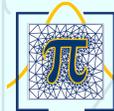




UNAH
UNIVERSIDAD NACIONAL
AUTÓNOMA DE HONDURAS



Maestría en
Matemática

BOLETÍN DIVULGATIVO

MAYO 2021

TEMÁTICA:

- SISTEMAS DINÁMICOS
- MODELACIÓN MATEMÁTICA
- AUTÓMATAS CELULARES
- ECUACIONES DIFERENCIALES ESTOCÁSTICAS
- MODELOS ESTOCÁSTICOS Y MODELOS VAR

Presentación

Este documento fué desarrollado por la Coordinación de Investigación y Vinculación de la Maestría en Matemáticas de la UNAH, presenta artículos divulgativos desarrollados por profesionales del Seminario de Investigación de Ingeniería Matemática y de Estadística Matemática de la tercera promoción del programa, cursos desarrollado durante el primer período académico del año 2021. Se abarca una temática bastante amplia: modelación matemática, sistemas dinámicos, autómatas celulares, ecuaciones diferenciales estocásticas, modelos estocásticos y modelos VAR; en algunos de los trabajos se desarrolló una revisión bibliográfica de trabajos pertinentes y se resumió según lo comprendido por cada autor, en otros casos, se realizó avances en sus trabajos de tesis que incluso incluyen experimentación.

El objetivo principal de desarrollar este documento es que a futuro, en base a la experiencia obtenida y después de tener varias experiencias similares, se transforme en una revista científica de Matemáticas, cuestión que requiere de mucho trabajo por parte del equipo de profesores investigadores del programa y otros colaboradores externos; además de ser una muestra de que en el programa se está desarrollando en los estudiantes un espíritu investigador.

Todas las revisiones bibliográficas y temas aquí presentados se encasillan dentro de las líneas de investigación de la UNAH, entre los temas prioritarios abarcados se encuentran: ciencia, cambio climático y vulnerabilidad, productividad, infraestructura y desarrollo territorial. Esto evidencia que el programa está sumamente interesado en colaborar con las prioridades investigativas de la universidad y mantiene un compromiso con vincularse con la sociedad.

En esta ocasión también se presenta una publicación original realizada por el matemático hondureño Oswaldo Sevilla, Licenciado en Matemática de la UNAH, Máster

en Matemática de CIMAT y pronto a culminar su Doctorado en Matemática también en CIMAT, entre otros estudios e investigaciones realizadas. Actualmente es Profesor de matemáticas de la Carrera de Matemática y de la Maestría en Matemática de la UNAH. Nuestro programa le agradece enormemente esta contribución, ya que contribuye al crecimiento de la publicación y es un paso importante hacia la realización del objetivo de transformarse en una revista de matemáticas.

Mayo del año 2021, Ciudad Universitaria
Tegucigalpa, M.D.C., Honduras

© Maestría en Matemáticas - UNAH
Edificio F1, Segundo Piso, Ciudad Universitaria
Tegucigalpa, M.D.C. Honduras.
<https://mm.unah.edu.hn/>
maestria.matematica@unah.edu.hn
Tel. 2216-3000 Ext. 100647

Contenido

1. Una cota inferior para el rango de Mordell-Weil de la fibra genérica de una superficie K3 elíptica dada como el cubriente doble ramificado de una superficie elíptica racional particular - Oswaldo Sevilla (p. 1 - 18)
2. Un método FWI usando el método de optimización de Chen - Victor Armijo (p. 19 - 31)
3. Identificación de modelos esparcidos en dinámica económica de alta dimensionalidad - Kerin Cardona (p. 32 - 42)
4. Identificación aproximada de señales en reconocimiento de patrones - William Funez (p. 43 - 54)
5. Análisis dinámico en reacciones químicas modeladas con sistemas dinámicos no lineales de orden fraccionario - Oscar Molina (p. 55 - 67)
6. Criticidad auto-organizada, el modelo de la pila de arena - Norman R. Sabillón (p. 68 - 79)
7. Modelo estocástico presa-predador - Elvis R. Arrazola (p. 80 - 94)
8. Ecuaciones diferenciales estocásticas en el estudio del desarrollo de arbovirus en Honduras a partir de un modelos susceptible-infectado-susceptible - Carlos A. Henríquez (p. 95 - 108)

9. Aplicaciones de redes neuronales - Cristian F. Juárez
.....(p. 109 - 123)

10. Aplicación de modelos VAR al análisis de riesgo financiero (Honduras) - Angel
Matute
.....(p. 124 - 134)

11. Modelos estadísticos de análisis de factores que afectan el rendimiento escolar
en matemática y español en tercer y sexto grado en Honduras - Lesky I. Rivas
.....(p. 135 - 152)

Una cota inferior para el rango de Mordell–Weil de la fibra genérica de una superficie K3 elíptica dada como el cubriente doble ramificado de una superficie elíptica racional particular

Oswaldo Sevilla

Resumen

Se encuentra una cota inferior para el rango de Mordell–Weil de una fibración elíptica semi-estable de una superficie K3 construida como un cubriente ramificado de una superficie elíptica racional con sección. Esto se obtiene aplicando la fórmula de Shioda-Tate y obteniendo una inyección del grupo de secciones de la superficie racional elíptica en el grupo de secciones de la superficie K3.

1. Introducción

Las superficies K3 son un tema de estudio muy relevante, encontrándose en la intersección de estudios en Geometría Compleja, Geometría Algebraica y Geometría Aritmética. Las superficies K3 aparecen también en algunos estudios de Teoría de Cuerdas en Física. Son variedades de Calabi–Yau de dimensión 2, una generalización natural de las curvas elípticas. Algunas de sus propiedades algebro-geométricas son notablemente difíciles de calcular, particularmente sus números de Picard y el comportamiento de los números de Picard en familias de superficies K3.

Las superficies K3 tienen una relación interesante con las curvas elípticas. En particular, toda superficie K3 con número de Picard al menos 5 posee una fibración elíptica, que se conoce como *superficie K3 elíptica*. La fórmula de

Shioda–Tate muestra una relación admirable entre la aritmética de las curvas elípticas y la geometría de las superficies K3 elípticas, y es una herramienta valiosa para poder estudiar el número de Picard de estas superficies.

En este trabajo se presenta un caso especial de superficies K3, y se calcula para un caso específico una cota inferior de su número de Picard usando una fibración elíptica.

2. Curvas elípticas

Para detalles sobre variedades algebraicas, y términos que por el enfoque del artículo no se han discutido, véase [5]. Como referencia sobre curvas elípticas, véase [10].

Una curva algebraica es una variedad algebraica de dimensión (algebraica) 1.

El género g de una curva algebraica compleja lisa coincide con su género topológico: es su número de asas, i.e., la curva es homeomorfa a una suma conexa de g toros topológicos $S^1 \times S^1$.

Definición 1 *Una curva elíptica sobre un campo K es un par (E, O) donde E es una curva algebraica completa y lisa de género 1 definida sobre K y O es un punto K -racional distinguido en E .*

Note que una variedad sobre \mathbb{C} es completa si y sólo si es compacta en la topología usual de \mathbb{C} (llamada topología trascendente en este contexto).

En este trabajo, a menos que se indique lo contrario, se trabajará con curvas elípticas sobre \mathbb{C} .

Observe que frecuentemente se llama curva elíptica a una curva lisa de dimensión 1 sin tomar en cuenta un punto distinguido.

Para una curva elíptica E definida sobre \mathbb{Q} , el conjunto de los puntos racionales de la curva E , denotado por $E(\mathbb{Q})$, tiene estructura de grupo. Este grupo es llamado el *grupo de Mordell–Weil* de la curva E . El punto distinguido es el elemento neutro del grupo.

El teorema de Mordell–Weil dice que el grupo de Mordell–Weil de una curva elíptica es un grupo abeliano finitamente generado (ver [10]). El rango de la parte libre de este grupo se llama el *rango de Mordell–Weil* de la curva elíptica.

En general, calcular el rango de MW de una curva elíptica es una tarea que puede volverse complicada. Se puede simplificar este cálculo en algunos casos específicos. Para ejemplos de estos cálculos puede verse [10].

Observación 2 *Note que para superficies elípticas sobre K se define el grupo de Mordell–Weil de una forma que coincide con el grupo de puntos $K(C)$ -racionales de la curva sobre el punto genérico de la curva base C . En ese caso no estudiamos los puntos sobre \mathbb{Q} .*

3. Superficies algebraicas

Aquí se precisan algunos términos sobre superficies algebraicas a ser usados en el artículo. Para detalles sobre superficies algebraicas complejas véase [2].

Una superficie algebraica sobre un campo K es una variedad algebraica de dimensión 2 definida sobre K . Se puede ver desde el punto de vista clásico y de manera más intuitiva como un conjunto que posee una cubierta por abiertos que dados por variedades algebraicas afines, tal que el pegado de estas cartas funciona bien. En el resto de este trabajo consideramos los campos \mathbb{C} y \mathbb{Q} , se trabajará sobre \mathbb{C} a menos que se especifique el campo.

Ejemplos sencillos de superficies sobre \mathbb{Q} son:

- el espacio afín de dimensión 2 sobre \mathbb{Q} (puede verse como \mathbb{Q}^2).
- el conjunto de puntos racionales que son ceros de un polinomio en tres variables con coeficientes racionales.
- el plano proyectivo racional $\mathbb{P}_{\mathbb{Q}}^2$.

La relación de birracionalidad es muy importante en geometría algebraica. Puede decirse de manera breve que dos variedades X, Y son birracionales si existen mapeos birracionales $X \rightarrow Y$ y $Y \rightarrow X$ tales que al restringirse a abiertos de Zariski dados en ambas variedades definen un isomorfismo. Para detalles sobre relación birracional de superficies véase [5].

Definición 3 *Una superficie se llama racional si es birracional a \mathbb{P}^2 (el plano proyectivo de dimensión 2).*

Definición 4 *La irregularidad $q(S)$ de una superficie compleja compacta S es la dimensión de $H^1(S, \mathbb{C})$.*

Si la superficie es simplemente conexa entonces $q(S) = 0$.

Definición 5 *El género geométrico $p_g(S)$ de una superficie compleja compacta S es la dimensión de $H^2(S, \mathcal{O}_S)$*

Este número coincide con la dimensión del espacio de secciones globales de 2-formas holomorfas.

Definición 6 *Cuando este espacio de secciones globales está generado por una 2-forma holomorfa global, se dice que su clase canónica es trivial*

Cuando la clase canónica de S es trivial, $p_g(S) = 1$.

Observaciones 7 *Los la irregularidad y el género geométrico de una superficie S se pueden definir a partir de sus números de Hodge: $q(S) = h^{0,1}(S)$ y $p_g(S) = h^{0,2}(S)$.*

Se escribirá simplemente q y p_g cuando no haya ambigüedad al respecto de la superficie de la que se esté hablando.

Los blowups de puntos de una superficie son la base de la relación birracional entre superficies algebraicas. Encontrar invariantes birracionales es sumamente importante en teoría de superficies.

Denotemos por $\text{Bl}_p(S)$ el blowup de una superficie S en un punto p . La cohomología de la superficie S tiene un comportamiento interesante bajo blowups:

Lema 8 (*[1, Lemma 11.1.5]*) *Para una superficie algebraica S*
 $H^1(\text{Bl}_p(S)) \cong H^1(S)$ y $H^2(\text{Bl}_p(S)) \cong H^2(S) \oplus \mathbb{Z}$

La irregularidad y el género geométrico son invariantes bajo ciertas relaciones. En particular bajo blowups y blowdowns:

Proposición 9 *q y p_g son invariantes bajo blowups y blowdowns de un punto en una superficie.*

De demostración: Por el lema 8, cada blowup (blowdown) deja invariante $q(S) = h^{0,1}(S)$, y aumenta (reduce) el segundo número de Betti $b^2(S)$ en uno.

Por la descomposición de Hodge de S se tiene que

$$b^2(S) = h^{1,1}(S) + 2h^{0,2}(S) = h^{1,1}(S) + 2p_g(S) \quad .$$

Luego, la única posibilidad es que $h^{1,1}(S)$ aumenta en uno con cada blowup. y cada blowdown lo reduce en uno. En consecuencia, $p_g(S)$ se mantiene constante. Q.E.D.

La irregularidad y el género geométrico son invariantes bajo un tipo más general de relaciones entre superficies: son invariantes birracionales.

Proposición 10 *q y p_g son invariantes birracionales de superficies.*

Deomstración: Esto sucede como consecuencia del teorema de contracción de Castelnuovo para superficies algebraicas birracionales (cf. [6], [1] o [5] y las referencias en ellos). El teorema de Castelnuovo nos dice que dos superficies birracionales pueden transformarse una en otra mediante un número finito de blowups y blowdowns. En consecuencia, tanto la irregularidad como el género geométrico de una superficie son invariantes birracionales. Q.E.D.

Definición 11 *El grupo de divisores de Weil de una variedad algebraica es el grupo abeliano libre generado por las subvariedades de codimensión 1.*

Como recordatorio, en una superficie las subvariedades de codimensión 1 son curvas.

Definición 12 *Un divisor principal es un divisor asociado a una función meromorfa f en la superficie, dado por $(f) = (f)_0 - (f)_\infty$, donde $(f)_0$ es la subvariedad de ceros de f y $(f)_\infty$ es la subvariedad de polos de f .*

Definición 13 *Dos divisores son linealmente equivalentes si su diferencia es un divisor principal.*

Definición 14 *Un divisor de Cartier de una variedad algebraica es un divisor localmente principal.*

Para referencia sobre divisores localmente principales, el lector puede recurrir a [9] y [5].

Definición 15 *El grupo de Picard $\text{Pic}(X)$ de una variedad algebraica X está definido como el grupo de haces lineales invertibles.*

Alternativamente, el grupo de Picard se puede definir como el grupo de divisores de Cartier módulo equivalencia lineal.

Definición 16 *El grupo de Néron–Severi $\text{NS}(X)$ de una variedad X es el grupo de Picard de X módulo equivalencia algebraica.*

Definición 17 *El número de Picard de una variedad es el rango de su grupo de Néron–Severi.*

4. Superficies elípticas

Las superficies elípticas son de gran interés y poseen muchas aplicaciones en geometría. Un ejemplo interesante es el artículo de Schoen [8], donde se construyen variedades Calabi–Yau a partir de superficies elípticas, y se calculan sus números de Hodge usando propiedades de esas fibraciones.

Definición 18 *Una superficie elíptica es una superficie proyectiva lisa \mathcal{S} que posee una fibración sobre una curva base C , cuyas fibras son todas, excepto un número finito, curvas lisas de género 1.*

Sea $\psi : \mathcal{S} \rightarrow \mathbb{P}^1$ una superficie elíptica. Denote por $\mu_{\text{crit}}(\mathcal{S})$ (o simplemente μ_{crit}) al conjunto de valores críticos de ψ , i.e., el conjunto de valores en \mathbb{P}^1 a los que les corresponde una fibra singular. Este conjunto también se conoce como el conjunto discriminante de ψ .

De particular interés son las superficies elípticas minimales:

Definición 19 *Se dice que una fibración elíptica es minimal si ninguna de sus fibras posee un componente que sea curva -1 .*

Como recordatorio, las *curvas -1* son las curvas cuya autointersección es -1 según la forma de intersección de la superficie ambiente.

Definimos ahora los divisores verticales, los cuales tienen un papel importante en el estudio de una superficie elíptica:

Definición 20 *Un divisor D sobre una superficie elíptica \mathcal{S} es vertical si para toda fibra F se cumple*

$$D \cdot F = 0.$$

Para superficies elípticas minimales, tenemos el siguiente resultado sobre su clase canónica:

Proposición 21 *Dada una superficie elíptica minimal \mathcal{S} , su clase canónica tiene autointersección nula, i.e.*

$$K_{\mathcal{S}}^2 = 0.$$

De demostración: Para una fibra general F , se tiene, al usar la fórmula de adjunción, que $K_{\mathcal{S}} \cdot F = 0$.

Usando el lema de Zariski (cf. [2]), se obtiene que $K_{\mathcal{S}}$ es suma de múltiplos racionales de fibras. Por tanto $K_{\mathcal{S}}^2 = 0$. Q.E.D.

En lo que resta del artículo sólo se trabajará con superficies elípticas minimales.

Si la superficie elíptica admite una sección, se llamará *superficie elíptica con sección*.

Definición 22 *Para una superficie elíptica $\pi : \mathcal{S} \rightarrow C$, una sección se define como un morfismo*

$$\sigma : C \rightarrow \mathcal{S} \text{ tal que } \pi \circ \sigma = \text{id}_C \quad .$$

En lo que sigue de esta sección, S denota una superficie algebraica que es de otra manera arbitraria, \mathcal{S} una superficie elíptica y \mathfrak{S} una superficie elíptica racional (a definirse adelante). Además, a menos que se indique lo contrario, se consideran solamente superficies elípticas con sección.

Sea $\pi : \mathcal{S} \rightarrow C$ una fibración elíptica con una sección dada σ . Este σ elige en cada fibra un punto, definiendo así una fibración por curvas elípticas.

Llame $\text{MW}(\mathcal{S})$ al conjunto de todas las secciones de la superficie elíptica \mathcal{S} (es decir, secciones de π). La adición fibra por fibra induce una ley de grupo en $\text{MW}(\mathcal{S})$ con σ como elemento neutro.

Definición 23 *El grupo $\text{MW}(\mathcal{S})$ de secciones de una superficie elíptica con sección $\pi : \mathcal{S} \rightarrow C$ se llama el grupo de Mordell–Weil de \mathcal{S} (o de π).*

Observación 24 *Para superficies elípticas compactas el grupo de Mordell–Weil es finitamente generado.*

Uno de los objetivos de este trabajo es calcular el grupo de Mordell–Weil de superficies elípticas racionales semi-estables.

Definición 25 *Una fibración elíptica es semi-estable si todas las fibras singulares son de tipo Kodaira I_n .*

Definición 26 *Una superficie elíptica racional (SER) \mathfrak{S} es una superficie elíptica birracional a \mathbb{P}^2 .*

Lema 27 *Para una superficie elíptica racional \mathfrak{S} se tiene $q(\mathfrak{S}) = 0$ y $p_g(\mathfrak{S}) = 0$*

.

De demostración: Observe que al ser la irregularidad y el género geométrico invariantes birracionales, y al ser \mathfrak{S} racional, se tiene:

$$q(\mathfrak{S}) = q(\mathbb{P}^2) = 0$$

y

$$p_g(\mathfrak{S}) = p_g(\mathbb{P}^2) = 0$$

Q.E.D.

Tenemos el siguiente resultado para SER semi-estables:

Proposición 28 *El número total de componentes de las fibras singulares en una superficie elíptica racional semi-estable es 12.*

De demostración: Usando la fórmula de Noether:

$$\chi(\mathcal{O}_{\mathfrak{S}}) = \frac{K_{\mathfrak{S}}^2 + e}{12}.$$

Como $K_{\mathfrak{S}}^2 = 0$ para una superficie elíptica minimal, tenemos que

$$\chi(\mathcal{O}_{\mathfrak{S}}) = \frac{e}{12}.$$

Note que para superficies elípticas racionales,

$$\chi(\mathcal{O}_{\mathfrak{S}}) = 1 - q(\mathfrak{S}) + p_g(\mathfrak{S}) = 1 \quad .$$

Como consecuencia

$$e(\mathfrak{S}) = 12. \tag{1}$$

Para $\mathcal{S} \rightarrow C$, una fibración elíptica semi-estable cualquiera sobre una curva C , llame E a una fibra general y E_c a la fibra sobre c . Tomando en cuenta que $e(E) = 0$, podemos calcular su característica topológica de Euler:

$$\begin{aligned} e(\mathcal{S}) &= e(E)e(C \setminus \mu_{\text{crit}}) + \sum_{c \in \mu_{\text{crit}}} (e(E_c) - e(E)) \\ &= \sum_{c \in \mu_{\text{crit}}} e(E_c) \end{aligned}$$

Para una superficie elíptica racional semi-estable esto se simplifica a:

$$\begin{aligned} e(\mathfrak{S}) &= \sum_{c \in \mu_{\text{crit}}} e(E_c) \\ &= \sum_{c \in \mu_{\text{crit}}} \#\text{componentes}(E_c) \end{aligned}$$

Esto sucede pues el número de componentes de una fibra de tipo Kodaira I_n y su número de nodos n coinciden. Así, usando (1) obtenemos

$$12 = \sum_{c \in \mu_{\text{crit}}} \#\text{componentes}(E_c) \quad . \quad (2)$$

Por tanto, el número total de componentes de las fibras singulares en una superficie elíptica racional semi-estable es 12. Q.E.D.

Observaciones 29 *Este resultado es importante al simplificar la fórmula de Shioda–Tate para calcular el número de Picard de una superficie elíptica racional.*

Observe además que este resultado es importante en otras situaciones, e.g. al estudiar las simplificaciones para las fórmulas en [8].

Es interesante notar que si la superficie no es minimal, el total de componentes de las fibras singulares puede ser mayor que 12.

Ahora calcularemos el número de Picard de una superficie elíptica racional.

Proposición 30 *Una superficie elíptica racional tiene número de Picard 10.*

De demostración: Sea \mathfrak{S} una superficie elíptica racional. Para una superficie lisa cualquiera S , sean $q(S)$ su irregularidad y $p_g(S)$ su género geométrico.

Denote por $b^n(S)$ the n -ésimo número de Betti de S , y sea ρ_S el número de Picard de esta superficie. Llame a $\lambda(S) = b^2(S) - \rho_S$ el *número de Lefschetz* de S .

λ , q y p_g son invariantes birracionalmente, como consecuencia del teorema de contracción de Castelnuovo para superficies algebraicas birracionalmente. En particular, observe que b^2 aumenta en uno con cada blowup y ρ también aumenta en uno con cada blowup, por tanto λ es invariante bajo blowups.

Dado que \mathfrak{S} es una superficie racional, $q(\mathfrak{S}) = q(\mathbb{P}^1) = 0$, $p_g(\mathfrak{S}) = p_g(\mathbb{P}^1) = 0$ y $\lambda(\mathfrak{S}) = \lambda(\mathbb{P}^2) = 0$.

Usando esta información y la proposición 4, el número de Euler está relacionado con el número de Picard por las ecuaciones

$$\begin{aligned}
 e(\mathfrak{S}) &= 12 = 2b^0(\mathfrak{S}) - 2b^1(\mathfrak{S}) + b^2(\mathfrak{S}) \\
 &= 2 - 4q(\mathfrak{S}) + \rho_{\mathfrak{S}} \\
 &= 2 + \rho_{\mathfrak{S}} \\
 &= 12
 \end{aligned}$$

Por tanto, se sigue que

$$\rho_{\mathfrak{S}} = 10.$$

Q.E.D.

De aquí en adelante, nos enfocaremos solamente en fibraciones elípticas *semi-estables*.

Ahora encontraremos una relación entre el rango de Mordell–Weil y el número de valores críticos. Note que el rango de MW de una superficie elíptica es igual al rango de MW de la fibra sobre el punto generico (en el sentido de esquemas, véase [5]).

Proposición 31 *Para una fibración elíptica $\mathcal{S} \rightarrow C$ con sección, se satisface la fórmula de Shioda–Tate:*

$$\rho_{\mathcal{S}} = r + 2 + \sum_{c \in \Delta} (\#\text{componentes}(\mathcal{S}_c) - 1).$$

donde r es el rango de Mordell–Weil de la fibración, \mathcal{S}_c es la fibra sobre $c \in C$ y Δ es el conjunto discriminante de la fibración.

Para su demostración, véase [7, corolario VII.2.4].

Teorema 32 *El rango de Mordell–Weil r de (la fibra genérica de) \mathfrak{S} satisface:*

$$r = \#\mu_{\text{crit}} - 4$$

Deomstración: La fórmula de Shioda–Tate para \mathfrak{S} es

$$\rho_{\mathfrak{S}} = r + 2 + \sum_{c \in \mu_{\text{crit}}} (\#\text{componentes}(X_c) - 1).$$

En consecuencia tenemos que

$$10 = r + 2 + \sum_{c \in \mu_{\text{crit}}} (\#\text{componentes}(X_c) - 1) \quad .$$

Usando (2) esto se simplifica a:

$$10 = r + 2 + 12 - \#\mu_{\text{crit}} \quad .$$

Por tanto

$$r = \#\mu_{\text{crit}} - 4.$$

Q.E.D.

Usando los números de Hodge anteriormente calculados, concluimos que el diamante de Hodge de una superficie elíptica racional \mathfrak{S} es

$$\begin{array}{ccccc} & & & & 1 \\ & & & & 0 & & 0 \\ & & & & 0 & 10 & 0. \\ & & & & 0 & & 0 \\ & & & & & & 1 \end{array}$$

Observe que $q(\mathfrak{S}) = 0$, $K_{\mathfrak{S}} \neq 0$ y $K_{\mathfrak{S}}^2 = 0$; entonces \mathfrak{S} es una superficie de Enriques.

5. Superficies K3

Definición 33 *Una superficie K3 definida sobre \mathbb{C} es una variedad proyectiva lisa de dimensión 2 con clase canónica trivial y de irregularidad cero.*

Las superficies K3 pueden verse como una generalización de las curvas elípticas en dos dimensiones.

El diamante de Hodge de una superficie K3 X es

$$\begin{array}{ccccc} & & & & 1 \\ & & & & 0 & & 0 \\ & & & & 1 & 20 & 1. \\ & & & & 0 & & 0 \\ & & & & & & 1 \end{array}$$

La sucesión exponencial en este caso muestra que el grupo de Picard $\text{Pic}(X)$ se inyecta en la cohomología intermedia $H^2(X, \mathbb{Z})$. Así, $\text{Pic}^0(X)$ es trivial, luego el grupo de Picard y el de Néron–Severi son isomorfos. Por el teorema (1,1) de Lefchetz, tenemos que el rango de Néron–Severi de X (i.e. su número de Picard) está acotado superiormente por la dimensión de $H^{1,1}(X)$, que es 20.

Las superficies K3 cuyo número de Picard es 20 se llaman *singulares*.

Es notable que existen variedades K3 que alcanzan un número de Picard 22. Estas se llaman *súper-singulares*, pero sólo pueden construirse sobre campos finitos.

6. Algunas superficies K3 elípticas

Es conocido que si el número de Picard de una superficie K3 es al menos 5, esta posee al menos una fibración elíptica. Las fibraciones elípticas son herramientas de cálculo muy valiosas, brindan información topológica y geométrica. También tienen un papel importante en Teoría-F (Física).

Construiremos superficies K3 como un cubriente doble ramificado sobre fibras no singulares de superficies elípticas racionales.

Dada S , una superficie K3, denote su número de Picard por

$$\rho_S := \rho(S) := \text{rank}(\text{NS}(S)) \quad .$$

Defina S como el pullback de una fibración elíptica de una superficie elíptica racional con sección $\psi: \mathfrak{S} \rightarrow \mathbb{P}^1$ dado por el diagrama:

$$\begin{array}{ccc} S & \xrightarrow{\pi} & \mathfrak{S} \\ \hat{\psi} \downarrow & & \psi \downarrow \\ \mathbb{P}^1 & \xrightarrow[2:1]{} & \mathbb{P}^1 \end{array} \quad (3)$$

Donde el cubriente doble π no se ramifica sobre ninguna curva singular.

Sea Y_t (resp. X_t) la fibra de S (resp. \mathfrak{S}) sobre t .

Proposición 34 *S es una superficie K3.*

Deomstración: Sólo necesitamos demostrar que la irregularidad es cero y la clase canónica es trivial.

Sea \tilde{X}_t una fibra S que se mapea a la fibra X_t de \mathfrak{S} , y sean a, b los puntos de ramificación del cubriente doble de \mathbb{P}^1 . De esta forma tenemos:

$$K_S = \pi^* K_{\mathfrak{S}} + \text{Ramificación} = -2\tilde{X}_{\infty} + \tilde{X}_a + \tilde{X}_b \sim -2\tilde{X}_{\infty} + \tilde{X}_{\infty} + \tilde{X}_{\infty} \sim 0$$

Por tanto la clase canónica es trivial.

Se puede usar la sucesión espectral de Leray para demostrar $q = 0$.

Una forma alternativa para probar $q = 0$ es tomar en cuenta que las fibras singulares se duplican todas, así que la característica topológica de Euler de la fibración es $e(S) = 2e(\mathfrak{S}) = 24$. Usando la fórmula de Noether y el hecho de que $K_S^2 = 0$ (pues S es minimal), obtenemos:

$$\begin{aligned}\chi(\mathcal{O}_S) &= 24/12 = 2 \\ \chi(\mathcal{O}_S) &= 1 - q + p_g = 2 - q\end{aligned}$$

De esta manera $q = 0$. Q.E.D.

Usando los resultados de la sección 4, el rango MW $r_{\mathfrak{S}}$ de la fibra genérica de \mathfrak{S} satisface:

$$r_{\mathfrak{S}} = \#\mu_{\text{crit}}(\mathfrak{S}) - 4$$

donde $\mu_{\text{crit}}(f)$ es el conjunto de valores críticos de la fibración f . Podemos escribir $\mu_{\text{crit}}(\text{superficie})$ si la fibración de la superficie es conocida y si no hay riesgo de confusión.

Teorema 35 *El rango de MW de \mathfrak{S} da una cota inferior para el rango de MW de S , i.e., $r_S \geq r_{\mathfrak{S}}$.*

De demostración: Dada una sección σ de \mathfrak{S} , existe un mapeo natural a una sección τ de S , dado por

$$\sigma \mapsto \tau_{\sigma} := (\sigma \circ F, \text{id}_{\mathbb{P}}^1).$$

Esto da un homomorfismo del grupo de secciones de \mathfrak{S} al grupo de secciones de S .

Este homomorfismo es inyectivo: suponga que se tienen dos secciones σ_1 y σ_2 tales que $\tau_{\sigma_1} = \tau_{\sigma_2}$. Entonces, para cada $z \in \mathbb{P}^1$ se cumple

$$\sigma_1(F(z)) = \sigma_2(F(z)) \quad .$$

Como F es sobreyectiva, $\sigma_1 = \sigma_2$, y en consecuencia el homomorfismo es inyectivo.

Dado que este homomorfismo es inyectivo, da una cota inferior para el grupo de secciones de S .

Q.E.D.

Teorema 36 Para una superficie K3 elíptica S definida como en el diagrama 3,

$$\rho_S \geq 22 - \#\mu_{\text{crit}}(\mathfrak{S}).$$

Deomstración: La fórmula de Shioda-Tate para S es:

$$\rho_S = r_S + 2 + \sum_{c \in \mu_{\text{crit}}(S)} (\#\text{componentes}(Y_c) - 1).$$

Dado que π no se ramifica en fibras singulares, esta fórmula puede simplificarse, pues el cubriente doble duplica todas las fibras singulares:

$$\rho_S = r_S + 2 + 2 \sum_{c \in \mu_{\text{crit}}(\mathfrak{S})} (\#\text{componentes}(X_c) - 1) = r_S + 2 + 2(12 - \#\mu_{\text{crit}}(\mathfrak{S}))$$

Además, dado que $r_S \geq r_{\mathfrak{S}} = \#\mu_{\text{crit}}(\mathfrak{S}) - 4$, obtenemos una expresión simple para la cota inferior del rango de Mordell–Weil de S :

$$\rho_S \geq \#\mu_{\text{crit}}(\mathfrak{S}) - 4 + 2 + 2(12 - \#\mu_{\text{crit}}(\mathfrak{S})) = 22 - \#\mu_{\text{crit}}(\mathfrak{S}) \quad (4)$$

Q.E.D.

7. Cálculo de la cota inferior del rango de Mordell–Weil para un caso específico

Usaremos esta cota para calcular explícitamente una cota para una superficie K3 asociada a una superficie elíptica racional.

Defina en coordenadas locales de $\mathbb{P}^1 \times \mathbb{P}^1$ el siguiente pincel:

$$\phi = (x + 1/x)(y + 1/y) = \mu \quad .$$

donde x e y corresponden a $(x : 1)$ y $(y : 1)$ respectivamente.

Este es un pincel de curvas elípticas en $\mathbb{P}^1 \times \mathbb{P}^1$. Esto puede comprobarse usando la fórmula de adjunción considerando que toda curva en este pincel es parte del sistema lineal anti-canónico, o considerando que $\mathbb{P}^1 \times \mathbb{P}^1$ es una variedad tórica definida por un polígono reflexivo y usando directamente los resultados de Batyrev en [3]. Este pincel posee 8 puntos base distintos, con coordenadas:

$$\begin{aligned} (\text{lugar base}) = \{ & ((0 : 1), (\pm i : 1)), ((\pm i : 1), (0 : 1)), \\ & ((1 : 0), (1 : \pm i)), ((1 : \pm i), (1 : 0)) \} \end{aligned}$$

Este pincel puede resolverse al hacer blowup de los 8 puntos base, lo cual incrementa en 8 el número de Picard de la superficie. Así, si llamamos \mathfrak{S} a la superficie después de estos blowup, tenemos

$$\rho(\mathfrak{S}) = \rho(\mathbb{P}^1 \times \mathbb{P}^1) + 8 = 2 + 8 = 10$$

Conocemos de antemano que $\rho(\mathbb{P}^1 \times \mathbb{P}^1) = 2$, ya sea usando la fórmula de Künneth, o considerando que $\mathbb{P}^1 \times \mathbb{P}^1$ es una superficie tórica.

Observe que el número de Picard de \mathfrak{S} concuerda con el número de Picard de una superficie elíptica racional. Claramente esta superficie elíptica es racional pues es blowup de $\mathbb{P}^1 \times \mathbb{P}^1$ que es racional.

\mathfrak{S} posee al menos una sección. Esto es claro pues el divisor excepcional correspondiente a cada punto base define una sección.

Para ver explícitamente una de estas secciones, considere la expresión del pincel en una carta afín de $\mathbb{P}^1 \times \mathbb{P}^1 = \{(x : z), (y : w) \mid (x : z), (y : w) \in \mathbb{P}^1\}$ dada por $z = 1$ y $w = 1$:

$$F = (x^2 + 1)(y^2 + 1) - \mu xy = 0 \quad .$$

Ahora, cambiemos variables para trasladar el punto base $x = i, y = 0$ al origen, escogiendo $x = u + i, v = y$.

Realizamos el blowup en el origen. El divisor excepcional E es una recta proyectiva. Dando las coordenadas naturales a los puntos del divisor excepcional, $(s : t) \in E$, al tomar el abierto de E dado por $s = 1$ obtenemos un abierto afín de la variedad. Ahí, el pincel tiene la siguiente expresión:

$$\begin{cases} F = ((u + i)^2 + 1)(v^2 + 1) - \mu(u + i)v = 0 \\ v = tu \end{cases}$$

Observe que en cada fibra se cumple

$$F = u(u + 2i)(v^2 + 1) - \mu(u + i)v = 0$$

Luego en la cerradura, se tiene

$$2i - i\mu t = 0 \quad ,$$

es decir, $\psi(t) = 2/t$.

De esto se obtiene una expresión para una sección σ :

$$\sigma(t) = 2/t$$

Es claro que $\psi(\sigma(t)) = t$ para toda $t \neq 0, \infty$. Como esto se cumple en un abierto de Zariski, concluimos que σ es una sección.

Se comprobará ahora que \mathfrak{S} es minimal.

Un cálculo explícito muestra que el conjunto discriminante de esta superficie es $\mu_{\text{crit}} = \mu_{\text{crit}}(\mathfrak{S}) = \{0, \pm 4, \infty\}$.

Analizando las matrices hessianas de las singularidades, se concluye que todas son nodales. Un cálculo directo muestra que los componentes de las fibras singulares son todos reducidos, por tanto la fibración es semi-estable.

Los tipos de fibras singulares son:

| | | | | |
|-------|-------|-------|-------|----------|
| μ | 0 | -4 | -4 | ∞ |
| tipo | I_4 | I_2 | I_2 | I_4 |

Al hacer blowup en un punto sobre una curva disminuye la autointersección de esa curva por 1.

Los 8 puntos base están distribuidos de esta forma: hay 2 puntos base en cada componente de la fibra $\mu = \infty$.

En $\mathbb{P}^1 \times \mathbb{P}^1$ cada componente C de la fibra $\mu = \infty$ tiene autointersección 0, luego en \mathfrak{S} cada uno tiene autointersección -2.

Los componentes de la fibra $\mu = 0$ son todas rectas, teniendo cada una dos puntos base en $\mathbb{P}^1 \times \mathbb{P}^1$, siendo cada una linealmente equivalente en $\mathbb{P}^1 \times \mathbb{P}^1$ a un componente de $\mu = \infty$, por tanto cada uno de los componentes tiene autointersección -2 en \mathfrak{S} .

Hay dos componentes irreducibles para cada una de las fibras $\mu = \pm 4$. Explícitamente, para la fibra sobre $\mu = 4$, las ecuaciones de las componentes en coordenadas locales x, y son

$$\begin{aligned} xy - 1 + i(x - y) &= 0 \\ xy - 1 - i(x - y) &= 0 \end{aligned}$$

Cada una de estas componentes toca 4 puntos base. En $\mathbb{P}^1 \times \mathbb{P}^1$, son curvas de bigrado (1,1), luego su autointersección es 2. De esta forma en \mathfrak{S} cada una de estas componentes tiene autointersección $2 - 4 = -2$.

