones son únicamente de los autores y contribuyentes individuales y no de MDPI ni de los editores. MDPI y/o los editores renuncian a toda responsabilidad por cualquier daño a personas o propiedad que resulte de cualquier idea, método, instrucción o producto mencionado en el contenido.

Machine Translated by Google

AppliedMath

MDPI

Artículo

De las soluciones álgebras geométricas de la ecuación de Toda a Fórmulas Sato

Pierre Gaillard

Faculté des Sciences Mirande, Université de Bourgogne Franche-Comté, 2100 Dijon, Francia;
pierre.gaillard@u-bourgogne.fr

Resumen: Sabemos que la degeneración de soluciones de PDE, dadas en términos de funciones theta en superficies de Riemann, proporciona resultados importantes sobre soluciones particulares, como en el caso de la ecuación NLS. Aquí, degeneramos las llamadas soluciones de espacios finitos de la ecuación reticular de Toda a partir de la formulación general en términos de funciones abelianas cuando los espacios tienden a puntos. Esta degeneración nos permite recuperar las fórmulas de Sato sin utilizar la teoría de la dispersión inversa ni métodos geométricos o teóricos de representación.

Palabras clave: Superficies de Riemann; funciones theta; integrales abelianas; Funciones de Baker Akhiezer; fórmulas sato

MSC: 35C99, 35Q35, 35Q53

1. Introducción

Se pueden elegir diferentes formulaciones para representar la ecuación de Toda. En el trabajo original de 1967 [1], se estudió la vibración de una cadena uniforme de partículas con interacción no lineal , y Toda consideró la siguiente ecuación para la n-ésima partícula de la

$$\text{cadena: } m\partial^2 x(un) = -\partial r(\varphi)(un - un-1) + \partial r(\varphi)(un+1 - un), \quad (1)$$

donde m representa la masa de las partículas y φ(r) es la energía de interacción entre partículas adyacentes. Se construyeron algunas soluciones para interacciones particulares φ.

La misma ecuación fue considerada con m = 1 por Date y Tanaka en 1976 [2], y construyeron algunas soluciones generales pero en términos de integrales que son difíciles de usar.

En el apéndice del artículo de Dubrovin [3], publicado en 1981, Krichever consideró la versión no abeliana de esta ecuación y proporcionó soluciones en términos de la función theta en superficies de Riemann.

En 1995, Matveev y Stahlhoffen [4] consideraron la siguiente versión de la ecuación de Toda:

$$\partial^2 x(un) = \exp(un-1 - un) - \exp(un - un+1). \quad (2)$$

Usaron la transformación de Darboux para construir soluciones a esta ecuación en términos de Casoratis.

En 2014, Zhang y Zhou [5] utilizaron la siguiente representación de la ecuación de Toda

$$\partial x\,\partial t(un) = (\partial t(un) + \alpha(t))(un-1 - 2un + un+1). \quad (3)$$

Utilizaron la generalización de la función exp para construir soluciones multiondas a esta ecuación.

Más recientemente, Sun, Ma y Yu [5] examinaron la siguiente representación de la ecuación de Toda:

$$\Delta(un) = 4(\exp(un-1 - un) - \exp(un - un+1)). \quad (4)$$

check for updates

Cita: Gaillard, P. De Álgebro Soluciones geométricas del Toda Ecuación a fórmulas de Sato. Matemáticas aplicadas 2024, 4, 856–867. https://doi.org/10.3390/appledmath4030046

Editor académico: Carlo Bianca

Recibido: 26 de abril de 2024
Revisado: 2 de julio de 2024
Aceptado: 2 de julio de 2024
Publicado: 9 de julio de 2024

Copyright: © 2024 por el autor. Licenciatario MDPI, Basilea, Suiza. Este artículo es un artículo de acceso abierto. distribuido bajo los términos y condiciones de los Creative Commons Licencia de atribución (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

Machine Translated by Google

Usaron una transformación logarítmica para obtener algunas soluciones particulares en términos de logaritmos.

En 2020, Duarte [6] consideró la siguiente representación de la ecuación de Toda:

$$\alpha\partial^2_x (u(x, y, n)) + \beta\partial^2 (u(x, y, n)) =$$
$$\exp(u(x, y, n - 1) - u(x, y, n)) - \exp(u(x, y, n) - u(x, y, n + 1)).$$

Usó un ansatz particular combinado con las propiedades de la ecuación de Laplace para Construir algunas soluciones en términos de funciones trigonométricas.

También en 2020, Schiebold y Nilson [7] estudiaron otra versión de la ecuación de Toda en la forma

$$\partial^2_{xy} (\log(1 + u_n)) = u_{n+1} - 2u_n + u_{n+1}. \quad (5)$$

Construyeron soluciones en el marco del álgebra lineal mediante determinantes.

Aquí, consideramos la ecuación de Toda en la siguiente representación [8-11]:

$$\frac{r_n}{r_{n+1}}\sqrt{c_n}\psi_{n+1} + v_n\psi_n + \frac{r_n}{r_{n-1}}\sqrt{c_{n-1}}\psi_{n-1} = \lambda\psi_n$$
$$\partial_t\psi_n = \sqrt{c_n}\psi_{n+1} + \partial_t(\ln r_n)\psi_n - \sqrt{c_{n-1}}\psi_{n-1}, \quad \frac{r_n}{2r_{n+1}}\frac{r_n}{2r_{n-1}} \quad (6)$$

con

$$c_n = \frac{{}_{m}2I_0\,\theta^2(tV'' + nD + Z)\,\theta(tV'}{' + (n + 1)D + Z)\theta(tV'' + (n - 1)D + Z)\,\theta(tV'' + (n - 1)D +} \quad (7)$$

$$nD + Z) \qquad Z)\,v_n = -R_0 + \partial_t \ln \theta(tV'' + \quad (8)$$

donde θ es la función theta de Riemann clásica.

Sabemos que la degeneración de las soluciones de las PDE dadas en términos de funciones theta de Riemann proporciona algunas soluciones particulares importantes. En el caso de la ecuación NLS, hemos logrado construir una solución cuasi-racional que involucra el determinante de orden N para cada entero positivo N dependiendo de 2N − 2 parámetros reales [12]. En el caso de la ecuación KdV, construimos soluciones en términos de determinantes de Fredholm y Wronskianos [13], a partir de funciones de Baker-Akhiezer.

En este estudio, recuperamos fórmulas de Sato para la ecuación de Toda utilizando este método.

A partir de soluciones de espacios finitos dadas en términos de funciones theta de Riemann, construimos algunas soluciones cuasi-racionales y recuperamos las fórmulas de Sato, usando la degeneración, como se indica en el marco de la ecuación NLS [14].

Precisamente, derivamos soluciones multisolitones a partir de soluciones de espacios finitos mediante una transición límite , es decir, haciendo que los espacios tiendan hacia puntos en una determinada superficie de Riemann. Esto se logró en el espíritu de [14] o más recientemente [13]. Una ventaja de este enfoque es que no se basa en la teoría de la dispersión inversa ni en métodos geométricos y teóricos de representación, lo que ofrece una nueva perspectiva del problema.

Consideramos la superficie de Riemann Γ representada por $\overline{k=1}\ a_k b_k u_n{}^{-1}{}_k{}^{-1}{}_{segundo}{}_k$ del algebraico curva definida por [8,9]

$$2\,\omega = \prod_{j=1}^{2g+2} (z - E_j),$$

con $E_j = E_k$, $j = k$.

Consideremos Ω″ y Ω'_1 $_0$ [11] integrales abelianas, verificando

$$\int_{A_k} d\Omega''_1(P) = \int_{A_k} d\Omega'_0(P) = 0 \qquad k = 1, \ldots, g, \quad (9)$$

con el siguiente comportamiento asintótico para P = (ω, z) Γ [11],

Machine Translated by Google

$$\Omega''_1(P) = \pm \frac{1}{2}(z + R0 + O(z^{-1})), \qquad PAG \to PAG_{\infty}^{\pm} \tag{10}$$

$$\Omega'_0(P) = \pm(\ln z - I0 + O(z^{-1})), \qquad PAG \to PAG_{\infty}^{\pm}, \tag{11}$$

$$\omega = \pm(z^{g+1} + O(z^{g})), \qquad PAG \to PAG_{\infty}^{\pm}. \tag{12}$$

Los vectores V'', D, X y Z están definidos por

$$V_j'' = \frac{1}{2\pi i b_j} d\Omega''_1, \tag{13}$$

$$re = U(P^{+}_{\infty}) - U(P^{-}_{\infty}), \tag{14}$$

$$X = K + \sum_{j=1}^{\bar{g}} U(P_j), \qquad Z = U(P^{+}_{\infty}) - X, \tag{15}$$

donde K es el vector de constantes de Riemann.

U(P) es la integral abeliana clásica dU.

^{PAG}

P0 Las soluciones al sistema (6) se pueden escribir [10,11] como

$$\psi_n(P, t) = \tilde{r}_n(t) e^{n\text{orte}\Omega_0'(P) + t\Omega_1''(P)} \frac{\theta(U(P) + tV'' + nD - X)}{\theta(U(P) - X)}, \tag{16}$$

con

$$\tilde{r}_n(t) = r_n(t) \frac{\theta(Z)\theta(Z - D)}{\theta(tV'' + nD + Z)\theta(tV'' + (n-1)D + Z)}^{\bar{T}_2}. \tag{17}$$

2. Degeneración de las soluciones

Supongamos que Ej son reales, Em < Ej si m < j, e intentemos evaluar los límites de todos los objetos en la Fórmula (16) cuando E2m, E2m−1 tiende a −αm, αm = $\kappa_{metro.}^{2}$ κm > 0, para 1 ≤ m ≤ g.

Como en la sección anterior, θ se construye a partir de la matriz de los períodos B del superficie Γ, los coeficientes cjk están relacionados con los diferenciales abelianos dUj por

$$dU_j = \frac{dz}{2^{g} \prod_{k=1}^{g-k} \sum_{k=1} c_{jkz}^{g+2} (z - E_k)}, \tag{18}$$

y los coeficientes cjk se pueden obtener resolviendo el sistema de ecuaciones lineales

$$dU_j = \delta_{jk}, \ 1 \le j \le g, \ 1 \le k \le g. \tag{Alaska}$$

En el resto de este artículo, utilizamos las siguientes notaciones:

$$\kappa_{mj} = (\alpha_m + E_{2g+1})(\alpha_j + E_{2g+2}), \kappa_{jm} = (\alpha_j + E_{2g+1})(\alpha_m + E_{2g+2}), C_k(z) = (z - E_{2g+1})(\alpha_k - E_{2g+2}), D_k(z) = (z - E_{2g+2})(\alpha_k - E_{2g+1}), L_k = \alpha_k + E_{2g+1}, M_k = \alpha_k + E_{2g+2}, F(z) = z - E_{2g+1}, G(z) = z - E_{2g+2}. \tag{19}$$

2.1. Límite de $P(z) = \prod_{j=1}^{2g+2}(z - E_j)$

Ahora estudiamos el límite de $P(z) = \prod_{j=1}^{2g+2}(z - E_j)$.

Machine Translated by Google

El límite de $P(z) = \prod_{j=1}^{2g+2} (z - E_j)$ es obviamente igual a $P\tilde{}(z) = \prod_{j=1}^{\bar{}} {}^{2}_{(z+\alpha_j)} (z - E_{2g+1})(z - E_{2g+2})$ o, con (19),

$$P\tilde{}(z) = \prod_{j=1}^{\bar{}} ((z + \alpha_j)F(z)G(z))^2 . \tag{20}$$

2.2. Límite de $dU_m = \frac{\sum_{k=1}^{\bar{}} c_{mk}\, z^{g-k}}{2g\sqrt{\prod_{k=1}^{2} (z-E_k)}}\, dz$

Ahora estudiamos el límite de $dU_m = \frac{z^{g}\, z^{-k} \sum_{k=1}^{\bar{}} c_{mk} z}{dz \cdot 2g+2\ (z - E_k)\ \prod_{k=1}^{m(z)} (z + \alpha_j)}\, dz$, donde $(z$

El límite de dUm es igual a $dU_m = g \prod_{j=1}^{\tilde{}} - E_{2g+1})(z - E_{2g+2})$

$metro(z) = \sum_{k=1}^{\bar{}} c\tilde{}_{mk} z^{g-k}$. La condición de normalización toma la forma en el límite.

$$dU_j \to = \delta_{kj}, \prod_{m=k}^{\text{Alaska}} (\alpha_m - \alpha_k)\ (\alpha_k + E_{2g+1})(\alpha_k + E_{2g+2}) \tag{21}$$

lo que demuestra que los números $-\alpha_m$, $m = k$ son los ceros de los polinomios $_k(z)$; por lo tanto, $_k(z)$ se puede escribir como $_k(z) = c\tilde{}_{k1} \prod_{m=k} (z + \alpha_m)$.

Por (21), obtenemos en el límite

$$c\tilde{}_{k1} = \frac{2\pi i}{(\alpha_k + E_{2g+1})(\alpha_k + E_{2g+2})}.$$

Entonces,

$$_k(z) = \frac{(\alpha_k + E_{2g+1})(\alpha_k + E_{2g+2})}{2\pi i} \prod_{m=k} (z + \alpha_m).$$

Además,

$$dU\tilde{}_k(z) = \frac{_k(z)dz}{(z - E_{2g+1})(z - E_{2g+2}) \prod_{metro=1}^{\bar{}} (z + \alpha_m)};$$

en otras palabras,

$$dU\tilde{}_k(z) = dz \frac{(\alpha_k + E_{2g+1})(\alpha_k + E_{2g+2})}{2\pi i\ (z - E_{2g+1})(z - E_{2g+2})(z + \alpha_k)}$$

y con (19),

$$dU\tilde{}_k(z) = dz \frac{L_k M_k}{2\pi i F(z)G(z)(z + \alpha_k)} \tag{22}$$

2.3. Límite de Bmk El

tema de esta subsección es el estudio del límite de Bmk.

Tenemos

$$yo = \int_{soy}^{E_{2g+2}} \frac{1}{2}\, dU_k \to B_{mk}.$$

La integral I se puede evaluar fácilmente a lo largo del eje real en la hoja superior de la superficie Γ, y obtenemos

Machine Translated by Google

$$B\tilde{}mk = \frac{i}{\pi} en \frac{\sqrt{(\alpha m + E2g+1)(\alpha k + E2g+2)} + \sqrt{(\alpha k + E2g+1)(\alpha m + E2g+2)}}{\sqrt{(\alpha m + E2g+1)(\alpha k + E2g+2)} - \sqrt{(\alpha k + E2g+1)(\alpha m + E2g+2)}},$$

o con las notaciones anteriores (19),

$$B\tilde{}mk = \frac{i}{\pi} En \frac{\kappa mk + \kappa km}{\kappa mk - \kappa km}. \quad (23)$$

Entonces, iBkk tiende a −∞.

Es más, tenemos

$$\int_{E2g+2}^{PAG} dU\tilde{}k = \frac{-yo}{2\pi} en \frac{\sqrt{(z - E2g+1)(\alpha k + E2g+2)} + \sqrt{(z - E2g+2)(\alpha k + E2g+1)}}{\sqrt{(z - E2g+1)(\alpha k + E2g+2)} - \sqrt{(z - E2g+2)(\alpha k + E2g+1)}},$$

o con las notaciones anteriores (19),

$$\int_{E2g+2}^{PAG} dU\tilde{}_k = \frac{-yo}{2\pi} en \frac{Ck(z) + Dk(z)}{Ck(z) - Dk(z)}. \quad (24)$$

2.4. Límite de Ω ″ 1

En esta subsección, estudiamos el límite de Ω Ω 1 ″ 1 .

Ω ″ es una integral abeliana de segundo tipo que satisface las condiciones

$$\int_{Alaska} d\Omega\ ''_1(P) = 0, \ 1 \le k \le g,$$

tal que

$$1\ \Omega''_{(P)} = \pm (z + R0 \mp O(z\ 1\ 2^{-1})), \quad PAG \to PAG_{\infty}^{\pm}$$

El límite de dΩ ″ (P) es igual a dΩ˜ “ 1 (P) = 1 $\frac{\varphi(z)}{\prod_{j=1}^{gramo} dz, (z + \alpha j)\ (z - E2g+1)(z - E2g+2)}$

donde $\varphi(z) = \sum (P)\ \sum_{k=0}^{g+1} c\tilde{}kz\ g+1-k$.

Es más, las condiciones Ω ″ (P), satisfaciendo la condición Ω ″ 2 $\frac{1}{2}z + O(1)$ cuando z → ±∞, tenemos $c\tilde{}0 = \frac{1}{1}$

$$\int_{Alaska} d\Omega''_1(P) = 0, \ 1 \le k \le gramo$$

demuestre que −α1, . . . , −αg son los ceros de φ; por lo tanto, $\varphi(z) = \frac{1}{1\ 2}\prod_{metro=1}^{} (z + \alpha m)(z - \beta)$.

Tenemos que $P(z) = \prod (z + \sum_{k=1}^{2g+2} (z - Ek)$ tiende a $P\tilde{}(z) = (z - E2g+1)(z - E2g+2) \prod_{metro=1}^{}$ αm), y obtenemos

$$d\Omega\tilde{}\ “\ (z) = dz;\ 1\ 2\ \frac{z}{(z - E2g+1)(z - E2g+2)}$$

$$\Omega'' \ (P) \text{ tiende a } 1$$

$$\Omega\tilde{}\ ''_{(P) = 1} \int_{E2g+2}^{z} \frac{tu}{2\ (u - E2g+1)(u - E2g+2)} Du.$$

Machine Translated by Google

Esto se puede evaluar y da

$$\Omega_1(P) = \frac{E_{2g+1} + E_{2g+2}}{4}\Omega'' \text{ en } \frac{\sqrt{z - E_{2g+1}} + \sqrt{z - E_{2g+2}}}{\sqrt{z - E_{2g+1}} - \sqrt{z - E_{2g+2}}} + \frac{1}{2}\sqrt{(z - E_{2g+1})(z - E_{2g+2})}.$$

Con las notaciones definidas en (19), se puede reescribir como

$$\Omega_1(P) = En\frac{E_{2g+1} + E_{2g+2}}{4} + \frac{F(z) + G(z)\Omega'}{F(z) - G(z)} + 2\pi i F(z)G(z). \tag{25}$$

2.5. Límite de Ω'_0

Consideramos Ω'_0, y estudiamos su límite.

Ω'_0 es una integral abeliana de tercer tipo que satisface las condiciones

$$\int_{A_k} d\Omega'_0(P) = 0, \quad 1 \le k \le g,$$

tal que

$$\Omega'_0(P) = \pm(\ln z - \ln 0 + O(z^{-1})), \quad P_{A_k} \to P_{A_k^\pm}.$$

El límite de $d\Omega'_0(P)$ es igual a

$$d\Omega'_0(P) = g\prod_{j=1}\frac{\sum_{k=0}^{g} c'_k z^{g-k}}{(z + \alpha_j)\sqrt{(z - E_{2g+1})(z - E_{2g+2})}}dz.$$

$\Omega'_0(P)$ satisface la condición $\Omega'_0(P) = \pm \ln(z) + O(1)$, y cuando $z \to \pm\infty$, tenemos $c'_0 = 1$.

Es más, las condiciones

$$\int_{A_k} d\Omega'_0(P) = 0, \quad 1 \le k \le g$$

demuestre que $-\alpha_1, \ldots, -\alpha_g$ son los ceros de φ; por lo tanto, $\varphi(z)$, definido por $\varphi(z) = \sum_{k=0}^{g} c'_k z^{g-k}$, puede escribirse como $\varphi(z) = \prod_{m=1}^{g}(z + \alpha_m)$.

Como $P(z) = \prod_{k=1}^{2g+2}(z - E_k)$ tiende a $P'(z) = (z - E_{2g+1})(z - E_{2g+2})\prod_{m=1}^{g}(z + \alpha_m)$, obtenemos

$$d\Omega'_0(z) = \frac{1}{\sqrt{(z - E_{2g+1})(z - E_{2g+2})}}dz,$$

y $\Omega'_0(P)$ tiende a $\Omega'_1(P) = \int_{E_{2g+2}}^{z} \frac{1}{\sqrt{(u - E_{2g+1})(u - E_{2g+2})}}Du.$

Esto se puede calcular y da

$$\Omega'_0(P) = En\frac{\sqrt{z - E_{2g+1}} + \sqrt{z - E_{2g+2}}}{\sqrt{z - E_{2g+1}} - \sqrt{z - E_{2g+2}}}.$$

Se puede escribir con (19) como

$$\Omega'_0(P) = En\frac{F(z) + G(z)}{F(z) - G(z)}. \tag{26}$$

2.6. Límite de X

Deducimos el límite de X.

De las secciones anteriores,

$$B'_{mk} = \frac{i}{\pi} En\frac{\kappa_{mk} + \kappa_{km}}{\kappa_{mk} - \kappa_{km}},$$

Machine Translated by Google

$$\tilde{U}_k(P) = \frac{i}{\ln 2\pi} \frac{Ck(z) + Dk(z))}{Ck(z) - Dk(z)}.$$

Como K está definido por $Km = \sum_{j=1}^{\leftarrow} Bjm - \overline{m2}$, y $X = K + \sum_{j=1}^{g} U(Pj)$, podemos escribir

$$Xm \to \tilde{X}m = \frac{i}{\pi} \sum_{j=m}^{\leftarrow} \ln \kappa mj \frac{\kappa mj + \kappa jm}{-\kappa jm} - \frac{i}{2} \text{metro} + \frac{i}{2\pi} \sum_{j=1}^{\leftarrow} \ln \frac{Cm(zj) + Dm(zj)}{Cm(zj) - Dm(zj)} + yo\infty. \quad (27)$$

2.7. Límite de $U(P \infty^{+})$, $U(P - \infty)$, D y Z En

esta subsección, estudiamos el límite de $U(P \infty)$, $U(P \pm\infty)$, D y Z.
De las secciones anteriores, es fácil obtener los límites de $\tilde{U}_k$ como

$$\lim_{P \to P^{\infty+}} \tilde{U}_k(P) = \ln 2\pi^{-yo} \frac{\overline{\alpha k + E2g+2} + \alpha k + \overline{E2g+1\ \alpha k +}}{\overline{E2g+2} - \alpha k + E2g+1},$$

o

$$\lim_{P \to P^{\infty+}} \tilde{U}_k(P) = \ln 2\pi^{-yo} \frac{Mk + Lc}{Mk + Lc}. \quad (28)$$

$$Dk = \text{Reino Unido}(P^{\infty+}) - \text{Reino Unido}(P_{\infty}) \to \tilde{D}_k = \frac{i}{\ln 2\pi} \frac{Mk + iLk}{Mk - iLk}. \quad (29)$$

De este modo,

$$\lim_{P \to P^{\infty-}} \tilde{U}_k(P) = \ln 2\pi^{-yo} \frac{\overline{\alpha k + E2g+2} + \alpha k + \overline{E2g+1\ \alpha k +}}{\overline{E2g+2} - \alpha k + E2g+1} - No,$$

o

$$\lim_{P \to P^{\infty-}} \tilde{U}_k(P) = \ln 2\pi^{-yo} \frac{Mk + Lc}{Mk + Lc} + \frac{i}{2\pi} \ln \frac{Mk + iLk}{Mk + iLk}. \quad (30)$$

Por tanto, el límite de $Z = U(P \infty) - X$ viene dado por

$$Zk \to \tilde{Z}_k = \frac{i}{\pi} \sum_{j=1,\ j=k}^{\leftarrow} \ln \frac{\kappa kj + \kappa jk}{\kappa kj - \kappa jk} + \frac{k}{2} - \frac{i}{2\pi j} \sum_{=1,}^{\leftarrow} \ln \frac{Ck(zj) + Dk(zj)}{Ck(zj) - Dk(zj)} - i\infty - \ln \frac{i}{2\pi} \frac{Mk + Lk}{Mk - Lk} \quad (31)$$

2.8. Límite de V″ En este

párrafo, el límite de $V\, d\Omega''\, V$ se define en (13). k'' se estudia.

$$'' = \frac{1}{2\pi i\, bk} \quad 1$$

tenemos yo $= \int_{-\alpha k}^{E2g+2} d\Omega\, 1\, '' \to \frac{1}{2\, bk\, 1}\, d\tilde{\Omega}''$

Como en la determinación del límite de $\Omega''_{1,}$ tenemos

$$\int_{-\alpha k}^{E2g+2} d\tilde{\Omega}\, \text{“}(P) = -1 \frac{E2g+1 + E2g+2}{4} \text{en} \frac{\overline{\alpha k + E2g+1} + \alpha k + \overline{E2g+2\ \alpha k +}}{\overline{E2g+1} - \alpha k + E2g+2} - \frac{1}{2} \overline{(\alpha k + E2g+1)(\alpha k + E2g+2)}.$$

Entonces,

$$V''_k = \frac{1}{2\pi i\, bk} \text{“}\, d\Omega| \to \tilde{V} = -\frac{E2g+1 + E2g+2}{4\pi i} \text{en} \frac{\overline{\alpha k + E2g+1} + \alpha k + \overline{E2g+2\ \alpha k +}}{\overline{E2g+1} - \alpha k + E2g+2} - \frac{1}{2\pi i} \overline{(\alpha k + E2g+1)(\alpha k + E2g+2)},$$

$$\tilde{V}\text{“}_k = \frac{E2g+1 + E2g+2}{4\pi i} \ln \frac{Lk - Mk}{Lc + Mk} - \frac{1}{2\pi i} LkMk. \quad (32)$$

Machine Translated by Google

2.9. Límite de θ(p)

Determinamos el límite de θ(p).

Denotemos a A como el argumento de θ(p) = ∑k Zg exp{πi(Bk|k) + 2πi(k|p)}.
A se puede reescribir en la forma

$$A = \pi i \sum_{j=1}^{\leftarrow} Bjjkj(kj - 1) + 2\pi i \sum_{j>m} Bmjkmkj + \sum_{j=1}^{\leftarrow} \pi i(2pj + Bjj)kj .$$

Usando la desigualdad kj(kj − 1) ≥ 0 para todo k Z g y el hecho de que iBkk tiende a −∞, tal que cada podemos reducir el límite ˜θ de θ(p) a una suma finita tomada sobre los vectores k Z ,
g kj debe ser igual a 0 o 1.

En esta sección, calculamos el límite de todos los términos en la expresión de la solución. ψ(P, t) en θ. Denotamos pi como los argumentos de estas diferentes expresiones.

Primero estudiamos los términos de r˜n correspondientes a los argumentos $p1 = \overset{+}{Z} = U(P_{\infty}) - X$, p2 = Z − D, p3 = tV″ + nD − Z, p4 = tV″ + (n − 1)D - Z.

Luego, estudiamos los términos restantes de ψn correspondientes a los argumentos p5 = U(P) − X, p6 = U(P) + tV″ + nD − X.

Con estas notaciones, la solución ψ(n, , t, z) se puede escribir como

$$\psi n(P, z) = rn(t) \frac{\theta(p1)\theta(p2)}{\theta(p3)\theta(p4)} \left(\frac{1}{2}\frac{\theta(p6)}{\theta(p5)}\right)^{norteΩ_0 (P)+tΩ'' 1 (PAG)}.$$

2.10. Límite de θ(p1)

Denotamos p1 como el término $Z = U(P_{\infty}^{+}) - X$.

Estudiamos su límite.

Entonces, θ(p1) → ˜θ1:

$$\tilde{\theta}1 = \sum \exp \sum k \quad Zg,\ kj = 0\ o\ 1\ j>l \ln \frac{\kappa lj - \kappa jl\ 2}{kjkl + \kappa lj + \kappa jl} \sum_{j=1}^{\leftarrow} kj \ln \frac{Mj + Lj}{Mj - Lj}$$

$$\sum 2 \ln \kappa jl - \frac{\kappa jl + \kappa lj +}{\kappa lj\ l \quad = j} j\pi i + \sum_{l=1}^{\leftarrow} \ln + \frac{Cj(zl) + Dj(zl)}{Cj(zl) - Dj(zl)} .$$

2.11. Límite de θ(p2)

Consideramos p2 = Z − D. Estudiamos su límite.

Entonces, θ(p2) → ˜θ2. Como Dj → D˜ es fácil ver que j ,

$$\tilde{\theta}2 = \sum \exp \sum k \quad Zg,\ kj = 0\ o\ 1 \ln \frac{\kappa lj - \kappa jl\ 2}{kjkl + \kappa lj + \kappa jl} \sum_{j=1}^{\leftarrow} kj \ln \frac{Mj + Lj}{Mj - Lj}$$

$$\sum 2 \ln \kappa jl - \frac{\kappa jl + \kappa lj +}{\kappa lj\ l \quad = j} j\pi i + \sum_{l=1}^{\leftarrow} \ln + \frac{Cj(zl) + Dj(zl)}{Cj(zl) - Dj(zl)} + \ln \frac{Mj + iLj}{Mj - iLj} .$$

2.12. Límite de θ(p3)

Sea p3 tV ″ + nD + Z. Estudiamos su límite.

Entonces, θ(p3) → ˜θ3:

$$\tilde{\theta}3 = \sum_{k} Zg,\ kj = 0\ o\ 1 \exp \sum_{j>l} \ln \frac{\kappa lj - \kappa jl\ 2}{kjkl + \kappa lj + \kappa jl} \sum_{j=1}^{\leftarrow} kj \ln \frac{Lj + Mj}{Lj - Mj}$$

$$+t \frac{Ej+1 + E2g+2\ 2}{} en \frac{Lj - Mj}{Lj + Mj} - LjMj + n\ln \frac{Mj + iLj}{Mj - iLj}$$

Machine Translated by Google

$$+\sum_{l=j} \ln + \frac{\kappa jl + \kappa lj}{j\pi i + \kappa jl - \kappa lj} \sum_{l=1}^{\leftarrow} en \frac{Cj(zl) + Dj(zl)}{Cj(zl) - Dj(zl)} .$$

2.13. Límite de θ(p4)

Consideramos p4 = tV″ + (n − 1)D + Z; θ(p4) → ˜θ4. Determinamos su límite.

Como Dj → D˜ es fácil verificar que j ,

$$\tilde{\theta}4 = \sum \exp \sum k \quad Zg, kj = 0 \text{ o } 1 \quad \kappa lj + \frac{\kappa lj - \kappa jl}{\kappa jl\, j=1} kjkl + 2 \ln \quad \overleftarrow{kj} \ln \quad \frac{Lj + Mj \sum}{Lj - Mj}$$

$$+t \quad \frac{Ej+1 + E2g+2}{2} \quad en \frac{Lj - Mj}{Lj + Mj} + LjMj + (n-1)\ln \quad \frac{Mj + iLj}{Mj - iLj}$$

$$+\sum_{l=j} \kappa lj \quad \frac{\kappa jl + \kappa lj \sum}{2 \ln + j\pi i + \kappa jl -} \overleftarrow{l=1} \; en \frac{Cj(zl) + Dj(zl)}{Cj(zl) - Dj(zl)} .$$

2.14. Límite de θ(p5)

El término p5 = U(P) − X tiende a p˜5 = U˜ (P) − X˜ .
Determinamos su límite.

Entonces, θ(p5) → ˜θ5:

$$\tilde{\theta}5 = \sum_{k \quad Zg, kj = 0 \text{ o } 1} \exp \sum_{j>l} {}^{2 \text{ litros}} \frac{\kappa lj - \kappa jl}{\kappa jl} kjkl + \kappa lj + \sum_{j=1}^{\leftarrow} kj \ln \frac{Cj(z) - Dj(z)}{Cj(z) + Dj(z)}$$

$$+\sum_{l=j} \kappa lj \quad \frac{\kappa jl + \kappa lj \sum}{2 \ln + j\pi i + \kappa jl -} \overleftarrow{l=1} \; en \frac{Cj(zl + Dj(zl)}{Cj(zl - Dj(zl)} .$$

2.15. Límite de θ(p6)

El término p6 = U(P) + tV″ + nD − X tiende a p˜6 = U˜ (P) + tV˜ “ − X˜ + nD˜ .
Luego, determinamos su límite. θ(p6) → ˜θ6 :

$$\tilde{\theta}6 = \sum \exp \sum k \quad Zg, kj = 0 \text{ o } 1 \quad \ln \frac{\kappa lj - \kappa jl\, 2}{kjkl + \kappa lj + \kappa jl\, j=1} \quad \overleftarrow{kj} \ln \quad \frac{Lj + Mj \sum}{Lj - Mj}$$

$$+t \quad \frac{Ej+1 + E2g+2}{2} \quad en \frac{Lj - Mj}{Lj + Mj} - LjMj + n\ln \quad \frac{Mj + iLj}{Mj - iLj}$$

$$+\sum_{l=j} \ln \frac{\kappa jl + \kappa lj\, 2}{\kappa jl - \kappa lj} + j\pi i - \sum_{l=1}^{\leftarrow} en \frac{Cj(zl) + Dj(zl)}{Cj(zl) - Dj(zl)} + en \frac{Cj(z) - Dj(z)}{Cj(z) + Dj(z)} .$$

2.16. Límite de soluciones Ψ

De la sección anterior, podemos dar el límite ψ˜ de ψ. Toma la siguiente forma:

$$\psi\tilde{}(z, n, t)) = \frac{rn(t)\ \tilde{\theta}1\ \tilde{\theta}2}{\tilde{\theta}5} \times \frac{{}^{mi} \quad {}^{norte \ln \frac{F(z)+G(z)}{F(z)-G(z)}} + t \quad \frac{E2g+1+E2g+2}{4} en \frac{F(z)+G(z)}{F(z)-G(z)} \quad {}^{1+2} F(z)G(z)}{\tilde{\theta}3\ \tilde{\theta}4} \quad \tilde{\theta}6 .$$

En la expresión anterior, ˜θ1, ˜θ2 y ˜θ5 son independientes de n y t.
Sólo ˜θ3, ˜θ4 y ˜θ6 dependen de n y t.
Elegimos el caso particular en el que t = 0.
Reemplazamos n por x.
Denotamos c como el coeficiente definido por

Machine Translated by Google

$$c = \frac{rn(0)\,\tilde{}\theta 1\,\tilde{}\theta 2^{\tilde{T}2}}{\tilde{}\theta 5}.$$

Denotamos H como la función definida por

$$H(z) = \frac{F(z) - GRAMO(z)}{F(z) + G(z)}.$$

En las diferentes sumas que involucran a θ, sólo quedan términos con kj = 1. Entonces, las sumas se pueden reducir sólo en los subconjuntos J de [1; g] del número entero N. Denotamos aj como el término $a_j = 2^{-1}\,\frac{Mj+iLj}{Mj-iLj}$. Entonces, el término ˜θ6 se puede escribir como

$$\tilde{}\theta 6 = \sum_J \prod_{\{1,\ldots,g\}j,l\ J,\ j<l} en\ \frac{\kappa lj - \kappa jl}{\kappa lj + \kappa jl}^2 \prod_{j\ J} \frac{Lj + Mj}{Lj - Mj}$$

$$\prod_{l\ Jl\ =j} \kappa jl\ \frac{\kappa jl + \kappa lj\ ln}{-\kappa lj}^2 \prod_{j\ J} \frac{Cj(zl) + Dj(zl)}{Cj(zl) + Dj(zl)}\ \frac{Cj(z) - Dj(z)}{Cj(z) + Dj(z)}^{mi\sum j\ J\ j\pi i^{mi-2x\sum j\ J\ aj}}.$$

Denotamos τ como la función definida por

$$\tau(x) = \tilde{}\theta 3(x) = \sum \prod ln\ \kappa lj + \kappa jl\ J\ \{1,\ldots,g\}_{j,l\ J,\ j<l} \frac{\kappa lj - \kappa jl}{}^2 \prod_{j\ J} \frac{Lj + Mj}{Lj - Mj}$$

$$\prod_{l\ Jl\ =j} \kappa jl\ \frac{\kappa jl + \kappa lj\ ln}{-\kappa lj}^2 \prod_{j\ J} \frac{Cj(zl) + Dj(zl)}{Cj(zl) - Dj(zl)}^{mi\sum j\ J\ j\pi i^{mi-2x\sum j\ J\ aj}}.$$

Por tanto, la solución Ψ˜ se puede escribir como

$$\Psi\tilde{}\,(x, z) = \frac{\exp(xH(z))\chi(x, z)}{c\,\tau(x)\tau(x-1)}. \tag{33}$$

Aquí volvemos a la formulación de Sato de la solución, como se indica en [15].

Esta expresión es muy similar a la solución expresada en la página 5830, y la expresión sión de la solución está dada por

$$\Psi\tilde{}\,(x, z) = \frac{\exp(xz)\chi(x, z)}{\tau(x)\tau(x-1)}. \tag{34}$$

La diferencia entre estas dos afirmaciones proviene del hecho de que estas dos las ecuaciones se tratan de manera diferente:

Van Diejen considera la ecuación

$$a(x)f(x+1) + b(x)f'x) + a(x-1)f(x-1) = \lambda\, f(x);$$

aquí consideramos (1)

$$\frac{rn}{rn-1}\sqrt{cn\psi n+1 + vn\psi n + rn+1} \quad \frac{rn}{rn-1}\sqrt{cn-1\psi n-1} = \lambda\psi n.$$

2.17. Límite de los potenciales asociados

En esta subsección, determinamos el límite de los potenciales asociados.

Los potenciales un = $\sqrt{cn}$ y vn se describen en (6) y (7) como

$$cn = \frac{mi2l0\ \theta^2(tV'' + nD + Z)}{\theta(tV'' + (n+1)D + Z)\theta(tV'' + (n-1)D + Z)}$$

Machine Translated by Google

$$-R0 + \partial t \ln \theta(tV'' + \frac{\theta(tV'' + (n-1)D + Z)\, vn =}{nD + Z)}.$$

Entonces, el límite de un es igual a

$$u\tilde{}n = k1 \frac{\tau(x)}{\tau(x+1)\tau(x-1)},$$

que es lo mismo que en [15] (p. 5830), hasta la constante k1.

Denotando $\sigma(x) = \partial t\theta(tV\tilde{} \text{“} + xD\tilde{} + Z\tilde{})t=0$, el límite de vn es igual a $\sigma(x -$

$$k2 + \tau(x - \frac{1)\, v\tilde{}n =}{1)} - \frac{\sigma(x)}{\tau(x)},$$

que es similar al escrito en [15] (p. 5830), hasta la constante k2.

3. Conclusiones

Hemos utilizado la degeneración de soluciones de algunas PDE dadas en términos de funciones theta de Riemann para obtener algunas soluciones importantes. En particular, en el caso de la ecuación NLS, hemos logrado construir una solución cuasi-racional que involucra el determinante de orden N para cada entero positivo N dependiendo de 2N − 2 parámetros reales [12]. En el caso de la ecuación KdV, a partir de funciones abelianas, construimos soluciones en términos de los determinantes de Fredholm y los Wronskianos [13].

En este estudio hemos conseguido recuperar fórmulas de Sato utilizando este método para la ecuación de Toda. A partir de soluciones dadas en términos de las funciones de Baker-Akiezer, logramos construir por degeneración, como se da en el marco de la ecuación NLS [14], algunas soluciones cuasi-racionales, y recuperamos las fórmulas de Sato para la ecuación de Toda.

Precisamente, derivamos soluciones multisolitones a partir de soluciones de espacios finitos mediante una transición límite , es decir, haciendo que los espacios tiendan hacia puntos en una determinada superficie de Riemann. Una ventaja de este enfoque es que no se basa en la teoría de la dispersión inversa ni en métodos geométricos y teóricos de representación, lo que ofrece otra perspectiva del proble

La degeneración de las soluciones de la ecuación NLS ha permitido construir nuevas soluciones cuasi-racionales de orden N en función de 2N − 2 parámetros reales y su construcción hasta el orden 23, que antes eran desconocidas. En el caso de la ecuación de KdV, la degeneración de las soluciones permitió encontrar las soluciones dadas por el método de Darboux, que constituyó un puente entre el enfoque del álgebra geométrica y el marco de las transformaciones de Darboux. En este artículo, esta es más o menos la misma situación que la de la ecuación KdV, donde vinculamos el enfoque álgebro-geométrico al marco del método de dispersión inversa.

Financiamiento: Esta investigación no recibió financiamiento externo.

Declaración de disponibilidad de datos: Los datos que respaldan los hallazgos de este estudio están disponibles en el artículo.

Conflictos de intereses: El autor declara no tener conflictos de intereses.

Referencias

1. Toda, M. Vibración de una cadena con interacción no lineal. J. Física. Soc. Japonés. 1967, 6, 22–431. [Referencia cruzada]
2. Fecha, E.; Tanaka, S. Soluciones periódicas de múltiples solitones de la ecuación de Korteweg-de Vries y la red Toda. Supl. Prog. Teor. Física. 1976, 59, 107–125. [Referencia cruzada]
3. Dubrovin, BA Funciones theta y ecuaciones no lineales. Ruso. Matemáticas. Sobrevivir. 1981, 36, 11. [Referencia cruzada]
4. Stahlhofen, AA; Matveev, VB Posiciones para la red de Toda y problemas espectrales relacionados. Física. Una Matemática. Gen. 1995, 28, 1957–1965. [Referencia cruzada]
5. Zhang, S.; Zhou, YY Soluciones multiondas para la ecuación reticular de Toda generalizando el método de la función exp. Int. J. Aplica. Matemáticas. 2014, 44, 1–6.
6. Duarte, VN Una clase de soluciones de la ecuación reticular bidimensional de Toda. Física. Letón. 2021, 385, 126979. [Referencia cruzada]
7. Nilson, T.; Schiebold, C. Fórmulas de solución para la red de Toda bidimensional y solitones en forma de partículas con inesperados comportamiento asintótico. J. Nonliner. Matemáticas. Física. 2020, 27, 57–94. [Referencia cruzada]

Machine Translated by Google

8. Bellokolos, ED; Bobenko, AI; Enol'skii, VZ; Es, AR; Matveev, VB Enfoque álgebro-geométrico de ecuaciones integrables no lineales; Series en Dinámica No Lineal; Springer-Verlag: Berlín/Heidelberg, Alemania, 1996.
9. Dubrovin, BA; Matveev, VB; Novikov, SP Sistemas integrables: ecuaciones no lineales del tipo Korteweg-De Vries, operadores lineales de zona finita y variedades abelianas; Serie de notas de conferencias; Sociedad Matemática de Londres: Londres, Reino Unido, 1976; Volumen 60, págs. 53-140.
10. Krichever, I. La red periódica no abeliana de Toda y la generalización bidimensional. Uspekhi Math Nauka 1981, 36, 72–77.
11. Smirnov, AO Solitones elípticos de ecuaciones no lineales integrables; Universidad Estatal de San Petersburgo: San Petersburgo, Rusia, 2002; págs. 1–170.
12. Gaillard, P. Familias de soluciones cuasirracionales de la ecuación NLS y ondas múltiples. J. Física. Una metanfetamina. Teor. 2011, 44, 435204. [Referencia cruzada]
13. Gaillard, P. Funciones theta degeneradas de Riemann, representaciones de Fredholm y wronskianas de las soluciones de la ecuación de KdV y el caso racional degenerado. J. Geom. Física. 2021, 161, 104059. [Referencia cruzada]
14. Su, AR; Rybin, AV; Salle, MA Integración exacta de la ecuación NLS. Teor. Matemáticas. Física. 1988, 74, 29–32. [Referencia cruzada]
15. Diejen, JFV Una fórmula de Sato para operadores de diferencias finitas sin reflexión. J. Matemáticas. Física. 1999, 40, 5822–5834. [Referencia cruzada]

Descargo de responsabilidad/Nota del editor: Las declaraciones, opiniones y datos contenidos en todas las publicaciones son únicamente de los autores y contribuyentes individuales y no de MDPI ni de los editores. MDPI y/o los editores renuncian a toda responsabilidad por cualquier daño a personas o propiedad que resulte de cualquier idea, método, instrucción o producto mencionado en el contenido.

Machine Translated by Google

AppliedMath

MDPI

Artículo

Modelo de circuito básico de convertidores de fuente de voltaje: Metodología y Modelado

Christian Bipongo Ndeke *, Marco Adonis y Ali Almaktoof

Departamento de Ingeniería Eléctrica, Electrónica e Informática, Universidad Tecnológica de la Península del Cabo, Ciudad del Cabo 7535, Sudáfrica; adonisma@cput.ac.za (MA); almaktoofa@cput.ac.za (AA)
* Correspondencia: ndekechristian@gmail.com o 219385017@mycput.ac.za; Teléfono: +27-62-836-2466

Resumen: Los convertidores de fuente de voltaje (VSC) se han convertido en los componentes clave de los sistemas eléctricos modernos, ya que facilitan la conversión de energía eficiente y el control flexible del flujo de energía. Comprender el modelo de circuito fundamental de los VSC es esencial para su modelado y análisis precisos en estudios de sistemas de energía. Un modelo de circuito convertidor de fuente de voltaje básico conectado a un filtro LC es esencial porque reduce las distorsiones armónicas y mejora la calidad general de la energía de la Esto garantiza un suministro de energía limpio y estable, necesario para la integración de múltiples fuentes de energía renovables y cargas sensibles. En este artículo se presenta una metodología integral para desarrollar un modelo de circuito básico de VSC, centrándose en los componentes clave y los principios involucrados. La metodología incluye el modelado de la modulación de ancho de pulso por vector espacial (SVPWM), así como el marco de referencia síncrono cero en cuadratura directa. Diferentes controles de diseño, incluido el diseño del bucle de control de corriente en el dominio S, el diseño del bucle de control de voltaje del bus de corriente continua (CC) en el dominio S y el diseño del bucle de control de voltaje de corriente alterna (CA) en el dominio S, se exploran para capturar con precisión el comportamiento dinámico y las estrategias de control de los VSC.

La metodología propuesta proporciona un marco sistemático para modelar VSC, lo que permite a ingenieros e investigadores analizar su desempeño y evaluar su impacto en la estabilidad y operación del sistema eléctrico. Se pueden realizar estudios futuros utilizando estudios de casos y escenarios de simulación para mostrar la eficiencia y aplicabilidad de los modelos desarrollados en el análisis de aplicaciones de electrónica de potencia basadas en VSC, incluidos sistemas de transmisión de corriente continua de alto voltaje (HVDC) y sistemas flexibles de transmisión de corriente alterna (FACTS).). La importancia de este trabajo radica en su potencial para avanzar en la comprensión y la aplicación de los VSC, contribuyendo a sistemas energéticos más resilientes y eficientes. Al proporcionar una base sólida para futuras investigaciones y desarrollo, este estudio respalda la integración continua de fuentes de energía renovables y el avance de la infraestructura eléctrica moderna.

Palabras clave: convertidor de fuente de voltaje; filtro LC; control actual; control de voltaje CC; control de voltaje CA

check for updates

Cita: Ndeke, CB; Adonis, M.; Almaktoof, A. Modelo de circuito básico de Convertidores de Fuentes de Voltaje: Metodología y Modelado. Matemáticas aplicadas 2024, 4, 889–907. https://doi.org/10.3390/appliedmath4030048

Editor académico: Jun-Juh Yan

Recibido: 21 de mayo de 2024
Revisado: 16 de julio de 2024
Aceptado: 18 de julio de 2024
Publicado: 29 de julio de 2024

Copyright: © 2024 por los autores. Licenciatario MDPI, Basilea, Suiza.
Este artículo es un artículo de acceso abierto. distribuido bajo los términos y condiciones de los Creative Commons Licencia de atribución (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

1. Introducción

Los convertidores de fuente de voltaje (VSC) se han convertido en componentes integrales de los sistemas eléctricos modernos, desempeñando un papel crucial a la hora de facilitar la conversión eficiente de energía y permitir un control flexible del flujo de energía [1]. A medida que continúan creciendo las demandas de integración de energías renovables, estabilización de la red y mejora de la confiabilidad de la red, el modelado y análisis precisos de los VSC se han vuelto primordiales [2,3]. La principal interfaz de alimentación que conecta las generaciones distribuidas (GD) con las microrredes es el convertidor AC-DC. El convertidor de fuente de voltaje funciona en dos modos distintos, operando como un rectificador activo para el control del bus de corriente continua (CC) y, como se muestra en la Figura 1, actuando como un inversor durante las operaciones de alimentación y formación de red utilizando distintos bucles de control [4]. Comprender el modelo de circuito básico de los VSC, junto con las metodologías para su control y modelado, es esencial para que ingenieros e investigadores diseñen, analicen y optimicen eficazmente aplicaciones de electrónica de potencia basadas en VSC [5]. El objetivo de este trabajo es presentar un

Machine Translated by Google

descripción general de la metodología, estrategias de control y técnicas de modelado asociadas con el modelo de circuito básico de VSC. Este artículo profundiza en los principios fundamentales que subyacen al funcionamiento de los VSC, centrándose en los componentes clave y sus interacciones dentro del circuito. Se hace hincapié en dilucidar diferentes metodologías, como la modulación de ancho de pulso por vector espacial (SVPWM), el marco de referencia síncrono cero en cuadratura directa y su integración en circuitos VSC.

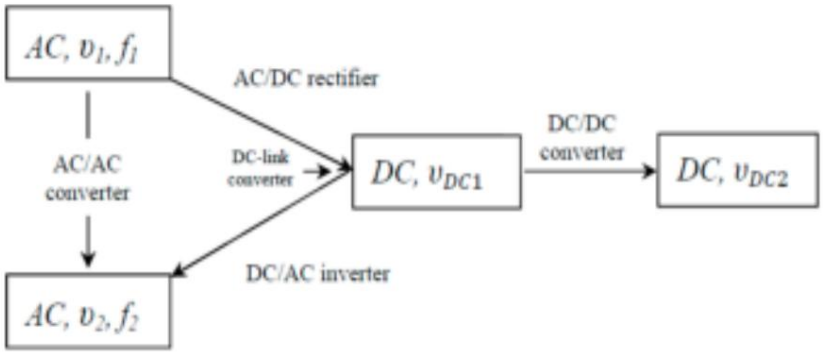


Figura 1. Grupos de convertidores de potencia de estado sólido en función de su transformación operativa [6].

Un modelo de circuito convertidor de fuente de voltaje fundamental conectado al filtro LC es crucial porque ayuda a reducir las distorsiones armónicas y mejorar la calidad general de la energía en la microrred. Esto garantiza un suministro de energía estable y limpio, lo cual es esencial para cargas sensibles y la integración de diversas fuentes de energía renovables. También desempeña un papel clave en la conversión de energía CC procedente de fuentes renovables, como paneles solares o baterías, a energía CA, que es necesaria para la mayoría de las cargas y para la sincronización con la red principal. El filtro LC ayuda a suavizar la salida, haciendo que el proceso de conversión sea más eficiente. Comprender el modelo de circuito VSC y su interacción con el filtro LC es esencial para diseñar sistemas que puedan integrarse perfectamente con la red principal u operar en modo isla. Esto garantiza que la microrred pueda manejar las fluctuaciones de la oferta y la demanda sin comprometer la estabilidad. En esencia, la importancia de introducir el modelo de circuito fundamental VSC conectado al filtro LC radica en su capacidad para abordar desafíos clave en la conversión de energía, la estabilidad, la integración y la gestión de sistemas híbridos de microrredes renovables.

Además, este artículo explora varias estrategias de control empleadas en VSC para regular el voltaje, la corriente y el flujo de energía en aplicaciones independientes y conectadas a la red. Se analizan varias técnicas de control que utilizan modulación de ancho de pulso (PWM) y control integral proporcional (PI). Además, también se introducen los aspectos de modelado del diseño y control de VSC para desarrollar modelos precisos y dinámicos que capturen el comportamiento transitorio y de estado estable de estos dispositivos. Para garantizar la relevancia y aplicabilidad de los modelos presentados en escenarios del mundo real, es necesario abordar consideraciones prácticas, como la estimación de parámetros, la validación y los desafíos de implementación . Se recomienda realizar estudios de casos y simulaciones para mostrar la eficiencia y precisión de las metodologías y estrategias de control presentadas en el análisis de aplicaciones de electrónica de potencia basadas en VSC, como sistemas flexibles de transmisión de corriente alterna (FACTS), energía renovable conectada a la red. y sistemas de transmisión de corriente continua de alto voltaje (HVDC).

Las siguientes secciones describen las partes restantes del documento: En la Sección 2, se presenta el modelado VSC de dos niveles conectado a la red con un filtro LC. El diseño del bucle de control de corriente en el dominio S se presenta en la Sección 3. La Sección 4 presenta el diseño de un bucle de control de voltaje del bus de CC en el dominio S. En la Sección 5, se proporciona el diseño de un bucle de control de voltaje CA en el dominio S. La Sección 6 ofrece una breve discusión y la Sección 7 concluye el artículo.