Por tanto, la fibración es minimal.

De esta forma se tiene que el rango de Mordell-Weil de \mathfrak{S} es

$$r_{\mathfrak{S}} = \#\mu_{\text{crit}} - 4 = 4 - 4 = 0$$

Con lo que obtenemos

$$r_S \geq 0$$

Y también obtenemos una cota inferior para el número de Picard de S :

$$\rho_S \geq 22 - \#\mu_{\text{crit}} = 18$$

Como consecuencia de este cálculo, podemos ver que esta construcción particular puede indicar una posible forma de encontrar superficies K3 singulares (es decir, con $\rho_S = 20$), o más en general, una forma de construir superficies K3 en las que es de interés analizar la modularidad de su función L cohomológica.

Observación 37 *Es interesante notar que pueden usarse otras técnicas para afinar la cota del número de Picard de una superficie K3 obtenida de esta forma. En particular se pueden usar métodos computacionales para calcular la función Zeta y con ello tener una estimación del valor sobre \mathbb{C} , pueden usarse los métodos en campos finitos de [4] para encontrar una cota superior y puede luego obtenerse un refinamiento de esta cota superior mediante la técnica de [11].*

Referencias

- [1] D. Arapura, *Algebraic Geometry over the Complex Numbers*, 1 ed., Universitext, Springer-Verlag New York, 2012.
- [2] W. Barth, K. Hulek, C. Peters, and A. Van de Ven, *Compact complex surfaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete, no. 3. Folge, Bd. 4., Springer-Verlag Berlin, 2004.
- [3] V. Batyrev, *Dual Polyhedra and Mirror Symmetry for Calabi-Yau Hypersurfaces in Toric Varieties*, Journal of Algebraic Geometry (1994), no. 3, 493–535.
- [4] F. Charles, *On the picard number of k3 surfaces over number fields*, 2011.
- [5] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics, no. 52, Springer-Verlag, 1977.
- [6] J. Kollár and S. Mori, *Birational geometry of algebraic varieties*, Cambridge Tracts in Mathematics, no. 134, Cambridge University Press, 1998.

- [7] R. Miranda, *The basic theory of elliptic surfaces: Notes of lectures*, January 1989.
- [8] C. Schoen, *On fiber products of rational elliptic surfaces with section*, *Math. Z.* **197** (1988), no. 2, 177–199.
- [9] I. Shafarevich, *Basic Algebraic Geometry I*, 2 ed., Springer-Verlag Berlin Heidelberg, 1994.
- [10] J. Silverman, *The Arithmetic of Elliptic Curves*, 1 ed., Graduate Texts in Mathematics, vol. 106, Springer-Verlag New York, 1986.
- [11] R. van Luijk, *$K3$ surfaces with picard number one and infinitely many rational points*, *Algebra & Number Theory* **1** (2007), no. 1, 1–15.

UN MÉTODO FWI USANDO DEL MÉTODO DE OPTIMIZACIÓN DE CHEN

RESUMEN. Los métodos de inversión de forma de onda completa (FWI por sus siglas en inglés), son técnicas para la caracterización geotécnica del suelo y se basan en la minimización de una función de costo, para ello se emplean métodos de optimización, que pueden ser métodos basados en búsqueda (usualmente iterativos). En este trabajo se estudia la solución un problema prototípico de inversión mediante el uso del método de búsqueda en curva de Chen derivado del método de Shi y Shen para minimizar la función de costo de un problema FWI, éste es un método computacionalmente eficiente pues requiere poca memoria y no necesita el cálculo de matrices relacionadas con la matriz Hessiana; además reduce el número de evaluaciones de la función y de su gradiente. Además se comparan los resultados de utilizar este método con los obtenidos al utilizar el método de gradiente conjugado Fletcher–Reeves.

ABSTRACT. The full waveform inversion methods (FWI) are techniques for the geotechnical characterization of the soil and are based on the minimization of a cost function, to do this, optimization methods are used; which can be methods based on search (usually iterative). In this work the solution of a prototypical inversion problem is studied by using the Chen curve search method derived from the Shi and Shen method to minimize the cost function of an FWI problem, this is a computationally efficient method as it requires little memory and does not need the calculation of matrices related to the Hessian matrix; also reduce the number of evaluations of the function and its gradient. In addition, the results of using this method are compared with those obtained using the Fletcher–Reeves conjugate gradient method.

1. INTRODUCCIÓN

Los métodos de inversión de forma de onda completa (FWI por sus siglas en inglés) están dentro de las técnicas más recientes para la caracterización geotécnica de un sitio [10]. En dichos métodos se usan ondas sísmicas para evaluar el sitio en investigación, colocando vibradores sísmicos en la superficie; las ondas generadas por los vibradores se propagan a través del suelo, ocurriendo múltiples reflexiones y refracciones; la respuesta del sitio a las ondas enviadas, es medido por geófonos también colocados en la superficie de la tierra. Para realizar un perfilado basado en este método hacen falta dos componentes principales: un modelo computacional para la simulación de ondas en un dominio semi-infinito [7]; y un método de optimización a través del cual las respuestas calculadas son ajustadas a las respuestas medidas, mediante ajustes iterativos de un modelo inicial asumido. Desde una perspectiva matemática, un FWI es un problema de inversión no lineal que puede ser interpretado como un problema de optimización local o global [2].

Fecha: 19 de abril de 2021.

Palabras y frases clave. Waveform inversion – Inverse theory – Numerical modelling – Wave propagation.

Para resolver este problema de optimización, es posible utilizar métodos iterativos, comúnmente métodos basados en búsquedas o métodos basados en regiones de confianza, en [2] propone el uso de métodos AGO (optimización por gradiente adaptativo, por sus siglas en inglés) para la solución del problema de optimización derivado del método FWI.

En este trabajo se estudiará la aplicación, a la solución del problema de optimización derivado de resolver un problema FWI, de un método de optimización basado en búsqueda presentado por Chen en 2018 [4], y que es derivado del método de Shi y Shen de 2005. Estos son métodos de minimización utilizando una regla de búsqueda en curva [13], que se asemeja a la regla de búsqueda de Wolfe [15]. Además se analizará de forma comparativa con los resultados obtenidos utilizando el método de gradiente conjugado de Fletcher–Reeves (FB), con los resultados obtenidos con el método de Chen.

2. JUSTIFICACIÓN

Este trabajo está ubicado dentro de la línea de investigación de Optimización Numérica de la Orientación en Ingeniería Matemática de la Maestría en Matemáticas de la Universidad Nacional Autónoma de Honduras. Los estudios en el campo de la optimización resultan de mucha importancia en múltiples disciplinas, ya que siempre es de interés conocer los valores óptimos de algún fenómeno o problemas. Además, muchos problemas de interés pueden ser reescritos como problemas de optimización.

Los métodos de optimización tienen muchas aplicaciones en campos como la economía, ingeniería y de la física, por ejemplo, los inversores buscan crear portafolios que eviten riesgos excesivos consiguiendo las máximas tasas de retorno, los manufactureros buscan la máxima eficiencia en el diseño y operación de sus procesos de producción y los sistemas físicos tienden a estados con mínima energía [9].

Si bien es cierto existen muchos métodos que se han estudiado anteriormente en otros trabajos y que pueden ser aplicados para resolver estos problemas de optimización, resulta significativo resaltar la utilidad que tiene contar con métodos de optimización innovadores y eficientes, y validar su aplicación y eficiencia en la solución de problemas específicos, debido a que la eficiencia de un método puede variar de acuerdo al problema.

En este trabajo, se estudia la aplicación del método optimización con regla de búsqueda en curva de Chen, basado en el método de Shi y Shen [13], dado que éste es adecuado para resolver problemas a gran escala [4]. Se estudiará como este método puede aplicarse para la solución de problemas FWI.

Los métodos FWI, es un tema de estudio de vanguardia en el campo de la geofísica, y resultan de mucha utilidad práctica para el estudio de suelos en Ingeniería Civil, éstos métodos se utilizan para la caracterización geotécnica de los suelos [10]. Por tanto este trabajo contribuye al eje de investigación de desarrollo económico y social de UNAH, teniendo importantes aplicaciones en el tema de infraestructura y desarrollo territorial.

3. ANTECEDENTES

La inversión de forma de onda completa tiene como objetivo derivar un modelo mediante la minimización de la diferencia entre las formas de onda observadas y las

formas de ondas medidas, utilizando una norma, usualmente la L_2 -norma [2]. Por tanto este método puede verse como un problema de minimización sin restricciones donde la función objetivo es una función de ajuste basada en una norma.

Un *problema de minimización sin restricciones*, UP por sus siglas en inglés, se define de la siguiente forma:

$$x^* = \arg \min f(x), x \in \mathbb{V}^n \quad (3.1)$$

donde $f : \mathbb{V}^n \rightarrow \mathbb{R}$ es una función continuamente diferenciable y \mathbb{V}^n es un espacio vectorial.

La mayoría de los algoritmos iterativos más conocidos para la solución de UP, basados en estrategias de búsqueda, tienen la forma:

$$x_{k+1} = x_k + \alpha_k d_k \quad (3.2)$$

donde d_k es la dirección de búsqueda y α_k es un tamaño de paso positivo; por lo general, estos algoritmos iterativos dependen del gradiente de la función ∇f y de una matriz relacionada con la matriz Hessiana, usualmente una aproximación de la inversa de la misma. La implementaciones concretas de estos algoritmos depende de como se escojan d_k y α_k .

Sea f la función objetivo presentada en el problema de minimización (3.1), $x_k \in \mathbb{R}^n$ la aproximación de la solución de (3.1) en la k -ésima iteración y $x^* \in \mathbb{R}^n$ el valor donde la función f alcanza un mínimo, y por tanto una solución de (3.1). Por simplicidad, llamamos g_k a $\nabla f(x_k)$, el gradiente de la función objetivo evaluada en x_k ; f_k a $f(x_k)$, la función objetivo evaluada en x_k y f^* a $f(x^*)$, la función objetivo evaluada en x^* .

Para solucionar un UP con un algoritmo basado en búsquedas existen muchos métodos que implementan la ecuación (3.2). Estos son algunos métodos que consideraremos clásicos, pues son de pleno conocimiento en esta área.

Si la dirección de búsqueda se define como $d_k = -g_k$, el método se conoce como método de máximo descenso (*Steepest descent method*). Este método tiene una baja tasa de convergencia en muchos casos; sin embargo, no requiere el cálculo y almacenamiento de matrices relacionadas con la matriz Hessiana de la función objetivo.

Si la dirección de búsqueda se define como: $d_k = -H_k g_k$, donde H_k es la inversa de la matriz Hessiana o, usualmente, una matriz que la aproxima, el método se conoce como método de Newton (*Newton-like method*). A diferencia del método del máximo descenso, aquí si se necesita calcular y almacenar matrices relacionadas con la inversa de la matriz Hessiana. Sin embargo, este método tiene una razón de convergencia mucho más rápida que el método del máximo descenso, e incluso de otros métodos que se cometan más adelante.

Si tomamos d_k como:

$$d_k = \begin{cases} -g_k & \text{si } k = 1 \\ -g_k + \beta_k d_{k-1} & \text{si } k > 1, \end{cases}$$

se define una familia de métodos llamados métodos de gradiente conjugado, el parámetro β_k define cada uno de los diferentes métodos dentro de esta familia. Estos métodos resultan ser una herramienta apropiada para la solución de problemas a gran escala; ya que, al igual que el método del máximo descenso, evita el cálculo y almacenamiento de matrices. Sin embargo, muchos de estos métodos no tienen una convergencia global [13].

Por otro lado, tenemos métodos más recientes, que aplican nuevas técnicas para abordar la solución de estos problemas de optimización. Miele y Cantrell en 1969 [11] estudiaron el método de gradiente con memoria para la solución de UP; este método utiliza información de la iteración anterior para determinar la dirección de búsqueda. Cragg y Levy en 1969 [6], propusieron un método de supermemoria que es una generalización del método de Miele y Cantrell utilizando la información de varios pasos iterativos anteriores para determinar la dirección de búsqueda.

Este método también fue investigado por Wolfe y Viazminsky [14]. Tanto el método de gradiente con memoria como el método de gradiente con super memoria son más eficientes que el método de gradiente conjugado expuesto anteriormente. Sin embargo, hay que resaltar que hay muy pocos resultados sobre convergencia para funciones objetivos no cuadráticas para estos métodos.

Otra familia de métodos utilizados en esta área son los métodos de búsqueda en curva para solución de problemas sin restricción, como los métodos ODE [3] (o métodos dinámicos), en estos métodos se determina la dirección de búsqueda en cada iteración mediante la solución de una ecuación diferencial ordinaria. Dado que estos métodos utilizan un camino de búsqueda curvilíneo, que es más general que el camino de búsqueda lineal que se usa de forma tradicional [12]; es posible obtener mejores resultados de convergencia.

Cabe mencionar el artículo publicado por Ford [8] donde se estudia una nueva clase de método casi-newtoniano multi pasos, que ofrece excelentes resultados de convergencia; pero de igual forma requiere calcular y almacenar algunas matrices en cada iteración.

Esta rápida vista a otros métodos nos da un panorama sobre las características que son deseables para un nuevo método para la solución de UP mediante estrategias de búsqueda para problemas a gran escala. El método debe tener una tasa de convergencia alta, debe evitar el cálculo y almacenamiento de matrices relacionadas con la matriz Hessiana, y debe poseer una convergencia global.

El algoritmo presentado por Shi y Shen [13], cumple con todas estas características, si bien no tiene una convergencia mayor que la de los métodos newtonianos, su tasa de convergencia es lineal. Shi y Shen prueban en su artículo que dicho método tiene convergencia global. Además, el método, al igual que método del gradiente conjugado, solo necesita del gradiente de la función objetivo en cada paso iterativo, por lo que no es necesario el cálculo y almacenamiento de matrices asociadas a la matriz Hessiana.

Este es un método de búsqueda en curva; sin embargo, no es necesario resolver un sistema de ecuaciones diferenciales ordinarias en cada iteración; la búsqueda se hace a través de una curva dada. Sin embargo, en cada iteración, el siguiente punto iterativo se determina por una regla de búsqueda en curva que se asemeja a la regla de búsqueda lineal de Wolfe [13], lo que requiere que se hagan muchas evaluaciones de la función y de su gradiente para determinar el valor del tamaño de paso que cumple con la regla de búsqueda de Wolfe.

El método de Chen [4] parte del método de Shi y Shen; y generaliza algunas ideas presentadas en [16], con este nuevo método se presenta un versión más eficiente del algoritmo; donde el tamaño de paso y la dirección de búsqueda no se calculan simultáneamente en todas los pasos iterativos, esto representa una mejora considerable en el tiempo de ejecución del método. La regla modificada de Wolfe que se mencionó anteriormente en el método de Shi y Shen, solo se usa en el primer

paso iterativo, en los siguientes pasos se usa otra estrategia, con una fórmula basada en la dirección de búsqueda y el gradiente de la función, la dirección de búsqueda se sigue tomando en una curva, siendo función del tamaño de paso; sin embargo se utiliza el tamaño de paso de la iteración anterior para calcular la dirección de búsqueda, lo que implica también una mejora en tiempo de ejecución del algoritmo.

4. PROBLEMA FWI EN EL DOMINIO DEL TIEMPO

En esta sección se revisará de forma breve que es un problema FWI en el dominio del tiempo. Sea $X \in \mathbb{R}^n$, un vector que define las coordenadas espaciales, con $n \in \{1, 2, 3\}$ y $t \in [0, T]$ una coordenada temporal; además sea $\mathbb{M} = \{m : \mathbb{R}^n \rightarrow \mathbb{R}^k\}$ una familia de funciones que representarán parámetros físicos dependientes de la posición (velocidad, densidad, etc). Definiremos $V_{obs}(X_g, t|X_s)$ y $V_{apx}(X_g, t|X_s)$ como los datos observados y aproximados respectivamente; donde los datos observados se obtiene utilizando fuentes de ondas sísmicas colocadas en las coordenadas $X_s \subseteq X$; registrándolas con geófonos colocados en las coordenadas $X_g \subseteq X$ en el tiempo t . Los datos aproximados son generados mediante el modelo directo que depende del vector de parámetros físicos, que en general depende de la posición, $\vec{m}(X) = [m_1, \dots, m_p]^T$; donde $m_i \in \mathbb{M}$; además, se define $V_{res}(X_g, t|X_s)$ como $V_{apx}(X_g, t|X_s) - V_{obs}(X_g, t|X_s)$ [2].

El objetivo del método FWI es minimizar una función de ajuste, generalmente basada en la L_2 -norma, de la diferencia entre los datos observados y los datos aproximados, usando un criterio de ajuste por mínimos cuadrados. Por lo que podemos definir una función de coste $\epsilon : \mathbb{M}^n \rightarrow \mathbb{R}$ como:

$$\epsilon(\vec{m}) = \frac{1}{2} \sum_{s=1}^{N_s} \sum_{g=1}^{N_g} \int_0^T \|V_{res}(X_g, t|X_s)\|^2. \quad (4.1)$$

Por tanto el problema de minimización a resolver de forma iterativa será:

$$\vec{m}^* = \arg \min_{\vec{m} \in \mathbb{M}^p} \epsilon(\vec{m}). \quad (4.2)$$

Para resolver el este problema de forma iterativa, se selecciona un modelo inicial para el vector de parámetros físicos $\vec{m}_0(X)$ y se emplea un proceso de búsqueda iterativo similar al de la ecuación (3.2):

$$m_{k+1} = m_k + \alpha_k d_k. \quad (4.3)$$

Problema prototípico: Con el objetivo de observar el comportamiento del nuevo algoritmo de optimización de Chen, se plantea el siguiente problema de inversión simplificado:

Sea $L, T \in \mathbb{R}$, la longitud y el tiempo. $x \in [0, L]$ y $t \in [0, T]$; además $c : [0, L] \rightarrow \mathbb{R}$, $U : [0, L] \times [0, T] \rightarrow \mathbb{R}$ y $F : [0, L] \times [0, T] \rightarrow \mathbb{R}$. Tenemos el siguiente modelo que representa una onda unidimensional, con condiciones de borde nulas:

$$c(x)^2 \frac{\partial^2 U(x, t)}{\partial x^2} - \frac{\partial^2 U(x, t)}{\partial t^2} = F(x, t) \quad (4.4)$$

$$\begin{aligned} U(0, t) &= U(L, t) = 0 \\ U(t, 0) &= U_t(t, 0) = 0, \end{aligned}$$

este puede ser interpretado como una cuerda fija en los extremos que inicialmente no está moviéndose y que es sometida a un forzamiento externo que la hará vibrar.

Notemos que que dada una función $c(x)$, es posible encontrar U correspondiente a $c(x)$, utilizando (4.4), por tanto diremos que U es función de c y lo representamos como $U(c)$. Dada un data medida experimentalmente, que llamaremos U_m , siendo esta una constante respecto de c , se desea encontrar $c^*(x)$, de tal forma que al utilizarlo para resolver (4.4) reproduzca los datos experimentales. Esto se hará mediante el siguiente problema de optimización:

$$c^* = \arg \min \frac{1}{2} \int_0^T \|\{U(x, t)\}(c) - U_m(x, t)\|^2 dt.$$

En general $c(x)$ es función de x , sin embargo en este problema consideramos que c no varía con x . Para resolver el problema de optimización necesitamos encontrar el gradiente de $\epsilon = \frac{1}{2} \|\{U(x, t)\}(c) - U_m(x, t)\|^2$ como función de c :

$$\nabla \epsilon = \frac{\partial \epsilon}{\partial c} = \frac{\partial U}{\partial c} U.$$

Utilizamos (4.4) para derivar una ecuación para el gradiente de U :

$$\begin{aligned} \frac{\partial}{\partial c} \left(c^2 \frac{\partial^2 U}{\partial x^2} \right) - \frac{\partial}{\partial c} \left(\frac{\partial^2 U}{\partial t^2} \right) &= \frac{\partial}{\partial c} (F) \\ 2c \frac{\partial^2 U}{\partial x^2} + c^2 \frac{\partial}{\partial c} \left(\frac{\partial^2 U}{\partial x^2} \right) - \frac{\partial}{\partial c} \left(\frac{\partial^2 U}{\partial t^2} \right) &= 0 \\ c^2 \frac{\partial^2 \left(\frac{\partial U}{\partial c} \right)}{\partial x^2} - \frac{\partial^2 \left(\frac{\partial U}{\partial c} \right)}{\partial t^2} &= -2c \frac{\partial^2 U}{\partial x^2} \end{aligned} \quad (4.5)$$

Para encontrar c^* , damos una función c_0 , usualmente una función constante; luego para aplicar nuestro método de optimización, necesitaremos evaluar la función objetivo y su gradiente; hacer dichas evaluaciones corresponderá a resolver las ecuaciones diferenciales (4.4) y (4.5) en cada pasa iterativo, se utilizará el método de diferencias finitas para dicho propósito.

5. DESCRIPCIÓN DEL MÉTODO DE CHEN

El método de Chen [4], es una extensión del método de Shi y Shen [13]; estos métodos son innovadores en su tipo; sin embargo están motivados por otros métodos ya existentes, la principal característica de éste, es su regla de búsqueda en curva, la cual es una derivación de la *regla de búsqueda de Wolfe* [15] propuesta por Philip Wolfe en 1969, conocida también como *condición de Wolfe*. La regla de búsqueda de Wolfe se compone de dos partes:

La *regla de Armijo* [1]:

$$f(x_k + \alpha_k d_k) - f_k \leq \mu_1 g_k^T d_k. \quad (5.1)$$

Y la *condición de curvatura*:

$$g(x_k + \alpha_k d_k)^T d_k \geq \mu_2 g_k^T d_k \quad (5.2)$$

donde $d_k \in \mathbb{R}^n$ es la dirección de búsqueda, $\alpha_k \in \mathbb{R}^+$ es el tamaño de paso y $0 < \mu_1 < \mu_2 < 1$ son dos constantes que se pueden adecuar de acuerdo al problema estudiado. El objetivo de esta regla es delimitar el valor de α_k que se debe utilizar en cada iteración de tal forma que se obtenga el mejor avance posible hacia la solución del problema. Más adelante se mostrará la versión de esta regla presentada por Shi y Shen en su método y luego utilizada por Chen en su versión avanzada del algoritmo.

Para garantizar la convergencia del método de Shi y Shen se asumen las siguientes hipótesis:

Hipótesis 1. La función tiene una cota inferior en:

$$L_0 = \{x \in \mathbb{R}^n \mid f(x) \leq f(x_0)\} \quad (5.3)$$

Hipótesis 2. El gradiente g es Lipschitz continuo en un conjunto convexo abierto B , que contiene a L_0 . Es decir, existe un $L > 0$ tal que:

$$\|g(x) - g(y)\| \leq L\|x - y\|, \quad \forall x, y \in B. \quad (5.4)$$

Las hipótesis del algoritmo de Chen son las mismas que las del algoritmo de Shi y Shen, presentadas en (5.3) y (5.4).

Hay muchas formas de escoger el tamaño de paso α , como el expuesto por Cohen en 1981 [5]; Shi y Shen en [13] proponen la siguiente regla de búsqueda, basadas en la regla de Wolfe (5.1) y (5.2):

En cada iteración, fijado $s_k > 0$, el tamaño de paso α debe satisfacer:

$$f(x_k + \alpha d_k(\alpha)) - f_k \leq \mu_1 \alpha d_k g_k^T d_k(\alpha) \quad (5.5)$$

$$g(x_k + \alpha d_k(\alpha))^T d_k(\alpha) \geq \mu_2 g_k^T d_k(\alpha) \quad (5.6)$$

Donde $\mu_1 < \frac{1}{2} < \mu_2 < 1$ y:

$$d_k(\alpha) = - \left(1 - \frac{\alpha s_k}{1 + \alpha} \right) g_k - \frac{\alpha s_k}{1 + \alpha} d_{k-1}. \quad (5.7)$$

El algoritmo de Chen, utiliza esta regla solo en la primera iteración, luego utiliza un fórmula basada en el en la dirección de búsqueda, el valor de gradiente y la constante L presentada en (5.4). A continuación presentamos el algoritmo de Chen.

Algoritmo de Chen: Dado un punto inicial x_1 , una constante positiva ϵ y tres parámetros μ_1, μ_2 y ρ tales que: $0 < \mu_1 < \frac{1}{2} < \mu_2 < 1, \frac{1}{2} < \rho < 1$. Definimos $k := 1$.

Paso 1: Si $\|\nabla f(x_k)\| < \epsilon$ parar. Sino ir al paso 2.

Paso 2: Se define:

$$s_k = \begin{cases} \rho & \text{si } k = 1 \\ \frac{\rho \|\nabla f(x_k)\|^2}{\rho \|\nabla f(x_k)\|^2 + (1-\rho) |\nabla f(x_k)^T d_{k-1}|} & \text{si } k \geq 2 \end{cases}$$

Paso 3: Calcular la dirección de búsqueda:

$$d_k(\alpha_k) = \begin{cases} -s_k \nabla f(x_k) & \text{si } k = 1 \\ -\rho \left(1 - \frac{\alpha_{k-1} s_k}{1 + \alpha_{k-1}}\right) \nabla f(x_k) - (1 - \rho) \frac{\alpha_{k-1} s_k}{1 + \alpha_{k-1}} d_{k-1} & \text{si } k \geq 2 \end{cases}$$

Paso 4:

- Si $k = 1$ calcular α_k usando la regla de Wolfe, las desigualdades 5.5 y 5.6
- Si $k > 1$ calcular α_k usando la siguiente relación:

$$\alpha_k = -\frac{\nabla f(x_k)^T d_k}{L_k \|d_k\|^2}$$

Donde: $\rho L \leq L_k \leq m_k L$ y $\{m_k\}$ es una secuencia positiva con una cota superior suficientemente grande y L es la constante requerida en 5.4.

Paso 5: Calcular $x_{k+1} = x_k + \alpha_k d_k$

Paso 6: $k := k + 1$, ir al paso 1

Por tanto se propone aplicar el algoritmo de Chen, como de describió en esta sección, para resolver el problema prototípico de inversión propuesto en la sección 4.

6. EXPERIMENTOS NUMÉRICOS

6.1. Comparación de desempeño de algoritmos de minimización basados en búsqueda. Inicialmente se hará una comparación de desempeño del método de Chen, para ello se usará el siguiente problema de minimización comparando el rendimiento del algoritmo de Chen, con el algoritmo de Shi y Shen y con un algoritmo de gradiente conjugado de tipo FR:

$$x^* = \arg \min_{x \in \mathbb{R}^n} \frac{1}{2} \|Ax - b\|_2^2$$

para $A \in \mathbb{R}^{m \times n}$ y $b \in \mathbb{R}^m$ dados.

Para comparar el rendimiento de los algoritmos, se realizaron 50 experimentos, donde la matriz A de 30 por 50, se definió de la siguiente forma:

$$A = R + I$$

donde R es una matriz con entradas aleatorias normalizadas de 30 por 50 e I es la matriz identidad conformarble para la suma. Además, se escogió de forma aleatoria los vectores b e x_0 para cada experimento, el vector x_0 representa el vector inicial en el proceso iterativo.

Se obtuvieron los resultados que se muestran en las figuras 1 y 2; donde se observa que el método de Chen no es superior a ninguno de los otros dos métodos el términos de la cantidad de iteraciones a realizar, sin embargo debido a que las cantidad de evaluaciones de la función objetivo y su gradiente es mucho menor, si

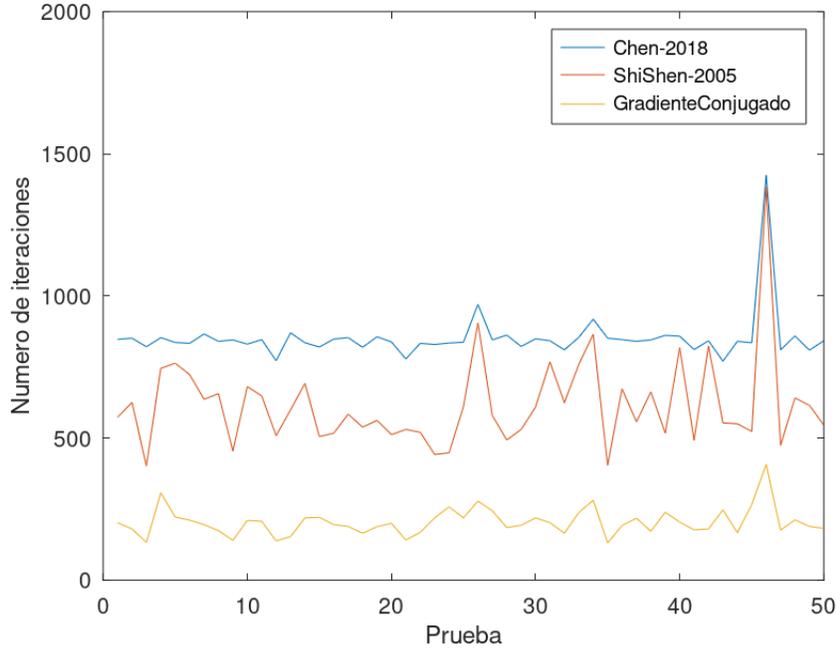


FIGURE 1. Comparación de iteraciones

se observa una ventaja significativa en términos del tiempo de ejecución.

6.2. Solución de problema prototípico con una fuente. En el siguiente experimento numérico se aplicó el algoritmo de Gradiente Conjugado y el algoritmo de Chen para la solución del problema de inversión prototípico propuesto en la sección 4. Se consideró una cuerda de longitud 1, vibrando durante 10 unidades de tiempo; además un valor de $c = 0.55$ y una función de forzamiento $Q(0, t) = 500\sin(x)$, se generaron datos de prueba $U_m(x, t)$ que simularan los datos medidos en problema real. Luego, se procedió a resolver el problema de inversión; que consiste en recuperar el valor de c a partir de los datos U_m .

En la sección anterior se observó que el método de Chen tenía un mejor desempeño que el método de gradiente conjugado FR para el problema de prueba utilizado; aquí se pretende comparar ambos métodos en la solución de un problema de inversión.

Al aplicar el método Chen, se utilizaron los siguiente parámetros $\mu_1 = 0.38$, $\mu_2 = 0.85$, $\rho = 0.58$ y $L = 10$. Al aplicar el método de gradiente conjugado FR se utilizan los siguiente parámetros $\mu_1 = 0.38$, $\mu_2 = 0.85$. Además, en cada iteración fue necesario resolver la ecuación (4.4); para ello se utilizó el método de diferencia finitas de Crank-Nicolson, utilizando una partición en 200 unidades para el espacio y en 100 unidades para el tiempo.

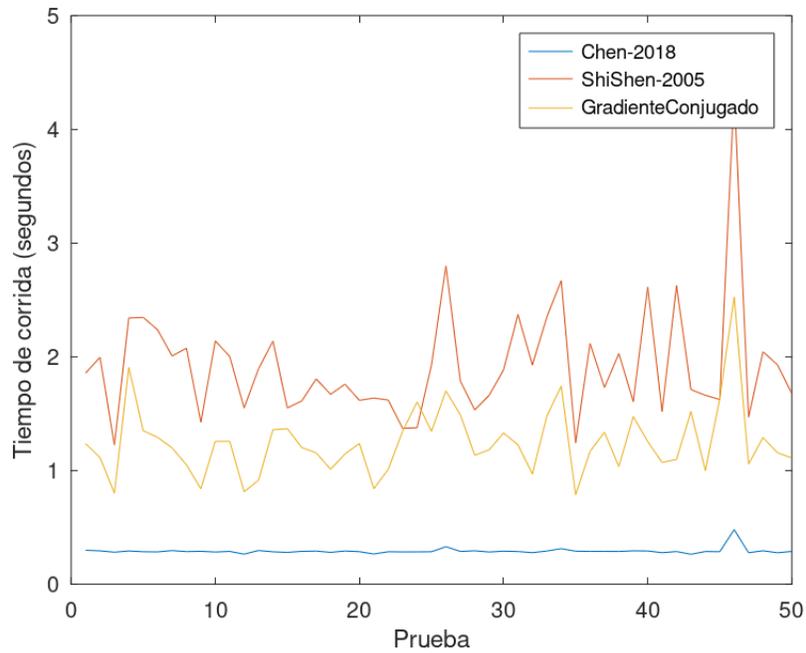


FIGURE 2. Comparación de tiempo de ejecución

Los resultados se obtienen ejecutando el siguiente código de Octave, donde se ejecuta el archivo **SingleTest**, disponible en código fuente de este trabajo.

```

1 > SingleTest
2 Getting matrix:
3 Elapsed time is 0.01438 seconds.
4 Getting testing data:
5 Elapsed time is 0.00756001 seconds.
6 Solving inversion problem with Chen:
7 Elapsed time is 5.92566 seconds.
8 Error on recovered data:
9 ans = 7.1683e-12
10 Iterations:
11 k = 37
12 Solving inversion problem with GC:
13 Elapsed time is 65.0365 seconds.
14 Error on recovered data:
15 ans = 6.9522e-12
16 Iterations:
17 k = 33

```

Los resultados obtenidos fueron los siguientes:

TABLE 1. Comparación del método de Chen con el método de gradiente conjugado

| | Gradiente conjugado | Chen |
|--------------------------|---------------------|-------------|
| Iteraciones | 33 | 37 |
| Tiempo de ejecución | 64.1226 seg | 5.71083 seg |
| Error en el valor de c | 6.9522e-12 | 7.16e-12 |

TABLE 2. Ejecución del método de Chen en paralelo

| | |
|-------------------------|------------|
| Tiempo de ejecución | 31.5808 |
| Error en valores de c | 2.0911e-11 |

6.3. Solución de problema prototípico con múltiples fuentes. Se realizó además otro experimento numérico, donde se explora la solución de N problemas prototípicos en paralelo; dado que una extensión del problema prototípico es que en lugar de una cuerda, se tenga una superficie que se está haciendo vibrar en N puntos diferentes; en cuyo caso es necesario resolver N problemas como los presentados en en la sección anterior.

Para ellos se aplican técnicas de computo de alto rendimiento, mediante el uso del paquete **Parallel** de Octave. En este caso, se considera un problema con 10 fuentes aplicadas en 10 puntos diferentes. En cada punto de tienen los siguientes valores para c : 0.55, 0.78, 0.82, 0.61, 0.79, 0.63, 0.7, 0.8, 0.9 y 0.6. Se utiliza la misma función de forzamiento, distancia y tiempo de medición que en la sección anterior para simular los datos para cada uno de los valores de c .

Además, se utiliza el método de optimización de Chen con los mismos parámetros que en la sección anterior. Los resultados se obtienen ejecutando el siguiente código de Octave, donde se ejecuta el archivo **MultiTest**, disponible en código fuente de este trabajo.

```

1 > MultiTest
2 Getting testing data..
3 Solving invertion problem:
4 Elapsed time is 31.5808 seconds.
5 ans = 2.0911e-11

```

Los resultado obtenidos fueron los siguientes:

Notemos que si resolvemos de forma secuencial los 10 problemas, utilizando los resultados de la sección anterior, esperaríamos un tiempo de solución de 57.1 segundos.

7. DISPONIBILIDAD DE DATOS

El código Octave desarrollado para este trabajo esta disponible por consulta vía correo electrónico a la dirección victor@unah.hn

8. CONCLUSIONES

- Para resolver un problema de optimización por búsqueda, utilizando la regla de Wolfe, es necesario evaluar varias veces la función objetivo y el gradiente de la misma en cada iteración.
- El método de Chen solo utiliza la regla de Wolfe en la primera iteración, lo que reduce la cantidad de evaluaciones necesarias; dado que en un método de inversión evaluar la función objetivo y su gradiente implica resolver una ecuación diferencial en cada caso; el método de Chen reduce mucho la cantidad de evaluaciones necesarias, por lo que resulta más conveniente para resolver este tipo de problemas.
- Cuando el problema de inversión se generaliza a un problema con varias fuentes, resulta conveniente aplicar técnicas de programación de alto rendimiento para resolverlo, ya que el problema se puede paralelizar. Según los experimentos numéricos realizados el tiempo de ejecución se reduce a la mitad cuando el problema se paraleliza.

9. TRABAJOS FUTUROS

En este trabajo solo se estudió un problema prototípico, para una dimensión, considerando un único parámetro que no varía con la posición. Además, se estudió la extensión del problema prototípico cuando se tienen varias fuentes. En trabajos futuros se estudiará el problema de inversión en dos dimensiones, con parámetros variables con la posición y con varias fuentes. Además se compararán los resultados que se obtengan, con los que se obtendrían utilizando un método de gradiente adaptativo.