Machine Translated by Google

2. Modelado VSC de dos niveles conectado a la red con filtro LC En la Figura

2 se presenta un modelo de circuito típico de un inversor de dos niveles con un filtro LC conectado a una carga trifásica . La señal aplicada a los terminales a, a`, b, `b, cy c` regula la potencia de los seis interruptores que forman la salida, denominados S1 a S6. Se supone que los interruptores S1 a S6 son complementarios [7]. Normalmente, un convertidor de fuente de voltaje con un filtro LCL está regulado por un controlador lineal convencional aumentado por un bucle de amortiguación para garantizar la estabilidad del sistema [8]. Además, tanto el sistema de CA trifásico como la fuente de voltaje del lado de CC pueden recibir energía directamente del VSC. ff En consecuencia, hay dos modos diferentes de operación para el VSC, ya sea el modo inversor (DC/AC) el modo rectificador (AC/DC).

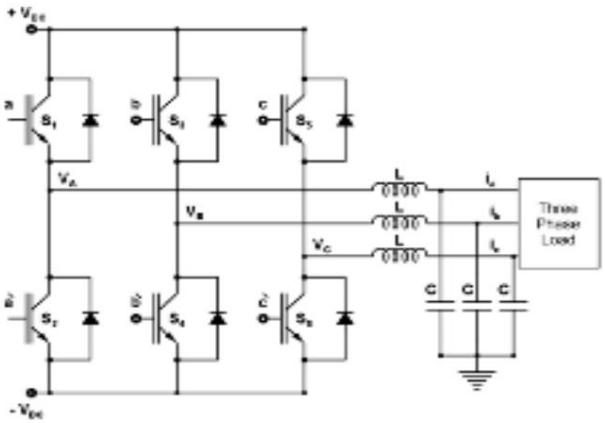


Figura 2. Convertidor de fuente de tensión trifásico-2 niveles [7].

Para conectar los sistemas de generación distribuida a la red, comúnmente se emplean convertidores de fuente de voltaje con modulación de ancho de pulso trifásicos (PWM-VSC) debido a su superior controlabilidad y mayor eficiencia en comparación con otros convertidores [9]. Para habilitar el control de voltaje del lado de CC y el factor de potencia en el punto de acoplamiento común (PCC), se utiliza una topología PWM-VSC de dos niveles. Sin embargo, la implementación del PWM da como resultado que aparezcan armónicos de banda base y banda lateral en el espectro de armónicos de voltaje de CA, lo que requiere el uso de filtros para reducir los armónicos de la corriente de la red [10].

2.1. Metodología de modulación de ancho de pulso de vector espacial (SVPWM)

El uso de la estrategia de comparación de triángulos simplifica la ejecución de este procedimiento.
En comparación con la comparación triangular PWM, logra un aumento del 15% en el voltaje de CA y reduce la distorsión armónica total (THD) de voltaje y corriente. En la Figura 3 [11] se presenta un diagrama vectorial espacial estándar de un inversor de fuente de voltaje (VSI) de dos niveles con estados de conmutación en los seis sectores . La Tabla 1 ilustra que un inversor de fuente de voltaje trifásico típico de dos niveles tiene ocho estados de conmutación, con seis estados activos que producen vectores de voltaje de +VCC o −VCC, y dos estados nulos que generan vectores de voltaje con amplitud cero. En la Figura 3, los estados de conmutación 000 y 111 representan los estados nulos, mientras que el plano del vector espacial está dividido en seis sectores iguales por los estados activos. Esta metodología emplea un vector de referencia giratorio, como se muestra en la Figura 3, muestreado una vez durante cada subciclo, Ts [12]. El objetivo principal es estimar el vector de la tensión de referencia Vre f utilizando ocho modelos de conmutación.

La Figura 3 ilustra que dos de los vectores potenciales de voltaje de salida ($\vec{V1}$ y vectores nulos $\vec{V7}$) son o cero, representados por valores iguales en las tres fases (000 o 111).
Los seis vectores restantes ($\vec{V1}$, ..., $\vec{V6}$) son distintos de cero y todos están separados espacialmente por 60◦ .

Se supuso arbitrariamente que el tiempo del vector de voltaje nulo se divide en partes iguales entre t0 y t7 para el enfoque SVPWM, lo que significa que T0 = T7.

Machine Translated by Google

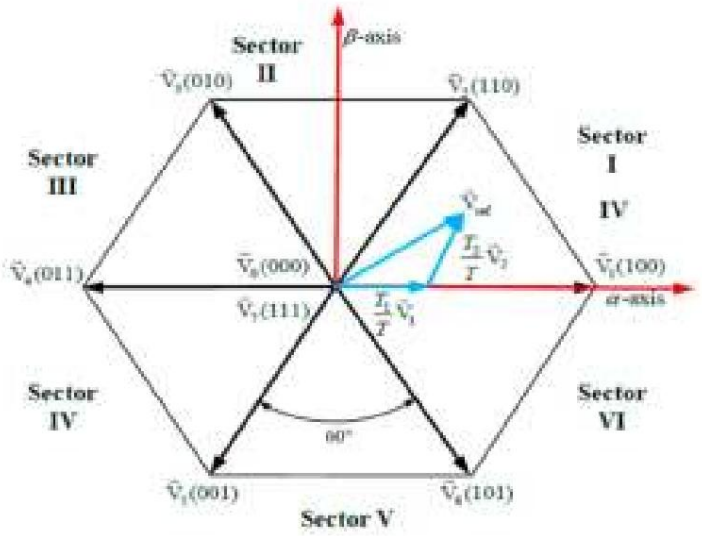


Figura 3. Un diagrama SVPWM típico para un inversor de fuente de voltaje con dos niveles.

Tabla 1. Tensiones de fase, tensiones de salida línea a línea y vectores de conmutación.

Vectores de voltaje	Vectores de conmutación			Tensión línea a neutro			Voltaje línea a línea		
	T1	T2	T3	Furgoneta	vbn	vcn	vab	Vbc	vca
$\vec{V}_0$	0	0	0	0	0	= + + 0	0	0	0
$\vec{V}_1$	1	0	0	$\frac{2}{3}$ VCC	$\frac{-1}{3}$ VCC	$\frac{-1}{3}$ VCC	VCC	0	−VCC
$\vec{V}_2$	1	1	0	$\frac{1}{3}$ VCC	$\frac{1}{3}$ VCC	= + $\frac{-2}{3}$ VCC	0	VCC	−VCC
$\vec{V}_3$	0	1	0	$\frac{-1}{3}$ VCC	$\frac{2}{3}$ VCC	$\frac{-1}{3}$ VCC	−VCC	VCC	0
$\vec{V}_4$	0	1	1	$\frac{-2}{3}$ VCC	$\frac{1}{3}$ VCC	= + 1 $\frac{}{3}$ VCC	−VCC	0	VCC
$\vec{V}_5$	0	0	1	$\frac{-1}{3}$ VCC	$\frac{-1}{3}$ VCC	$\frac{2}{3}$ VCC	0	−VCC	VCC
$\vec{V}_6$	1	0	1	$\frac{1}{3}$ VCC	$\frac{-2}{3}$ VCC	$\frac{1}{3}$ VCC	VCC	−VCC	0
$\vec{V}_7$	1	1	1	0	0	0	0	0	0

Para calcular el intervalo de muestreo se utiliza lo siguiente:

$$Ts = T0 + T1 + T2 \quad (1)$$

o

$$TS = T0 + Tmod \quad (2)$$

con

$$Tmod = T1 + T2 \quad (3)$$

donde Tmod es el tiempo de modulación; T0, T1 y T2 son los intervalos de tiempo; y Ts es el intervalo de muestreo. Al examinar el primer sector, el vector de voltaje de referencia. $\vec{V}_{ref}$ mayo presentarse como vectores de funciones:

$$\vec{v}_{ref} = \frac{T1}{TS}\vec{V}1 + \frac{T2}{TS}\vec{V}2 \quad (4)$$

$$\vec{v}_{ref}\ \begin{matrix}\text{porque } \alpha \\ \text{pecado } \alpha\end{matrix} = \frac{T1}{TS}\vec{V}1\ \begin{matrix}\text{porque (0)} \\ \text{pecado (0)}\end{matrix} + \frac{T2}{TS}\vec{V}2\ \begin{matrix}\text{porque } (\pi/3) \\ \text{pecado } (\pi/3)\end{matrix} \quad (5)$$

Machine Translated by Google

La ecuación (5) se puede dividir en dos partes y expresar de la siguiente manera:

$$TS \vec{}_{vref} \quad \text{porque } \alpha = T1 \vec{} V1 \cos(0) + T2 \vec{} V2 \cos(\pi/3) \tag{6}$$

$$TS \vec{}_{vref} \quad \text{sen } \alpha = T1 \vec{} V1 \operatorname{sen}(0) + T2 \vec{} V2 \operatorname{sen}(\pi/3) \tag{7}$$

Las ecuaciones de intervalo de muestra se obtienen resolviendo T2 usando la ecuación (7) y luego sustituyendo los resultados en la Ecuación (6).

$$T1 = mTs \frac{\text{pecado}\left(\frac{\pi}{3} - \alpha\right)}{\text{pecado}\left(\frac{\pi}{3}\right)} \tag{8}$$

$$T2 = mTs \frac{\text{pecado}(\alpha)}{\text{sen}\left(\frac{\pi}{3}\right)} \tag{9}$$

$$0 \le \le 60° \qquad T0 = TS - T1 - T2 \tag{10}$$

El índice de modulación, $m = \frac{vref}{Vcc}$, se representa como $0 \le \alpha \le 60^{\circ}$, y TS es el medio período para cambiar de operador. Las siguientes fórmulas representan los períodos de permanencia para cada sector:

$$T1 = mTSsin\left(\frac{3\,n\pi}{3} - \alpha\right) \tag{11}$$

$$T2 = mTssin\left(\alpha - \frac{\pi(norte - 1)}{3\,3}\right) \tag{12}$$

La Figura 4 muestra la secuencia de conmutación de salida SVPWM en el primer sector.

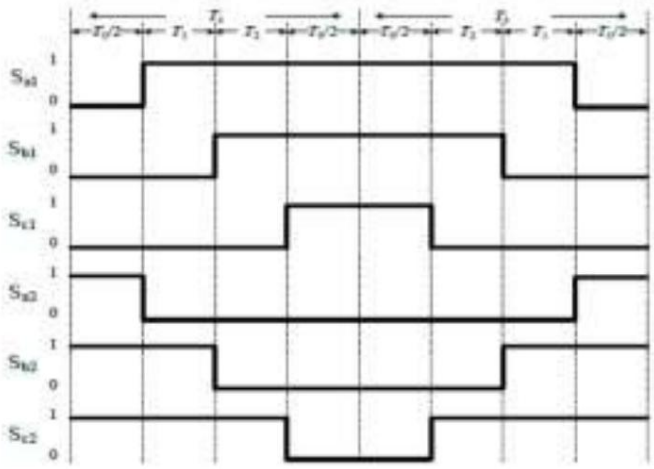


Figura 4. Secuencia de conmutación de salida en el primer sector.

Los períodos de conmutación para los interruptores superior e inferior de cada sector se resumen en la Tabla 2 [13].

Tabla 2. Cálculo del tiempo de conmutación para cada sector.

Sectores	Interruptores superiores (T1,T3,T5)	Interruptores inferiores (T4, S6, S2)
Sector 1	S1 = T1 + T2 + T0/2 S3 = T2 + T0/2 S5 = T0/2	S4 = T0/2 S6 = T1 + T0/2 S2 = T1 + T2 + T0/2
Sector 2	S1 = T1 + T0/2 S3 = T1 + T2 + T0/2 S5 = T0/2	S4 = T2 + T0/2 S6 = T0/2 S2 = T1 + T2 + T0/2

Machine Translated by Google

Tabla 2. Cont.

Sectores	Interruptores superiores (T1,T3,T5)	Interruptores inferiores (T4, S6, S2)
Sector 3	S1 = T0/2 S3 = T1 + T2 + T0/2 S5 = T2 + T0/2	S4 = T1 + T2 + T0/2 S6 = T0/2 S2 = T1 + T0/2
Sector 4	S1 = T0/2 S3 = T1 + T0/2 S5 = T1 + T2 + T0/2	S4 = T1 + T2 + T0/2 S6 = T2 + T0/2 S2 = T0/2
Sector 5	S1 = T2 + T0/2 S3 = T0/2 S5 = T1 + T2 + T0/2	S4 = T1 + T0/2 S6 = T1 + T2 + T0/2 S2 = T0/2
Sector 6	S1 = T1 + T2 + T0/2 S3 = T0/2 S5 = T1 + T0/2	S4 = T0/2 S6 = T1 + T2 + T0/2 S2 = T2 + T0/2

2.2. Metodología del marco de referencia síncrono cero en cuadratura directa

La metodología de marco de referencia síncrono Direct Quadrature Zero (DQZ) es una técnica de control ampliamente utilizada en electrónica de potencia, particularmente en VSI, para lograr el control preciso y eficiente de sistemas de CA trifásicos [14]. Esta metodología se aplica ampliamente en sistemas de energía renovable conectados a la red (como inversores eólicos y solares), variadores de velocidad para motores, filtros de potencia activos para mitigación de armónicos y otras aplicaciones que requieren el control preciso del voltaje de CA y la forma de onda de la corriente . 15].

La metodología del marco de referencia síncrono DQZ implica transformar las señales de CA trifásicas (normalmente voltajes o corrientes) en un marco de referencia giratorio alineado con la frecuencia de CA del sistema. Esta transformación simplifica el control de los VSI al desacoplar los componentes de interés (como potencia activa, potencia reactiva y componentes armónicos) entre sí.

Las señales de CA trifásicas se transforman desde el dominio del tiempo (marco de referencia a, b, c) al marco de referencia giratorio DQZ utilizando transformaciones de Clarke y Park [16].

La transformación de Park rota los componentes αβ a un marco de referencia DQZ estacionario (d, q, 0), donde los componentes d y q representan los componentes directo y en cuadratura de las señales, respectivamente, y 0 representa el componente de secuencia cero [17]. La transformación de Clarke convierte cantidades trifásicas (a, b, c) en cantidades bifásicas (α, β), que representan los componentes simétricos de las señales de CA. En el marco de referencia DQZ, se aplican estrategias de control como controladores integrales proporcionales (PI) o algoritmos más avanzados para regular los componentes d y q de las señales de referencia. Estos controladores ajustan los patrones de conmutación de los dispositivos semiconductores de VSI, típicamente transistores bipolares de puerta aislada (IGBT), para generar la forma de onda de voltaje de salida de CA deseada, asegurando el control preciso de la magnitud del voltaje, la frecuencia y la calidad de la forma de onda [18]. La Figura 5 muestra un fasor espacial con el sistema de referencia αβ0 .

Los valores de corriente trifásicos revisados (ia, ib , ic) se muestran en la Figura 5 como proyecciones sobre el nuevo eje de referencia (iα, iβ). Cuando un sistema trifásico de tres hilos está equilibrado y el eje cero representa el componente de modo común, la corriente total de fase se anula. Por tanto, las variables α y β pueden caracterizar completamente un sistema definido en el sistema de referencia ABC.

Las variables de salida ABC se describen en la Ecuación (13) a continuación:

$$ia = IMcos(\omega t);\ ib = IMcos\ \omega t - \frac{2\pi}{3}\ 2\pi\ ;\ ic = IMcos\ \omega t + \frac{}{3} \quad (13)$$

donde t es la cantidad de tiempo necesaria para alcanzar el estado estable y ω es la frecuencia síncrona constante. La ecuación (14) expresa la transformación αβ0 en forma compleja, mientras que la ecuación (15) la expresa en forma matricial.

Machine Translated by Google

$$\vec{Es} = i\alpha + j i\beta = \frac{2}{3}\left(ia + ib\, e^{\frac{j2\pi}{3}} + hielo\, 3^{\frac{-j2\pi}{3}}\right) \quad (14)$$

$$\begin{matrix} i\alpha \\ i\beta \\ i0 \end{matrix} = \frac{2}{3}\begin{matrix} 1 & \frac{-1}{2} & \frac{-1}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{-\sqrt{3}}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{matrix}\begin{matrix} lowa \\ ib \\ ic \end{matrix} \quad (15)$$

Las corrientes del eje de referencia son iα , iβ e i0, mientras que las corrientes trifásicas son iα, ib e ic.

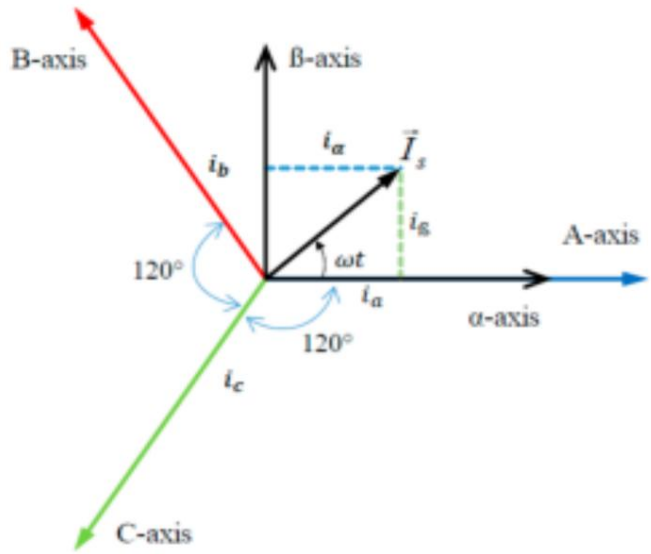


Figura 5. Fasor espacial y marco de referencia αβ0 .

Usando la transformación de Park con un desplazamiento de −90◦ , el marco gira alrededor del eje 0 a la misma frecuencia que las sinusoides que definen los fasores, mientras el sistema está en el marco αβ0 . Como se puede observar en la Figura 6, el componente se transforma al eje Q, el cual se posiciona 90◦ en un ángulo de cuadratura con respecto a la componente directa, y la componente −β se transforma α β en el eje d, que está alineado con el vector de rotación.

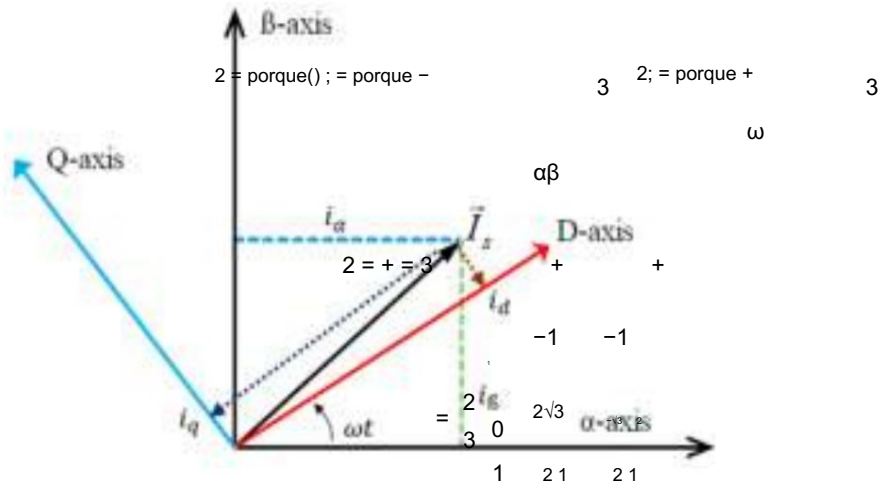


Figura 6. Fasor espacial y marco de referencia DQ0.

Las ecuaciones (16) y (17) expresan la transformación DQ0 en representaciones complejas y matriciales , respectivamente.

$$\vec{I}_s \quad ia + jib\, e \quad = id + jiq \quad = 0 \quad \alpha\beta \quad (16)$$

$$= \quad 3 \quad 2 \quad \sin(\quad)\sin(\quad + \quad 32(\quad -2\cos(\quad)\cos(\omega t - \quad -\beta \quad 2 \quad)\text{porque}(\quad + 3 \quad)$$

Machine Translated by Google

$$\begin{bmatrix} \\ \\ \end{bmatrix} = \begin{matrix} \text{pecado}(\omega t) & -\cos(\omega t) & 0 \\ \cos(\omega t) & \text{sen}(\omega t) & 0 \\ 0 & 0 & 1 \end{matrix} \begin{matrix} i\alpha \\ i\beta \\ i0 \end{matrix} \quad (17)$$

En consecuencia, la rotación genera valores DC a partir de las señales periódicas. Dado que el bastidor de hilado sigue al eje a 90 grados, en t = 0, los ejes Q y A coinciden. Cuando Is está alineado con precisión con el ángulo de referencia ωt, id = 1, iq = 0 e i0 = 0 son los componentes DQ0 . Las transformaciones completas de DQ0 a ABC y sus correspondientes inversas se presentan a continuación:

$$\begin{bmatrix} \\ \\ i0 \end{bmatrix} = \begin{matrix} \text{pecado}(\omega t) & \text{pecado}(\omega t - \frac{2\pi}{3}) & \text{pecado}(\omega t + \frac{2\pi}{3}) \\ \cos(\omega t) & \cos(\omega t - \frac{2\pi}{3}) & \cos(\omega t + \frac{2\pi}{3}) \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{matrix} \begin{matrix} \text{lowa} \\ ib \\ ic \end{matrix} \quad (18)$$

$$\begin{matrix} \text{lowa} \\ ib \\ ic \end{matrix} = \begin{matrix} \text{pecado}(\omega t) & \cos(\omega t) & 1 \\ \text{pecado}(\omega t - \frac{2\pi}{3}) & \cos(\omega t - \frac{2\pi}{3}) & 1 \\ \text{pecado}(\omega t + \frac{2\pi}{3}) & \cos(\omega t + \frac{2\pi}{3}) & 1 \end{matrix} \begin{bmatrix} \\ \\ i0 \end{bmatrix} \quad (19)$$

3. Diseño del bucle de control de corriente en el dominio S El diseño

del bucle de control de corriente en el dominio S para un VSI trifásico de dos niveles es un aspecto crucial para garantizar el funcionamiento preciso y estable del inversor. Esta metodología de diseño aprovecha la teoría de control clásica en el dominio S para lograr la regulación efectiva de las corrientes de salida del inversor. Según Wencong et al., este bucle de control modifica la tensión de salida del inversor para permitir la inyección de la corriente necesaria en la red pública [19]. El VSI trifásico generalmente se modela utilizando un conjunto de ecuaciones diferenciales que describen la relación entre el voltaje CC de entrada, los estados de conmutación del inversor y las corrientes CA de salida. Estas ecuaciones se transforman en el marco de referencia síncrono (marco dq) para desacoplar las corrientes alternas trifásicas en dos componentes ortogonales (eje d y eje q), simplificando el diseño de control. La dinámica del sistema está representada por funciones de transferencia en el dominio S, que relacionan las entradas de control (tensiones de referencia) con las salidas (corrientes reales). Para una VSI, la función de transferencia generalmente incluye la dinámica del filtro LC del inversor y la influencia de la impedancia de carga. La Figura 7 ilustra la representación esquemática de un VSC trifásico de dos niveles con un filtro L. Rf y Lf representan la resistencia y la inductancia, respectivamente, entre los interruptores del convertidor y el

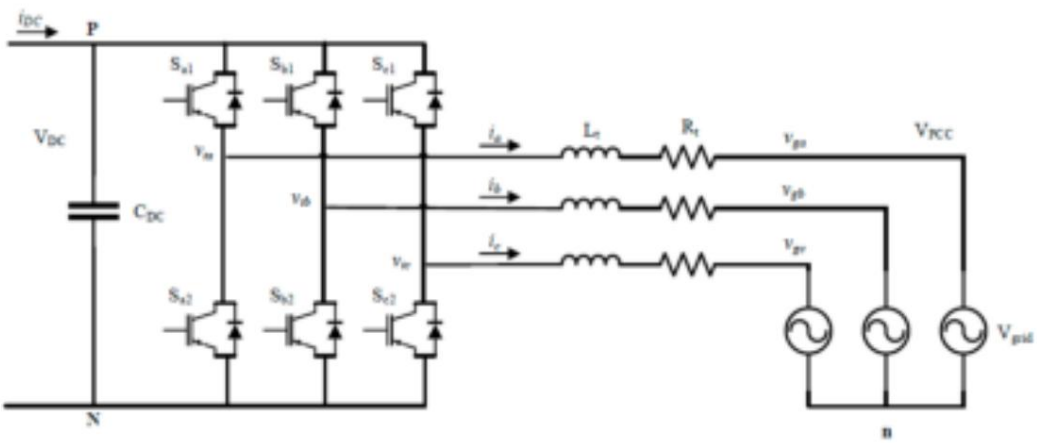
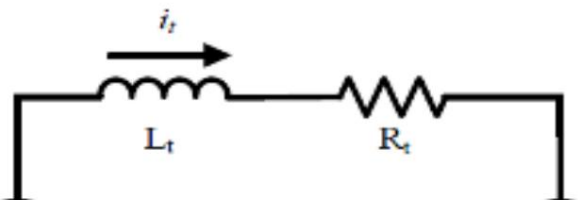


Figura 7. Diagrama de circuito VSC de dos niveles con un filtro L incluido.

Se puede producir un circuito por fase equivalente reducido aplicando la ley de voltaje de Kirchhoff al circuito que se muestra arriba, como se ilustra en la Figura 8.



Machine Translated by Google

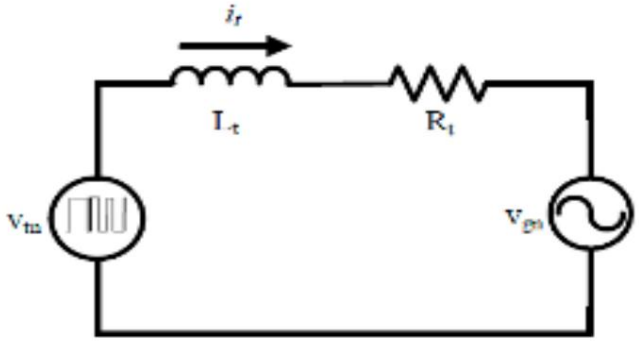


Figura 8. Un filtro L en un circuito por fase VSC simplificado.

El voltaje de la red, vgn, se muestra en la Figura 8 como un componente relativamente fijo que puede considerarse disruptivo. Como resultado, la corriente, it, se puede regular alterando el voltaje de salida del inversor vtn. En consecuencia, el VSC funciona como una fuente de energía que inyecta electricidad a la red. La ley de voltaje de Kirchhoff, cuando se aplica a los tres circuitos por fase, produce lo siguiente:

$$\begin{aligned} vtan &= Lf \frac{dia}{dt} + Rf\, ia + vgan \\ vtbn &= Lf \frac{dib}{dt} + Rf\, ib + vgbn \\ vtcn &= Lf \frac{dic}{dt} + Rf\, ic + vgcn \end{aligned} \tag{20}$$

variables vtan, vtbn y vtcn representan los tres voltajes de salida del inversor por fase; Lf y Rf denotan la inductancia y resistencia entre los interruptores del convertidor y el punto de acoplamiento común (PCC); vgan, vgbn y vgcn representan los tres voltajes de red por fase; e ia, ib e ic representan las tres corrientes del inversor por fase. Consolidando las tres ecuaciones de (20) en una sola ecuación, obtenemos lo siguiente:

$$vt,abc = Lf \frac{diabc}{dt} + Rf\, iabc + vg,abc \tag{21}$$

La ecuación (15) se utiliza para traducir la ecuación (21) al sistema de referencia αβ0, lo que da como resultado

$$vt,\alpha\beta = Lf \frac{di\alpha\beta}{dt} + Rf\, i\alpha\beta + vg,\alpha\beta \tag{22}$$

Usando el marco de referencia dq0 en (16) y repitiendo el proceso para la ecuación (22), el resultado es

$$\begin{aligned} vt,dq &= vt,\alpha\beta e^{-j(\omega t - \frac{\pi}{2})} \\ vt,dq &= Lf \frac{didq}{dt} + Rf\, idq + vg,dq \end{aligned} \tag{23}$$

El marco a, b, c se puede dividir en dos componentes, d y q, como se indicó en la Sección 2.1.

$$Lf \frac{hizo}{dt}d = vtd + Lf\,\omega iq - Rf\, id - vgd \tag{24}$$

$$Lf \frac{diq}{dt} = vtq + Lf\,\omega iq - Rf\, iq - vgq \tag{25}$$

con $vtd = \frac{V}{\sqrt{DC}}\, 3\, md$ y $vtq = \frac{V}{\sqrt{DC}}\, 3\, mq$, donde mdq es el índice de modulación, y $\frac{V}{\sqrt{CC}\, 3}$ es el Convertidor de ganancia SVPWM.

Las dinámicas id e iq están relacionadas debido al término Lf ω que se encuentra en las ecuaciones (24) y (25). Para desacoplar los subsistemas d y q y tener en cuenta las entradas de interrupción de la red vgd y vgq, el índice de modulación para cada uno debe establecerse de la siguiente manera:

$$MD = \frac{\sqrt{3}}{VCC}\, ud - Lf\,\omega iq + vgd \tag{26}$$

Machine Translated by Google

$$\text{metros cuadrados}\frac{\sqrt{3}}{VCC}\ uq - Lf\ \omega id + vgq \tag{27}$$

Para el subsistema d , el término anticipatorio es −Lf ωiq + vgd, mientras que para el subsistema q , es −Lf ωid + vgq. La ecuación (26) se puede sustituir en (24), obteniendo

$$Lf = \frac{\text{hizo}}{dt} Rf \text{ identificación} \tag{28}$$

La ecuación (27) se puede sustituir en (25) de la misma manera para obtener

$$Lf = \frac{diq}{uq} - Rf\ iq\ dt \tag{29}$$

con uq como salida del compensador PI.

Los sistemas de control de los bucles de control de corriente son los mismos en los ejes d y q , como se ve en (28) y (29). Por tanto, los compensadores vinculados pueden ser los mismos. Como resultado, sólo se determinarán las ganancias actuales del compensador de control en el eje d. Cuando se transforma la Ecuación (29) utilizando el método de Laplace, se obtienen los siguientes resultados:

$$Lf\ sId(s) = ud(s) - Rf\ Id(s)\ ud\ (s)$$
$$\text{Identificación(es)} = \frac{}{Lf + Rf} \tag{30}$$

Por tanto, la planta de control se puede expresar de la siguiente manera:

$$\text{Médico(s)} = \frac{1}{Lf\ s + Rf} \tag{31}$$

La siguiente Ecuación (32) indica que el compensador es un controlador PI:

$$Gc(s) = \frac{(\ ) = kps + ki}{s} \tag{32}$$

La Figura 9 ilustra la función de transferencia del circuito cerrado en el eje d:

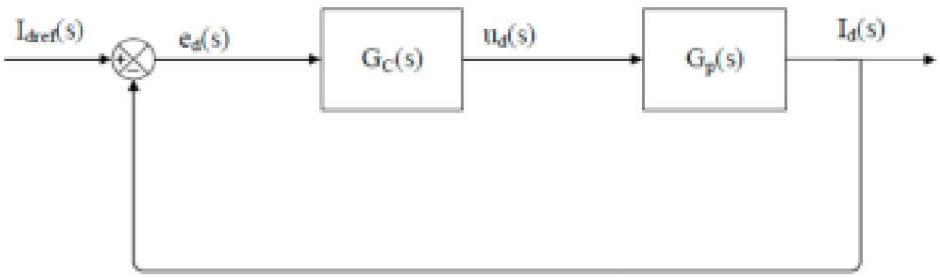


Figura 9. Diagrama de bloques del sistema VSC controlado por corriente en forma simplificada.

La ecuación (33) proporciona la siguiente ganancia en bucle abierto:

$$Gol(s) = GC(s)GP(s)\ (\ ) = \tag{33}$$

Reemplazando las Ecuaciones de la planta compensadora y de control por sus términos, la Ecuación (33), quedará:

$$\text{Gol(es)} = + (\ ) = \frac{kps + ki}{s} \times \frac{1\ 1}{Lf\ s + Rf} = \frac{kps + ki + 2}{si + Rf\ s} = \frac{kp + \frac{ki\ s}{si}}{1 + \frac{kp}{RF\ s + si}} \tag{34}$$

Una vez que los términos $\frac{ki}{si} = \frac{Rf}{}$ se ecualizan aplicando el enfoque de cancelación de polos, el kp ganancia en bucle abierto, la ecuación (34) se reduce de la función de transferencia de segundo a primer orden de la siguiente manera:

$$(\ ) =$$

Machine Translated by Google

$$Gol(es) = \frac{kp}{si} \quad (35)$$

La ecuación (36) representa la ganancia de control actual en el circuito cerrado en el eje d:

$$Gi(s) = Gcl(s) = \frac{Gol(es)}{1 + gol(es)} \quad (36)$$

La ecuación (35) se puede sustituir en (36) para obtener la siguiente expresión para la ganancia de mirada cerrada:

$$\frac{\text{Identificación(es) kp}}{\text{Idre f(s) kp + Lf s}} = Gi(s) = \quad (37)$$

Reemplazando $kp = \frac{si}{\tau i}$ y $ki = \frac{RF}{\tau i}$, la ganancia de mirada cerrada se convertirá

$$Gi(s) = \tau i \frac{1}{s + 1} \quad (38)$$

El tiempo constante de control de circuito cerrado resultante se escribe como τi . Por lo tanto, lo siguiente La fórmula proporciona el tiempo de estabilización de primer orden en el sistema de circuito cerrado:

$$\text{sedimentación} = 4{,}6\tau i \text{ ti,} \quad (39)$$

4. Diseño del lazo de control para el voltaje del bus de CC en el dominio S

Para mantener un flujo de energía equilibrado, la técnica de control de voltaje del bus de CC es esencial. La implementación de este control es imperativa ya que permite que el convertidor de fuente de voltaje funcione como un rectificador, convirtiendo la energía de CA en energía de CC y facilitando el control del voltaje del capacitor del enlace de CC para suministrar una carga de salida. Además, para fuentes de energía CC intermitentes, como sistemas fotovoltaicos, esta técnica de control ayuda a ajustar la referencia del bucle de control actual para regular la entrada de energía CC de manera efectiva [20]. El circuito de la Figura 7 se muestra en la Figura 10, con una resistencia de purga RB y señales de retroalimentación adicionales (iDC e iL) agregadas para simplificar el modelado de control de CC.

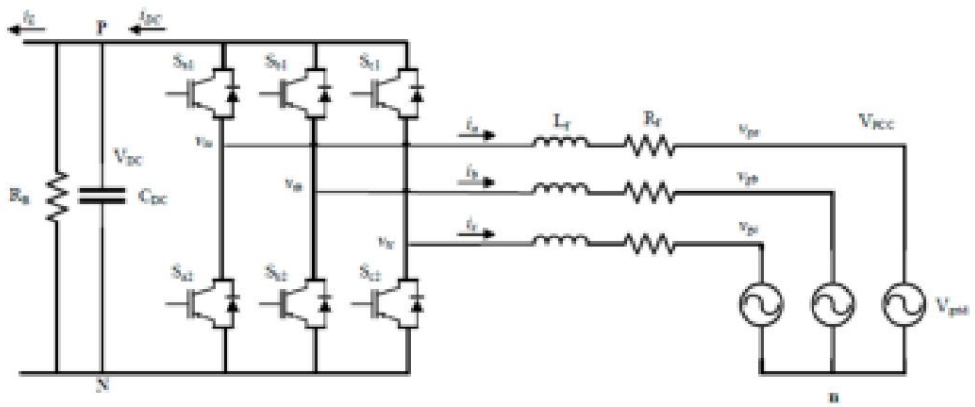
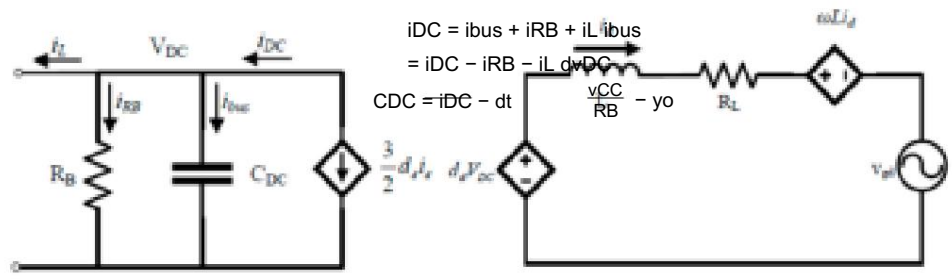


Figura 10. VSC de dos niveles con un filtro L y una resistencia de purga del circuito del enlace de CC.

La resistencia y la inductancia entre el PCC y la salida del inversor son Rf y Lf . Dado que la transferencia de potencia activa permite cargar y descargar el condensador del enlace de CC, el circuito que se muestra en la configuración del marco de referencia síncrono DQ solo puede modelarse para el componente del eje d. Como resultado, no se necesita una corriente reactiva instantánea, ni el componente del eje q para este procedimiento, como muestra la Figura 11 .

Las siguientes ecuaciones ilustran la ecuación dinámica de voltaje del capacitor del enlace de CC:

$$\begin{aligned} iDC &= ibus + iRB + iL \\ ibus &= iDC - iRB - iL \\ CDC \frac{dvDC}{dt} &= iDC - \frac{vCC}{RB} - yo \end{aligned} \quad (40)$$



Machine Translated by Google

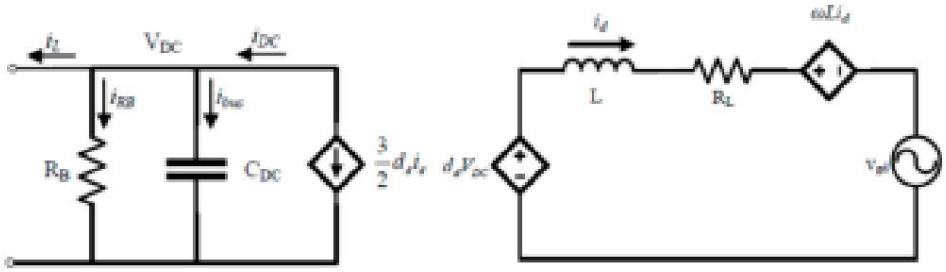


Figura 11. Configuración del componente VSC del eje d con una resistencia de purga y un filtro L.

La ecuación (40) se puede transformar usando la transformada de Laplace para obtener

$$VCC(es) = \frac{1\ CDC + \frac{}{RB}}{CDC + \frac{1\ sCDCRB + 1}{RB}} VDC(s) = IDC(S) - IL(s) = + + ud(s) = \frac{RB}{} ud(es) \quad (41)$$

La fórmula para calcular la salida del compensador PI, ud(s), es ud(s) = IDC(s) − IL(s), donde IL(s) es el término de retroalimentación. Esto permite que el controlador evite inventar un valor medido. Considerando esto, el sistema de control se escribe de la siguiente manera:

$$+ PG = \frac{RB}{( ) = ( ) - ( ) sCDCRB + 1} \quad (42)$$

La relación entre estas corrientes se representa de la siguiente manera, suponiendo un hecho de que el iDC solo puede controlarse directamente con la corriente extraída de la red:

$$+ P3 = PDC \quad (43)$$

La potencia activa viene dada como

$$P3 = \frac{3}{2} vgdid + vgqiq \quad (44)$$

La potencia del bus de CC se obtiene reemplazando la ecuación (44) por (43), lo que produce lo siguiente:

$$\frac{3}{2} vgdid = -iDCVDC \quad (45)$$

La siguiente ecuación describe la relación entre la constante de tiempo τDC del bucle de control para la tensión del bus de CC y la constante de tiempo τi del bucle de control de corriente:

$$\tau i \quad \tau CC \quad (46)$$

La relación entre estos dos bucles de control puede expresarse mediante la siguiente ecuación (47) si se cumple la condición de la ecuación (46), id = Idre f , y los dos bucles de control están desacoplados.

$$Idre\ f = \frac{2}{3} \frac{VCC}{vgd} IDC \quad (47)$$

El coeficiente de ganancia interna se obtiene de la relación entre Idre f e IDC y se expresa de la siguiente manera:

$$Kinner = \frac{IDC}{Idre\ f} = \frac{3}{2} \frac{vgd}{VCC} \quad (48)$$

Con base en lo siguiente, se presume que el compensador es un controlador PI expresado como

$$CG(es) = \frac{kps + ki}{s} \quad (49)$$

El método seleccionado para el control de voltaje de CC adopta un enfoque en cascada, que incluye un bucle externo para el control de voltaje del bus de CC y un bucle interno para el control de corriente. Esta táctica garantiza que el tiempo de reacción del controlador actual sea más rápido que el de

Machine Translated by Google

el controlador de la tensión CC de acuerdo con la Ecuación (46). El bucle de tensión CC considera el bucle de corriente como una ganancia constante, lo que permite un cambio gradual en la referencia de corriente y, por tanto, una respuesta libre de errores del controlador de corriente. En consecuencia, la dinámica (+) ( ) = de los dos circuitos representados en la Figura 11 está efectivamente desconectada. La Figura 12 × + 1 muestra un diagrama de tensión del enlace de CC de un sistema VSC en forma simplificada.

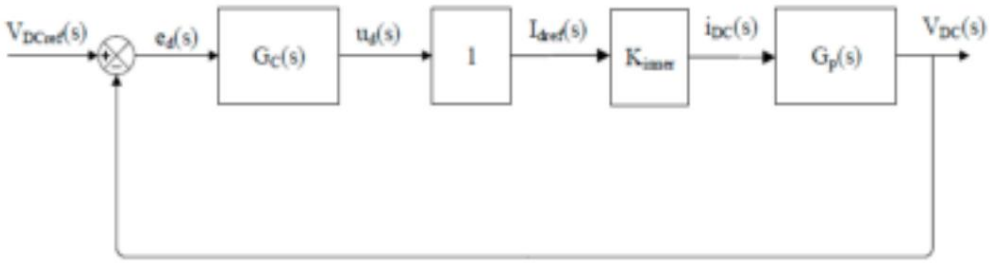


Figura 12. Modelo de bucle de tensión del enlace CC simplificado del sistema VSC.

La ecuación (50) proporciona la ganancia en el bucle abierto según la Figura 12.

$$Gol(s) = GC(S)GP(S)Kinner \tag{50}$$

Las ecuaciones (42) y (49) se pueden sustituir en (50) para obtener lo siguiente:

$$Gol(es) = \frac{kps + ki}{s} \times \frac{RB}{2CDCRB + s} \times Kinner = \frac{KinnerRB\ kps + ki}{sCDCRB + 1} \tag{51}$$

Con respecto al control de corriente en el eje d, la ganancia de bucle cerrado se expresa en Ecuación (52):

$$GDC(S) = Gcl(s) = \frac{Gol(es)}{1 + gol(es)} \tag{52}$$

La ecuación (51) sustituida en (52) dará como resultado

$$CDG(es) = \frac{Kinner\ kps + ki\ /CDC}{S^2 + \frac{KinnerRBkp+1}{CDCRB} + \frac{Kinnerki}{\ldots}} \tag{53}$$

Viendo que la Ecuación (53) tiene un numerador de primer orden, se establecerá un prefiltro para eliminar este término, permitiendo alinear (53) con la función de transferencia estándar de segundo orden, que es:

$$GPF(s) = \frac{GDC,\ deseado(s)}{GDC,antiguo(s)} = \frac{ki}{kps + ki} \tag{54}$$

La ganancia en lazo cerrado del control de corriente en el eje d se puede obtener multiplicando las Ecuaciones (53) y (54). El resultado es el siguiente:

$$CDG(es) = \frac{\frac{Kinnerki}{\ldots}}{S^2 + \frac{KinnerRBkp+1}{CDCRB} S + \frac{Kinnerki}{\ldots}} \tag{55}$$

La ganancia en lazo cerrado del control de corriente en el eje d se puede obtener multiplicando las Ecuaciones (53) y (54). El resultado es el siguiente:

$$KinnerRB = 4{,}6 \times \frac{2\xi\omega nCDCRB - 1\ kp}{} \tag{56}$$

$$ki = \frac{\omega n\ 2CDC}{Kinner} \tag{57}$$

Machine Translated by Google

Por lo tanto, el siguiente es el segundo tiempo de establecimiento del sistema de circuito cerrado:

$$\tau DC = \frac{1}{\xi \omega n}$$
$$tDC, \text{ ajuste} = 4{,}6 \times \tau DC \quad (58)$$

5. Diseño del lazo de control para el voltaje CA en el dominio S

Cuando una falla o un corte de energía impide que la red principal suministre energía, un VSC El sistema en modo de formación de cuadrícula crea una referencia de cuadrícula y facilita la transferencia de datos activos. o potencias reactivas durante el funcionamiento de la microrred en modo aislado. Esto permite que la microrred siga funcionando [21]. El VSI trifásico se modela mediante ecuaciones diferenciales. que describen la relación entre el voltaje CC de entrada, los estados de conmutación del inversor y los voltajes de CA de salida. La dinámica del sistema VSI, incluido el LC.
El filtro (que generalmente se usa para suavizar la salida del inversor), como se presenta en la Figura 13, son representado en el dominio S mediante funciones de transferencia [22].

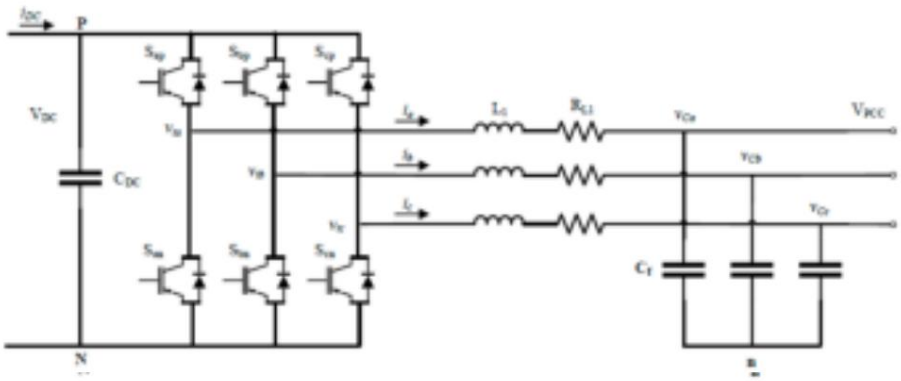


Figura 13. VSC de dos niveles con filtro LC.

Las resistencias entre el convertidor y el condensador del filtro se presentan como RL1 y L1. Un circuito equivalente condensado por fase producido usando la ley de voltaje de Kirchhoff se muestra en la Figura 14.

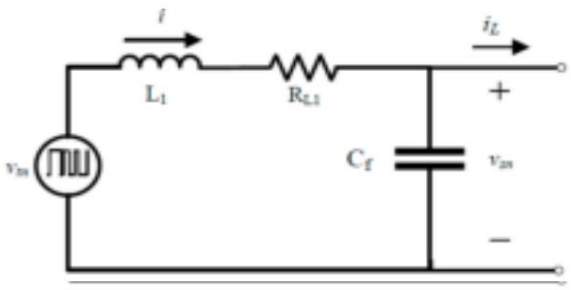


Figura 14. Un VSC con un circuito por fase de filtro LC de forma simplificada.

Las corrientes de fase para cada uno de los tres circuitos por fase se pueden obtener usando la actual ley de Kirchhoff de la siguiente manera:

$$ia = iC\,fa + iLa$$
$$ib = iC\,fb + iLb$$
$$ic = iC\,fc + iLc \quad (59)$$

Combinando las tres ecuaciones de (59) se obtiene lo siguiente:

$$iabc = iC\,f,abc + iL,abc \quad (60)$$

Utilizando la transformada ABC-αβ0 en la Ecuación (60), se obtiene el siguiente resultado:

$$iC\,f,\alpha\beta = \alpha\beta + iL,\alpha\beta \quad (61)$$

Machine Translated by Google

Usando la ecuación (16), la transformada αβ0-dq0 se puede convertir a

$$iC\,f\,,dq = iC\,f\,,\alpha\beta e^{-j(\omega t-\frac{\pi}{2})} \tag{62}$$

$$iC,dq = idq - iL,dq \tag{63}$$

En consecuencia, como se describe en la Sección 2.1, el marco de referencia a, b, c puede dividirse en componentes d y q y conectarse en cascada a través de los bucles de control de corriente en los ejes d y q , lo que da como resultado

$$cf\ \frac{DVD}{} = Cf\,\omega vsq + id - iLd\,dt \tag{64}$$

$$cf\ \frac{dvsq}{dt} = -Cf\,\omega vsd + iq - iLq \tag{65}$$

La existencia de los términos Cf ω en las Ecuaciones (64) y (65) hace que se acoplen las dinámicas vsd y vsq . Los términos de avance −Cf ωvsq + iLd para el componente del eje d y Cf ωvsd + iLq para el componente del eje q se introducen para desacoplarlos y tener en cuenta las entradas de perturbación de carga iLd e iLq. Además, las unidades de control de los componentes del eje d y del eje q son las mismas, como se muestra en las Ecuaciones (64) y (65); de este modo,

$$\frac{dvsd}{dt} = \frac{1}{cf}[ud\,] \tag{66}$$

ud es la salida del compensador PI. La ecuación (66) se resuelve aplicando la fórmula de Laplace. transformar, que produce lo siguiente:

$$Vsd(s) = \frac{\underline{ud(es)}}{sCf} \tag{67}$$

Como resultado, la planta de control es la siguiente:

$$GP(es) \frac{1}{} = sCf \tag{68}$$

Además, se supone que la constante de tiempo del bucle de control de la tensión CA, τv, será cinco veces mayor que la del bucle de control de corriente, τi . Si se cumple esta suposición, los dos bucles de control se desacoplarán. La circunstancia es la siguiente:

$$\tau i \quad \tau v \tag{69}$$

Por lo tanto, las entradas de los bucles de control de corriente se presentan en las Ecuaciones (70) y (71), e Idq(s) = Idqre f(s) para el bucle de control de tensión CA.

$$Idre\,f(s) = ud(s) - Cf\,\omega Vsq(s) + ILd(s) \tag{70}$$

$$Iqre\,f(s) = uq(s) + Cf\,\omega Vsd(s) + ILq(s) \tag{71}$$

La ecuación (72) proporciona la ganancia del controlador PI.

$$CG(es) = \frac{kps + ki}{s} \tag{72}$$

Dado que las plantas de control de la ecuación (68) para los componentes d y q son las mismas, los compensadores correspondientes también pueden ser los mismos. Con referencia a la Figura 15, la única información que es necesario recopilar a continuación son las ganancias de los compensadores de control de voltaje en el eje d. La siguiente es la ganancia en bucle abierto:

$$Gol(s) = GC(s)GP(s) \tag{73}$$

Machine Translated by Google

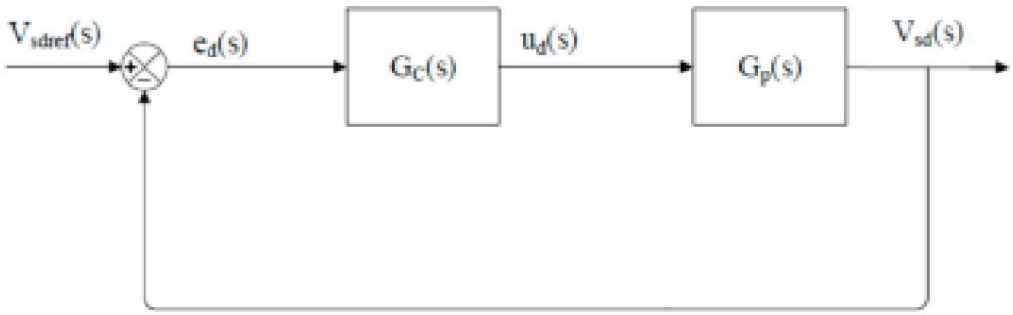


Figura 15. Un diagrama de bloques para el bucle de voltaje de CA del sistema VSC en forma simplificada.

Los GC(s) y GP(s) se pueden sustituir por sus valores correspondientes para obtener el siguiente resultado:

$$Gol(es) = \frac{kps + ki}{s} \times \frac{1}{scf,} = \frac{kps + ki\,s}{2Cf} \quad (74)$$

La ganancia del control actual en el eje d en circuito cerrado se puede representar de la siguiente manera:

$$Gv(s) = Gcl(s) = \frac{Gol(es)}{1 + gol(es)} \quad (75)$$

Cuando la ecuación (74) se sustituye en (75), esto da como resultado

$$Gv(s) = \frac{\frac{(kps+ki)}{cf}}{s^2 + 2\frac{kp}{cf}Kis + \frac{}{cf}} \quad (76)$$

El diagrama de bloques del bucle de voltaje CA del sistema VSC en forma simplificada se muestra en la Figura 15.

Dado que el numerador de primer orden de la ecuación (76) debe cancelarse, es necesario desarrollar un prefiltro para permitir que la ecuación (76) corresponda a la función de transferencia convencional de segundo orden. Como se indica a continuación, se requiere un prefiltro.

$$GPF(s) = \frac{GDC,\ deseado(s)}{GDC,\ antiguo(s)} = \frac{ki}{kps + ki} \quad (77)$$

El siguiente es el resultado de multiplicar las ecuaciones (76) y (77):

$$Gv(s) = \frac{ki}{s\,2Cf + kps + ki} \quad (78)$$

Al resolver los resultados de los coeficientes del denominador de la ecuación (78), las ganancias de los controladores proporcional e integral se resuelven para

$$kp = 2\xi\omega nCf \quad (79)$$

$$ki = \omega n2Cf \quad (80)$$

Como resultado, la constante de tiempo del bucle de control del voltaje de CA se proporciona como

$$\tau v = \frac{1}{\xi\omega n} \quad (81)$$

Para desacoplar los dos bucles de control, se considera que el tiempo de establecimiento del sistema para este segundo orden en el circuito cerrado es cinco veces mayor que el tiempo de establecimiento del bucle de control existente. En otras palabras,

Machine Translated by Google

$$t_{v,\text{asentándose}} = \frac{4.6}{\xi\omega_n} = 4{,}6 \times 5 \times \tau_i = 5 \times t_{i,\text{asentándose}} \quad (82)$$

6. Discusión

Esta sección de Discusión explora los aspectos clave del modelo de circuito básico de VSC. Las diferentes metodologías y estrategias de control empleadas en los VSC, incluida la metodología SVPWM, la metodología del marco de referencia síncrono DQZ, el diseño de la corriente, el voltaje del bus de CC y los bucles de control de voltaje de CA en el dominio S, se presentan en la Tabla 3. .

Tabla 3. Resumen del modelo de circuito básico de convertidores de fuente de voltaje.

Metodología SVPWM	SVPWM es una técnica ampliamente utilizada para controlar VSC porque puede producir excelentes formas de onda de salida con menos distorsión armónica. Al sintetizar eficazmente el vector de voltaje de salida deseado, SVPWM permite el control preciso de la salida VSC. Sin embargo, los desafíos, como la complejidad computacional y las limitaciones de la frecuencia de conmutación, deben abordarse cuidadosamente para garantizar un rendimiento óptimo.
Metodología del marco de referencia síncrono DQZ	La metodología del marco de referencia síncrono DQ0 proporciona una poderosa herramienta para analizar y controlar VSC, particularmente en aplicaciones conectadas a la red. Al transformar el sistema trifásico en un espacio bidimensional con ejes directos (d) y en cuadratura (q) , esta metodología simplifica los procesos de control y modelado. Permite la representación precisa de la dinámica del sistema y facilita el diseño de algoritmos de control, mejorando el rendimiento general y la estabilidad de los sistemas basados en VSC.
Diseño del bucle de control de la corriente en el dominio S.	El control de corriente eficaz es crucial para garantizar una respuesta precisa y dinámica en los VSC, especialmente en aplicaciones que requieren una regulación precisa del flujo de corriente. Al diseñar bucles de control actuales en el dominio S, los ingenieros pueden aprovechar técnicas de control avanzadas para lograr los objetivos de rendimiento deseados, como una respuesta rápida, un bajo sobreimpulso y una estabilidad sólida. La implementación de algoritmos de control sofisticados, como controladores integrales proporcionales o control predictivo de modelos (MPC), puede mejorar aún más la eficiencia de los bucles de control actuales en los VSC.
Diseño del lazo de control para el Tensión del bus CC en el dominio S	El bucle de control para el voltaje del bus de CC desempeña un papel fundamental a la hora de mantener un funcionamiento estable y garantizar un rendimiento fiable de los VSC. Al regular el voltaje del bus de CC, este bucle de control permite una conversión de energía eficiente y una transferencia de energía fluida entre el VSC y la red o carga. El diseño del bucle de control de voltaje del bus de CC en el dominio S permite el ajuste preciso de los parámetros del controlador y las características de respuesta dinámica, lo que garantiza un rendimiento óptimo en diferentes condiciones operativas y perfiles de carga.
Diseño del lazo de control para el Tensión alterna en el dominio S	En aplicaciones VSC conectadas a la red, el bucle de control del voltaje de CA es esencial para regular el voltaje de salida para cumplir con los requisitos de la red, como la magnitud y la frecuencia del voltaje. Al diseñar el circuito de control para el voltaje de CA en el dominio S, los ingenieros pueden desarrollar estrategias de control sólidas para mantener la estabilidad de la red, mitigar los armónicos de voltaje y garantizar la integración perfecta de las fuentes de energía renovables. Se pueden emplear técnicas de control avanzadas, como el control predictivo y el control adaptativo, para mejorar la eficiencia y la flexibilidad del circuito de control de voltaje de CA en los VSC.

7. Conclusiones

En este artículo se presenta el modelo básico de circuito VSC con un filtro LC de salida. La principal interfaz de energía entre los generadores distribuidos y la microrred es el convertidor AC-DC. Se han observado dos formas de operación para un convertidor de fuente de voltaje. Actúa como un rectificador activo cuando se utiliza como control de bus de CC. Utiliza una variedad de bucles de control para funcionar como inversor cuando se encuentra en los modos de alimentación y formación de red. También se ha presentado un manual completo que cubre una variedad de técnicas, incluido el SVPWM, el marco síncrono DQZ y el bucle de bloqueo de fase. Además, este artículo sirve como una guía completa para comprender el modelo de circuito básico de los VSC, las metodologías para su control y las técnicas para su modelado preciso. Al proporcionar información sobre los principios fundamentales y las consideraciones prácticas asociadas con los VSC, este documento tiene como objetivo contribuir al avance de las aplicaciones de electrónica de potencia basadas en VSC y su integración en los sistemas de energía modernos para una mayor eficiencia.

Machine Translated by Google

confiabilidad y sostenibilidad. Es necesario realizar simulaciones y estudios de casos para ilustrar la eficacia y precisión de las metodologías y estrategias de control presentadas en el análisis de aplicaciones de electrónica de potencia basadas en VSC.

Contribuciones de los autores: Conceptualización; metodología; análisis formal; redacción: preparación del borrador original, CBN; redacción: revisión y edición; supervisión.; administración de proyectos, MA y AA

Todos los autores han leído y aceptado la versión publicada del manuscrito.

Financiamiento: Esta investigación fue financiada por la subvención número PMDS22052313773 de la Fundación Nacional de Investigación (NRF) y la APC fue financiada por la Universidad Tecnológica de Península del Cabo.

Declaración de la Junta de Revisión Institucional: Esta investigación no involucra sujetos humanos o animales; por lo tanto, no fue necesario el consentimiento de los participantes para recopilar datos para este proyecto.

Declaración de disponibilidad de datos: Las solicitudes de acceso a los conjuntos de datos se pueden realizar al autor correspondiente.

Agradecimientos: Como autor correspondiente, reconozco que este artículo se completó como parte de una investigación de estudio de doctorado bajo la supervisión y el apoyo de Marco Adonis, y que Ali Almaktoof proporcionó dirección, edición y revisión del manuscrito antes de su publicación.

Conflictos de intereses: No se revelan conflictos de intereses relacionados con esta publicación.

Referencias

1. Jing, T.; Maklakov, AS Revisión de convertidores de fuentes de voltaje para aplicaciones de energía. En Actas de la Conferencia Internacional de los Urales sobre Energía Verde de 2018 (UralCon), Chelyabinsk, Federación de Rusia, 4 a 6 de octubre de 2018; págs. 275–281.
2. Brodzicki, M.; Klucznik, J.; Czapp, S. Evaluación del impacto de VSC en el sistema eléctrico utilizando una curva de capacidad PQ adecuada. Electrónica 2023, 12, 2462. [CrossRef]
3. Ndeke, CB; Adonis, M.; Almaktoof, A. Estrategia de gestión energética para un sistema de microrred híbrido que utiliza energía renovable. Descubrimiento. Energía 2024, 4, 1. [CrossRef]
4. Monteiro, V.; Monteiro, LF; Franco, Florida; Mandrioli, R.; Ricco, M.; Grandi, G.; Afonso, JL El papel de los convertidores CA/CC frontales en hogares inteligentes híbridos CA/CC: análisis y validación experimental. Electrónica 2021, 10, 2601. [CrossRef]
5. Xiong, L.; Liu, X.; Liu, Y.; Zhuo, F. Problemas de modelado y estabilidad de sistemas de energía dominados por convertidores de fuente de voltaje: una revisión. CSEE J. Sistema de energía eléctrica. 2022, 8, 1530–1549.
6. Almaktoof, AMA Inversores multinivel que utilizan control de corriente predictivo de modelo de conjunto finito para sistemas de energía renovable Aplicaciones. Doctor en Filosofía. Tesis, Universidad de la Península del Cabo, El Cabo, Sudáfrica, marzo de 2015.
7. Gaballá, M.; El-Bardini, M. Generación de señales digitales de bajo costo para accionar un inversor PWM de vector espacial. Ain Shams Ing. J. 2013, 4, 763–774. [Referencia cruzada]
8. Jeong, M.; Gfrörer, T.; Biela, J. Ampliación del rango de parámetros estables de filtros LCL para convertidores conectados a la red mediante amortiguación inherente del control predictivo del modelo (MPC). En actas del 23.º taller del IEEE de 2022 sobre control y modelado de electrónica de potencia (COMPEL), Tel Aviv, Israel, 20 a 23 de junio de 2022.
9. Bernet, D.; Hiller, M. Convertidores de media tensión conectados a la red con filtros activos de fuente de tensión en paralelo. IET Electr. Aplicacion de energia 2019, 13, 1507-1513. [Referencia cruzada]
10. Kouchaki, A.; Nymand, M. Diseño analítico de filtro LCL pasivo para rectificadores de corrección del factor de potencia trifásicos de dos niveles. Traducción IEEE. Electron de potencia. 2018, 33, 3012–3022. [Referencia cruzada]
11. Kumar Peddapelli, S. 2. Técnica de modulación de ancho de pulso de vector espacial. En Modulación de Ancho de Pulso: Análisis y Rendimiento en Inversores Multinivel; De Gruyter Oldenbourg: Berlín, Alemania; Boston, MA, EE. UU., 2017; págs. 13-22. [Referencia cruzada]
12. Lewicki, A.; Odeh, Cl; Kondratenko, D.; Morawiec, M. Modulación de ancho de pulso de vector espacial hibridado para inversor de fuente de voltaje multifásico de dos niveles. Traducción IEEE. Electron de potencia. 2022, 37, 7663–7674. [Referencia cruzada]
13. Doan, NS; Tsvetkov, AN; Nguyen, TH Estudio e implementación del inversor de modulación de ancho de pulso de vector espacial en un Arduino. Conferencia web E3S. 2021, 288, 4–9. [Referencia cruzada]
14. Waqas, M.; Ahmed, T.; Elavarasan, RM; Waqar, A.; León, K.; Pugazhendhi, R.; Das, N.; Jeelani, MW DQ Control basado en transformación de inversor monofásico conectado a la red. En Actas de la 31.ª Conferencia de Ingeniería Energética de las Universidades de Australasia (AUPEC) de 2021, virtual, 26 a 30 de septiembre de 2021.
15. Venkatramanan, D.; Singh, MK; Ajala, O.; Domínguez-García, A.; Dhople, S. Modelos de sistemas integrados para redes con generadores e inversores. arXiv 2022, arXiv:2203.08253.
16. Transformaciones de Altahir, AAR Park y Clark: una breve reseña. Abril de 2020, págs. 2 a 5. Disponible en línea: https://www.researchga te.net/profile/Ali-Altahir/publication/340681625_Park_and_Clark_Transformations_Park_and_Clark_Transformations_A_S hort_Review/links/5e98bb8a4585150839e3768b/Park-and-Clark-Transformations-Park-and-Clark-Transformations-A-Short -Revisión.pdf (consultado el 27 de mayo de 2024).

Machine Translated by Google

17. Casado-Machado, F.; Martínez-Ramos, JL; Barragán-Villarejo, M.; Maza-Ortega, JM; Rosendo-Macias, JA Transformada de marco de referencia reducida: deconstrucción de sistemas trifásicos de cuatro hilos. Acceso IEEE 2020, 8, 143021–143032. [Referencia cruzada]
18. O'Rourke, CJ; Qasim, MM; Overlin, señor; Kirtley, JL Una interpretación geométrica de marcos de referencia y transformaciones: Dq0, Clarke y Park. Traducción IEEE. Conversaciones de energía. 2019, 34, 2070–2083. [Referencia cruzada]
19. Hollweg, GV; Khan, SA; Chaturvedi, S.; Fan, Y.; Wang, M.; Su, W. Convertidores conectados a la red: un breve estudio de topologías, filtros de salida, control de corriente y funcionamiento de redes débiles. Energías 2023, 16, 3611. [CrossRef]
20. Ali, M.; Yaqoob, M.; Cao, L.; Loo, KH Mejora de la regulación de voltaje del bus de CC en un sistema convertidor en cascada mediante un nuevo Esquema de control anticipativo de corriente de carga sin sensores. Electrón de potencia IET. 2021, 14, 1457–1467. [Referencia cruzada]
21. Wu, H.; Wang, X. Control de VSC que forman redes: una perspectiva de una fuente de voltaje interno adaptable rápida/lenta. Traducción IEEE. Electron de potencia. 2023, 38, 10151–10169. [Referencia cruzada]
22. Sadeque, F.; Fateh, F. Sobre esquemas de control para inversores formadores de red. En Actas de la Conferencia de Energía y Energía de Kansas (KPEC) del IEEE de 2022, Manhattan, KS, EE. UU., 25 y 26 de abril de 2022.

Descargo de responsabilidad/Nota del editor: Las declaraciones, opiniones y datos contenidos en todas las publicaciones son únicamente de los autores y contribuyentes individuales y no de MDPI ni de los editores. MDPI y/o los editores renuncian a toda responsabilidad por cualquier daño a personas o propiedad que resulte de cualquier idea, método, instrucción o producto mencionado en el contenido.

Machine Translated by Google

AppliedMath

MDPI

Artículo

Resolver problemas complejos de optimización mediante aprendizaje automático †

Steven Prestwich

Facultad de Ciencias de la Computación y Tecnología de la Información, University College Cork, T12 XF62 Cork, Irlanda; s.prestwich@cs.ucc.ie; Teléfono: +353-36-21-420-5911

† Este artículo es una versión ampliada de nuestro artículo publicado en la 9.ª Conferencia Internacional sobre Aprendizaje Automático, Optimización y Ciencia de Datos, LOD 2023 (Grasmere, Lake District, Reino Unido, 22 al 26 de septiembre de 2023).

Resumen: La mayoría de las investigaciones sobre optimización se centran en casos relativamente simples: un tomador de decisiones, un objetivo y posiblemente un conjunto de restricciones. Sin embargo, los problemas de optimización del mundo real a menudo conllevan complicaciones: pueden ser multiobjetivo, multiagente, multietapa o multinivel, y pueden tener incertidumbre, conocimiento parcial u objetivos no lineales. Cada uno de ellos ha llevado a áreas de investigación con métodos de solución específicos. Sin embargo, cuando se encuentran nuevos problemas híbridos, normalmente no hay una solución disponible. Definimos una clase amplia de problema de optimización discreta llamado programa de influencia y describimos un algoritmo liviano basado en aprendizaje por refuerzo multiobjetivo y multiagente con muestreo. Mostramos que se puede utilizar para resolver problemas de una amplia gama de literaturas: programación de restricciones, redes bayesianas, programación estocástica, diagramas de influencia (estándar, memoria limitada y multiobjetivo) y teoría de juegos (programación multinivel, juegos bayesianos). y razonamiento de nivel k). Esperamos que sea útil para la creación rápida de prototipos de métodos de solución para nuevos problemas híbridos.

Palabras clave: multiobjetivo; multiagente; aprendizaje por refuerzo; mejoramiento

1. Introducción

Los problemas de optimización son omnipresentes en la inteligencia artificial, la investigación de operaciones y una amplia gama de áreas de aplicación. En su forma más simple, requieren que tomemos un conjunto de decisiones para optimizar una función objetivo, posiblemente mientras se satisfacen algunas restricciones. Una decisión suele representarse mediante la asignación de un valor de dominio a una variable de decisión, donde el dominio podría ser un conjunto de números reales, enteros, categorías o estructuras de datos. Diferentes dominios, limitaciones y objetivos están respaldados por diferentes métodos. La programación lineal entera mixta (MILP) y, más generalmente, la programación matemática, la programación de restricciones (CP), la programación dinámica y las metaheurísticas (como la búsqueda local y los algoritmos genéticos) se pueden aplicar a estos problemas.

Sin embargo, muchos problemas de optimización son más complejos que esto. En particular, podrían implicar incertidumbre, que puede modelarse mediante variables aleatorias que no están controladas por quien toma las decisiones. En lugar de ello, toman valores según alguna distribución de probabilidad. La programación estocástica es una extensión de MILP que modela este tipo de problemas.
En programas estocásticos de múltiples etapas, las variables de decisión y de azar pueden estar entrelazadas de modo que debamos tomar decisiones sin conocer los valores de algunas variables de azar.
Una vez tomadas decisiones, podemos observar algunas variables aleatorias, pero luego debemos tomar decisiones para la siguiente etapa, y así sucesivamente. Una solución a un problema de múltiples etapas no es un simple conjunto de decisiones, sino una política que toma la forma de un árbol para permitir reacciones ante eventos aleatorios. Generalmente se supone que las distribuciones de probabilidad no se ven afectadas por las decisiones (incertidumbre exógena), pero algunas aplicaciones tienen distribuciones dependientes de la decisión (incertidumbre endógena). En este último caso se suele utilizar la programación dinámica estocástica.

Otra posible complicación es que quien toma las decisiones, o el agente, pueda tener múltiples objetivos: por ejemplo, maximizar las ganancias y minimizar el daño ambiental. Se debe llegar a alguna forma de compromiso para tales problemas. Otra complicación más es que un

check for updates

Cita: Prestwich, S. Resolución de problemas complejos de optimización mediante aprendizaje automático. Matemáticas aplicadas 2024, 4, 908–926. https://doi.org/10.3390/appliedmath4030049

Editores académicos: Libor Peka˘r y Jinyun Yuan

Recibido: 8 de mayo de 2024
Revisado: 24 de julio de 2024
Aceptado: 29 de julio de 2024
Publicado: 31 de julio de 2024

Copyright: © 2024 por el autor. Licenciatario MDPI, Basilea, Suiza.
Este artículo es un artículo de acceso abierto. distribuido bajo los términos y condiciones de los Creative Commons Licencia de atribución (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

Machine Translated by Google

El problema puede involucrar a múltiples agentes. En la programación multinivel, un agente debe tomar una decisión sabiendo cómo reaccionará otro agente, quien a su vez podría tener que tener en cuenta a otros agentes. El conocimiento puede ser parcial: un agente puede no saber exactamente cómo actuará o ha actuado un adversario, y debe tomar una decisión basada en suposiciones.

Los diagramas de influencia (ID) son un formalismo particularmente expresivo que puede modelar problemas de decisión de varias etapas con incertidumbre endógena. Las variables de azar pueden ser independientes o estar relacionadas mediante tablas de probabilidad condicional, formando una red bayesiana. También se permiten observaciones o los valores observados de variables aleatorias, y la inferencia bayesiana se puede utilizar para inferir las distribuciones de variables aleatorias no observadas basadas en estas observaciones. Hay varios métodos exactos y aproximados disponibles para resolver identificaciones.

Un enfoque de aprendizaje automático para problemas de optimización complejos es el aprendizaje por refuerzo (RL) [1] , que se puede utilizar para resolver problemas de decisión de múltiples etapas en entornos estocásticos . La RL multiagente se ha utilizado para aprender a jugar juegos como Go y Chess a niveles sobrehumanos a través del autojuego [2]. También se han ideado algoritmos de RL multiobjetivo y estas extensiones se han combinado en el campo de la RL multiobjetivo y multiagente . Se han producido desarrollos similares en la literatura sobre computación evolutiva.

Así, contamos con una amplia gama de tecnologías disponibles, cada una capaz de resolver cierto tipo de problemas de manera eficiente. Sin embargo, ante un nuevo problema de optimización con funciones híbridas, es posible que no haya una solución disponible. Se han investigado algunos híbridos, por ejemplo, el aprendizaje por refuerzo multiobjetivo con restricciones [3] y la programación lineal entera estocástica multiobjetivo de dos niveles [4]. Los ID se han ampliado para manejar utilidades no lineales [5], múltiples agentes [6-10], múltiples objetivos [11,12], restricciones estrictas [13] y variables parcialmente observadas [14], aunque no todos en un solo sistema. Pero en general no existe ningún método para híbridos complejos. Para tomar un ejemplo hipotético, no existe una forma obvia de modelar y resolver un problema de optimización discreto de dos niveles cuyo líder es un programa estocástico no lineal restringido al azar y cuyos seguidores son un problema de satisfacción de restricciones bioobjetivo y un Max ponderado. -Problema del SAT. Un investigador que se enfrenta a un problema de este tipo debe inventar un nuevo enfoque o hacer suposiciones simplificadoras para adaptarlo a un marco existente. Un método de uso general capaz de manejar todas estas complicaciones sería muy útil, al menos para una investigación preliminar, aunque eventualmente podría ser necesario un método especializado más eficiente.

Otro inconveniente de la fragmentación de las tecnologías es que resulta difícil comparar diferentes enfoques. Supongamos que nos enfrentamos a una nueva aplicación con dos tomadores de decisiones pero no estamos seguros de cuál es la mejor manera de modelarla. ¿Deberíamos tratarlo como un programa de dos niveles (ver Sección 3.8), un juego bayesiano (ver Sección 3.9), usar el pensamiento de nivel k (ver Sección 3.10), modelar a un tomador de decisiones como una variable aleatoria e implementar un programa estocástico (ver Sección 3.6) o tomar un valor esperado de un agente y reducirlo a un programa de restricciones (ver Sección 3.1)? Algunas simplificaciones podrían suponer poca diferencia para el objetivo, mientras que otras podrían hacer imposible la viabilidad. Cada enfoque requiere una tecnología diferente que requiere un tiempo considerable para dominarla y, posiblemente, un costo financiero. Un método que sea lo suficientemente general para resolver todos estos modelos sería de gran beneficio para la creación rápida de prototipos y para explorar las consecuencias de las elecciones de diseño del modelo.

En este artículo, describimos una clase muy general de problema de optimización discreta que llamamos Programa de Influencia (IP) que generaliza una amplia gama de problemas de investigación de operaciones, inteligencia artificial y teoría de juegos (el nombre está inspirado en la flexibilidad de los ID). También presentamos un solucionador simple para IP basado en una combinación de aprendizaje por refuerzo multiagente y multiobjetivo y muestreo Markov Chain Monte Carlo. RL se ha propuesto como un enfoque unificador para la toma de decisiones secuencial en condiciones de incertidumbre [15] y ha resuelto juegos adversarios complejos, lo que lo convierte en un candidato natural para un solucionador de propósito

Este artículo amplía nuestro trabajo anterior. En [16] se describió una versión anterior que utiliza programación lógica , y en [17] aparece una versión más reciente basada en C, donde nuestra clase de problema se denominó "diagrama de influencia mixta". Este artículo amplía [17] formalizando el problema, proporcionando más trabajo relacionado, reemplazando el muestreo de rechazo simple por muestreo de Gibbs y resolviendo una gama más amplia de problemas. La sección 2 formaliza el problema.

Machine Translated by Google

y describe el algoritmo. La Sección 3 lo aplica a una variedad de problemas de diferentes literaturas de optimización para demostrar su flexibilidad y facilidad de uso. Finalmente, la Sección 4 resume los resultados y analiza el trabajo futuro.

2. El marco de programación de influencia

Ahora presentamos nuestro marco de propiedad intelectual y un algoritmo, después de discutir el trabajo relacionado.

2.1. Trabajo relacionado

Se ha realizado un trabajo considerable en RL multiobjetivo (MORL). La literatura MORL es demasiado extensa para examinarla aquí, pero en [18] se proporciona una encuesta reciente sobre los algoritmos MORL con una discusión sobre estos temas . Algunos métodos utilizan la escalarización para convertir MORL en RL, lo que lo convierte en un método de política única como muchos otros. La escalarización ha sido criticada porque conduce a una sola política, y algunos algoritmos MORL se aproximan a los frentes de Pareto y aprenden políticas para una variedad de ponderaciones. La escalarización también ha sido criticada con el argumento de que puede resultar difícil elegir los pesos adecuados y que la elección pone el poder de decisión en manos de los ingenieros que ejecutan el algoritmo. La mayoría de estos métodos utilizan escalarización lineal. A pesar de las deficiencias de la escalarización, demostraremos que puede dar buenos resultados en una variedad de problemas. También es fácil de implementar y tiene poca sobrecarga de tiempo de ejecución, lo cual es importante para nuestro enfoque liviano. Los resultados que utilizan métricas ponderadas también son menos sensibles a la elección de las ponderaciones [19].

También se ha trabajado mucho sobre RL multiagente (MARL), y se puede encontrar una encuesta reciente de MARL en [20]. Se ha realizado menos trabajo en la intersección de MARL y MORL: RL multiobjetivo y multiagente (MOMARL). En [21] se proporciona una encuesta sobre MOMARL y algoritmos relacionados . Debido a la complejidad de los problemas de MOMARL, no existe ni siquiera una única definición acordada de lo que constituye una solución. MO-MIX [22] es un algoritmo MOMARL que utiliza escalarización lineal y RL profunda en el que se utiliza una red neuronal artificial para la agregación de estados.

2.2. La clase problemática

Definimos un programa de influencia (IP) (discreto) como una tupla V, D, A, U,L, O,P donde:

- V = v1, . . . , vn es una lista ordenada de variables vi ; • D = D1, . . . , Dn es una lista de sus correspondientes dominios finitos Di de valores posibles, normalmente rangos de números enteros o conjuntos de nombres simbólicos;
- A = A1, . . . , An es una lista de sus agentes correspondientes (tomadores de decisiones) Ai , que tienen nombres simbólicos (en negrita);
- U es una función que mapea una asignación de variable total a un vector de utilidad para cada agente; • L es un conjunto de enlaces dirigidos vi → vj (i < j) entre variables; • O es un conjunto de observaciones, donde una observación es una asignación vi = dj de un valor dj Di a una variable aleatoria xi ;
- P es una función que asigna una probabilidad a cada variable aleatoria vj , dada una asignación para cada variable vi tal que vi → vj L: normalmente se representa mediante una tabla de probabilidad (condicional).

Cada variable está asociada exactamente a un agente, y permitimos la posibilidad de variables aleatorias cuyo agente tenga el nombre azar. Los enlaces L definen qué variables son visibles para decisiones posteriores y distribuciones aleatorias de variables. Una utilidad es una función que asigna una asignación de variable total (un valor por variable) a un valor real: permitimos que las utilidades sean programables, lo que requiere que el usuario escriba una pequeña función para calcular las utilidades a partir de una asignación de variable total. Cada agente tiene al menos una utilidad, aparte del azar. El objetivo de una IP es encontrar una política para cada agente que optimice Pareto sus utilidades en el contexto de observaciones y visibilidad entre variables.

Machine Translated by Google

2.3. Un algoritmo

Ahora describimos el algoritmo INFPROG que se muestra en el Algoritmo 1. Es un prototipo de investigación liviano que utiliza técnicas conocidas, usado solo para demostrar la flexibilidad de nuestro enfoque, y muchas otras combinaciones de métodos son potencialmente posibles.

INFPROG se basa en un algoritmo RL simple: SARSA tabular de pasos infinitos con selección de acción codiciosa y tasa de aprendizaje α [1]. El paso infinito indica que la recompensa se respalda por igual en todos los valores del episodio, lo que es más sólido que el Q-learning en presencia de variables no observadas [1]. Tabular indica que los pares estado-acción tienen valores en una tabla. El factor de descuento γ se establece en 1 ya que nuestro problema RL es episódico. Si un agente tiene múltiples utilidades, estas se escalan de modo que haya una utilidad por agente (ver Sección 2.6).
Los objetivos escalarizados se utilizan como recompensas, se calculan al final de cada episodio cuando a todas las variables se les han asignado valores actualizados y se respaldan en estados anteriores para esos agentes (a[v] denota el agente a correspondiente a la variable v): el usuario debe proporcionar código para este paso. Sin embargo, el usuario probablemente no esté interesado en el valor del objetivo escalarizado, por lo que se imprimen versiones suavizadas de los objetivos originales: estos serán los valores que reportaremos.

Algoritmo 1 El algoritmo INFPROG para resolver IP Requiere: números enteros

```
H, E y un hiperparámetro de utilidad Y inicializar      = 1, α = 1, qH,v,i = 0 para el
    episodio e = 1 E hacer para v = v1 . . . vn hacer
                                    . . .

                si v es probabilidad entonces

                    muestrear xv de su distribución, de lo contrario,
                si U[0, 1] <     , entonces
                    muestreará aleatoriamente xv      Dv ,
                de lo contrario

                    xv = argmini    Dv (qH,v,i) end if
                end for
            para cada

            agente a calcula el objetivo
                escalarizado fa(    v) usando Y end for for v = v1 . . . vn hacer
            qH,v,xv =

            α fa[v] (    v) + (1 − α)qH,v,xv


            fin de la

            actualización      y α
    fin de
```

La utilidad se calcula al final de un episodio: en términos de RL, se trata de una recompensa escasa que puede dificultar el aprendizaje. En trabajos futuros, podríamos permitir múltiples nodos de valor como en los ID, o compensar las escasas recompensas agregando técnicas de RL como la repetición de experiencias en retrospectiva [23]. Esa técnica fue diseñada para algoritmos RL fuera de la política, pero SARSA está dentro de la política, por lo que sería necesaria alguna modificación del algoritmo: por ejemplo, reemplazar SARSA por Q-learning o un algoritmo RL profundo más reciente. Hay muchas posibilidades y esta área está lista para ser explorada.

Los parámetros y α comienzan en 1 y decaen a 0. Se han propuesto muchos esquemas de decaimiento en RL, y elegimos arbitrariamente α = = ((E − e)/E) [3]. En nuestro uso de RL, un estado es una asignación a variables v1 . . . vi para algunos i = 1 . . . n, una acción es la asignación de un valor a una variable de decisión vi+1 , un episodio asigna todas las variables y la recompensa es la utilidad (valor de la función objetivo) calculada al final de un episodio. Los qH,v,xv son valores de estado-acción utilizados para definir la política, que debería optimizar la recompensa esperada.

Tenga en cuenta que es probable que para algunos problemas sea mejor utilizar SARSA con un seguimiento de elegibilidad y 0 < λ < 1 (λ es un hiperparámetro utilizado con seguimientos de elegibilidad [1]) y investigaremos esto en trabajos futuros. Sin embargo, para nuestra investigación, el prototipo que

Machine Translated by Google

establezca efectivamente λ = 1 eligiendo SARSA de pasos infinitos, simplificando así el algoritmo al eliminar un hiperparámetro que requiere ajuste.

2.4. Agregación de

estados Cada decisión puede depender de todas las decisiones anteriores y eventos aleatorios (pero consulte la Sección 3.5), por lo que una política puede involucrar una gran cantidad de estados distintos. Para combatir este problema, los algoritmos RL agrupan estados mediante métodos de agregación de estados . Elegimos una forma simple llamada codificación aleatoria de mosaicos [1], específicamente hash Zobrist [24] con entradas de la tabla hash H para algún número entero grande H. Esto funciona de la siguiente manera. A cada par variable-valor (de decisión o azar) v, x , le asignamos un entero aleatorio rvx que permanece fijo. En cualquier momento durante un episodio, tenemos un conjunto S de asignaciones v, x , y tomamos el -or exclusivo de los valores rvx (es decir, sus patrones de bits) asociados con las asignaciones XS = v,x S rvx. Finalmente, usamos XS para indexar una matriz V con H entradas: el valor de qH,v,xv se almacena en V[XS mod H] (en todos nuestros experimentos, fijamos H = 2 × 108).

El algoritmo INFPROG toma dos hiperparámetros numéricos: un número entero H usado para la agregación de estados y un número entero E que es el número de iteraciones utilizadas por el solucionador. Si H es lo suficientemente grande, entonces las colisiones hash son poco probables y tendremos un elemento de matriz único para cada estado encontrado. Se podría esperar que el hash de Zobrist tenga un rendimiento deficiente cuando el número de estados se acerque o exceda el tamaño de la tabla hash, porque las colisiones hash confundirán los valores de diferentes pares estado-acción. Sorprendentemente, puede funcionar bien incluso cuando las colisiones de hash son frecuentes y se ha utilizado en la programación de ajedrez [25].

2.5. Muestreo

de INFPROG muestrea variables aleatorias utilizando un algoritmo Monte Carlo de cadena de Markov (MCMC) como en la programación probabilística. Por lo tanto, cada episodio de SARSA es también un barrido de las variables aleatorias. Nuestro trabajo anterior [16,17] utilizó un muestreo de rechazo: al final de un episodio de SARSA, si las variables de probabilidad no coincidían con las observaciones, entonces el episodio era rechazado, en el sentido de que las recompensas no estaban respaldadas por estados anteriores. Esto funcionó en algunos problemas, pero no es práctico cuando las probabilidades son muy bajas, y se propuso el uso de un algoritmo MCMC como trabajo futuro. INFPROG utiliza muestreo de Gibbs y puede manejar estos casos.

En los experimentos, encontramos que algunos valores de la tabla de probabilidad condicional que se establecieron en 0 o 1 requerían ajuste: este es un modo de falla conocido del muestreo de Gibbs. De hecho, para problemas de optimización sin observaciones (lo cual es cierto para la mayoría de los problemas), reemplazamos a Gibbs con un muestreo de transformada inversa simple.

2.6. Múltiples objetivos

Muchas aplicaciones del mundo real tienen múltiples objetivos y debemos encontrar un equilibrio. Los objetivos se pueden combinar de más de una manera en RL, y un enfoque simple y popular es la escalarización lineal: tomar una combinación lineal de los objetivos, dando un único objetivo que puede usarse como recompensa de RL. Esto tiene el inconveniente de que no puede generar ninguna solución en una región no convexa del frente de Pareto. También puede resultar difícil elegir los pesos adecuados, especialmente cuando los objetivos utilizan unidades diferentes.

Otro enfoque es clasificar los objetivos por importancia y luego buscar el mejor resultado lexicográficamente. Esto tiene el mismo problema que la escalarización lineal (aunque un trabajo reciente lo aborda [26]), que puede solucionarse poniendo umbrales en todos los objetivos excepto en los menos importantes , un método llamado ordenamiento lexicográfico umbralizado [27]. Sin embargo, tiene el inconveniente de utilizar un algoritmo RL específico. Además, en algunas de nuestras aplicaciones, el objetivo más importante es minimizar las violaciones de restricciones, y esto no debe establecerse como umbral.

Machine Translated by Google

El método que elegimos es reducir los múltiples objetivos a un único objetivo mediante escalarización métrica ponderada [19] en la que se minimiza la distancia entre el vector de valores fo para los objetivos o y un punto utópico u en el espacio multiobjetivo:

$$\min_x \sum_o w_o |f_o(x) - u|^p \quad {}_{o|pag}$$

$$1/p$$

para algunos p ≥ 1 y pesos wo elegidos por el usuario. La necesidad de elegir pesos es una desventaja en términos de facilidad de uso, pero una ventaja es que podemos ajustarlos para encontrar diferentes puntos en el frente de Pareto. Teóricamente se garantiza que el caso especial de la escalarización de Chebyshev (p = ∞) hace que todo el frente de Pareto sea alcanzable, pero las normas Lp (p = 1, 2) pueden funcionar mejor en la práctica [28]. En los experimentos encontramos mejores resultados con L1 que con L2 o L∞ a pesar de las ventajas teóricas de Chebyshev, por lo que usaremos L1.
Sin embargo, en trabajos futuros, podría ser mejor utilizar Chebyshev y encontrar formas de mejorar sus resultados.

El punto utópico a menudo se ajusta durante la búsqueda para que esté justo más allá del mejor punto encontrado hasta el momento, pero INFPROG usa una u fija proporcionada por el usuario. Dependiendo de si cada objetivo debe maximizarse o minimizarse, elegimos un valor que sea óptimo o lo suficientemente alto/bajo como para ser inalcanzable.

Por lo tanto, para aplicar INFPROG a un problema de optimización, también debemos proporcionar un hiperparámetro Y para guiar la escalarización: una lista de sublistas de pares. Cada par es un valor utópico que expresa una utilidad deseada más un peso, y cada sublista corresponde a un agente y contiene un par para cada una de las utilidades del agente. Siempre elegiremos pesos que sumen 1 para cada agente, aunque esto no es estrictamente necesario. El uso de Y se ilustrará a continuación mediante ejemplos.

2.7. Una explicación intuitiva

Para ayudar al lector, ahora proporcionamos una explicación intuitiva de cómo funciona el algoritmo. El bucle externo del algoritmo realiza E episodios (un término tomado de RL) donde E es un hiperparámetro entero elegido por el usuario. Cada episodio recorre las variables en el orden especificado por la lista v1 ,. . . , vn , asignando a cada uno un valor.

A una variable de decisión se le asigna un valor como en el algoritmo SARSA: durante los primeros episodios, la asignación es en gran medida aleatoria, pero en episodios posteriores, los valores se asignan con más avidez para maximizar la recompensa esperada estimada (otro término de RL, correspondiente a la utilidad) . en diagramas de influencia y al valor de la función objetivo en optimización general).
El grado de avidez está controlado por el parámetro , que desciende de 1 (completamente aleatorio) en el primer episodio a 0 (completamente avaricioso) en el último episodio. En los algoritmos RL, esta estrategia conduce teóricamente a una política óptima.

A una variable aleatoria se le asigna un valor mediante muestreo, por lo que un episodio intercala SARSA con muestreo. En términos de RL, la asignación de variables de probabilidad es simplemente parte del entorno en el que opera el agente. En el caso especial en el que todas las variables son variables aleatorias, la IP es una red bayesiana y un episodio se reduce a un barrido MCMC a través de las variables.

Al final de un episodio, se calcula una recompensa para cada agente y se realiza una copia de seguridad de todos los estados anteriores como en SARSA, con la ligera complicación de que en nuestra versión multiagente cada recompensa debe tener una copia de seguridad del agente apropiado.

La elección del SARSA tabular de pasos infinitos fue en gran medida arbitraria y tiene ventajas y desventajas. Un algoritmo dentro de políticas como SARSA se considera más consistente y estable que un algoritmo fuera de políticas como Q-learning, pero también más lento y menos eficiente. La elección de un algoritmo tabular se realizó por su simplicidad de implementación.
La elección de pasos infinitos corresponde a establecer λ = 1 en una traza de elegibilidad de RL, lo que la convierte en un método de Monte Carlo en lugar de un método de diferencia temporal. Los primeros se consideran más robustos en presencia de variables no observadas, ya que no se basan en la propiedad de Markov.

Machine Translated by Google

La restricción de recompensas que ocurren solo al final de un episodio fue diseñada para simplificar la implementación y la interfaz de usuario. En algunos casos, sería preferible permitir recompensas intermedias, si éstas ocurren naturalmente en el problema. Sin embargo, esto requeriría que el usuario escriba varios fragmentos de código y especifique cuándo se ejecutará cada uno.
Para nuestro prototipo de investigación tomamos el camino más sencillo.

Respecto a la necesidad de que el usuario proporcione funciones de utilidad programables: aunque esto pueda parecer menos fácil de usar que alguna forma de lenguaje de especificación, es extremadamente flexible. Observamos que se adopta un enfoque similar en el campo de la programación probabilística [29,30].

3. Aplicaciones En

esta sección, tomamos pequeños problemas de una variedad de estudios en la literatura, los modelamos como IP y los resolvemos usando INFPROG. No compararemos nuestro método con otros en términos de eficiencia, ya que éste no es el objetivo del artículo. Por la misma razón, informamos pocos tiempos de ejecución, aunque son bastante cortos: normalmente unos segundos y, como máximo, unos minutos. De hecho, no esperamos que sea competitivo en ninguna clase particular de problema, aunque como método basado en RL, debería funcionar razonablemente bien en aplicaciones de la literatura de RL. Nuestro objetivo aquí es sólo demostrar que un único solucionador puede resolver una amplia variedad de problemas de optimización, llenando así un vacío en la tecnología de optimización. Tenemos la intención de implementar solucionadores de IP más rápidos, escalables y fáciles de usar en trabajos futuros, y esperamos que otros investigadores también encuentren mejores algoritmos.

3.1. Programación de restricciones

Primero consideramos un problema de optimización “simple”: un agente y un objetivo. Tomamos un conocido problema de satisfacción de restricciones (CSP, por sus siglas en inglés) conocido como ocho reinas: colocamos ocho reinas de ajedrez en un tablero de ajedrez estándar de 8 × 8 de tal manera que dos reinas no se ataquen entre sí (por estar en la misma fila, columna o diagonal). Este es un problema popular en la programación de restricciones [31], y el problema más pequeño de las cuatro reinas se ilustra en la Figura 1. El primer ejemplo es una solución porque no hay dos reinas en la misma fila, columna o diagonal. El segundo ejemplo no es solución porque viola tres restricciones: las reinas de las dos primeras filas están en la misma columna, las de las dos últimas filas están en la misma diagonal y las de la primera y tercera filas están en otra. diagonal.

Tratamos a ocho reinas como un problema de Max-CSP en el que el objetivo es maximizar el número de restricciones satisfechas (en este caso, minimizar el número de violaciones de restricciones a 0). Consideraremos dos posibles modelos de IP.

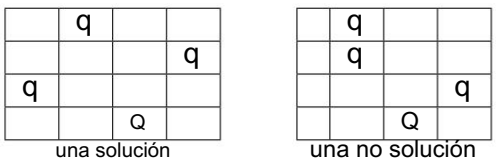


Figura 1. El problema de las 4 reinas.

3.1.1. Un modelo disperso

El problema se puede modelar de la siguiente manera: V = v1, . . . , v8 , cada Di = {1, . . . , 8}, A = a, . . . , a (todas las variables pertenecen al mismo agente a), U = viol donde viola cuenta violaciones de restricciones:

$$\sum_{\sum vi = vj \quad |i-j| = |vi-vj|}^{8} {}^{yo-1}_{yo=2\ j=1}$$

· es el corchete de Iverson que toma el valor 1 si su argumento es verdadero y 0 si es falso. P es la función vacía (no hay tablas de probabilidad porque no hay variables aleatorias) y L = O = . El hiperparámetro de utilidad es Y = 0, 1 : un agente con una utilidad cuyo punto utópico es 0 y el peso asociado es 1.

Machine Translated by Google

Aunque no informaremos sobre los tiempos de ejecución en general, podría ser de interés para el lector ver un ejemplo. Nuestro solucionador está implementado en C y ejecutado en una CPU Intel(R) Core(TM) i5-6500 a 3,20 GHz con 8 GB de RAM del sistema (Intel, Santa Clara, CA, EE. UU.). Con 107 episodios, el solucionador tarda 10,3 s por ejecución y, en 100 ejecuciones, encontró una solución correcta 70 veces. Las otras 30 veces, la solución contenía una violación de restricción, lo que ilustra que INFPROG puede quedar atrapado en óptimos locales. También tiene bastante éxito utilizando sólo 106 episodios y encontrando 61 soluciones correctas.

Este resultado no es competitivo con métodos como la programación de restricciones, la programación entera o la búsqueda local, que pueden resolver el problema de manera más rápida y confiable. Pero, como se indicó anteriormente, nuestro objetivo es la flexibilidad en una amplia gama de problemas, no la eficiencia en una clase de problema en particular.

3.1.2. Un modelo denso

Nos referimos al modelo anterior como disperso porque tenía pocos enlaces (0). Como modelo denso alternativo , podemos agregar enlaces al IP de cada variable de decisión a cada una de las anteriores: L = {vi → vj |i < j}. Estos no son lógicamente necesarios pero aportan más información a las variables de decisión. Con la densa propiedad intelectual, encontramos un 100% de éxito en ocho reinas usando solo 105 episodios, e incluso con solo 30,000 episodios, logró un 69% de éxito. Esto sugiere que hacer que las decisiones sean visibles para todas las decisiones posteriores mejorará el desempeño. Sin embargo, agregar estos enlaces crea una cantidad de estados que son exponenciales en la cantidad de variables, lo que causa muchas colisiones de hash en INFPROG, lo que a su vez podría degradar el rendimiento.

Este fenómeno es bien conocido en RL. Como ejemplo, considere el problema de despacho de ascensores en el que debemos controlar un conjunto de ascensores para minimizar los tiempos de espera esperados. Incluso si restringimos el problema al tiempo discreto, tiene un gran número de estados, debido a las muchas combinaciones posibles de posiciones del ascensor, la dirección del movimiento y qué botones se han presionado. En principio, sería mejor tratar cada estado por separado, y el rendimiento superior de nuestro modelo denso respalda esta opinión, pero la explosión combinatoria lo hace poco práctico para problemas grandes.

Una forma de evitar la explosión es tratar cada ascensor de forma independiente en un enfoque RL distribuido ; ver, por ejemplo, [32]. Hay un agente que toma decisiones para cada ascensor y no se comunican, pero tienen un objetivo común e indirectamente aprenden a cooperar. En [33] se ofrece una encuesta reciente sobre RL distribuida con aplicaciones . Nuestra IP dispersa funciona de la misma manera, aunque solo usamos un agente: en su lugar, podríamos usar un agente separado para cada variable, pero no hay diferencia porque tienen el mismo objetivo.

Controlar los vínculos entre variables facilita que INFPROG emule métodos RL que están total, parcialmente y no distribuidos. Para problemas posteriores, usaremos una mezcla de modelos dispersos y densos.

3.2. Redes Bayesianas

De un problema con todas las variables de decisión, pasamos a un problema con todas las variables de probabilidad: el ejemplo de alarma de Pearl [34]. Este es un ejemplo de inferencia probabilística, tal como se realiza en el campo de la programación probabilística: inferir una probabilidad condicional a partir de un programa probabilístico.

En este pequeño problema, una casa tiene una alarma que puede activarse ante un terremoto o un robo, cada uno con una probabilidad previa. También tenemos probabilidades condicionales para la alarma en diferentes circunstancias: cuando hay/no hay robo y cuando hay/no hay terremoto. Dado que se ha activado la alarma, ¿cuál es la probabilidad de que se haya producido un robo? La respuesta afecta nuestras acciones.

Machine Translated by Google

Podemos modelar esto como una IP. Las variables son V = {robo, terremoto, alarma}, los dominios D son todos {sí, no}, los agentes A son todos azar, los enlaces son L = {robo → alarma, terremoto → alarma} como en la red bayesiana , las probabilidades P son:

Pr[robo = sí] = 0.7
Pr[terremoto = sí] = 0.2
Pr[alarma = sí|robo = sí,terremoto = sí] = 0.9 Pr[alarma = sí|robo = sí,terremoto = no] = 0.8 Pr[alarma = si|robo = no,terremoto = si] = 0.1 Pr[alarma = si|robo = no,terremoto = no] = 0.2

hay una observación O = {alarma = sí} y el vector de utilidad es

U = terremoto = sí

Por lo tanto, la tarea es calcular la probabilidad condicional Pr[terremoto = sí|alarma = sí]. Podemos tratar esto como una utilidad, aunque el único agente es el azar, por lo que no intentamos optimizarlo tomando decisiones. No existen agentes que tomen decisiones, por lo que el hiperparámetro de utilidad es Y = . INFPROG aplica el muestreo de Gibbs y encuentra una "utilidad" esperada de aproximadamente 0,23, que es la probabilidad correcta.

No utilizaremos observaciones en ejemplos posteriores ya que no son una característica de la mayoría de los problemas de optimización. Sin embargo, se pueden utilizar para calcular el valor de la información y determinar qué variables tienen más probabilidades de reducir la incertidumbre en una variable de interés.

3.3. Diagramas de influencia

Pasamos ahora a problemas que contienen variables de decisión y de probabilidad. Los ID [35] son modelos gráficos populares en el análisis de decisiones y pueden modelar relaciones importantes entre incertidumbres, decisiones y valores. Inicialmente fueron concebidos como herramientas para la formulación de problemas, pero también han surgido como herramientas computacionales eficientes. Tienen muchas aplicaciones, incluido el diagnóstico médico [36], la ciberseguridad [37] y la gestión de riesgos [38].

Un ID es un gráfico acíclico dirigido con tres tipos de nodos: los nodos de decisión corresponden a variables de decisión y se dibujan como rectángulos; los nodos aleatorios (o inciertos) corresponden a variables aleatorias y se dibujan como óvalos; Los nodos de valor corresponden a preferencias u objetivos y se dibujan como rectángulos redondeados o polígonos como diamantes.

Cada variable de probabilidad está asociada con una tabla de probabilidad condicional que especifica su distribución para cada combinación de valores de sus variables principales en el gráfico. Cada variable de decisión también tiene un conjunto de variables principales en el gráfico y su valor depende únicamente de sus valores. Generalmente se considera que las variables de decisión están ordenadas temporalmente y las variables aleatorias se observan en diferentes puntos del ordenamiento. Una suposición de identificación estándar es sin olvido, lo que significa que los padres de cualquier variable de decisión son todos sus antepasados en el gráfico: por lo tanto, una decisión puede depender de todo lo que ha ocurrido antes. Todas las variables son discretas.

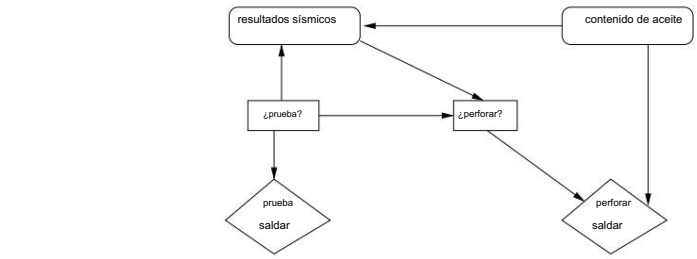
Cada nodo de valor está asociado con una tabla que muestra la utilidad (un número real) de cada combinación de valores de variables principales. Una política de decisión es una regla para cada variable de decisión que indica cómo elegir su valor entre los de sus padres. Cualquier política tiene una utilidad esperada total (es una expectativa debido a las variables de probabilidad) y resolver un ID significa calcular su política óptima, que tiene la utilidad esperada máxima.

Existen varios métodos para resolver ID. Algunos son exactos y se basan en la eliminación de variables [39-43], mientras que otros son aproximados [44-48]. Sin embargo, aparte de nuestro trabajo ([16,17] y este artículo), no conocemos otras aplicaciones de RL para resolver ID, lo que parece sorprendente ya que tanto ID como RL pueden usarse para modelar y resolver problemas de decisión secuencial La conexión principal que generalmente se establece entre los dos es que las ID pueden modelar problemas que pueden ser abordados por RL; por ejemplo, las ID causales se han utilizado para modelar marcos de seguridad de inteligencia artificial general que a menudo utilizan RL [49]. Nuestro enfoque basado en RL se encuentra en alguna parte

Machine Translated by Google

entre métodos exactos y aproximados: dada suficiente memoria y tiempo de entrenamiento, tiene
tiene el potencial de encontrar una solución óptima, pero no está garantizado que lo haga. no esperamos
Es tan eficiente como los algoritmos especializados, pero puede abordar identificaciones con extensiones como
memoria limitada (ver Sección 3.5), múltiples agentes y múltiples objetivos. Además, el
Los éxitos recientes de la RL profunda hacen que su aplicación a grandes ID sea una dirección de investigación prometedora.
Sin embargo, en este artículo nos limitamos a demostrar que RL realmente puede resolver ID.

Usamos el ID de Oil Wildcatter que se muestra en la Figura 2, un problema bien conocido publicado
en [50]. Un buscador de petróleo debe decidir si perforar o no en busca de petróleo en un sitio específico.
Antes de perforar, podrán realizar una prueba sísmica que ayudará a determinar la situación geológica.
estructura del sitio. El resultado de la prueba puede ser cerrado (lo que indica una cantidad significativa de aceite), abierto (lo que indica
algo de aceite) o difuso (probablemente sin aceite). La nota de valor especial significa que los resultados de la prueba son
no disponible si no se realiza la prueba sísmica. La decisión de la prueba no depende de ningún
otra variable, pero la decisión de perforación depende de si se realizó una prueba y, de ser así, de su
resultado. La variable del petróleo no es observable, por lo que ninguna decisión depende de ella (esto es distinto de
olvidándose de su valor que se aborda en el apartado 3.5).



contenido	Pro]		
	remojo	húmedo	seco
	0,5	0.3	0,2

resultado de la prueba U1(T)	
prueba	
si no	−10 0

Resultados		Pr[S|O, T]			
sísmicos	contenido	prueba cerrado	abierto	nota difusa	
seco	Sí	0,1	0,3	0,6	0
seco	No	0	0	0	1
húmedo	Sí	0,3	0,4	0,3	0
remojo	No	0	0	0	1
mojado	Sí	0,5	0,4	0,1	0
remojo	No	0	0	0	1

Ejercicio de contenido		U2(O, D)
de resultados de	ejercicios	
seco	Sí	−70
seco	No	0
húmedo	Sí	50
remojo	No	0
mojado	Sí	200
remojo	No	0

Figura 2. Diagrama de influencia de los contaminantes salvajes del petróleo.