REFERENCES

1. Armijo, L. (1966). Minimization of functions having Lipschitz continuous first partial derivatives. *Pacific Journal of mathematics* 16, 1–3.
2. Bernal-Romero, M., and Iturrarán-Viveros, U. (2021). Accelerating full-waveform inversion through adaptive gradient optimization methods and dynamic simultaneous sources. *Geophysical Journal International* 225, 97–126.
3. Botsaris, C. A. (1978). Differential gradient methods. *Journal of Mathematical Analysis and Applications* 63, 177–198.
4. Chen, C., Luo, L., Han, C., and Chen, Y. (2018). Global convergence of an extended descent algorithm without line search for unconstrained optimization. *Journal of Applied Mathematics and Physics* 6, 130–137.
5. Cohen, A. (1981). Stepsize analysis for descent methods. *Journal of Optimization Theory and Applications* 33, 187–205.
6. Cragg, E., and Levy, A. (1969). Study on a supermemory gradient method for the minimization of functions. *Journal of Optimization Theory and Applications* 4, 191–205.
7. Fathi, A., Poursartip, B., and Kallivokas, L. F. (2015). Time-domain hybrid formulations for wave simulations in three-dimensional PML-truncated heterogeneous media. *International Journal for Numerical Methods in Engineering* 101, 165–198.

8. Ford, J., and Tharmlikit, S. (2003). New implicit updates in multi-step quasi-Newton methods for unconstrained optimisation. *Journal of Computational and Applied Mathematics* 152, 133–146.
9. Jorge, N., and Stephen, J. W. (1999). Numerical optimization. *Springerverlang, USA*.
10. Kallivokas, L., Fathi, A., Kucukcoban, S., Stokoe II, K., Bielak, J., and Ghattas, O. (2013). Site characterization using full waveform inversion. *Soil Dynamics and Earthquake Engineering* 47, 62–82.
11. Miele, A., and Cantrell, J. (1969). Study on a memory gradient method for the minimization of functions. *Journal of Optimization Theory and Applications* 3, 459–470.
12. Schropp, J. (1997). A note on minimization problems and multistep methods. *Numerische Mathematik* 78, 87–101.
13. Shi, Z.-J., and Shen, J. (2005). A new descent algorithm with curve search rule. *Applied mathematics and computation* 161, 753–768.
14. Wolfe, M., and Viazminsky, C. (1976). Supermemory descent methods for unconstrained minimization. *Journal of Optimization Theory and Applications* 18, 455–468.
15. Wolfe, P. (1969). Convergence conditions for ascent methods. *SIAM review* 11, 226–235.
16. Zhou, G. (2009). A descent algorithm without line search for unconstrained optimization. *Applied Mathematics and Computation* 215, 2528–2533.

DEPARTAMENTO DE MATEMÁTICAS, FACULTAD DE CIENCIAS, UNIVERSIDAD NACIONAL AUTÓNOMA DE HONDURAS

Dirección de correo electrónico: victor@unah.hn

IDENTIFICACIÓN DE MODELOS ESPARCIDOS EN DINÁMICA ECONÓMICA DE ALTA DIMENSIONALIDAD

KERIN CARDONA

RESUMEN. En este documento, se presentan aplicaciones de algunas técnicas teóricas y computacionales, para la aproximación estructurada de sistemas dinámicos basados en datos económicos. La investigación realizada en este artículo está enfocada a modelos de alta dimensionalidad con aplicaciones en economía y finanzas. Las técnicas antes mencionadas se comparan, por medio de algunas simulaciones numéricas, con los métodos más recientes y efectivos en el área de modelación esparcida de alta dimensionalidad.

1. INTRODUCCIÓN

La construcción de modelos basados en datos es una pieza fundamental en aplicaciones en ciencias e ingeniería. La identificación de sistemas representa un conjunto de técnicas desarrolladas para construir modelos matemáticos de sistemas dinámicos a partir de datos observados.

Es de nuestro interés modelos de alta dimensionalidad, los cuales han ganado una importante relevancia en muchas áreas de la economía [5], finanzas [6] y aprendizaje automático [16]. Por ejemplo los modelos autorregresivos vectoriales (VAR) son considerados la clave para el análisis de información estructurada en la evolución conjunta de series de tiempo macroeconómicas.

Los modelos VAR estándar no suelen incluir más de diez variables [3], sin embargo es de interés para los econométricos poder observar cientos de series de datos. Dado que el número de parámetros crece de manera cuadrática con respecto al tamaño del modelo, fue necesario extender los modelos VAR a modelos de factores estimados (FAVAR) [17]. Bernanke utilizó modelos FAVAR en [4], para estimar los efectos de las políticas monetarias.

Algunos ejemplos de aplicación de modelos de alta dimensionalidad son los datos de tipo panel a gran escala [18], matrices de volatilidad [10], riesgo y construcción de portafolio y modelos de riesgo crediticio. En general los problemas de alta dimensión proponen un gran reto en la estimación de parámetros, debido a que pequeños errores en la estimación de los elementos dan como resultado grandes errores en las matrices.

Fecha: May 3, 2021 .

Palabras y frases clave. modelos esparcidos, identificación de sistemas, matrices estructuradas.

Los modelos esparcidos representan una de las mejores alternativas para tratar problemas de alta dimensión [15]. La idea principal es asumir que el vector de parámetros contiene muchas componentes exactamente cero o lo suficientemente pequeñas. Dicha suposición es de gran importancia para lograr la identificación del verdadero modelo subyacente, especialmente si solo se cuenta con una muestra pequeña.

El propósito de este documento es aplicar algunas técnicas teóricas y computacionales presentadas en [31], para la aproximación de sistemas basados en datos económicos de alta dimensionalidad. Nuestro enfoque es implementar los métodos y algoritmos presentados en [30] y [31] que nos permitan identificar, basados en datos, los modelos que posteriormente pueden ser usados para un análisis detallado de la estructura y dinámica del fenómeno.

2. JUSTIFICACIÓN

Extraer información de datos económicos, tomando en cuenta características especiales de dichos datos y la información a priori de la teoría económica, es uno de los objetivos principales que persigue la econometría [14, 12]. El estudio de series temporales y ecuaciones en diferencias se han convertido en una fuente de inspiración para el desarrollo de la estadística en los últimos años.

Debido al gran avance en las ciencias computacionales, los datos de alta dimensión han surgido en muchos campos de la ciencia, ingeniería, humanidades y la economía. Diversos temas como ser comercio electrónico, mercadeo, genética, imágenes biomédicas y muchos más, proporcionan datos a gran escala que requieren de nuevas técnicas teóricas y computacionales para un tratamiento más efectivo.

Los sistemas dinámicos aplicados a los mercados financieros actualmente reciben mucha atención a medida que la información de los mercados se vuelve disponible [10]. Construir modelos de los mercados financieros es un problema inverso, dado que no se cuenta con principios fundamentales, como en el caso de la física e ingeniería.

El tema presentado en este documento se enmarca en el eje prioritario de investigación de la Universidad Nacional Autónoma de Honduras (UNAH): población y condiciones de vida, dentro del tema cultura, ciencia y educación. Además la Modelación Matemática es una de las líneas de investigación de la Maestría en Matemática con orientación en Ingeniería Matemática, por lo cual este proyecto contribuye a la solución de problemas de gran relevancia para el país.

3. ANTECEDENTES

La econometría ha inspirado el desarrollo de modelos de series de tiempo y ecuaciones en diferencias. Por ejemplo los primeros trabajos en series de tiempo en [34, 32], el desarrollo de una teoría asintótica para el estimador de mínimos

cuadrados (LS) para ecuaciones lineales estocásticas en diferencias (sistemas AR) presentada en [21] y su posterior extensión a sistemas simultáneos, donde LS no es consistente, desarrollada en [19].

Importantes avances en los modelos ARMAX fueron propuestos en [9] y posteriormente en [11]. Adicionalmente los modelos de errores en variables, que suceden cuando hay alteraciones en las medidas de entrada y salida tienen un origen en el análisis econométrico.

Los vectores autorregresivos VAR, introducidos por Christopher Sims en [26], definieron un nuevo marco macroeconómico muy prometedor. Un modelo VAR es un modelo lineal de n -ecuaciones y n -variables donde el valor actual de una variable es proporcionado por sus valores anteriores más los valores pasados de las $n - 1$ variables restantes. Este sencillo proceso permitió una mejor visión de la dinámica en series de tiempo y además las herramientas estadísticas que conllevan los VARs son fáciles de usar e interpretar.

Bernanke, Boivin y Eliasz investigaron en [4] soluciones a problemas que surgieron en la implementación de los modelos VAR. Estos problemas se deben a la información limitada que generalmente se utiliza en estos modelos empíricos. La metodología implementada FAVAR, es decir, VAR con factor aumentado, fue aplicada con éxito en la identificación de los efectos de política monetaria en la economía.

Durante la última década los métodos basados en datos se han desarrollado de gran manera, esto debido a la gran cantidad de datos y evolución de técnicas analíticas. En [24] se estudian modelos cuadráticos esparcidos basados en polinomios, en los cuales la dinámica depende de un número reducido de variables e interacciones de segundo orden. En particular, en dicho trabajo se detallan estrategias de muestreo que conllevan a la identificación exacta de sistemas dinámicos de primer orden, aun cuando solo se cuenta con un número reducido de muestras.

En [25] se presentó una aplicación del algoritmo iterativo de Douglas-Rachford como una opción eficiente para resolver el problema de mínimos cuadrados penalizado que surge en [24]. Para un tratamiento completo del método de Douglas-Rachford revisar, por ejemplo [1].

En el caso del modelo de regresión lineal

$$(3.1) \quad y = X\beta + \varepsilon,$$

donde $y = (y_1, \dots, y_n)^T$ es un vector n -dimensional de respuesta, $X = (x_1, \dots, x_n)^T$ es una matriz de $n \times p$, y ε es un vector de ruido n -dimensional. Consideremos el caso de un modelo canónico lineal con matriz ortogonal reescalada $X^T X = nI_p$. El problema de mínimos cuadrados penalizado (PLS) es

$$(3.2) \quad \min_{\beta \in \mathbb{R}^p} \left\{ \frac{1}{2n} \|y - X\beta\|_2^2 + \sum_{j=1}^p p_\lambda(|\beta_j|) \right\},$$

donde $\|\cdot\|_2$ denota la norma en L_2 y p_λ es una función de penalización para $\lambda \geq 0$.

Con este proceso de regularización del método convencional de mínimos cuadrados, se espera una mejor estimación de los coeficientes de regresión con estimadores esparcidos.

4. MODELOS AUTORREGRESIVOS

En esta sección se presentan algunas técnicas utilizadas para modelar y pronosticar series temporales. El pronóstico de series temporales es una de las áreas más importantes en el estudio de variables económicas [35], donde observaciones pasadas de una o varias variables son recolectadas y analizadas para desarrollar modelos que describan una relación subyacente. Los modelos autorregresivos son ampliamente utilizados en muchos campos de las ciencias e ingeniería, por ejemplo en la predicción de precios de acciones financieras [7], en la predicción de la dinámica de la vegetación utilizando datos satelitales remotos [23], entre otros.

Un *proceso autorregresivo* de orden p , el cual en lo sucesivo denotaremos por $AR(p)$, es un proceso donde el valor de una serie de tiempo en el tiempo t , y_t , es la suma ponderada de los p valores pasados más un término de ruido denotado por ε_t . El proceso lo podemos representar por la siguiente ecuación en diferencias

$$(4.1) \quad y_t = a_1 y_{t-1} + a_2 y_{t-2} + \dots + a_p y_{t-p} + \varepsilon_t,$$

donde $\varepsilon_t \sim N(0, \sigma^2)$. Si utilizamos el operador *lag*, denotado por L , el cual aplicado a la variable en el tiempo t se obtiene el valor de la variable en el tiempo $t - 1$, esto es $Ly_t = y_{t-1}$, el proceso (4.1) lo podemos expresar de la forma

$$(1 - a_1 L - a_2 L^2 - \dots - a_p L^p) y_t = \varepsilon_t$$

o de forma compacta

$$(4.2) \quad a(L) y_t = \varepsilon_t.$$

donde el polinomio autorregresivo $a(L)$ está definido por $a(L) = 1 - a_1 L - a_2 L^2 - \dots - a_p L^p$. Para estimar los coeficientes $a_i, i = 1, \dots, n$, podemos utilizar el método de mínimos cuadrados haciendo el cuadrado del error alcanzar su mínimo [33]

$$\begin{aligned} A &= \arg \min_{\hat{A} \in \mathbf{R}^{1 \times p}} \|Y - \hat{A}X\|_F^2 \\ &= (X^T X)^{-1} X^T Y, \end{aligned}$$

donde

$$X = \begin{bmatrix} y(p) & y(p-1) & \cdots & y(1) \\ y(p+1) & y(p) & \cdots & y(2) \\ \vdots & \vdots & \vdots & \vdots \\ y(N-1) & y(N-2) & \cdots & y(N-p) \end{bmatrix}, Y = \begin{bmatrix} y(p+1) \\ y(p+2) \\ \vdots \\ y(N) \end{bmatrix},$$

N representa el tamaño de la serie de tiempo.

Un proceso de *medias móviles* de orden p , el cual de forma abreviada denotaremos por $MA(p)$, es un proceso de sumas ponderadas de q elementos autorregresivos de disturbio más un término contemporáneo de disturbio, de la forma

$$(4.3) \quad y(t) = b_0 \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q},$$

lo cual se puede representar en forma abreviada

$$(4.4) \quad y_t = b(L) \varepsilon_t,$$

donde $\varepsilon_t \sim N(0, \sigma^2)$, b_0, \dots, b_q son constantes y $b(L) = b_0 + b_1 L + \dots + b_q L^q$ representa el polinomio de medias móviles de grado q con $b_0 \neq 0$. Los procesos AR y MA presentados anteriormente se pueden unir para generar el proceso *autorregresivo de medias móviles*, el cual denotaremos por $ARMA(p, q)$ y estará representado por

$$(4.5) \quad y_t = a_1 y_{t-1} + a_2 y_{t-2} + \dots + a_p y_{t-p} + \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

o de forma compacta

$$(4.6) \quad a(L)y_t = b(L)\varepsilon_t.$$

Es posible identificar un modelo $ARMA(p, q)$ utilizando una técnica autorregresiva de dos pasos. El método funciona de la siguiente manera:

1. Se inicia realizando una identificación de un modelo autorregresivo de orden m para y_t , es decir encontramos $\tilde{y}_t = \sum_{j=1}^m A'_j y_{t-j}$. Luego definimos \hat{u}_t de la forma

$$\hat{u}_t = y_t - \tilde{y}_t,$$

lo cual representa una estimación de los residuos u_t .

2. Utilizando la estimación encontrada para \hat{u}_t se procede a realizar una identificación del modelo autorregresivo para y_t de la forma

$$y_t = \sum_{j=1}^p A_j y_{t-j} + \sum_{j=1}^q A'_j \hat{u}_{t-j}.$$

Los procesos ARMA pueden dar un modelo reducido con relativamente menos parámetros desconocidos. Nos dan una alternativa a los modelos AR y MA de alto orden utilizando una combinación de ambos procesos. Si se incluyen variables determinadas de forma exógena, esto es que su valor está determinado por factores externos al modelo, se conoce como proceso ARMA con variables exógenas y lo denotaremos por ARMAX. Si utilizamos la variable x_t para la variable exógena, el proceso ARMAX lo podemos representar de la forma

$$(4.7) \quad a(L)y_t = b(L)\varepsilon_t + g(L)x_t.$$

El grado del polinomio $g(L)$ nos proporciona el número de elementos pasados de la variable exógena x_t , que influyen en la variable endógena y_t , es decir que su valor está determinado por factores dentro del modelo. Una extensión multivariada del modelo AR es el proceso autorregresivo vectorial VAR, el cual tiene la siguiente representación general

$$(4.8) \quad \mathbf{x}_t = \mathbf{A}_1 \mathbf{x}_{t-1} + \mathbf{A}_2 \mathbf{x}_{t-2} + \dots + \mathbf{A}_p \mathbf{x}_{t-p} + \mathbf{s}_t + \varepsilon_t,$$

donde $\mathbf{x}_t = (x_{1,t}, \dots, x_{n,t})^T$ es una serie de tiempo multivariada; A_i , $i = 1, \dots, n$ son matrices de $n \times n$; $\varepsilon_t = (\varepsilon_{1,t}, \dots, \varepsilon_{n,t})^T$ es ruido blanco multivariado con matriz de covarianza Ω y $\mathbf{s}_t = (s_{1,t}, \dots, s_{n,t})^T$ es un vector de términos determinísticos. En términos del operador L , un modelo VAR se puede escribir de la forma

$$\mathbf{x}_t = (A_1 L + A_2 L^2 + \dots + A_p L^p) \mathbf{x}_t + \mathbf{s}_t + \varepsilon_t.$$

En la mayoría de las aplicaciones, el término determinista \mathbf{s}_t consiste en términos constantes, por ejemplo $\mathbf{s}_t = \mathbf{v}$, donde \mathbf{v}_t es un vector constante. Un modelo de series de tiempo univariado está limitado por solo una parte de toda la información y por tanto una extensión natural son los modelos VAR. Sin embargo, en el caso que se tenga disponible mucha información en series de tiempo, no resulta conveniente incluir todas las variables en el modelo. Tratar de estimar más parámetros puede llevar a resultados imprecisos, especialmente en los modelos VAR, donde el número

de parámetros a estimar crece de manera cuadrática con respecto al número de variables.

En la siguiente sección se presenta una alternativa para incluir más información en el modelo utilizando técnicas de reducción de orden. El modelo aproximado factorial dinámico presentado en [27] nos permite hacer predicciones de variables macroeconómicas de alta dimensionalidad.

5. MODELOS FACTORIALES DINÁMICOS APROXIMADOS

Un modelo utilizado para pronósticos en macroeconomía se basa en el llamado *Modelo factorial dinámico aproximado*. En este modelo series de tiempo que dependen solo de un número esencial de factores son observadas con cierto ruido. Los modelos factoriales descomponen el comportamiento de una variable económica x_{it} en componentes controladas por algunos factores F_t , los cuales son comunes para todas las variables, pero con efectos específicos en cada una de ellas [27]. La variable a pronosticar depende de una combinación lineal de sus valores pasados y también depende de los siguientes factores a estimar

$$(5.1) \quad x_t = \Lambda F_t + \tilde{x}_t,$$

donde x_t es el vector de observaciones en el tiempo t ; $\Lambda = (\lambda_1, \dots, \lambda_N)^T$ es una matriz de cargas de $N \times r$; F_t es el vector de factores de $r \times 1$ y \tilde{x}_t es un vector que representa ruido o incertidumbre de medición. En nuestro caso x_t es medido, F_t y \tilde{x}_t son desconocidos. Generalmente en modelos macroeconómicos la dimensión de x_t es mayor que la dimensión de F_t . La variable a ser pronosticada y_{t+1} se asume que sigue el siguiente modelo dinámico

$$(5.2) \quad y_{t+1} = \beta^T F_t + \gamma(L)y_t + v_t,$$

donde y_{t+1} es la variable a ser pronosticada; $\beta = (\beta_1, \dots, \beta_r)^T$; F_t es el mismo vector de factores de (5.1); $\gamma(L)$ es un polinomio autorregresivo similar a (4); $x_t = (x_{1t}, \dots, x_{Nt})^T$; v_t es un vector que representa ruido aleatorio o errores de medición. La ecuación (5.2) implica que los y_{t+1} están formados por los factores previos, sus elementos autorregresivos y ruido. La ecuación (5.1) implica que la variación de las N variables está determinada por r factores. En [8] se utilizó este tipo de modelo para pronosticar la tasa de crecimiento de producción industrial, la tasa de desempleo y la tasa de inflación en la economía de Taiwan. Los datos contienen 95 variables macroeconómicas y 87 series de tiempo internacionales durante 1981-2006.

La presencia de alta dimensionalidad conlleva un gran reto en el desarrollo de la teoría econométrica. Debido a la habilidad de resumir información de conjuntos de datos grandes, los modelos factoriales se han convertido en una herramienta

indispensable en economía y finanzas. Una revisión importante del aporte de los modelos factoriales en estimación e inferencia estadística se puede encontrar en [2].

6. PROBLEMAS DE ERROR EN VARIABLES (EIV)

La identificación de sistemas dinámicos lineales, donde los datos de salida del sistema provienen de mediciones distorsionadas por ruido, ha sido ampliamente estudiado. Diferentes soluciones a dicho problema se han presentado, por ejemplo en [20, 29]. Por otro lado, la identificación de sistemas cuando los datos de entrada también están contaminados por errores de medición se considera un problema con mayor grado de dificultad. Este tipo de problemas es conocido como modelos de error en variables (EIV, por sus siglas en inglés). Una motivación para el estudio de modelos EIV resulta de la situación presentada en la sección anterior, donde datos con alta dimensionalidad son aproximados por un pequeño número de factores, lo cual constituye la motivación del análisis factorial.

Se considera el siguiente sistema dinámico lineal como ejemplo prototipo, con medidas de entrada-salida distorsionadas por ruido como se muestra en la siguiente figura

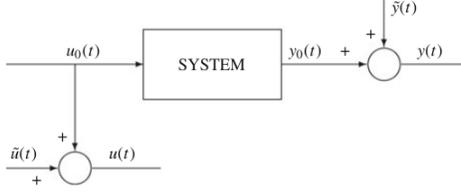


FIGURA 1. Prototipo para el problema EIV, Fuente: [28].

Sea $u(t) = u_0(t) + \tilde{u}(t)$, $y(t) = y_0(t) + \tilde{y}(t)$, donde $u_0(t)$ y $y_0(t)$ son señales libres de ruido que están relacionadas por una ecuación lineal de diferencias de la forma

$$A(L)y_0(t) = B(L)u_0(t),$$

donde $A(L)$ y $B(L)$ son polinomios autorregresivos en términos del operador L definido en la sección (4) de la forma

$$\begin{aligned} A(L) &= 1 + a_1L + a_2L^2 + \dots + a_{n_a}L^{n_a} \\ B(L) &= 1 + b_1L + b_2L^2 + \dots + b_{n_b}L^{n_b} \end{aligned}$$

El problema consiste en determinar las características del sistema, es decir, dadas las muestras de señales de entrada-salida con ruido $\{y_i, u_i\}_{i=1}^N$, determinar una estimación del vector de parámetros del sistema

$$\theta = (a_1, \dots, a_{n_a}, b_1, \dots, b_{n_b})^T.$$

Luego si consideramos el sistema dinámico de la forma

$$y(t) = \varphi^T(t)\theta + \varepsilon(t),$$

con $\varphi(t) = (-y(t-1), \dots, -y(t-n_a), u(t-1), \dots, u(t-n_b))^T$, donde el efecto del ruido se puede expresar como $\varepsilon(t) = A(L)\tilde{y}(t) - B(L)\tilde{u}(t)$. El sistema se puede representar en su forma matricial como:

$$\Phi\theta = \mathbf{Y},$$

donde

$$\Phi = \begin{bmatrix} -y(n) & \cdots & -y(1-n_a+n) & u(n) & \cdots & u(1-n_b+n) \\ -y(n+1) & \cdots & -y(2-n_a+n) & u(n+1) & \cdots & \\ \vdots & \ddots & & \vdots & \ddots & \\ \vdots & & & \vdots & & \\ -y(N-1) & \cdots & -y(N-n_a) & u(N-1) & \cdots & u(N-n_b) \end{bmatrix}$$

$$\mathbf{Y} = \begin{bmatrix} y(n+1) \\ \vdots \\ y(N) \end{bmatrix}$$

con $n = \max(n_a, n_b)$.

El método de mínimos cuadrados total (TLS, por sus siglas en inglés) se puede utilizar para resolver el sistema sobredeterminado de ecuaciones anterior, donde tanto la matriz de coeficientes, como el lado derecho de la ecuación están sujetos a errores de medición. El problema consiste en encontrar θ_{TLS} tal que

$$(6.1) \quad \theta_{TLS} = \arg \min_{\theta \in \mathbb{R}^{n_a+n_b}} \|\Phi\theta - \mathbf{Y}\|_F^2.$$

El siguiente resultado presentado en [13], nos permite aproximar de forma explícita el problema anterior de mínimos cuadrados total

$$\theta_{TLS} = (\Phi^T \Phi - \sigma^2 \mathbf{I})^{-1} \Phi^T \mathbf{Y},$$

donde σ es el mayor valor singular de Φ .

Una desventaja de utilizar el método TLS para encontrar la solución de (6.1) es que no explota la estructura de la matriz Φ . Con el objetivo de aprovechar tal aspecto estructural, algunas extensiones al método TLS se han desarrollado. Por

ejemplo en [22] se introduce el método mínimos cuadrados totales estructurados (STLS, por sus siglas en inglés).

7. CONCLUSIONES Y TRABAJO FUTURO

En este documento se hace una revisión de los modelos clásicos univariados para series de tiempo basados en procesos autorregresivos AR, además su extensión a modelos multivariados VAR. Como un primer tratamiento en la identificación de modelos económicos, los modelos clásicos aportan suficientes elementos para el pronóstico de variables macroeconómicas. A medida crece el número de variables y los datos se vuelven de alta dimensionalidad, los modelos clásicos no son adecuados y es necesario la implementación de modelos estructurados, por ejemplo el modelo factorial dinámico aproximado. Dicho modelo es un caso especial de una clase de modelos llamados error en variables EIV. El problema general EIV se puede solucionar mediante el método de mínimos cuadrados total TLS, pero presenta el inconveniente que no garantiza que la solución sea esparcida. Un enfoque por explorar en el futuro es la extensión del TLS al método de mínimos cuadrados total estructurado STLS. En todos los casos presentados, el problema de la estimación de los parámetros de los modelos, al final se reduce a encontrar la solución de un sistema sobredeterminado de ecuaciones de la forma $AX \approx B$, donde A es una matriz de $m \times n$. En general el tipo de soluciones del sistema anterior depende del método seleccionado. En trabajos futuros se buscará desarrollar métodos teóricos y computacionales que permitan generar soluciones esparcidas en la estimación de los parámetros de modelos económicos de alta dimensionalidad.

REFERENCIAS

- [1] Artacho, F. J. A., Campoy, R., Tam, M. K. (2019). The douglas-rachford algorithm for convex and nonconvex feasibility problems.
- [2] Bai, J., Wang, P. (2015). Econometric analysis of large factor models. *Annual Review of Economics*.
- [3] Berdiev, A. N., Saunoris, J. W. (2016). Financial development and the shadow economy: A panel var analysis. *Economic Modelling*.
- [4] Bernanke, B. S., Boivin, J., Eliasziw, P. (2005). Measuring the effects of monetary policy: A factor-augmented vector autoregressive (favar) approach. *The Quarterly Journal of Economics*, 120(1).
- [5] Brum, J., Scheidegger, S. (2017). Using adaptive sparse grids to solve high-dimensional dynamic models. *Econometrica*, 85(5).
- [6] Cai, T. T., Zhang, C.-H., Zhou, H. H. (2010). Optimal rates of convergence for covariance matrix estimation. *The Annals of Statistics*, 38(4): 2118 – 2144. doi:10.1214/09-AOS752.
- [7] Cao, J., Wang, J. Stock price forecasting model based on modified convolution neural network and financial time series analysis. *International Journal of Communication Systems*, 32(12): e3987. doi:https://doi.org/10.1002/dac.3987. E3987 IJCS-18-0961.R2.
- [8] Chen, C.-C. (2011). The advantages of dynamic factor models as techniques for forecasting: Evidence from taiwanese macroeconomic data. *International Journal of Economics and Finance*.
- [9] Dillon, J. R. A. J. L. (1971). On estimating allocative efficiency in cross-sectional analysis of production. *Australian Journal of Agricultural Economics*.
- [10] Fan, J., Kim, D. (2018). Structured volatility matrix estimation for non-synchronized high-frequency financial data. *Journal of Econometrics*.

- [11] Hannan, E. J. (1970). The asymptotic theory of linear time series models. *Journal of Applied Probability*.
- [12] Heinrich, T. (2008). Identification problems in growth econometrics. *Stockholm University*.
- [13] Huffel, S. V., Cheng, C.-L., Mastronardi, N. (2007). Total least squares and errors-in-variables modeling. *Computational and statistic data analysis*.
- [14] Johansen, S. (2010). Some identification problems in the cointegrated vector autoregressive model. *Journal of Econometrics*, 158(2): 262–273. doi: <https://doi.org/10.1016/j.jeconom.2010.01.007>. Specification Analysis in Honor of Phoebus J. Dhrymes.
- [15] Kadah, N., Özbek, N. S. (2020). Model investigation of nonlinear dynamical systems by sparse identification. *European Journal of Science and Technology*.
- [16] Kaiser, E., Kutz, J. N., Brunton, S. L. (2017). Sparse identification of nonlinear dynamics for model predictive control in the low-data limit. *The Royal Society A Mathematical Physical and Engineering Sciences*.
- [17] Kaufmann, S., Schumacher, C. (2013). Bayesian estimation of sparse dynamic factor models with order-independent identification. *The Royal Society A Mathematical Physical and Engineering Sciences*.
- [18] Kock, A. B., Tang, H. (2018). Uniform inference in high-dimensional dynamic panel data model with approximately sparse fixed effects. *Econometric Theory*.
- [19] Koopmans, T., Reiersol, O. (1950). The identification of structural characteristics. *Cowles Commission for research in Economics*.
- [20] Ljung, L. (1999). *System Identification: Theory for the User*. Prentice Hall PTR.
- [21] Mann, H. B., Wald, A. (1943). On stochastic limit and order relationships. *Columbia University*.
- [22] Markovskiy, I., Willems, J. C., Huffel, S. V., Moor, B. D., Pintelon, R. (2005). Application of structured total least squares for system identification and model reduction. *IEEE Transaction on automatic control*.
- [23] Reddy, D. S., Prasad, P. R. C. (2018). Prediction of vegetation dynamics using ndvi time series data and lstm. *Modeling Earth Systems and Environment*, 4: 409–419. doi: <https://doi.org/10.1007/s40808-018-0431-3>.
- [24] Schaeffer, H., Tran, G., Ward, R. (2018). Extracting sparse high-dimensional dynamics from limited data. *SIAM J. APPL. MATH.*, 78(6).
- [25] Schaeffer, H., Tran, G., Ward, R. (2018). Extracting sparse high-dimensional dynamics from limited data.
- [26] Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*.
- [27] Stock, J. H., Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2): 147–162. doi:10.1198/073500102317351921.
- [28] Söderström, T. (2018). Errors-in-variables methods in system identification. *Springer*.
- [29] Söderström, T., Stoica, P. (1989). *System Identification*. Prentice Hall.
- [30] Vides, F. (2020). On cyclic finite-state approximation of data-driven systems. *IEEE*.
- [31] Vides, F. (2020). Universal algebraic controllers and system identification.
- [32] Wold, H. (1938). On the inversion of moving averages. *Scandinavian Actuarial Journal*, 3-4.
- [33] Xu, W., Peng, H., Zeng, X., Zhou, F., Tian, X., Peng, X. (2019). A hybrid modelling method for time series forecasting based on a linear regression model and deep learning. *Applied Intelligence*, 49. doi:10.1007/s10489-019-01426-3.
- [34] Yule, G. U. (1927). On a method of investigating periodicities in disturbed series, with special reference to wolfer’s sunspot numbers. *The Royal Society*.
- [35] Zhang, G. (2003). Time series forecasting using a hybrid arima and neural network model. *Neurocomputing*, 50: 159–175. doi:[https://doi.org/10.1016/S0925-2312\(01\)00702-0](https://doi.org/10.1016/S0925-2312(01)00702-0).

DEPARTAMENTO DE MATEMÁTICA APLICADA, UNIVERSIDAD NACIONAL AUTÓNOMA DE HONDURAS

Dirección de correo electrónico: kerin.cardona@unah.edu.hn

IDENTIFICACIÓN APROXIMADA DE SEÑALES EN RECONOCIMIENTO DE PATRONES

WILLIAM FUNEZ

RESUMEN. En este artículo estudiamos un sistema de identificación y clasificación de señales que utiliza el algoritmo de descomposición en valores singulares (SVD) y análisis de componentes principales (PCA). El algoritmo se basa en un enfoque de “eigenfaces” para los primeros experimentos realizados, que representa un método de PCA en el que se utiliza un pequeño conjunto de características significativas para describir la variación entre las imágenes. Se muestran resultados experimentales para diferentes números de “eigenfaces” para verificar la viabilidad del método propuesto.

ABSTRACT. In this article we study a signal identification and classification system which uses the Singular Value Decomposition (SVD) and Principal Component Analysis (PCA) algorithm. The algorithm is based on an eigenfaces approach for the first performed experiments, which represents a PCA method in which a small set of significant features is used to describe the variation between images. Experimental results are shown for different numbers of eigenfaces to verify the feasibility of the proposed method.

1. INTRODUCCIÓN

Una de las demostraciones más llamativas de la descomposición en valores singulares (SVD) y análisis de componentes principales (PCA) es la identificación o clasificación de imágenes. En este problema, PCA (es decir, SVD en datos de resta de la media) se aplica a un conjunto de imágenes para extraer las correlaciones más dominantes entre las imágenes. El resultado de esta descomposición es un conjunto que se le llama “eigenfaces” que definen un nuevo sistema de coordenadas donde las imágenes pueden representarse en estas coordenadas tomando el producto escalar con cada una de las principales componentes [9]. El problema llamado “eigenface” fue estudiado por primera vez por Sirovich y Kirby en [11]. Su aplicación al reconocimiento facial automatizado fue presentado por Turk y Pentland en [10]. Este problema se resuelve mediante PCA, la cual es una técnica de proyección que encuentra un conjunto de vectores de proyección diseñados para que los datos proyectados retengan la mayor parte de la información sobre los datos originales. Los vectores más representativos son los autovectores correspondientes a los autovalores más altos de la matriz de covarianza [9]. Este método reduce la dimensionalidad del espacio de datos al proyectar datos desde el espacio de dimensión M al espacio de dimensión N , donde $N \ll M$. En la siguiente sección se presenta la justificación

Fecha: Abril 19, 2021.

Palabras y frases clave. Descomposición en valores singulares, análisis de componentes principales, identificación, eigenfaces.

Asesor Dr. Fredy Vides.

del proyecto, donde se dará la importancia y líneas de investigación relacionadas con el proyecto.

2. JUSTIFICACIÓN

La importancia de este trabajo recae en la creciente intersección de los métodos basados en datos, optimización, estadística y otros temas de la matemática. La mayoría de los programas de pregrado en ingeniería y los campos de la ciencia en la UNAH, tienen poca exposición a los métodos de datos y optimización, el objetivo es proporcionar una introducción a la ciencia de datos. Se ha elegido SVD y PCA por su relevancia, simplicidad y generalidad, esto ya que cada vez más los investigadores están utilizando enfoques basados en datos.

El descubrimiento basado en datos está revolucionando actualmente la forma en que modelamos, predecimos y controlamos sistemas complejos. Los problemas científicos y de ingeniería más apremiantes de la era moderna no son basados en principios básicos, esto hace que cada vez más los investigadores recurran a enfoques basados en datos para una amplia gama de sistemas complejos como, la turbulencia, el cerebro, el clima, la epidemiología, las finanzas, la robótica, la autonomía, entre otros [9]. Estos sistemas son típicamente no lineales, dinámicos, de escala múltiple en espacio y tiempo, de alta dimensión, con patrones dominantes que deben ser modelados para el objetivo eventual de detección, predicción, estimación y control [2, 7].

Las aplicaciones que se presentarán en este trabajo será la identificación y detección de imágenes, ya que uno de los enfoques de PCA más simples y efectivos utilizados en los sistemas de identificación de imágenes es el llamado enfoque de “eigenfaces”. Este enfoque transforma los rostros en un pequeño conjunto de características esenciales, las “eigenfaces”, que son los componentes principales del conjunto inicial de imágenes de aprendizaje (conjunto de formación), en la sección 5 se brinda más detalles sobre la idea principal de lo son las “eigenfaces”.

La identificación se realiza proyectando una nueva imagen en el subespacio de “eigenfaces”, después de lo cual la persona se clasifica comparando su posición en el espacio de eigenfaces con la posición de los individuos conocidos [11]. La ventaja de este enfoque sobre otros sistemas de identificación es su simplicidad, velocidad e insensibilidad a los cambios pequeños o graduales en el rostro. El problema se limita a los archivos que se pueden utilizar para identificar el rostro. Es decir, las imágenes deben ser vistas de manera frontal y vertical. También se utiliza la modificación de los algoritmos en [3] y un resultado presentado en [1].

La teoría de aproximación esparcida (representación esparcida) trata con soluciones esparcidas para sistemas de ecuaciones lineales, las técnicas para encontrar éstas soluciones y explotarlas en aplicaciones han encontrado un amplio uso en el procesamiento de imágenes en [4], en este caso se presentará una aplicación mediante SVD para clasificación de imágenes. Para mejorar ciertas debilidades de este método se desarrollará análisis robusto de componentes principales (RPCA), que busca descomponer una matriz de datos A en una matriz estructurada L de “low-rank” y una matriz esparcida S [5]. A continuación se mostrará la líneas de investigación entre las cuales este proyecto está relacionado.