Un modelo de IP es el siguiente. Las variables son V = petróleo,prueba,sísmica, perforación , los dominios
son

D = {seco, húmedo,remojo}, {sí, no}, {cerrado, abierto, difuso, nota}, {sí, no}

los agentes son A = chance, empresa, chance, empresa , los enlaces son L = {simulacro →
sísmico, perforación → prueba} (los enlaces de identificación a los servicios públicos no son parte de una IP), no hay observaciones
(O =), el vector de utilidad es U = t + d donde t es el resultado de la prueba y d es el taladro
pago, y las probabilidades son como se muestran en la Figura 2. El hiperparámetro de utilidad es
Y = 1000, 1 : un agente con un objetivo a maximizar.

La política óptima conocida dadas las utilidades y probabilidades anteriores es la siguiente:
aplicar la prueba sísmica y perforar si el resultado de la prueba es abierto o cerrado. INFPROG encuentra esto
solución e informa una aproximación cercana a la utilidad esperada correcta: 22.5.

3.4. Diagramas de influencia multiobjetivo

Como primer ejemplo de optimización multiobjetivo, utilizamos un wildcatter de petróleo bioobjetivo.
Identificación de [12]. Además de maximizar los beneficios, el objetivo es minimizar los daños medioambientales.
daño. El IP es como en la Sección 3.3, excepto por pagos adicionales: la utilidad de las pruebas es
ahora (−10, 10) mientras que la utilidad de la perforación es (−70, 18), (50, 12) y (200, 8) para seco, húmedo,

Machine Translated by Google

y remojo, respectivamente: en cada caso, el primer valor es la utilidad original mientras que el segundo valor se refiere a la nueva utilidad. En lugar de una solución óptima, hay cuatro soluciones óptimas de Pareto, cada una con dos valores de utilidad: (1) probar y luego explorar si el resultado es cerrado o abierto, con utilidad (22.5, 17.56); (2) no probar sino perforar, con la utilidad (20, 14.2); (3) probar, luego practicar si el resultado es cerrado, con utilidad (11, 12.78); (4) no probar ni perforar, con utilidad (0, 0).

Las utilidades ahora son U = p, d , donde p es el pago como antes yd es el daño ambiental. Con un hiperparámetro de utilidad Y = 1000, 0.55 , −1000, 0.45 , en múltiples ejecuciones, INFPROG encontró las políticas (1), (2) y (4), pero no (3).

3.5. Diagramas de influencia de la memoria limitada

Las identificaciones estándar están diseñadas para manejar situaciones que involucran a un único agente que no se olvida. Los diagramas de influencia de memoria limitada (LIMID) [14] son generalizaciones de ID que permiten la toma de decisiones con información limitada y decisiones simultáneas, y pueden tener políticas mucho más pequeñas. Relajan la regularidad (ordenamiento total de las variables) y los supuestos de no olvidar las identificaciones. Los LIMID se consideran más difíciles de resolver de manera óptima que los ID. Manejamos la característica de memoria limitada de una manera simple: durante el hash de Zobrist, el conjunto S contiene solo asignaciones que son visibles para la variable de decisión (según lo especificado por los enlaces IP L).

Por ejemplo, utilizamos un problema de cría de cerdos de [14]. Un criador de cerdos cría cerdos durante cuatro meses y luego los vende. Durante este período, el cerdo puede desarrollar o no una enfermedad. Si tiene la enfermedad cuando debe venderse, entonces debe venderse para el sacrificio y su precio de mercado esperado es 300. Si está libre de la enfermedad, entonces su precio de mercado esperado es 100 Una vez al mes, un veterinario examina al cerdo para detectar la enfermedad. Si está enfermo, entonces el test lo indica con una probabilidad de 0,80, y si está sano, entonces el test lo indica con una probabilidad de 0,90. En cada visita mensual, el cirujano puede tratar o no al cerdo, y el tratamiento cuesta 100. Un cerdo padece la enfermedad en el mes 1 con una probabilidad de 0,10. Un cerdo sano desarrollará la enfermedad en el próximo mes con una probabilidad de 0,20 sin tratamiento y 0,10 con tratamiento. Un cerdo enfermo seguirá estando enfermo durante el próximo mes con una probabilidad de 0,90 sin tratamiento y de 0,50 con tratamiento. La identificación se muestra en la Figura 3.

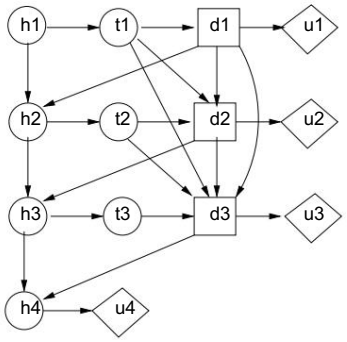


Figura 3. Diagrama de influencia de la cría porcina.

Una IP para modelar este problema es la siguiente. Las variables son V = hi , ti , di ,. . . , h4 (i = 1, 2, 3), los dominios D son {saludable, no saludable} para hi , {pos, neg} para ti y {tratar, salir} para di , hay un agente en A llamamos criador, los enlaces son L = {hi → hi+1 , hi → tj(i < j), ti → dj(i ≤ j), di → hi+1}, P contiene las probabilidades mostradas, U = u1 + u2 + u3 + u4 , y O = .

Usando un hiperparámetro de utilidad Y = 1000, 1 , INFPROG casi siempre encuentra la política óptima con una utilidad esperada de aproximadamente 727: ignorar la primera prueba y no tratar, luego seguir los resultados de las otras dos pruebas (tratar si es positivo). Como se señala en [16], en [14] se da una política diferente: tratar en el mes 3 si las pruebas 1, 2 o 3 son positivas. Encontramos que su política tiene una utilidad esperada de 725,884 mientras que la nuestra es óptima. Informan de la misma utilidad esperada que nosotros, por lo que creemos que se trata simplemente de un error tipográfico.

Machine Translated by Google

3.6. Programación estocástica de múltiples etapas

La programación estocástica [51] modela y resuelve problemas que involucran variables de decisión y de azar, con distribuciones conocidas para estas últimas. Los problemas pueden tener una o más etapas: en cada etapa se toman decisiones y luego se observan variables aleatorias. Una solución no es una simple asignación de variables, sino una política que nos dice cómo tomar decisiones dadas las asignaciones a variables de etapas anteriores. La programación estocástica se remonta a la década de 1950 y ahora es un área importante de investigación en programación matemática e investigación de operaciones.

Ahora mostramos que los programas estocásticos discretos se pueden modelar y resolver como IP. Usamos un programa estocástico de dos etapas de [52]:

max z = 1.5x1 + 4x2 + E[Q(x1, x2)] donde
Q(x1, x2) es el valor de max 16y1 +
19y2 + 23y3 + 28y4 tal que 2y1 + 3y2 + 4y3 + 5y4 ≤
ξ1 − x1 6y1 + y2 + 3y3 + 2y4 ≤ ξ2 − x2
donde 0 ≤ x1, x2 ≤ 5, yi {0, 1} (i = 1,
2, 3, 4) ξ1, ξ2 distribuidos uniformemente en {5, 5.5, . . . , 15}

Una solución a este problema consiste en asignaciones fijas a x1, x2 y asignaciones a yi que dependen de los valores aleatorios de ξ . La política óptima tiene un objetivo de 61,32.

Modelamos el problema como una IP de la siguiente manera. Las variables son

$$V = \quad x1, x2, \xi1, \xi2, y1, y2, y3, y4$$

los agentes son

$$A = \quad \text{optar, optar, oportunidad, oportunidad, optar, optar, optar, optar}$$

(elegimos el nombre y optamos por el agente no aleatorio), los dominios son

$$D = \quad D1, D1, D2, D2, D3, D3, D3, D3$$

donde D1 = {0, . . . , 5}, D2 = {5, 5,5, . . . , 15} y D3 = {0, 1}, P asigna probabilidades iguales a los valores del dominio ξi y O = .

Como en la Sección 3.1, modelamos las restricciones minimizando sus violaciones como un objetivo adicional, eligiendo una ponderación grande para priorizar la viabilidad. Sin embargo, en este problema también tenemos un objetivo, por lo que usaremos dos utilidades: la primera para minimizar el número de violaciones de restricciones a 0 y la segunda para maximizar z. Por tanto, U = viol, z . El hiperparámetro de utilidad es Y = −10, 0.99 , 100, 0.01 .

3.6.1. Un modelo disperso

En este IP, los enlaces son L = {ξi → yj | i, j}. En múltiples ejecuciones de hasta 107 episodios, INFPROG encontró políticas con una utilidad esperada de aproximadamente 55. Esto es bastante bueno, pero inferior al óptimo conocido de 61,32. Aumentar a 109 episodios no hizo ninguna diferencia.

3.6.2. Un modelo denso

En un intento por mejorar los resultados, agregamos los siguientes enlaces a la IP dispersa: x1 → x2, xi → yj (i, j} y yi → yj (i < j). En múltiples ejecuciones de 108 episodios, INFPROG encontró políticas con una recompensa esperada de aproximadamente 60, que está cerca del óptimo.

3.7. Programación restringida al azar

Una restricción de probabilidad (o probabilística) es una restricción que debe satisfacerse con algún umbral de probabilidad: Pr[C(x,r)] ≥ θ para una restricción C sobre las variables de decisión x y las variables de probabilidad r. La programación restringida al azar [53] es un método de optimización bajo incertidumbre y tiene muchas aplicaciones, ya que es una forma natural de modelar la incertidumbre. Las restricciones aleatorias suelen ser desigualdades, pero permitimos cualquier forma de restricción. Restricciones de probabilidad

Machine Translated by Google

También se han utilizado en el área de RL segura (o restringida) utilizando varios enfoques [54,55], y nuestro enfoque está relacionado con el de [3]. No parecen haber sido agregados a las identificaciones.

Modelamos restricciones de azar agregando un nuevo objetivo para cada uno, con una recompensa de 1 por satisfacción y 0 por violación, y estableciendo el valor utópico para ese objetivo en el umbral de probabilidad deseado. El peso fijado al objetivo debe ser alto.

Por ejemplo, modificamos el programa estocástico de la Sección 3.6 adjuntando probabilidades a las dos restricciones estrictas:

$$\Pr[2y1 + 3y2 + 4y3 + 5y4 \leq \xi 1 - x1] = 0{,}7 \text{ (a)}$$
$$\Pr[6y1 + y2 + 3y3 + 2y4 \leq \xi 2 - x2] = 0{,}9 \text{ (b)}$$

Observe que las restricciones de probabilidad son de la forma Pr[·] = p en lugar de la habitual Pr[·] ≥ p. No se garantiza que INFPROG satisfaga los umbrales de probabilidad debido a su enfoque multiobjetivo: en lugar de forzar la satisfacción de una restricción de probabilidad para exceder un umbral de probabilidad, intenta igualar el umbral en una compensación con otros objetivos. Sin embargo, se puede utilizar como herramienta exploratoria para encontrar una política con las características deseadas y el usuario puede aumentar los umbrales de forma iterativa.

Para modelar este problema, modificamos el modelo denso de la Sección 3.6, agregando un nuevo objetivo para cada restricción de probabilidad. El vector de utilidad ahora es U = viola, violb , z . El hiperparámetro de utilidad es

$$Y = \quad 0.7, 0.496 \quad , \quad 0.9, 0.496 \quad , \quad 100, 0.008$$

Para resolver este IP de tres objetivos, experimentamos con diferentes pesos y encontramos una variedad de políticas, con diferentes compromisos entre las restricciones de azar y el objetivo original. No todas las soluciones fueron útiles, pero utilizando ponderaciones (0,496, 0,496, 0,008), encontramos una política con objetivos (0,74, 0,95, 71) que satisface los requisitos. Relajar las restricciones estrictas a las restricciones de azar nos permitió aumentar el objetivo z de 60 a 71.

3.8. Programación multinivel

Muchos problemas en economía, diplomacia, guerra, política, industria, juegos y otras áreas involucran a múltiples agentes, que forman parte del entorno de cada uno. La RL multiagente se puede aplicar a estos problemas, al igual que varias formas de identificación: identificaciones biagente [7], identificaciones multiagente [8], identificaciones basadas en teoría de juegos [10], redes de identificaciones [6], e identificaciones dinámicas interactivas [9 Los LIMID pueden modelar múltiples agentes, pero solo el caso cooperativo en el que todos tienen el mismo objetivo.

El caso más común es un programa de dos niveles o un juego de Stackelberg. Generalmente involucran variables continuas y se ha realizado relativamente poco trabajo en programas discretos de dos niveles [56], por lo que nuevos métodos son de interés. Sin embargo, como caso aún más desafiante, utilizamos un programa discreto de tres niveles estudiado en al menos dos publicaciones [57,58] y que se muestra en la Figura 4.

$$\begin{aligned}
&\max_{x1} z1 = (x1 + x2 + 2x3 + 4)(-x1 - x2 + x3 + 2x4 + 1)\\
&\text{donde } x2\\
&\text{resuelve } \max_{x2} z2 = 2x2 +\\
&x3 + 3x4 \text{ donde } x3, x3, x3, x3, x3, x3 \text{ resuelve, para dado } x1, x2\\
&\max_{x3\ldots x8} z3 = \frac{2x1+3x2+2x3-3x4}{5x1+11x2+x5+29} \text{ sujeto}\\
&a\ -3x1 + 7x2\\
&\quad + x3 + x5 = 10\ 14x1 + 4x2 + x6 = 6\ x1 + x2 + x3 - x4 + x7 = 5\ 2x1 +\\
&\quad x2 + 2x4 + x8 = 8\ 5 \leq x3 \leq 20\ 4 \leq x4 \leq 30\ 0 \leq x1 \leq 5\ 0 \leq x2 \leq 2\ 1 \leq x5 \leq 20\\
&\qquad\qquad 0 \leq x6 \leq 30\ 0 \leq x7 \leq 15\ 0 \leq x8 \leq 20
\end{aligned}$$

Figura 4. Un programa no lineal de tres niveles.

Se han utilizado programas de tres niveles para modelar problemas en la gestión de la cadena de suministro, defensa de redes, planificación, logística y economía. Implican a tres agentes: el primero es el líder de nivel superior, el segundo es el seguidor de nivel medio y el tercero es el seguidor (de nivel inferior) . Están organizados jerárquicamente: el líder toma las decisiones primero, el

Machine Translated by Google

El seguidor de nivel medio reacciona, luego el seguidor de nivel inferior reacciona. En [59] se proporciona un estudio sobre la programación multinivel , que incluye una sección sobre el caso de tres niveles. Los problemas de toma de decisiones descentralizada en un sistema de gestión jerárquico a menudo contienen más de dos niveles. Como aplicación diferente, mencionamos [60] que utiliza programación de tres niveles para modelar un problema defensor-atacante-defensor en la defensa de una red eléctrica.

Este problema también es de interés porque sus objetivos son cuadráticos para el líder, lineales para el seguidor de nivel medio y fraccionarios para el seguidor de nivel inferior. Se agregaron objetivos no lineales a las ID en [5] utilizando aproximaciones de control óptimo no lineales.

Modelamos el problema usando la siguiente IP. Las variables son V = x1, . . . , x8 , y sus dominios son rangos de valores enteros D = 0 − 5, 0 − 2, 5 − 20, 4 − 30, 1 − 20, 0 − 30, 0 − 15, 0 − 20 . Los agentes son A = líder, medio, abajo ,. . . , abajo : v1 es una variable líder de nivel superior, v2 es una variable líder de nivel medio y v3–v8 son variables seguidoras.

Los vínculos son L = {xi → xj |i < j} (un modelo denso en el sentido de la Sección 3.1), las utilidades son

$$U = \text{violetas, z1} \quad , \quad \text{violetas, z2} \quad , \quad \text{violetas, z3}$$

y O = . El primer objetivo utilizado por todos los agentes es el número de violaciones de restricciones, que debe ser 0: las restricciones aseguran que cualquier solución sea factible en tres niveles, por lo que todos los agentes deben ser penalizados por violaciones (en este problema, ningún agente puede hacer una decisión que conduce a una violación de la restricción). Los objetivos secundarios son z1, z2, z3. Usando un hiperparámetro de utilidad

$$Y = \text{0, 0.99999} \quad , \quad \text{1000, 0.00001} \quad , \quad \text{0, 0.999} \quad , \quad \text{100, 0.001} \quad , \quad \text{0, 0.99} \quad , \quad \text{1, 0.01} \quad)$$

Este es el valor Y más complejo que utilizamos, por lo que proporcionamos una explicación. El problema de tres niveles se ha modelado utilizando tres agentes, cada uno con dos objetivos (uno para los objetivos del problema y otro para maximizar la satisfacción de las restricciones). Usamos escalarización métrica ponderada para reducir esto a tres agentes con un objetivo cada uno, por lo que debemos especificar tres puntos utópicos (uno por agente). Cada punto utópico tiene dos valores objetivos ideales y un peso asociado para cada valor. Entonces, por ejemplo, el agente 1 tiene un punto utópico 0, 0.99999 con pesos asociados 1000, 0.00001 .

En ejecuciones repetidas con 107 episodios, INFPROG encuentra de manera confiable la política correcta (0, 0, 9, 4, 1, 6, 0, 0) con objetivos ((0, 396), (0, 21), (0, 0.2)). Como se señala en [17], el primer objetivo se da como 612 en [58], pero se puede verificar que la solución produce z1 = 396. Las soluciones subóptimas que a veces encuentra se pueden reconocer por sus valores más bajos de z1 .

3.9. Juegos bayesianos

Los supuestos de la teoría de juegos clásica requieren información completa, que a menudo no es realista. Los juegos bayesianos fueron un desarrollo que permitió información incompleta (privada o secreta) evitando cálculos infinitos [61]. Una característica importante es que un jugador puede tener un "tipo" que otros jugadores no conocen, lo que representa su estado de ánimo.

Como ejemplo sencillo, tomemos un problema bien conocido llamado el dilema del Sheriff. Hay dos jugadores: un sheriff y un sospechoso armado. El sospechoso tiene dos tipos posibles, criminal o civil, y sólo él sabe de qué tipo es. El sheriff debe saber si debe disparar sin conocer el tipo de sospechoso, mientras que al sospechoso se le permite usar esa información. Las matrices de pagos para las diversas posibilidades se muestran en la Figura 5, donde en cada caso, la primera cifra es un pago sospechoso y la segunda un pago del sheriff.

Podemos modelar esto como una IP de la siguiente manera. Las variables son V = tipo,sospechoso,sheriff con dominios

$$D = \{\text{criminal, civil}\}, \{\text{disparar, no}\}, \{\text{disparar, no}\}$$

los agentes son A = oportunidad, sospechoso, sheriff . Los enlaces son L = {tipo → sospechoso}, lo que permite al sospechoso, pero no al sheriff, conocer el tipo de sospechoso. Las probabilidades de tipo P son p para Pr[criminal] y 1 − p para Pr[civil], las utilidades son U = ususpechoso , usheriff y O = .

Machine Translated by Google

Usando un hiperparámetro de utilidad Y = 100, 1 , 100, 1 y 1000 episodios, INF -PROG encuentra de manera confiable la política correcta: el sospechoso dispara si y solo si es un criminal, y el el sheriff dispara si y sólo si p > entonces se necesitan más episodios para reducir el error de muestreo.

del sheriff: disparo criminal	tipo de acción	
sospechosa disparo	0,0 2,−2	no −2,−1 −1,1

del sheriff: disparar a civiles	Tipo de acción	
acción sospechosa	disparar −3,−1 −1,−2	no −2,−1 0,0

Figura 5. Beneficios del dilema del sheriff.

3.10. Los juegos bayesianos

de razonamiento de nivel k también han sido criticados por hacer suposiciones poco realistas que no siempre predicen el comportamiento en el mundo real. Otro enfoque de la teoría de juegos para la información incompleta es el razonamiento de nivel k [62,63], que supone que los jugadores juegan estratégicamente pero con racionalidad limitada. Recientemente se ha utilizado en el análisis de riesgos adversarios para modelar problemas en la lucha contra el terrorismo [64] y otros campos.

Consideremos un problema simple de nivel k: el concurso de belleza keynesiano. Se pide a los participantes del concurso que elijan un número que esperan que sea lo más cercano posible a alguna fracción p del promedio de las elecciones de todos los participantes. La teoría de juegos clásica predice que todos los jugadores elegirán 0, que es el equilibrio de Nash. El razonamiento es que, si todos los jugadores eligen al azar, es mejor elegir p multiplicado por la media: 50p o aproximadamente 33 cuando p = (que es un $\frac{2}{3}$ valor popular). Pero todos los jugadores saben esto, por lo que deberían elegir p multiplicado por ese valor aproximadamente 25. Pero todos los jugadores saben esto, por lo que la conclusión lógica es que 0 es la única opción razonable; sin embargo, en los experimentos, pocas personas eligen 0.

En el razonamiento de nivel k, hacemos suposiciones sobre los otros jugadores. Podríamos suponer que todos los demás jugadores son pensadores de nivel 0 que no piensan estratégicamente. Simplemente eligen un número uniformemente al azar: c0 U[0, 100] (restringiremos nuestros números a números enteros ya que nuestro método es discreto). Bajo este supuesto, deberíamos usar el pensamiento de nivel 1 y elegir p multiplicado por la elección esperada del nivel 0: 33. Si en cambio asumimos que todos los demás jugadores son de nivel 1, deberíamos usar el pensamiento de nivel 2 y elegir 22. De manera similar, Si asumimos que todos los demás son de nivel 2, deberíamos usar el pensamiento de nivel 3 y elegir 15.

Podemos modelar y resolver estos problemas como IP, como ahora mostramos usando el concurso de belleza. Las variables representan las elecciones de los jugadores hipotéticos de nivel k (k = 0, 1, 2) y del jugador real de nivel 3: V = c0, c1, c2, c3 . Los dominios en D son todos {0, 1, 2, . . . , 100}. Los agentes son A = chance, P1, P2, P3 . Las probabilidades aleatorias son todas $\frac{1}{101}$. No hay vínculos ya que todas las elecciones se hacen simultáneamente, sin conocer las otras elecciones: L = . No hay observaciones entonces O = . El jugador de nivel i debe adivinar lo más cerca posible de p multiplicado por el valor que espera del jugador de nivel (i − 1) , es decir, |ci − pci−1 | debe minimizarse, por lo que las utilidades son |c1 − pc0| , |c2 − pc1| , |c3 − pc2| . Aplicando INFPROG con 106 episodios y un hiperparámetro de utilidad Y = 0, 1 , 0, 1 , 0, 1 para minimizar todas las utilidades, encontramos c1, tomando valores aproximadamente 29– 37, c2 19–25 y c3 13–15. El uso de más episodios reduce la variación en estos números.

3.11. Implicaciones prácticas

Hemos demostrado que una amplia gama de problemas de diversas áreas de la literatura se pueden modelar como IP y que estos se pueden resolver mediante el solucionador INFPROG . Por supuesto, no afirmamos que nuestro enfoque pueda abordar ningún problema de investigación de operaciones, teoría de juegos, etc. , especialmente porque actualmente solo puede manejar variables discretas.
Sin embargo, creemos que tiene implicaciones prácticas interesantes.

En primer lugar, nuestro enfoque es útil para la creación rápida de prototipos. Pocos investigadores dominan la programación estocástica, la programación dinámica, la teoría de juegos, el aprendizaje automático y muchas otras áreas. Cuando nos enfrentamos a un nuevo problema de optimización complejo, no siempre está claro cómo modelarlo y, en la práctica, a menudo hacemos suposiciones simplificadoras basadas en nuestros antecedentes. Un programador estocástico podría simplificar un problema de múltiples agentes

Machine Translated by Google

modelando a un adversario usando variables aleatorias, o pretender que las decisiones no pueden afectar las distribuciones de variables aleatorias, o aproximar un problema de múltiples etapas por uno de dos etapas. Un programador de restricciones o un programador de números enteros podría aproximar una variable aleatoria según una expectativa. Un teórico de juegos podría suponer un juego de suma cero o un conocimiento completo, incluso cuando esto no sea realista. En general, es difícil saber si nuestras elecciones de diseño son razonables. Creemos que una herramienta como INFPROG es útil durante las primeras etapas del modelado , ya que nos permite explorar las consecuencias de varias simplificaciones sin la necesidad de dominar múltiples áreas de investigación: simplemente podemos cambiar la formulación del problema y observar el resultado. . Algunas simplificaciones pueden causar pocos cambios en la calidad de la solución, mientras que otras pueden reducir en gran medida la calidad o incluso hacer que un problema sea inviable. Los modelos de programa de ocho reinas y estocástico pueden verse como ejemplos: en ambos casos, implementamos diferentes modelos y comparamos los valores objetivos resultantes, lo que llevó a la conclusión de que los modelos densos dan mejores resultados.

En segundo lugar, esperamos que nuestro enfoque sea útil por derecho propio. Se basa en RL multiagente y multiobjetivo, un paradigma muy general con amplias aplicaciones. Para problemas específicos, no esperamos que INFPROG sea competitivo con algoritmos especializados, pero para problemas que son inherentemente complejos, podría ser una herramienta útil, ya que hay poco software disponible capaz de abordar problemas que involucran múltiples agentes, múltiples objetivos y eventos aleatorios. y conocimiento parcial.

Cabe señalar que INFPROG es simplemente un prototipo de investigación y ciertamente no lo consideramos un producto terminado. Esperamos que en el futuro surjan algoritmos IP más competitivos, especialmente basados en RL profundo e implementados en hardware altamente paralelo. Esto debería mejorar en gran medida la escalabilidad del enfoque IP para problemas de optimización complejos, ya que RL ya ha demostrado su capacidad para abordar grandes problemas de múltiples agentes, como aprender a jugar Go.

3.12. Un pequeño estudio

de caso Para ilustrar nuestro enfoque con más detalle, mostramos la utilidad programable para el problema Oil Wildcatter de la Sección 3.3 en la Figura 6. El usuario debe proporcionar una función C llamada utilidad con dos argumentos: una matriz unidimensional v de número entero variables de probabilidad y decisión y una matriz bidimensional u de utilidades. Esta función se llama al final de un episodio, por lo que podemos suponer que a todas las variables en v se les han asignado valores. (Los * y ** son notaciones C que indican la dimensionalidad de la matriz).

```
utilidad nula (int* v, float** u) {

    int petróleo=v[0], prueba=v[1], sísmica=v[2], perforación=v[3], no=0,
        sí=1, seco=0, húmedo=1, remojo=2, empresa =1; flotador
    Drill_payoff, test_payoff;

    si (prueba==sí) test_payoff= –10.0; de lo contrario test_payoff=0.0;

    if (petróleo==seco && perforación==sí) Drill_payoff= –70.0; else if
    (petróleo==húmedo && taladro==sí) Drill_payoff= 50.0; else if (aceite==remojo
    && taladro==sí) Drill_payoff= 200.0; demás
                                        Drill_payoff= 0.0;

    u[empresa][0]=drill_payoff+test_payoff;
}
```

Figura 6. Utilidad programable para el ejemplo del petróleo salvaje.

Para facilitar la lectura, hemos creado variables enteras para nombrar las variables de IP (por ejemplo, oil=v[0]), sus valores posibles (por ejemplo, sí=0) y el único agente (compañía; tenga en cuenta que a las variables de probabilidad se les asigna el número de agente 0).). El resultado total es la suma de los resultados de la perforación y la prueba, y esta suma se asigna a la única utilidad u[1][0] en el problema. El segundo índice es 0 porque las utilidades de un agente están numeradas a partir de 0: para un problema bioobjetivo como el de la Sección 3.4, también necesitaríamos asignar un valor a u[empresa][1].

Machine Translated by Google

El usuario también debe proporcionar un pequeño archivo de texto que describa los parámetros del problema (número de agentes, variables, tamaños de dominio, etc.). No mostramos un ejemplo ya que INFPROG es actualmente un prototipo de investigación y carece de un lenguaje de especificación fácil de usar.

4. Conclusiones

Este artículo hace dos contribuciones. En primer lugar, define una nueva clase de problema de optimización llamado programa de influencia (IP) que es lo suficientemente general como para abarcar una amplia variedad de problemas de investigación de operaciones, inteligencia artificial, teoría de juegos y análisis de riesgos adversarios. En segundo lugar, propone un algoritmo llamado INFPROG basado en aprendizaje por refuerzo más muestreo. INFPROG es un solucionador liviano que no necesita una unidad de procesamiento de gráficos ni otro hardware especializado. Demostramos la flexibilidad de IP utilizando ejemplos tomados de diversas literaturas: hasta donde sabemos, ningún otro solucionador puede abordar esta gama de problemas de optimización. Nuestro trabajo también muestra que muchos problemas existentes de otras literaturas pueden servir como problemas de referencia para algoritmos de aprendizaje por refuerzo multiobjetivo y/o multiagente , de los cuales faltan [18].

Esperamos que las IP sean útiles para problemas de optimización híbrida que no encajan bien en ninguna clase en particular y no tienen solucionadores disponibles. Una fuente fructífera de aplicaciones podría ser el aprendizaje por refuerzo seguro, en el que se utilizan restricciones para garantizar políticas sólidas y seguras. También podría ser útil para la creación rápida de prototipos de diferentes modelos, para investigar cómo modelar un nuevo problema, para probar los efectos de ocultar información o hacer suposiciones simplificadoras. Finalmente, podría utilizarse como herramienta educativa para comprender los diferentes modelos de problemas de optimización.

En trabajos futuros, exploraremos la aplicación de nuestro método a problemas complejos del mundo real. La implementación actual de INFPROG es solo un prototipo de investigación, destinado a demostrar el potencial de un enfoque unificado para una amplia gama de problemas complejos de optimización. Sólo puede manejar variables discretas, utiliza una forma simple y bastante débil de agregación de estados (codificación aleatoria de mosaicos) y carece de una interfaz de usuario clara. Esperamos diseñar una versión más fácil de usar que también incluya variables continuas y utilizar una agregación de estados más poderosa: las dos últimas mejoras podrían lograrse pasando de SARSA a un algoritmo actor-crítico como la optimización de políticas proximales [65].

Financiamiento: Este material se basa en trabajos apoyados por la Science Foundation Ireland bajo la subvención No. 12/RC/2289-P2, que está cofinanciada por el Fondo Europeo de Desarrollo Regional. También nos gustaría agradecer el apoyo del Centro CONFIRM para la Fabricación Inteligente de la Science Foundation Ireland, Código de Investigación 16/RC/3918, y el proyecto TAILOR ICT48 H2020 de la UE bajo el contrato n.º 952215.

Declaración de la Junta de Revisión Institucional: No aplicable.

Declaración de Consentimiento Informado: No aplicable.

Declaración de disponibilidad de datos: No se crearon ni analizaron nuevos datos en este estudio. El intercambio de datos no se aplica a este artículo.

Conflictos de intereses: Los financiadores no tuvieron ningún papel en el diseño del estudio; en la recopilación, análisis o interpretación de datos; en la redacción del manuscrito; o en la decisión de publicar los resultados.

Referencias

1. Sutton, RS; Barto, AG Aprendizaje por refuerzo: una introducción; Prensa del MIT: Cambridge, MA, EE. UU., 1998. 2.
Plata, D.; Hubert, T.; Schrittwieser, J.; Antonoglou, I.; Lai, M.; Guez, A.; Lanctot, M.; Sifré, L.; Kumaran, D.; Graepel, T.; Lillicrap, T.; Simonyan, K.; Hassabis, D. Un algoritmo general de aprendizaje por refuerzo que domina el ajedrez, el shogi y el juego autónomo. Ciencia 2018, 362, 1140–1144. [Referencia cruzada] [PubMed]
3. Huang, SH; Abdolmaleki, A.; Vezzani, G.; Brakel, P.; Mankowitz, DJ; Neunert, M.; Bohez, S.; Tasa, Y.; Heess, N.; Riedmiller, MA; Hadsell, R. Un marco de aprendizaje por refuerzo multiobjetivo restringido. En Actas de la Quinta Conferencia sobre Aprendizaje de Robots, Londres, Reino Unido, 8 a 11 de noviembre de 2021; págs. 883–893.
4. Elshafei, MMK; El-Sherberry, MS Problema interactivo de programación lineal entera estocástica multiobjetivo de dos niveles. Aplicación de tendencias . Ciencia. Res. 2008, 3, 154–164.
5. Kratochvíl, V.; Vomlel, J. Diagramas de influencia para la optimización del perfil de velocidad. Int. J. Aprox. Razón. 2017, 88, 567–586. [Referencia cruzada]

Machine Translated by Google

6. Gal, Y.; Pfeffer, A. Diagramas de redes de influencia: un formalismo para representar las creencias de los agentes y los procesos de toma de decisiones. J. Artif. Intel. Res. 2008, 33, 109–147. [Referencia cruzada]
7. González-Ortega, J.; Insúa, RD; Cano, J. Análisis de riesgo adversario para diagramas de influencia de dos agentes: un enfoque algorítmico. euros. J. Ópera. Res. 2019, 273, 1085–1096. [Referencia cruzada]
8. Koller, D.; Milch, B. Diagramas de influencia de múltiples agentes para representar y resolver juegos. Economía de juegos. Comportamiento. 2001, 45, 181–221. [Referencia cruzada]
9. Polich, K.; Gmytrasiewicz, G. Diagramas de influencia dinámica interactiva. En Actas de la Sexta Conferencia Internacional Conjunta sobre Agentes Autónomos y Sistemas Multiagentes, Comunicaciones en Ciencias de la Información y la Computación, Honolulu, Hawaii, EE. UU., 14 a 18 de mayo de 2007; Volumen 288, págs. 623–630.
10. Zhou, LH; Kevin, L.; Liu, WY Diagramas de influencia basados en la teoría de juegos. Sistema experto. 2013, 30, 341–351. [Referencia cruzada]
11. Diehl, M.; Haimes, Y. Diagramas de influencia con múltiples objetivos y análisis de compensaciones. Traducción IEEE. Sistema. Hombre Cibernético. Parte A 2004, 34, 293–304. [Referencia cruzada]
12. Marinescu, R.; Razak, A.; Wilson, N. Diagramas de influencia multiobjetivo. En Actas de la 28.ª Conferencia sobre la incertidumbre en la inteligencia artificial, Isla Catalina, California, EE. UU., 5 al 17 de agosto de 2012; págs. 574–583.
13. Jenzarli, A. Diagramas de influencia de información/relevancia. En actas de la 11.ª Conferencia sobre incertidumbre en el ámbito artificial. Intelligence (UAI), Montreal, QC, Canadá, 18 a 20 de agosto de 1995; págs. 329–337.
14. Lauritzen, SL; Nilsson, D. Representación y resolución de problemas de decisión con información limitada. Gestionar. Ciencia. 2001, 47, 1238-1251. [Referencia cruzada]
15. Powell, Aprendizaje por refuerzo y optimización estocástica del Banco Mundial: un marco unificado para decisiones secuenciales; Wiley: Hoboken, Nueva Jersey, Estados Unidos, 2022.
16. Prestwich, SD; Toffano, F.; Wilson, N. Un lenguaje de programación probabilístico para diagramas de influencia. En Actas de la XI Conferencia Internacional sobre Gestión Escalable de la Incertidumbre, Granada, España, 4 al 6 de octubre de 2017.
17. Prestwich, SD Resolución de diagramas de influencia mixta mediante aprendizaje por refuerzo. En actas de la novena conferencia internacional sobre aprendizaje automático, optimización y ciencia de datos, Grasmere, Reino Unido, 22 a 26 de septiembre de 2023.
18. Hayes, CF; R˘adulescu, R.; Bargiacchi, E.; Källström, J.; Macfarlane, M.; Reymond, M.; Verstraeten, T.; Zintgraf, LM; Dazeley, R.; Heintz, F.; et al. Una guía práctica para la planificación y el aprendizaje por refuerzo multiobjetivo. Auton. Agente. Sistema multiagente. 2022, 36, 26. [Referencia cruzada]
19. van Moffaert, K.; Drugan, MM; Nowé, A. Aprendizaje por refuerzo escalarizado multiobjetivo: nuevas técnicas de diseño. En Actas del Simposio IEEE sobre programación dinámica adaptativa y aprendizaje por refuerzo, Singapur, 16 a 19 de abril de 2013; págs. 191-199.
20. Zhang, K.; Yang, Z.; Ba¸sar, T. Aprendizaje por refuerzo de agentes múltiples: una descripción general selectiva de teorías y algoritmos. En Manual de control y aprendizaje por refuerzo; Vamvoudakis, KG, Wan, Y., Lewis, FL, Cansever, D., Eds.; Estudios en Sistemas, Decisión y Control; Springer: Cham, Suiza, 2021; Volumen 325.
21. R˘adulescu, R.; Manion, P.; Roijers, DM; Nowé, A. Toma de decisiones multiobjetivo y multiagente: un análisis basado en la utilidad y Encuesta. Auton. Agentes Sistema multiagente. 2020, 34, 10. [Referencia cruzada]
22. Hu, T.; Luo, B.; Yang, C.; Huang, T. MO-MIX: Toma de decisiones cooperativa de múltiples agentes y objetivos múltiples con refuerzo profundo. Aprendizaje mental. Traducción IEEE. Patrón Anal. Mach. Intel. 2023, 45, 12098–12112. [Referencia cruzada] [PubMed]
23. Andrychowicz, M.; Wolski, F.; Rayo, A.; Schneider, J.; Fong, R.; Welinder, P.; McGrew, B.; Tobin, J.; Abbeel, P.; Zaremba, W. Repetición de la experiencia en retrospectiva. En Actas de la 31.ª Conferencia Internacional sobre Sistemas de Procesamiento de Información Neural, Long Beach, CA, EE. UU., 4 a 9 de diciembre de 2017; págs. 5055–5065.
24. Zobrist, AL Un nuevo método de hash con aplicación para juegos. ICGA J. 1990, 13, 69–73.
25. Hyatt, RM; Cozzie, A. El efecto de las colisiones de firmas hash en un programa de ajedrez. ICGA J. 2005, 28, 131-139. [Referencia cruzada]
26. Skálse, J.; Hammond, L.; Grifo, C.; Abate, A. Aprendizaje por refuerzo lexicográfico multiobjetivo. En Actas del 31 Conferencia internacional conjunta sobre inteligencia artificial, Viena (Austria), 23 a 29 de julio de 2022; págs. 3430–3436.
27. Gabor, Z.; Kalmar, Z.; Szepesvári, C. Aprendizaje por refuerzo multicriterio. En actas de la 15ª Conferencia Internacional sobre Machine Learning, Madison, WA, EE.UU., 24 a 27 de julio de 1998; págs. 197-205.
28. Giagkiozis, I.; Fleming, PJ Métodos para la optimización multiobjetivo: un análisis. inf. Ciencia. 2015, 293, 338–350. [Referencia cruzada]
29. Gordon, AD; Henzinger, TA; Nori, AV; Rajamani, SK Programación probabilística. En Actas de la Internacional Conferencia sobre ingeniería de software, Hyderabad, India, 31 de mayo a 7 de junio de 2014.
30. Pfeffer, A. Programación probabilística práctica; Publicaciones de Manning: Shelter Island, Nueva York, EE. UU., 2016.
31. Rossi, F.; van Beek, P.; Walsh, T. (Eds.) Manual de programación de restricciones; Elsevier: Ámsterdam, Países Bajos, 2006.
32. Crites, RH; Barto, AG Control de grupos de ascensores mediante múltiples agentes de aprendizaje por refuerzo. Mach. Aprender. 1998, 33, 235–262. [Referencia cruzada]
33. Useng, M.; Avdulrahman, S. Una encuesta sobre aprendizaje por refuerzo distribuido. Mesopotamia J. Big Data 2022, 2022, 44–50. [Referencia cruzada]
34. Pearl, J. Razonamiento probabilístico en sistemas inteligentes: redes de inferencia plausible; Serie en Representación y Razonamiento; morgan Editores Kaufman: Burlington, MA, EE. UU., 1988.
35. Howard, RA; Matheson, JE Diagramas de influencia. En Lecturas en Análisis de Decisiones; Grupo de Decisiones Estratégicas: Menlo Park, CA, EE.UU., 1981; Capítulo 38, págs. 763–771.

Machine Translated by Google

36. Nease, RF, Jr.; Owens, DK Uso de diagramas de influencia para estructurar decisiones médicas. Medicina. Decide. Mak. 1997, 17, 263–275. [Referencia cruzada] [PubMed]
37. Chockalingam, S.; Maathuis, C. Diagramas de influencia en seguridad cibernética: conceptualización y aplicaciones potenciales. En Actas de la 22.ª Conferencia Europea sobre Seguridad y Guerra Cibernética, El Pireo (Grecia), 22 y 23 de junio de 2023; Volumen 22.
38. Carriger, JF; Newman, MC Diagramas de influencia como herramientas de toma de decisiones para la gestión del riesgo de plaguicidas. Integral Reinar. Evaluar. Gestionar. 2011, 8, 339–350. [Referencia cruzada]
39. Dechter, R. Una nueva perspectiva sobre algoritmos para optimizar políticas en condiciones de incertidumbre. En Sistemas de Planificación de Inteligencia Artificial; Asociación para el Avance de la Inteligencia Artificial: Washington, DC, EE.UU., 2000; págs. 72–81.
40. Jensen, F.; Jensen, V.; Dittmer, S. De los diagramas de influencia a los árboles de unión. En Incertidumbre en la Inteligencia Artificial; Morgan Kaufmann: Burlington, MA, EE.UU., 1994; págs. 367–373.
41. Shachter, RD Evaluación de diagramas de influencia. Ópera. Res. 1986, 34, 871–882. [Referencia cruzada]
42. Shenoy, P. Sistemas basados en valoración para el análisis de decisiones bayesiano. Ópera. Res. 1992, 40, 463–484. [Referencia cruzada]
43. Tatman, JA; Shachter, RD Programación dinámica y diagramas de influencia. Traducción IEEE. Sistema. Hombre Cibernético. 1990, 20, 365–379. [Referencia cruzada]
44. Canó, A.; Gómez, M.; Moral, S. Un método de Monte Carlo hacia adelante y hacia atrás para resolver diagramas de influencia. Int. J. Aprox. Razón. 2006, 42, 119-135. [Referencia cruzada]
45. Charnes, JM; Shenoy, PP Método Monte Carlo de varias etapas para resolver diagramas de influencia mediante computación local. Gestionar. Ciencia. 2004, 50, 405–418. [Referencia cruzada]
46. Marinescu, R.; Lee, J.; Dechter, R. Un nuevo esquema delimitador para diagramas de influencia. En Actas de la 35.ª Conferencia sobre Inteligencia Artificial, Vancouver, BC, Canadá, 2 a 9 de febrero de 2021; Asociación para el Avance de la Inteligencia Artificial: Washington, DC, EE. UU., 2021; págs. 12158-12165.
47. Watthayu, W. Representación y resolución de diagramas de influencia en la toma de decisiones con criterios múltiples: un método de propagación de creencias descabellado. En Actas del Simposio Internacional sobre Ciencias de la Computación y sus Aplicaciones, Hobart, TAS, Australia, 13 a 15 de octubre de 2008; págs. 118-125.
48. Yuan, C.; Wu, X. Resolución de diagramas de influencia mediante búsqueda heurística. En actas del Simposio Internacional sobre Artificial Inteligencia y Matemáticas, Sanya, China, 23 y 24 de octubre de 2010.
49. Everitt, T.; Kumar, R.; Krakovna, V.; Legg, S. Modelado de marcos de seguridad AGI con diagramas de influencia causal. En Actas del Taller sobre Seguridad de la Inteligencia Artificial, Taller CEUR, Honolulu, HI, EE. UU., 27 de enero de 2019; Volumen 2419.
50. Raiffa, H. Análisis de decisiones; Addison-Wesley: Reading, MA, EE. UU., 1968.
51. Birge, JR; Louveaux, FV Introducción a la programación estocástica; Springer: Nueva York, NY, Estados Unidos, 2011.
52. Ahmed, S.; Tawarmalani, M.; Sahinidis, NV Un algoritmo finito de rama y límite para programas enteros estocásticos de dos etapas. Matemáticas. Programa. 2004, 100, 355–377. [Referencia cruzada]
53. Charnes, A.; Cooper, WW Programación restringida al azar. Gestionar. Ciencia. 1959, 6, 73–79. [Referencia cruzada]
54. García, J.; Fernández, F. Una encuesta integral sobre el aprendizaje por refuerzo seguro. J. Mach. Aprender. Res. 2015, 16, 1437–1480.
55. Gu, S.; Yang, L.; Du, Y.; Chen, G.; Walter; Wang, J.; Yang, Y.; Knoll, AC Una revisión del aprendizaje por refuerzo seguro: métodos, teoría y aplicaciones. arXiv 2022, arXiv:2205.10330.
56. Kovacs, A.; Kis, T. Enfoque de programación con restricciones para un problema de programación binivel. Restricciones 2011, 16, 317–340. [Referencia cruzada]
57. Arora, R.; Arora, SR Un algoritmo para problemas de programación entera multinivel no lineal. Int. J. Computación. Ciencia. Matemáticas. 2010, 3, 211–225. [Referencia cruzada]
58. Mishra, S.; Verma, AB Un enfoque no diferencial para resolver problemas de programación de tres niveles. Soy. Int. J.Res. Ciencia. Tecnología. Matemáticas. 2015.
59. Lu, J.; Han, J.; Hu, Y.; Zhang, G. Toma de decisiones multinivel: una encuesta. inf. Ciencia. 2016, 346–347, 463–487. [Referencia cruzada]
60. Alguacil, N.; Delgadillo, A.; Arroyo, JM Un enfoque de programación de tres niveles para la planificación de la defensa de la red eléctrica. Computadora. Ópera. Res. 2014, 41, 282–290. [Referencia cruzada]
61. Harsanyi, JC Juegos con información incompleta jugados por jugadores "bayesianos", I-III: Parte I. El modelo básico. Gestionar. Ciencia. 2004, 50, 1804–1817. [Referencia cruzada]
62. Nagel, R. Desentrañar en juegos de adivinanzas: un estudio experimental. Soy. Economía. Rev. 1995, 85, 1313–1326.
63. Stahl, DO; Wilson, PW sobre los modelos de otros jugadores: teoría y evidencia experimental. Economía de juegos. Comportamiento. 1995, 10, 218–254. [Referencia cruzada]
64. Rothschild, C.; Alberto, L.; Guikema, S. Análisis de riesgo adversario con información incompleta: un enfoque de nivel k. Anal de riesgo. 2011, 32, 1219–1231. [Referencia cruzada]
65. Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; Klimov, O. Algoritmos de optimización de políticas próximas. arXiv 2017, arXiv:1707.06347.

Descargo de responsabilidad/Nota del editor: Las declaraciones, opiniones y datos contenidos en todas las publicaciones son únicamente de los autores y contribuyentes individuales y no de MDPI ni de los editores. MDPI y/o los editores renuncian a toda responsabilidad por cualquier daño a personas o propiedad que resulte de cualquier idea, método, instrucción o producto mencionado en el contenido.

Machine Translated by Google

AppliedMath

MDPI

Artículo

Análisis de datos multidimensionales del lenguaje profundo en JRR Tolkien y CS Lewis revelan estrechas conexiones matemáticas

Emilio Matricciani

Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB), Politecnico di Milano, 20133 Milán, Italia; emilio.matricciani@polimi.it

Resumen: Los estudiosos de la literatura inglesa dicen unánimemente que JRR Tolkien influyó en los escritos de CS Lewis . Por primera vez, hemos investigado este tema matemáticamente utilizando un análisis multidimensional original de parámetros lingüísticos, basado en variables superficiales profundas del lenguaje y canales lingüísticos. Para situar nuestra investigación en el marco de la literatura inglesa, hemos considerado algunas novelas escritas por autores anteriores, como C. Dickens, G. MacDonald y otros. Las variables profundas del lenguaje y los canales lingüísticos, discutidos en el artículo, probablemente se deban al diseño inconsciente de los escritores y revelan conexiones entre textos que van mucho más allá de la conciencia de los escritores. En resumen, la capacidad de la memoria extendida a corto plazo requerida por los lectores, el índice de legibilidad universal de los textos, la representación geométrica de los textos y el ajuste fino de los canales lingüísticos dentro de los textos (todas herramientas ampliamente discutidas en el artículo) revelaron fuertes conexiones entre El Señor de los Anillos (Tolkien), Las Crónicas de Narnia, La Trilogía Espacial (Lewis) y novelas de MacDonald, coincidiendo por tanto con lo que dicen los estudiosos de la Literatura Inglesa.

Palabras clave: lenguas alfabéticas; memoria extendida a corto plazo; comunicación humana; mente humana ; oraciones: modelado matemático; índice de legibilidad universal

check for updates

Cita: Matricciani, E. El análisis de datos multidimensionales del lenguaje profundo en JRR Tolkien y CS Lewis revela estrechas conexiones matemáticas. Matemáticas aplicadas 2024, 4, 927–949. https://doi.org/10.3390/appledmath4030050

Editor académico: Libor Peka˘r

Recibido: 17 de junio de 2024
Revisado: 29 de julio de 2024
Aceptado: 31 de julio de 2024
Publicado: 1 de agosto de 2024

Copyright: © 2024 por el autor. Licenciatario MDPI, Basilea, Suiza. Este artículo es un artículo de acceso abierto. distribuido bajo los términos y condiciones de los Creative Commons Licencia de atribución (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

1. Introducción

Unánimemente, en un gran número de artículos (algunos de los cuales se recuerdan aquí [1–8] de la vasta literatura sobre el tema), los estudiosos de la literatura inglesa afirman que JRR Tolkien influyó en los escritos de CS Lewis. El propósito del presente artículo no es revisar la gran riqueza de literatura basada en el enfoque típico utilizado por los estudiosos de la literatura (que no es nuestra especialidad), sino investigar este tema matemática y estadísticamente (un estudio que nunca antes se ha realizado). mediante el uso de métodos recientes ideados al investigar el impacto de las variables superficiales del lenguaje profundo [9,10] y los canales lingüísticos [11] en los textos literarios. Dado que los estudiosos mencionan la influencia de George MacDonald en ambos, consideramos algunas novelas escritas por este autor anterior. Para situar todas estas novelas en el marco de la literatura inglesa, consideramos algunas novelas escritas por otros autores anteriores, como C. Dickens y otros.

Tras esta introducción, en el apartado 2, presentamos los textos literarios (novelas) considerados. En la Sección 3, informamos la serie de palabras, oraciones e interpunciones versus capítulos de algunas novelas, y definimos un índice útil para describir sintéticamente la regularidad debido a lo que pensamos es un diseño consciente de los autores. En la Sección 4, comenzamos a explorar las cuatro variables profundas del lenguaje; Para evitar malentendidos, estas variables y los canales lingüísticos derivados de ellas se refieren a la estructura “superficial” de los textos, no a la estructura “profunda” mencionada en la teoría cognitiva. En la Sección 5, informamos resultados relacionados con la memoria extendida a corto plazo y un índice de legibilidad universal; Ambos temas abordan los buffers de la memoria humana a corto plazo. En la Sección 6, representamos textos literarios geométricamente en el plano cartesiano definiendo combinaciones lineales de variables lingüísticas profundas y calculamos la probabilidad de que un texto pueda confundirse con otro. En la Sección 7, mostramos las relaciones lineales existentes entre variables lingüísticas en las novelas consideradas. En la Sección 8,

Machine Translated by Google

Presentamos la teoría de los canales lingüísticos. En la Sección 9, lo aplicamos a las novelas actualmente estudió. Finalmente, en la Sección 10, resumimos los principales hallazgos y concluimos. Varios Los apéndices reportan datos numéricos.

2. Base de datos de Textos Literarios (Novelas)

Primero introduzcamos la base de datos de textos literarios utilizados en el presente artículo. Tabla 1 enumera algunas estadísticas básicas de las novelas de Tolkien, Lewis y MacDonald. Para configurar estos textos en el marco de la literatura inglesa anterior, consideramos las novelas de Charles Dickens (Tabla 2) y otros autores (Tabla 3).

Tabla 1. Novelas escritas por Tolkien, Lewis y MacDonald, con año de publicación. Número de capítulos (es decir, el número de muestras consideradas al calcular las líneas de regresión que se informan a continuación), número total de caracteres contenidos en las palabras (C), número total de palabras W y oraciones (S). Se han eliminado títulos, notas a pie de página y otro material superfluo presente en los textos digitales.

John R. R. Tolkien (1892-1973)	Capítulos (M)	Personajes (C)	Palabras (W)	Oraciones (S)
El hobbit (1937)	19	394,154	95,914	5890
El señor de los anillos (1954-1955)	62	1.906.531	472,173	34.601
El Silmarillion (póstumo, 1977)	24	429,639	101.627	3346
Clive S. Lewis (1898-1963)				
Las cartas de cinta adhesiva (1942)	31	135.204	31.040	1330
La trilogía espacial (1938-1945)	123	1.243.141	295,240	20,124
Las crónicas de Narnia (1950-1956)	110	1.318.482	322,544	23.515
George MacDonald (1824-1905)				
Phantastes: Un romance de hadas para hombres y mujeres (1858)	25	283.676	67.551	3274
A espaldas del viento del norte (1871)	38	349.041	90.697	5017
La princesa y el duende (1872)	32	208,325	51.090	3205
Lilith: un romance (1895)	47	386,522	94.127	6271

Tabla 2. Novelas de Charles Dickens, con año de publicación. Número de capítulos (M, es decir, el número de muestras consideradas al calcular las líneas de regresión que se informan a continuación), número total de caracteres contenidos en las palabras (C), número total de palabras W y oraciones (S).