La Universidad Nacional Autónoma de Honduras (UNAH), estipula como uno de sus ejes de investigación el estudio de la Población, dentro de éste se define como

tema prioritario de investigación, entre otros, ciencia, de esta manera, el presente estudio se enmarca en este eje prioritario. Considerando que éste artículo cubre el análisis de datos, este es un tema de interés enfocado en la línea de investigación de optimización matemática, análisis numérico de la Maestría en Matemática con Orientación a Ingeniería de la UNAH, ya que los sistemas suelen ser no lineales, dinámicos y de alta dimensión, con patrones dominantes que deben caracterizarse y modelarse para el objetivo eventual de detección, predicción, estimación y control. En la siguiente sección se expone una perspectiva histórica sobre como se ha desarrollado la temática del artículo desde sus inicios hasta como se logró llegar a los conceptos mas modernos.

3. ANTECEDENTES

La SVD tiene una larga y rica historia, que va desde los primeros trabajos de desarrollo de los fundamentos teóricos hasta el trabajo moderno sobre estabilidad y eficiencia computacional. Hay una excelente revisión histórica de Stewart en [6], que proporciona contexto y muchos detalles importantes. La revisión se centra en los primeros trabajos teóricos de Beltrami y Jordan (1873), son los progenitores de la SVD, Beltrami en virtud de la primera publicación y Jordan por la integridad y elegancia de su tratamiento. La contribución de Beltrami apareció en el *Journal of Mathematics for the Use of the Students of the Italian Universities* y su propósito era alentar a los estudiantes a familiarizarse con las formas bilineales. Sylvester (1889), escribió dos documentos sobre el tema de la SVD, también una nota al pie de página que aparece al final de un artículo en *The Messenger of Mathematics*, en el documento Sylvester describe un algoritmo iterativo para reducir una forma cuadrática a una forma diagonal, en la nota señala que se puede usar una iteración análoga para diagonalizar una forma bilineal. Schmidt (1907), logra que la historia ahora pase del dominio del álgebra lineal a las ecuaciones integrales, uno de los temas candentes de las primeras décadas del siglo XX, en el tratamiento de ecuaciones integrales con núcleos asimétricos, Erhard Schmidt (de la fama de Gram-Schmidt y un estudiante de Hilbert) introdujo el análogo de dimensión infinita de la SVD. Pero fue más allá de la mera existencia de la descomposición al mostrar cómo se puede utilizar para obtener aproximaciones óptimas de low-rank a un operador, al hacerlo, transformó la SVD de una curiosidad matemática a una importante herramienta teórica y computacional. Autonne (1913), extendió la descomposición a matrices complejas. Hotelling (1933), como alternativa al análisis factorial (técnica estadística de reducción de datos usada para explicar las correlaciones entre las variables observadas en términos de un número menor de variables no observadas llamadas factores). Eckart (1936) y Young (1939) lo extendieron a matrices rectangulares y redescubrieron el teorema de aproximación de Schmidt, que a menudo (e incorrectamente) se llama el teorema de Eckart-Young. La SVD fue introducida en el análisis numérico por Golub y Kahan (1965). En las siguientes secciones se dará una breve descripción de SVD y propiedades que se utilizarán para el desarrollo del proyecto.

4. DESCRIPCIÓN DE SVD

Como en este proyecto la información que necesitamos requiere ser almacenada en una matriz A , que a su vez el propósito principal es representar e identificar la

información de dicha matriz, se hará uso de la técnica de SVD. En términos generales la SVD es una factorización matricial que nos proporciona ciertas características de la matriz A , también es una generalización del teorema espectral, ya que en este caso podemos hacer la descomposición para cualquier matriz rectangular y no simétrica, nos permite describir el mapeo lineal asociado a la matriz A mediante el sistema $Ax = b$.

En términos generales para cualquier matriz $A \in \mathbb{R}^{n \times m}$ se puede obtener una factorización de la forma $A = U\Sigma V^T$, donde $U \in \mathbb{R}^{n \times n}$, $V \in \mathbb{R}^{m \times m}$ son matrices ortogonales y $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_r)$ son los valores singulares de A , se tiene que $r = \text{rank}A$, esto equivale al Teorema 2.6.3 de [1]. Una vez obtenida la SVD es posible acceder a una gran cantidad de información de la matriz A , por ejemplo el rango de la matriz, la norma espectral y el número de condición [1, 2]. También podemos resolver sistemas de ecuaciones que consideren a la matriz A como la matriz de coeficientes y es importante en algunos problemas de optimización, estos problemas nos ayudarán a obtener una mejor aproximación de “low-rank”.

Otra propiedad que se puede extraer de la SVD es la asociada con sus normas, ya que para la SVD los valores singulares juegan un papel muy importante, se tendrá que hacer uso de lo siguiente [1]:

1. Norma Espectral: $\|A\|_2 = \sigma_1$.
2. Norma de Frobenius: $\|A\|_F^2 = \sum_{i=1}^n \sigma_i^2$.

En la siguiente sección se presenta la idea principal detrás de las “eigenfaces”, su cálculo y como se representan en una base.

5. IDEA PRINCIPAL DETRÁS DE LAS EIGENFACES

Suponga que Γ es un vector de tamaño $N^2 \times 1$, correspondiente a una imagen I de tamaño $N \times N$, donde la imagen tiene N píxeles de alto y N píxeles de ancho. La idea es representar Γ ($\Phi = \Gamma - \text{cara media}$) en un espacio de baja dimensión:

$$(5.1) \quad \tilde{\Phi} - \text{media} = \sum_{i=1}^K w_i u_i$$

Cálculo de las Eigenfaces. La ecuación (5.1) representa las caras sobre una base se da más detalle de esta ecuación al final de esta sección, a continuación se muestra el proceso de como realizar el cálculo de las “eigenfaces”:

Paso 1: Se obtendrá las imágenes de las caras I_1, I_2, \dots, I_M .

*Obs. Las imágenes de la cara deben estar centradas y deben tener el mismo tamaño [4].

Paso 2: Representar cada imagen I_i como un vector Γ_i .

Paso 3: Calcular el vector de cara promedio Ψ :

$$\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i$$

Paso 4: Restar la cara media

$$\Phi_i = \Gamma_i - \Psi.$$

Paso 5: Calcular la matriz de covarianza C :

$$C = \frac{1}{M} \sum_{i=1}^M \Phi \Phi^T = AA^T, \quad (\text{Matrix de } N^2 \times N^2),$$

$$\text{donde } A = [\Phi_1 \cdots \Phi_M], \quad (\text{Matrix de } N^2 \times M).$$

Paso 6: Calcular los autovectores u_i de AA^T

Paso 6.1: Considerando la matriz $A^T A$ (de tamaño $M \times M$).

Paso 6.2: Calcular los autovectores v_i de $A^T A$.

$$A^T A v_i = \mu_i v_i$$

$$AA^T A v_i = \mu_i A v_i$$

$$C A v_i = \mu_i u_i \text{ donde } u_i = A v_i.$$

Por tanto, AA^T y $A^T A$ tienen los mismos valores propios y sus vectores propios están relacionados de la siguiente manera: $u_i = A v_i$.

Obs. 1. AA^T puede tener hasta N^2 autovalores y autovectores.

Obs. 2. $A^T A$ puede tener hasta M autovalores y autovectores.

Obs. 3. Los M autovalores de $A^T A$ (junto con sus correspondientes autovectores) corresponden a los M autovalores más grandes de AA^T (junto con sus correspondientes autovectores).

Paso 6.3: Calcular los M mejores vectores propios de AA^T : $u_i = A v_i$.

Obs. Normalizar u_i tal que $\|u_i\| = 1$

Paso 7: Mantener solo K autovectores (correspondientes a los K autovalores más grandes)

Representar las caras sobre una base. Cada cara (menos la media) Φ_i en el conjunto se puede representar como una combinación lineal de los mejores vectores propios de K :

$$\tilde{\Phi}_i - \text{media} = \sum_{i=1}^K w_i u_i, \quad (w_i = u_i^T \Phi_i),$$

cada cara normalizada Φ_i está representada en una base por un vector

$$\Omega_i = [w_1^i \cdots w_K^i]^T, \quad i = 1, \dots, M.$$

En la siguiente sección se presenta el uso de “eigenfaces” para identificación y detección de imágenes.

6. IDENTIFICACIÓN Y DETECCIÓN

El análisis de componentes principales (PCA) transforma un conjunto de datos obtenidos de variables posiblemente correlacionadas en un conjunto de valores de variables no correlacionadas llamadas componentes principales [9]. El número de componentes puede ser menor o igual al número de variables originales. La primera componente principal tiene la varianza más alta posible, y cada uno de las componentes siguientes tiene la varianza más alta posible bajo la restricción de que debe ser ortogonal a la componente anterior. Se desea encontrar las componentes principales, en este caso vectores propios de la matriz de covarianza de las imágenes. Lo primero que debemos hacer es crear un conjunto de datos de entrenamiento. La imagen I_i se puede representar como un vector concatenando filas [10]. Lo antes

mencionado es lo que se mostró en la sección 5, a continuación se presentará el uso de “eigenfaces” para identificación y detección de imágenes con sus correspondientes pasos.

Algorithm 1 Identificación.

Input: Dada una imagen desconocida Γ (centrada y del mismo tamaño que el conjunto ingresado al inicio), se debe seguir lo siguiente:

- 1: Normalizar Γ : $\Phi = \Gamma - \Psi$
- 2: Proyectar en el eigenspace

$$\tilde{\Phi}_i - \text{media} = \sum_{i=1}^K w_i u_i, \quad (w_i = u_i^T \Phi_i)$$

- 3: Representar Φ como: $\Omega = [w_1 \cdots w_K]^T$
 - 4: Encontrar $e_r = \min_l \|\Omega - \Omega^l\|$
 - 5: Si $e_r < T_r$, entonces Γ se reconoce como la cara l del conjunto.
-

Algorithm 2 Detección.

Input: Dada una imagen desconocida Γ (centrada y del mismo tamaño que el conjunto ingresado al inicio), se debe seguir lo siguiente:

- 1: Calcular $\Phi = \Gamma - \Psi$
- 2: Calcular

$$\tilde{\Phi}_i - \text{media} = \sum_{i=1}^K w_i u_i, \quad (w_i = u_i^T \Phi_i)$$

- 3: Calcular $e_d = \|\Phi_i - \tilde{\Phi}_i\|$
 - 4: Si $e_d < T_d$, entonces Γ es la cara.
-

En la siguiente sección se presenta un experimento numérico que muestra como resolver la identificación y detección de imágenes.

7. EXPERIMENTO NUMÉRICO

El experimento numérico muestra como resolver la identificación de imágenes, se toma un grupo de imágenes de un hombre en este caso del actor Sylvester Stallone y otro de una mujer en este caso la cantante Taylor Swift se toma 20 imágenes de cada persona (ver figura 1), las cuales fueron extraídas de internet, recordar que todas las imágenes deben estar centradas y tener el mismo tamaño, cada imagen tiene 200 píxeles de alto y 175 píxeles de ancho donde cada una de las imágenes que se utiliza se ha reformado en un vector columna con $200 \times 175 = 35,000$ elementos, a continuación se presenta el algoritmo 1 el cual nos permite identificar las imágenes en dos grupos (ver figura 4), mediante las técnicas descritas en las secciones anteriores, también nos permite detectar una imagen ingresada después del conjunto inicial, para lo antes mencionado se utiliza GNU Octave con técnicas computacionales utilizadas en [12].

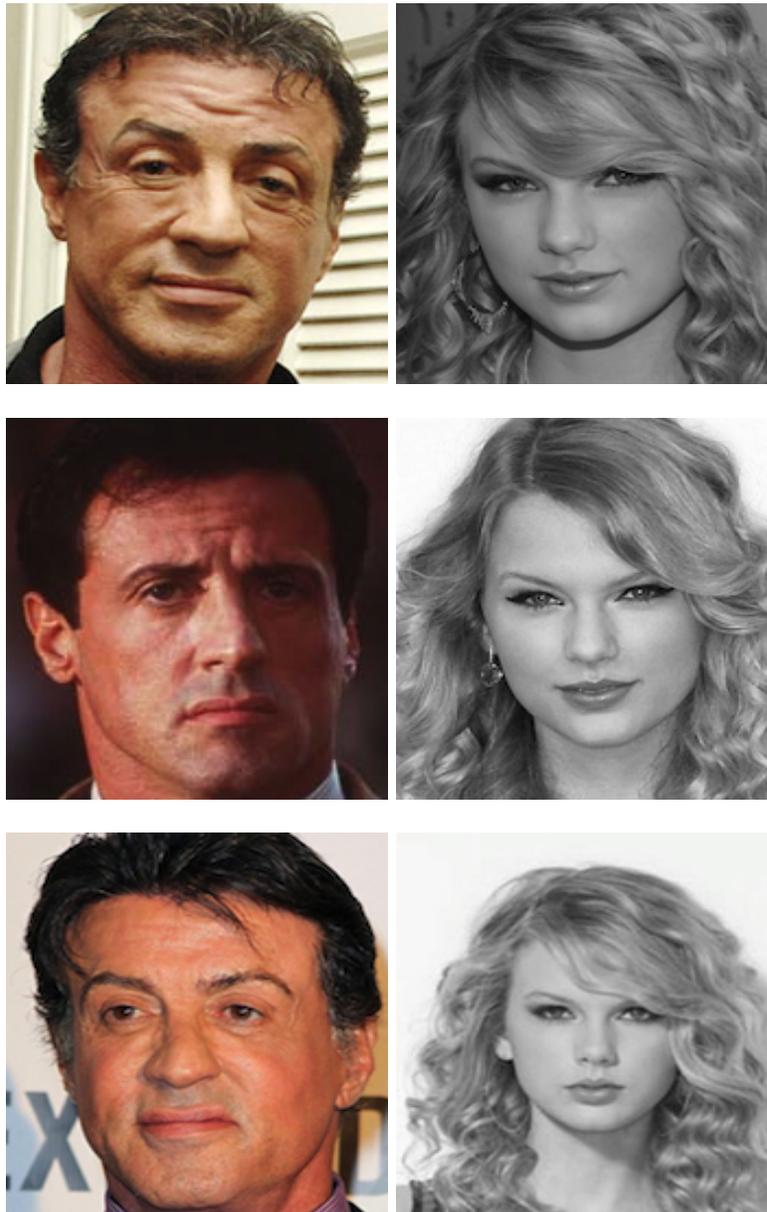


FIGURA 1. Algunas imágenes del conjunto de las 20 imágenes que se cargan para cada grupo para realizar el experimento.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%  
% ALGORITMO 1  
% Identificacion  
% Deteccion
```

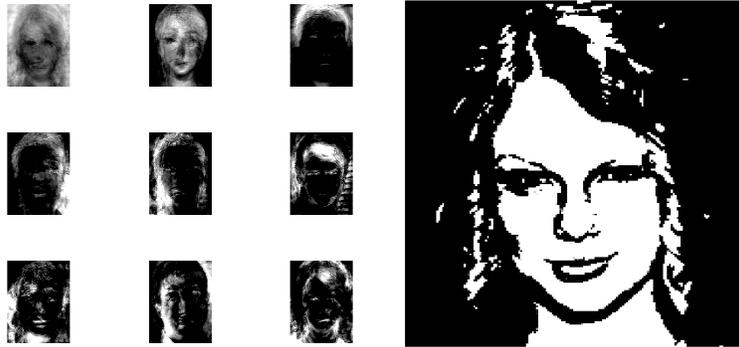
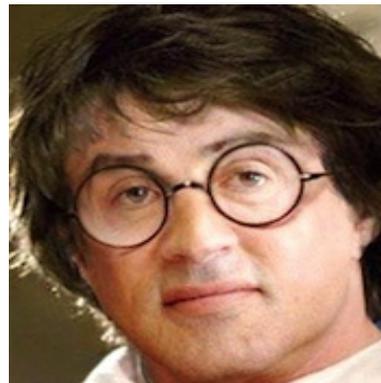


FIGURA 2. Proyecta la imagen sobre una base.



(A) Imagen de una mujer la cual no es Taylor Swift, además cuenta con daños en su rostro, notaremos que sé aleja en la identificación.



(B) Imagen de Stallone, solo que está imagen cuenta con rasgos distintos al grupo de 20 imágenes que se ingresó al inicio.

FIGURA 3. Imágenes que se desean detectar, para cada uno de los distintos grupos.

```
clear all
close all
clc

% Tamaño de cada imagen
m = 200;
n = 175;

% Numero de imagenes (muestra)
N = 20;

avg = zeros(m*n,1); % la cara promedio
```

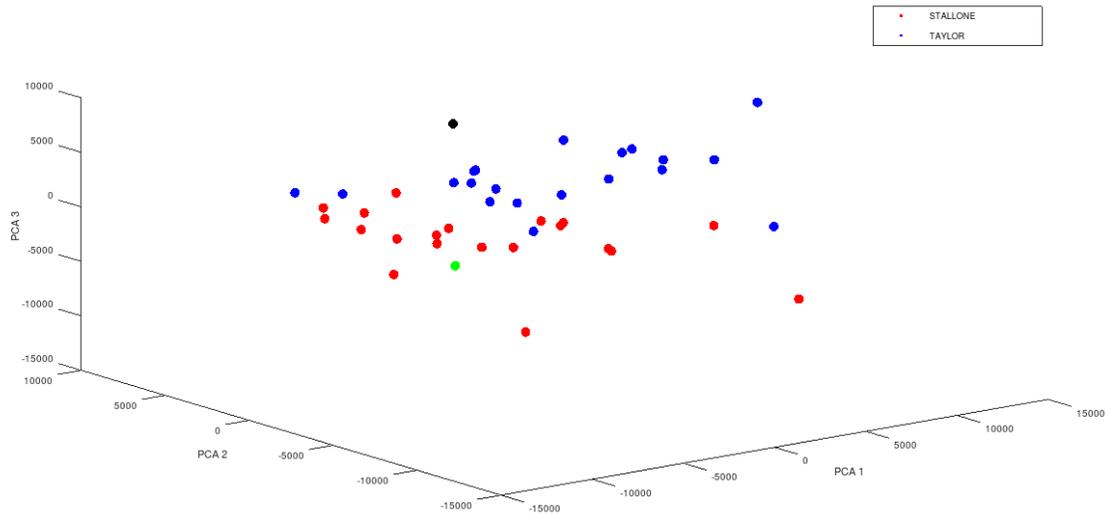


FIGURA 4. En el siguiente gráfico se muestra la identificación de los dos grupos, los puntos de color rojo son los correspondientes a Stallone y los de color azul a Taylor Swift, donde las imágenes ingresadas al final están de colores distintos para la imagen de Stallone con rasgo distinto se tiene en color verde la cual se mantiene en el grupo, así que detecta bien las características principales del rostro, para la imagen de la mujer con daños en el rostro que no es Taylor Swift se tiene en color negro podemos ver que se aleja del grupo, lo cual está bien, ya que no corresponde al grupo de imágenes de Taylor Swift.

```
A = [];
```

```
% Cargar imagenes de Sylvester Stallone
count = 0;
for j = 1:N
    figure(1)
    ff = ['faces/stallone',num2str(j,'%02d'),' .jpg'];
    u = imread(ff); % Leer la imagen en una matriz
    imshow(u)
    if(size(u,3)==1)
        M=double(u);
    else
        M=double(rgb2gray(u));
    end
    pause(0.1);
    R = reshape(M,m*n,1);
    A = [A,R];
    avg = avg + R;
    count = count + 1;
```

```

end
%% Cargar imagenes de Taylor Swift
for j = 1:N
    figure(1)
    ff = ['faces/taylor',num2str(j,'%02d'),' .jpg'];
    u = imread(ff); % Leer la imagen en una matriz
    imshow(u)
    M = double(u(:,:,1));
    R = reshape(M,m*n,1);
    A = [A,R];
    pause(0.1);
    avg = avg + R;
    count = count + 1
end
avg = avg /count;

%% Calcular el "promedio" de cada cara
avgTS = uint8(reshape(avg,m,n));
figure(1), imshow(avgTS);

%% Centrar las imagenes de la muestra en el "origen"
figure(1)
for j = 1:2*N
    A(:,j) = A(:,j) - avg;
    R = reshape(A(:,j),m,n);
    imshow(R);
    set(gcf,'Position',[100 900 500 500])
    pause(.1);
end

%% Calculando la SVD
[U,S,V] = svd(A,'econ');
Phi = U(:,1:2*N);
Phi(:,1) = -1*Phi(:,1);
figure(2)
count = 1;
for i=1:3
    for j=1:3
        subplot(3,3,count)
        imshow(uint8(25000*reshape(Phi(:,count),m,n)));
        count = count + 1;
    end
end

%% Proyectar cada imagen sobre la base
for j = 1:N

```

```

        imvec = A(:,j);
        STAL(:,j) = imvec'*Phi(:,1:3);
    end
for j = 1:N
    imvec = A(:,N+j);
    TLR(:,j) = imvec'*Phi(:,1:3);
end

figure(3)

plot3(STAL(1,:),STAL(2,:),STAL(3,:), 'r.', 'MarkerSize', 30)
hold on
plot3(TLR(1,:),TLR(2,:),TLR(3,:), 'b.', 'MarkerSize', 30)
xlabel('PCA 1')
ylabel('PCA 2')
zlabel('PCA 3')
legend('STALLONE', 'TAYLOR')

%% Agrega algunas imagenes fuera del conjunto
u = imread('faces/pruebastallone1.jpg');
figure(4)
subplot(1,2,1)
imshow(u);
u = double(rgb2gray(u));
ustal = reshape(u,m*n,1)-avg;
stalpts = ustal'*Phi(:,1:3);
v = imread('faces/pruebaconjunto2.jpg');
subplot(1,2,2)
imshow(v);
v = double(rgb2gray(v));
vterm = reshape(v,m*n,1)-avg;
termpts = vterm'*Phi(:,1:3);
%%
figure(3)
plot3(stalpts(1),stalpts(2),stalpts(3), 'g.', 'MarkerSize', 30)
plot3(termpts(1),termpts(2),termpts(3), 'k.', 'MarkerSize', 30)

```

En la siguiente sección se presenta las conclusiones del trabajo realizado para resolver la identificación y detección de imágenes, utilizando el análisis de componentes principales.

8. CONCLUSIÓN

Se utilizó una base de datos que contiene 40 imágenes de 2 personas diferentes (20 imágenes por persona), utilizando el análisis de componentes principales para realizar este proceso, de los resultados obtenidos mediante el experimento numérico podemos notar que la identificación y detección de imágenes es aceptable, pero se puede realizar una mejora a este método, haciendo una pequeña observación también que la tasa de identificación aumenta con el número de imágenes por persona. Es obvio que si la distancia mínima entre la imagen de prueba y otras

imágenes es cero, la imagen de prueba coincide completamente con la imagen de la base. Si la distancia es mayor que cero, pero menor que cierto umbral, es una persona conocida con otra expresión facial, de lo contrario es una persona desconocida.

9. TRABAJO A FUTURO

Para mejorar el sistema de identificación de señales utilizado en este trabajo, en esta mejora se utilizará el algoritmo de descomposición en modo dinámico (DMD), donde esta técnica se basa en la descomposición ortogonal adecuada (POD), que utiliza la (SVD) computacionalmente eficiente de modo que escala bien para proporcionar reducción de dimensionalidad en sistemas de alta dimensión. La DMD es una técnica que permite extraer las características de flujo dinámicamente relevantes de datos experimentales o numéricos. En esta técnica, un mapeo lineal se construye a partir de una serie de datos resueltos en el tiempo, para este tipo de experimento se considerará la descomposición de Koopman en relación con la DMD, considerando la ecuación no lineal de Schrödinger (NLS)

REFERENCIAS

1. Horn, R. A. (1986). *Topics in Matrix Analysis*. USA: Cambridge University Press.
2. Golub, G. H. (1996). *Matrix Computation*. USA: Johns Hopkins University Press.
3. Liberty, E., Woolfe F., Tygert M., Martinsson P., Rokhlin V. (2007). Randomized algorithms for the low-rank approximation of matrices. *Proceedings of the National Academy of Sciences - PNAS*. 104(51): 20167-20172. doi.org/10.1073/pnas.0709640104
4. Yang, A., Ganesh, A., Wright, J., Sastry, S., Ma, Y. (2009). Robust Face Recognition via Sparse Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 31(2):210-27. doi.org/10.1109/TPAMI.2008.79
5. Candès, E. J., Wright, J., Li, X., Ma, Y. (2011). Robust principal component analysis?. *Journal of the ACM*. 58(3):1-37. doi.org/10.1145/1970392.1970395
6. Stewart, G. W. (1993). On the early history of the singular value decomposition. *SIAM Review*. 35(4):551-566. doi.org/10.1137/1035134
7. Trefethen, L. (1997). *Numerical Linear Algebra*. USA: Society for Industrial and Applied Mathematics (SIAM).
8. Ayubi, J. (2014). Lossy Color Image Compression Based on Singular Value Decomposition and GNU GZIP. *ACSIJ Advances in Computer Science: an International Journal*. 3(3): 2322-5157.
9. Nathan Kutz, J. (2013). *Data-Driven Modeling Scientific Computation. Methods for Complex Systems Big Data*. Oxford University Press.
10. Pentland, A., Turk, M. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*. 3(1):71-86. doi.org/10.1162/jocn.1991.3.1.71.
11. Sirovich, L., Kirby M. (1987). Low-dimensional procedure for the characterization of human faces. *Journal of the Optical Society of America A*. 4(3):519-524. doi.org/10.1364/JOSAA.4.000519
12. Vides, F. (2021). *Métodos Numéricos y Modelación Computacional*. <https://caddslab.github.io/HPCMMModeling.html>.

UNIVERSIDAD NACIONAL AUTÓNOMA DE HONDURAS - UNAH, ESCUELA DE MATEMÁTICA Y CIENCIAS DE LA COMPUTACIÓN, MAESTRIA EN MATEMÁTICA
 Dirección actual: Tegucigalpa M.D.C, Honduras
 Dirección de correo electrónico: william.funez@unah.hn

ANÁLISIS DINÁMICO EN REACCIONES QUÍMICAS MODELADAS CON SISTEMAS DINÁMICOS NO LINEALES DE ORDEN FRACCIONARIO

OSCAR MOLINA

RESUMEN. La construcción de modelos para la resolución de los problemas de la sociedad actual es un tema que ha ido en auge en la ciencia contemporánea y ha marcado tendencia en la investigación. Podemos ver como la matemática relacionada con otras disciplinas, generalmente apoyadas de la informática e inteligencia artificial, pueden generar modelos que ayudan a dar solución a los problemas más complejos de la ciencia y la sociedad. En base a lo anterior podemos afirmar que la ciencia contemporánea tiene un interés dominante en construir modelos matemáticos que permitan comprender la dinámica de los sistemas y como mejorar el control sobre los mismos. Es por ello que en este artículo se expone un caso de la ingeniería química, en la que primero se toma la cinética de una reacción química, representada por un sistema dinámico de orden entero y luego se convierte a un sistema de orden fraccionario aproximado. Para realizar esto se utilizan las definiciones de las derivadas de orden fraccionario, así como sus propiedades para convertir dichos modelos. Además, se realiza una simulación numérica en la que se obtiene la respuesta del sistema dinámico así como los retratos de fase y su diagrama de bifurcación.

ABSTRACT. The construction of models to solve the problems of current society is a topic that has been on the rise in contemporary science and it has set a trend in research. We can see how mathematics related to other disciplines, generally supported by computer science and artificial intelligence, can generate models that help to solve the most complex problems of science and society. Based on the above we can affirm that contemporary science has a dominant interest in building models that allow the understanding of the dynamics of systems and how to improve control over them. That is why this article presents a case of chemical engineering, in which first a kinetic chemical reaction represented by a dynamic system of integer order is taken and then it is converted to an approximate fractional order system. To do this, the definitions of the derivatives of fractional order are used, as well as their properties to convert said models. In addition, a numerical simulation is carried out in which the response of the dynamic system is obtained as well as the phase portraits and its bifurcation diagram.

1. INTRODUCCIÓN

La representación matemática de un fenómeno ya sea físico o químico, se modela usualmente utilizando ecuaciones diferenciales; sin embargo existen problemas en que dichas ecuaciones no tienen solución analítica, especialmente las que no son lineales. Éstas generalmente presentan comportamiento caótico, de ahí el

Date: abril, 2021.

Key words and phrases. Cinética, cálculo fraccionario, caos.

interés en apoyarse en el uso de diversas herramientas para entender estos comportamientos y poder controlarlos [1].

Un sistema no necesita ser complejo o de difícil medición experimental para mostrar un comportamiento caótico, pero si ocurre cuando es de tipo no lineal. Tal es el caso que la fuente del caos son precisamente este tipo de sistemas no lineales. El caos también es caracterizado por una alta sensibilidad a las condiciones iniciales, el paso de la unidad individual a la simetría recursiva entre niveles de escala y el poseer mecanismos de realimentación. Sin importar con que precisión se midan estos sistemas, la predicción de su futuro se vuelve radicalmente erróneo luego de cierto tiempo [2].

Las ecuaciones diferenciales fraccionarias, han despertado un interés considerable en la matemática aplicada, han sido utilizadas en muchas aplicaciones físicas y procesos químicos. Los tópicos matemáticos y sus métodos de resolución para este tipo de ecuaciones han sido discutidos ampliamente por muchos autores, los cuales han arrojado soluciones más exactas comparado con otros tipos de sistemas; para el caso que nos compete, se emplea una herramienta computacional para realizar un análisis dinámico del sistema como ser: eigenvalores, bifurcación, retratos de fase, entre otros. Así mismo, se hace una comparación con un sistema dinámico no lineal de ecuaciones diferenciales de orden entero [3].

Este trabajo pretende mostrar que el el cálculo fraccionario es una herramienta más precisa para resolver este tipo de sistemas derivados del modelamiento matemático de la cinética de reacciones químicas.

2. JUSTIFICACIÓN

Las reacciones químicas están siempre presente en la materia y son responsables de innumerables fenómenos de constante transformación de una cosa en otra, tales como: los cambios atmosféricos, la fotosíntesis, la combustión de la gasolina, la corrosión de los metales, la digestión de los alimentos, la fermentación de licores, por mencionar algunos ejemplos básicos. Consecuentemente, el conocimiento sobre las reacciones químicas es importante no solo para comprender las formas y métodos de supervivencia de los humanos, sino también porque, para poder atender las necesidades y/o comodidades de los humanos, las industrias buscan nuevos métodos que permitan hacer más rentables sus procesos físico-químicos. Para poder hacer más eficientes dichos procesos, es necesario tomar en cuenta el modelado matemático de reacciones cinéticas químicas de algunos órdenes que conduce a la formación de un sistema de ecuaciones diferenciales. Estos sistemas pueden ser lineales o no lineales, dependiendo del orden de la cinética en la reacción química [4].

Algunas reacciones químicas se comportan de forma caótica, siendo modeladas por medio de un sistema dinámico no lineal. Esta dinámica no lineal se encuentra en casi todas las áreas de trabajo en la industria química, por ejemplo: operaciones unitarias no convencionales como la separación con membranas (ultra filtración y diálisis) y, procesos convectivos y difusivos en medios tortuosos (poros absorbentes, empaques de torres, granos de catalizador, filtros y lechos fluidizados, entre otros).

Es importante comprender que la reacción química, es uno de los fenómenos donde la irreversibilidad termodinámica permite la auto-organización para sistemas abiertos y lejos del equilibrio. La sola ley de “acción de masas” hace no lineales las ecuaciones de la cinética química, máxime si se introduce el factor exponencial de Arrhenius, para dar cuenta de los efectos de la temperatura. Así mismo, las reacciones químicas y el flujo turbulento son ejemplos en donde las leyes del caos prometen importantes avances para la industria de procesos químicos. El estudio del caos podría facilitar la respuesta a muchas inquietudes vigentes al día de hoy, como ser el comportamiento variable de una reacción química a pequeñas variaciones en sus condiciones iniciales [2].

Este estudio se enmarca dentro del eje de investigación: Ambiente, Biodiversidad y Desarrollo, de las líneas de investigación de la Universidad Nacional Autónoma de Honduras (UNAH); en el cual se pretende realizar un análisis sistemático de un sistema dinámico no lineal, además se puede encuadrar dentro de la línea de investigación de Modelación Matemática de la Orientación en Matemática Aplicada, de la Maestría en Matemática de la Facultad de Ciencias de la UNAH, de modo que pueda aportar al desarrollo académico e industrial de este país.

3. ANTECEDENTES

La dinámica química o cinética química comprende el estudio de las interacciones entre las partículas presentes en las sustancias utilizadas en una reacción química, en un sistema o en la naturaleza. Por otra parte, un sistema dinámico observa como estos sistemas evolucionan en función del tiempo, cuyo comportamiento se puede caracterizar determinando los límites del sistema, los elementos y sus relaciones; de esta forma se pueden elaborar modelos que buscan representar la estructura del mismo sistema [5].

El modelado matemático de reacciones cinéticas químicas de algunos órdenes conduce a la formación de un sistema de ecuaciones diferenciales ordinarias de primer orden. Por tanto, toda relación entre dos variables que no responde a una relación de proporcionalidad como la anterior será no lineal. Es fácil ver que la mayoría de los sistemas que se encuentran en la naturaleza son no lineales y, de hecho, se han encontrado fraccionalidades, es por ello que, las derivadas de orden fraccionario, nos brindan un número infinito de grados de libertad para predecir el comportamiento futuro de las cantidades deseadas, como ser la concentración $C(t)$ de un reactivo con respecto al tiempo [4].

Fue hasta en 1921 que se publicó el primer trabajo sobre una reacción oscilante, y fue desarrollado por Willian C. Bray, quien mediante una reacción cuyo producto es la descomposición catalítica del peróxido de hidrógeno bajo la influencia de iones yodito, presentaba oscilaciones en cuanto al cambio de color [6]. La respuesta de muchos químicos fue de probar que la causa de las oscilaciones se debía a impurezas heterogéneas desconocidas. Esto hizo pensar a la mayoría de los químicos de la época que las reacciones oscilantes eran imposibles. Sin embargo, en las décadas de los 30 a los 50 se logró establecer que no había diferencia entre sistemas homogéneos y heterogéneos en cuanto a comportamientos oscilantes. Y hasta el año de 1973 el

problema de las reacciones químicas oscilantes con el ion yodato fue explicado con los trabajos de Briggs y Rauscher [7]. Las implicaciones de estos descubrimientos experimentales han sido llevadas al plano de la modelación matemática, en la que destaca la reacción Belousov-Zhabotinsky o reacción de B-Z.

Después de la reacción de B-Z, han sido reportadas un número importante de reacciones químicas y bioquímicas, especialmente en oxidación orgánica: una de ellas es la oxidación de p-xileno a ácido tereftálico; esta reacción es caótica, a menudo precedida por oscilaciones regulares; también existen reacciones de oxidación en catálisis heterogénea, procesos de oxidación con O_2 electro-disolución de metales, estas generalmente presentan oscilaciones irregulares; asimismo hay reacciones enzimáticas oxidasa-peroxidasa, polimerización de radical libre del vinil acetato que han permitido descubrir la presencia de caos por doblamiento de período e histéresis, etc. [2].

Asimismo, las aplicaciones del cálculo fraccionario se han empleado en muchos campos de la química aplicada. Por ejemplo, Kashif y colaboradores [8] desarrollaron una investigación utilizando operadores diferenciales de orden fraccionario llamados Caputo-Fabrizio y Atangana-Baleanu, en la que analizaron la distribución de temperatura para diferentes tipos de nanopartículas suspendidas en etilenglicol, cobre y óxido de aluminio. Kashif y Aguilar [9] llevaron a cabo un análisis comparativo del modelo de alcohol en sangre utilizando también operadores fraccionarios Caputo-Fabrizio y Atangana-Baleanu. Otro ejemplo donde estos operadores fraccionarios fueron utilizados se percibe en Qureshi y colaboradores [10], quienes investigaron las ecuaciones de gobierno utilizadas en el modelado de la cantidad de alcohol en sangre humana en cualquier momento.

Continuando con los últimos avances del cálculo fraccionario en la química, Qureshi y Atangana [11] propusieron tres nuevos modelos no clásicos para el dengue y mostraron su mejor desempeño en comparación con el modelo clásico. También validaron sus nuevos modelos a través de datos reales. Qureshi y Yusuf [12,13] analizaron modelos no clásicos de epidemiología y al presentar sus datos se encontró que tenían una mejor eficiencia que el modelo clásico existente. Yusuf y colaboradores [14] investigaron la versión no clásica del modelo epidémico de dos cepas y presentó la singularidad de su solución, mientras que Atangana y Bonyah [15] presentaron un enfoque estocástico no clásico para capturar la heterogeneidad entre varios fenómenos naturales y físicos. Bonyah y colaboradores [16] presentaron tres nuevos enfoques basados en operadores no clásicos para modelar diferentes tipos de atractores caóticos no encontrados antes en la literatura disponible. Se pueden consultar otras aplicaciones del mundo real del cálculo fraccionario incluidas en [17] para familiarizarse con este nuevo campo de la ciencia matemática emergente. Además, no sólo se investigan las ecuaciones diferenciales ordinarias bajo el dominio del cálculo fraccional, sino que también se han estudiado varios modelos matemáticos interesantes basados en ecuaciones diferenciales parciales, ecuaciones integrales y ecuaciones integrodiferenciales para obtener nuevos resultados [18].

4. CONCEPTOS PRELIMINARES

Se introduce en esta sección algunos conceptos básicos del cálculo fraccionario, mismos que serán utilizados en secciones posteriores.

Definición 1. (*Integral fraccionaria de Riemann-Liouville*) Sea $\alpha > 0$. Se define la integral fraccionaria de Riemann-Liouville para una función $y(t)$ de orden α como:

$$(4.1) \quad J_{0,t}^{\alpha}y(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-w)^{\alpha-1}y(w)dw, \quad t > 0$$

donde Γ es la función gamma definida por $\Gamma(z) = \int_0^{\infty} x^{z-1}e^{-x}dx$.

Definición 2. (*Derivada fraccionaria de Riemann-Liouville*) Sea $\alpha > 0$. Se define la derivada fraccionaria de Riemann-Liouville para una función $y(t)$ de orden α como:

$$(4.2) \quad D_{0,t}^{\alpha}y(t) = \frac{1}{\Gamma(n-\alpha)} \frac{d^n}{dt^n} \int_0^t (t-w)^{n-\alpha-1}y(w)dw, \quad t > 0$$

donde $(n-1) \leq \alpha < n \in \mathbb{N}$.

Definición 3. (*Derivada fraccionaria de Caputo*) Sea $\alpha > 0$. Se define la derivada fraccionaria de Caputo para una función $y(t)$ de orden α como:

$$(4.3) \quad D_{0,t}^{\alpha}y(t) = \frac{1}{\Gamma(n-\alpha)} \int_0^t (t-w)^{n-\alpha-1}y^{(n)}(w)dw, \quad t > 0$$

Propiedades del operador de derivada de orden fraccionario:

- (i) $D^0 f(t) = f(t)$
- (ii) $D^n f(t) = \frac{d^n f(t)}{dt^n}$ cuando n es un entero positivo.
- (iii) $D^{\alpha} D^{\beta} f(t) = D^{\alpha+\beta} f(t)$

donde $(n-1) < \alpha \leq n \in \mathbb{N}$, Las derivadas fraccionarias de Riemann-Liouville y Caputo están relacionadas por:

$$(4.4) \quad D_{0,t}^{\alpha}y(t) = D_{0,t}^{\alpha}y(t) - \sum_{k=0}^{n-1} \frac{y^{(k)}(c)(t-c)^{k-\alpha}}{\Gamma(k-\alpha+1)},$$

donde $(n-1) < \alpha < n \in \mathbb{N}$, $y(t) \in C^{n-1}[c, t]$ y $y^{(n)}$ es una función integrable en $[c, t]$. También vale la pena señalar que la derivada fraccionaria de Caputo se prefiere a la de Riemann-Liouville en problemas aplicados por la razón de tener orden de números enteros en condiciones iniciales suministradas con el primero [4].

Definición 4. La transformada de Laplace para el operador derivada fraccional de Caputo de orden real arbitrario $\alpha > 0$ de una función $y(t)$ es la siguiente [4]:

$$(4.5) \quad \mathcal{L}[D_{0,t}^\alpha y(t)] = s^\alpha \mathcal{L}[y(t)] - \sum_{k=0}^{n-1} s^{\alpha-k-1} y^{(k)}(0).$$

Las derivadas de orden fraccionario tanto de Caputo como de Riemann-Liouville son útiles en este trabajo ya que para convertir la cinética de la reacción, representada por un sistema dinámico de orden entero, se tiene que derivar el modelo original, usando cualquiera de las dos derivadas de orden fraccionario, además de sus propiedades para poder obtener el sistema dinámico de orden fraccionario aproximado. La transformada de Laplace de orden fraccionario se utilizará en el futuro para poder obtener una sistema de orden fraccionario exacto.

5. MODELADO MATEMÁTICO DE LA REACCIÓN CINÉTICA CASO LINEAL

Las velocidades de las reacciones químicas se investigan generalmente bajo el dominio de la cinética química, en la que el tratamiento matemático de esta cinética se logra mediante el uso de modelos matemáticos basados en ecuaciones diferenciales sujetas a algunas condiciones iniciales o de frontera. El interés está orientado a determinar los perfiles de concentración dependientes del tiempo de las diferentes especies que reaccionan entre si. En aras de la simplicidad, asumiremos que no hay participación de reactivos intermedios, entonces la reacción avanza solo un paso de la siguiente manera:



donde k_1 se conoce como la constante de velocidad de la reacción que brinda información sobre la lentitud o rapidez de la reacción, esta constante puede depender de la reacción en sí y de otros parámetros ambientales como la temperatura, pero no depende de las concentraciones de las especies que participan en la reacción química; por otro lado, si la especie a se transforma en un producto c a través de una especie intermedia b , entonces este tipo de mecanismo de transformación secuencial está representado por:



Como la velocidad de la primera reacción no depende de la concentración de la especie b , entonces:

$$(5.3) \quad C_a(t) = -k_1 C_a(t).$$

por el contrario, C_b estará dado por:

$$(5.4) \quad C_b(t) = k_1 C_a(t) - k_2 C_b(t).$$

por último $C_c(t)$ está dado por:

$$(5.5) \quad C_c(t) = k_2 C_b(t).$$

El modelado anterior conduce al siguiente sistema dinámico lineal autónomo homogéneo de ecuaciones gobernantes, sujeto a las siguientes condiciones iniciales:

$$(5.6) \quad \frac{d}{dt}C(t) = AC(t), \quad C(t=0) = C_0,$$

al mismo tiempo el sistema dinámico que se obtiene es el siguiente:

$$(5.7) \quad \begin{aligned} C_a(t) &= -k_1 C_a(t), & C_a(t=0) &= 1 \\ C_b(t) &= k_1 C_a(t) - k_2 C_b(t), & C_b(t=0) &= 0 \\ C_c(t) &= k_2 C_b(t), & C_c(t=0) &= 0. \end{aligned}$$

y por el balance de masa $C_a + C_b + C_c = 1$, luego se obtiene:

$$(5.8) \quad A = \begin{bmatrix} -k_1 & 0 & 0 \\ k_1 & -k_2 & 0 \\ 0 & k_2 & 0 \end{bmatrix}.$$

Obteniendo los tres eigenvalores $\lambda_1 = 0$, $\lambda_2 = -k_1$, $\lambda_3 = -k_2$. Hay un eigenvalor igual a cero por ser A una matriz singular. Sin embargo, los otros dos eigenvalores negativos sugieren que la solución se acerca al punto de equilibrio a lo largo de las líneas paralelas a sus respectivos eigenvectores asociados. También se puede observar que la suma de todas las ecuaciones en (5.7) es constante e igual a 0. Por esta razón se debe utilizar el operador fraccionario de tipo de Caputo, ya que la derivada de Caputo de una función constante es 0.

Luego aplicando el operador fraccionario de tipo Caputo al sistema (5.7) para fraccionalizarlo, se obtiene el siguiente sistema:

$$(5.9) \quad \begin{aligned} c\left[\frac{d^\alpha}{dt^\alpha}C_a(t)\right] &= \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-w)^{-\alpha} C'_a(w)dw = -k_1^\alpha C_a(t), & C_a(t=0) &= 1 \\ c\left[\frac{d^\alpha}{dt^\alpha}C_b(t)\right] &= \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-w)^{-\alpha} C'_b(w)dw = k_1^\alpha C_a(t) - k_2^\alpha C_b(t), & C_b(t=0) &= 0 \\ c\left[\frac{d^\alpha}{dt^\alpha}C_c(t)\right] &= \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-w)^{-\alpha} C'_c(w)dw = k_2^\alpha C_b(t), & C_c(t=0) &= 0. \end{aligned}$$

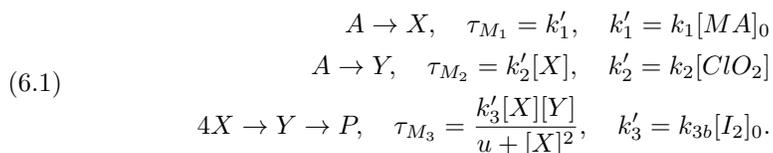
Luego, para resolverlo se aplica la transformada de Laplace y su inversa para el operador fraccional de Caputo en (5.7), esto genera las siguientes soluciones [5]:

$$\begin{aligned}
 C_a(t) &= e^{(k_1 t)}, \\
 C_b(t) &= \frac{k_1}{k_1 - k_2} (e^{(-k_2)t} - e^{(-k_1)t}), \\
 C_c(t) &= \frac{1}{k_1 - k_2} (k_2 e^{(-k_1)t} - k_1 e^{(-k_2)t} + 1).
 \end{aligned}
 \tag{5.10}$$

las soluciones anteriores muestran cómo evolucionan las especies reactantes a , b y c con respecto al tiempo para el caso lineal, en la próxima sección se verá la modulación para el caso no lineal.

6. MODELADO MATEMÁTICO DE LA REACCIÓN CINÉTICA CASO NO-LINEAL

Para el caso que nos compete tomaremos como punto de partida el modelo cinético ya existente de la reacción B-Z, que se desarrolla de la siguiente forma:



Utilizando las siguientes transformaciones $X = ax$, $Y = by$, $\tau = ct$, $a = \sqrt{u}$, $b = \frac{uk'_2}{k_3}, \frac{1}{k_2}$, $\gamma = \frac{ck'_1}{a}$ y $\beta = \frac{a}{b}$. Obtenemos el siguiente sistema bidimensional [19]:

$$\begin{aligned}
 \frac{dx}{dt} &= \gamma - x - \frac{4xy}{1 + x^2} \\
 \frac{dy}{dt} &= \beta x \left(1 - \frac{y}{1 + x^2}\right),
 \end{aligned}
 \tag{6.2}$$

donde:

$x, y \mapsto$ determinan las concentraciones de las especies que reaccionan inicialmente,
 $\gamma \mapsto$ parámetro cinético,
 $\beta \mapsto$ parámetro cinético.

para este sistema dinámico se obtiene un único punto positivo de equilibrio y es:

$$P = \left(\frac{\gamma}{5}, 1 + \frac{\gamma^2}{25} \right),$$

resultando la siguiente derivada fraccionaria de Caputo para el sistema:

$$(6.3) \quad \begin{aligned} D^{\alpha-1}\dot{x} &= D^{\alpha-1}\left(\gamma - x - \frac{4xy}{1+x^2}\right), \\ D^{\alpha-1}\dot{y} &= D^{\alpha-1}\left(\beta x\left(1 - \frac{y}{1+x^2}\right)\right), \end{aligned}$$

Debido a la propiedad del operador de la derivada de orden fraccionario se convierte en:

$$(6.4) \quad \begin{aligned} D^{\alpha}x &= D^{\alpha-1}\left(\gamma - x - \frac{4xy}{1+x^2}\right), \\ D^{\alpha}y &= D^{\alpha-1}\left(\beta x\left(1 - \frac{y}{1+x^2}\right)\right), \end{aligned}$$

o en su equivalente:

$$(6.5) \quad \begin{aligned} D^{\alpha}x &= \int D^{\alpha}\left(\gamma - x - \frac{4xy}{1+x^2}\right)dt, \\ D^{\alpha}y &= \int D^{\alpha}\left(\beta x\left(1 - \frac{y}{1+x^2}\right)\right)dt, \end{aligned}$$

Tomando en cuenta que $D^{\alpha}x \simeq \frac{dx}{dt}$ para un $0 \leq \alpha \leq 1$, la solución aproximada del sistema de orden fraccionario es:

$$(6.6) \quad \begin{aligned} D^{\alpha}x &= \int D^{\alpha}\left(\gamma - x - \frac{4xy}{1+x^2}\right)dt \simeq \gamma - x - \frac{4xy}{1+x^2}, \\ D^{\alpha}y &= \int D^{\alpha}\left(\beta x\left(1 - \frac{y}{1+x^2}\right)\right)dt \simeq \beta x\left(1 - \frac{y}{1+x^2}\right), \end{aligned}$$

la cual, en su forma reducida se expresa como:

$$(6.7) \quad \begin{aligned} D^{\alpha}x &= \left(\gamma - x - \frac{4xy}{1+x^2}\right), \\ D^{\alpha}y &= \left(\beta x\left(1 - \frac{y}{1+x^2}\right)\right), \end{aligned}$$

Cercano ambos modelos al punto de equilibrio, se utilizó el algoritmo FO-Solver, en el Software Octave, el cual presenta el comportamiento caótico mostrado en las figuras 1, 2 y 3. En la figura 4 se muestra el diagrama de bifurcación en el tiempo para la especie reactante x .

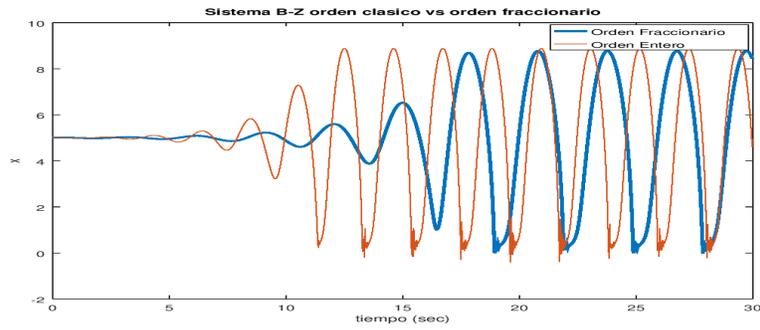


FIGURE 1. Evolución en el tiempo para la especie reactante x con $\gamma = 25$ y un $\beta = 10$

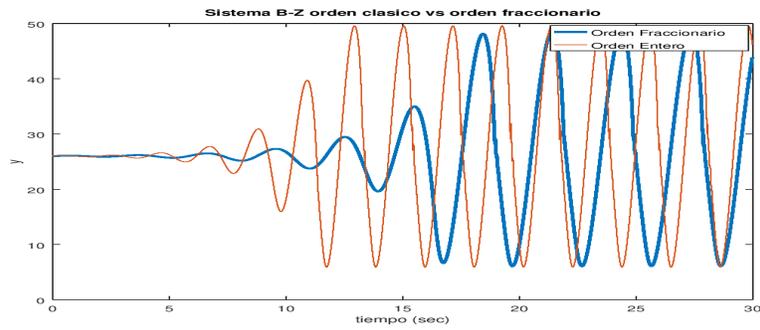


FIGURE 2. Evolución en el tiempo para la especie reactante y con $\gamma = 25$ y un $\beta = 10$

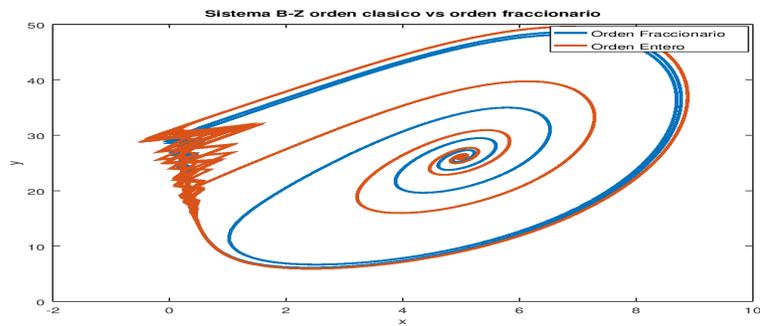


FIGURE 3. Retrato de fase para ambos sistemas con un $\gamma = 25$ y un $\beta = 10$

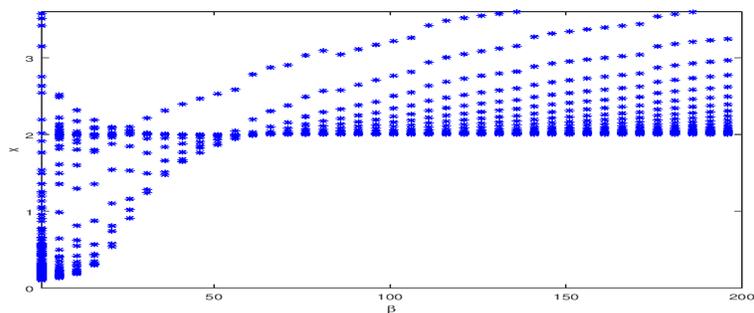


FIGURE 4. Bifurcación en el tiempo para la especie reactante x con $\gamma = 25$.

Tanto en la figura 1 como en la figura 2 se puede ver la evolución en el tiempo para la especie reactante x , donde se aprecia la comparación de la respuesta del sistema de orden fraccionario en comparación con la respuesta del sistema de orden entero. Al poder modelar la cinética de una reacción química, los parámetros de la misma, así como el de orden fraccionario, se puede aproximar más a la representación del sistema químico en la realidad que con un sistema dinámico de orden entero.

Se puede ver en la figura 3 el comportamiento caótico de las variables reactantes x y y . Se puede notar como las oscilaciones generan un ciclo límite inestable cuyo dominio de atracción es el punto de equilibrio del sistema. Este ciclo límite inestable puede causar resultados adversos cuando se realice un experimento de laboratorio, ya que la población de ambas especies podrían colapsar.

Las regiones de la figura 4 de la gráfica de bifurcación muestran dónde el sistema dinámico entra en régimen caótico, básicamente cuando $\beta \simeq 50$ se entra en una transición de caos.

7. CONCLUSIONES Y TRABAJOS FUTUROS

En el retrato de fase se pudo observar que el dominio de la atracción de las variables de estado tiende al único punto de equilibrio, por lo que se concluye que es un atractor autoexcitado (self-excited).

En este trabajo se realiza una presentación de las aplicaciones del cálculo fraccionario a los sistemas dinámicos derivados de la cinética de reacciones químicas, en el cual se implementaron las propiedades del cálculo de orden fraccionario para convertir el sistema de orden entero a orden fraccionario aproximado, mostrando una comparación entre el modelo clásico y el modelo fraccionario derivado del modelo clásico.

En el futuro se pretende resolver el sistema dinámico de manera exacta, por medio del siguiente procedimiento: linealizar por series de Taylor, aproximar la parte no lineal por funciones de Bessel y resolver el sistema con aproximaciones

numéricas, creando las rutinas específicas en Octave y/o Matlab, para la solución del sistema, para luego compararlo con la solución teórica. Además, en el futuro se pretende realizar un análisis dinámico más exhaustivo en el que se deberá incluir los exponentes de Lyapunov y Kaplan-Yorke, de modo que estos sirvan para la obtención de un supresor de caos por retroalimentación de estados.

En este artículo se demuestra tanto analítica como numéricamente las ventajas de los modelos o sistemas dinámicos de orden fraccionario en comparación de los sistemas de orden entero, ya que con dichos modelos se puede representar con más fidelidad muchos sistemas físicos que se encuentran tanto en las ciencias como en las ingenierías; por ejemplo, los sistemas dinámicos de orden fraccionario nos sirven para poder modelar problemas de la química, física, biología, sistemas mecánicos, sistemas eléctricos y de energía, sistemas económicos, sistemas sociales, etc. Por esta razón los resultados presentados en este reporte proveen evidencia de cómo la matemática aplicada pueden ser empleada tanto en la academia así como en la investigación y desarrollo.

REFERENCES

1. Yan, X., Du, W., Quian, F. (2004). *Development of a Kinetic Model for Industrial Oxidation of p-Xylene by RBF-PLS and CCA*. Shanghai, China: Automation Institute, East China University of Science and Technology.
2. Aguirre, J., Chejne, F., Daro lvarez, H. (2004). *Una aproximación a los comportamientos caóticos en procesos químicos*. Dyna, vol. 71, 111-124.
3. Khalil, I., Teoman Naskali, A., Sabanovic, Asif S. (2009). *On Fraction Order Modeling and Control of Dynamical Systems. IFAC 24956*. Turkey Turkey: Faculty of Engineering and Natural Science, Sabanci University.
4. Qureshi, S., Aziz, S., (2019). *Fractional modeling for a chemical kinetic reaction in a batch reactor via nonlocal operator with power law kernel*. Jamshoro 76062. Sindh, Pakistan: Department of Basic Sciences and Related Studies, Mehran Iniversity of Engineering and Technology.
5. Dinámica química. (2019). Wikipedia, La enciclopedia libre. es.wikipedia.org/w/index.php?title=Din%C3%A1mica_qu%C3%ADmica&oldid=120760227.
6. Bray, W. C., (1921). *Periodic reaction in homogeneous solution and its relation to catalysis*. J. Am. Chem. Soc., Vol. 43, p. 1262-1267. doi.org/10.1021/ja01439a007.
7. Briggs, T. S., Rauscher, W. C. (1973). *An Oscillating Iodine Clock*. J. Chem. Ed. 50, p. 496. doi.org/10.1021/ed050p496.
8. Abro, K.A., Chandio, A.D., Abro, I.A., Khan, I. (2019). *Dual thermal analysis of magneto-hydrodynamic flow of nanofluids via modern approaches of Caputo-Fabrizio and Atangana-Baleanu fractional derivatives embedded in porous medium*, J. Therm. Anal. Calorim. 135 (4): 21972207.
9. Abro, K.A., Gómez-Aguilar, J.F. (2019). *Analysis of blood alcohol model via non-integer order derivatives, in: Fractional Derivatives with Mittag-Leffler Kernel*. Springer, Cham, 6979.
10. Qureshi, S., Yusuf, A., Shaikh, A., Inc, M., Baleanu, D. (2019). *Fractional modeling of blood ethanol concentration system with real data application*. Chaos 29 (1) 013143.
11. Qureshi, S., Atangana, A. (2019). *Mathematical analysis of dengue fever outbreak by novel fractional operators with field data*. Physica A 121127.
12. Qureshi, S., Yusuf, A. (2019). *Fractional derivatives applied to MSEIR problems: Comparative study with real world data*. Eur. Phys. J. Plus 134 (4): 171.
13. S. Qureshi, S., Yusuf, A. (2019). *Modeling chickenpox disease with fractional derivatives: From caputo to atangana-baleanu*. Chaos Solitons Fractals 122: 111118.
14. Yusuf, A., Qureshi, S., Inc, M., Aliyu, A.I., Baleanu, D., Shaikh, A.A. (2018). *Two-strain epidemic model involving fractional derivative with Mittag-Leffler kernel*. Chaos 28 (12): 123121.
15. Atangana, A., Bonyah, E. (2019). *Fractional stochastic modeling: New approach to capture more heterogeneity*. Chaos 29 (1): 013118.

16. Bonyah, E., Atangana, A., Elsadany, A.A. (2019). *A fractional model for predatorprey with omnivore*. Chaos 29 (1): 013136.
17. Atangana, A., Qureshi, S. (2019). *Modeling attractors of chaotic dynamical systems with fractal-fractional operators*. Chaos Solitons Fractals 123: 320337.
18. Abu Arqub, O., Al-Smadi, M. (2018). *Numerical algorithm for solving time-fractional partial integrodifferential equations subject to initial and Dirichlet boundary conditions*. Numer. Methods Partial Differential Equations 34 (5): 15771597.
19. Din, Q., Donchev, T., & Kolev, D. (2018). *Stability, bifurcation analysis and chaos control in chlorine dioxideiodinemalonic acid reaction*. NMATCH Commun. Math. Comput. Chem, 79(3), 577-606.

ESCUELA DE MATEMÁTICA, UNIVERSIDAD NACIONAL AUNTÓNOMA DE HONDURAS
Current address: Departamento de Matemática Aplicada, Facultad de Ciencias, Universidad Nacional Autónoma de Honduras, Ciudad Universitaria
E-mail address: ceriex@gmail.com

CRITICIDAD AUTO-ORGANIZADA “EL MODELO DE LA PILA DE ARENA”

NORMAN R. SABILLON

RESUMEN. Este artículo introduce el significado de SOC (criticidad autoorganizada). SOC es un método de sistemas interactivos a gran escala que se aproxima a los fenómenos naturales, anteriormente no tenían una explicación secuencial. El modelo trata de cómo integrar la teoría, a lo físico. Al introducir conocimientos importantes sobre el prototipo científico, comenzaremos a estudiar el “modelo de pila de arena” y veremos cada uno de los estados, hasta que alcanza un punto de equilibrio. Su composición y configuración también se basará en sus reglas básicas, la existencia de todo el modelo y sus definiciones de recepción grupal, que nos permiten ver muchas características del sistema tema.

Abstract. This article introduces the meaning of SOC (self-organized criticality). SOC is a method of large-scale interacting systems approach to natural phenomena, which previously had no sequential explanation. The model deals with how to integrate theory into the physical. By introducing important knowledge about the scientific prototype, we will begin to study the "sand pile model." and look at each of the states, until it reaches a point of equilibrium. Its composition and configuration will also be based on its basic rules, the existence of the whole model and its definitions of group reception, which allow us to see many features of the subject system.

1. INTRODUCCIÓN

Este artículo estudia la teoría de la criticidad autoorganizada, que tiene como objetivo explicar fenómenos naturales que antes no podían ser analizados. En la primera parte, comprenderemos algo de la descripción de la teoría. Esta es una historia nueva, pero por su Desarrollo rápido y forma especial: Escalar fenómenos muy grandes en el eje del espacio y el eje del tiempo para encontrar alguna relación entre los dos ejes. Por tal razón, muchas ramas de la ciencia han decidido adoptar esta técnica con el fin de observar el comportamiento de fenómenos físicos importantes.