Novela (Año de publicación)	Capítulos (M)	Personajes (C)	Palabras (W)	Oraciones (S)
Las aventuras de Oliver Twist (1837-1839)	53	679.008	160.604	9121
David Copperfield (1849-1850)	64	1.469.251	363,284	19.610
Casa desolada (1852-1853)	64	1.480.523	350.020	20,967
Una historia de dos ciudades (1859)	45	607.424	142.762	8098
Nuestro amigo mutuo (1864-1865)	67	1.394.753	330.593	17.409

Hemos utilizado el texto digital de una novela (archivo WinWord) y contamos, para cada capítulo, el número de caracteres, palabras, oraciones e interpunciones (signos de puntuación). Antes Al hacerlo, hemos eliminado los títulos, notas a pie de página y otro material extraño presente en los textos digitales, un trabajo pesado. El recuento es muy sencillo, aunque requiere mucho tiempo. Winword proporciona directamente la cantidad de caracteres y palabras. El número de frases se calculó usando WinWord para reemplazar cada punto con un punto: por supuesto, esta acción no cambia el texto, pero da el número de estas sustituciones y por lo tanto, el número de puntos. Se repitió el mismo procedimiento para los signos de interrogación. y signos de exclamación. La suma de los tres totales da el número total de oraciones. en el texto analizado. El mismo procedimiento da el número total de comas, dos puntos y punto y coma. La suma de estos últimos valores con el número total de oraciones da el número total de interpunciones.

Machine Translated by Google

Tabla 3. Novelas por autores de literatura inglesa, con año de publicación. Número de capítulos (M, es decir, el número de muestras consideradas al calcular las líneas de regresión que se informan a continuación), total número de caracteres contenidos en las palabras (C), número total de palabras W y oraciones (S).

Novela (Autor, Año)	Capítulos (M)	Personajes (C)	Palabras (W)	Oraciones (S)
Orgullo y prejuicio (J. Austen, 1813)	61	537.005	121,934	6013
Feria de las vanidades (W. Thackeray, 1847–1848)	66	1.285.688	277,716	13.007
Moby Dick (H. Melville, 1851)	132	92,2351	203,983	9582
El molino sobre el hilo dental (G. Eliot, 1860)	57	888.867	207,358	9018
Las aventuras de Alicia en el país de las maravillas (L. Carroll, 1865)	12	107.452	27.170	1629
Aventuras de Huckleberry Finn (M. Twain, 1884)	42	427,473	110.997	5887
El libro de la selva (R. Kipling, 1894)	9	209,935	51.090	3214
La guerra de los mundos (HG Wells, 1897)	27	265.499	60.556	3306
El maravilloso mago de Oz (LF Baum, 1900)	22	156.973	39.074	2219
El sabueso de los Baskerville (AC Doyle, 1901-1902)	15	245,327	59.132	4080
Peter Pan (JM Barrie, 1902)	17	194,105	47.097	31,77
Martin Eden (J. Londres, 1908-1909)	45	601.672	139,281	9173

Se puede notar cierta homogeneidad en las novelas del mismo autor. Las historias en El La Trilogía Espacial y Las Crónicas de Narnia, de Lewis, se cuentan con aproximadamente el mismo número de capítulos, palabras y frases, como también ocurre con un par de novelas de MacDonald, como At la espalda del viento del norte y Lilith: un romance. Se puede encontrar cierta homogeneidad en David. Copperfield, Bleak House y Our Mutual Friend (de Dickens) y en Las aventuras de Oliver Twist y Historia de dos ciudades. Creemos que estos valores numéricos no se deben al azar sino gestionado conscientemente por los autores, que es un tema que profundizaremos en la siguiente sección.

3. Diseño Consciente de Textos: Palabras, Oraciones e Interpunciones versus Capítulos

En primer lugar, estudiamos las variables lingüísticas que creemos que los autores diseñaron deliberadamente. En los detalles, mostramos la serie de palabras, oraciones e interpunciones. versus capítulo.

Consideremos una obra literaria (una novela) y su subdivisión en bloques inconexos. de texto lo suficientemente largo como para dar valores promedio confiables. Sea nS el número de oraciones contenido en un bloque de texto, nW el número de palabras contenidas en las oraciones nS , nC el número de caracteres contenidos en las nW palabras y nI el número de signos de puntuación (interpunciones) contenidas en las oraciones nS .

10/110 = 0.09

La Figura 1 muestra la serie nW versus el número de capítulo normalizado de El Señor de la Anillos, Las Crónicas de Narnia, La Trilogía Espacial.

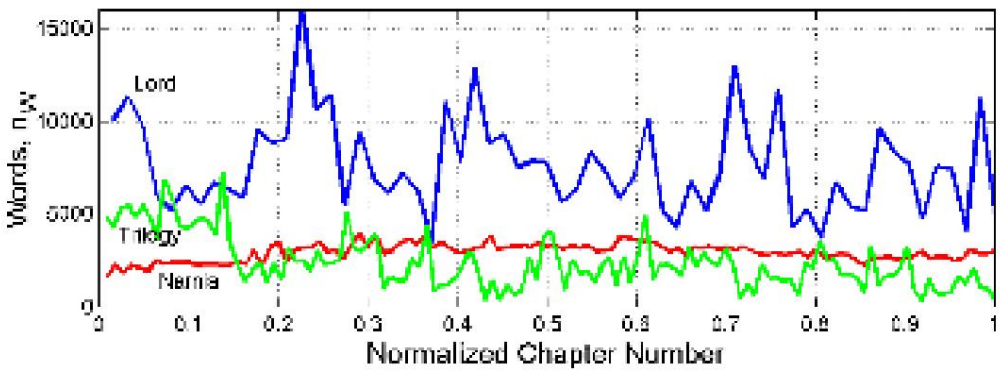


Figura 1. Serie de palabras versus número de capítulo normalizado. Línea azul: El Señor de los Anillos (Caballero); línea roja: Las Crónicas de Narnia (Narnia); Línea verde: La Trilogía Espacial (Trilogía).

Machine Translated by Google

Por ejemplo, el valor normalizado del capítulo 10 de Las Crónicas de Narnia, es 10/110 = 0,09 en la escala x de la Figura 1. Esta normalización permite la muestra sinóptica de novelas con un diferente número de capítulos.

En Las Crónicas de Narnia (en adelante, Narnia, para abreviar), podemos notar una valor prácticamente constante nW en comparación con El Señor de los Anillos (Lord) y El Espacio Trilogía (Trilogía).

Definamos un índice sintético para describir la serie dibujada en la Figura 1, a saber el coeficiente de variación δ, dado por la desviación estándar σnW dividida por la media valor < nW >

$$\delta = \frac{\sigma nW}{< nW >} \quad (1)$$

Las tablas 4 y 5 informan δ para nW, nS y nI . Dado que nS y nI están muy bien correlacionados con nS, los tres coeficientes de dispersión son aproximadamente iguales.

Tabla 4. El coeficiente de dispersión en las series de palabras, oraciones e interpunciones en la novelas indicadas de Tolkien, Lewis y MacDonald.

Novedoso	Palabras	Oraciones	Interpunciones	Promedio
El hobbit	0,49	0,48	0,50	0,49
El Señor de los Anillos	0,34	0,36	0,34	0,35
El Silmarillion	0,73	0,86	0,80	0,80
Las cartas de cinta de rosca	0,07	0,17	0,14	0,13
La trilogía espacial	0,60	0,61	0,58	0,59
Las crónicas de Narnia	0,16	0,20	0,20	0,19
A espaldas del viento del norte	0,54	0,61	0,56	0,57
Phantastes: un romance de hadas para hombres y mujeres	0,66	0,73	0,63	0,67
Lilith: un romance	0,43	0,53	0,46	0,47
La princesa y el duende	0,53	0,75	0,62	0,64

Tabla 5. El coeficiente de dispersión en las series de palabras, oraciones e interpunciones en la novelas indicadas.

Novedoso	Palabras	Oraciones	Interpunciones	Promedio
Las aventuras de Oliver Twist	0,31	0,33	0,32	0,32
David Copperfield	0,37	0,38	0,37	0,37
Casa sombría	0,28	0,31	0,30	0,30
Una historia de dos ciudades	0,52	0,57	0,52	0,54
Nuestro amigo mutuo	0,26	0,29	0,27	0,27
Martín Edén	0,29	0,33	0,29	0,31
El sabueso de los Baskerville	0,26	0,29	0,25	0,27
Peter Pan	0,29	0,41	0,33	0,34

En Narnia δ = 0,16, en Lord δ = 0,34 y en Trilogy δ = 0,60. Notemos también el valor mínimo δ = 0,07 en The Screwtape Letters (Cinta de rosca).

El valor medio general (palabras, oraciones e interpunciones mezcladas) es < δ >= 0,44 y la desviación estándar σδ = 0,18. Por tanto, Screwtape es prácticamente más que 2 × σδ de la media, como también lo está Silmarillion en el otro lado, y Narnia está aproximadamente 1,5 × σδ . En cambio, Trilogía, El Señor y El Hobbit (Hobbit) están dentro de 1 × σδ .

A partir de estos resultados, parece que Lewis diseñó los capítulos de Narnia y Screwtape. con una distribución casi uniforme de palabras, oraciones e interpunciones, muy probablemente

Machine Translated by Google

debido a la audiencia prevista en Narnia (es decir, niños) y la herramienta de ficción "cartas" utilizada en Screwtape. En Trilogía el diseño parece muy diferente (δ = 0,60, muy dentro de 1 × σδ) probablemente debido al desarrollo de la historia de ciencia ficción narrada.

Tolkien actuó de manera diferente a Lewis, porque parece haber diseñado los capítulos de manera más aleatoria y dentro de 1 × σδ , como muestran Hobbit y Lord . Una excepción es El Silmarillion, publicado póstumamente, que es un texto lejos de ser una "novela".

Finalmente, observe que las novelas de MacDonald muestran valores más homogéneos, muy similares a Hobbit y Trilogía y al resto de novelas enumeradas en la Tabla 5.

En conclusión, el análisis de series de palabras, oraciones e interpunciones por capítulo no indica conexiones probables entre Tolkien, Lewis y MacDonald. Cada autor estructuró su uso de palabras, oraciones y puntuación según planes distintos, que variaban no sólo entre autores sino también entre diferentes novelas del mismo autor.

Sin embargo, hay variables lingüísticas que –como hemos informado para los textos literarios modernos y antiguos– no son diseñadas/gestionadas conscientemente por los autores; por lo tanto, estas variables son las mejores candidatas para revelar conexiones matemáticas/estadísticas ocultas entre textos. En la siguiente sección, comenzamos a tratar estas variables, con el propósito específico de comparar a Tolkien y Lewis, aunque esta comparación se sitúa en el marco más general de los autores mencionados en la Sección 2.

4. Variables superficiales del lenguaje profundo

Comenzamos explorando las cuatro variables estocásticas que llamamos variables de lenguaje profundo, siguiendo nuestra teoría estadística general sobre lenguajes alfabéticos [9-11]. Para evitar posibles malentendidos, estas variables, y los canales lingüísticos que de ellas se derivan, se refieren a la estructura "superficial" de los textos, no a la estructura "profunda" mencionada en la teoría cognitiva.

Al contrario de las variables estudiadas en la Sección 3, las variables del lenguaje profundo probablemente se deban a un diseño inconsciente. Como se muestra en [9-11], revelan conexiones entre textos que van mucho más allá de la conciencia de los escritores; por lo tanto, la representación geométrica de textos [10] y el ajuste fino de los canales lingüísticos [11] son herramientas más adecuadas para revelar conexiones. También es probable que indiquen la influencia de un autor sobre otro.

Definimos la cantidad de caracteres por capítulo nC y la cantidad de IP por capítulo y las cuatro variables del lenguaje profundo son [9] la cantidad de caracteres CP: nIP ,

$$PC = \frac{\text{Cantidad de facturas}}{\text{noroeste}} \tag{2}$$

el número de palabras por oración PF:

$$FP = \frac{\text{noroeste}}{\text{nS}} \tag{3}$$

el número de interpunciones por palabra, denominado intervalo de palabras, IP:

$$PI = \frac{\text{yo}}{\text{noroeste}} \tag{4}$$

el número de intervalos de palabras por oración MF:

$$MF = \frac{\text{cortar}}{\text{nS}} \tag{5}$$

La ecuación (5) también se puede escribir como MF = PF/IP.

Las tablas 6 a 9 informan la media y la desviación estándar de estas variables. Tenga en cuenta que estos valores se han calculado ponderando cada capítulo con su número de palabras para evitar que los capítulos cortos pesen tanto como los largos. Por ejemplo, el capítulo 1 de Lord tiene 10097 palabras; por lo tanto, su peso estadístico es 10097/472173 ≈ 0,021, no 1/62 ≈ 0,016. Observe también que el coeficiente de dispersión utilizado en la Sección 2 se calculó pesando cada capítulo 1/62, no 10097/472173, para concordar visualmente con la serie dibujada en la Figura 1.

Machine Translated by Google

Tabla 6. John R.R. Tolkien. Valor medio y desviación estándar (entre paréntesis) de <CP> , <PF> , < IP >, < MF > en las novelas indicadas. La media y la desviación estándar se han calculado mediante sopesando cada capítulo con su número de palabras.

Novedoso	Cp	FP	IP	MF
El hobbit	4,11 (0,06)	16,54 (2,03)	7,93 (0,98)	2,09 (0,12)
El Señor de los Anillos	4,04 (0,08)	13,92 (1,98)	6,68 (0,51)	2,08 (0,20)
El Silmarillion	4,23 (0,08)	31,21 (5,32)	8,58 (0,58)	3,62 (0,42)

Tabla 7. Clive S. Lewis. Valor medio y desviación estándar (entre paréntesis) de <CP> , <PF> , < IP >, < MF > en las novelas indicadas. La media y la desviación estándar se han calculado mediante sopesando cada capítulo con su número de palabras.

Novedoso	Cp	FP	IP	MF
Las cartas de cinta de rosca	4,36 (0,12)	23,95 (3,82)	9,72 (1,00)	2.47 (032)
La trilogía espacial	4,21 (0,16)	15,25 (3,05)	7,47 (0,98)	2,03 (0,22)
Las crónicas de Narnia	4,09 (0,09)	13,97 (1,94)	7,10 (0,89)	1,97 (0,15)

Tabla 8. George MacDonald. Valor medio y desviación estándar (entre paréntesis) de <CP> , <PF> , < IP >, < MF > en las novelas indicadas. La media y la desviación estándar se han calculado mediante sopesando cada capítulo con su número de palabras.

A espaldas del viento del norte	3,85 (0,11)	15,52 (3,34)	6,76 (0,77)	2,29 (0,35)
Phantastes: un romance de hadas para hombres y mujeres	4,20 (0,12)	21,15 (3,58)	6,43 (0,49)	3,28 (0,45)
Lilith: un romance	4,11 (0,25)	15,87 (3,89)	6,43 (0,57)	2,45 (0,42)
La princesa y el duende	4,08 (0,14)	17,81 (4,05)	7,09 (1,22)	2,46 (0,53)

Tabla 9. Otros autores. Valor medio y desviación estándar (entre paréntesis) de <CP> , <PF> , < IP >, < MF > en las novelas indicadas. La media y la desviación estándar se han calculado mediante sopesando cada capítulo con su número de palabras.

Novedoso	<CP>	<FP>	<IP>	<MF>
Las aventuras de Oliver Twist	4,23 (0,09)	18,04 (3,13)	5,70 (0,52)	3,16 (0,34)
David Copperfield	4,04 (0,12)	18,83 (2,50)	5,61 (0,30)	3,35 (0,33)
Casa sombría	4,23 (0,14)	16,95 (2,21)	6,59 (0,49)	2,57 (0,21)
Una historia de dos ciudades	4,26 (0,12)	18,27 (4,24)	6,19 (0,46)	2,93 (0,46)
Nuestro amigo mutuo	4,22 (0,12)	16,46 (2,01)	6,03 (0,37)	2,73 (0,27)
Orgullo y prejuicio	4,40 (0,14)	21,31 (5,02)	7,16 (0,46)	2,95 (0,46)
Feria de la vanidad	4,63 (0,08)	21,95 (3,67)	6,73 (0,63)	3,25 (0,39)
Moby Dick	4,52 (0,16)	23,82 (7,44)	6,45 (0,99)	3,64 (0,80)
El molino en el hilo dental	4,29 (0,13)	23,84 (4,99)	7,09 (0,69)	3,35 (0,48)
Las aventuras de Alicia en el país de las maravillas	3,96 (0,08)	17,19 (3,20)	5,79 (0,55)	2,95 (0,28)
Las aventuras de Huckleberry Finn	3,85 (0,10)	19,39 (3,12)	6,63 (0,67)	2,94 (0,48)
El libro de la selva	4,11 (0,09)	16,46 (3,09)	7,14 (0,53)	2,29 (0,30)
La guerra de los mundos	4,38 (0,18)	19,22 (4,13)	7,67 (0,92)	2,48 (0,31)
El maravilloso mago de Oz	4,02 (0,10)	17,90 (2,23)	7,63 (0,64)	2,34 (0,15)
El sabueso de los Baskerville	4,15 (0,12)	15,07 (3,16)	7,83 (0,94)	1,91 (0,22)
Peter Pan	4,12 (0,09)	15,65 (3,98)	6,35 (0,92)	2,44 (0,37)
Martín Edén	4,32 (0,13)	15,61 (2,71)	6,76 (0,64)	2,30 (0,26)

Machine Translated by Google

Específicamente, sea M el número de muestras (es decir, capítulos), entonces el valor medio <PF> viene dado por

$$<PF> = \sum_{k=1}^{METRO} PF,k \times nW,k/ \sum_{k=1}^{METRO} norteW,k \quad (6)$$

Por tanto, fíjate, para no dejarte engañar, que $< PF > = \frac{1}{Mk} \sum_{=1}^{METRO} FP,k = \sum_{k=1}^{METRO} norteW,k/ \sum_{k=1}^{METRO} nS,k=A/S.$

En otras palabras, <PF> no viene dado por el número total de palabras W dividido por el número total de oraciones S, ni asignando el peso 1/M a cada capítulo. Los tres valores coinciden sólo si todos los bloques de texto contienen el mismo número de palabras y el mismo número de frases, lo que no ocurrió. Las mismas observaciones se aplican a todas las demás variables.

Las siguientes características se pueden observar en las Tablas 6 a 9. Lord y Narnia comparten el mismo <PF> . Silmarillion es claramente diferente de Lord y Hobbit, lo que concuerda con el diferente coeficiente de dispersión. Screwtape es claramente diferente de Narnia y Trilogy. Hay una gran homogeneidad en las novelas de Dicken y una gran homogeneidad en <CP> en todas las novelas .

En las siguientes secciones, utilizamos < PF >, < IP > y < MF > para calcular índices interesantes relacionados con la memoria a corto plazo de los lectores.

5. Memoria extendida a corto plazo de escritores/lectores e índice de legibilidad universal

En este apartado nos ocupamos de las variables lingüísticas que, muy probablemente, no son gestionadas conscientemente por los escritores que, por supuesto, actúan también como lectores de su propio texto. Primero informamos los hallazgos relacionados con la memoria extendida a corto plazo y luego los relacionados con un índice de legibilidad universal. Ambos temas abordan los buffers de la memoria humana a corto plazo.

5.1. Memoria extendida a corto plazo y factor de multiplicidad

En [12,13], hemos conjeturado que la memoria humana a corto plazo es sensible a dos variables independientes, que aparentemente involucran dos buffers de memoria a corto plazo en serie, constituyentes de lo que hemos llamado la memoria extendida a corto plazo (E-STM).
El primer buffer se modela según el número de palabras entre dos interpunciones consecutivas, es decir, la variable IP, el intervalo de palabras, que sigue la ley 7 ± 2 de Miller [14]; el segundo búfer se modela según el número de intervalos de palabras, IP, contenidos en una oración (es decir, la variable MF), que varía aproximadamente de 1 a 7.

En [13], estudiamos los patrones (que dependen del tamaño de los dos buffers) que determinan el número de oraciones que teóricamente se pueden registrar en el E-STM de una capacidad determinada. Luego se compararon estos patrones con el número de frases realmente encontradas en novelas de literatura italiana e inglesa. Hemos descubierto que la mayoría de los autores escriben para lectores con poca memoria y, en consecuencia, se ven obligados a reutilizar patrones de oraciones para transmitir múltiples significados. Este comportamiento se cuantifica mediante el factor de multiplicidad α, definido como la relación entre el número de oraciones en una novela y el número de oraciones teóricamente permitidas por los dos buffers, una función de IP y MF.

Descubrimos que α > 1 es más probable que α < 1 y, a menudo, α 1. En el último caso, los escritores reutilizan muchas veces el mismo patrón de número de palabras. Pocas novelas muestran α < 1; en este caso, los escritores no utilizan algunos o la mayoría de ellos. Los valores de α encontrados en las novelas actualmente estudiadas se informan en las Tablas 10 y 11.

Tabla 10. Factor de multiplicidad α, índice de legibilidad universal < GU > y número de años escolares Y en las novelas indicadas de Tolkien, Lewis, MacDonald.

Novedoso	α	<GU>	Y
El hobbit	39,4	52,4	9.9
El Señor de los Anillos	368.1	64.2	7.4
El Silmarillion	0,2	38,7	>13

Machine Translated by Google

Tabla 10. Cont.

Novedoso	α	<GU>	Y
Las cartas de cinta de rosca	1.4	33,5	>13
La trilogía espacial	186,3	56.2	9.0
Las crónicas de Narnia	297,7	61.1	7.9
A espaldas del viento del norte	26.3	63,9	7.5
Phantastes: un romance de hadas para hombres y mujeres	24.3	56,6	8.9
Lilith: un romance	1.4	63.0	7.5
La princesa y el duende	8.5	58.2	8.4

Tabla 11. Factor de multiplicidad α, índice de legibilidad universal < GU > y número de años escolares en las indicadas novelas de la literatura inglesa.

Novedoso	α	<GU>	Y
Las aventuras de Oliver Twist	9.46	63.19	7.5
David Copperfield	12.63	64,78	7.2
Casa sombría	56,98	58,74	8.3
Una historia de dos ciudades	11.89	59,91	8.0
Nuestro amigo mutuo	43.41	62,68	7.6
Orgullo y prejuicio	5.20	50.33	10.4
Feria de la vanidad	5.26	49,74	10.5
Moby Dick	1,56	52,63	9.9
El molino en el hilo dental	2.17	50.22	10.5
Las aventuras de Alicia en el país de las maravillas	2,90	59,27	8.1
Las aventuras de Huckleberry Finn	7.05	60,42	8.0
El libro de la selva	14.10	57,59	8.6
La guerra de los mundos	6.72	49.05	10.8
El maravilloso mago de Oz	7.02	53,83	9.5
El sabueso de los Baskerville	43,87	54,87	9.2
Peter Pan	13.07	63,60	7.5
Martín Edén	46.33	58,53	8.2

5.2. Índice de legibilidad universal

En la referencia [14], hemos propuesto un índice de legibilidad universal dado por

$$GU = 89–10kCP + 300/PF–6(IP–6) \quad (7)$$

$$k = < CP,ITA >/< CP,Eng > \quad (8)$$

En la Ecuación (8), < Cp,ITA >= 4.48, < Cp,ENG >= 4.24. Utilizando las ecuaciones (7) y (8), el valor medio < kCP > de cualquier idioma está obligado a ser igual al encontrado en italiano, es decir, 4,48. La razón de esta elección es que CP es un parámetro típico de un lenguaje. que, si no se escala, sesgaría a GU sin cuantificar realmente la dificultad de lectura para lectores, que en su lengua están acostumbrados, en promedio, a leer palabras más cortas o más largas que en italiano. Esta escala, por lo tanto, evita cambiar GU por la única razón de que un El idioma tiene, en promedio, palabras más cortas (como el inglés) o más largas que el italiano. En todo caso, Cp afecta la ecuación (7) mucho menos que PF o IP.

Los valores de < GU >—calculados como las otras variables lingüísticas, es decir, ponderando Los capítulos (muestras) según el número de palabras se presentan en las Tablas 10 y 11.

Machine Translated by Google

El lector puede verse tentado a calcular la ecuación (7) introduciendo los valores medios informados en las tablas 6 a 9. Esto, por supuesto, se puede realizar, pero debe tenerse en cuenta que los valores así obtenidos son siempre menores o iguales (por lo tanto, son límites inferiores) a las medias calculadas a partir de las muestras (ver Apéndice A). Por ejemplo, para Lord, en lugar de 64,9, obtendríamos 61,9.

Es interesante "decodificar" estos valores medios en el número mínimo de años escolares, Y necesarios para que una novela sea "fácil" de leer, según el sistema escolar italiano, que se asume como referencia, ver Figura 1 de [15]. Los resultados también se enumeran en las Tablas 10 y 11.

5.3. Discusión

Se pueden extraer varias observaciones interesantes de los resultados presentados en las subsecciones anteriores. (a).

Silmarillion con α = 0,2 es bastante diferente de otros escritos de Tolkien. Matemáticamente, esto se debe a su gran < MF >= 3,62 y < IP >= 8,58. En la práctica, el número de oraciones teóricas permitidas por el E-STM para leer este texto es solo 1/α = 5 veces el número de patrones de oraciones realmente utilizados en el texto. El lector necesita un E–STM potente y capacidad de lectura, ya que GU = 38,7 e Y > 13. Esto no ocurre con Hobbit (α = 39,4, GU = 52,4, Y =9,9) y Lord (α = 368,1, GU = 64,2 , Y = 7,4) en el que Tolkien reutiliza patrones muchas veces, especialmente en Lord. (b). Lord y Narnia muestran valores muy grandes, α = 368,1 y α = 297,7, y GU y años escolares muy similares: GU = 64,2, Y = 7,4 y GU = 61,1, Y = 7,9, respectivamente.

Lewis reutiliza muchas veces los patrones de oraciones en esta novela, pero no en Screwtape (α = 1,4), que es más difícil de leer (GU = 33,5) y requiere más años de escolarización, Y > 13. Además, Lord y Narnia tienen prácticamente lo mismo < PF >≈ 14. (c). En general, Narnia está más cerca de Lord que de Trilogy, aunque el número de palabras y frases en Trilogy y Narnia son bastante similares (Tabla 1). Esta diferencia entre Trilogía (GU = 56,2, Y = 9) y Narnia (GU = 61,1, Y = 7,9) podría depender de los diferentes lectores a los que se dirige, niños para Narnia y adultos para Trilogía, con diferente capacidad lectora, como indica GU . (d). Las novelas de MacDonald muestran valores de α y GU muy similares a los de las otras

Novelas inglesas.

(mi). Observe la homogeneidad en las novelas de Dicken, que requieren alrededor de Y = 7 8 años de Índice de escuela y legibilidad < GU >= 59 65.

En conclusión, Lord y Narnia son las novelas que se dirigen a lectores con buffers E-STM muy similares, reutilizan patrones de oraciones de manera similar, contienen la misma cantidad de palabras por oración y requieren la misma capacidad de lectura y años escolares en comparación con otras novelas de Tolkien y Lewis. Las conexiones matemáticas entre Lord y Narnia se profundizarán en la siguiente sección, donde se utilizan los cuatro parámetros profundos del lenguaje para representar textos geométricamente.

6. Representación Geométrica de Textos

Los valores medios de las tablas 6 a 9 se pueden utilizar para evaluar en qué medida los textos son "cercanos", o matemáticamente similares, en el plano de coordenadas cartesianas, definiendo combinaciones lineales de variables del lenguaje profundo. Luego los textos se modelan como vectores; la representación se analiza en detalle en [9,10] y se recuerda brevemente aquí. Una extensión de esta representación geométrica de los textos permite calcular la probabilidad de que un texto pueda confundirse con otro , una extensión en dos dimensiones del problema comentado en [16]. Los valores de la probabilidad condicional entre dos textos (autores) pueden considerarse un índice que indica quién influyó en quién.

Machine Translated by Google

6.1. Representación vectorial de textos

Consideremos los siguientes seis vectores de los componentes indicados del lenguaje profundo. variables: $\vec{R1}$ = (< CP >, < PF >), $\vec{R2}$ = (< MF >, < PF >), $\vec{R3}$ = (<IP> , <PF>) , $\vec{R4}$ = (< CP >, < MF >), $\vec{R5}$ = (<IP> , <MF>) , $\vec{R6}$ = (< IP >, < CP >) y sus suma vectorial resultante:

$$\vec{R} = \sum_{k=1}^{6} \vec{Rk} = x\,\vec{yo} + y\,\vec{j} \quad (9)$$

La elección de qué parámetro representa la componente en los ejes de abscisas y ordenadas no es importante porque, una vez hecha la elección, de ello dependerán los resultados numéricos, pero no las comparaciones relativas y conclusiones generales.

En el primer cuadrante del plano de coordenadas cartesiano, es probable que dos textos estén matemáticamente conectados (muestran puntos finales cercanos del vector (9)) si su distancia pitagórica relativa es pequeña. Una distancia pequeña significa que los textos comparten una estructura matemática similar, según las cuatro variables profundas del lenguaje.

Al considerar los componentes vectoriales x e y de la ecuación (9), obtenemos el diagrama de dispersión que se muestra en la Figura 2 , donde X e Y son coordenadas normalizadas calculadas estableciendo a Lord en el origen (X = 0, Y = 0) y Silmarillion en (X = 1, Y = 1), según las transformaciones lineales:

$$X = \frac{x - xSeñor}{xSilma - xSeñor} \quad (10)$$

$$X = \frac{y - ySeñor}{ySilma - ySeñor} \quad (11)$$

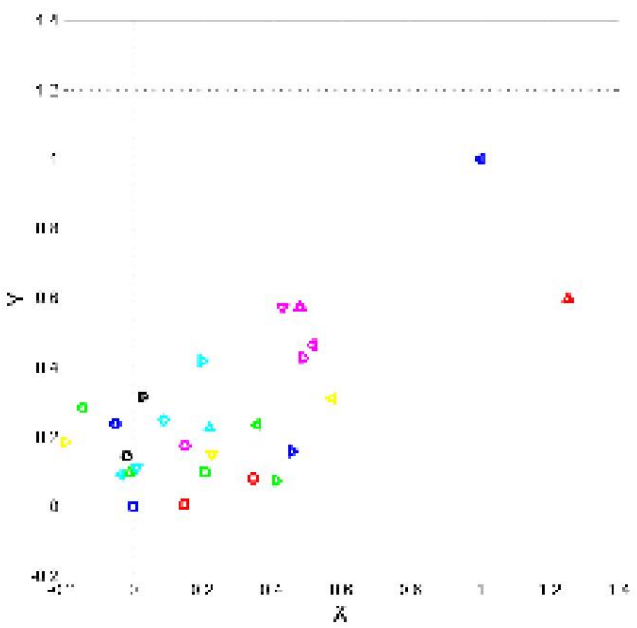


Figura 2. Coordenadas normalizadas X e Y del punto final del vector (5) de modo que Lord, cuadrado azul, está en (0,0) y Silmarillion, triángulo azul que apunta hacia la izquierda, está en (1,1). Narnia: cuadrado rojo; Trilogía: círculo rojo; Hobbit: triángulo azul que apunta hacia la derecha; Cinta de rosca: triángulo rojo apuntando hacia arriba; Atrás: triángulo cian apuntando hacia la izquierda; Lilith: triángulo cian apuntando hacia abajo; Atrás: triángulo cian apuntando hacia la izquierda; Fantastes: triángulo cian apuntando hacia la derecha; Princesa: triángulo cian apuntando hacia arriba; Oliver: círculo azul; David: círculo verde; Cuento: círculo cian; Desolado: círculo magenta; Mutua: círculo negro; Orgullo: triángulo magenta apuntando hacia la derecha; Vanidad: triángulo magenta apuntando hacia la izquierda; Moby: triángulo magenta apuntando hacia abajo; Molino: triángulo magenta apuntando hacia arriba; Alice: triángulo amarillo apuntando hacia la derecha; Selva: triángulo amarillo apuntando hacia abajo; Guerra: triángulo amarillo apuntando hacia la derecha; Oz: triángulo verde apuntando hacia la izquierda; Tomar el sol: triángulo verde que apunta hacia la derecha; Peter: triángulo verde apuntando hacia arriba; Martín: cuadrado verde; Finn: triángulo negro que apunta hacia la derecha.

< 0,2

Machine Translated by Google

En la Figura 2, podemos observar que Silmarillion y Screwtape están claramente muy lejos de todos los demás textos examinados, lo que marca su sorprendente diversidad, como ya se señaló; por lo tanto, en los siguientes análisis los descuidamos. Además, Pride, Vanity, Moby y Floss se agrupan y lejos de Trilogy, Narnia y Lord; por lo tanto, en los siguientes análisis no los consideraremos más a fondo.

El conjunto completo de la distancia pitagórica d entre pares de textos se presenta en el Apéndice B. Estos datos describen sintéticamente la proximidad de los textos y pueden indicar a los estudiosos de la literatura conexiones entre textos no consideradas antes.

La Figura 3 muestra un ejemplo de estas distancias relativas a Lord, Narnia y Trilogy. Haciendo referencia a los casos en los que d < 0,2, podemos observar lo siguiente: (a).

Los textos más cercanos a Lord son Narnia, Back, Lilith, Mutual y Peter. (b).
Los textos más cercanos a Narnia son Lord, Lilith, Bleak, Martin y Peter. (do).
Los textos más cercanos a la Trilogía son Hobbit, Martin y Peter.

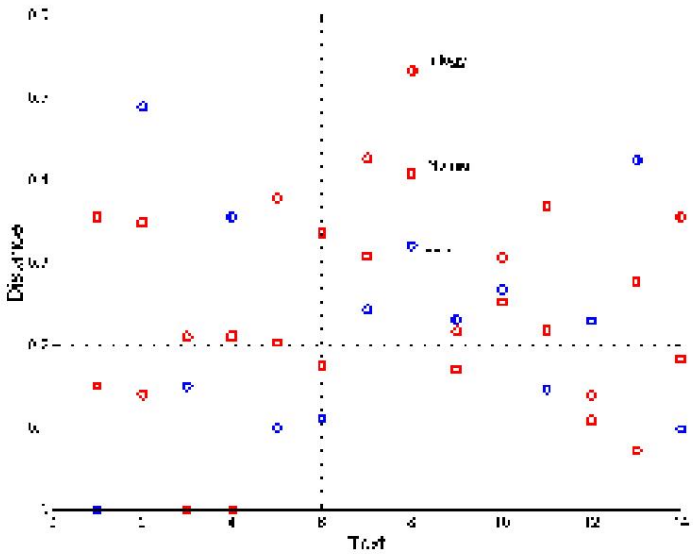


Figura 3. Distancia pitagórica d entre pares de textos considerando Lord (las distancias referidas a este caso están marcadas con círculos azules), Narnia (cuadrados rojos) y Trilogía (círculos rojos). Clave: Lord 1, Hobbit 2, Narnia 3, Trilogy 4, Back 5, Lilith 6, Oliver 7, David 8, Bleak 9, Tale 10, Mutual 11, Martin 12, Bask 13, Peter 14.

Además de la proximidad con novelas anteriores, Lord y Narnia muestran una estrecha proximidad entre sí y con dos novelas de MacDonald.

Estas observaciones, sin embargo, se refieren a la visualización "promedio" de vectores cuyo punto final depende sólo de los valores medios. La desviación estándar de las cuatro variables del lenguaje profundo, presentadas en las tablas 6 a 9, introduce dispersión de datos; por lo tanto, en la siguiente subsección estudiamos y discutimos este tema calculando la probabilidad (llamada probabilidad de “error”) de que un texto pueda confundirse matemáticamente con otro.

6.2. Probabilidad de error: un índice para evaluar quién influyó en quién

además del vector $\vec{R}$ de la ecuación (9), debido a los valores medios, podemos considerar otra vector $\vec{\rho}$ debido a la desviación estándar de las cuatro variables profundas del lenguaje que se suma a R. ρ , + En este caso, vector aleatorio final que describe un texto viene dado por

$$\vec{t} = \vec{R} + \vec{\rho} \quad (12)$$

Ahora, para obtener una idea de esta nueva descripción, consideramos el área de un círculo. centrado en el punto final de R. $\vec{R}$

$$= 0{,}5 \times (+)$$

$$=$$

–

Machine Translated by Google

Fijamos la magnitud (radio) ρ de la siguiente manera. Primero, sumamos las varianzas de la y profunda . variables del lenguaje que determinan los componentes x e y de $\vec{R}$, sean σ y $^{2}_{2\sigma x}$. Entonces, calculamos el valor medio σ $^{2\,2\,2 = 0{,}5 \times \sigma + \sigma y}_{\rho}$ finalmente fijamos

$$\rho = \sigma\rho \tag{13}$$

Ahora, dado que al calcular las coordenadas x e y de $\vec{R}$, una variable del lenguaje profundo se puede sumar dos o más veces, sumamos su desviación estándar (denominada sigma) dos o más veces antes de elevarla al cuadrado. Por ejemplo, en el componente x, IP aparece tres veces; por lo tanto, su contribución a la varianza total en el eje x es 9 veces la varianza calculada a partir de la desviación estándar informada en las tablas 6 a 9. Para Lord, por ejemplo, es 9 × 0,512 . Después de estos cálculos, los valores del círculo 1-sigma se transforman en las coordenadas normalizadas X, Y según las Ecuaciones (10) y (11).

La Figura 4 muestra un ejemplo significativo que involucra a Lord, Narnia, Trilogy, Back y Peter. Vemos que Lord puede confundirse casi por completo con Narnia, y parcialmente con Trilogy, pero no al revés. Lord también se puede confundir con Peter y Back, lo que indica fuertes conexiones con estas novelas anteriores.

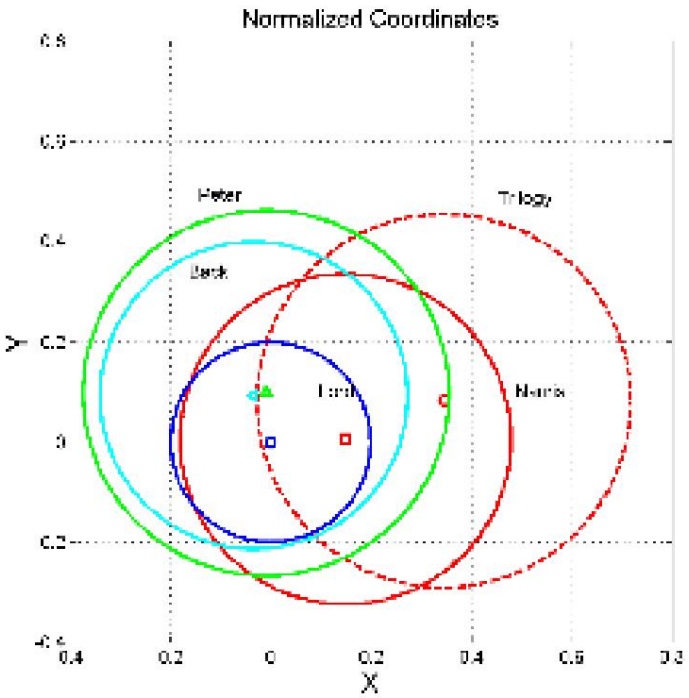


Figura 4. Coordenadas normalizadas X e Y del punto final del vector (5) y los círculos 1-sigma, tales que Lord, el cuadrado azul, está en (0,0) y Silmarillion, el triángulo azul que apunta hacia la izquierda, está en (1,1). Señor: cuadrado azul (círculo azul 1-sigma); Narnia: cuadrado rojo (círculo rojo 1-sigma); Trilogía: círculo rojo (círculo rojo discontinuo 1-sigma); Atrás: triángulo cian apuntando hacia la izquierda (círculo cian 1-sigma); Peter: triángulo verde apuntando hacia arriba (círculo verde 1-sigma).

Ahora podemos estimar la probabilidad (condicional) de que un texto se confunda con otro calculando la proporción de áreas. Este procedimiento es correcto si suponemos que la densidad bivariada de las coordenadas normalizadas ρX, ρY, centradas en $\vec{R}$, es uniforme. Al asumir esta hipótesis, podemos calcular las probabilidades como la relación de áreas [17,18].

La hipótesis de una uniformidad sustancial alrededor de R debe justificarse observando que = (/) las $\vec{R}$ coordenadas X, Y probablemente estén distribuidas según una densidad bivariada log-normal porque el logaritmo de las cuatro variables profundas del lenguaje, que se combinan en la Ecuación (9) (/) ( ) linealmente, se puede modelar como el teorema del límite central, deberíamos esperar = (/) ( ) aproximadamente un modelo gaussiano en los valores con un valor significativamente mayor.

= (a) = (b) ( ) ( ) gaussiano. Para lineales; pero

Machine Translated by Google

desviación estándar que la de las variables individuales. Por lo tanto, en la zona cercana a $\vec{R}$, el La función de densidad bivariada no debe tener un máximo, de ahí el modelado de densidad uniforme.

Ahora podemos calcular las siguientes probabilidades. Sea A el área común de dos 1-círculos sigma (es decir, el área proporcional a la probabilidad conjunta de dos textos), sea A1 el área del círculo 1-sigma del texto 1 y A2 el área del círculo 1-sigma del texto 2. Ahora, dado que las probabilidades son proporcionales a las áreas, obtenemos las siguientes relaciones:

$$\frac{A}{A1} = \frac{P(A1, A2)}{PAG(A1)} = \frac{P(A2/A1)P(A1)}{PAG(A1)} = P(A2/A1) \quad (14)$$

$$\frac{A}{A2} = \frac{P(A1, A2)}{P(A2)} = \frac{P(A1/A2)P(A2)}{P(A2)} = P(A1/A2) \quad (15)$$

En otras palabras, A/A1 da la probabilidad condicional P(A2/A1) de que la parte del texto 2 puede confundirse (o "contenerse") con el texto 1; A/A2 da la probabilidad condicional P(A1/A2) esa parte del texto 1 se puede confundir con el texto 2. Observe que estas probabilidades condicionales dependen de la distancia entre dos textos y de los radios 1-sigma (Apéndice

Por supuesto, estas probabilidades conjuntas se pueden extender a tres o más textos, por ejemplo, en En la Figura 4 podríamos calcular el área compartida por Lord, Narnia y Trilogy y la probabilidad conjunta correspondiente, que no se realiza en el presente artículo.

Pensamos que las probabilidades condicionales y la presentación visual de los círculos 1-sigma dar pistas útiles para establecer posibles conexiones ocultas entre textos y, tal vez, incluso entre autores, porque las variables involucradas no son manejadas conscientemente por ellos.

En la Tabla 12, la probabilidad condicional P(A2/A1) se informa en las columnas; por lo tanto, A1 se refiere al texto indicado en la fila superior. P(A1/A2) se informa en las filas; por lo tanto, A2 se refiere al texto indicado en la columna de la izquierda.

Tabla 12. Probabilidad condicional entre las novelas indicadas. P(A2/A1) se informa en las columnas; por lo tanto, A1 se refiere al texto indicado en la fila superior. P(A1/A2) se informa en las filas; por lo tanto, A2 se refiere al texto indicado en la columna de la izquierda. Por ejemplo, asumiendo Señor como texto 1 (columna 1 de Tabla 12) y Narnia como texto 2 (fila 3), encontramos P(A2/A1) = 0,974 y viceversa. Si asumimos Narnia como texto 1 (columna 3) y Lord como texto 2 (fila 1), encontramos P(A2/A1) = 0,356.

	Novela Lord Hobbit	Narnia	Trilogía	Espalda	Lilith	Oliver	David	Bleak	Tale	Mutual	Martin	Bask	Peter
1	Lord 1 Hobbit	0,031	0,356	0,142	0,423	0,511	0,277	0,041	0,307	0,231	0,619	0,301	0,078	0,299
2	0,099 Narnia 0,974		0,421	0,731	0,171	0,225	0,074		0,592	0,376	0,060	0,665	0,833	0.227
3	Trilogía 0,498 Back	1		0,550	0,647	0,781	0,489	0	0,927	0,625	0,886	0,949	0,462	0.602
4	Lililth Oliver 0,352	0,354	1		0,400	0,510	0,297		0,921	0,608	0,473	0,978	0,908	0,421
5	David	0,786	0,704	1		0,997	0,866	0,166	0,763	0,692		0,566	0,179	0.707
6	0,024 1 0,891	0,124	0,559	0,270	1		0,741	0	0,762	0,661		0,546	0,169	0.521
7	Desolado 0,307	0,121	0,498	0,254	0,735	1		0,796	0,444	0,579	1 1	0,239	0,043	0,378
8	Tale 0,330 Mutual		0,227		0,466	0,540	1		0,029	0,173	0,917	0,001		0,168
9	0,390 Martin 0,490	0,029	0,035	0,108	0,195	0,186	0,416	0,559		0,598	0,262	0,558	0	0,296
10	Bask 0,250 Peter	0	0,339	0	0,323	0,437	0,350	0,913	1		0,526	0,484	0,208	0.385
11	1	0,182	0,327	0,264	0,417	0,541	0,650	1	0,852	1		0,205	0,176	0,188
12		0,165	0,204	0,248	0,266	0,361	0,454	0,051	0,331	0,342	0,774		0,031	0.384
13		0,012	0,565	0,085	0,389	0,509	0,307		0,906	0,551	1	1		0,289
14		0,333	0,826	0,345	0,736	0,455	0,838	0,244	1	0,312	1	0,110	0,996	0,427	0,285	0,669	0,988	0,397	0,904	0,529	0,160	0,785	0,790	0,396	1	0,300	1

Observe que P(A2/A1) = 1 significa A = A1; por lo tanto, el texto 1 puede confundirse completamente con texto 2. P(A1/A2) = 1 significa A = A2; por lo tanto, el texto 2 se puede confundir completamente con el texto 1.

Por ejemplo, asumiendo Señor como texto 1 (columna 1 de la Tabla 12) y Narnia como texto 2 (fila 3), encontramos P(A2/A1) = 0,974 y viceversa. Si asumimos Narnia como texto 1 (columna 3) y Señor como texto 2 (fila 1), encontramos P(A2/A1) = 0,356. Estos datos indican que Lord puede ser Se confunde con Narnia con una probabilidad cercana a 1, pero no al revés. En otras palabras, en En el banco de datos considerado en este artículo, si una máquina extrae aleatoriamente un capítulo de Lord, otra máquina, inconsciente de esta elección, podría atribuirla a Lord, pero también con disminución probabilidad de Back, Peter, Narnia y Lilith.

Machine Translated by Google

Por el contrario, si el texto está extraído de Narnia, entonces lo más probable es que se atribuya a Pedro o Trilogía que al Señor u otros textos.

Creemos que estas probabilidades condicionales indican quién influyó más en quién. En otras palabras, Tolkien influyó más en Lewis que todo lo contrario.

Ahora podemos definir un parámetro sintético que pone de relieve en qué medida, en promedio, dos textos pueden confundirse erróneamente entre sí. El parámetro es el promedio = (/) ( ) + (/) ( ) probabilidad condicional (ver [16] para un problema similar):

( ) = ( ) = 0,5

$$pe = P(A2/A1)P(A1) + P(A1/A2)P(A2) = 0,5 \times [P(A2/A1) + P(A1/A2)] \quad (16)$$

= 0

Ahora bien, como al comparar dos textos podemos suponer que P(A1) = P(A2) = 0,5, obtenemos

$$pe = 0,5 \times [P(A2/A1) + P(A1/A2)] \quad (17)$$

= 1

Si pe = 0, no hay intersección entre los dos círculos 1-sigma. Los dos textos no pueden confundirse; por lo tanto, no existe ninguna conexión matemática que involucre los parámetros profundos del lenguaje (esto sucede con Screwtape y Silmarillion, que pueden confundirse entre sí, pero no con los otros textos). Si pe = 1, los dos textos pueden confundirse totalmente y los dos círculos 1-sigma coinciden. El Apéndice D informa los valores de pe para todos > 0,5 ≤ 0,5 los pares de novelas.

= 0,5

Ahora, sólo para permitir un análisis aproximado, es razonable suponer pe = 0,5 como umbral de referencia, es decir, la probabilidad de obtener cara o cruz al lanzar una moneda justa. Si pe > 0,5, entonces dos textos pueden confundirse no por casualidad; si pe ≤ 0,5, entonces no es probable que se confundan dos textos .

> 0,5

Para visualizar pe , la Figura 5 dibuja pe cuando el texto 1 es Señor (columna 1 de la Tabla 12), Narnia (columna 3) o Trilogía (columna 4). Observamos que pe > 0,5 en los siguientes casos: (a).
Señor como texto 1: Narnia, Back, Lilith, Mutual, Peter. (b).
Narnia como texto 1: Lord, Trilogy, Back, Lilith, Bleak, Mutual, Martin, Peter. (do).
Trilogía como texto 1: Hobbit, Narnia, Bleak, Martin, Bask.

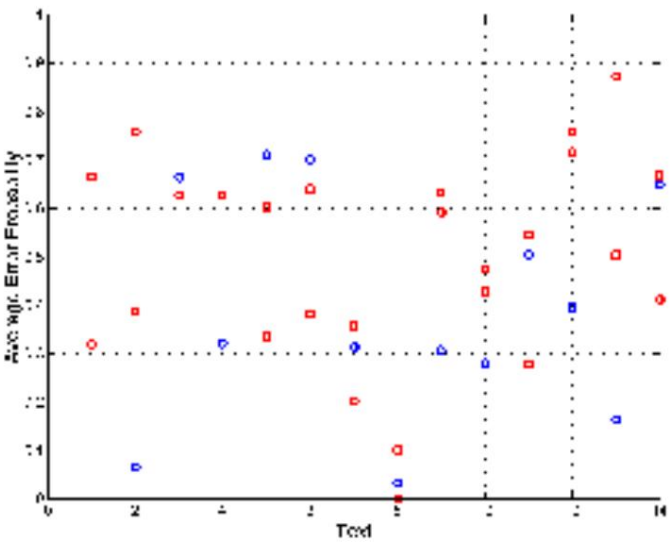


Figura 5. Probabilidad de error pe versus texto 2. Lord (las probabilidades referidas a este caso están etiquetadas con círculos azules), Narnia (cuadrados rojos) y Trilogy (círculos rojos). Clave de texto: Lord 1, Hobbit 2, Narnia 3, Trilogy 4, Back 5, Lilith 6, Oliver 7, David 8, Bleak 9, Tale 10, Mutual 11, Martin 12, Bask 13, Peter 14.

Podemos reiterar que Tolkien (Lord) aparece significativamente conectado con Lewis (Narnia), MacDonald (Back, Lilith) y Barrie (Peter), pero no con las novelas de Dicken donde, por el contrario, Lewis aparece conectado.

Machine Translated by Google

En la siguiente sección, se destacan las cuatro variables profundas del lenguaje para considerarlas. Canales lingüísticos existentes en los textos. Este es el análisis que hemos llamado el "ajuste fino" de textos [11].

7. Relaciones lineales en textos literarios

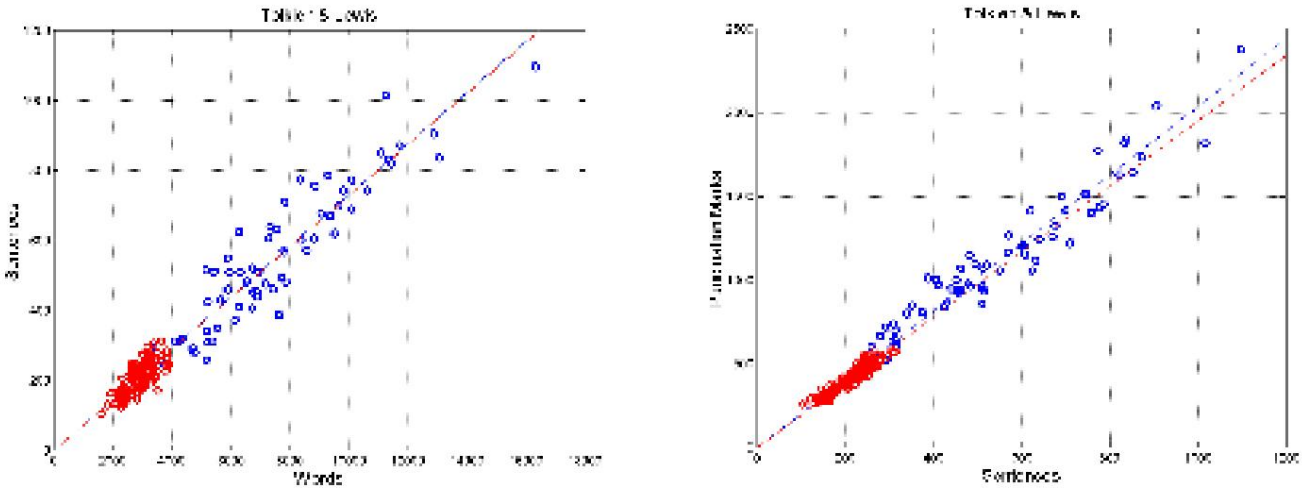
La teoría de los canales lingüísticos, que revisaremos en la siguiente sección, es basado en la línea de regresión entre variables lingüísticas:

$$y = mx = \quad 0.0729 \quad = 0,7610. \tag{18}$$

$= 0,0731$ $= 0,9199$

Por ello, mostramos ejemplos de estas relaciones lineales que se encuentran en Lord y Narnia.

La Figura 6a muestra el diagrama de dispersión de nS $< >$ versus nW de Lord y Narnia. En Narnia, el la pendiente de la recta de regresión es m = 0,0729 y el coeficiente de correlación r = 0,7610. En Señor, m = 0,0731 y r = 0,9199. Dado que las relaciones promedio, es decir, la ecuación (18), son prácticamente idéntico (ver también los valores de < PF > en las Tablas 6 y 7), mientras que la correlación $= 2.0372$ $= 0.9609$ $= 1.9520$ $=$ coeficientes (es decir, la dispersión de los datos) no lo son, este hecho afectará la oración canal discutido en la Sección 9.



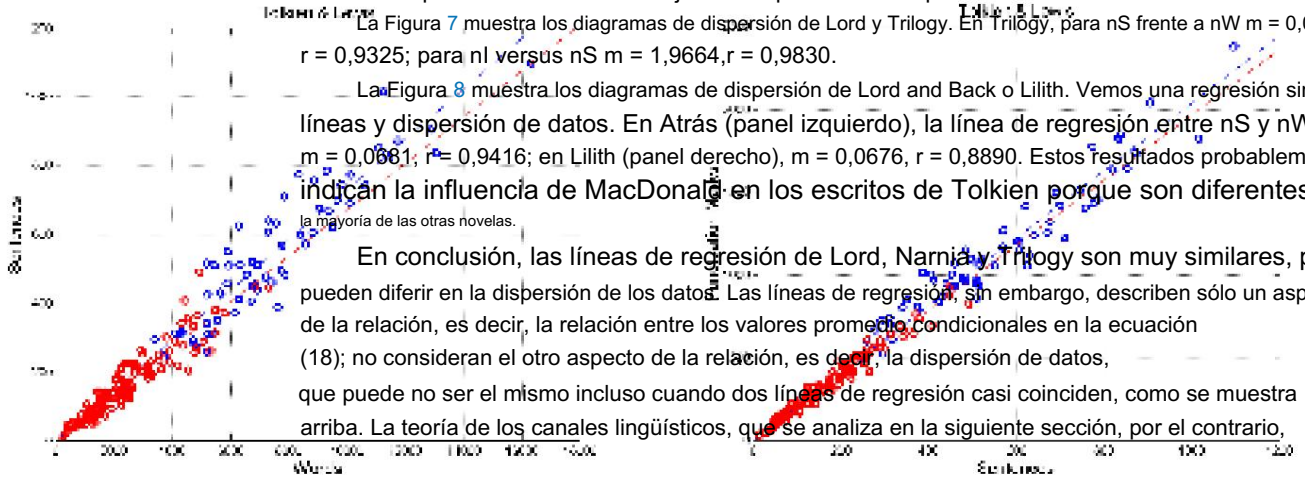
(a) (b)

Figura 6. (a) Diagrama de dispersión de nS frente a nW en Lord (azul) y Narnia (rojo); (b) nI versus nS en Lord (azul) y Narnia (rojo).

Se pueden realizar observaciones similares para la Figura 6b, que muestra nI versus nS en Señor y Narnia. Encontramos $m = 2.0372$, r = 0.9609 en Lord, y m = 1.9520 y r = 0.9384 en Narnia. El Apéndice E informa el conjunto completo de estos parámetros. $=$ $0.0672 = 0.9325 = 1.9664 = 0.9830$

La Figura 7 muestra los diagramas de dispersión de Lord y Trilogy. En Trilogy, para nS frente a nW m = 0,0672, r = 0,9325; para nI versus nS m = 1,9664,r = 0,9830.

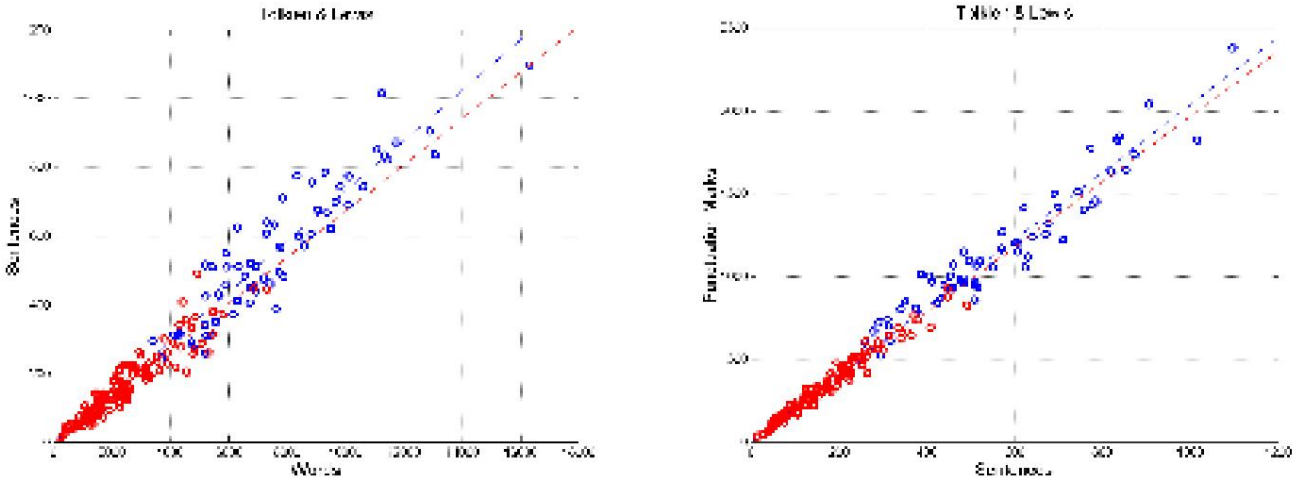
La Figura 8 muestra los diagramas de dispersión de Lord and Back o Lilith. Vemos una regresión similar líneas y dispersión de datos. En Atrás (panel izquierdo), la línea de regresión entre nS y nW da m = 0,0681, r = 0,9416; en Lilith (panel derecho), m = 0,0676, r = 0,8890. Estos resultados probablemente indican la influencia de MacDonald en los escritos de Tolkien porque son diferentes de la mayoría de las otras novelas.



En conclusión, las líneas de regresión de Lord, Narnia y Trilogy son muy similares, pero pueden diferir en la dispersión de los datos. Las líneas de regresión, sin embargo, describen sólo un aspecto de la relación, es decir, la relación entre los valores promedio condicionales en la ecuación (18); no consideran el otro aspecto de la relación, es decir, la dispersión de datos, que puede no ser el mismo incluso cuando dos líneas de regresión casi coinciden, como se muestra arriba. La teoría de los canales lingüísticos, que se analiza en la siguiente sección, por el contrario,

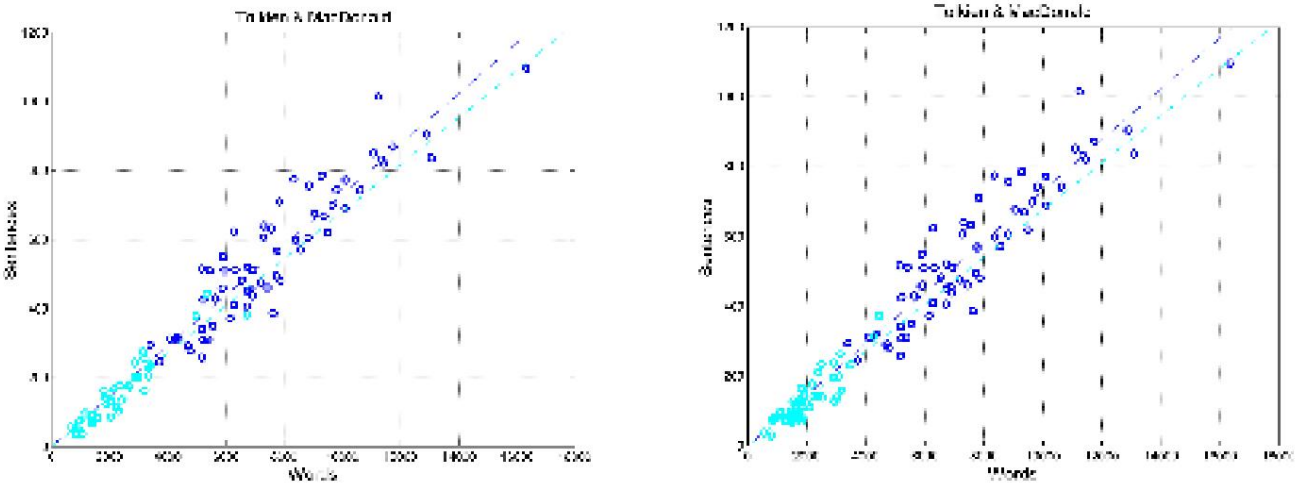
Machine Translated by Google

considera tanto las pendientes como los coeficientes de correlación y proporciona una herramienta de "ajuste fino" para cada par de obras, optimizando cada uno de los cuatro parámetros profundos del lenguaje.



(a) = 0,0681 = 0,9416 (b) = 0,0676 =

Figura 7. (a) Diagrama de dispersión de nS versus nW en Lord (azul) y Trilogy (rojo); (b) nI versus nS en Lord (azul) y Trilogy (rojo).



(a) (b)

Figura 8. Diagrama de dispersión del número de oraciones nS versus el número de palabras nW: (a) Lord (azul) y Back (cian); (b) Lord (azul) y Lilith (cian).

8. Teoría de los canales lingüísticos En este

apartado recordamos la teoría general de los canales lingüísticos [11]. En una obra literaria, una variable independiente (de referencia) x (p. ej., nW) y una variable dependiente y (p. ej., nS) pueden relacionarse mediante la recta de regresión dada por la ecuación (18).

Consideremos dos bloques de texto diferentes Yk e Yj , por ejemplo, los capítulos de la obra k y la obra j. La ecuación (18) no proporciona la relación completa entre dos variables porque vincula sólo valores promedio condicionales. Podemos escribir relaciones lineales más generales, que se ocupan de la dispersión de los datos (medidos por los coeficientes de correlación rk y rj , respectivamente) alrededor de los valores promedio (medidos por las pendientes mk y

$$yk = mkx + nk \quad (19)$$

)

$$yj = mjx + nj \quad (20)$$

)

Machine Translated by Google

Los modelos lineales de las Ecuaciones (19) y (20) introducen “ruido” aditivo a través de las variables estocásticas nk y nj , con valor medio cero [9,11,15]. El ruido se debe al coeficiente de correlación |r| = 1.

Podemos comparar dos obras literarias eliminando x; por lo tanto, comparamos la variable de salida y con el mismo número de la variable de entrada x. Por ejemplo, podemos comparar el número de oraciones en dos novelas (para un número igual de palabras) considerando no solo la relación promedio, ecuación (18), sino también la dispersión de los datos, medida por el coeficiente de correlación, ecuaciones (19) y (20). Nos referimos a este canal de comunicación como “canal de oraciones”, canal S, y a este procesamiento como “sintonía fina” porque profundiza el análisis de los datos y puede proporcionar más información sobre la relación entre dos obras literarias o cualquier otro texto. .

Al eliminar x de las ecuaciones (19) y (20), obtenemos la relación lineal entre el número de oraciones de entrada en el trabajo Yk (ahora el texto de entrada de referencia) y el número de oraciones en el texto Yj (ahora el texto de salida):

$$mk\ mk\frac{mj}{}\ mj\ nk + \frac{}{}nj\ yj = yk– \tag{21}$$

En comparación con la nueva obra de referencia Yk , la pendiente mjk viene dada

$$\text{por } mjk = mj/mk \tag{22}$$

La fuente de ruido que produce el coeficiente de correlación entre Yk e Yj viene dada por

$$\frac{mj}{}nk + nj = –mjknk + nj\ njk = – mk \tag{23}$$

La “relación ruido-señal de regresión”, Rm, debida a mjk = 1, del nuevo canal viene dada por

$$salón = mjk–12 \tag{24}$$

El coeficiente de correlación desconocido rjk entre yj e yk viene dado por

$$rjk = \text{porque arcos } rj –arcos(rk) \tag{25}$$

La “relación ruido-señal de correlación”, Rr texto Yk a , debido a rjk < 1, del nuevo canal de texto Yj viene dada por

$$Rr = \frac{1–r\ jk^{2}}{jk}\frac{}{m2}jk\ 2\ r \tag{26}$$

Debido a que las dos fuentes de ruido son independientes y aditivas, la relación total ruido-señal del canal que conecta el texto Yk con el texto Yj viene dada por

$$R = Rm + Rr \tag{27}$$

Observe que la ecuación (27) se puede representar gráficamente [10]. Finalmente, el total y el Las relaciones parciales señal-ruido están dadas por

$$ΓdB = –10 × log10R \tag{28}$$

$$Γm,dB = –10 × log10Rm \tag{29}$$

$$Γr,dB = –10 × log10Rr \tag{30}$$

Por supuesto, esperamos que ningún canal pueda producir rjk = 1 y mjk = 1; por lo tanto, ΓdB = ∞, caso denominado canal ideal, a menos que se compare un texto consigo mismo. En la práctica, siempre encontramos rjk < 1 y mjk = 1. La pendiente mjk mide el “sesgo” multiplicativo de la variable dependiente en comparación con la variable independiente; el coeficiente de correlación rjk mide qué tan "preciso" es el mejor ajuste lineal.

En conclusión, la pendiente mjk es la fuente del ruido de regresión Rm, y la correlación El coeficiente rjk es principalmente la fuente del ruido de correlación del canal Rr .

Machine Translated by Google

9. Canales lingüísticos

En textos largos (como novelas, ensayos, etc.), podemos definir al menos cuatro líneas lingüísticas. canales [11], a saber:

(a). Canal de oración (canal S)
(b). Canal de interpunciones (canal I)
(do). Canal de intervalo de palabra, canal WI
(d). Canal de personajes (canal C).

En los canales S, el número de frases de dos textos se compara con el mismo número de palabras. Estos canales describen cuántas oraciones escribe el autor del texto j , en comparación al escritor del texto k (texto de referencia), utilizando el mismo número de palabras. Por lo tanto, estos Los canales están más vinculados al PF que a otros parámetros. Es muy probable que reflejen la estilo del escritor.

En I–channels se compara el número de intervalos de palabras de dos textos para el mismo número de frases. Estos canales describen cuántos textos breves entre dos signos de puntuación contiguos (de longitud IP) utilizan dos autores; por lo tanto, estos canales son más vinculado a MF que a otros parámetros. Dado que es muy probable que MF esté relacionado con el E-STM, Los canales I están más relacionados con el segundo buffer del E-STM de los lectores que con el estilo del escritor.

En los canales WI, se compara el número de palabras contenidas en un intervalo de palabras (es decir, IP) para el mismo número de interpunciones. Estos canales están más vinculados a IP que a otros parámetros. Dado que es muy probable que IP esté conectado con E-STM, los canales WI son más más relacionado con el primer buffer de E-STM de los lectores que con el estilo del escritor.

En los canales C, el número de caracteres de dos textos se compara con el mismo número de palabras. Están más relacionados con el idioma utilizado, por ejemplo, el inglés, que con el otro. parámetros, a menos que se consideren ensayos o textos científico/académicos porque estos últimos los textos utilizan, en promedio, palabras más largas [9].

Como ejemplo, la Tabla 13 presenta las relaciones señal/ruido total y parcial ΓdB, Γm,dB, Γr,dB en los cuatro canales considerando a Lord como texto de referencia (entrada). en otros En palabras, el texto j se compara con el texto k (texto de referencia, es decir, Señor).

Tabla 13. Relación señal-ruido total y parcial ΓdB, Γm,dB, Γr,dB en los cuatro canales considerando a Lord como texto de referencia (entrada).

	Canal S			I–Canal			WI–Canal			Canal C		
Novedoso	ΓdB	Γm,dB	Γr,dB	ΓdB	Γm,dB	Γr,dB	ΓdB	Γm,dB	Γr,dB	ΓdB	Γm,dB	Γr,dB
hobbit	14.60	15.48	21,94	15.04	30.08	15.18	15.61	15,77	30.19	28,87	35,90	29,83
narnia	10.12	51,26	10.12	21,83	27,57	23.18	7,94	27.08	7,99	23.14	37,47	23.30
Trilogía	21.27	21,86	30.24	20.07	29.18	20.64	18.74	19.21	28.63	27,61	27,85	40.30
Volver	21.10	23.30	25.11	20,96	24.12	23,82	21,96	57,42	21,96	26,52	26,68	41.05
Lililth	19,92	22.47	23.44	18.47	19.15	26,87	20,86	26.28	22.33	19.17	33,98	19.31
Oliver	12,87	12.93	31,51	5.72	5.74	29.30	16.54	16.56	40,83	26,50	26,63	41,73
David	11.43	11.52	28.37	4.18	4.19	30,77	14.43	16.10	19.37	32,88	51,53	32,94
Bleak	14.91	14,93	38.01	12.10	12.29	25,80	29.00	36.33	29,88	21,76	26,69	23.45
Tale 12.66 Mutual		13,31	21.25	8.26	8,56	19,99	16,13	22,76	17.20	25,88	25,89	55.01
16.13 Martin 16.00		16,22	33,05	9,97	10,00	31,20	19,40	20,33	26,55	23,24	26,87	25,70
Bask 10.92 Peter		20,01	18,20	18,04	20,74	21,38	29,30	40,14	29,68	21,11	23,16	25.34
18.10		23,84	11,14	18,77	20,81	23,03	11,77	15,70	14,02	26,34	32,54	27,53
		24,41	19,25	17,31	17,53	30,45	16,67	22,23	18,09	31,19	34,62	33,81

El Apéndice F informa ΓdB para todas las novelas consideradas en el artículo.

Hagamos algunas observaciones fundamentales sobre la Tabla 13, aplicables a cualquiera que sea el texto de referencia. Las relaciones señal-ruido de los canales C son prácticamente las más grandes, que van desde 19,17 dB (Lilith) a 31,19 dB (Atrás). Estos resultados simplemente dicen que todos Los autores utilizan el mismo idioma y escriben textos del mismo tipo, es decir, novelas, no ensayos.

Machine Translated by Google

o artículos científicos/académicos. Estos canales no son aptos para distinguir o evaluar grandes diferencias entre textos o autores.

En los otros tres canales, podemos notar que Trilogy, Back y Lilith tienen las relaciones señal-ruido más grandes, alrededor de 19 a 22 dB; por lo tanto, estas novelas son muy similares a Lord. En otras palabras, estos canales parecen confirmar la probable influencia de MacDonald tanto en Lord como en Trilogy y la conexión entre Lord y Trilogy.

Por el contrario, Narnia muestra valores pobres en el Canal S (10,12 dB) y el Canal WI (7,94 dB). Estos valores bajos están determinados por el ruido de correlación porque R = Rm + Rr ≈ Rr .

Si consideramos solo Γm,dB (es decir, solo la línea de regresión), entonces notamos una fuerte conexión con Lord ya que Γm,dB = 51,26 dB. Como ya hemos observado en la Figura 6, las líneas de regresión son prácticamente idénticas pero la dispersión de los datos no lo es. Lewis en Narnia es menos “regular” que en Trilogy o Tolkien en Lord a la hora de dar forma (inconscientemente) a estos dos canales lingüísticos.

10. Resumen y conclusiones

Los estudiosos de la literatura inglesa dicen unánimemente que JRR Tolkien influyó en CS Los escritos de Lewis. Por primera vez, hemos investigado este tema matemáticamente utilizando un análisis multidimensional original de parámetros lingüísticos, basado en las variables superficiales profundas del lenguaje y los canales lingüísticos.

Para situar nuestra investigación en el marco de la literatura inglesa, también hemos considerado algunas novelas escritas por autores anteriores, como Charles Dickens y otros, incluido George MacDonald, porque los estudiosos mencionan su probable influencia en Tolkien y Lewis.

En nuestro análisis multidimensional, en nuestra opinión, sólo la serie de palabras, oraciones e interpunciones por capítulo fueron planificadas conscientemente por los autores y, específicamente, no indican conexiones fuertes entre Tolkien, Lewis y MacDonald. Cada autor distribuyó palabras, frases e interpunciones de manera diferente de un autor a otro y, en ocasiones, incluso de una novela a otra del mismo autor.

Por el contrario, las variables profundas del lenguaje y los canales lingüísticos, discutidos en el artículo, probablemente se deban a un diseño inconsciente y pueden revelar conexiones entre textos que van mucho más allá de la conciencia de los escritores.

En resumen, los amortiguadores de la memoria extendida a corto plazo que requieren los lectores, el índice de legibilidad universal de los textos, la representación geométrica de los textos y el ajuste fino de los canales lingüísticos (todas herramientas ampliamente discutidas en el artículo) han revelado fuertes conexiones entre El Señor de los Anillos (Tolkien), Las Crónicas de Narnia y La Trilogía Espacial (Lewis) por un lado, y la fuerte conexión también con algunas novelas de MacDonald por el otro, por lo tanto sustancialmente de acuerdo con lo que dicen los estudiosos de la literatura inglesa.

Financiamiento: Esta investigación no recibió financiamiento externo.

Declaración de la Junta de Revisión Institucional: No aplicable.

Declaración de Consentimiento Informado: No aplicable.

Declaración de disponibilidad de datos: los datos están contenidos en el artículo.

Agradecimientos: El autor desea agradecer a los numerosos académicos que, con gran cuidado y amor, mantienen textos digitales disponibles para lectores y académicos de diferentes disciplinas académicas, como la Biblioteca Digital Perseus y el Proyecto Gutenberg.

Conflictos de intereses: El autor declara no tener conflictos de intereses.

Apéndice A. Límite inferior del índice medio de legibilidad universal

El valor medio de GU viene dado por

$$< GU >= 89–10k< CP > +300 < 1/PF >–6(< IP > –6) \quad \text{(A1)}$$

Machine Translated by Google

El valor calculado introduciendo la media de las variables viene dado por

$$GU,media = 89–10k< CP > +300/< PF >–6(< IP > –6) \quad (A2)$$

Por lo tanto

$$< GU > –GU,media = 300 \times \frac{1}{<PF>} – < \frac{1}{FP} > \quad (A3)$$

Ahora bien, se puede demostrar con la desigualdad de Cauchy-Schwarz que 1/< x >≥< 1/x >; por lo tanto < GU > –GU,media ≥ 0; por lo tanto < GU >≥ GU,media.

Apéndice B. Distancia pitagórica d entre pares de textos

Tabla A1. Distancia pitagórica d entre pares de textos.

Novela Lord Hobbit Narnia Trilogía Espalda Lilith Oliver David Bleak										Cuento Mutual Martin Bask Peter			
señor 0													
Hobbit 0,488 Narnia	0												
0,150 Trilogía 0,355	0,348	0											
Atrás 0,099 Lililth	0,140	0,211	0										
0,112 Oliver 0,244	0,499	0,203	0,379	0									
David 0,320 Desolado	0,453	0,174	0,336	0,048	0								
0,231 Cuento 0,267	0,518	0,307	0,426	0,146	0,141	0							
Mutual 0,146 Martin	0,620	0,407	0,532	0,222	0,234	0,106	0						
0,230 Bask 0,424	0,312	0,170	0,217	0,201	0,153	0,211	0,316	0					
Peter 0,098	0,381	0,252	0,305	0,201	0,161	0,143	0,239	0,096	0				
	0,479	0,218	0,369	0,055	0,044	0,099	0,190	0,170	0,151	0			
	0,2612	0,109	0,139	0,241	0,197	0,294	0,400	0,096	0,192	0,230	0		
	0,0964	0,277	0,072	0,451	0,408	0,496	0,602	0,286	0,371	0,441	0,211	0	
	0,4744	0,183	0,355	0,025	0,024	0,146	0,232	0,177	0,182	0,048	0,217	0,427	0

Apéndice C. Área común entre círculos

Listamos el código de Matlab para calcular el área común entre texto, descargado de https://it.mathworks.com/matlabcentral/answers/273066%E2%80%93overlapping%E2%80%93área%E2%80%93entre%E2%80%93dos%E2%80%93círculos (accedido en 15 de junio de 2024).

Sea d la distancia entre los centros de dos círculos y sus dos radios sean r1 y r2.

Entonces, el área, A, de la región de superposición de los dos círculos se puede calcular de la siguiente manera usando Función a'tan2' de Matlab:

```
t = raíz cuadrada ((d+r1+r2)*(d+r1-r2)*(d-r1+r2)*(-d+r1+r2));

A = r1^2*atan2(t,d^2+r1^2-r2^2)+r2^2*atan2(t,d^2-r1^2+r2^2)-t/2;
```

Apéndice D. Probabilidad de error condicional

Tabla A2. Probabilidad de error entre los textos indicados.

	Caballero	Trilogía Hobbit Narnia Volver	Lilith	óliver david		Desolado	Cuento	Mutual Martín Bask			Pedro
Lord	—-										
Hobbit 0,065 Narnia		—-									
0,665 Trilogía 0,320		0,387 —-									
0,711 0,701 0,314		0,759 0,627 —-									
Atrás	David	0,148 0,603 0,335 —-									
Lilith	0,033	0,173 0,639 0,382 0,866 —-									
óliver		0,051 0,358 0,203 0,666 0 0,101 0 0,496 0,387	0,641	—-							
Desolado 0,307		0,633 0,593 0,543 0,270 0,475 0,428 0,554	0,372	0,664	—-						
Cuento 0,280 Mutual		0,036 0,545 0,279 0,633 0,499 7 0,716 0,478	0,560	0,397	0,040	—-					
0,505 Martin 0,395		0,830 0,504 0,873 0,211 0,231 0,669 0,412	0,601	0,615	0,300	0,725	—-				
Bask 0,164 Peter		0,853	0,681	0,686	0,273	0,428	0,558 —-				
0,649			0,528	0,273	0,003	0,732	0,517	0,367 —-			
			0,241	0,077		0,439	0,286	0,095	0,591 —-		
			0,761	0,687	0 0,568	0,642	0,644	0,594	0.587	0,294 —-	

Machine Translated by Google

Apéndice E. Pendiente y coeficiente de correlación de las líneas de regresión

Cuadro A3. Pendiente/coeficiente de correlación de la recta de regresión y = mx, Ecuación (18), modelando la variables indicadas (dependientes/independientes). Mantenemos cuatro dígitos porque algunas novelas sólo difieren en el tercer y cuarto dígito.

Novedoso	Oraciones/Palabras	Interpunciones/oraciones	Palabras/Interpunciones	Personajes/Palabras
Caballero	0,0731/0,9199	2,0372/0,9609	6,6134/0,9609	4,0367/0,9982
hobbit	0,0608/0,9532	2,1010/0,9936	7,6902/0,9532	4,1014/0,9996
narnia	0,0729/0,7610	1,9520/0,9384	6,9062/0,7991	4.0907/0.9919
Trilogía	0,0672/0,9325	1,9664/0,9830	7,3380/0,9696	4.2002/0.9976
Atrás	0,0681/0,9416	2,1640/0,9759	6,6045/0,9799	3,8496/0,9976
Lilith	0,0676/0,8890	2,2619/0,9488	6,2926/0,9800	4,1174/0,9863
óliver	0,0566/0,9059	3,0893/0,9544	5,6302/0,9638	4,2248/0,9977
David	0,0537/0,9390	3,2949/0,9657	5,5775/0,9882	4,0474/0,9966
Desolado	0,0600/0,9258	2,5324/0,9715	6.51250.9694	4,2235/0,9923
Cuento	0,0573/0,9574	2,7972/0,9785	6,1323/0,9912	4,2417/0,9983
Mutual	0,0618/0,9299	2,6814/0,9549	5,9766/0,9740	4,2197/0,9940
Martín	0,0658/0,8583	2,2243/0,9364	6,6785/0,9514	4,3171/0,9939
Disfrutar	0,0684/0,7706	1,8517/0,9366	7,6984/0,9005	4,1320/0,9949
Pedro	0,0687/0,8686	2,3080/0,9679	6.1018/0.9152	4,1117/0,9968

Apéndice F. Relación señal-ruido total ΓdB, en los cuatro canales lingüísticos

Las tablas A4 a A7 informan la relación señal-ruido ΓdB en los canales entre el texto de entrada k (referencia) se informa en la primera fila y el texto de salida j se informa en la columna de la izquierda. Para Por ejemplo, en la Tabla A4, si el texto de entrada es Lord y el texto de salida es Trilogy, entonces ΓdB = 21,27 dB; viceversa, ΓdB = 20,44. Una ligera asimetría es típica de los canales lingüísticos [12,15].

Cuadro A4. Relación señal/ruido total ΓdB, Canales S.

Novela	Señor	Hobbit	Narnia	Trilogía	Volver	Lilith	Oliver	David	Señor desolado	∞	12,65	10,08	20,44	18,92	10,61	Cuento	Mutual	Martin	Bask	Peter
8,68	13,19	10,18	14,63	14,51	Hobbit	14,60	∞	8,21	19,11	14,25	15,97	17,00	1,64	24,02	23,05	12,73	Narnia	10,12	5,30	∞	8,21	11,54	6,81	4,19	6,85	9,79	17,14
4,15	7,11	13,32	Trilogía	21,27	18,00	9,58	∞	30.77	19.49	13.81	19.02	18.29	14.21	21.14	15.10	Back	21.10	17.94	8.87	30.56	∞	17.45	12.67	11.43	16.83	8.50	13.08
14.07	19.31	13.65	Lililth	19.92	13.16	12.79	19.39	17.58	∞	13.97	20.37	15.86	10.98	16.86	23.03	Oliver	12.87	17.07	10.19	15.51	14.60	15.61	∞	19.50	23,39	13,54
22.67	16.83	19.93	15.65	David	11.43	18,20	8,42	13,92	13,49	12,83	20,29	∞	19,17	21,60	17,53	12,38	9,69	16,57
8,86	15,12
13,29	26,92
11.20	14.50
9,09	11,86
Desolado	14,91	21,86	9,79	19.30	18.05	17,27	21,95	18,12	∞	19,17	30,17	15,67	10,43	15.34
Cuento	12,66	24,57	7,84	15,85	15,69	13.21	16,61	20,78	19,89	∞	19,45	11,91	8.23	11.93
Mutual	16,13	22,78	9.70	21,87	20,24	18.07	18,92	16,27	29,88	18,43	∞	15,64	10,19	15,76
Martín	16,00	11,43	14,86	15,46	14,23	23,46	13,93	9,79	14.29	9,78	14,62	∞	16,34	26,65
Tomar el sol	10,92	6.53	23,97	9.38	8.78	13.09	8.49	5.56	8.33	5.38	8.51	15,69	∞	15,21
Pedro	18.10	11.21	14.52	16.19	14,96	26,66	12,56	9.04	13.59	9.38	14,24	26,20	15.13	∞

Cuadro A5. Relación señal/ruido total ΓdB, I–Canales.

Novela	Lord	Hobbit	Narnia	Trilogía	Atrás	Lilith	Oliver	David	Bleak	Lord	∞	15,57	21,18	9,35	Cuento	Mutual	Martin	Bask	Peter
19.50	21,74	19.50	8.36	14,05	11,16	12,37	19,14	17,61	18.44
Hobbit	15.04	∞	11.25	19.77	19.33	13.61	9.23	8.52	13,97	11,74	11,60	12,29	10,07	16,07
Narnia	21,83	12,48	∞	15,46	16.18	17.10	8.65	7,72	12.20	9.93	11,23	18,24	25,28	15.11
Trilogía	20,07	20,66	15,33	∞	20,28	15.03	8.61	7,83	12,86	10,53	11,05	14,25	14.16	15,95

Machine Translated by Google

Cuadro A5. Cont.

Novela Lord Hobbit Narnia Trilogía Volver Lilith Oliver David Desolado Volver 20,96 18,83 14,72 19,35 ∞ 19,45	Cuento Mutual Martin Bask Peter
10,31 16,69 12,90 13,81 9.26	17.18 12.61 23.09
Lilith 18,47 12,40 15,76 13,22 18,74 ∞ 11,43 10,00 17,75 13,45 16,06 27,65	12,92 23,22
Óliver 5.72 5.21 4.61 4.42 7.06 8,72 ∞ 22,73 12,64 16,56 16,36	8.04 3.42 9.25
David 4.18 4.22 3.04 3.25 5.57 6.66 22,01 ∞ 10,38 14,45 12,63	5,99 1,96 7.38
Desolador 12,10 11,84 9.57 10.56 15.30 16.42 14.53 12,69 ∞ 20,10 21,84 14,23	7.91 20.13
Cuento 8.26 9.01 6.36 7.46 10.66 11.20 17,84 16,02 19,15 ∞ 19,40	9,88 5.04 13.18
Mutuo 9.97 8.68 8.40 8.02 11.71 14.57 17,59 14,48 21,08 20,07 ∞ 13,28	6.83 15.34
Martín 18,04 11,32 17,10 12.46 16,71 27,92 10,97 9.59 15,87 12,48 15,04 ∞ 13,93	19.38
Tomar el sol 18,77 12,04 25,74 15,14 14,57 14,71 7,92 7.11 10,98 9.06 10,12 15,52 ∞ 13,36	
Pedro 17,31 14,69 13,28 14.34 22,40 22,88 11,85 10,47 20,95 14,94 16,77 18,77 11,10 ∞	

Cuadro A6. Relación señal/ruido total ΓdB, WI–Canales.

Novela Lord Hobbit Narnia Trilogía Volver Lilith Oliver David Bleak Lord ∞ 16,96 8,69 Hobbit 15,61 ∞ 7,80	Cuento Mutual Martin Bask Peter
19,72 21,94 20,11 15,14 12,42 28,75 14,96 18,33 29,47 13,83	15.46
22.03 13.72 11.77 8.65 7.24 14.25 9.46 10,34 16,38	16,86 10.64
Narnia7.94 9,59 ∞ 7,96 6.03 5.48 5.46 3.05 6,96 3.80 5.44 8.92	13,77 10,57
Trilogía 18,74 22,69 Atrás 6,92 ∞ 18,24 15.14 10.32 9.40 17,94 12,40 12,81 18,24 13,92	10.84
21,96 15,48 Lililth 20,86 13,90 6.80 19,30 ∞ 26,10 14,37 14,32 26,03 19,46 19,22 19,02	11.69 12.07
Oliver 16,54 11,40 12,64 18,25 7.07 16.59 26,52 ∞ 17,01 17,17 24,86 22,59 24,34 18,32	11.15 12,89
∞ 18,55 17,28 16,43 23,12 15,86 David 8.64 10,55 12,03 .35,97 15,72 ∞ 15.66 20.68 20.29 13.45 6.50 Sombrío 29.00 15.86 7.97 18.98 26.27 24.36	10,46 16.18
15.99 13.95 ∞ 17.22 20.71 23.33 Cuento 6.13 1 2,16 5,82 15,93 20,40 23,01 15,15 19,82 18,16 ∞ 19,83 14,60 Mutual 19,40 12,72 7,83 14,60 20,15	9,04 11.74
22,43 19,37 21,49 20,26 ∞ 17,75 Martín 29. 30 17,61 9,50 19,33 18,82 Tomar el sol 11,77 16,84 12,10 13,11 9,38 Pedro 16,67 13,01 12,48 13,17	12,70 14.34
13,36	9,60 10.71
	10,94 14.62
17,40 14,28 11,19 22,92 13,25 16,43 ∞ 14,97 16,69	
8,32 7,06 4,92 10,38 6,29 7,77 13,14 ∞ 11,52	
13,42 14,94 10,26 15,40 10,80 14,27 17,92 13,59 ∞	

Cuadro A7. Relación señal/ruido total ΓdB, Canales C.

Novela Señor Hobbit Narnia Trilogía Volver Lilith Oliver David Señor desolado ∞ 29,12 23,37 27,97 26,10 Hobbit	Cuento Mutual Martin Bask Peter
28,87 ∞ 20,03 26,66 22,08 Narnia 23,14 20,07 ∞ 24,11 21,11 19,51 26,91 32,93 22,42 26,31 23,84 21,90 26,70 31,43	
28,25 23,63 30,01 22,43 29,10 25,22 Trilogía 27,61 26,30 23,69 ∞ 17,21 26,33 24,95 20,20 27,13 21,51 20,64 22,77 25,73	
20,81 19,94 44,44 27,92 25,21 36,70 27,87 26,30 Atrás 26,52 22,80 21,89 21,57 ∞ 19,06 21,03 25,94 19,85 20,63 20,44 18,83 Lilith 19,17 17,15 28,14	31,08 26,48
20,28 18,09 ∞ 20,09 21,25 26,43 19. 28 24,45 23,08	28,81 32,34
	22,56 23,80
	23,79 21,31
Óliver 26,50 25,91 23.13 44,38 20.22 19,65 ∞ 26,67 24,97 39,76 27,56 26,57 27,81 30,39	
David 32,88 25,17 26,81 28.28 25,48 21,54 27,08 ∞ 24,77 25,80 26,30 23,38	31,25 36,01
Desolador 21,76 19,71 29,73 25.11 18,84 26,05 24,98 24,22 ∞ 23,63 36,70 31,88	29,79 25,53
Cuento 25,88 26,71 21,87 36,57 19,78 18,78 39,69 25,31 23,56 ∞ 25,70 25,32	25,80 28,24
Mutual 23,24 21,04 28,76 27,79 19,51 24,05 27,59 25,83 36,72 25,79 ∞ 32,93	32,78 27,93
Martín 21,11 19,88 24,71 25,90 17,72 22,43 26,23 22,73 31,63 25,03 32,73 ∞ 26,77 24,53	
Tomar el sol 26,34 22,64 30,92 29,07 21,84 23,73 28,13 30,99 30,08 26,19 32,99 27,17 ∞ 33,27	
Pedro 31,19 25,68 26,39 32,56 23,22 21,33 30,67 35,87 25,93 28,59 28,28 25,08	33.36 ∞

Referencias

1. Carpenter, H. Los Inklings: CS Lewis, JRR Tolkien, Charles Williams y sus amigos; Houghton Mifflin: Boston, MA, EE. UU., 1979.
2. Glyer, DP La compañía que mantienen. CS Lewis y JRR Tolkien como escritores comunitarios; Prensa de la Universidad Estatal de Kent: Kent, OH, Estados Unidos, 2007.
3. Duríez, C.; Porter, D. El manual de los Inklings: una guía completa de la vida, el pensamiento y los escritos de CS Lewis, JRR Tolkien, Charles Williams, Owen Barfield y sus amigos; Chalice Press: Nashville, TN, EE. UU., 2001.
4. Isley, WLCS Lewis sobre la amistad. Inklings Forever 2008, 6, 9. Disponible en línea: https://pillars.taylor.edu/inklings_forever/vol6/iss1/9 (consultado el 30 de mayo de 2024).
5. Sammons, MC La guerra de los mundos de fantasía: CS Lewis y JRR Tolkien sobre el arte y la imaginación; Praeger: Westport, CT, EE. UU., 2010.
6. Duriez, C. Los Oxford Inklings. Lewis, Tolkien y su círculo; Lion Books: Oxford, Reino Unido, 2015.

Machine Translated by Google

7. Hooper, W. Los Inklings. En CS Lewis y su círculo. Ensayos y memorias de la Oxford CS Lewis Society; White, R., Wolfe, J., Wolfe, BN, Eds.; Oxford University Press: Nueva York, NY, EE.UU.; Oxford, Reino Unido, 2015; págs. 197-213.
8. Gokulapriya, TJRR Obras literarias de Tolkien: una revisión. Int. Rev. iluminado. Semental. 2022, 4, 31–39.
9. Matricciani, E. Estadísticas profundas del lenguaje italiano a lo largo de siete siglos de literatura y conexiones empíricas con Ley 7 2 de Miller y memoria a corto plazo. Abrir J. Stat. 2019, 9, 373–406. [Referencia cruzada]
10. Matricciani, E. Una teoría estadística de la traducción de idiomas basada en la teoría de la comunicación. Abrir J. Stat. 2020, 10, 936–997. [Referencia cruzada]
11. Matricciani, E. Múltiples canales de comunicación en textos literarios. Abrir J. Stat. 2022, 12, 486–520. [Referencia cruzada]
12. Matricciani, E. ¿La memoria a corto plazo está formada por dos unidades de procesamiento? Pistas de las literaturas italiana e inglesa en varios Siglos. Información 2024, 15, 6. [CrossRef]
13. Matricciani, E. Una estructura matemática subyacente a las oraciones y su conexión con la memoria a corto plazo. Aplica. Matemáticas 2024, 4, 120-142. [Referencia cruzada]
14. Miller, GA El número mágico siete, más o menos dos. Algunos límites a nuestra capacidad para procesar información. Psicólogo. Rev. 1956, 63, 81–97. [Referencia cruzada] [PubMed]
15. Matricciani, E. Los índices de legibilidad no lo dicen todo sobre la legibilidad de un texto. Análisis 2023, 2, 296–314. [Referencia cruzada]
16. Matricciani, E. Relaciones matemático-lingüísticas guardadas o perdidas al traducir textos: extensión de la teoría estadística de Traducción y su aplicación al Nuevo Testamento. Información 2022, 13, 20. [CrossRef]
17. Papoulis Papoulis, A. Probabilidad y estadística; Prentice Hall: Hoboken, Nueva Jersey, EE. UU., 1990.
18. Lindgren, Teoría estadística de BW, 2ª ed.; Compañía MacMillan: Nueva York, NY, Estados Unidos, 1968.

Descargo de responsabilidad/Nota del editor: Las declaraciones, opiniones y datos contenidos en todas las publicaciones son únicamente de los autores y contribuyentes individuales y no de MDPI ni de los editores. MDPI y/o los editores renuncian a toda responsabilidad por cualquier daño a personas o propiedad que resulte de cualquier idea, método, instrucción o producto mencionado en el contenido.

Machine Translated by Google

AppliedMath

MDPI

Artículo

Método pseudoespectral de Chebyshev para diferencial fraccional Ecuaciones en dominios particionados no superpuestos

Shina Daniel Oloniiju [1,*], Nancy Mukwevho [1,2], Yusuf Olatunji Tijani [1] y Olumuyiwa Otegbeye [3]

[1] Departamento de Matemáticas, Universidad de Rhodes, Makhanda, PO Box 94, Grahamstown 6140, Sudáfrica; nancy.mukwevho@univen.ac.za (Nuevo México); yusuf.tijani@ru.ac.za (YOT)
[2] Departamento de Ciencias Matemáticas y Computacionales, Universidad de Venda, Private Bag X5050, Thohoyandou 0950, Sudáfrica 3
Escuela de Ciencias de la Computación y Matemáticas Aplicadas, Universidad de Witwatersrand, Private Bag 3, Braamfontein, Johannesburgo 2050, Sudáfrica; olumuyiwa.otegbeye@wits.ac.za
* Correspondencia: s.oloniiju@ru.ac.za

Resumen: Los operadores diferenciales fraccionarios son inherentemente no locales, por lo que los métodos globales, como los métodos espectrales, son muy adecuados para manejar estos operadores no locales. La integración a largo plazo de modelos diferenciales, como los sistemas dinámicos caóticos, plantea desafíos y consideraciones específicos que hacen que los métodos numéricos multidominio sean ventajosos al abordar tales problemas. Este estudio propone un nuevo método pseudoespectral multidominio basado en el primer tipo de polinomios de Chebyshev y la cuadratura de Gauss-Lobatto para problemas de valores iniciales fraccionarios. La técnica propuesta implica dividir el dominio del problema en subdominios no superpuestos, calculando el diferencial fraccionario. operador en cada subdominio como la suma de las partes 'local' y 'memoria' y derivando las matrices de diferenciación correspondientes para desarrollar los esquemas numéricos. El análisis de estabilidad lineal indica que el esquema numérico es absolutamente estable para ciertos valores de orden arbitrario no entero y condicionalmente estable para otros. Los ejemplos numéricos, que van desde ecuaciones lineales simples hasta sistemas de ecuaciones no lineales, demuestran que el enfoque multidominio es más apropiado, eficiente y preciso que el esquema de dominio único, particularmente para problemas con dinámica de largo plazo.

Palabras clave: multidominio; método pseudoespectral; ecuaciones diferenciales fraccionarias; polinomios de Chebyshev ; cuadratura de Gauss-Lobatto; análisis de estabilidad

check for updates

Cita: Oloniiju, SD; Mukweho, N.; Tijani, YO; Otegbeye, O. Método pseudoespectral de Chebyshev para ecuaciones diferenciales fraccionarias en dominios particionados no superpuestos. Matemáticas aplicadas 2024, 4, 950–974. https://doi.org/10.3390/appliedmath4030051

Editores académicos: Renhai Wang y Junesang Choi

Recibido: 26 de junio de 2024
Revisado: 29 de julio de 2024
Aceptado: 30 de julio de 2024
Publicado: 2 de agosto de 2024

Copyright: © 2024 por los autores. Licenciatario MDPI, Basilea, Suiza. Este artículo es un artículo de acceso abierto. distribuido bajo los términos y condiciones de los Creative Commons Licencia de atribución (CC BY) (https:// creativecommons.org/licenses/by/4.0/).

1. Introducción

Las ecuaciones diferenciales fraccionarias (FDE) están ganando mucha atención en la ingeniería , la ciencia, las finanzas y otros campos. A diferencia de las ecuaciones diferenciales clásicas, que se limitan a derivadas de orden entero, las FDE introducen un mayor grado de complejidad en las ecuaciones diferenciales al incorporar derivadas arbitrarias que no son de orden entero. Los operadores diferenciales fraccionales se pueden definir de varias maneras, siendo las definiciones más comunes aquellas que involucran núcleos singulares, como los operadores diferenciales fraccionales de Riemann-Liouville y Caputo [1,2]. Estas definiciones se basan en integrales con núcleos de ley potencial, que a veces pueden conducir a singularidades y complicar las aproximaciones numéricas. Sin embargo, estos operadores diferenciales con núcleos de leyes de potencia son ventajosos para desarrollar métodos basados en espectros utilizando funciones básicas como los polinomios de Chebyshev y Legendre, ya que estos polinomios se pueden expresar en forma de series de potencias. Por lo tanto, la derivada de Caputo será utilizada en el método numérico propuesto en este estudio. Sin embargo, en los últimos años ha habido un interés significativo en los núcleos no singulares, que proporcionan un marco alternativo para definir operadores fraccionarios. Los ejemplos incluyen los operadores fraccionarios Caputo-Fabrizio y Atangana-Baleanu, que utilizan las funciones exponencial y Mittag-Leffler como núcleos, respectivamente [3,4]. Estos núcleos no singulares se introdujeron para ampliar el alcance y la aplicabilidad

Machine Translated by Google

Esta complejidad adicional permite a los operadores diferenciales fraccionarios capturar los efectos de la memoria de un sistema o describir fenómenos inherentes a la memoria. Como resultado, los FDE ofrecen un marco más sólido para modelar la dinámica intrincada de sistemas complejos en diversas disciplinas científicas [5,6]. Encontrar soluciones analíticas para ecuaciones diferenciales fraccionarias (FDE) es particularmente desafiante; sin embargo, varios estudios han propuesto soluciones y métodos analíticos para FDE, por ejemplo, Ahmadian y Darvishi [7] derivaron solitones y soluciones de ondas viajeras utilizando los métodos de expansión tanh-coth, tan-cot, sech-csch y sec-csc para (Ecuación fraccionaria de Biswas-Milovic 1 + 1) dimensional [8]. Kumar y cols. [9] propusieron un método de descomposición de Laplace modificado, que hibrida la transformada de Laplace y los métodos de descomposición de Adomian para analizar la ecuación fraccionaria de Navier-Stokes que modela el flujo inestable de un fluido newtoniano en un tubo. La ecuación de onda larga regularizada simétrica fraccionaria se ha resuelto analíticamente utilizando métodos como el método de expansión G'/G [10], el método de expansión G'/G [11], el método complejo extendido [11], el método algebraico directo método [12] y el método de la función elíptica Otros métodos analíticos que se han utilizado para FDE incluyen la transformada de Sumudu [14], el método de Kudryashov [15] y varias técnicas semianalíticas como el método de expansión exp [16,17] y el método seno-coseno [18].], entre otros. Los métodos numéricos también desempeñan un papel importante en la obtención de soluciones precisas para los FDE y proporcionan información valiosa sobre la dinámica de los sistemas descritos por los FDE [19, 20].

Se han propuesto muchos métodos numéricos para aproximar las soluciones de FDE. Por ejemplo, Hattaf et al. [21] introdujeron un enfoque numérico novedoso para aproximar la derivada fraccionaria generalizada de Hattaf (GHF), que implica un operador diferencial fraccionario con un núcleo no singular. El método numérico desarrollado se probó en FDE con soluciones analíticas para verificar su precisión comparando las soluciones aproximadas con las soluciones exactas. También se resolvieron FDE no lineales que describen la dinámica de la infección por VIH para demostrar la aplicabilidad del método a problemas del mundo real. Los métodos predictor-corrector (PCM) para FDE fueron propuestos por Kumar et al. [22], Jhinga y Daftardar-Gejji [23] y Lee et al. [24]. Estos estudios demostraron que los PCM resuelven FDE más rápido y reducen drásticamente el tiempo de cálculo. Kumar y cols. [25] introdujeron dos técnicas numéricas computacionalmente eficientes basadas en el punto medio y el método de Heun para problemas de valores iniciales fraccionarios (IVP) lineales y no lineales. El estudio concluyó que los métodos propuestos ofrecen ventajas significativas sobre los métodos tradicionales y mejorados de Euler.

La naturaleza no local de los operadores diferenciales fraccionarios a veces hace que el análisis y la búsqueda de soluciones de FDE sean tareas hercúleas [26]. Los métodos espectrales han sido una poderosa técnica numérica para aproximar las soluciones de FDE [27-29]. A diferencia de otras técnicas numéricas tradicionales, los métodos basados en espectros son globales y altamente precisos, lo que los hace eficientes para resolver ecuaciones diferenciales [30]. Numerosos estudios han aprovechado la naturaleza global de los métodos basados en espectros para discretizar operadores diferenciales fraccionarios y obtener soluciones numéricas de FDE [31]. Zayernouri y Karniadakis [32] y Zhao et al. [33] propusieron un método espectral de Galerkin exponencialmente convergente para FDE. Zayernouri y Karniadakis [32] compararon el rendimiento del método basado en espectros desarrollado y el método de diferencias finitas (FDM) establecido y concluyeron que el método propuesto funciona mejor que el FDM en términos de tasa de convergencia. Khader y Sweilam [34] utilizaron el método pseudoespectral de Legendre para resolver la ecuación fraccionaria de advección-dispersión con el operador diferencial fraccional de Caputo. El estudio analizó la precisión del nuevo operador diferencial fraccionario discreto examinando su tasa de convergencia y estimando el error máximo posible, y los resultados fueron consistentes con los resultados existentes de la literatura. Cardone et al. resolvieron ecuaciones diferenciales fraccionarias estocásticas . [35] utilizando el método de colocación espectral. El método implica obtener una versión totalmente discreta del operador continuo original y resolver un sistema de ecuaciones algebraicas no lineales. Las pruebas sobre varios problemas demostraron que el método propuesto era eficaz para las funciones básicas y los puntos

Machine Translated by Google

Los métodos espectrales exhiben convergencia exponencial para ecuaciones diferenciales con soluciones suficientemente suaves. Este alto nivel de precisión ha hecho que los métodos espectrales, particularmente el método pseudoespectral de Chebyshev, sean una opción popular para resolver una variedad de ecuaciones diferenciales fraccionarias. Por ejemplo, Doha [36] utilizó el método pseudoespectral de Chebyshev para aproximar las soluciones de ecuaciones diferenciales fraccionarias de varios términos . De manera similar, Sweilam [37] aplicó este método para resolver numéricamente ecuaciones integrodiferenciales fraccionarias. Otros estudios que han utilizado el método pseudoespectral de Chebyshev para aproximar la solución de ecuaciones diferenciales fraccionarias incluyen los de Dabiri [38], Oloniiju et al. [39,40,41], entre otros. Sin embargo, esta propiedad deseable de convergencia espectral puede deteriorarse con grandes dominios de tiempo o soluciones no uniformes [42]. El enfoque multidominio aborda estos desafíos ofreciendo una forma precisa y computacionalmente eficiente de gestionar soluciones potencialmente no fluidas. También puede mejorar la estabilidad numérica y la precisión al dividir el dominio de un problema, especialmente los problemas que dependen del tiempo, en subdominios más pequeños [43]. Samuel y Motsa [44] describieron un método de colocación espectral multidominio para resolver ecuaciones diferenciales parciales (PDE) hiperbólicas con un dominio de tiempo grande. El estudio demostró que dividir el dominio computacional en subdominios más pequeños proporciona resultados precisos y reduce el tiempo computacional. Birem y Klein [45] presentaron un método espectral multidominio para la ecuación de Schrödinger. Jung y Olmos-Liceaga [46] introdujeron un método espectral multidominio no conforme y no superpuesto para simular soluciones de ondas viajeras en las ecuaciones de reacción-difusión. El método tiene tres características clave: precisión de alto orden y tasa de convergencia, malla adaptativa y técnica multidominio. Qin y cols. [47] propusieron un método de mínimos cuadrados de Legendre-Galerkin multidominio para ecuaciones diferenciales lineales con coeficientes variables. El estudio estableció dos propiedades clave del método: continuidad, que garantiza coherencia y suavidad en los puntos finales de cada subdominio, y coercitividad, que garantiza estabilidad y convergencia. Además, el estudio estableció la estimación óptima del error en la norma H1 .