Se recorrerá la labor de ciertos actores importantes de esta teoría, sus trabajos y en qué rama de la ciencia se aplicaron, teniendo la consciencia de que existe muchos más. De autores que se dedicaron a examinar sobre esta línea de investigación.

En la Sección 3, veremos cuál es la definición y características de la Pila de Arena Abeliana. Se descubre los estados que atraviesa el sistema, se definirá las reglas más importantes de la teoría. Aunque las normas son simples, deben basarse en nuestro modelo. Qué realicen cambios estratégicos al sistema original.

Date: 4 de mayo de 2021.

Key words and phrases. SOC, Pilas de Arenas, Derrumbe.

En la cuarta parte se explica el modelo paso a paso, se precisa el subconjunto y las condiciones que se deben cumplir, cuál es el entorno de disipación, y también se define una de las normas básicas (regla de avalancha) de los elementos del sistema.

Se observan los conceptos básicos para comprender el término configuración y también se presenta la regla de la avalancha y dos grupos necesarios, la especificación de teoremas y lemas. La búsqueda de ejemplos que sean útiles para implementar cualquier idea definida.

2. JUSTIFICACIÓN

Una de las principales razones para estudiar este tema es que diversos fenómenos naturales pueden modelarse a gran escala, y por esta razón es que forma parte de la dirección de investigación de la maestría en matemáticas aplicadas de la UNAH, en el apartado 5 “Celular Automata” y en el tema del desarrollo de la energía en el tema prioritario número 13.

Este tipo de modelo se utiliza para el análisis 2D en una red de transmisión de energía. En la red, el efecto dominó aumentará por pequeñas perturbaciones en el estado crítico y pequeños eventos afectarán cualquier elemento del sistema. En abundantes casos, no se permite el estado de equilibrio. Una cosa que necesita especial atención es que muchas veces no se habla de agentes que no existen fuera del sistema. Existe el peligro de autoorganización en la red eléctrica [2, 13], ha habido décadas de investigación sobre la vulnerabilidad energética en las redes de corte de energía, este es un tema global, es muy importante poder controlar la red de transmisión que evite fallas de energía. [10, 17, 19, 29].

En las redes eléctricas también existen la modelación de la pila de arenas con el propósito de ver la evolución simultánea y a su vez afirmar la viabilidad de estudios de fallas eléctricas por simulación de SOC [30], una de las características de los sistemas SOC es pasar de un estado no crítico por un estado crítico por la autoorganización, lo cual se ha convertido en objeto de estudio.

3. ANTECEDENTES

Con el surgimiento de una teoría, la teoría intenta explicar cómo sistemas complejos con muchos grados de libertad observados en la naturaleza pueden generar relaciones de *ley de potencia* de acuerdo con sus patrones de estructura de distribución o reglas simples. Está relacionado con sus vecinos más cercanos, y esta idea apareció a mediados de los años ochenta.

El principal motivo de esta teoría es el escalamiento de fenómenos demasiados grandes, para llevarlo a un eje temporal y espacial, observado en una amplia gama de sistemas disipativos con enormes grados de libertad.

El concepto tardío de Per Bak de *Criticidad Auto-organizada (SOC)*, establecido por primera vez en sus artículos seminales por Tang y Wiesenfeld [3, 4].

En 1987, Per Bak y sus colaboradores presentaron un modelo que evoluciona hacia un estado crítico sin ningún ajuste externo. Este modelo a menudo se denomina

modelo de arena de Per Bak o el modelo de Bak-Tang-Wiesenfeld (BTW).

En primer lugar, la motivación de BTW para presentar la idea de SOC, que se expuso con mayor claridad en [11], la cual expone una representación de un modelo simple de la evolución de un sistema dinámico que evoluciona a una estructura sin escala. Los autores consideraban que esa no era la mejor representación ni la más realista pero que si era un modelo que funcionaba para el comportamiento de escalamiento, esto debido al surgimiento de los sistemas dinámicos disipativos extendidos. Este estudio fue llevado a cabo por la razón de que el universo tiene una distribución fractal [11].

Derivado de estos estudios se logró identificar que la teoría de SOC tiene las siguientes características.

- Los eventos son estadísticamente independiente, espacial y temporal;
- El tamaño o la distribución de frecuencia de ocurrencia es libre de escala y se puede caracterizar por una función de la ley de potencia en algún rango de tamaño;
- La evolución espacial y temporal detallada es compleja e implica una geometría fractal y características de tiempo fluctuantes estocásticas.

Después de todos estos descubrimientos, comenzó la investigación sobre terremotos. Los autores Carlson y Langer [8], al considerar el comportamiento de un modelo mecánico simple, en el lugar donde la característica no lineal es el deslizamiento entre las placas, también en su otro artículo [7], mostró cómo el elemento no lineal del modelo es la fricción deslizante entre el bloque y la superficie, el espectro de la ley de potencia generado por la simulación es similar al observado en la naturaleza. Estos artículos proporcionan importantes avances para la investigación sísmica, lo que nos permite profundizar en este campo.

En 1992, el modelo SOC de incendios forestales diseñado por los autores Turcotte y Malamud [26] comenzó a usarse en fenómenos que antes no tenían modelo. Bak [22] en su artículo probó la posibilidad de alcanzar un potencial punto de equilibrio dinámico, con la expresión de parámetros de control y la posibilidad de evolución dinámica. En otro artículo coeditado con Tang y Weisenfeld, revisaron el concepto de SOC y su posible aplicación en la investigación del transporte en plasma confinado magnéticamente [9].

El evento extremo que cambió la sociedad global se llama cisne negro. Estos se mencionan en la literatura de Taleb [23]. La distribución de frecuencia de muchos fenómenos naturales suele estar muy cerca de la distribución de la ley de potencia (fractal). El evento que supera estas inferencias se describe como el Rey Dragón, como lo califican los autores Sachs, Yoder, Turcotte, Rundle y Malamud. Este artículo considera eventos extremos como terremotos, erupciones volcánicas, incendios forestales, deslizamientos de tierra e inundaciones [21].

También se han desarrollado los temas en torno a los fenómenos en el espacio, la corona solar es una región muy dinámica que es la fuente de muchos fenómenos (por ejemplo, brotes solares, eyecciones de masa coronal, eventos de partículas energéticas solares), los brotes solares (incluidos los micro y nanoaflores) son uno de los ejemplos más convincentes del comportamiento de tipo SOC.

Unos de los fenómenos en el espacio más estudiados son las flores estelares, las distribuciones acumulativas y diferenciales de la tasa de destellos se construyeron y se ajustaron con diferentes métodos. Se encuentra que las leyes de poder describen aproximadamente las distribuciones descritas por Audard, Güdel, Drake y Vinay L. Kashyap [1]. Muchos de estos fenómenos pueden ser descritos por un modelo generalizado *FD-SOC*, “*Criticidad Auto-Organizada Fractal-Difusa*”.

Sachs en 2012, encuentra que las distribuciones de tamaño de frecuencia global de terremotos y erupciones volcánicas exhiben un comportamiento de ley de potencia para tamaños pequeños [24, 25]. SOC en las ciencias ambientales [6, 27], donde se ve la tierra como un sistema complejo donde muchos componentes biológicos y físicos interactúan tanto en espacio y tiempo, también se muestra la compleja evolución espacio-temporal de los cambios climáticos a escalas locales, tratando de reducir el complejo sistema que representa la tierra.

Existe más literatura sobre el desarrollo de este tema, que cubre una amplia gama de subtemas y muchos campos de interés en la investigación de fenómenos. Hay libros dedicados a este método, que muestran fenómenos y modelos que simulan el SOC. Esta es una literatura específica sobre SOC, comenzando con Bak (1996), Jensen (1998), Turcotte (1999), Charbonneau y otros. (2001), Hergarten (2002), Sornette (2004), Christensen y Moloney (2005), Aschwanden (2011), Crosby (2011) y Pruessner (2012) al final.



FIGURA 1. Estos son los temas sobre la teoría SOC de mayor consideración

4. CARACTERÍSTICAS DEL MODELO DE PILAS DE ARENAS

Consideremos a \mathbb{Z}^2 hay números enteros en cada vecindario o celda, iniciando una breve descripción de lo que es un modelo de pilas de arenas Abelian, primero se define el umbral que se convertirá en la regla más simple, donde si al menos tiene cuatro partículas, se permite derribar de la celda (x,y) , lo cual permite eliminar 4 partículas de esa vecindad y se agregan a las celda mas cercanas a (x,y) donde se repartirán 1 a $(x,y-1)$, $(x,y+1)$, $(x-1,y)$ y $(x+1,y)$, esto hasta que ninguna celda contenga no más de 3 partículas.

Las celdas con menos de 4 células se llaman estables y todo el proceso se llama estabilización. Este modelo es un autómata celular (Clase IV)¹ con regla de transición simétrica, lo cual se observa más claro en la Figura 2 y la Figura 3.

¹Es una combinación entre la clase II y la Clase III, donde el llenado del sistema es por secuencias caóticas (Complejas)[28]

La Figura 3 es una representación unidimensional de lo que sucederá en la Figura 2 que es la representación geométrica del \mathbb{Z}^2 .

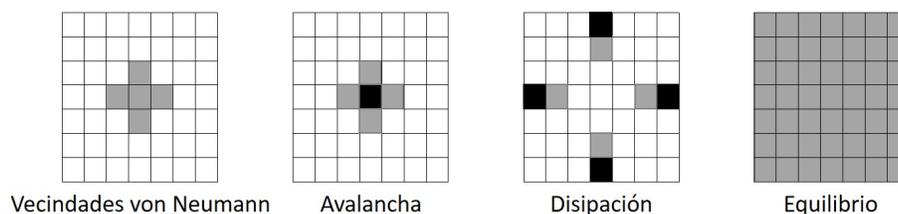


FIGURA 2. Estados de la pila de arena Abeliano en 2D

Se define cada estado de nuestra pila de arena, lo más sencillo posible:

- **Vecindades von Neumann:** La organización de nuestro sistema en vecindades o células, que van de izquierda a derecha, desde arriba hasta abajo.
- **Avalancha:** También conocida como activo o propagación, distribuye las unidades (granos o partículas) a las vecindades más cercanas de forma independiente, de modo que cada uno de los estado de las vecindades reciba una unidad.
- **Disipación:** Son las unidades que se pierden en los límites, donde una de las vecindades más cercana se encuentra fuera de nuestro sistema y es por ese motivo es que se pierde la partícula.
- **Configuración final en equilibrio:** Es el estado alcanzado por nuestro sistema al no contener en ninguna de sus vecindades no más de 3 partículas, evitando así posibles avalanchas o disipaciones.
- **Progresión del tiempo:** Es el progreso en el tiempo donde todos las vecindades se actualizan en paralelo, cuando todos los sistemas activos se actualizan a la vez [18].

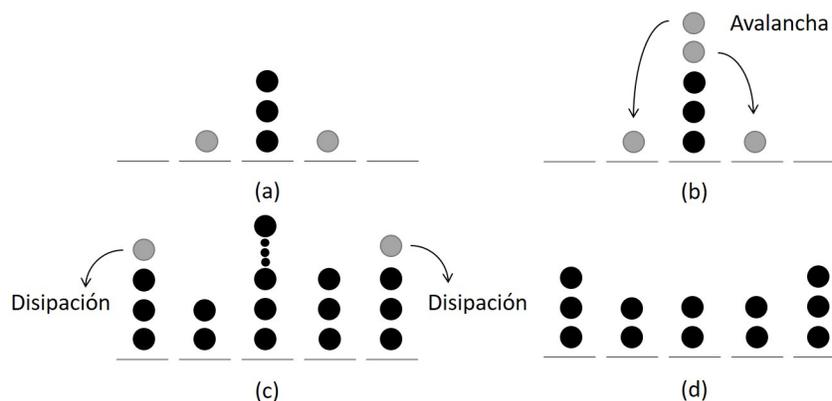


FIGURA 3. Estados de nuestra pila de arena, (a) Vecindades de von Neumann (b) Una Avalancha (c) Disipación (d) Configuración final en equilibrio

La regla más importante y simple dentro del sistema es el valor en el umbral, este es el que determina en un vecindario de von Neumann los estados de equilibrio.

Al momento en que una vecindad alcanza el umbral esta modifica su estado produciendo una avalancha, la cual puede propagarse a las vecindades si y solo si las vecindades también se encuentran en el umbral. Como consecuencia se da inicio a avalanchas en paralelo, por lo que a pesar de la simplicidad de la regla, puede producirse avalanchas de todos los tamaños en el sistema.

Se espera que, en la progresión del tiempo, se repliquen múltiples propagaciones en las vecindades, hasta que se alcance un equilibrio dinámico, dejando como resultado pilas de diferentes tamaños. Una de las formas sencillas de recordar la regla de avalancha es ver el número de vecindades o vértices conectados a nuestra vecindad de estudio, el número de partículas no debe exceder el umbral.

En este artículo se trabaja con $d = 2$, donde d viene de \mathbb{Z}^d con $d \geq 2$, también comprender que al hablar de “Abeliano”² es mencionar la *propiedad abeliana del modelo de pilas de arena*. Una de la característica esencial es la simetría y los sitios de disipación.

5. DESCRIPCIÓN DEL MODELO DE PILAS DE ARENAS

5.1. Composición del Modelo. Sea V un subconjunto finito de \mathbb{Z}^2 . Se dice que una matriz de valores enteros $\Delta V_{x,y}$, indexados por $x, y \in V$, es una matriz de derrumbamiento si cumple las siguientes condiciones:

- (1) $\forall x, y \in V, x \neq y, \Delta V_{x,y} = \Delta V_{y,x} \leq 0$,
- (2) $\forall x \in V, \Delta V_{x,x} \geq 1$,
- (3) $\forall x \in V, \sum_{y \in V} \Delta V_{y,x} \geq 0$,
- (4) $\sum_{x,y \in V} \Delta V_{x,y} \geq 0$.

La cuarta condición es sobre los límites donde ocurre la disipación, para las cuales la condición tercera es estricta. Esto es fundamental para tener la regla de la avalancha y que ΔV será las vecindades o celosías con otras condiciones, más explícitamente:

$$(5.1) \quad \Delta V = \begin{cases} \Delta V_{x,x} = 2d & \text{si } x \in V \\ \Delta V_{y,x} = -1 & \text{si } x, y \text{ son vecindades cercanas} \\ \Delta V_{y,x} = 0 & \text{en otro caso} \end{cases}$$

5.2. Configuración. Es un mapa natural es donde $V \in \mathbb{N}$, asignando a cada sitio un número natural $\eta(x) \geq 1$ (el número de partículas en el sitio x).

Una configuración $\eta \in \mathbb{N}^V$ es llamada estable si para todo $x \in V, \eta(x) \leq \Delta V_{x,x}$, de caso contrario es inestable. Se denota como Ω_V al conjunto de todas las celosías con una altura estable o en el umbral, y el numero máximo de Ω_V se va a denotar $\eta^{max}(x) = \Delta V_{x,x} \quad \forall x \in V$. Para $\eta \in \mathbb{N}^V$ y $V' \subset V, \eta | V'$ para denotar como la restricción de η a V' .

²Son aquellos donde se define una operación aditiva, Dhar llamó este modelo generalizado el Modelo Abeliano de Pilas de Arena (ASM por sus siglas en inglés: Abelian Sandpile Model). Mostró también una correspondencia 1-1 entre las configuraciones recurrentes del ASM [12].

5.3. Regla de Avalancha o Derrumbe. La avalancha en la matriz ΔV son las asignaciones T_x .

$$T_x : \mathbb{N}^V \rightarrow \mathbb{N}^V$$

Indexado por V , y definido por:

$$(5.2) \quad T_x((\eta)(x)) = \begin{cases} \eta(y) - \Delta_{x,y}^V & \text{si } \eta(x) > \Delta_{x,x}^V, \\ \eta(y) & \text{en otro caso.} \end{cases}$$

La regla de derrumbe conmuta entre configuraciones inestables. Lo que significa es que para $x, z \in V$ y η tal que $\eta(x) > \Delta_{x,x}^V$ y $\eta(z) > \Delta_{z,z}^V$.

$$(5.3) \quad T_x \circ T_z(\eta) = T_z \circ T_x(\eta)$$

Al elegir una enumeración x_1, \dots, x_n del conjunto V , la transformación que se derrumba es la asignación [20].

$$\tau_{\Delta^V} : \mathbb{N}^V \rightarrow \Omega_V$$

$$(5.4) \quad \tau_{\Delta^V}(\eta) = \prod_{i=1}^N T_{x_i}(\eta)$$

Ahora se define el grupo derrumbado, que es el conjunto de vecindades que fueron derrumbados al menos una vez durante la estabilización. Por otro lado, el grupo visitado, son las vecindades que recibieron una partícula después de un derrumbe, el grupo derrumbado y visitado durante la estabilización, están estrechamente relacionados entre sí.

Example 5.1. Imagina que el cálculo de ambos grupos ya existe con un número inicial n de partículas. Es importante pensar que una vez, al agregar más partículas al número original, está claro que ambos grupos crecerán también. Eso también debe ser probado.

Para poder comprender mejor como identificar estos dos grupos es necesario la definición de un radio, el cual es el conjunto de vecindades V y el radio máximo $\sup_{x,y \in V} \max(|x|, |y|) + 1$, si V esta vacío. El mayor de nuestro interés es el radio del grupo derrumbado de nuestra configuración inicial, su forma está determinado por el teorema 5.3, que fue probado por Fey-den Boer y Reding [15].

Theorem 5.2. *Para todo entero positivo $n \geq 4$ la configuración inicial con n partículas, en el origen y el nivel del suelo 2 termina y el derrumbamiento es un cuadro centrado cómo origen.*

El mejor límite conocido es $\sqrt{n} \geq r \geq 4/r$, sus extremos son triviales [14].

Theorem 5.3. *El grupo de derrumbamiento de las pilas de arenas Abelianas para n partículas en el origen y en el nivel de suelo 2 es un cuadrado, centrado que es el origen, y radio a mucho $4(1 + o(1))\sqrt{n}$.*

Lo mas difícil para demostrar este teorema es la elección para calcular el radio, y también el número de partículas en cada vecindad, como función explícita de n , pero la última configuración muestra ciertos patrones que se repiten para muchas de las vecindades (ver Figura 4).

Lemma 5.4. *El radio del grupo derrumbado de la pila de arena Abelianas para n partículas en el origen y al nivel suelo es a lo sumo $(1 + o(1))\sqrt{n}$.*

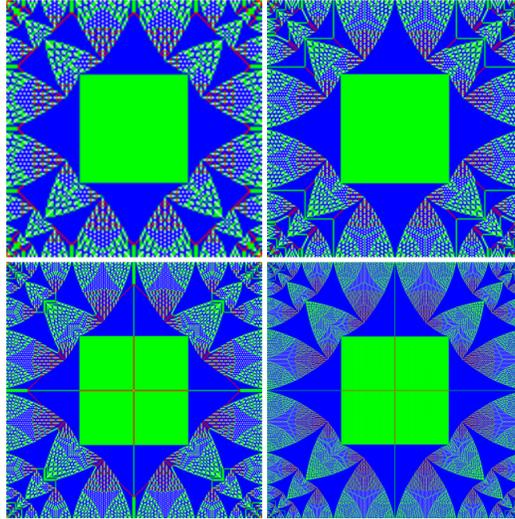


FIGURA 4. cuadrícula cuadrada para diferentes valores de N , $N = 128$ (arriba a la izquierda), 198 (arriba a la derecha), 243 (abajo a la izquierda) y 521 (abajo a la derecha). El esquema de color es el siguiente: naranja = 0 fichas, rojo = 1 ficha, verde = 2 fichas y azul = 3 fichas [16].

Lemma 5.5. *El cuadrado S_r lleno de 4 partículas en el nivel del suelo 2, crece a lo sumo de 2, es decir el grupo derrumbado de su estabilización tiene un radio de a lo sumo $2r$.*

Una vez que ya tenemos el lema 5.4 y lema 5.5 disponibles, demostramos el teorema 5.3.

5.4. **Parámetros.** [5] El comportamiento del modelo se analiza en función de diferentes parámetros, como:

- Tamaño de modelo,
- Número de dimensiones de vecindades,
- Conservación de masas, es decir, si el número de partículas extraídas de un derrumbe es igual a la suma de partículas que recibieron sus sitios vecinos,
- Condiciones de contorno.

Se pueden pensar diferentes tipos de condiciones de contorno:

- Abierto: si un sitio cercano a la frontera se derrumba, algunos de sus partículas abandonan el sistema (hay disipación),
- Cerrado: el sitio cercano a la frontera no se colapsa completamente, pero mantiene las partículas que se caerían en un caso abierto.
- Periódico: el sistema no tiene límites, es decir, el volcado cerca del borde está envuelto.
- Mezclado: por ejemplo, la celosía es periódica

6. EJEMPLOS

Example 6.1. En este primer ejemplo se utiliza un sistema de 5x5 con un recuento inicial de 16 partículas, donde se sabe que la vecindad tiene 4 cercanas que determinan el umbral del sistema. Una idea rápida para calcular el número de configuración es dividir el número de partículas iniciales por el umbral del sistema ($16/4 = 4$), sin embargo, se encuentra que a través del proceso de retroalimentación, nuestro grupo de colapso es 9 y el número posible de configuraciones varía desde un mínimo de 6 hasta un máximo de 9 configuraciones para alcanzar el equilibrio.

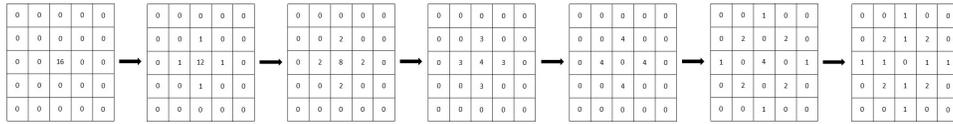


FIGURA 5. El modelo de pilas de arena de 5x5, con un configuración inicial de 16 partículas

Example 6.2.

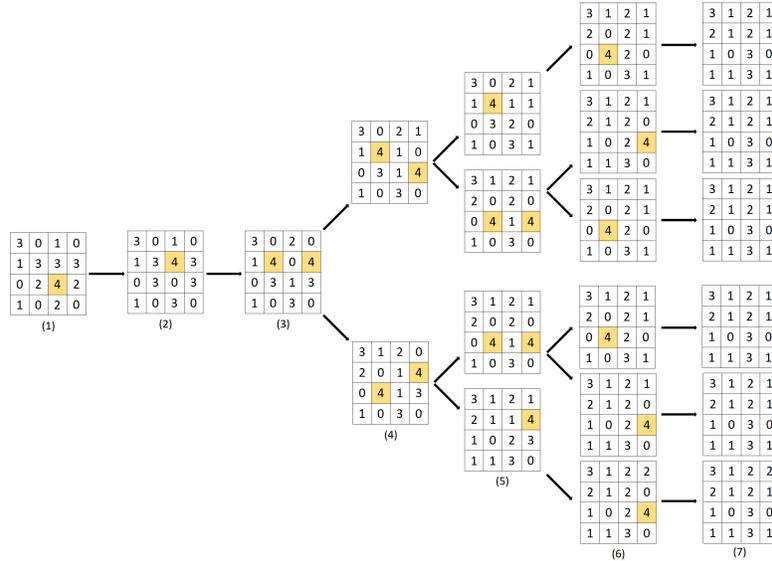


FIGURA 6. Una configuración inicial en común, 6 diferentes derribes, 7 posibles configuración, una misma configuración final de equilibrio con 23 partículas.

Este ejemplo muestra el teorema 5.3 donde comienzas con un sistema de 4 x 4 y un número aleatorio de partículas que forman la configuración original donde ves algunos cuadros amarillos que colapsan. Figura 6 muestra los diferentes escenarios que conducen a una única configuración final estable.

En la Figura 7 vemos que al agregar una partícula en cada uno de los vecindarios de nuestro sistema, el grupo de colapso crece en un factor de como máximo 2 y agregamos otra partícula de la misma manera a todos los distritos con los que se puede demostrar que el factor vuelve a crecer como máximo 2.

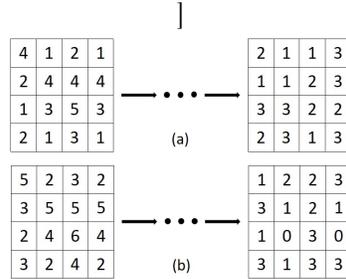


FIGURA 7. (a) es la configuración inicial original más una partícula en cada una de sus vecindades, (b) es la configuración inicial original más dos partículas en cada una de sus vecindades.

Example 6.3. En la Figura 8 vemos la representación gráfica de cada una de las posibles condiciones de contorno para nuestro sistema de pila de arena abeliana, las celdas amarillas representan el grupo de colapso y las celdas verdes el grupo visitante.

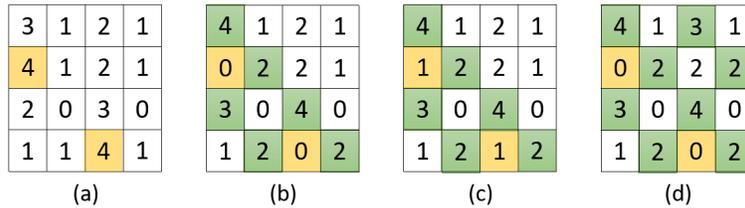


FIGURA 8. Condiciones de contorno en el modelo de pilas de arena, de muestra (a): abierto (b), cerrado (c) y periódico (d) [5]

7. CONCLUSIONES

- (1) Uno de los atributos del SOC (como el modelo de pila de arena) es el almacenamiento de información. Esto se debe a la repetición que se produce cuando alcanza su configuración de equilibrio final, porque se ha maximizado el número de patrones de avalancha repetidos.
- (2) Estos modelos se utilizan para parametrizar muchos fenómenos naturales porque se centran en los efectos espaciales y temporales del barrio más cercano al barrio examinado, y no tienen nada que ver con lo ocurrido en el pasado, porque son independientes en el tiempo y el espacio.

- (3) La importancia del grupo de derrumbe es muy importante, porque a menudo podemos comprender cuántos eventos han ocurrido en la naturaleza. El grupo de visitado ayuda a comprender cuánto afectará la retroalimentación al sistema para lograr una configuración final estable.
- (4) Su importancia en la investigación de fallas de energía lo convierte en un modelo de investigación muy importante en el campo de los autómatas celulares para comprender la posibilidad de desarrollar temas de investigación futuros.

REFERENCIAS

1. Marc Audard, Manuel Güdel, Jeremy J Drake, and Vinay L Kashyap, *Extreme-ultraviolet flare activity in late-type stars*, The Astrophysical Journal **541** (2000), no. 1, 396.
2. Per Bak, Kan Chen, and Chao Tang, *A forest-fire model and some thoughts on turbulence*, Physics letters A **147** (1990), no. 5-6, 297–300.
3. Per Bak, Chao Tang, and Kurt Wiesenfeld, *Self-organized criticality: An explanation of the $1/f$ noise*, Physical review letters **59** (1987), no. 4, 381.
4. ———, *Self-organized criticality*, Physical review A **38** (1988), no. 1, 364.
5. Arka Banerjee, *Self-organized criticality in sandpile models*, 2018.
6. Vincenzo Capparelli, Christian Franzke, Antonio Vecchio, Mervyn P Freeman, Nicholas W Watkins, and Vincenzo Carbone, *A spatiotemporal analysis of us station temperature trends over the last century*, Journal of Geophysical Research: Atmospheres **118** (2013), no. 14, 7427–7434.
7. Jean M Carlson and JS Langer, *Properties of earthquakes generated by fault dynamics*, Physical Review Letters **62** (1989), no. 22, 2632.
8. JM Carlson and JS Langer, *Mechanical model of an earthquake fault*, Physical Review A **40** (1989), no. 11, 6470.
9. BA Carreras, D Newman, VE Lynch, and PH Diamond, *Self-organized criticality as a paradigm for transport in magnetically confined plasmas*, Plasma Physics Reports **22** (1996), no. 9, 740–753.
10. Benjamin A Carreras, David E Newman, Ian Dobson, and A Bruce Poole, *Evidence for self-organized criticality in a time series of electric power system blackouts*, IEEE Transactions on Circuits and Systems I: Regular Papers **51** (2004), no. 9, 1733–1740.
11. Kan Chen and Per Bak, *Is the universe operating at a self-organized critical state?*, Physics Letters A **140** (1989), no. 6, 299–302.
12. Deepak Dhar, *Self-organized critical state of sandpile automaton models*, Physical Review Letters **64** (1990), no. 14, 1613.
13. Ian Dobson, Jie Chen, JS Thorp, Benjamin A Carreras, and David E Newman, *Examining criticality of blackouts in power system models with cascading events*, Proceedings of the 35th annual Hawaii international conference on system sciences, IEEE, 2002, pp. 10–pp.
14. Anne Fey, Lionel Levine, and Yuval Peres, *Growth rates and explosions in sandpiles*, Journal of statistical physics **138** (2010), no. 1-3, 143–159.
15. Anne Fey-den Boer and Frank Redig, *Limiting shapes for deterministic centrally seeded growth models*, Journal of Statistical Physics **130** (2008), no. 3, 579.
16. Alexander E Holroyd, Lionel Levine, Karola Mészáros, Yuyal Peres, James Propp, and David B Wilson, *Chip-firing and rotor-routing on directed graphs*, In and Out of Equilibrium 2, Springer, 2008, pp. 331–364 (9).
17. Nannan Jiang, Hui Ren, David Watts, Ying Liu, Haitao Lu, and Jiefu Tian, *The spatial statistics of self-organized in power system*, 2018 International Conference on Power System Technology (POWERCON), IEEE, 2018, pp. 3974–3980.
18. Damien Olivier Juan Luis Jimenez Laredo, Frédéric Guinand and Pascal Bouvry, *Load balancing at the edge of chaos: how self-organized criticality can lead to energy-efficient computing*, IEEE Transactions on Parallel and Distributed Systems **28** (2016), no. 2, 517–529.
19. Yan Limei, Jin Hongzhang, and Rong Panxiang, *Brittleness in electric transmission network*, 2005 IEEE/PES Transmission & Distribution Conference & Exposition: Asia and Pacific, IEEE, 2005, pp. 1–4.

20. Ronald Meester, Frank Redig, and Dmitri Znamenski, *The abelian sandpile; a mathematical introduction*, arXiv preprint cond-mat/0301481 (2003), 1–.
21. MK Sachs, MR Yoder, DL Turcotte, JB Rundle, and BD Malamud, *Black swans, power laws, and dragon-kings: Earthquakes, volcanic eruptions, landslides, wildfires, floods, and soc models*, The European Physical Journal Special Topics **205** (2012), no. 1, 167–182.
22. Didier Sornette and Ivan Dornic, *Parallel bak-sneppen model and directed percolation*, Physical Review E **54** (1996), no. 4, 3334.
23. Nassim Nicholas Taleb, *Black swans and the domains of statistics*, The American Statistician **61** (2007), no. 3, 198–200.
24. Terry E Tullis, Keith Richards-Dinger, Michael Barall, James H Dieterich, Edward H Field, Eric M Heien, Louise H Kellogg, Fred F Pollitz, John B Rundle, Michael K Sachs, et al., *A comparison among observations and earthquake simulator results for the allcal2 california fault model*, Seismological Research Letters **83** (2012), no. 6, 994–1006.
25. ———, *Generic earthquake simulator*, Seismological Research Letters **83** (2012), no. 6, 959–963.
26. Donald L Turcotte and Bruce D Malamud, *Landslides, forest fires, and earthquakes: examples of self-organized critical behavior*, Physica A: Statistical Mechanics and its Applications **340** (2004), no. 4, 580–589.
27. Nicholas W Watkins and Mervyn P Freeman, *Natural complexity*, Science **320** (2008), no. 5874, 323–324.
28. Stephen Wolfram, *Universality and complexity in cellular automata*, Physica D: Nonlinear Phenomena **10** (1984), no. 1-2, 1–35.
29. Qun Yu, Anqiang Feng, Qing He, and Na Cao, *Weak part identification in power grid based on self-organized criticality*, 2016 China International Conference on Electricity Distribution (CICED), IEEE, 2016, pp. 1–6.
30. Qing Zhong, Zhe Zhang, Feng Liu, Gang Wang, and Longjun Wang, *Self-organized criticality characteristics of unplanned power failures in distribution network*, 2016 IEEE International Conference on Power System Technology (POWERCON), IEEE, 2016, pp. 1–5.

MAESTRÍA EN MATEMÁTICAS, ORIENTACIÓN EN INGENIERÍA MATEMÁTICA, UNIVERSIDAD NACIONAL AUTÓNOMA DE HONDURAS, TEGUCIGALPA, HONDURAS
E-mail address: n.sabillon@unah.hn

MODELO ESTOCÁSTICO PRESA-PREDADOR

ELVIS RAFAEL ARRAZOLA

RESUMEN. En los últimos años, han surgido muchos problemas nuevos desafiantes y emocionantes en biología matemática y ecología. El horizonte de esta investigación se amplía y se profundiza día a día por la exótica utilidad e importancia para el ser humano. También ha habido una creciente preocupación por la preservación del equilibrio ecológico en la naturaleza, que juega un papel importante en la armonía entre las diferentes especies. Como resultado, los modelos matemáticos para poblaciones con interacción entre especies se han vuelto populares entre biólogos y científicos en general.

En este artículo trabajaremos con los modelos presa-depredador que están muy extendidos y vienen en muchos aspectos en la ecología de poblaciones.

Para nuestro aspecto pasaremos del modelo determinista al modelo estocástico y mediante métodos numérico para resolver ecuaciones diferenciales estocásticas mostraremos algunos resultados.

1. INTRODUCCIÓN

La adecuación de los modelos matemáticos puede aumentarse en gran medida teniendo en cuenta las propiedades estocásticas de los sistemas dinámicos. Los modelos estocásticos son ampliamente utilizados en hidrodinámica, estos basados en modelación estocástica de los caudales de cuencas hidrográficas, en dinámica de poblaciones y epidemiología aplicado a la modelación estocástica de la interacción entre células cancerosas y células de la BMU[1], en filtrado de señales basado en la modelación estocástica de las señales de acción sísmica en lugares donde podría representar riesgos de derrumbe[2], en matemáticas económicas y financieras, están aplicadas en diferentes áreas como ser modelación estocástica de la tasa de cambio de la moneda regional respecto U.S. Dólar[3], también para el modelo estocástico de equilibrio macroeconómico, particularmente en la acumulación de capital, inflación y política fiscal[4], así como diferentes campos de la física.

Las ecuaciones diferenciales estocásticas (EDE) [5], son el principal aparato matemático de tales modelos. A menudo en cursos de teoría de probabilidad y modelado matemático, se estudian procesos simples de nacimiento y muerte que proporcionan una introducción útil al modelado estocástico; el nacimiento, muerte y nacimiento-muerte [6]. Estos procesos son generalizaciones estocásticas naturales del modelo poblacional determinista para el crecimiento poblacional de una sola especie. El modelo clásico es formulado considerando dos poblaciones que interactúan a la vez en tiempo t , donde a las especies de depredadores se denota por $f(t)$ y una especie de presa $r(t)$.

Date: Mayo 5, 2021.

Key words and phrases. Ecuaciones diferenciales estocásticas, Procesos de Wiener, Ruido Gaussiano, Euler-Muruyama, Integral de Itô.

Para este proyecto de investigación se desarrollará de manera básica el modelo presa-predador mencionado en el párrafo anterior, el alcance deseado es resolver el modelo de manera aproximado utilizando métodos de Euler-Maruyama[7], para ecuaciones diferenciales estocásticas, así, posteriormente hacer comparaciones entre modelo presa-predador originalmente modelado por los sistemas dinámicos y el modelo estocástico presa-depredador; y así poder encontrar algunas diferencias de los mismos; más aun el análisis del caso original de la modelación presa-predador podría variar en muchos aspectos, ya que podemos analizar los casos en los que tenemos más predadores que presas y así determinar qué población será la mas afectada, analizar también si tenemos exceso de presas y pocos depredadores; por ejemplo si en una población existe un exceso de insectos en un determinado cultivo que afectan la cosecha y no tenemos los suficientes predadores para contrarrestar el daño a los alimentos.

2. JUSTIFICACIÓN

Como finalidad primordial de la Maestría en Matemáticas de la Universidad Nacional Autónoma de Honduras (UNAH), se plantea que el egresado sea capaz de resolver problemas que se presentan en la ciencia, y enmarcados en los ejes prioritarios de investigación de la misma universidad. Este trabajo se sitúa en el eje de investigación: población y condiciones de vida en el medio ambiente, específicamente en el tema prioritario correspondiente a cultura, ciencia y educación. Nuestro énfasis es poder analizar el comportamiento de dos tipos de poblaciones, las cuales están determinadas por las presas y los depredadores. En el marco de las condiciones de vida podemos ver como un sistema de virus (predadores) que se propaga en el cuerpo humano que ataca los glóbulos blancos (presas); en donde nuestro enfoque podría desarrollar al análisis de como reaccionaría un individuo con facilidad al acceso de recurso para combatir el virus y analizar que sucede con los individuos que no pueden acceder a estos recursos. En la misma maestría, la orientación en Estadística Matemática tiene sus propias líneas de investigación, dentro de las cuales podemos determinar nuestro trabajo dentro de dos áreas los cuales son: el análisis de sobrevivencia y confiabilidad puesto que enmarca el hecho de trabajar con temas de nacimiento-muerte [13], además de ellos por su implementación mediante EDE representa fundamental importancia en el área de procesos estocásticos ya que se analiza cómo se comportan las variables aleatorias respecto al tiempo.

Además de ello realizaremos un proceso de modelación de los métodos numéricos, podremos profundizar en el análisis de estabilidad en la línea de modelación matemática, mismas que hacen referencia a una de las líneas de investigación del programa. Esto nos permitirá analizar el comportamiento tanto en modelación estocástica como en modelación de los sistemas dinámicos,[14], siendo de gran interés para poder predecir la variación de las soluciones al problema que planteemos; esto al poder hacer cambios en las condiciones iniciales que permitan determinar de manera eficiente las predicciones.

3. ANTECEDENTES

El modelo presa-depredador fue originalmente propuesto por A.J. Lotka y V. Volterra en la década de 1920[8]. Es un modelo relativamente simple y el cual se

estudia en la mayoría de los cursos de ecuaciones diferenciales ordinarias. Las ecuaciones de Lotka-Volterra se desarrollaron para describir la dinámica de sistemas de la biología. Este sistema de ecuaciones diferenciales no lineales puede describirse como un método más general de la versión de un modelo de Kolmogorov porque se centra solo en las interacciones depredador-presa e ignora la competencia, la enfermedad y el mutualismo que incluye el modelo de Kolmogorov. Las ecuaciones de Lotka-Volterra se pueden escribir simplemente como un sistema de primer orden no lineal de ecuaciones diferenciales ordinarias (EDO). Como las ecuaciones son de naturaleza diferencial, estas soluciones son deterministas (no hay aleatoriedad involucrada, y las mismas condiciones iniciales producen el mismo resultado), y el tiempo es continuo (las generaciones de depredadores y presas son continuamente superpuestas). Como con muchos otros modelos matemáticos, muchos supuestos fueron hechos en la creación de las ecuaciones de Lotka-Volterra. Tales suposiciones incluyen:

1. No hay escasez de alimentos para la población de presas.
2. La cantidad de comida suministrada a la presa está directamente relacionada con el tamaño de población de la presa.
3. La tasa de cambio de la población es directamente proporcional a su tamaño.
4. El ambiente es constante y se supone que la adaptación genética sea insignificante.
5. Los depredadores nunca dejarán de comer.

Posteriormente tenemos el modelo de población presa-depredador estocástico con efecto Allee; en las poblaciones ecológicas, el crecimiento de las especies a menudo se ve amortiguado por un efecto Allee. El llamado efecto Allee se describe por la relación positiva entre la densidad de población y la tasa de crecimiento per-cápita. El modelo logístico supone que la tasa de crecimiento per-cápita disminuye monótonamente con la densidad. Sin embargo para la población sujeta a un efecto Allee, la tasa de crecimiento per-cápita muestra una curva jorobada que aumenta (de negativa a positiva) a baja densidad, hasta un máximo en densidad intermedia y luego disminuye.