Este estudio explora el método pseudoespectral de Chebyshev en un entorno multidominio como un enfoque novedoso para resolver FDE. Aquí, el dominio del problema se divide en un número finito de subdominios que no se superponen. El enfoque multidominio para las FDE se ha aplicado a problemas de biología [48] y finanzas [49,50]. Maleki y Kajani [48] utilizaron el método pseudoespectral multidominio de Legendre-Gauss para aproximar soluciones para el modelo fraccional de crecimiento poblacional de Volterra. El estudio demostró la eficacia del método propuesto para abordar una amplia gama de ecuaciones integrodiferenciales , incluidas aquellas con órdenes enteros y fraccionarios. Bambe Moutsinga et al. [49] desarrollaron un método basado en espectros para ecuaciones diferenciales fraccionarias que modela la valoración dentro de un marco de difusión de salto afín, sistemas caóticos e hipercaóticos y modelos de precios fraccionados para criptomonedas. El estudio estableció que el método espectral desarrollado supera a los paquetes numéricos de MATLAB basados en el método de Euler.
Kharazmi [50] investigó un enfoque basado en datos para resolver ecuaciones diferenciales parciales fraccionarias estocásticas (SFPDE) utilizando métodos espectrales multidominio. Zhao y cols. [51] propusieron un método de colocación espectral multidominio (MDSCM) para resolver FDE no lineales. El MDSCM mostró una ventaja significativa sobre los métodos espectrales de dominio único al lograr una alta precisión para ecuaciones diferenciales con soluciones de baja regularidad
Klein y Stoilov [52] propusieron una técnica numérica para el laplaciano fraccionario, basada en los operadores fraccionarios de Riesz, definidos en R. Los integrandos se transforman utilizando la curva Zq subyacente para garantizar un cálculo preciso dentro de cada subintervalo. Luego, las integrales transformadas dentro de cada subintervalo se resuelven utilizando el algoritmo de Clenshaw-Curtis. Tavassoli Kajani [53] propuso una técnica de colocación modificada de Müntz-Legendre en un solo dominio y en un entorno de múltiples dominios para la ecuación del pantógrafo fraccionario. Xu y Hesthaven [54] propusieron un método espectral multidominio estable y altamente preciso con un método de penalización para ecuaciones de difusión y advección parcial fraccionaria . El estudio concluyó que la alta precisión del esquema propuesto lo hace mu

Machine Translated by Google

para la integración a largo plazo de las EED. También se señaló que la tasa de convergencia se alinea con el análisis teórico del esquema.

El presente estudio tiene como objetivo presentar un método numérico novedoso, el método pseudoespectral de Chebyshev multidominio no superpuesto, basado en el primer tipo de polinomios de Chebyshev y la cuadratura de Gauss-Lobatto para ecuaciones diferenciales fraccionarias. La elección de los polinomios de Chebyshev se debe a su propiedad de ortogonalidad con respecto a su función de peso, que simplifica la proyección de la solución aproximada en el espacio polinomial. Además, el cálculo de los pesos y nodos de la cuadratura de Chebyshev-Gauss-Lobatto es bastante sencillo y esto reduce la sobrecarga computacional en comparación con otros polinomios ortogonales. El método propuesto funciona descomponiendo el dominio de una ecuación diferencial fraccionaria en subdominios más pequeños qu
La novedad de este estudio radica en su enfoque de dividir el dominio de un problema de valor inicial fraccionario en subintervalos no superpuestos y proyectar la solución del problema en el espacio de polinomios ortogonales de Chebyshev de primera especie. Esta técnica implica separar el operador diferencial fraccionario en la suma de los componentes "local" y "memoria" dentro de cada subdominio. Para demostrar la versatilidad y eficiencia del método propuesto, se llevan a cabo análisis de estabilidad y experimentación numérica sobre varios tipos de ecuaciones diferenciales fraccionarias. Se realiza un análisis de error numérico para demostrar las ventajas del enfoque de descomposición de dominio sobre el método pseudoespectral de Chebyshev de dominio único.

El resto del artículo está organizado de la siguiente manera: la Sección 2 proporciona ciertas definiciones preliminares que son necesarias para desarrollar los esquemas numéricos; La sección 3 está dedicada a desarrollar y analizar el esquema numérico para FDE utilizando los polinomios de Chebyshev en la cuadratura de Gauss-Lobatto en dominios particionados no superpuestos; La Sección 4 demuestra la precisión y eficacia del método numérico presentando algunos ejemplos que van desde ecuaciones lineales simples hasta sistemas de ecuaciones no lineales que requieren integración a largo plazo; y la Sección 5 finalmente concluye el estudio.

2. Preliminares

Esta sección presenta y define los conceptos básicos y los operadores fraccionarios necesarios para desarrollar el esquema numérico.

Definición 1. Supongamos que $u : [a, b] \to R$ es una función continua, $\gamma > 0$ y la integral de orden fraccional de Riemann-Liouville de u(t) se define como [55]

$$ {}^{RL}_{un}Y_t o\ \gamma\ tu(t) = \frac{1}{\Gamma(\gamma)}\ t\ \gamma-1\quad tu(t) = \frac{1}{\Gamma(\gamma)}\int_a^t (t-\tau)\ \gamma-1u(\tau)d\tau,\ t > 0,\ \gamma > 0, \quad u\ (t),\ \gamma = 0, \tag{1} $$

donde Γ es la función gamma de Euler y $t\ \gamma-1\quad u(t)$ es el producto de convolución de la función de potencia, $t\gamma-1$, y la función u(t).

Definición 2 (Integral fraccionaria de $(t - a) j$). Para la función de potencia $(t - a) j$ integral se , el fraccionario define como [56]

$$ {}^{RL}_{un}Y_t o\ \gamma\ (t-a)\ j\ =\ \frac{\Gamma(j+1)}{\gamma+1)}(t-a)\ j+\gamma\ \Gamma(j+\ \cdot \tag{2} $$

Definición 3. El operador diferencial fraccional de Caputo del lado izquierdo se define como [55]

$$ {}_{ca}re_t\gamma\ tu(t) = \frac{1}{\Gamma(\ \gamma\ -\gamma)}\ t\ \gamma\ -\gamma-1\ \ tu\ \ \gamma\ \ (t) = \frac{1}{\Gamma(\ \gamma\ -\gamma)}\int_a^t (t-\tau)\ \ \gamma-\gamma-1u\ \ \gamma\ \ (\tau)d\tau,\ \ \gamma\ \ -1 < \gamma \le\ \ \gamma\ \ , \tag{3} $$

Machine Translated by Google

donde γ es el techo de γ, γ ≥ γ y t γ−γ−1 función tu γ (t) es el producto de convolución de la de potencia, t γ−γ−1 , y la γ -ésima derivada de orden entero de la función u(t).

Definición 4 (Derivada fraccionaria de (t − a) j). La derivada fraccionaria de (t − a) j se define como [56]

$$ca^{re}{}_{\gamma}(t-a)^{j} = \begin{cases} 0 & j \in N_0, \quad j < \gamma \\ \frac{\Gamma(j+1)}{\Gamma(j+1-\gamma)}(t-a)^{j-\gamma} & j \in N_0, \quad j \geq \gamma \end{cases}. \tag{4}$$

Definición 5. El conjunto de puntos de colocación utilizado en este estudio es el de los puntos de Chebyshev–Gauss–Lobatto definidos como $t_r = -\cos\frac{\pi r}{N}, r = 0(1)N$ [57].

3. Discretización numérica y análisis mediante el método pseudoespectral

Esta sección presenta la discretización numérica de un problema de valor inicial fraccionario utilizando polinomios de Chebyshev de primera clase y la cuadratura de Gauss-Lobatto. Cubre las aproximaciones de la derivada fraccionaria de una función arbitraria, tanto en un solo dominio como en dominios divididos que no se superponen. En primer lugar, la aproximación polinómica de una función arbitraria u(t), continua en un intervalo [0, h], se presenta en términos de polinomios de Chebyshev desplazados de primer tipo. Estos polinomios se obtienen mediante la siguiente fórmula de recurrencia:

$$\Pi_{h,0}(t) = 1, \Pi_{h,1}(t) = \frac{2t}{h} - 1, \Pi_{h,j+1}(t) = 2\Pi_{h,1}(t)\cdot\Pi_{h,j}(t) - \Pi_{h,j-1}(t), j \geq 1. \tag{5}$$

La función u(t) se representa de la siguiente manera:

$$u(t) \quad U_h(t) = \hat{u}_j \sum_{j=0}^{\infty} \Pi_{h,j}(t), \quad t \in [0, h], \tag{6}$$

donde los coeficientes $\hat{u}_j$ en la ecuación (6) satisfacen la condición de ortogonalidad ℓ_2 de los polinomios de Chebyshev desplazados de primer tipo, que se definen con respecto al producto interno

$$\hat{u}_j = \langle U_h(t), \Pi_{h,j}(t) \rangle = v_j \frac{1}{h} \int_0^h \frac{U_h(t)\Pi_{h,j}(t)\,dt}{\sqrt{ht - t^2}} = \frac{1}{v_j} \sum_{i=0}^{\infty} U_h(t_i)\Pi_{h,j}(t_i) \omega_i, \tag{7}$$

dónde

$$v_j = \begin{cases} \pi & j = 0 \text{ en} \\ \frac{\pi}{2} & \text{caso contrario.} \end{cases} \tag{8}$$

Si se considera una aproximación truncada de orden N de la función y se coloca en N + 1 puntos distintos de Chebyshev-Gauss-Lobatto (CGL), con $t_r \in [0, h]; r = 0(1)N$, entonces

$$t_u(t) \quad \sum_{i=0}^{N} U_h(t_i) \omega_i \sum_{j=0}^{N} \frac{1}{v_j} \Pi_{h,j}(t_i)\Pi_{h,j}(t_r), \quad r = 0(1)N, \tag{9}$$

donde ω_i es el número de Christoffel asociado con la cuadratura de Chebyshev-Gauss-Lobatto y se define de la siguiente manera:

$$\omega_i = \begin{cases} \frac{\pi}{2N} & i = 0, N \\ \frac{\pi}{N} & \text{de lo contrario.} \end{cases} \tag{10}$$

Usando las siguientes representaciones matriciales

Machine Translated by Google

$$\rho = \begin{matrix} 0 & & \\ & \ddots & \\ & & N \end{matrix}, \quad = \begin{matrix} \frac{1}{v0} & & \\ & \ddots & \\ & & \frac{1}{\sqrt{1}\,vN} \end{matrix}, \quad \Pi h = \begin{matrix} \Pi h,0(t0) & \Pi h,0(t1) & \dots & \Pi h,0(tN) \\ \Pi h,1(t0) & \Pi h,1(t1) & \dots & \Pi h,1(tN) \\ \vdots & \vdots & \ddots & \vdots \\ \Pi h,N(t0) & \Pi h,N(t1) & \cdots & \Pi h,N(tN) \end{matrix}, \tag{11}$$

La ecuación (9) se puede expresar como

$$u(t) \quad \sum_{yo=0}^{norte} Uh\,(t)Ei(tr) = E\,TUh\,, \tag{12}$$

donde $E = \rho\,\Pi T \quad \Pi h$ y $Uh = [Uh\,(t0)\ Uh\,(t1) \quad \dots\ Uh\,(tN)]T.$

3.1. La discretización del operador diferencial fraccionario en un solo dominio

Ahora, aproximaremos el orden arbitrario no entero de la función u(t). Suponiendo que u(t) Cp ([0, h]), donde p ≥ N + γ + 1, que representa el espacio de funciones continuamente diferenciables, t > 0 y γ > 0, entonces la aproximación para las funciones arbitrarias no - la derivada entera de u(t) viene dada por

$$\begin{matrix} c \\ 0 \end{matrix} D\,\gamma\,tu(t) \quad \sum_{yo=0}^{norte} Uh\,(ti) \quad i \sum_{j=0}^{norte} \Pi h,j\,(ti)\frac{1}{\sqrt{vj}} \quad \begin{matrix} do \\ 0 \end{matrix} D\,\gamma\,t\,\Pi h,j\,(t). \tag{13}$$

La derivada de orden fraccionario del polinomio de Chebyshev $\begin{matrix} c \\ 0 \end{matrix}$ Se obtiene $D\,\gamma\,t\,\Pi h,j\,(t)$ utilizando la forma en serie del polinomio [56], de la siguiente manera:

$$\begin{matrix} c \\ 0 \end{matrix} D\,\gamma\,t\,\Pi h,j\,(t) = j \quad (j \quad \frac{(-1)\,j-k\,(j+k-1)!22k}{k)!(2k)!h} \sum_{k} \begin{matrix} c \\ 0 \end{matrix} re\,\gamma\,t^{k} = \sum_{jk=} \frac{(-1)\,j-k\,(j+k-1)!22k\,\Gamma(k+1)\,t\,k-\gamma\,(j-}{k)!(2k)!h\,\Gamma(k-\gamma+1)}. \tag{14}$$

Procediendo a aproximar $t^{k-\gamma}$ como una combinación lineal de polinomios de Chebyshev de primer tipo desplazados en los puntos CGL, la ecuación (14) ahora se puede expresar como [39]:

$$\begin{matrix} c \\ 0 \end{matrix} re\,\gamma\,t\,\Pi h,j\,(t) \quad hD\,(\gamma)\Pi h\,, \tag{15}$$

donde $hD(\gamma) = \Delta\,\Omega\,\Phi\,\Delta T$ y Δ, Ω y Φ son matrices cuyas entradas están definidas como sigue:

$$\Delta j,k = \begin{cases} \frac{1}{\frac{j(-1)\,j-k\,(j+k-1)!22j\,(j-k)!}{(2k)!h}\ \frac{}{k}} & j = 0,\ k = 0\ j \\ = 1(1)N,\ k = 0(1)j & \\ 0 & \text{de lo contrario} \end{cases}, \quad \Omega k,k = \begin{cases} 0 & k < \gamma \\ \frac{\Gamma(k+1)}{\Gamma(k-\gamma+1)} & k = \gamma(1)norte \end{cases},$$

$$\Phi p,k = \begin{cases} \frac{0\,\sqrt{}\pi h\ k-\gamma+p\Gamma(k-\gamma+p+\ \frac{1}{2})}{\Gamma(k-\gamma+p+1)} & k < \gamma \\ & k = \gamma(1)N,\ p = 0(1)s,\ s = 0(1)N. \end{cases} \tag{16}$$

Por lo tanto, ahora $\begin{matrix} do \\ 0 \end{matrix} re\,\gamma$ u(t) se aproxima a 0.

$$\begin{matrix} c \\ 0 \end{matrix} D\,\gamma\,t\,u(t) \quad (0,hD\gamma\,) \quad t \quad Oh\,, \tag{17}$$

donde $0,hD\gamma = \rho\,\Pi T \quad h \quad hD(\gamma)\,\Pi h$. Nos referimos a $(0,hD\gamma\,)^{t}$ como la diferenciación fraccionaria matriz en el dominio [0, h].

Machine Translated by Google

3.2. Discretización del operador diferencial fraccionario en dominios particionados no superpuestos

Supongamos que u(t) se define en t [0, T], dividido en q subintervalos no superpuestos, tales que 0 = t0 < t1 < . . . < tq = T y hl = tl − tl−1 , l = 1(1)q. Supongamos t [tl , tl+1] y γ − 1 < γ ≤ γ ; entonces,

$$
\begin{aligned}
{}_{0}^{C}re_{t}\gamma\, tu(t) &= \frac{1}{\Gamma(\ \gamma\ -\gamma)}\int_{0}^{t}(t-\tau)\quad \gamma-\gamma-1u\quad \gamma\quad (\tau)d\tau \\
&= \sum_{e=1}^{\nu} {}_{C\,te-1}D^{\gamma}\, tu(t) + te\quad {}_{C\,tl}re_{t}\gamma\, tu(t) \\
&= \sum_{e=1}^{\nu}\frac{1}{\Gamma(\ \gamma\ -\gamma)}\int_{te-1}^{te}(t-\tau)\quad \gamma-\gamma-1u\quad \gamma\quad (\tau)d\tau + \frac{1}{\Gamma(\ \gamma\ -\gamma)}\int_{tl}^{t}(t-\tau)\quad \gamma-\gamma-1u\quad \gamma\quad (\tau)d\tau.
\end{aligned}
\tag{18}
$$

El primer término de la ecuación (18) es la 'memoria', y el último término de la ecuación es la derivada local de orden arbitrario de u(t) con respecto a t [tl , tl+1]. La derivada de orden arbitrario en el primer subdominio es local y se calcula sin término de memoria. Los dos términos de la ecuación (18) se aproximan por separado. Si la longitud, hl , de cada subdominio, [tl , tl+1], es equivalente a h (de la discretizaciónen un solo dominio) en la Sección 3.1, y colocamos en N + 1 puntos CGL distintos, entonces t El segundo término de la ecuación (18) se aproxima de la siguiente manera: $\overset{\nu}{r}$ [tl , tl+1], con t $\overset{\nu}{norte} \equiv t\,0\,,^{l+1}$

$$
{}_{C\,tl}D^{\gamma}_{t}\, u(t) \quad (tl\ ,hlD\ {}^{\gamma\cdot}_{l})\ {}^{t}_{}\ Uhl\ ,l\ , \qquad t\quad [tl\ ,\ tl+1\ ]. \tag{19}
$$

El primer término de la ecuación (18), la 'memoria', que es la derivada arbitraria de orden no entero , ${}_{C\,te-1}D^{\gamma}\, u(t)$, se aproxima como te

$$
{}_{C\,te-1}re\ \gamma \sum\sum tu(t)\ {}_{yo=0}^{norte}\ Uh\ (t\ e\ i)\ \ i\ {}_{j=0}^{norte}\Pi hl\ ,j\ (t\ e\ )\ te \quad i\quad {}_{C\,te-1}D\ \gamma_{te}\ \Pi hl\ ,j\ (t),\qquad {}_{yo}^{y}\quad [te-1,\ te\ ]. \tag{20}
$$

La derivada te−1 $\Delta\ \overset{do}{\Theta},\quad D\ \gamma_{te}\ \Pi hl\ ,j\ (t)$ en los puntos t del CGL $\overset{\nu}{r}$ se aproxima como ${}_{C\,te-1}D\ \gamma_{te}\ \Pi hl\ ,j\ (t\ r\ )\ \overset{\nu}{}$ donde las entradas de Θ se definen como

$$
\Theta j,r = \frac{{}^{0}}{j!\ \Gamma(\ \gamma-\gamma)(j-\ \gamma\ )!}\quad te\ te-1\ (t\ \overset{\nu}{r}\ -\tau)\quad \gamma-1-\gamma\ (\tau-te-1)\ j-\quad \gamma d\tau, \qquad \begin{matrix} j<\ \gamma \\ {}^{l}r\quad [tl\ ,tl+1\ ] \\ j=\ \gamma(1)N,\ r=0(1)N. \end{matrix} \tag{21}
$$

La integral en la ecuación (21) se puede calcular usando cualquier fórmula de cuadratura, como las fórmulas de Newton-Cotes o de cuadratura gaussiana. Por lo tanto, el primer término de la ecuación (18) se calcula como

$$
\sum^{\nu}\ {}^{do}_{te\ e=1}D^{\gamma}\, u(t)\quad te-1\ \ \sum_{e=1}^{\nu}(te-1\ ,teD\ \ {}^{\gamma}_{mi}\ )\ {}^{t}_{}\ Uhl\ ,\ e\ , \tag{22}
$$

donde Uhl ,e [te−1, = [Uhl (t e $_{0}$) Uhl (t e) $\overline{\tau}$ . . . Uhl (t e $_{norte}$)]T, t $^{mi}_{i}$ ¿Están los puntos CGL asignados a te] y te−1 ,teD $^{\gamma}_{mi}$ ρ ΠT hl $\Delta\ \Theta$.

3.3. Discretización pseudoespectral de FIVP en dominios particionados no superpuestos

Considere el problema del valor inicial fraccionario (FIVP)

$$
\begin{aligned}
&{}_{0}^{C}D^{\gamma}_{t}\, tu(t) = \kappa u(t) + f(t), \qquad t\quad (0,\ T], \\
&u(0) = u0,
\end{aligned}
\tag{23}
$$

donde κ C y γ (0, 1). Supongamos que el dominio [0, T] se divide en q dominios no superpuestos, de modo que 0 = t0 < t1 < t2 < . . . < tq = T, y como antes, definimos hl = tl − tl−1 ,

Machine Translated by Google

l = 1(1)q. Para obtener la solución de la ecuación (23) en un dominio arbitrario [tl , tl+1], reescribimos la ecuación como

$$\sum_{e=1}^{do} te-1\, D\gamma\, tu(t) + c\, D\gamma\, tu(t) = \kappa u(t) + f(t), \quad t \in [tl , tl+1], \quad l = 1(1)q - 1, \tag{24}$$

$$u(0) = u0.$$

Supongamos que la solución u(t) de la ecuación (24) puede aproximarse como una combinación lineal de funciones que están en el espacio de polinomios de Chebyshev desplazados de grado N como máximo en la cuadratura CGL. Luego, utilizando las discretizaciones del operador diferencial dadas en las Secciones 3.1 y 3.2, el FIVP (24) ahora es equivalente a los siguientes sistemas algebraicos:

$$\begin{aligned} &(0,hlD\gamma 1\,)\,TUhl\,,1 = E\, f\, t1 \\ &Uhl\,(t0\,) = u0, \quad t \in [t0, t1 \equiv t0 + hl\,], \end{aligned} \tag{25}$$

y

$$\begin{aligned} &\sum_{e=1} (te-1\,,teD\gamma e\,)\,TUhl\,,e + (tl\,,hlD\gamma\,)\,TUhl\,,l \quad T = \kappa E_l\, Uhl\,,l + E\, T.f. \\ &Uhl\,(t0\,) = Uhl\,(t\,norte\,^{l-1}), \quad t \in [tl , tl+1 \equiv tl + hl\,], \quad l = 2(1)q - 1, \end{aligned} \tag{26}$$

donde f = [f(t 0), f(t 1), . . . , f(t norte)]T y l = 1(1)q.

3.4. La estabilidad numérica del esquema

La estabilidad del método numérico propuesto se demuestra considerando el siguiente problema de valor inicial fraccionario lineal

$$c0\, D\gamma\, tu(t) = \kappa u(t), \quad \kappa \in C \text{ st } Re(\kappa) \le 0, \ \gamma \in (0, 1), \ t \in [0, h], \ h \in R+, \tag{27}$$

cuya solución exacta es u(t) = u0Eγ,1(κt γ), donde Eγ,1(·) es la función de Mittag-Leffler de dos parámetros . Si consideramos una partición del dominio [0, h] en q = h subdominios iguales, la discretización del modelo Ecuación (27) a través del método numérico propuesto en un subdominio arbitrario [tl , tl+1] ≡ [tl , tl + hl] se obtiene como

$$\frac{1}{\gamma}\,(0,hlD\,h\,\gamma)\,t\,T - \kappa E_l\, Uhl\,,l = -\sum_{e=1} (te-1\,,teD\gamma e\,)\,TUhl\,,e\,. \tag{28}$$

Aquí, (te−1 , teDe) γ t es la matriz de diferenciación de memoria como en la Ecuación (22), (0,hlD la matriz de γ)T es diferenciación fraccionaria en un solo dominio [0, hl] y (tl ,hlD γ·l)T ≡ 1 h γ (0,hlD γ) T.

Para estudiar la estabilidad del esquema numérico pseudoespectral propuesto, definimos la matriz de estabilidad

$$S(w) = [P - wQ] - 1R, \tag{29}$$

donde w = κh γ y P, Q y R son las matrices definidas, respectivamente, como l

Machine Translated by Google

$$P = \begin{bmatrix} (0,hlD^{\frac{\gamma}{2}})^t & 0 & \dots & 0 \\ 0 & (0,hlD^{\frac{\gamma}{3}})^t & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (0,hlD^{\frac{\gamma}{q}})^t \end{bmatrix}, \quad J = \begin{bmatrix} E\,0 & \dots & 0\,2T \\ 0\,E\,3 & & \dots\,0 \\ \vdots & \ddots & \vdots \\ 0\,0 & \dots & mi^t_q \end{bmatrix},$$

$$R = \begin{bmatrix} -h^{\gamma}_{(t0,t1D}\,l^{\gamma\ t} & 0 & \dots & 0 \\ -h^{\gamma}\,l_{(t0,t1D}\,l^{\frac{1}{\gamma 2})^t} & -h^{\gamma}_{(t1,t2D}\,l^{\frac{\gamma}{2}})^t & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -h^{\gamma}_{(t0,t1D}\,l^{\gamma}_{)\,q-1}{}^t & -h^{\gamma}_{(t1,t2D}\,l^{\gamma}_{)\,q-1}{}^t & \dots & -h^{\gamma}\,l_{(tq-2,tq-1D}\,l^{\gamma}_{)\,q-1}{}^t \end{bmatrix} \quad (30)$$

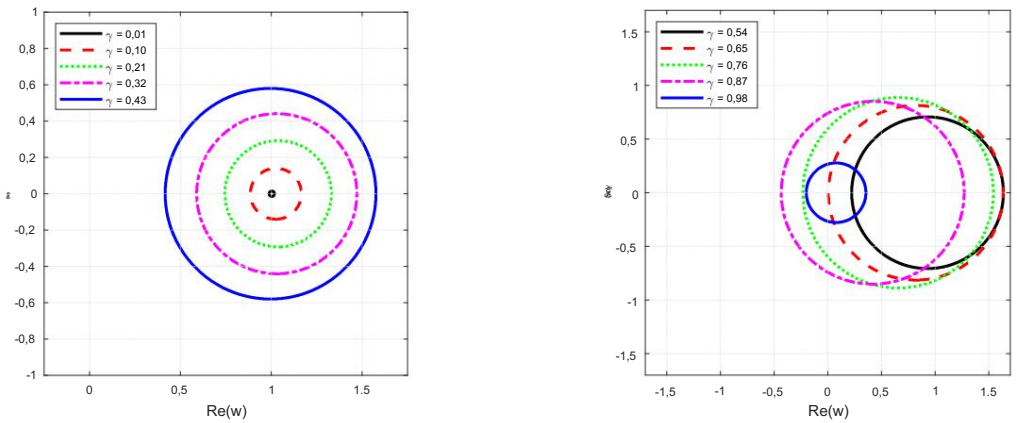
La estabilidad del esquema numérico está determinada por los valores propios de S(w). La región de estabilidad del esquema numérico (28) generalmente se define por el conjunto de puntos en el plano C para los cuales el radio espectral o el módulo del valor propio dominante es menor o igual a 1. Si hay más de un no- valor propio cero con potencial para cruzar el umbral de estabilidad de 1, la región de estabilidad se determina considerando estos valores propios y las regiones donde estos valores propios tienen una magnitud menor o igual a 1. La región de estabilidad general del esquema numérico sería entonces la intersección de estas regiones. Si la región de estabilidad incluye Re(w) ≤ 0 (toda la mitad izquierda del plano C), entonces se dice que el esquema numérico es absolutamente estable (A-estable) [58]. Si, además de la estabilidad A, el límite del valor propio dominante tiende a cero cuando w tiende a −∞, entonces se dice que el esquema numérico es estable de Lubich (L-estable) [59]. La propiedad de estabilidad L significa que cualquier componente que pueda causar inestabilidad en el esquema numérico se amortigua a medida que aumenta la longitud del subdominio, hl , lo cual es particularmente importante cuando se maneja la integración a largo plazo de ecuaciones diferenciales rígidas. Para ilustrar, la Tabla 1 muestra el valor propio dominante, la región de estabilidad y la propiedad de estabilidad para una discretización lineal (N = 1) de la Ecuación (28) con dos subdominios (q = 2) para algunos valores seleccionados de orden fraccionario γ. Como se muestra en la tabla, la región de estabilidad del esquema numérico es el exterior de algunos discos cerrados en C; Los límites de estos discos se ilustran en la Figura 1. Como lo indica la figura, las regiones de estabilidad para γ = 0,01, 0,10, 0,21, 0,32, 0,43, 0,54, 0,65 satisfacen las condiciones de estabilidad absoluta y de Lubich, por lo que el esquema numérico porque estos valores de γ son absolutamente estables y Lubich. Sin embargo, los esquemas numéricos para γ = 0,76, 0,87, 0,98 no son absolutamente estables y, por extensión, no son L-estables.

En la Figura 2, se presentan las regiones de estabilidad para una aproximación cuadrática con q = 3 y q = 5 para γ = 0,21, 0,54, 0,98. La región de estabilidad en la figura es el área fuera de los discos elípticos, lo que muestra que las regiones de estabilidad para q = 3 y q = 5 subdominios son las mismas. Sin embargo, en ambos casos se obtiene un número variable de valores propios con potencial para cruzar el umbral de estabilidad. Para q = 3, hay un valor propio dominante distinto de cero para γ = 0,21 y 0,54, y hay cuatro valores propios distintos de cero para γ = 0,98 con el potencial de cruzar el umbral de estabilidad de 1. Para q = 5, γ = 0,21, 0,54, 0,98 tienen ocho valores propios distintos de cero que pueden cruzar el umbral de estabilidad. Sin embargo, cuando se toma la intersección de las regiones de estabilidad resultantes de estos valores propios, se obtiene la misma región de estabilidad con q = 3 para cada valor de γ ilustrado , como se muestra en la Figura 2. La figura muestra que los esquemas numéricos para γ = 0,21 y γ = 0,54 son absolutamente estables y el esquema para γ = 0,98 no es absolutamente estable (estabilidad condicional). Aunque no se puede llegar a una conclusión general sobre la estabilidad del esquema numérico, se puede inferir de este análisis que si el esquema numérico es absoluta o condicionalmente estable depende del valor del orden arbitrario no entero de la ecuación diferencial fraccionaria.

Machine Translated by Google

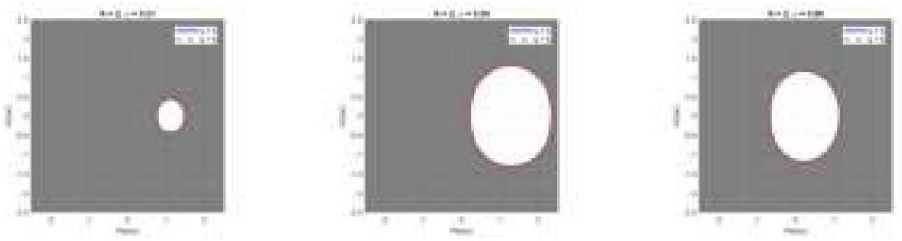
Tabla 1. Valor propio dominante, región de estabilidad y propiedad de estabilidad para algunos valores seleccionados de orden arbitrario, γ, para una aproximación lineal (N = 1) con q = 2.

γ	Valor propio dominante	Región de estabilidad	Propiedad de estabilidad
0,01	$\frac{-124958025860977}{9007199254740992w - 9034579827725136}$	$\frac{9034579827725136\,w - 124958025860977}{9007199254740992} \geq 1$	A–estable L–estable
0,10	$\frac{-1254334826011059}{9007199254740992w - 9235870117935490}$	$w - \frac{9235870117935490}{9007199254740992} \geq \frac{1254334826011059}{9007199254740992}$	A–estable L–estable
0,21	$\frac{-2627472107360489}{9007199254740992w - 9350065552559051}$	$\frac{9350065552559051\,w - 2627472107360489}{9007199254740992} \geq 1$	A–estable L–estable
0,32	$\frac{-3959902260456486}{9007199254740992w - 9279409270142741}$	$w - \frac{9279409270142741}{9007199254740992} \geq \frac{3959902260456486}{9007199254740992}$	A–estable L–estable
0,43	$\frac{-10428492653806604}{18014398509481984w - 17946001659717557}$	$w - \frac{17946001659717557}{18014398509481984} \geq \frac{10428492653806604}{18014398509481984}$	A–estable L–estable
0,54	$\frac{-12700697864360454}{18014398509481984w - 16735571476860095}$	$w - \frac{16735571476860095}{18014398509481984} \geq \frac{12700697864360454}{18014398509481984}$	A–estable L–estable
0,65	$\frac{-7316166143825643}{9007199254740992w - 7382436898791357}$	$\frac{7382436898791357w - 7316166143825643}{9007199254740992} \geq 1$	A–estable L–estable
0,76	$\frac{-15952263845377693}{18014398509481984w - 11810807580156978}$	$w - \frac{11810807580156978}{18014398509481984} \geq \frac{15952263845377693}{18014398509481984}$	
0,87	$\frac{-30618033637964479}{36028797018963968w - 15079171789616800}$	$w - \frac{15079171789616800}{36028797018963968} \geq \frac{30618033637964479}{36028797018963968}$	
0,98	$\frac{-9934148207590571}{36028797018963968w - 2781302084999484}$	$w - \frac{2781302084999484}{36028797018963968} \geq \frac{9934148207590571}{36028797018963968}$	



(a) (b)

Figura 1. Los límites de los discos cerrados en la Tabla 1. La región de estabilidad de la discretización lineal con dos subdominios para los valores indicados de orden arbitrario no entero, (a) γ = 0,01, 0,10, 0,21, 0,32, 0,43 y (b) γ = 0,54, 0,65, 0,76, 0,87, 0,98, se encuentra fuera de estos círculos.



(a) (b) (do)

Figura 2. Las regiones de estabilidad de aproximaciones cuadráticas con q = 3 y q = 5 para (a) γ = 0,21, (b) γ = 0,54 y (c) γ = 0,98. Las regiones de estabilidad son las áreas sombreadas ubicadas fuera de los discos elípticos.

Machine Translated by Google

4. Experimentación numérica y resultados.

Esta sección presenta una serie de ejemplos numéricos, que van desde ecuaciones lineales simples hasta sistemas de ecuaciones no lineales, junto con los resultados. Análisis detallados de la Los resultados de cada ejemplo se proporcionan para ilustrar las observaciones clave. Para una integral comprensión de los métodos numéricos utilizados en estos ejemplos, consulte la descripción previsto en la Sección 3.

Ejemplo 1. Considere la ecuación diferencial fraccionaria lineal [60]

$$ {}_{0}^{do}re_{t}^{\gamma}tu(t) + tu(t) = \frac{2t^{2-\gamma}}{\Gamma(3-\gamma)} - \frac{t^{1-\gamma}}{\Gamma(2-\gamma)} + t^{2} - t,\ 0 < \gamma \le 1,\ t \quad [0, T], \tag{31} $$

con condición inicial u(0) = 0 y solución exacta u(t) = t 2 - t.

La Tabla 2 presenta el error absoluto entre las soluciones aproximadas y exactas usando el método pseudoespectral multidominio no superpuesto con la cuadratura CGL y los polinomios de Chebyshev como funciones base. La tabla también incluye resultados numéricos. obtenido sin descomponer el dominio. Este análisis comparativo destaca la precisión y eficacia del enfoque multidominio para reducir errores y mejorar la precisión de la solución numérica. De la tabla se puede ver que, para el fraccionario orden de γ = 0,5, aunque el error absoluto máximo en la configuración multidominio no no disminuye en términos de ceros significativos, se reduce a medida que q aumenta. Esto indica que, aunque mínimo, existe una ventaja al utilizar más subdominios. Comparando el error al resultado de dominio único, hay un cero significativo adicional en el error del enfoque multidominio, lo que indica que el enfoque multidominio es más efectivo para este problema.

Tabla 2. Error absoluto máximo para el ejemplo 1 en configuraciones de dominio único y multidominio.

	norte = 10, γ = 0,5			N = 10, T = 6,0		
q	T = 1,0	T = 3,0	T = 6,0	γ = 0,2	γ = 0,6	γ = 0,8
	7,7426 × 10-5	6,2253 × 10-4	2,7903 × 10-3	1,3625 × 10−3	3,6161 × 10−3	5,2468 × 10−3
5 10	4,0800 × 10-5	5,2107 × 10-4	2,2192 × 10-3	6,7562 × 10−4	1,9440 × 10−3	2,8829 × 10−3
20	2,3538 × 10-5	4,4940 × 10-4	1,8850 × 10-3	4,4076 × 10−4	1,5245 × 10−3	2,1135 × 10−3
50	2,0489 × 10-5	4,1298 × 10-4	1,6596 × 10-3	3,8984 × 10−4	1,2966 × 10−3	1,5794 × 10−3
100	2,0286 × 10-5	4,0279 × 10-4	1,6132 × 10-3	3,8131 × 10−4	1,2746 × 10−3	1,4322 × 10−3
200	2,0374 × 10-5	4,0190 × 10-4	1,6047 × 10-3	3,8224 × 10−4	1,2946 × 10−3	1,3898 × 10−3
			Dominio único			
	3,3518 × 10−4	1,0362 × 10−3	2,1415 × 10-3 9,0794 × 10-4		2,5580 × 10−3	2,9839 × 10−3

De la Tabla 2, la efectividad de la técnica numérica propuesta para obtener una Una solución aproximada precisa para un FDE lineal es evidente, por lo que en los ejemplos siguientes, considere FDE no lineales para validar aún más el método numérico.

Ejemplo 2. Considere la ecuación diferencial fraccionaria no lineal [60]

$$ {}_{0}^{do}re_{t}^{\gamma}tu(t) - tu^{2}(t) = -\frac{2}{(t+1)^{2}},\quad 0 < \gamma \le 1,\ t \quad [0, T], \tag{32} $$

donde la condición inicial es u(0) = −2 y la solución exacta cuando γ = 1 viene dada por u(t) = $-\frac{2}{(t+1)}$.

Machine Translated by Google

La Tabla 3 muestra la solución aproximada obtenida usando dos subdominios diferentes. en el método propuesto. Las soluciones en estos subdominios coinciden con los valores numéricos. informado en Odibat y Momani [60]. Cabe señalar que las soluciones de forma cerrada para γ = 0,2 y γ = 0,4 no están disponibles. Sin embargo, para γ = 1, la precisión de la propuesta El método es incuestionable. Para dar aún más credibilidad a la precisión del método propuesto, La Tabla 4 demuestra que el error residual es consistentemente cercano a cero para varios valores. de γ. Esto sugiere que el método propuesto resuelve efectivamente el problema del Ejemplo 2.

Tabla 3. Solución aproximada del Ejemplo 2 usando dos subdominios diferentes, q, y el numérico valores informados en Odibat y Momani [60].

t	γ = 0,2			γ = 0,4			γ = 1,0		
	q = 10	q = 20	Reportado [60]	q = 10	q = 20	Reportado [60]	q = 10	q = 20	Exacto
0,0	−2,0000	−2,0000	−2.0000	−2,0000	−2,0000	−2.0000	−2,0000	−2,0000	−2,0000
0,2	−1,4454	−1,4455	−1,4864	−1,4874	−1,4874	−1,4899	−1,6666	−1,6666	−1,6666
0,4	−1,3295	−1,3296	−1,3665	−1,3506	−1,3506	−1,3525	−1,4285	−1,4285	−1,4285
0,6	−1,2477	−1,2478	−1,2822	−1,2525	−1,2526	−1,2541	−1,2500	−1,2500	−1,2500
0,8	−1,1867	−1,1868	−1,2192	−1,1774	−1,1774	−1,1787	−1,1111	−1,1111	−1,1111
1,0	−1,1396	−1,1397	−1,1704	−1,1177	−1,1178	−1,1189	−0,9999	−1,0000	−1,0000

Tabla 4. Los errores residuales del Ejemplo 2 en puntos seleccionados usando dos subdominios diferentes, q, para diferentes valores de orden fraccionario γ.

t	γ = 0,2		γ = 0,4		γ = 0,6	
	q = 10	q = 20	q = 10	q = 20	q = 10	q = 20
0,2	4,4230 × 10−9	4,8561 × 10−13	9,7841 × 10−11	9,2539 × 10−12	3,2469 × 10−10	5,2865 × 10−11
0,4	9,3414 × 10−13	2,1760 × 10−13	1,6637 × 10−11	3,8897 × 10−12	9,3717 × 10−11	2,0417 × 10−11
0,6	4,9305 × 10−13	1,2423 × 10−13	8,9744 × 10−12	2,1446 × 10−12	4,9785 × 10−11	1,1396 × 10−11
0,8	2,8477 × 10−13	7,1831 × 10−14	5,4355 × 10−12	1,3228 × 10−12	3,0330 × 10−11	7,1187 × 10−12
1.0	1,7807 × 10−13	4,6629 × 10−14	3,5078 × 10−12	8,6397 × 10−13	1,9900 × 10−11	4,6949 × 10−12

Ejemplo 3. A continuación, considere la ecuación diferencial fraccionaria no lineal [61]

$$ {}_0^{do}re_t^{\gamma} tu(t) + tu(t) \quad ^2 = \frac{2t\ 2-\gamma}{\Gamma(3-\gamma)} - \frac{t\ 1-\gamma}{\Gamma(2-\gamma)} + (t \quad ^2 - t + 1) \quad ^2, \quad 0 < \gamma \le 1, \qquad t \quad [0, T], \ (33) $$

con la condición inicial u(0) = 1 y cuya solución exacta es u(t) = t $^2 - t + 1$.

La Tabla 5 presenta el error absoluto máximo de la diferencia entre el exacto y el Soluciones aproximadas obtenidas utilizando los métodos multidominio y monodominio. Él Se observa en la tabla que existe conformidad con la tendencia observada en la Tabla 2. Como el dominio del FDE se hace más grande, el error en el esquema de dominio único se deteriora; sin embargo, en comparación, la magnitud de los errores es mejor para el esquema multidominio. Observamos que la magnitud del error aumenta a medida que aumenta γ , similar a la tendencia observado en la Tabla 2.

Machine Translated by Google

Tabla 5. El error absoluto máximo de la diferencia entre las soluciones exacta y aproximada del Ejemplo 3 para diferentes números de subdominios, q.

	norte = 10, γ = 0,5			N = 10, T = 6,0		
q	T = 1,0	T = 3,0	T = 6,0	γ = 0,2	γ = 0,6	γ = 0,8
	6,5833 × 10-5	4,5300 × 10-4	7,7010 × 10-4	1,6897 × 10−4	9,9126 × 10−4	1,0256 × 10−3
5 10	3,6216 × 10-5	2,6020 × 10-4	4,7125 × 10-4	1,0248 × 10−4	6,3577 × 10−4	7,4647 × 10−4
20	1,9347 × 10-5	1,3492 × 10-4	2,8073 × 10-4	5,5876 × 10−5	3,9608 × 10−4	5,1975 × 10−4
50	1,6865 × 10-5	8,2634 × 10-5	1,3621 × 10-4	2,3143 × 10−5	2,0636 × 10−4	3,2248 × 10−4
100	1,6547 × 10-5	6,7163 × 10-5	8,2558 × 10-5	1,0361 × 10−5	1,1932 × 10−4	2,1211 × 10−4
200	1,6576 × 10-5	5,9548 × 10-5	6,7326 × 10-5	7,5375 × 10−6	7,6523 × 10−5	1,3480 × 10−4
			Dominio único			
	3,1847 × 10−4	9,0415 × 10−4	1,7639 × 10−3 6,1773 × 10−4		1,9684 × 10−3	2,2669 × 10−3

Ejemplo 4. Considere el valor inicial de la ecuación diferencial fraccionaria homogénea de Bagley-Torvik

$$^{do}_{0}re_{t}{}^{\gamma}u(t) + u\text{“}(t) + u(t) = 0, \text{ para } t \quad (0, T] \text{ y } \gamma \quad (1, 2), \tag{34}$$

sujeto a las condiciones iniciales u(0) = 1,0 y u′ (0) = 0.

Si reemplazamos el lado derecho de la ecuación (34) con $2 + 4\ \overline{\frac{t}{\pi}}\ 2 + t$ y establezca t (0, 1], sujeto a las condiciones iniciales u(0) = 0,0 y u′ (0) = 0, la solución de forma cerrada es dado como $u(t) = t^{2}$ cuando γ = 1,5. La Tabla 6 presenta los errores obtenidos al tomar el valor absoluto. valor de la diferencia entre las soluciones numéricas obtenidas a través del multidominio técnica y las soluciones exactas, y estas se comparan con los errores obtenidos mediante el método evolutivo de redes neuronales artificiales en el estudio de Raja et al. [62].

Tabla 6. Los errores absolutos de la diferencia entre las soluciones exacta y numérica del Ejemplo 4 a diferentes valores de t usando N = 12 y q = 400 en comparación con los errores reportados en el trabajo de Raja et al. [62].

t	Solución exacta	Resultado actual	Error presente	Error informado [62]
0,1	0,01	0,01000	1,83183 × 10-7	5,48 × 10−5
0,3	0,09	0,09005	1,26966 × 10-6	2,66 × 10−3
0,5	0,25	0,25014	6,69388 × 10-6	1,67 × 10−2
0,7	0,49	0,49026	3,40281 × 10-5	5,80 × 10−2
0,9	0,81	0,81040	1,25968 × 10-4	1,50 × 10−1
1,0	1,00	1,00048	2,12565 × 10-4	2,25 × 10−1

La Tabla 6 compara las soluciones exactas y aproximadas obtenidas utilizando el método multidominio en varios valores de t y los errores asociados en cada t. Los errores son en comparación con los obtenidos en Raja et al. [62]. Se observa en la tabla que el Las soluciones aproximadas coinciden mejor con las soluciones exactas que los resultados de Raja et al. [62], y esto valida la eficiencia del método propuesto ya que supera a una red neuronal enfoque conocido por ser eficiente y confiable.

Para validar aún más la precisión del método multidominio propuesto, el infinito Se presenta la norma del residuo para diferentes números de subdominios. La tabla 7 da la norma infinita de los errores residuales para el ejemplo 4. El error residual se obtiene mediante el equivalente discreto de la ecuación (34), definida por

$$\text{Residual} = \sum_{e=1}^{\text{“}} (te{-}1\ ,teD_{e}^{\gamma})\ TUhl\ ,e \quad + (tl\ ,hlD \quad {}^{\gamma}_{\text{“}})\ TUhl\ ,l + (tl\ ,hlD2\ l \quad)\ TUhl\ ,l + E \quad {}^{t}_{lUhl\ ,\ l}\ . \tag{35}$$

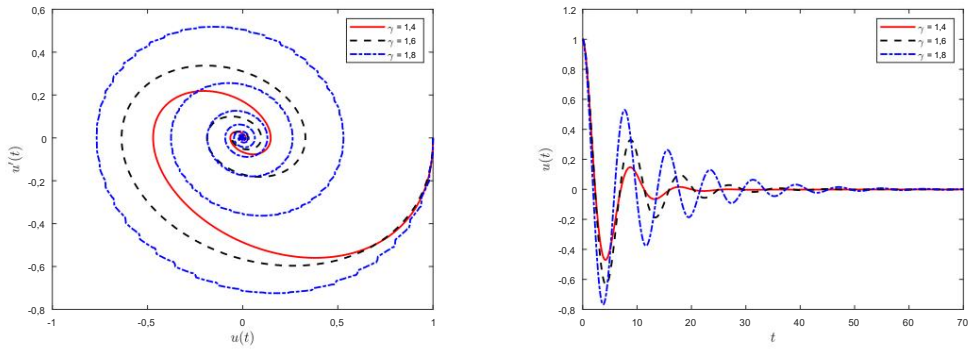
Machine Translated by Google

La Tabla 7 muestra que el método es eficiente para resolver diferenciales de orden fraccionario.
ecuaciones de valores variables de orden arbitrario no entero, ya que el error residual permanece
muy pequeño para varios órdenes fraccionarios, lo que demuestra la eficacia de la técnica multidominio.

Tabla 7. Los errores residuales máximos para el Ejemplo 4 para diferentes valores de T , órdenes fraccionarios y diferente número de multidominios, q.

	norte = 10, γ = 1,5			N = 10, T = 6,0		
q	T = 1,0	T = 3,0	T = 6,0	γ = 1,2	γ = 1,6	γ = 1,8
10	1,115 × 10−10	1,158 × 10−11	1,933 × 10−12	1,031 × 10−12	1,870 × 10−12	3,445 × 10−12
20	7,346 × 10−10	6,618 × 10−11	1,158 × 10−11	1,078 × 10−11	2,104 × 10−11	2,765 × 10−11

Para validar aún más los resultados obtenidos, reproducimos los perfiles de solución obtenidos.
por Zafar et al. [63] y observan una excelente concordancia en el resultado, como se muestra en la Figura 3.



(a) (b)

Figura 3. La dinámica de la ecuación de Bagley-Torvik (34) para varios valores de γ con T = 70,
N = 10 y q = 400: (a) el retrato de fase para u − u ′ plano y (b) el perfil u(t) .

En los ejemplos restantes de esta sección, consideramos sistemas de fracciones acopladas.
Problemas de valor inicial que describen sistemas caóticos. Estas ecuaciones a menudo exhiben una dinámica
de largo plazo e, invariablemente, requieren una integración a largo plazo para comprender completamente la
dinámica subyacente involucrada. Por lo tanto, una técnica numérica, como la multidominio
El método pseudoespectral de Chebyshev, que puede manejar la integración a largo plazo, encaja perfectamente
para resolver tales sistemas, como lo corroboran los resultados de los siguientes cuatro ejemplos.

Ejemplo 5. Considere el sistema caótico fraccionario de ecuaciones diferenciales que describen una
sistema eléctrico inherente a la memoria [64]

$$
\begin{aligned}
{}^{do}_{0}re_{t}^{\gamma}u1(t) &= \frac{u2(t)}{do}, \\
{}^{do}_{0}re_{t}^{\gamma}u2(t) &= -\frac{1}{l}\,u1(t) + \beta u2(t)(u3(t))2 - \beta u2(t), \\
{}^{do}_{0}re_{t}^{\gamma}u3(t) &= -u2(t) - \alpha u3(t) + u2(t)u3(t), \qquad t \in (0, T], \gamma \in (0, 1],
\end{aligned}
\tag{36}
$$

sujeto al siguiente estado inicial, (0.1, 0.0, 0.1). Aquí, c = 0,70, L = 3,00, β = 1,10 y
α = 0,60. El voltaje a través del capacitor, c, está representado por la variable u1(t), y la corriente
a través del inductor, L, está representado por la variable u2(t), y el estado interno de la memristiva
El sistema se denota por la variable u3(t).

El error residual máximo promedio para el sistema (36) se da en la Tabla 8. Observamos
aquí que los errores residuales máximos presentados en la Tabla 8 y tablas posteriores en este

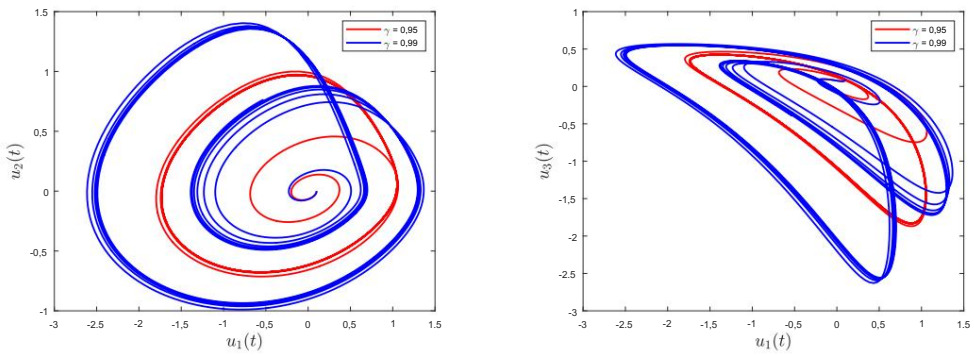
Machine Translated by Google

sección se obtienen como el promedio de los residuos máximos de las ecuaciones del sistema. Los residuos se calculan de manera similar a la Ecuación (35). Se puede observar en la Tabla 8 que para varios órdenes fraccionarios y dominios aparentemente grandes, [0, T], se requiere un alto nivel de precisión. Se obtiene cuando se utiliza el método pseudoespectral de Chebyshev multidominio. debido a la naturaleza inherentemente caótica del problema, la dinámica del sistema generalmente requiere una gran dominio físico para desarrollarse plenamente y capturar eficazmente estos complejos procesos de largo plazo. comportamientos, el dominio físico se divide en una gran cantidad de subdominios más pequeños. La tabla muestra que el promedio de los errores residuales máximos disminuye marginalmente a medida que aumenta el número de subdominios. Esta relación inversa entre el número de subdominios y el error residual resalta la aplicabilidad del multidominio técnica cuando se trata de este tipo de sistema que requiere una integración a largo plazo. Este Es particularmente evidente cuando el error se compara con el error asociado con el método de dominio único. Como ya se observó en los ejemplos anteriores y ahora se corrobora Según la Tabla 8, el error asociado con el método de dominio único es peor que el error se obtiene cuando se utiliza el enfoque multidominio.