El efecto Allee puede surgir de varias fuentes, como las dificultades para encontrar parejas cuando la densidad de población es baja, la disfunción social en poblaciones pequeñas, la explotación de alimentos, la depresión endogámica y la evitación o defensa de los depredadores. Este efecto generalmente se satura o desaparece a medida que la población aumenta. El efecto Allee se ha modelado de diferentes maneras, como el efecto Allee multiplicativo y el efecto Allee aditivo. Desde un punto de vista ecológico, el efecto Allee se ha modelado en casos fuertes y débiles. El fuerte efecto Allee introduce un umbral de población, y la población debe superar este umbral para crecer. En contraste, la población con un efecto Allee débil no tiene un umbral. El fuerte impacto potencial del efecto Allee en la dinámica de la población ha atraído mucha atención recientemente. Se pueden encontrar investigaciones detalladas relacionadas con el efecto Allee en [9].

Por último mencionaremos la dinámica global del modelo estocástico presa , depredador con interferencia mutua y defensa de presas en donde se pretende analizar dinámicas de los modelos biológicos[10]. Factores no lineales como la respuesta funcional y la interferencia mutua se encuentran entre varios factores cruciales que afectan los modelos biológicos. Los tipos de interferencia mutua han sido de gran interés últimamente en donde definimos interferencia mutua como las interacciones de comportamiento entre los organismos de alimentación, que reducen el tiempo que cada individuo pasa obteniendo alimentos, o la cantidad de alimentos que cada individuo consume. Ocurre más comúnmente donde la cantidad de alimentos es escasa, o la población de organismos de alimentación es grande. Wang estudió la existencia, la permanencia y estabilidad asintótica global para un sistema presa-depredador con interferencia mutua[11]. Wang y su grupo también obtuvieron algunas condiciones suficientes, para la permanencia y la atracción global de la solución periódica positiva para un modelo Volterra, y modelo presa-depredador retrasado con interferencia mutua.

Todos los procesos biológicos se ven afectados fundamentalmente por el ruido blanco ambiental. En modelado estocástico, el ruido ambiental se concibe como fluctuaciones estocásticas. Podemos analizar un sistema biológico bajo fluctuaciones estocásticas, considerando el ruido ambiental forzado para la población. Podemos analizar además que cuando la población se desvía del punto de equilibrio, el sistema muestra un comportamiento irregular (es decir, inestabilidad). Así, después de que uno investiga un sistema determinista, se procede a la extensión de los resultados al modelo estocástico.

4. EL MODELO PRESA - DEPREDADOR Y ALGUNAS VARIACIONES

Supongamos que en un ecosistema cerrado (es decir, no se permite la migración hacia adentro o hacia afuera del sistema) solo hay 2 tipos de animales: el depredador y la presa. Forman una cadena alimentaria simple donde la especie depredadora caza a la especie presa, mientras que la presa se alimenta de la vegetación. El tamaño de las dos poblaciones se puede describir mediante un sistema simple de dos ecuaciones diferenciales de primer orden no lineales (también conocidas como las ecuaciones de Lotka-Volterra, que se originaron en el estudio de las poblaciones de peces del Mediterráneo durante e inmediatamente después de la Primera Guerra Mundial). El modelo clásico es formulado considerando dos poblaciones que interactúan en un tiempo t , un especie depredadora, denotada $f(t)$ y una especie de presa, $r(t)$. En ausencia de depredadores, las especies de presas exhiben un crecimiento exponencial. En particular,

$$(4.1) \quad \frac{dr(t)}{dt} = ar(t)$$

donde $r(t) > 0$ para todo t y a es la tasa de natalidad de presas. Se supone que aquí siempre hay suficiente comida para la presa; de manera similar, en ausencia de presas, los depredadores mueren exponencialmente. Es decir,

$$(4.2) \quad \frac{df(t)}{dt} = -cf(t)$$

donde $f(t) > 0$ para todo t y $c > 0$ es la tasa de muerte de depredadores. Se puede suponer que hay otro alimento para los depredadores, pero no es suficiente para

mantener a la población. Cuando ambas poblaciones están presentes, la interacción de las especies se modela como un producto de los tamaños de la población (es decir, $r(t)f(t)$). El término $r(t)f(t)$ se toma prestado de modelos químicos de tasas de reacción donde las moléculas en solución interactúan chocando al azar unas con otras. En el modelo clásico, una interacción presa-depredador da como resultado una presa muerta a una tasa $br(t)f(t)$, donde $b > 0$. Dado que se supone que las presas son la principal fuente de alimento para los depredadores, los depredadores tienen una tasa de natalidad $dr(t)f(t)$, donde $c > 0$. El modelo clásico depredador-presa, con los supuestos discutidos anterior, produce el siguiente sistema de ecuaciones:

$$(4.3) \quad \begin{cases} \frac{dr(t)}{dt} = ar(t) - br(t)f(t) \\ \frac{df(t)}{dt} = -cf(t) + dr(t)f(t) \end{cases}$$

donde a , b , c y d son constantes positivas; a , c representan las tasas de crecimiento intrínsecas de las presas y depredadores respectivamente y b , d miden la interacción entre las dos especies. De manera general, un modelo de Lotka-Volterra determinista clásico de dos especies se puede expresar de la siguiente manera:

$$(4.4) \quad \begin{cases} \frac{dX(t)}{dt} = X[r_1 - a_{11}X - a_{12}Y] \\ \frac{dY(t)}{dt} = Y[r_2 - a_{21}X - a_{22}Y] \end{cases}$$

aquí $X(t)$ y $Y(t)$ representan los tamaños de las poblaciones. Tenga en cuenta que la diferencia entre este sistema y Volterra (4.3) son los nuevos coeficientes intra-específicos a_{ii} . Mediante este nuevo sistema podemos determinar tres clasificaciones con respecto a las presas y depredadores.

- 1) Cooperación entre las especies; esto sucede cuando $r_1 > 0$, $r_2 > 0$, $a_{12} < 0$, $a_{21} < 0$.
- 2) Competencia entre dos especies; esto sucede cuando $r_1 > 0$, $r_2 > 0$, $a_{12} > 0$, $a_{21} > 0$.
- 3) Presa-depredador; esto sucede cuando $r_1 > 0$, $r_2 < 0$, $a_{12} > 0$, $a_{21} < 0$.

Esta dinámica puede verse en [12].

Sin embargo, en estos modelos deterministas los parámetros son constantes y no incorporan muchos factores impredecibles que estén relacionados con los factores ambientales y la variabilidad demográfica (variabilidad alimentaria, enfermedades ...). En general, la variabilidad ambiental se trata modificando los parámetros del modelo. Este enfoque consiste en definir los parámetros con una función lineal de ruido blanco gaussiano (WN) por sus siglas en inglés [5], es decir, una función lineal de la derivada de un Proceso de Wiener [5]. Es decir, suponemos que el ruido blanco aparece en r_1 y r_2 entonces generamos un cambio de la siguiente manera:

$$(4.5) \quad r_j \rightarrow r_j + \sigma_j \frac{dW_j(t)}{dt}$$

en este caso σ_j es una constante positiva y $W_j(t)$ se le conoce como movimiento Browniano [13]. Entonces se describe el siguiente modelo estocástico:

$$(4.6) \quad \begin{cases} \frac{dX(t)}{dt} = X[r_1 - a_{11}X - a_{12}Y] + \sigma_1 X \frac{dW_1(t)}{dt} \\ \frac{dY(t)}{dt} = Y[r_2 - a_{21}X - a_{22}Y] + \sigma_2 Y \frac{dW_2(t)}{dt} \end{cases}$$

donde el término estocástico es lineal en las variables estocásticas X y Y . Bajo la hipótesis simple $\sigma_1, \sigma_2 > 0$, se muestra que esta solución es positiva y global. Otra posibilidad considerada sucede cuando ruidos blancos (WN) aparecen en los parámetros a_{ij} de manera que:

$$(4.7) \quad a_{ij} \rightarrow a_{ij} + \sigma_{ij} \frac{dW(t)}{dt}$$

con las condiciones $\sigma_{ii} > 0$ y $\sigma_{ij} \geq 0$ si $i \neq j$. En este caso, el término de difusión en el sistema diferencial estocástico sería cuadrático. Luego el sistema de ecuaciones diferenciales estocásticas (SEDE) para la dinámica de dos poblaciones que interactúan tiene la forma:

$$(4.8) \quad d\mathbf{X} = \mu(t, \mathbf{X}) + \mathbf{B}(t, \mathbf{X})d\mathbf{W}$$

donde $\mathbf{X} = (X, Y)^T$ y $\mathbf{W} = (W_1, W_2)^T$ son dos procesos de Wiener independientes. Además de ello se define $\mu(t, \mathbf{X})$ que es una función vectorial y $\mathbf{B}(t, \mathbf{X})$ es la matriz de difusión correspondiente. Una punto clave en biología de poblaciones es comprender las limitaciones que conducen a la extinción o coexistencia de las especies.

4.1. Análisis del modelo estocástico presa-predador. Consideramos un sistema muy simple de presa-depredador de Lotka-Volterra .

Supongamos que las presas pueden ser conejos y que disfrutan de un suministro infinito de alimento y los depredadores pueden ser zorros. Este problema se puede modelar mediante ecuaciones diferenciales ordinarias no lineales. Sea $R(t)$ el tamaño de la población de presas en el momento t y $F(t)$ el tamaño de la población de depredadores en el momento t . En el modelo que se va a formular, se va a suponer que en lugar de una tasa (determinista) de nacimientos y muertes de depredadores y presas, existe la probabilidad de que nazcan o mueran depredadores y presas. Por tanto, $R(t)$ y $F(t)$ son variables aleatorias dependientes del tiempo. Como en el proceso simple de nacimiento-muerte para una sola especie, los nacimientos y muertes en este proceso ocurren proporcionalmente al tamaño de la población. Suponemos que la probabilidad infinitesimal del nacimiento de una presa individual durante el intervalo de tiempo pequeño Δt es $ar(t)\Delta t$, donde para $a > 0$ tenemos la tasa de natalidad de la presa. De manera similar, asumimos que la probabilidad infinitesimal de la muerte de un depredador individual durante el intervalo de tiempo pequeño Δt es $cf(t)\Delta t$, donde para $c > 0$ tenemos la tasa de muerte del depredador. Veamos el siguiente diagrama para entender el comportamiento que tendrá nuestro sistema:

| transiciones | tasas |
|--|--------|
| $r \rightarrow r + 1, f \rightarrow f$ | ar |
| $r \rightarrow r - 1, f \rightarrow f$ | brf |
| $r \rightarrow r, f \rightarrow f + 1$ | drf |
| $r \rightarrow r, f \rightarrow f - 1$ | cf . |

El argumento estándar que usa las ecuaciones de Kolmogorov hacia adelante [5] se usa para obtener $P_{r,f}(t)$, considerando la probabilidad $P_{r,f}(t + \Delta t)$ donde $P_{r,f}(t)$ es la probabilidad que existan depredadores y presas al mismo tiempo. Esta probabilidad se obtiene como la suma de las probabilidades de los siguientes eventos mutuamente excluyentes:

- a) Hay r presas y f depredadores en el tiempo t y no se producen nacimientos ni muertes de ninguna de las especies en $(t, t + \Delta t)$.
- b) Hay $r - 1$ presas y f depredadores en el tiempo t y una presa nace en $(t, t + \Delta t)$.
- c) Hay r presas y $f - 1$ depredadores en el tiempo t y el nacimiento de un depredador ocurre en $(t, t + \Delta t)$.
- d) Hay $r + 1$ presas y f depredadores en el tiempo t y se produce la muerte de una presa en $(t, t + \Delta t)$.
- e) Hay presas r y depredadores $f + 1$ y la muerte de un depredador ocurre en $(t, t + \Delta t)$.

Se asume que Δt es suficientemente pequeño para garantizar que solo uno de estos eventos pueda ocurrir en $(t, t + \Delta t)$.

Haciendo un análisis de la variabilidad demográfica y abordando las variaciones del proceso podemos determinar lo siguiente: supongamos que Δt es un pequeño intervalo de tiempo.

$$\begin{aligned}\Delta R &= R(t + \Delta t) - R(t), \\ \Delta F &= F(t + \Delta t) - F(t),\end{aligned}$$

como mencionamos anteriormente tenemos cinco probabilidades descritas a continuación:

$$(4.9) \quad P(\Delta R = 1, \Delta F = 0) = aR(t)\Delta t$$

$$(4.10) \quad P(\Delta R = -1, \Delta F = 0) = bR(t)F(t)\Delta t$$

$$(4.11) \quad P(\Delta R = 0, \Delta F = 1) = dR(t)F(t)\Delta t$$

$$(4.12) \quad P(\Delta R = 0, \Delta F = -1) = cF(t)\Delta t$$

$$(4.13) \quad P(\Delta R = 0, \Delta F = 0) = 1 - P(t)\Delta t.$$

donde $P(t) = aR(t) + cF(t) + dR(t)F(t) + bR(t)F(t)$. Luego para obtener las ecuaciones diferenciales estocásticas calculamos la esperanza y la covarianza de las variables estocásticas ΔR y ΔF (ver [14]). En este caso, el sistema diferencial estocástico es:

$$(4.14) \quad \begin{cases} dR = R(a - bF)dt + \sqrt{R(a - bF)}dW_1(t) \\ dF = F(-c + dR)dt + \sqrt{F(c + dR)}dW_2(t) \end{cases}$$

aquí $dW_1(t)$ y $dW_2(t)$ son dos procesos de Wiener independientes [5].

4.2. Algunas variaciones del modelo para parámetros definidos. Suponemos que nuestro modelo está determinado por los valores de $a = 2$, $c = 1$, $b = d = \alpha$ entonces el sistema modelado a partir de (4.10) es:

$$(4.15) \quad \begin{cases} dR = R(2 - \alpha F)dt + \sqrt{R(2 - \alpha F)}dW_1(t) \\ dF = F(-1 + \alpha R)dt + \sqrt{F(\alpha R + 1)}dW_2(t) \end{cases}$$

notemos que los procesos de Wiener siguen siendo los mismos y por ende son independientes. Este modelo nos permite tener un solo parámetro α . Como la variabilidad ambiental se introduce modificando los parámetros del modelo, consideremos la siguiente modificación del parámetro α :

$$(4.16) \quad \alpha \rightarrow \alpha + \sigma \frac{dW(t)}{dt}.$$

donde $W(t)$ es un movimiento browniano y $\sigma > 0$ es una constante. Por lo tanto, si $R(t)$ y $F(t)$ representan el tamaño de la población de presas y depredadores respectivamente, entonces satisfacen el sistema diferencial estocástico:

$$(4.17) \quad \begin{cases} dR = R(2 - \alpha F)dt - \sigma RF dW_1(t) \\ dF = F(-1 + \alpha R)dt + \sigma RF dW_2(t) \end{cases}$$

el término σRF representa el término de contacto ponderado con la tasa de intensidad de ruido σ . Este tipo de sistema ha sido estudiados en la literatura[15], sin embargo, siempre se ha asumido que los coeficientes $a_{jj} > 0$ en (4.4), una suposición que no es razonable en el modelo estocástico anterior. Consideramos un tercer modelo donde introducimos distintas perturbaciones lineales a nuestras variaciones. Obtenemos el siguiente sistema estocástico:

$$(4.18) \quad \begin{cases} dR = R(2 - \alpha F)dt + \sigma_1 R dW_1(t) \\ dF = F(-1 + \alpha R)dt + \sigma_2 F dW_2(t) \end{cases}$$

donde $dW_1(t)$ y $dW_2(t)$ siguen siendo dos procesos de Wiener independientes, los términos $\sigma_1 R$ y $\sigma_2 F$ representan las tasas de intensidad de ruido para las presas y depredadores respectivamente. Este modelo es un caso particular de (4.6) y por lo tanto sus soluciones son globales y positivas cuando $\sigma_1, \sigma_2 > 0$. Este modelo también se estudia en [16].

La importancia es que para estos tres modelos es que podemos ver el comportamiento que pueden tener las presas y los depredadores bajo ciertas condiciones iniciales aplicadas a cada uno de los modelos y además dándole valores a los parámetros correspondientes.

4.3. Ecuaciones de Kolmogorov hacia atrás. Se ha desarrollado un modelo estocástico de manera sencilla partiendo del modelo determinista presa-predador de Locka-Volterra. En esta parte de nuestra investigación queremos ver el como es el planteamiento de nuestro modelo mediante la ecuación de Kolmogorov hacia atrás partiendo de (2.8) y recordando que donde $\mathbf{X} = (X, Y)^T$ y $\mathbf{W} = (W_1, W_2)^T$ son dos procesos de Wiener independientes, $\mu(t, \mathbf{X})$ es una función vectorial y $\mathbf{B}(t, \mathbf{X})$ es la matriz de difusión correspondiente. Definamos la variable aleatoria T que indica el tiempo de persistencia, es decir:

$$T = \inf \{t \geq 0 : X(t) = 0 \text{ o } Y(t) = 0\}.$$

En este caso aunque no se indica explícitamente, T depende del valor inicial $X(0), Y(0)$.

En [1] se discute que el tiempo medio de persistencia $\tau(r, f) = E(T_r, f)$ para (4.8) satisface la ecuación estacionaria hacia atrás de Kolmogorov siguiente:

$$(4.19) \quad L(\tau) = \mu_1 \frac{\partial \tau}{\partial X} + \mu_2 \frac{\partial \tau}{\partial Y} + \frac{1}{2} \sum_{i=1}^2 \sum_{j=1}^2 d_{ij} \frac{\partial^2 \tau}{\partial X \partial Y} + 1 = 0$$

donde el término de difusión es la matriz $\mathbf{D} = \mathbf{B}\mathbf{B}^T = (d_{ij})$ y con condiciones de

frontera dadas por:

$$(4.20) \quad \begin{cases} \tau(x, 0) = \tau(0, y) = 0 \\ \frac{d\tau}{dr}(M_1, y) = 0 \\ \frac{d\tau}{df}(x, M_2) = 0 \end{cases}$$

esto sucede siempre que los valores de x e y no pueda exceder algunos valores M_1 y M_2 respectivamente. Además, el operador L es elíptico porque la matriz D es definida positiva en el dominio, esto es; $\mathbf{x}^T \mathbf{D} \mathbf{x} = \mathbf{x}^T B B^T \mathbf{x} = (\mathbf{B} \mathbf{x})^T (\mathbf{B} \mathbf{x})^T = \|\mathbf{B}^T \mathbf{x}\|_2^2 \geq 0$, por lo tanto sus autovalores son positivos.

Partiendo de la ecuación de Kolmogorov hacia atrás(4.8) podemos encontrar el término de la matriz de difusión para las ecuaciones (4.11), (4.13) y (4.14). Tomemos el caso del modelo estocástico (4.11) entonces el término de difusión está definido por la matriz:

$$(4.21) \quad \mathbf{D}_1 = \frac{1}{2} \begin{pmatrix} r(2 + \alpha f) & 0 \\ 0 & f(\alpha r + 1) \end{pmatrix}$$

para este caso el tiempo medio de extinción $\tau(r, f) = E(T_r, f)$ satisface la ecuación de Kolmogorov hacia atrás estacionaria descrita por:

$$(4.22) \quad L(\tau) = r(2 - \alpha f) \frac{\partial \tau}{\partial r} + f(\alpha r - 1) \frac{\partial \tau}{\partial f} + \frac{1}{2} r(2 + \alpha f) \frac{\partial^2 \tau}{\partial r^2} + \frac{1}{2} f(\alpha r + 1) \frac{\partial^2 \tau}{\partial f^2} + 1 = 0.$$

Para este ecuación tenemos las condiciones de frontera $\tau(0, f) = \tau(r, 0) = 0$, además, como $r(2 + \alpha f) > 0$ y $f(\alpha r + 1) > 0$ cuando $r, f > 0$, el operador $L(\tau)$ es localmente elíptico y su solución es positiva.

Se puede hacer el mismo análisis para poder encontrar las matrices de (4.13) y (4.14); esto para luego poder encontrar las ecuaciones de Komogorov hacia atrás correspondientes. Partiendo de estas ecuaciones podemos hacer el análisis correspondientes para el modelo determinista y el modelo estocástico presa-depredador y así identificar e comportamiento de dichos modelos y verificar si presentan diferencias significativas.

5. PROCESO DE WIENER

El proceso de Wiener $W(t)$ depende continuamente de $t \in [0, T]$ y cumple las siguientes condiciones:

- $W(0) = 0$ (con probabilidad 1).
- Para $0 \leq s < t$, el incremento $W(t) - W(s)$ es una variable aleatoria normalmente distribuida con media cero y varianza $t - s$. Es equivalente a $W(t) - W(s) \sim \sqrt{t - s} N(0, 1) = N(0, t - s)$, donde $N(0, 1)$ denota una variable aleatoria normalmente distribuida con media cero y varianza de la unidad.
- Para $0 \leq s < t < u < v \leq T$, los incrementos $W(t) - W(s)$ y $W(v) - W(u)$ son variables aleatorias independientes.

La distribución normal se define como una distribución de probabilidad continua con una función de densidad de probabilidad en forma de campana, que se conoce como función gaussiana:

$$(5.1) \quad f(x, \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp^{-\frac{1}{2} \left(\frac{(x-\mu)^2}{\sigma^2} \right)}$$

donde μ es la media ó el valor esperado, σ^2 es la varianza. Cuando $\mu = 0$ y $\sigma^2 = 1$ decimos que tenemos una distribución normal estándar.

En general, $N(\mu, \sigma^2)$ denota una distribución normal con valor esperado μ y varianza σ^2 . Si la variable aleatoria X se distribuye normalmente con valor esperado μ y varianza σ^2 , entonces podemos escribir $x \sim N(\mu, \sigma^2) \sim \mu + \sigma N(0, 1)$.

5.1. Integrales de Itô y Stratonovich. Para describir el modelo estocástico que investigamos, necesitamos definir las dos formas diferentes de ecuaciones diferenciales estocásticas: la forma Itô y la forma Stratonovich [17].

Definición 5.1. Sea $[0, T] \subseteq \mathfrak{R}$ y $0 = t_0 < t_1 < \dots < t_n = T$ una discretización del intervalo de tiempo $[0, T]$ siendo el paso más grande $\Delta t = \max (t_{k+1} - t_k)$ para $k \in 1, \dots, n - 1$ entonces

$$(5.2) \quad \int_0^T f(t) dWt = \lim_{\Delta \rightarrow 0} \sum_{i=0}^{n-1} f(t_i) [W(t_{i+1}) - W(t_i)]$$

se llama integral de Itô para la función f .

Definición 5.2. Sea $[0, T] \subseteq \mathfrak{R}$ y $0 = t_0 < t_1 < \dots < t_n = T$ una discretización del intervalo de tiempo $[0, T]$ siendo el paso más grande $\Delta t = \max (t_{k+1} - t_k)$ para $k \in 1, \dots, n - 1$ entonces

$$(5.3) \quad \int_0^T f(t) \circ dWt = \lim_{\Delta \rightarrow 0} \sum_{i=0}^{N-1} f\left(\frac{t_{i+1} + t_i}{2}\right) [W(t_{i+1}) - W(t_i)]$$

se llama integral de Stratonovich de la función f

Por lo tanto, existe una diferencia importante entre las integrales de Itô y de Stratonovich, ya que f se evalúa en el extremo izquierdo del intervalo en el caso de Itô y en el punto medio del caso Stratonovich[3].

La forma Itô de una ecuación diferencial estocástica se puede escribir en forma diferencial:

$$(5.4) \quad dx(t) = f(t, x(t))dt + g(t, x(t))dW_t$$

para $0 \leq t \leq T$ donde $W(t)$ es un proceso de Wiener.

Podemos escribir el modelo de ecuación diferencial estocástica para la interacción depredador-presa agregando un término estocástico al modelo (2.3) como

$$(5.5) \quad \begin{cases} dr = ar - brf + \sigma r dW_t \\ df = -cf + drf + \sigma f dW_t \end{cases}$$

donde $W(t) = W_t$ es un proceso de Wiener estándar, o un movimiento browniano, en el intervalo de tiempo $[0, T]$ y σ es un parámetro que representa la fuerza del ruido. Los modelos estocásticos son modelos más precisos para estudiar procesos biológicos importantes. Un ejemplo de tal aplicación biológica en ecología es el estudio de la interacción presa-depredador. El modelo determinista no describe

los aspectos básicos de un sistema natural en un entorno cambiante que puede causar variaciones aleatorias en la tasa de crecimiento de las presas depredadoras y la tasa de mortalidad. El ruido blanco gaussiano, que es un concepto útil para modelar fenómenos que fluctúan rápidamente y la principal fuente de ruido es que existen incertidumbres inherentes en un sistema ecológico, como estaciones variables o aportes de nutrientes.

6. MÉTODOS NUMÉRICOS PARA MODELOS ESTOCÁSTICOS CONTINUOS

Muchas áreas de la ciencia y la ingeniería se basan en el análisis cuantitativo, a medida que se dispone de modelos matemáticos más complejos de los fenómenos del mundo real. Dado que la mayoría de estos modelos no tienen una solución exacta de forma cerrada, las aproximaciones numéricas son las únicas herramientas disponibles para analizarlos. En particular, el método de Euler-Maruyama (ver por ejemplo Higham [18]) se usa ampliamente para aproximar la solución de modelos de ecuaciones diferenciales estocásticas que surgen en las aplicaciones.

Consideremos una ecuación diferencial estocástica de Itô en la forma general:

$$(6.1) \quad dx(t) = f(x(t))dt + g(x(t))dW_t$$

para $0 \leq t \leq T$. Aquí f y g son funciones vectoriales, y W es un solo proceso de Wiener. El intervalo $[0, T]$ se discretiza como $0 = t_0 < t_1 < \dots < t_n = T$ donde $t_j = j\Delta t$ y $\Delta t = \frac{T-t_0}{n}$ para algún número entero n . El método de Euler-Maruyama (EM) aplicado a la ecuación (4.1) se puede escribir de la siguiente forma:

$$(6.2) \quad x_j = x_{j-1} + f(x_{j-1})\Delta t + g(x_{j-1})(W(t_j) - W(t_{j-1}))$$

para $j=1, \dots, n$.

Dependiendo de las propiedades deseadas de la solución numérica, la solución exacta de la SDE puede aproximarse mediante métodos numéricos fuertes o débiles (ver, por ejemplo, Higham [17]).

Definición 6.1. Si x_k es la aproximación numérica de $x(t_k)$ en una cuadrícula $0 = t_0 < t_1 < \dots < t_n = T$ con el paso más grande $\Delta t = \max(t_{k+1} - t_k)$ con $k \in 0, 1, \dots, n-1$ entonces se dice que la aproximación numérica $(x_k)_k$ converge a $x(t)$ con un orden global fuerte $\psi > 0$ si existe una constante $c > 0$ que no depende de Δt o en la cuadrícula y $\Delta t_0 > 0$ tal que

$$E|x_n - x(t_n)| \leq c\Delta t^\psi$$

para cualquier cuadrícula con $\Delta t < \Delta t_0$.

Tenga en cuenta que ψ puede ser fraccional, ya que el orden cuadrático medio de un incremento de Wiener $\Delta W = W(t + \Delta t) - W(t)$ es $\Delta t^{1/2}$. Las aproximaciones fuertes son necesarias cuando se requiere que la solución numérica siga de cerca la solución exacta en cada camino. Sin embargo, cuando estemos interesados en obtener una aproximación precisa de los momentos de la solución exacta, se emplearán métodos numéricos débiles.

Definición 6.2. Si x_k es la aproximación numérica de $x(t_k)$ en una cuadrícula $0 = t_0 < t_1 < \dots < t_n = T$ con el paso más grande $\Delta t = \max(t_{k+1} - t_k)$ con $k \in 0, 1, \dots, n-1$ entonces se dice que la aproximación numérica $(x_k)_k$ converge a

$x(t)$ con un orden global débil $\psi > 0$ para $T > 0$ y cualquier polinomio p , se cumple lo siguiente:

$$|E(p(x_n)) - E(p(x(T)))| \leq c\Delta t^\psi$$

para cualquier $\Delta t < \Delta t_0$ para algún $\Delta t_0 > 0$ donde c es una constante que es independiente de Δt y de la cuadrícula.

Observamos que el método de Euler-Maruyama es de orden de convergencia fuerte $1/2$ y orden de convergencia débil 1 . Recordamos que el método de Euler en el marco determinista era de orden de convergencia 1 . El orden de convergencia $1/2$ fuerte se debe a la presencia de los incrementos de Wiener que se comportan como:

$$(6.3) \quad \begin{cases} dr = ar - brf + \sigma r dW_t \\ df = -cf + drf + \sigma f dW_t \end{cases}$$

donde W_t es un proceso Wiener. Consideramos los siguientes parámetros $a = 0,05$, $b = 1$, $c = 0,2$ y $d = 0,3$. Las condiciones iniciales son $r(0) = 0,1$ e $f(0) = 0,1$.

Euler-Maruyama para el modelo estocástico presa-depredador: El método Euler-Maruyama (EM)[18] es uno de los métodos computacionales más simples para la aproximación de EDEs, es el análogo al método de Euler para ecuaciones diferenciales ordinarias. Desarrollamos brevemente este método para pasar al caso aleatorio pensando en el problema general de valor inicial.

Aplicamos el método de Euler-Maruyama al sistema (4.3) y obtenemos:

$$(6.4) \quad \begin{cases} r(i+1) = ar(i) - br(i)f(i) + \sigma r(i)\sqrt{\Delta t}N(0,1) \\ f(i+1) = -cf(i) + dr(i)f(i) + \sigma f(i)\sqrt{\Delta t}N(0,1) \end{cases}$$

para $i = 0, 1, \dots, n$ donde $r(0) = 3$, $f(0) = 2$, $c = 0,37$, $d = 0,05$, $a = 0,4$, $b = 0,3$. El número de pasos en la malla es $n = 1000$, $T = 100$; $\Delta t = T/n$; El valor del parámetro σ varía de $\sigma = 0$ a $\sigma = 0,3$.

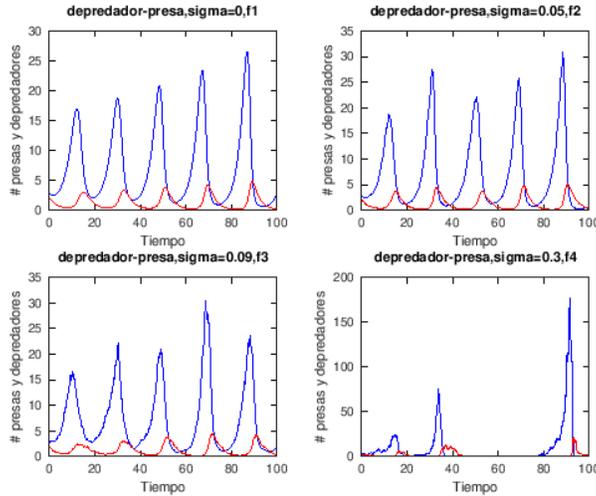


FIGURA 1. La evolución en el tiempo del depredador y de la presa con una secuencia creciente de ruidos.

Cuando la fuerza del ruido aumenta, las fluctuaciones de la Curva aumentan. La figura 1 muestra la evolución de los depredadores y presas a través del tiempo, notando que a medida aumenta el ruido blanco la población de presas (azul) aumenta respecto a los depredadores (rojo).

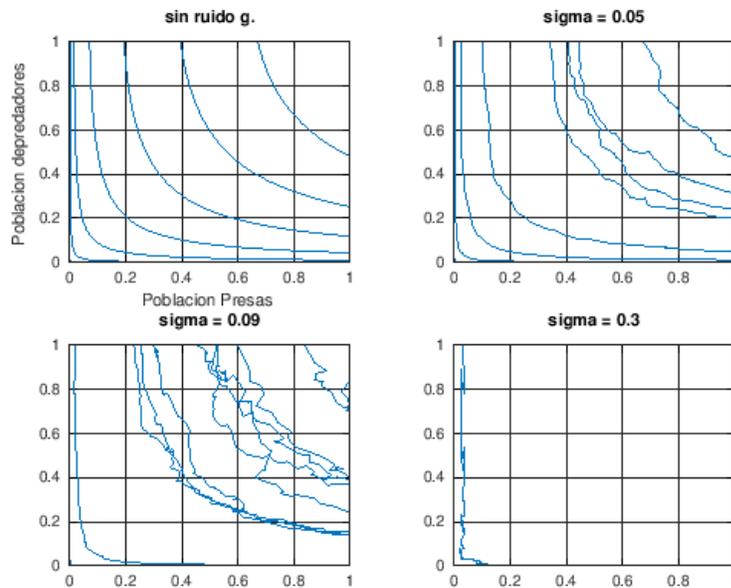


FIGURA 2. Secuencia de retratos de fase del modelo estocástico depredador-presa.

En la figura 2 podemos ver que a medida aumentamos el ruido blanco (σ) el retrato de fase del modelo va desapareciendo hasta perder su forma. Podemos observar que las soluciones en equilibrio ya no son tan notorias. Por lo tanto la población en equilibrio se ve afectada por el ruido blanco gaussiano.

7. CONCLUSIONES

1. Mediante la elaboración de este trabajo de investigación se pretendió analizar cuales son las variaciones que presenta la modelación de un problema; para el cual en este caso dicho problema está planteado de dos formas mediante el modelos determinista Lotka-Volterra y un modelo estocástico.
2. Nuestro propósito también es que podamos identificar que tipo de modelo estocástico se está presentando, ya que al pasar del modelo determinista este genera diferentes formas de modelarlo estocásticamente, la cuales se desarrollaron en este trabajo mediante las ecuaciones (4.11), (4.13) y (4.14).
3. Vemos también que podemos desarrollar nuestro modelo mediante la ecuación de Kolmogorov hacia atrás entonces queremos concluir que tipo de modelación estocástica nos proporcionaría mejor información de acuerdo a los datos que utilicemos, esto mediante los parámetros correspondientes a cada ecuación.
4. La ecuación depredador-presa de Lotka-Volterra es un modelo crítico en ecología. El estudio de la dinámica en el modelo Lotka-Volterra y su generalización es un problema clave en ecología.
5. El ruido blanco gaussiano se utilizó para observar oscilaciones en un entorno estocástico. La predicción de la coexistencia de las poblaciones de dos especies se puede hacer con la ayuda de la observación de fluctuaciones en el modelo estocástico.

REFERENCIAS

- [1] Jose A. Canto. Modelación estocástica de la interacción entre células cancerosas y células de la bmu. Master's thesis, Centro de investigación matemática. (CIMAT), 2016.
- [2] J.E. Hutado . Modelación estocástica de la acción sísmica. Technical report, 1999.
- [3] Carlos A. Grajales. Modelación estocástica de la tasa de cambio peso colombiano / u.s. dólar. Master's thesis, universidad de Medellin, 2006.
- [4] Francisco Venegas. Un modelo estocástico de equilibrio macroeconómico: acumulación de capital, inflación y política fiscal. *SciElo Analytics*, 208.
- [5] Bernt Oksendal. *Stochastic Differential Equations*. New York: Springer-Verlag Heidelberg, 2003.
- [6] H. C. Tuckwell. *Elementary Applications of Probability Theory*. Chapman & Hal , London:, 1995.
- [7] E.Platen P.E. Kloeden. *Numerical Solution of Stochastic Differential Equations*. New York: Springer,, 1995.
- [8] W. Hu Tyler Blaszak. Lotka-volterra models of predator-prey relationships. 2019.
- [9] B. Dennis. Allee effects: population growth, critical density, and the chance of extinction. *Universidad de Idaho, Moscú*, 3(4), 1989.
- [10] Feng Rao. Dynamical analysis of a stochastic predator-prey model with an allee effect. *Hindawi*, 2013.
- [11] K. Wang. Existence and global asymptotic stability of positive periodic solution for a predator prey system with mutual interference. *ScienceDirec*, 10:481–538, 2009.
- [12] L.J.S. Allen. *An Introduction to Stochastic Processes with Applications to Biology*. Person Prentice Hall, 2003.
- [13] Charles J. Stone Paul G. Hoel, Sidney C. Port. *Introduction to Stochastic Processes*. University of California, Los Angeles, 2003.
- [14] F. Vadillo. ¿modelos deterministas o estocásticos ? *MatemáticaIa*, pages 1–7, 2011.
- [15] M. Fan M. Liu. Persistence and extinction in stochastic non-autonomous logistic systems. *J. Math*, 275, 2011.
- [16] M. Fan M. Liu. Permanence of stochastic lotka-volterra systems,. *J. Nonlinear Sci*, 27, 2017.
- [17] E. Allen. *Modelling with Itô stochastic differential equations*. Springer, Dordrecht,Netherlands, 2007.
- [18] D. J. Higham. An algorithmic introduction to numerical simulation of stochastic differential equation. *SIAM Review*, 43, 2001.

ESCUELA DE MATEMÁTICA, UNIVERSIDAD NACIONAL AUTONOMA DE HONDURAS, TEGUCIGALPA
 Email address: elvis.arrazola@unah.edu.hn

**ECUACIONES DIFERENCIALES ESTOCÁSTICAS EN EL
ESTUDIO DEL DESARROLLO DE ARBOVIRUS EN HONDURAS
A PARTIR DE UN MODELO
SUSCEPTIBLE-INFECTADO-SUSCEPTIBLE**

CARLOS A. HENRÍQUEZ

RESUMEN. Este artículo presenta un estudio sobre la aplicabilidad de ecuaciones diferenciales estocásticas. Dichas ecuaciones son importantes ya que modelan fenómenos inestables debido a la aleatoriedad de algunos de sus componentes en función de series de tiempo. En primer lugar, se estudia el movimiento browniano como proceso estocástico básico, su construcción y propiedades. Consecuentemente, se desarrolla la teoría de integración estocástica de Itô y el Teorema de existencia y unicidad de solución. Se trata también la fórmula de Itô y ejemplos clásicos de ecuaciones diferenciales estocásticas. Finalmente, se presentan aplicaciones numéricas orientadas a la solución de problemas en epidemiología, haciendo uso del modelo Euler.

1. INTRODUCCIÓN

Las ecuaciones algebraicas, diferenciales e integrales, son empleadas de forma aplicada en algunos campos de estudio: ingeniería, economía y en ciencias sociales y de la salud, entre otras, para caracterizar el estado actual de un sistema físico, económico o social y predecir su evolución en una serie de tiempo.

En general, algunos componentes de estas ecuaciones no se conocen con precisión debido a una información insuficiente o a fenómenos aleatorios inherentes del problema. Así, los problemas que surgen a partir de estas ecuaciones cuyos coeficientes son dados al azar se conocen como *Problemas Estocásticos*, surgiendo así de forma consecuente, el estudio y análisis de las Ecuaciones Diferenciales Estocásticas (EDE's) [2].

Mediante la modelación de diversos fenómenos aleatorios, el cálculo estocástico tiene amplias aplicaciones en diferentes áreas de investigación, regidos por EDE's y son definidos sobre diversos espacios. Por ejemplo: en el área de la geofísica, podemos considerar a la Tierra como una esfera de dimensión 2 y modelar una EDE sobre la Tierra; en farmacocinética, se estudia los procesos físico-químicos que sufre un fármaco cuando es administrado en algún organismo, considerando liberación, absorción, distribución y eliminación, los que están expuestos a influencias que no son del todo factibles para modelar explícitamente. Por lo tanto, existe una necesidad de extender los modelos deterministas a modelos donde se incluya un componente estocástico; en el área de ingeniería, las fuerzas que actúan en un avión

Palabras y frases clave. Movimiento browniano, Integral estocástica, Fórmula de Itô, Método de Euler.

desde el despegue hasta el aterrizaje dependen de manera compleja de las condiciones ambientales y el patrón de vuelo, estas fuerzas pueden ser modeladas por EDE's; en biología, las propiedades óseas necesarias para desarrollar articulaciones artificiales confiables varían significativamente con los individuos y sus edades; en finanzas, donde los precios de las acciones y su evolución en el tiempo dependen de un gran número de factores que no pueden ser descritos por modelos determinísticos y a menudo están regidos por una EDE [2].

En este artículo, se inicia con una extensión de lo que se conoce de la teoría de las ecuaciones diferenciales ordinarias en ecuaciones diferenciales donde se presentan elementos aleatorios. Además, se desarrolla la teoría de integración estocástica que nos permitirá dar una definición para las integrales del tipo $\int GdW$, donde G es un proceso estocástico. Estas integrales, aparecen con gran frecuencia en el estudio y resolución de estas ecuaciones. Otra herramienta útil para el estudio de las EDE's es la fórmula de Itô [4]. Por ejemplo, la solución para $dX = XdW$, cuando $X(0) = 1$ sería $X(t) = e^{W(t) - \frac{1}{2}t}$ y no $\bar{X}(t) = e^{W(t)}$, como se esperaría.

Además, se describe el teorema de existencia y unicidad de solución de EDE's definidas sobre variedades Riemannianas [2]. Estas ecuaciones modelan numerosos fenómenos en física, química, finanzas, biología y demografía, entre otros. De tal manera que, se estudiarán modelos basados en EDE's para distintos campos de aplicación. Se trabajará como objeto de estudio y análisis en el área de la epidemiología, usando un modelo SIS (sospechoso-infectado-sospechoso) [4]. Se discutirá por qué una modelización con EDE's se ajusta más que una modelización con ecuaciones diferenciales ordinarias. Finalmente, se aplicará un método de aproximación Euler, este es una adaptación del método para EDO's con la salvedad que para simular los crecimientos del movimiento browniano se utiliza una distribución normal, lo que permitirá estudiar la dinámica de propagación de enfermedades denominadas como arbovirus en zonas de mayor tendencia a considerarse como epidemias en todo nuestro país.

2. JUSTIFICACIÓN

La Universidad Nacional Autónoma de Honduras (UNAH), estipula como uno de sus ejes de investigación a la población y condiciones de vida, dentro de éste se define uno de los temas prioritarios de investigación, la salud, estado y sociedad, de esta manera, el presente estudio tiene este objetivo primordial. Además, considerando que esta investigación plantea un modelo de predicción sobre el comportamiento en el contagio de arbovirus en nuestro país, se ubica en la línea de investigación sobre Probabilidad y Procesos Estocásticos de la Orientación en Estadística de la Maestría en Matemática de la Facultad de Ciencias de la Universidad Nacional Autónoma de Honduras.