Tabla 8. Errores residuales del ejemplo 5 para dos valores de T usando N = 10 y para números diferentes de los subdominios q, con c = 0,70, L = 3,00, β = 1,10 y α = 0,60.

	T = 20,0			T = 50,0		
q	γ = 0,7	γ = 0,8	γ = 0,9	γ = 0,7	γ = 0,8	γ = 0,9
100	2,515 × 10−11	7,810 × 10−12	8,503 × 10−12	8,781 × 10−11	3,336 × 10−11	8.054 × 10−10
200	9,780 × 10−12	2,590 × 10−12	2,144 × 10−12	3,434 × 10−11	1,113 × 10−11	1,935 × 10−10
300	5,544 × 10−12	1,365 × 10−12	9,795 × 10−13	1,973 × 10−11	5,817 × 10−12	8,468 × 10−11
400	3,717 × 10−12	8,632 × 10−13	5,718 × 10−13	1,319 × 10−11	3,701 × 10−12	4,723 × 10−11
500	2,720 × 10−12	6,125 × 10−13	4,095 × 10−13	9,769 × 10−12	2,592 × 10−12	3.004 × 10−11
			Dominio único			
	2,397 × 10−9	2,439 × 10−9	1,920 × 10−7 3,101 × 10−8		4,108 × 10−7	5,427 × 10−7

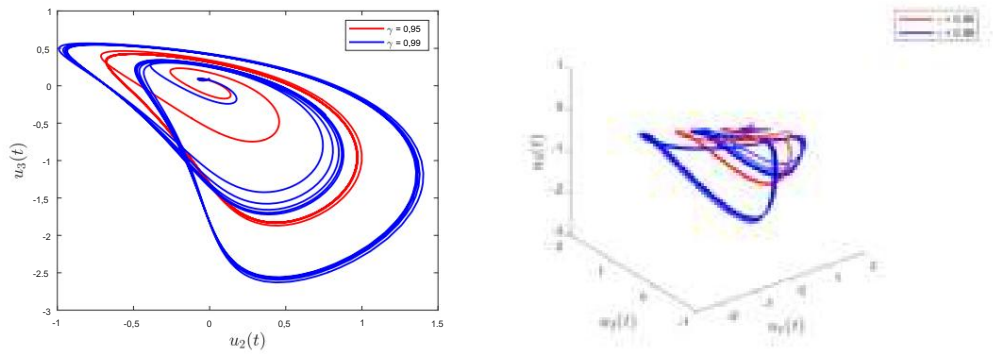
Para dar mayor credibilidad a los resultados numéricos de este ejemplo, reproducimos el dinámica, en forma de retratos de fase, del sistema eléctrico caótico memristivo como informado en el trabajo de Sene [64]. Como se muestra en la Figura 4, existe un excelente acuerdo entre estos retratos de fase y los de la obra de Sene [64].



(a) (b)

Figura 4. Continuación.

Machine Translated by Google



(cd)

Figura 4. Los retratos de fase de los sistemas caóticos (36) para el (a) plano u1 − u2 , (b) plano u1 − u3 , (c) plano u2 − u3 y (d) superficie u1 − u2 − u3 con γ = 0,95 y 0,99, T = 250, N = 13 y q = 850.

Ejemplo 6. Considere las ecuaciones diferenciales de estiramiento, torsión y pliegue modificadas fraccionarias [65]

$$\begin{aligned} {}_{0}^{c}D_{t}^{\gamma}\, u1(t) &= \Omega 1 u3(t) - 8u1(t)u2(t), \quad t \in (0, T], \gamma \in (0, 1], \\ {}_{0}^{c}D_{t}^{\gamma}\, u2(t) &= 11(u1(t))^2 + 3(u2(t))^2 + (u3(t))^2 + \Omega 2 u1(t)u3(t) - 3\Omega 2^{3}, \\ {}_{0}^{c}D_{t}^{\gamma}\, u3(t) &= 2\,1 - \Omega 4\, u2(t)u3(t) - \Omega 1 u1(t) - \Omega 2 u1(t)u2(t), \end{aligned} \tag{37}$$

bajo el estado inicial (0.2, −0.2, 0.2), con γ = 1.00, Ω1 = 0.2, Ω2 = 2.0, Ω3 = 1.0 y Ω4 = 0.1.

De manera similar al ejemplo anterior, este sistema generalmente requiere un dominio físico grande para estabilizarse y desarrollarse completamente, y cuando se resuelve en un solo dominio, la solución numérica es inestable, como lo demuestran los errores residuales promedio en un solo dominio presentados en la Tabla 9. Presentamos los errores residuales máximos promedio para varios órdenes fraccionarios y dominios relativamente grandes, [0, T], para diferentes números relativamente grandes de subdominios en la Tabla 9. Existe una concordancia entre el nivel de precisión observado para los resultados en la Tabla 9 y los obtenidos en los ejemplos anteriores. Se muestra que el error obtenido para los diversos números de subdominio utilizados es pequeño, lo que indica que el método pseudoespectral multidominio de Chebyshev propuesto resuelve eficientemente este sistema de FDE. A medida que q aumenta, se observa que el error residual promedio mejora, lo que resulta en una mayor precisión. La eficacia superior del enfoque multidominio sobre el esquema numérico de dominio único se vuelve marginalmente más clara en este ejemplo. En este caso, se observa que el esquema numérico de dominio único produce errores numéricos bastante grandes, mientras que la técnica de dominio múltiple puede lograr un nivel de precisión más alto. Esta disparidad en el nivel de desempeño resalta la limitación en el alcance del esquema de dominio único, mientras que el enfoque de dominio múltiple demuestra robustez y versatilidad, especialmente para sistemas complejos como el considerado en este ejemplo.

La Figura 5 muestra el comportamiento de los perfiles de solución a medida que t evoluciona. Los resultados son los esperados y están completamente de acuerdo con el resultado de Azam et al. [65]. También se observa en la Figura 6 que el sistema caótico fluye a través de los distintos estados girando y girando alrededor de los cuadrantes. Esta dinámica también fue capturada por Azam et al. [65] y valida los resultados numéricos que se presentan en este estudio.

Machine Translated by Google

Tabla 9. Los errores residuales máximos promedio para el Ejemplo 6 para dos valores de T y diferentes números de subdominios q usando N = 10. Se utilizan los siguientes valores de parámetros: Ω1 = 0,2, Ω2 = 2,0, Ω3 = 1,0 y Ω4 = 0,1.

	T = 20,0			T = 50,0		
q	γ = 0,7	γ = 0,8	γ = 0,9	γ = 0,7	γ = 0,8	γ = 0,9
100	1,292 × 10−8	4,793 × 10−8	2,171 × 10−8	1,364 × 10−10	1,059 × 10−8	4,952 × 10−8
200	1,883 × 10−8	1,192 × 10−8	6,393 × 10−9	1,060 × 10−8	2,928 × 10−8	4,607 × 10−8
300	8,108 × 10−9	5,539 × 10−9	3,267 × 10−9	1,826 × 10−8	3,502 × 10−8	1,144 × 10−8
400	4,546 × 10−9	3,223 × 10−9	1,790 × 10−9	2,750 × 10−8	1,284 × 10−8	1.065 × 10-8
500	3,082 × 10−9	2,098 × 10−9	1,182 × 10−9	1,883 × 10−8	1,194 × 10−8	6,393 × 10−9
			Dominio único			
	1.730 × 101	1.999 × 101	2.419 × 101	8.388 × 101	2,195 × 10−1	4.131 × 101

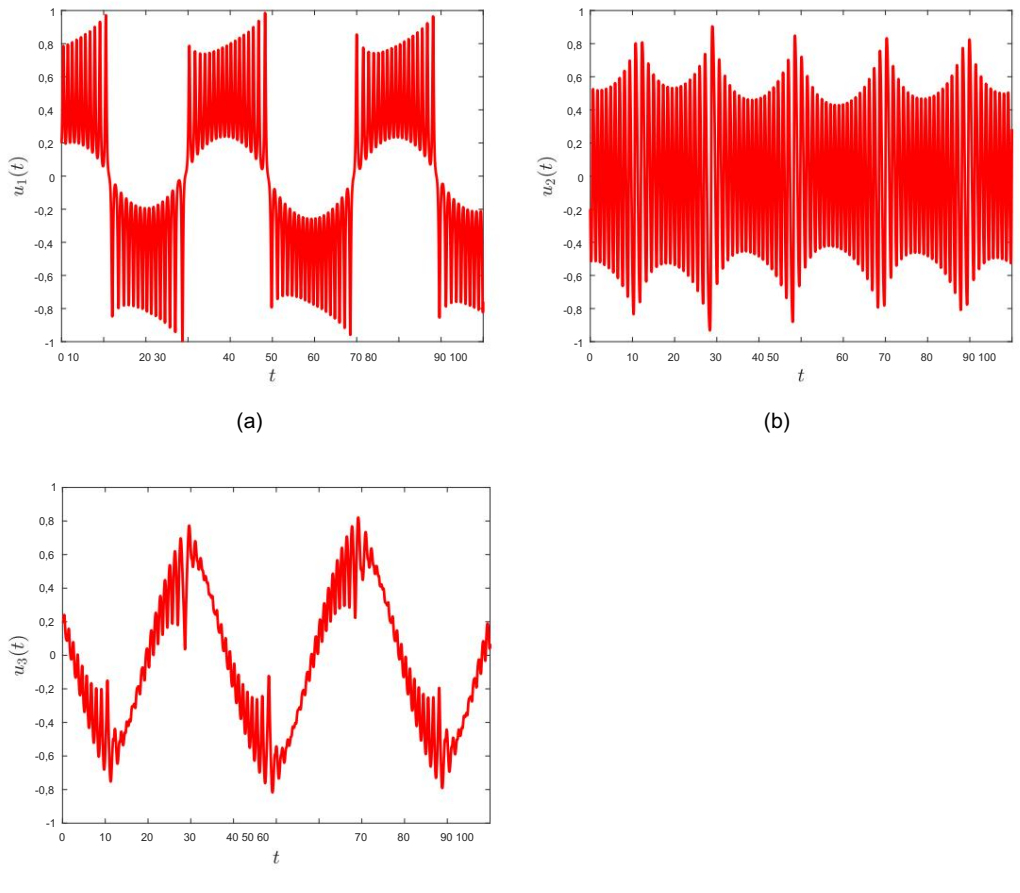


Figura 5. La evolución temporal de u1(t), u2(t), u3(t) del sistema caótico (37) con γ = 1,00, Ω1 = 0,2, T = 100, N = 13 y q = 350.

Machine Translated by Google

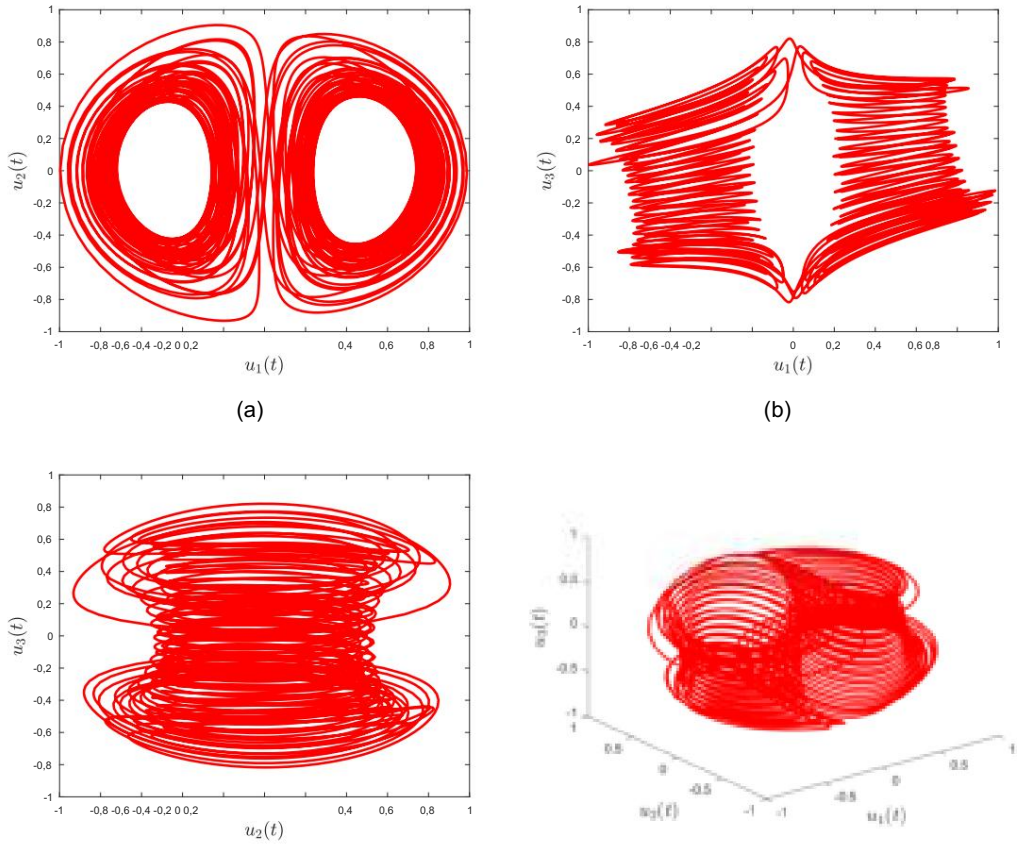


Figura 6. Los retratos de fase del sistema caótico (37) para el (a) plano u1 − u2 , (b) plano u1 − u3 , (c) plano u2 − u3 y (d) superficie u1 − u2 − u3 con γ = 1,00, Ω1 = 0,2, T = 100, N = 13 y q = 350.

Ejemplo 7. Considere las ecuaciones diferenciales fraccionarias de Rossler [66]

$$\begin{aligned} {}_{0}^{do}re_{t}^{\gamma}u1(t) &= -u2(t) - u3(t),\ t \quad (0, T],\ \gamma \quad (0, 1], \\ {}_{0}^{do}re_{t}^{\gamma}u2(t) &= u1(t) + \Psi 1u2(t), \\ {}_{0}^{do}re_{t}^{\gamma}u3(t) &= 0{,}2 + u1(t)u3(t) - 10u3(t), \end{aligned} \tag{38}$$

con las siguientes condiciones iniciales: (1, 1, 1), con Ψ1 = 0,4 y γ = 0,9.

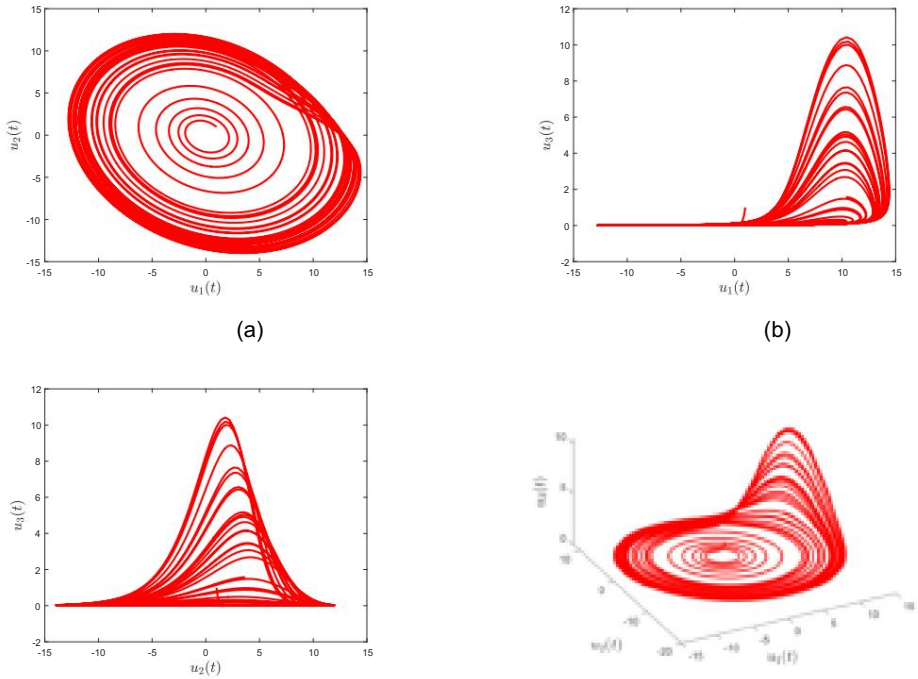
La Tabla 10 muestra los errores residuales máximos promedio para la Ecuación (38) para valores de orden arbitrario no entero, γ, y para dos dominios, (0, T]. Los resultados son presentado para diferentes números de descomposición de dominio. En la Tabla 10 vemos que el Los errores residuales son pequeños, lo que indica que las soluciones aproximadas son precisas. como lo ha hecho validado varias veces en este estudio, un aumento en el número de subdominios utilizados da como resultado soluciones más precisas. También se muestra que utilizando el enfoque de dominio único da como resultado una menor precisión en las soluciones aproximadas obtenidas.

Machine Translated by Google

Tabla 10. Errores residuales para el ejemplo 7 en dos tiempos finales T usando el número de puntos de la cuadrícula N = 10 para diferentes números de subdominios q y Ψ1 = 0,40.

	T = 20,0			T = 50,0		
q	γ = 0,7	γ = 0,8	γ = 0,9	γ = 0,7	γ = 0,8	γ = 0,9
100	1,073 × 10−7	4,973 × 10−8	1,171 × 10−8	1,469 × 10-7	1,469 × 10−7	1,893 × 10−7
200	5,544 × 10−8	1,857 × 10−8	3,569 × 10−9	1,269 × 10-7	6,666 × 10−8	4,696 × 10−8
300	3,479 × 10−8	1,015 × 10−8	1,740 × 10−9	9,241 × 10-8	3,853 × 10−8	3.299 × 10-8
400	2,459 × 10−8	6,520 × 10−9	1,042 × 10−9	6,979 × 10-8	2,574 × 10−8	1,933 × 10−8
500	1,859 × 10−8	4,653 × 10−9	7,023 × 10−10	5,544 × 10-8	1,857 × 10−8	1,189 × 10−8
			Dominio único			
	1,594 × 10−4	2,218 × 10−4	6,733 × 10−5 1,725 × 10−4		2,518 × 10−4	3,726 × 10−4

La Figura 7 muestra los retratos de fase del sistema caótico, y el resultado concuerda con los resultados obtenidos en Li et al. [66], validando así las presentes soluciones numéricas.



(a) (b)

(cd)

Figura 7. Los retratos de fase de los sistemas caóticos (38) para el (a) plano u1 -u2 , (b) el plano u1 - u3 , (c) plano u2 − u3 y (d) superficie u1 − u2 − u3 con γ = 0,90, Ψ1 = 0,4, T = 200, N = 10 y q = 900.

Ejemplo 8. Considere el siguiente sistema caótico fraccionario no lineal [67]

$$\begin{aligned} {}_{0}^{do}re_{t}^{\gamma}u1(t) &= \Psi 1u1(t) + u2(t) + u2(t)u3(t),\ t \quad (0, T],\ \gamma \quad (0, 1], \\ {}_{0}^{do}re_{t}^{\gamma}u2(t) &= -u1(t)u3(t) + u2(t)u3(t), \\ {}_{0}^{do}re_{t}^{\gamma}u3(t) &= -\Psi 2u3(t) - \Psi 3u1(t)u2(t) - m1, \end{aligned} \tag{39}$$

con el estado inicial dado como (0.2, 0.2, 0.2), Ψ1 = Ψ2 = Ψ3 = 1.0, m1 = 0.0 y γ = 0.9.

En la Tabla 11, los errores residuales máximos promedio obtenidos para la Ecuación (39) son presentado. La tabla muestra que un aumento en el número de dominios particionados resulta en una mayor precisión de las soluciones aproximadas obtenidas para los órdenes fraccionarios seleccionados y dominios. Se observa que los errores residuales máximos promedio son pequeños, lo que confirma que

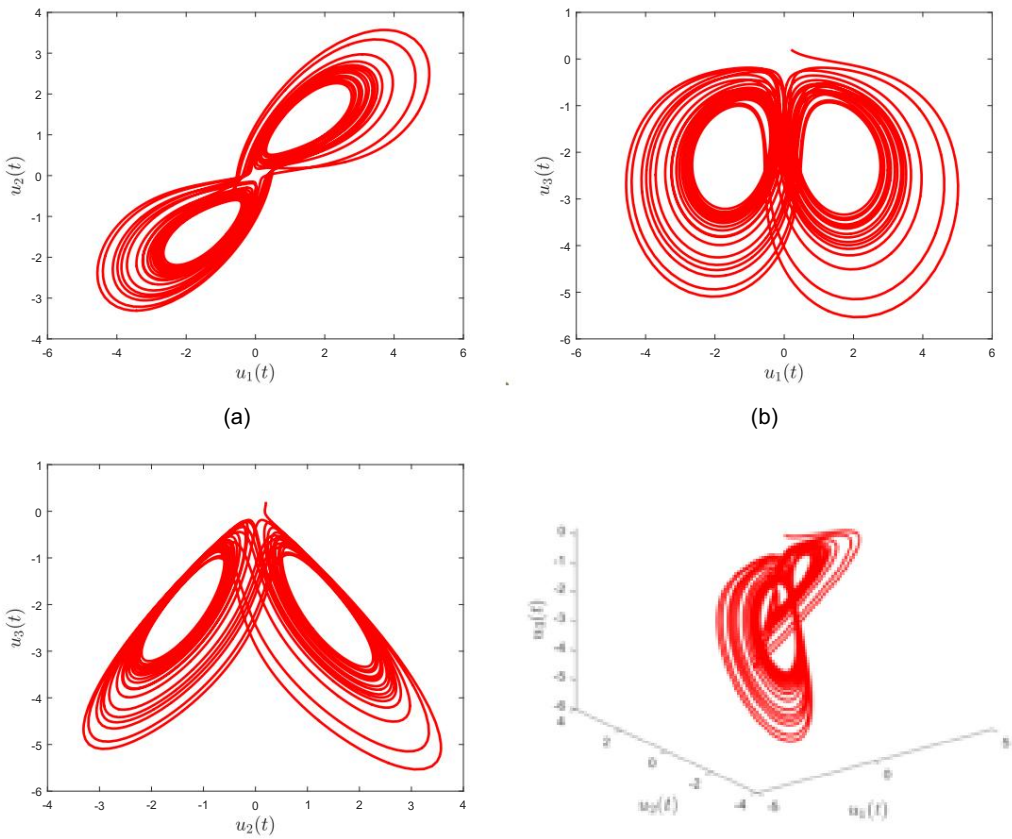
Machine Translated by Google

el método multidominio propuesto es altamente eficiente para este modelo diferencial fraccionario, a diferencia del enfoque de dominio único, que no es tan preciso.

Tabla 11. Los errores residuales máximos promedio para el Ejemplo 8 en dos dominios diferentes (0, T] con N = 10 para diferentes números de subdominios q, Ψ1 = Ψ2 = Ψ3 = 1,0 y m1 = 0,0.

	T = 20,0			T = 50,0		
q	γ = 0,7	γ = 0,8	γ = 0,9	γ = 0,7	γ = 0,8	γ = 0,9
100	4,114 × 10−8	6,012 × 10−8	5,802 × 10−8	2,142 × 10−7	1,226 × 10−7	2,711 × 10−7
200	1,305 × 10−8	1,673 × 10−8	1,559 × 10−8	5,335 × 10−8	6,845 × 10−8	7.236 × 10-8
300	6,155 × 10−9	7,361 × 10−9	7,034 × 10−9	3,422 × 10−8	4,357 × 10−8	4.049 × 10-8
400	3,398 × 10−9	4,196 × 10−9	3,935 × 10−9	2,050 × 10−8	2,575 × 10−8	2,436 × 10−8
500	2,195 × 10−9	2,720 × 10−9	2,540 × 10−9	1,304 × 10−8	1,676 × 10−8	1,559 × 10−8
			Dominio único			
	3.055 × 10-2	1,225 × 10−2	9,667 × 10−1 2,279 × 10−2		6.453 × 101	4.708 × 101

Presentamos los retratos de fase del sistema caótico en la Figura 8. La figura valida la dinámica informada por Sene [67] y está en excelente acuerdo con la resultado publicado. La concordancia entre los retratos de fase presentados aquí y aquellos documentado en el estudio de Sene [67] sirve para corroborar aún más la precisión y fiabilidad del método numérico actual.



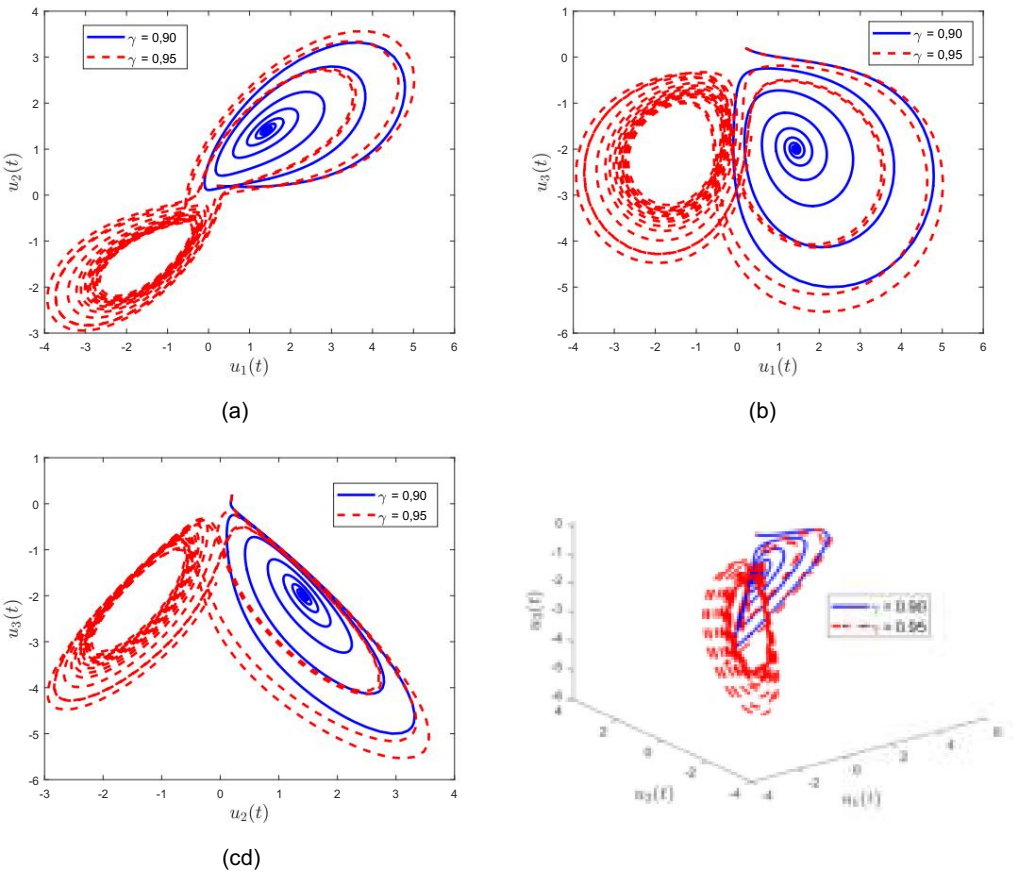
(a) (b)

(cd)

Figura 8. Los retratos de fase de los sistemas caóticos (39) para el (a) plano u1 − u2 , (b) plano u1 − u3 , (c) plano u2 − u3 y (d) superficie u1 − u2 − u3 con γ = 0,95, T = 150, N = 10 y q = 800.

Machine Translated by Google

La Figura 9 muestra los retratos de fase del sistema caótico en la Ecuación (39) usando dos diferentes valores de γ . La figura justifica que la dinámica del sistema caótico depende de El orden real arbitrario.



(a) (b)

(cd)

Figura 9. Los retratos de fase de los sistemas caóticos (39) para el (a) plano u1 − u2 , (b) plano u1 − u3 , (c) plano u2 − u3 y (d) superficie u1 − u2 − u3 con γ = 0,90 y 0,95, T = 80, N = 10 y q = 350.

La Figura 10 muestra la precisión de los resultados numéricos (el máximo residual absoluto en cada subdominio) de cada ecuación en los sistemas caóticos presentados en Ecuaciones (36)–(39). Las cifras muestran que la precisión de las soluciones obtenidas mediante el método pseudoespectral multidominio de Chebyshev es alto y en ocasiones alcanza errores menores que los reportados hasta ahora en las tablas. Como se puede observar claramente, como el A medida que aumenta el número de subdominios en el dominio computacional, el error residual disminuye correspondientemente. En algunos casos, el error residual se reduce al orden de 10-16. Sin embargo, los errores residuales con un pequeño número de subdominios ya muestran buena precisión, lo que demuestra que el uso del enfoque multidominio propuesto en este El estudio es muy eficiente para resolver ecuaciones diferenciales fraccionarias que normalmente requieren integración a largo plazo.

Machine Translated by Google

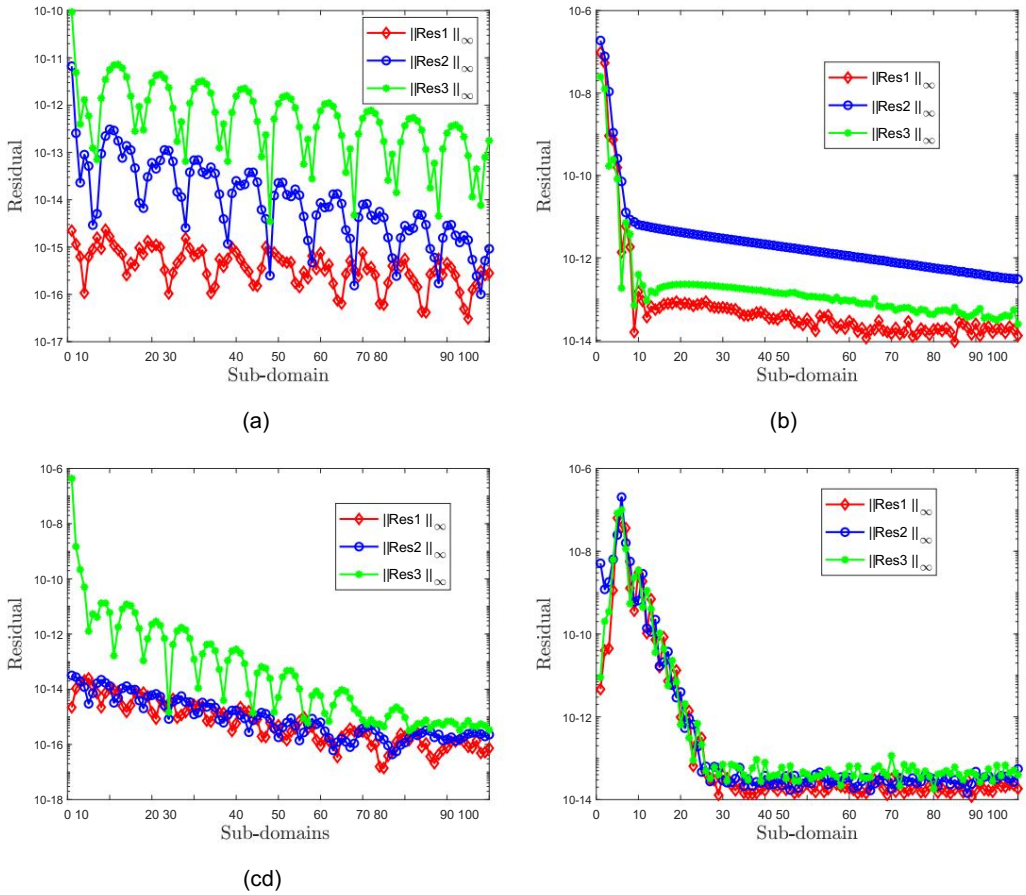


Figura 10. El error residual máximo en cada dominio para (a) Ecuación (36) (b) Ecuación (37) (c) Ecuación (38) y (d) Ecuación (39) usando γ = 0,8, T = 50, N = 10 y q = 100.

5. Conclusiones

Este estudio introdujo un enfoque novedoso, el método pseudoespectral multidominio de Chebyshev para resolver ecuaciones diferenciales fraccionarias. La novedad del método surgió del uso de multidominios no superpuestos, polinomios de Chebyshev de primer tipo y puntos de colocación de Gauss-Lobatto en el desarrollo del esquema, que se adaptó específicamente para manejar ecuaciones diferenciales fraccionarias. El método aprovecha las propiedades globales de los métodos basados en espectros para discretizar operadores no locales, como los operadores diferenciales fraccionarios, la convergencia exponencial del método pseudoespectral basado en polinomios de Chebyshev y la capacidad del enfoque multidominio para mantener la precisión de las soluciones a problemas con dinámica de largo plazo. La región de estabilidad del esquema multidominio se estableció mediante análisis de estabilidad lineal. El análisis de estabilidad mostró que es posible obtener un esquema numérico absoluto y estable según Lubich para algunos valores de orden arbitrario no entero. Se llevaron a cabo varios experimentos numéricos con FDE individuales lineales y no lineales y sistemas de FDE no lineales. Los ejemplos se eligieron cuidadosamente para abarcar distintos niveles de complejidad, incluidos problemas de valores iniciales fraccionarios con dinámicas de largo plazo, como sistemas caóticos, que demostraron beneficiarse significativamente de un esquema numérico multidominio. Estos ejemplos han establecido que el método numérico multidominio propuesto proporciona resultados muy precisos. Para validar el método, se realizaron comparaciones de los resultados numéricos obtenidos en este estudio con soluciones analíticas, cuando estuvieron disponibles, y soluciones aproximadas de la literatura. Se encontró que el método propuesto daba soluciones más precisas que las soluciones aproximadas de la literatura. Cuando no se disponía de una solución exacta, se realizó un análisis de error residual para establecer la precisión del método propuesto. En todos los ejemplos considerados, se encontró que el método produce soluciones comparat

Machine Translated by Google

En estudios futuros, exploraremos la aplicabilidad del método pseudoespectral de Chebyshev en dominios particionados superpuestos o no superpuestos a problemas de valores de frontera fraccionarios con dinámica no lineal.

Contribuciones del autor: SDO: conceptualización, investigación, metodología, software, redacción: borrador original, revisión y edición, supervisión. NM: Investigación, Metodología, Redacción: Borrador original, Revisión y Edición. YOT: investigación, metodología, redacción: borrador original, revisión y edición. OO: Conceptualización, Metodología, Redacción—Borrador Original, Revisión y Edición, Supervisión. Todos los autores han leído y aceptado la versión publicada del manuscrito.

Financiamiento: este trabajo se basa en la investigación apoyada en parte por la Fundación Nacional de Investigación de Sudáfrica (Número de referencia TTK2204163593). El APC fue financiado por la Universidad de Witwatersrand.

Declaración de la Junta de Revisión Institucional: No aplicable.

Declaración de Consentimiento Informado: No aplicable.

Declaración de disponibilidad de datos: los datos generados están contenidos en el artículo.

Agradecimientos: Yusuf Olatunji Tijani agradece el apoyo de la Universidad de Rodas.

Conflictos de intereses: Los autores declaran que no existen conflictos de intereses.

Abreviaturas

En este manuscrito se utilizan las siguientes abreviaturas:

CGL Chebyshev-Gauss-Lobatto
FDE Ecuaciones diferenciales fraccionarias
Método de colocación espectral multidominio MSCM

Referencias

1. Podlubny, I. Ecuaciones diferenciales fraccionarias: introducción a las derivadas fraccionarias, ecuaciones diferenciales fraccionarias y métodos de Su solución y algunas de sus aplicaciones; Elsevier: Ámsterdam, Países Bajos, 1998.
2. Oldham, K.; Spanier, J. La teoría del cálculo fraccional y aplicaciones de diferenciación e integración al orden arbitrario; Elsevier: Ámsterdam, Países Bajos, 1974.
3. Caputo, M.; Fabrizio, M. Una nueva definición de derivada fraccionaria sin núcleo singular. Prog. Fracta. Diferir de. Aplica. 2015, 1, 73–85.
4. Atangana, A.; Baleanu, D. Nuevas derivadas fraccionarias con núcleo no local y no singular: teoría y aplicación al calor modelo de transferencia. arXiv 2016, arXiv:1602.03408.
5. Javadi, R.; Mesgarani, H.; Nikan, O.; Avazzadeh, Z. Resolver ecuaciones diferenciales de orden fraccionario mediante el uso de una red neuronal de función de base radial fraccionaria. Simetría 2023, 15, 1275. [CrossRef]
6. Benchochra, M.; Karapinar, E.; Lazreg, JE; Salim, A. Ecuaciones diferenciales fraccionarias: nuevos avances para fracciones generalizadas Derivados; Springer: Cham, Suiza, 2023.
7. Ahmadian, S.; Darvishi, M. Versión fraccionaria de la ecuación de Biswas-Milovic (1 + 1) dimensional y sus soluciones. Optik 2016, 127, 10135–10147. [Referencia cruzada]
8. Manafian, J.; Lakestani, M. Un nuevo enfoque analítico para resolver algunas de las ecuaciones diferenciales parciales de orden fraccionario. Indio J. Física. 2017, 91, 243–258. [Referencia cruzada]
9. Kumar, S.; Kumar, D.; Abbasbandy, S.; Rashidi, M. Solución analítica de la ecuación fraccionaria de Navier-Stokes mediante el uso de Método de descomposición de Laplace. Ain Shams Ing. J. 2014, 5, 569–574. [Referencia cruzada]
10. Shakeel, M.; Mohyud-Din, ST Un nuevo método de expansión (G'/G) y su aplicación a la simetría fraccionaria del espacio-tiempo ecuación de onda larga regularizada (SRLW). Adv. Tendencias Matemáticas. 2015, 2, 1–16. [Referencia cruzada]
11. Zhu, Q.; Qi, J. Soluciones exactas de la onda larga regularizada simétrica diferencial parcial fraccional espacio-tiempo no lineal (SRLW) Ecuación empleando dos métodos. Adv. Matemáticas. Física. 2022, 2022, 8062119. [Referencia cruzada]
12. Senol, M. Nuevas soluciones analíticas de ecuación de onda larga regularizada simétrica fraccionaria. Rev. Méx. De Física 2020, 66, 297–307. [Referencia cruzada]
13. Ünal, S.Ç.; Da¸scıo ˘glu, A.; Bayram, DV Jacobi soluciones de funciones elípticas de espacio-tiempo fraccional simétrica regularizada larga ecuación de onda. Matemáticas. Ciencia. Aplica. Notas electrónicas 2021, 9, 53–63. [Referencia cruzada]
14. Tuluce Demiray, S.; Bulut, H.; Belgacem, FBM Método de transformada Sumudu para soluciones analíticas de tipo fraccionario ordinario ecuaciones diferenciales. Matemáticas. Problema. Ing. 2015, 2015, 131690. [Referencia cruzada]
15. Kumar, D.; Darvishi, M.; Joardar, A. Método Kudryashov modificado y su aplicación a la versión fraccionada de la variedad de Ecuaciones tipo Boussinesq en aguas poco profundas. Optar. Electrón cuántico. 2018, 50, 128. [Referencia cruzada]

Machine Translated by Google

16. He, JH Método de la función Exp para ecuaciones diferenciales fraccionarias. Int. J. Ciencia no lineal. Número. Simultáneo. 2013, 14, 363–366. [Referencia cruzada]
17. Ahmadian, S.; Darvishi, M. Un nuevo modelo fraccionario de Biswas-Milovic con sus soluciones periódicas de solitones. Optik 2016, 127, 7694–7703. [Referencia cruzada]
18. Turan, DA Un método wavelet seno-coseno para las soluciones de aproximación de la ecuación fraccionaria de Bagley-Torvik. Sigma J. Ing. Nat. Ciencia. Ve Fen Bilim. Derg. 2022, 40, 150–154.
19. Khan, A.; Shah, K.; Luo, D. Soluciones numéricas para algunas familias de ecuaciones diferenciales de orden fraccionario mediante polinomios de Laguerre. En Sistemas No Lineales: Aspectos Teóricos y Aplicaciones Recientes; IntechOpen: Londres, Reino Unido, 2020.
20. Sol, ZZ; Gao, GH Ecuaciones diferenciales fraccionarias: métodos de diferencias finitas; Walter de Gruyter GmbH & Co KG: Berlín, Alemania, 2020.
21. Hattaf, K.; Hajhouji, Z.; Ait Ichou, M.; Yousfi, N. Un método numérico para ecuaciones diferenciales fraccionarias con nuevos generalizados Derivada fraccionaria de Hattaf. Matemáticas. Problema. Ing. 2022, 2022, 3358071. [Referencia cruzada]
22. Kumar, M.; Jhinga, A.; Daftardar-Gejji, V. Métodos numéricos de orden superior para ecuaciones diferenciales de retardo fraccionario. Indio J. Pura aplicación. Matemáticas. 2024, 1–22. [Referencia cruzada]
23. Jhinga, A.; Daftardar-Gejji, V. Un nuevo método numérico para resolver ecuaciones diferenciales de retardo fraccionario. Computadora. Aplica. Matemáticas. 2019, 38, 166. [Referencia cruzada]
24. Lee, S.; Kim, H.; Jang, B. Un método numérico novedoso para resolver ecuaciones diferenciales de orden fraccionario no lineales y sus Aplicaciones. Fractal fractal. 2024, 8, 65. [Referencia cruzada]
25. Kumar, S.; Shaw, PK; Abdel-Aty, AH; Mahmoud, EE Un estudio numérico sobre ecuación diferencial fraccionaria con modelo de crecimiento poblacional. Número. Métodos de diferencia parcial. Ecuaciones 2024, 40, e22684. [Referencia cruzada]
26. Bhrawy, AH; Taha, TM; Machado, JAT Una revisión de matrices operativas y técnicas espectrales para cálculo fraccionario. Dyn no lineal. 2015, 81, 1023–1052. [Referencia cruzada]
27. Baleanu, D.; Bhrawy, A.; Taha, T. Un método espectral de Laguerre generalizado modificado para ecuaciones diferenciales fraccionarias en la media línea. En Actas del análisis abstracto y aplicado; Biblioteca en línea Wiley: Hoboken, Nueva Jersey, EE. UU., 2013; Volumen 2013, pág. 413529.
28. Delkhosh, M.; Parand, K. Un nuevo método computacional basado en funciones fraccionarias de Lagrange para resolver fracciones de varios términos ecuaciones diferenciales. Número. Algoritmos 2021, 88, 729–766. [Referencia cruzada]
29. Xu, Q.; Zheng, Z. Método de colocación espectral para ecuaciones diferenciales/integrales fraccionarias con operador fraccionario generalizado. Int. J. Difieren. Ecuaciones 2019, 2019, 3734617. [CrossRef]
30. Raslan, KR; Abd El Salam, MA; Ali, KK; Mohamed, método EM Spectral Tau para resolver diferenciales generales de orden fraccionario ecuaciones con argumento funcional lineal. J. Egipto. Matemáticas. Soc. 2019, 27, 33. [Referencia cruzada]
31. Shen, J.; Sheng, C. Métodos espectrales para ecuaciones diferenciales fraccionarias utilizando funciones de Jacobi generalizadas. Mano. Fracta. Calc. Aplica. Número. Métodos 2019, 3, 127–155.
32. Zayernouri, M.; Karniadakis, GE Métodos espectrales y de elementos espectrales exponencialmente precisos para EDO fraccionarias. J. Computación. Física. 2014, 257, 460–480. [Referencia cruzada]
33. Zhao, J.; Zhang, Y.; Xu, Y. Métodos implícitos de Runge-Kutta y Galerkin espectral para el espacio de Riesz de orden fraccionario/distribuido ecuación de difusión. Computadora. Aplica. Matemáticas. 2020, 39, 47. [Referencia cruzada]
34. Khader, M.; Sweilam, N. Soluciones aproximadas para la ecuación fraccionaria de advección-dispersión utilizando pseudoespectral de Legendre método. Computadora. Aplica. Matemáticas. 2014, 33, 739–750. [Referencia cruzada]
35. Cardone, A.; D'Ambrosio, R.; Paternoster, B. Un método espectral para ecuaciones diferenciales fraccionarias estocásticas. Aplica. Número. Matemáticas. 2019, 139, 115-119. [Referencia cruzada]
36. Doha, EH; Bhrawy, AH; Ezz-Eldien, SS Métodos espectrales eficientes de Chebyshev para resolver órdenes fraccionarios de varios términos ecuaciones diferenciales. Aplica. Matemáticas. Modelo. 2011, 35, 5662–5672. [Referencia cruzada]
37. Sweilam, Nuevo Hampshire; Khader, MM Un método pseudoespectral de Chebyshev para resolver ecuaciones integrodiferenciales de orden fraccionario. ANZIAM J. 2010, 51, 464–475. [Referencia cruzada]
38. Dabiri, A.; Karimi, L. Un método pseudoespectral de Chebyshev para resolver problemas de control óptimo de orden fraccionario. En Actas de la Conferencia Estadounidense de Control de 2019, Filadelfia, PA, EE. UU., 10 a 12 de julio de 2019.
39. Oloniiju, SD; Goqo, SP; Sibanda, P. Un método pseudoespectral geométricamente convergente para ecuaciones diferenciales parciales fraccionarias multidimensionales de dos lados. J. Aplica. Anal. Computadora. 2021, 11, 1699–1717. [Referencia cruzada] [PubMed]
40. Oloniiju, SD; Goqo, SP; Sibanda, P. Un método pseudoespectral de Chebyshev para las soluciones numéricas de ecuaciones diferenciales ordinarias fraccionarias de orden distribuido. Aplica. Matemáticas. Notas electrónicas 2022, 22, 132-141.
41. Oloniiju, SD; Goqo, SP; Sibanda, P. Un método pseudoespectral para ecuaciones diferenciales fraccionarias de espacio bilateral de orden distribuido en el tiempo. Taiwán. J. Matemáticas. 2021, 25, 959–979. [Referencia cruzada]
42. Adel, M.; Khader, MM; Assiri, TA; Kallel, W. Simulación numérica del modelo COVID-19 utilizando una técnica de relajación espectral multidominio. Simetría 2023, 15, 931. [CrossRef]
43. Mkhatshwa, M.; Khumalo, M.; Dlamini, P. Método de colocación espectral multivariante multidominio para (2 + 1) dimensional no lineal ecuaciones diferenciales parciales. Diferencia parcial. Aplicación de ecuaciones. Matemáticas. 2022, 6, 100440. [Referencia cruzada]
44. Samuel, F.; Motsa, S. Resolución de ecuaciones diferenciales parciales hiperbólicas utilizando un método de colocación espectral bivariada multidominio de alta precisión. Movimiento ondulatorio 2019, 88, 57–72. [Referencia cruzada]
45. Birem, M.; Klein, C. Método espectral multidominio para ecuaciones de Schrödinger. Adv. Computadora. Matemáticas. 2016, 42, 395–423. [Referencia cruzada]

Machine Translated by Google

46. Jung, JH; Olmos-Liceaga, D. Un método de malla adaptativa multidominio de Chebyshev para ecuaciones de reacción-difusión. Aplica. Número. Matemáticas. 2023, 190, 283–302. [Referencia cruzada]
47. Qin, Y.; Chen, Y.; Huang, Y.; Ma, H. Método de mínimos cuadrados multidominio de Legendre-Galerkin para ecuaciones diferenciales lineales con coeficientes variables. Número. Matemáticas. Métodos teóricos aplicados. 2020, 13, 665–688.
48. Maleki, M.; Kajani, MT Aproximaciones numéricas para el modelo de crecimiento poblacional de Volterra con orden fraccionario a través de un multi-método pseudoespectral de dominio. Aplica. Matemáticas. Modelo. 2015, 39, 4300–4308. [Referencia cruzada]
49. Bambe Moutsinga, CR Métodos espectrales de tiempo robustos para resolver ecuaciones diferenciales fraccionarias en finanzas. Doctor en Filosofía. Tesis, Universidad de Pretoria, Pretoria, Sudáfrica, 2021.
50. Kharazmi, E. Métodos espectrales de dominio único/multidominio basados en datos para PDE fraccionarias estocásticas. Doctor en Filosofía. Tesis, Estado de Michigan Universidad, East Lansing, MI, EE. UU., 2018.
51. Zhao, T.; Mao, Z.; Karniadakis, GE Método de colocación espectral multidominio para diferencial fraccional no lineal de orden variable ecuaciones. Computadora. Métodos de aplicación. Mec. Ing. 2019, 348, 377–395. [Referencia cruzada]
52. Klein, C.; Stoilov, N. Enfoque espectral multidominio para derivadas fraccionarias de orden racional. Semental. Aplica. Matemáticas. 2024, 152, 1110–1132. [Referencia cruzada]
53. Tavassoli Kajani, M. Solución numérica de ecuaciones de pantógrafo fraccionario mediante polinomios de Müntz-Legendre. Matemáticas. Ciencia. 2023, 18, 387–395. [Referencia cruzada]
54. Xu, Q.; Hesthaven, JS Métodos estables de penalización espectral multidominio para ecuaciones diferenciales parciales fraccionarias. J. Computación. Física. 2014, 257, 241–258. [Referencia cruzada]
55. Hayat, U.; Kamran, A.; Ambreen, B.; Yildirim, A.; Mohyud-din, ST Sobre el sistema de ecuaciones diferenciales parciales fraccionarias de tiempo. Walailak J. Ciencias. Tecnología. (WJST) 2013, 10, 437–448.
56. Oloniiju, S.; Nkomo, N.; Goqo, S.; Sibanda, P. Método espectral desplazado de Chebyshev para fracciones de espacio-tiempo bidimensionales ecuaciones diferenciales parciales. Preimpresiones 2022, 2022050138. [CrossRef]
57. Trefethen, LN Métodos espectrales y de diferencias finitas para ecuaciones diferenciales ordinarias y parciales; Universidad de Cornell: Nueva York, Estados Unidos, 1996.
58. Dahlquist, GG Un problema de estabilidad especial para métodos lineales de varios pasos. Número BIT. Matemáticas. 1963, 3, 27–43. [Referencia cruzada]
59. Ehle, BL Sobre las aproximaciones de Padé a la función exponencial y los métodos A-estables para la solución numérica de problemas de valores iniciales. Doctor en Filosofía. Tesis, Universidad de Waterloo, Waterloo, ON, EE.UU., 1969.
60. Odibat, ZM; Momani, S. Un algoritmo para la solución numérica de ecuaciones diferenciales de orden fraccionario. J. Aplica. Matemáticas. Informática 2008, 26, 15-27.
61. Li, C.; Yi, Q.; Chen, A. Métodos de diferencias finitas con mallas no uniformes para ecuaciones diferenciales fraccionarias no lineales. J. Computadora. Física. 2016, 316, 614–631. [Referencia cruzada]
62. Raja, MAZ; Khan, JA; Qureshi, IM Solución del sistema de orden fraccionario de la ecuación de Bagley-Torvik utilizando métodos evolutivos inteligencia computacional. Matemáticas. Problema. Ing. 2011, 2011. [Referencia cruzada]
63. Zafar, AA; Kudra, G.; Awrejcewicz, J. Una investigación de la ecuación fraccionaria de Bagley-Torvik. Entropía 2019, 22, 28. [CrossRef] [PubMed]
64. Sene, N. Introducción al sistema caótico de orden fraccionario bajo operador fraccionario en sentido Caputo. Alex. Ing. J. 2021, 60, 3997–4014. [Referencia cruzada]
65. Azam, A.; Aqeel, M.; Ahmad, S.; Ahmad, F. Comportamiento caótico del flujo de estiramiento-giro-pliegue (STF) modificado con propiedad fractal. Dyn no lineal. 2017, 90, 1–12. [Referencia cruzada]
66. Li, C.; Liao, X.; Yu, J. Sincronización de sistemas caóticos de orden fraccionario. Física. Rev.E 2003, 68, 067203. [CrossRef]
67. Sene, N. Estudio de un sistema caótico de orden fraccionario representado por el operador Caputo. Complejidad 2021, 2021, 5534872. [Referencia cruzada]

Descargo de responsabilidad/Nota del editor: Las declaraciones, opiniones y datos contenidos en todas las publicaciones son únicamente de los autores y contribuyentes individuales y no de MDPI ni de los editores. MDPI y/o los editores renuncian a toda responsabilidad por cualquier daño a personas o propiedad que resulte de cualquier idea, método, instrucción o producto mencionado en el contenido.