Honduras está localizada en Centroamérica, cuenta con una superficie de 112,492 Km^2 ; política y administrativamente está dividida en 18 departamentos, el idioma oficial es el español, aunque existe diversidad étnica donde algunos grupos conservan su lengua nativa como los garífunas y los misquitos ubicados en la zona costera atlántica del país. Según el Instituto Nacional de Estadística (INE), la población actual ronda las 9,416,634 personas al 11 de abril de 2021 de los cuales 4,582,981

son hombres y 4,833,653 son mujeres.

Para octubre del año 2013, la red de servicios de la Secretaria de Salud disponía de un total de 1,678 unidades prestadoras de servicios de los cuales hay: 6 hospitales nacionales, 6 hospitales regionales, 16 hospitales de área, 436 Centros Médicos Odontológicos (CMO), 1 072 centro de salud rurales (CSR), 74 Clínicas Materno-Infantil, 3 Clípper, 15 Centro Escolar Odontológicos (CEO) y 82 entre Laboratorios Regionales, IHSS y Hospitales privados que reportan a la Secretaria de Salud [10]. En la actualidad, según la Secretaría de Salud hay 1,799 Unidades de Salud.

En las últimas dos décadas, se han proliferado en nuestro país el número de casos de personas que padecen sobre alguna de las enfermedades consideradas como arbovirus, que son virus que se transmiten al ser humano o a otros vertebrados por ciertas especies de artrópodos hematófagos, especialmente insectos (moscas y mosquitos) y arácnidos (garrapatas). Los arbovirus no forman parte del sistema de clasificación actual de los virus, que se basa en la naturaleza y estructura del genoma viral.

Los mosquitos del genero *Aedes* (*aegypti* y *albopictus*) son los vectores de varios arbovirus como: Chikungunya, Dengue, Fiebre Amarilla, Occidente del Nilo y el Zika. El Dengue y el Chikungunya han pasado a ser dos enfermedades con repercusiones importantes en los servicios de atención de salud y en los programas de control de vectores por los impactos que han ocasionado. Las dos enfermedades están presentes en nuestro país y la prevención y control de las mismas dependen en gran medida del logro de la participación comunitaria en forma activa [12].

El territorio nacional tiene una temperatura anual promedio entre 18 y 30 °C, además de una humedad de alrededor del 80%, es altamente favorable para la longevidad, fecundidad y desarrollo de los mosquitos del genero *Aedes*. Las condiciones de pobreza y pobreza extrema, los servicios insuficientes para de recolección de basura, la falta de disponibilidad de agua potable las 24 horas del día, las condiciones insalubres de la vivienda, generan las condiciones para que los criaderos de *Aedes* proliferen al interior y en los alrededores de la viviendas.

La Secretaria de Salud tiene varias décadas luchando contra la enfermedad del dengue y desde el año 2013 estaba en estado de alerta sanitaria para prevenir y controlar la enfermedad del Chikungunya. En el año 2015, inicio la alerta para prevenir y controlar la enfermedad del Zika. Desde ese instante se han implementado algunas medidas, sin embargo, estas actividades no han producido los efectos esperados sobre el control de estas enfermedades. A pesar de tener una experiencia de 50 años en tratar de disminuir los índices de infestación de *Aedes* en las viviendas a niveles que no presenten riesgos de transmisión, hasta la fecha esto no ha sido posible.

En nuestro país, las cifras son alarmantes y si bien es cierto que la cantidad de decesos debido a estos padecimientos no es relativamente alto, es probable que los recursos del Estado se podrían invertir de mejor manera. Con este estudio, se pretende predecir a partir de los datos que se tienen sobre la cantidad de personas

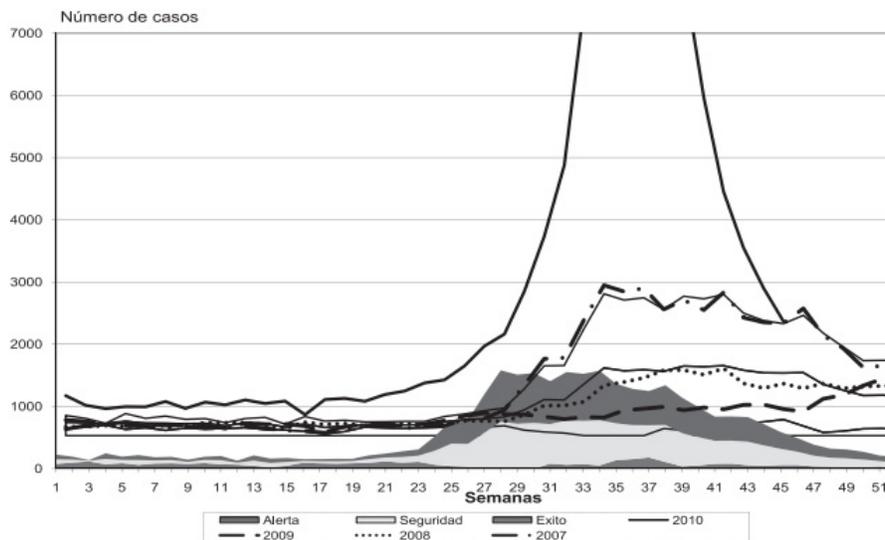
afectadas y un modelo basado en Ecuaciones Diferenciales Estocásticas el continuo comportamiento de dichas afecciones con el propósito de que a quien corresponda, puedan tomar decisiones de alto nivel en función a una administración más eficiente de los recursos y con ello, fortalecer todo el sistema de salud.

3. ANTECEDENTES

En Honduras, los primeros casos de dengue clásico se presentaron durante la epidemia de 1978 causada por el primer serotipo del mosquito *Aedes aegypti*, donde ocurrieron alrededor de 134,000 casos en todo el país, afectando principalmente a poblaciones de los departamentos de Islas de la Bahía, Cortés, Colón y Valle. También hubo brotes localizados en poblaciones del Valle de Comayagua y a lo largo de la carretera de Olancho. Cinco personas murieron en San Pedro Sula con enfermedad hemorrágica, dos de ellas con síndrome de choque; desafortunadamente no se hicieron cultivos, ni se tomaron muestras para confirmar el diagnóstico de dengue. Posteriormente en 1987, se presentó la segunda epidemia en la ciudad de Choluteca estimándose en 28,606 los casos ocurridos. El serotipo predominante aislado siguió siendo el mismo, pero también se aisló el cuarto serotipo de 32 pacientes febriles probablemente importado de El Salvador por la presencia de este serotipo en ese país. En 1988, se detecta un comportamiento epidémico en la ciudad de Yusguare, Choluteca con alrededor de 300 casos y en 1989 se registra la tercera epidemia en la ciudad de Tegucigalpa, con alrededor de 1,300 casos no pudiéndose cuantificar la magnitud del brote ni el serotipo circulante [11].

El patrón epidemiológico de la fiebre del dengue en Honduras durante los últimos 32 años ha mostrado dos distintas etapas: brotes epidémicos en áreas localizadas (1978-1994) y circulación viral endémica o epidémica a nivel nacional con predominio de las Regiones Sanitarias Metropolitanas de Tegucigalpa y San Pedro Sula (1995-2010). Durante el periodo 1991-2010 se puede observar que los casos de dengue clásico han experimentado un patrón ascendente con una tendencia de incremento lineal con presencia de brotes epidémicos cada tres a cuatro años a partir de 1991 (3,045 casos), 1995 (18,152 casos), 1998 (21,359 casos), 2002 (32,269 casos), 2007 (29, 328 casos) y 2010 (63,477 casos) hasta la semana epidemiológica 39, Fig(1).

En estudios más recientes hasta la semana epidemiológica 48 (24-30 de noviembre) de 2019 se notificaron 105,968 casos de los cuales, corresponden a: Dengue: 105,513 (99.57%), Dengue Sin Signos de Alarma: 86,185 (81.33%), Dengue Grave: 19,328 (18.24%), Chikungunya: 218 (0.21%) y Zika: 237 (0.22%) [7]. Según este informe, los casos de dengue sin signos de alarma representan epidemia a nivel regional en toda Honduras, de tal manera que 79,786 del total de casos (92.6%) ocurrieron en 12 regiones sanitarias: Yoro con 14,101 casos, la zona Metropolitana del Distrito Central con 11,733 casos, Cortés 10,571 casos, Olancho 8,218 casos, Atlántida con 7,684 casos, Colón con 5,173 casos, Choluteca con 5,262 casos, Santa Bárbara con 3,694 casos, Copán con 3,644 casos, Municipalidad de San Pedro Sula (MSPS) 3,573 casos, Comayagua con 3,333 casos y La Paz con 2,800 casos [7].



Fuente: Programa Nacional de Dengue, Secretaría de Salud Honduras
 *2010 hasta la semana epidemiológica 37

FIGURA 1. Canal endémico de dengue en Honduras, período 2005-2009. Curvas epidémicas de dengue 2007-2010*

Con base en estos datos alarmantes, en este estudio se realizará la simulación a través de modelos matemáticos que permitan explorar el impacto del comportamiento en la dinámica de la transmisión de arbovirus, proporcionando información valiosa para la toma de decisiones con el objetivo de controlarlas o erradicarlas.

Hay dos maneras de lograr la eficacia general de un programa para controlar el brote de una enfermedad: un análisis retrospectivo de datos y la simulación utilizando modelos matemáticos [16]. Diversas investigaciones demuestran que el uso de modelos matemáticos para describir la dinámica de transmisión de las enfermedades infecciosas son una herramienta necesaria para realizar un buen análisis de la relación costo-eficacia en la aplicación de medidas de control en la propagación de una enfermedad [14] y [15].

En ese sentido, las ecuaciones diferenciales ordinarias (EDO's) han constituido una herramienta matemática fundamental para modelar sistemas físicos. Las soluciones de estas ecuaciones son curvas regulares que representan el estado del sistema en cada instante. Sin embargo, parece razonable modificar el modelo dado de tal manera que se pueda incluir efectos aleatorios (ruido) que perturben el sistema en cada instante. Es decir, que el vector tangente a la curva, si existiese, estaría determinado por una parte determinista pero también por una componente aleatoria que debemos determinar caracterizada por el ruido, lo que es razonable representar con el movimiento Browniano.

4. PRELIMINARES

4.1. Movimiento Browniano. Conocer el movimiento browniano es fundamental para comprender las EDE's, ya que interviene en ellas dotándolas de un componente aleatorio. Para describir el movimiento, matemáticamente es común utilizar el concepto de un proceso estocástico $B_t(\omega)$, interpretado como la posición en el tiempo t de un grano de polen ω . Se generalizará ligeramente y considerará una analogía n -dimensional. Para construir $\{B_t\}_{t \geq 0}$ usamos el teorema de extensión de Kolmogorov: [8]

Teorema 4.1 (Teorema de extensión de Kolmogorov). *Para todo $t_1, \dots, t_k \in T, k \in N$, sea v_{t_1, \dots, t_k} una familia de medidas de probabilidad sobre R^{nk} tal que,*

$$(4.1) \quad v_{t_{\sigma(1)}, \dots, t_{\sigma(k)}}(F_1 \times \dots \times F_k) = v_{t_1, \dots, t_k}(F_{\sigma^{-1}(1)} \times \dots \times F_{\sigma^{-1}(k)})$$

Para toda permutación σ sobre $\{1, 2, \dots, k\}$, y,

$$(4.2) \quad v_{t_1, \dots, t_k}(F_1 \times \dots \times F_k) = v_{t_1, \dots, t_k, t_{k+1}, \dots, t_{k+m}}(F_1 \times \dots \times F_k \times R^m \times \dots \times R^m)$$

Para todo $k \in N$, donde el conjunto del lado derecho tiene $k+m$ factores. Entonces, existe un espacio de probabilidad (Ω, \mathcal{F}, P) y un proceso estocástico $\{X_t\}$ sobre Ω , $X_t : \Omega \rightarrow R^n$ tal que,

$$v_{t_1, \dots, t_k}(F_1 \times \dots \times F_k) = P[X_{t_1} \in F_1, \dots, X_{t_k} \in F_k]$$

Para todo $t_i \in T, k \in N$ y todo conjunto de Borel F_i .

Para trabajar con el movimiento browniano se modelará como una cadena de Markov. La idea intuitiva de cadena de Markov es que dado el valor $X(s)$ se puede predecir a partir de él, los futuros valores de $X(t)$ también como si se tuviera la historia completa de valores que toma el proceso, anterior al tiempo s . Por eso, los procesos solamente conocen su valor en el tiempo s pero no se sabe cómo han llegado hasta ese punto [4].

Definición 4.2. El proceso $\{B_t\}_{t \geq 0}$ sobre (Ω, \mathcal{F}, P) se dice que es un movimiento browniano [6], si cumple:

- (1) $B_0 = 0$, es decir, $P(\{w : B_0(w) \neq 0\}) = 0$
- (2) $B_t \sim N(0, t)$. Si $s < t$ se tiene que $B_t - B_s \sim N(0, t - s)$ o de forma alternativa, $B_t - B_s \sim \sqrt{t - s}N(0, 1)$.
- (3) $\{B_t\}_{t \geq 0}$ tiene incrementos independientes y estacionarios.
- (4) B_t es continua en casi toda parte. (Teorema de continuidad de Kolmogorov, vea Øksendal(2000, p. 14))

4.2. Simulación de un modelo browniano. Simular ecuaciones diferenciales estocásticas es en general difícil, pues no existe un método canónico para las ecuaciones de este tipo. Sin embargo, podemos aprovechar la forma de la ecuación en el sistema de Itô para el modelo y aproximarla con métodos conocidos.

Suponga que para un modelo epidémico se derivan de un proceso de difusión X_t que varía con una distribución normal y las variables aleatorias del sistema son continuas en un intervalo $[0, T]$ y que cumple las propiedades de la definición 4.2. Y que se pueden derivar de ecuaciones diferenciales parciales de Kolmogorov hacia adelante y hacia atrás para la función de densidad de probabilidad de transición

y, estas a su vez, conducen directamente a las ecuaciones diferenciales estocásticas [6], es decir, se modela su variación como,

$$dX_t = \sigma dB_t$$

para $\sigma > 0$. Los elementos de difusión de esas pequeñas partículas que conducen la epidemia en el seno de un fluido, son modeladas como la variación en el tiempo de un movimiento browniano B_t . Cuando σ es constante la solución viene dada por,

$$X_t = X_0 + \sigma B_t$$

Una derivación informal de las EDE's correspondientes al modelo epidémico es denotar el cambio en las variables del sistema. El sistema de las EDE's tiene una forma donde B_t es un vector con procesos independientes de Wiener.

$$\frac{dX_t}{X_t} = \sigma dB_t \implies dX_t = \sigma X_t dB_t$$

Integrando en ambos extremos tendríamos,

$$X_t = X_0 + \sigma \int_0^t X_s dB_s$$

Esta última expresión conduce a describir el concepto de integrales estocásticas que junto a la fórmula de Itô representan herramientas valiosas en la resolución de EDE's.

4.3. Integración estocástica. El proceso estocástico más destacado es sin duda el movimiento browniano, el cual no es diferenciable en ningún punto. En este artículo se estudiará la integración estocástica para ser capaces de comprender las EDE's [4]. Para ello, definamos la siguiente EDE,

$$\begin{cases} [l]dX = b(X, t)dt + B(X, t)dW \\ X(0) = X_0 \end{cases}$$

Lo que nos generaría una solución a partir de la expresión,

$$X_t = X_0 + \int_0^t b(X, s)ds + \int_0^t B(X, s)dW$$

No se puede entender $\int_0^T B(X, s)dW$ como una integral ordinaria. Para el caso, si $G = W$, generando un procedimiento para conocer $\int_0^T WdW$. El procedimiento para aproximar dicha integral es utilizar series de Riemann y a partir de ellas construir una aproximación [4],

$$(4.3) \quad R_n \rightarrow \frac{W^2(T)}{2} + (\lambda - \frac{1}{2})T$$

La construcción de la integral estocástica es similar a la construcción de la integral de Riemann. La aparición explícita de λ es debido a la falta de regularidad que existe en dicha integral, como consecuencia de la diferencial del browniano. Con lo cual, la definición de integral estocástica de Itô es el caso particular en el que $\lambda = 0$,

$$(4.4) \quad \int_0^T WdW = \frac{W^2(T)}{2} - \frac{T}{2}$$

Existen otras definiciones para diferentes valores de λ como la integral de Stratonovich, en el que $\lambda = \frac{1}{2}$,

$$(4.5) \quad \int_0^T W dW = \frac{W^2(T)}{2}$$

4.4. Fórmula de Itô. Esta fórmula es importante ya que es una herramienta útil para resolver EDE's y comprobar su solución. Para ello, se mostrará el siguiente teorema, véase [8] y [9].

Teorema 4.3. *Sea $X(t)$ un proceso estocástico que tiene diferencial estocástica,*

$$dX = Fdt + GdW$$

donde $F \in L^1(0, T)$, $G \in L^2(0, T)$. Sea $u : [0, T] \rightarrow \mathbb{R}$ una función continua tal que existen derivadas parciales $\frac{\partial u}{\partial t}$, $\frac{\partial u}{\partial x}$, $\frac{\partial^2 u}{\partial x^2}$ y son continuas. Denotamos,

$$Y(t) = u(X(t), t)$$

Entonces, Y tiene diferencial estocástica,

$$(4.6) \quad dY = \frac{\partial u}{\partial t} dt + \frac{\partial u}{\partial x} dX + \frac{1}{2} \frac{\partial^2 u}{\partial x^2} G^2 dt = \left(\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} F + \frac{1}{2} \frac{\partial^2 u}{\partial x^2} G^2 \right) dt + \frac{\partial u}{\partial x} G dW$$

Se invita al lector a ver la demostración en [4].

4.5. Existencia y unicidad. Vamos a analizar la existencia y unicidad de soluciones de ecuaciones diferenciales estocásticas del tipo,

$$(4.7) \quad \frac{dX_t}{dt} = b(t, X_t) + B(t, X_t)W_t, \quad b(t, x) \in \mathbb{R}, B(t, x) \in \mathbb{R}$$

donde W_t representa el ruido blanco 1-dimensional. Recordemos que la interpretación de Itô de la ecuación (4.7) es que X_t verifique la ecuación en forma diferencial, [1]

$$(4.8) \quad dX_t = b(t, X_t)dt + B(t, X_t)dB_t, \quad W_t = \frac{dB_t}{dt}$$

Por ello, existe una expresión que nos garantiza la existencia y unicidad para dichas soluciones,

Teorema 4.4. *Sean $b : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$ y $B : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$ funciones continuas que cumplen las siguientes condiciones:*

(i)

$$\begin{aligned} |b(x_1, t) - b(x_2, t)| &\leq l|x_1 - x_2| \\ |B(x_1, t) - B(x_2, t)| &\leq l|x_1 - x_2| \\ \forall x_1, x_2 \in \mathbb{R}, 0 \leq t \leq T \text{ y cierta constante } l > 0 \end{aligned}$$

(ii) X_0 es una v.a. que toma valores reales tal que $E(|X_0|^2) < +\infty$

(iii) X_0 es independiente de $\mathcal{W}^+(0) = \mathcal{U}(W(s) - W(0) : s \geq 0)$, con $W(t)$ un movimiento browniano unidimensional.

Entonces, existe una única solución $X(t) \in L^2(0, T)$ de la EDE,

$$(4.9) \quad dX = b(t, X)dt + B(t, X)dW, \quad X(0) = X_0$$

En *i*) del teorema de existencia y unicidad debemos resaltar que b y B son uniformemente Lipschitz continuas en la variable x , [4].

5. ESTOCASTICIDAD DEMOGRÁFICA

Existen tres métodos distintos de formular un modelo epidemiológico estocástico, los cuales se distinguen de acuerdo a los supuestos acerca del tiempo y el espacio de estados. En primer lugar, se describe la modelación con una cadena de Markov a tiempo discreto (CMTD), donde el tiempo y el espacio de estados son discretos. Se muestra también, un modelo usando una cadena de Markov a tiempo continuo (CMTC), donde el tiempo es continuo pero el espacio de estados es discreto. Por último, se aborda el problema mediante un sistema de ecuaciones diferenciales estocásticas (EDE's) basado en un proceso de difusión, donde tanto el tiempo como el espacio de estados son continuos [17]. En este artículo, se presenta el modelo de EDE's y la inferencia el modelos epidemiológicos compartimentales de este tipo. En la Sección 5.1, se describe brevemente los modelos compartimentales estocásticos y en la Sección 5.2, se describe el caso particular del modelo SIS (susceptible-infectado-susceptible).

5.1. Modelos compartimentales estocásticos. Son modelos que incorporan al modelo epidemiológico aleatoriedad a nivel poblacional, llamada estocasticidad demográfica. Esta consideración se refiere a que los tiempos de entrada y salida de los compartimentos tienen una distribución probabilística, lo cual se refleja en la aleatoriedad de la solución del sistema a un tiempo fijo t . Por ello, considerar la estocasticidad demográfica resulta sumamente relevante al trabajar con poblaciones relativamente pequeñas.

5.2. Modelo Susceptible-Infectado-Susceptible metapoblacional. Modelos SIS epidemiológicos de compartimentos se definen a partir de clases y subclases. Las tasas de transición entre las clases para estos modelos se estiman a partir del conocimiento cualitativo y evidente en la historia natural de la enfermedad [18]. Para este estudio, se eligió un modelo discreto metapoblacional de transmisión de enfermedades SIS. En este tipo de modelos, la población se divide en dos grupos de personas: las que han sido infectadas por la enfermedad y son infecciosas y las que son susceptibles de ser infectadas por la enfermedad [13].

Los modelos SIS se usan para enfermedades en las que no hay inmunidad, pues, una vez que las personas infectadas se recuperan, pasan a ser de nuevo susceptibles. Por lo tanto, la progresión de la enfermedad desde el punto de vista de un individuo es susceptible-infectado-susceptible. En este modelo, la población se podría dividir en dos parches que pueden representar ciudades o regiones geográficas o comunidades que están interconectadas por los procesos de inmigración y emigración [19]. En el caso especial de nuestro país, podríamos dividirlo en dos regiones sanitarias, la zona noroccidental y la centro-sur-oriente para llevar a cabo el análisis.

6. PRESENTACIÓN DEL MODELO

Los modelos matemáticos de EDE's tienen un papel muy importante en aplicaciones a múltiples áreas como en la epidemiología. Se propone dar un método numérico sencillo para simular EDE's con el soporte del programa MATLAB. Así mismo, se presentará una simulación aplicada. El método numérico que utilizaremos es el método de Euler. Dicho método es análogo al utilizado para ecuaciones diferenciales ordinarias.

6.1. **Método de Euler.** Sea una EDE de la forma (4.9) o su equivalente,

$$X(t) = X_0 + \int_0^t b(X, s)dt + \int_0^t B(X, s)dW, \quad 0 \leq t \leq T$$

Vamos a definir el método numérico de Euler para resolver la EDE anterior. Para ello, comenzamos discretizando el intervalo $[0, T]$ en el que estamos trabajando. Llamamos $\delta t = \frac{T}{N}$, donde N es un entero positivo para dividir el intervalo.

Entonces, el método de Euler aproxima la solución del siguiente modo,

$$(6.1) \quad X_j = X_{j-1} + b(X_{j-1}, t)\delta t + B(X_{j-1}, t)(W(T) - W(T-1)), \quad j = 1, 2, \dots, N$$

Se sabe que los incrementos del movimiento brownianos siguen una distribución $N(0, \delta t)$. Entonces, se podrán simular estos incrementos utilizando dicha distribución. Con lo cual, la aproximación resulta del siguiente modo,

$$(6.2) \quad X_j = X_{j-1} + b(X_{j-1}, t)\delta t + B(X_{j-1}, t)N(0, \delta t)$$

De esta manera, podemos observar que computacionalmente no resultará muy complejo simular una EDE, una vez conocido el método que emplearemos.

6.2. **Simulación en epidemiología.** La epidemiología estudia el control, distribución, frecuencia y factores determinantes de las enfermedades infecciosas. En ese sentido, existen una gran variedad de modelos matemáticos aplicados a epidemias. Nos centraremos en el modelo SIS adaptado a EDE's [4]. La población de individuos se divide en susceptibles $S(t)$ e infectados $I(t)$. Los individuos susceptibles pueden ser infectados al estar en contacto con la enfermedad. Además, en el modelo SIS, los individuos infectados pueden recuperarse, pasando a ser susceptibles. Así, se pueden plantear las ecuaciones,

$$\begin{aligned} \frac{dS}{dt} &= -\beta I(t)S(t) + \gamma I(t) + \mu N - \mu S(t) \\ \frac{dI}{dt} &= \beta I(t)S(t) - (\mu + \gamma)I(t) \end{aligned}$$

Donde β es la tasa de contagio, N el número total de individuos, γ es la tasa de recuperación de individuos infectados y μ la tasa de mortalidad por individuo.

Se pretende trasladarlo al caso de las EDE's debido a que en el mundo real estamos expuestos a múltiples factores que influyen en el comportamiento de las enfermedades y los organismos. Como consecuencia, no siempre es posible realizar una predicción de que va a suceder; y, por ello, en el caso de pequeñas poblaciones es más adecuado utilizar un modelo estocástico. El modelo que vamos a usar asume que los cambios que se producen en el crecimiento de la población son muy variables debido a la aleatoriedad de los nacimientos y muertes que inciden en la población. En ese sentido, el modelo estocástico extraído de [22] es el siguiente,

$$\begin{aligned} dS &= (-\beta I(t)S(t) + \gamma I(t) + \mu N - \mu S(t))dt + \sqrt{\beta I(t)S(t) + (\mu + \gamma)I(t)}dW \\ dI &= (-\beta I(t)S(t) - (\mu + \gamma)I(t))dt + \sqrt{\beta I(t)S(t) + (\mu + \gamma)I(t)}dW \end{aligned}$$

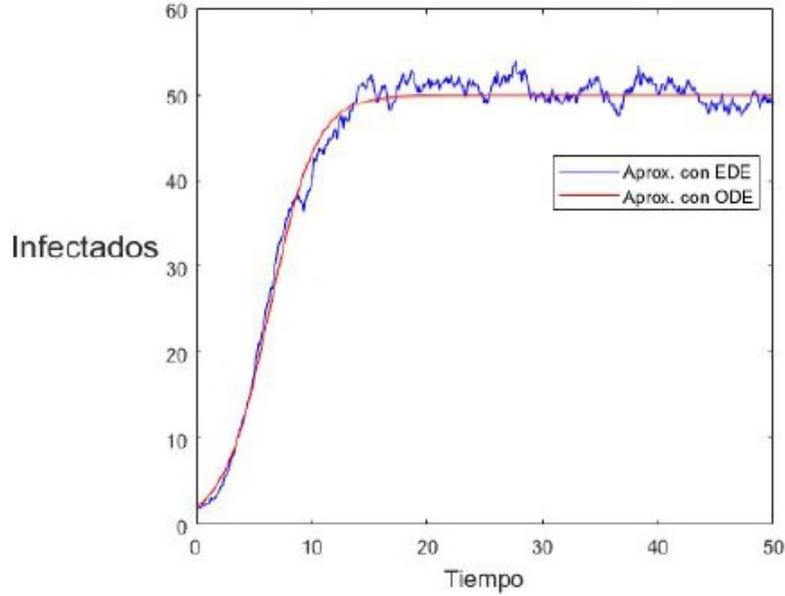


FIGURA 2. Simulación de individuos afectados en una población por un virus determinado

Así, combinando las dos ecuaciones y utilizando que $S + I = N$, llegamos a la EDE,

$$(6.3) \quad dI = (\beta I(t)(N - I(t)) - (\mu + \gamma)I(t))dt + \sqrt{\beta I(t)(N - I(t)) + (\mu + \gamma)I(t)}dW$$

Por tanto, la versión determinista de la ecuación (6.3) sería,

$$(6.4) \quad \frac{dI}{dt} = I(t)(\beta N - \beta I(t) - \mu - \gamma)$$

A continuación se hace una simulación para comparar la EDO en (6.4) con la EDE en (6.3), los valores de las constantes han sido extraídas de [22]. En la fig(2), podemos comparar los resultados obtenidos por el modelo determinístico a partir de la EDO en (6.4) y el modelo estocástico a partir de la EDE en (6.3) y la notable presencia de ruido blanco, que resultaría más conveniente con el propósito de aplicar este modelo a la realidad en nuestro país por la afección por arbovirus considerando el ruido blanco.

En la simulación se utilizaron los siguientes valores constantes, para $I_0 = 2$, $N = 100$, $\beta = 1$, $\mu = \gamma = 0.25$. Se puede observar que en este caso debido a los valores de las constantes, el número de infectados crece rápidamente hasta el instante $t = 15$. A partir de ese instante, el número de infectados se mantiene estable entorno a 50 infectados. Como se mencionó, el modelo de EDE's resalta, en su simulación, las fluctuaciones del crecimiento de una población pequeña. Y justamente, el modelo que se utilizará para el estudio del arbovirus será el definido en (6.3).

7. CONCLUSIONES

- (1) Una ecuación diferencial estocástica es una ecuación diferencial cuyos coeficientes son números aleatorios o funciones aleatorias de las variables independientes. La forma más común de una EDE es una ecuación diferencial ordinaria donde el extremo derecho es perturbado por un término que depende de una variable aleatoria o ruido blanco.
- (2) El uso de ecuaciones diferenciales estocásticas para modelar requiere de una cantidad de trabajo importante y de un extenso conocimiento de procesos estocásticos. En ocasiones, es posible agregar elementos probabilísticos, ya que hacen que el modelo aproxime de mejor forma a la realidad. Por tanto, las ecuaciones diferenciales estocásticas representan una herramienta adecuada para describir sistemas con ruido externo.
- (3) La simulación numérica es una herramienta muy útil para comprender mejor el comportamiento de la EDE's. Por esto, a pesar de que el campo de las matemáticas que se dedica a encontrar las soluciones numéricas de dichas EDE's, y especialmente las ecuaciones diferenciales parciales estocásticas, es relativamente joven, se han construido algoritmos rápidos para simular las soluciones de estos sistemas dinámicos. Sin embargo, estos algoritmos no son triviales y son mucho más que una simple generalización, lo que significa que aún es necesaria mayor investigación y nuevas ideas para tener algoritmos que tengan un menor error de aproximación que los actuales.

8. RETOS EN EL FUTURO CERCANO

La propagación de enfermedades infecciosas y sus medidas de control han sido objeto de diversos estudios, la mayoría lo han hecho en modelos para la dinámica en una población. El propósito de este artículo es utilizar modelos metapoblacionales por medio de ecuaciones diferenciales estocásticas y todas las herramientas matemáticas necesarias que permitan obtener resultados que refuercen una estrategia de Estado para el mejor aprovechamiento de los recursos.

Estos modelos permiten conocer la influencia de los flujos migratorios en la propagación de una enfermedad y comprender las características de la propagación en subpoblaciones, cada una con su propia dinámica pero conectadas por el movimiento de personas entre sí. Los datos han demostrado que el dengue ha causado, en volumen, la inversión de un sinnúmero de recursos del Estado y en la economía familiar; esta afección, como parte de las distintas enfermedades clasificadas como arbovirus está presente en nuestro diario vivir y no basta con la promoción de campañas preventivas, debemos controlar de mejor manera la forma en cómo combatirlo.

En este sentido, con este estudio se pretende establecer un modelo matemático que proporcione nuevos elementos para contrarrestar esta y, muy probablemente, otras afecciones. Por ello, existen retos que debemos solventar para obtener resultados óptimos:

- Se está realizando la gestión en la Secretaría de Salud para obtener datos que validen la situación real de dicha afección, más en este momento que

la mayor parte de los recursos están destinados al combate de la pandemia de SARS-Cov-2.

- Es indispensable comprender el fundamento teórico-práctico de los modelos de solución numérica, así como hacer un análisis previo de las EDE's propuestas para elegir adecuadamente el método numérico a utilizar. Esto implica que se deben realizar varias pruebas con métodos diferentes para determinar que técnica modela mejor el problema.
- El control de una enfermedad infecciosa se hace mediante la intervención directa o la reducción de la susceptibilidad de una población. Al analizar el modelo propuesto, observamos que se puede ejercer control sobre algunos parámetros que representen la aplicación de estrategias para un control óptimo de la enfermedad. Se puede controlar el flujo de personas infectadas entre las dos poblaciones, o zonas sanitarias como se propone en este estudio, capturado por los parámetros establecidos. Se puede tomar como variables de control la proporción de la población a la que se le aplicará algún tipo de tratamiento; el impacto de las medidas de intervención directa, como el distanciamiento social y la limpieza constante de los hogares, entre otros.
- Se podría considerar el estudio de modelos con cadenas de Markov de tiempo discreto y un modelo con cadenas de Markov de tiempo continuo. En el primer modelo, tanto el tiempo como las variables de estado son discretas. En el segundo modelo, el tiempo es continuo y las variables de estado siguen siendo discretas. Inclusive, considerar modelos SIR (susceptible-infectado-recuperado) y SEIR (susceptibles-expuestos-infectados-recueperados), de este último hay estudios predictivos para la situación pandémica provocada por el SARS-CoV-2.
- Los modelos epidemiológicos son métodos muy utilizados para la proyección de casos en enfermedades infecciosas; sin embargo, se debe tener en cuenta que son modelos deterministas que pueden utilizar parámetros supuestos y podrían generar resultados imprecisos. Por ello, será necesario realizar las pruebas pertinentes que permitan mitigar esos efectos.

REFERENCIAS

1. Baqueiro, L. (2018). Ecuaciones diferenciales estocásticas con aplicaciones en finanzas. Trabajo Fin de Grado. Departamento de Ecuaciones Diferenciales y Análisis Numérico. Universidad de Sevilla, España.
2. Begazo, F. (2019). Fórmulas de Itô Para Flujos Estocásticos y Sus Aplicaciones. Tesis de Posgrado. Unidad de Posgrado de la Facultad de Ciencias Naturales y Formales. Universidad Nacional de San Agustín de Arequipa, Perú.
3. Hull, John C. (2015). *Options, futures, and other derivatives*. Ninth edition. University of Toronto, Canada.
4. Mirones, O. (2019). Ecuaciones Diferenciales Estocásticas y Aplicaciones. Trabajo Fin de Grado. Facultad de Ciencias. Universidad de Cantabria, España.
5. Bravo, R. (2015). Ecuaciones diferenciales estocásticas en modelos de farmacocinética. Trabajo Fin de Grado. Facultad de Ciencias Físico-Matemáticas. Benemérita Universidad Autónoma de Puebla, México.
6. Rodríguez, Ó. (2016). Serie estocástica de Taylor-Itô y métodos numéricos para ecuaciones diferenciales estocásticas. Boletín de Matemáticas, Vol. I, Número 23, pp. 81-103. Universidad Nacional de Colombia. Bogotá, Colombia.
7. Unidad de Vigilancia de la Salud (2019). Boletín Informativo. Vol. 57, número 48, 8 de diciembre. Secretaría de Salud. Tegucigalpa, Honduras.

8. Øksendal, Bernt (2000). *Stochastic Differential Equations: An Introduction with Applications*. Vol. 82, Month 01, isbn 978-3-540-60243-9. Journal of the American Statistical Association. Doi 10.1007/978-3-662-03185-8.
9. Marín S., Freddy H. (2007). *Ecuaciones Diferenciales Estocásticas y casos de aplicación en Finanzas*. Tesis de posgrado. Departamento de Ciencias Básicas, Universidad EAFIT. Medellín, Colombia.
10. *Análisis de la situación de salud, Honduras 2014*. Departamento de Epidemiología, Subsecretaría de Riesgos Poblacionales, Secretaría de Salud. Tegucigalpa, Honduras.
11. Ávila, G.; Araujo, R.; Orellana, G. (2010) *Situación Epidemiológica del Dengue en Honduras, 1991-2010*. Revista Médica Hondureña. Vol. 78, N° 3. Programa Nacional de Dengue, Secretaría de Salud. Tegucigalpa, Honduras.
12. *Proyecto MSS Arbovirosis*. Proyectos Investigación Médica Servicio Social (2016). Biblioteca Virtual en Salud, Biblioteca Médica Hondureña. Universidad Nacional Autónoma de Honduras (UNAH). Tegucigalpa, Honduras.
13. Alpizar-Brenes, G. (2016) *Análisis de un modelo SIS para el estudio de la dinámica de propagación de la enfermedad al aplicar medidas de control*. Tecnología en Marcha. Edición especial. Matemática Aplicada, Mayo 2016. Pág 42-50. Escuela de Matemática. Instituto Tecnológico de Costa Rica.
14. Kretzschmar, M. & Wallinga, J. (2010). *Mathematical models in infectious disease epidemiology*. En Modern Infectious Disease Epidemiology (214-226). Springer. Alexander Krämer, Mirjam Kretzschmar, Klaus Krickeberg (Eds.), New York, 2010.
15. González, P., Lee, S., Velázquez, L. & Castillo-Chavez, C. (2011). *A note on the use of optimal control on a discrete time model of influenza dynamics*. Mathematical Biosciences and Engineering, 8, 183-197.
16. Welte, R., Leidl, R., Greiner, W. & Postma, M.(2010). *Health economics of infectious diseases*. En Modern Infectious Disease Epidemiology (pp. 253-279). Springer. Alexander Krämer, Mirjam Kretzschmar, Klaus Krickeberg (Eds.), New York, 2010.
17. Allen, L. J. (2008). *An introduction to stochastic epidemic models*. In Mathematical epidemiology, pages 81–130. Springer, 2008. (Citado en páginas 21 y 35).
18. Brauer, F., Castillo-Chavez, C., De la Pava, E., Castillo-Garsow, C., Chowell, D., Espinoza, B., González, P., Hernández, C. & Moreno, V. (2014). *Modelos de la propagación de enfermedades infecciosas*. Universidad Autónoma de Occidente. Cali, Colombia.
19. Arino, J. & Driessche, P. (2006). *Disease spread in metapopulations*. Fields Institute Communications, 48, 1-13.
20. Greenhalgh, D., Liang, Y. & Mao, X. (2015). *Demographic Stochasticity in the SDE SIS Epidemic Model*. University of Strathclyde, Glasgow, UK.
21. Ávila-Pozos, R., Jimenez-Munguía, R.R. & Telmotzi-Ávila, R. (2019). *Modelos estocásticos en epidemiología*. Boletín Científico del Instituto en Ciencias Básicas e Ingeniería. Publicación semanal PADI N. 12, Pág. 95-101. Universidad Autónoma del Estado de Hidalgo, México.
22. Ortega-Lenis, D., Arango-Londoño, D., Muñoz, E., Caicedo, D., Mena, J., Torres, M. & Méndez, F. (2020). *Predicciones de un modelo SEIR para casos de COVID-19 en Cali, Colombia*. Revista de Salud Pública N. 22(2), Pág. 1-6. Cali, Colombia. DOI: <https://doi.org/10.15446/rsap.V22n2.86432>

UNIVERSIDAD NACIONAL AUTÓNOMA DE HONDURAS (UNAH), CIUDAD UNIVERSITARIA, TEGUCIGALPA, HONDURAS

Dirección actual: Escuela de Matemática y Ciencias de la Computación, Facultad de Ciencias, UNAH

Dirección de correo electrónico: c.henriquez@unah.hn

APLICACIONES DE REDES NEURONALES

CRISTIAN F. JUÁREZ

ABSTRACT. In the following document, a bibliographic review about the first advances that served as the theoretical base for the development of artificial neuronal pathways was carried out. Including concepts and ideas that remain relevant and current for its respective computational implementation. Furthermore, the internal functioning and the conditions that must be met for optimal performance are shown. The prediction outlines used in temporal data series where established and consequently an example of its application to logistical mapping is exhibited, using one model of recurrent neuronal pathways

RESUMEN

En este trabajo, se realizó una revisión bibliográfica de los primeros avances que sirvieron como base teórica al desarrollo de las redes neuronales artificiales, conceptos e ideas que aún siguen vigentes para su respectiva implementación computacional. Además se muestra su funcionamiento interno y las condiciones que deben cumplirse para que éste sea óptimo. Se establecieron los esquemas de predicción utilizados en datos de series temporales y posteriormente se muestra un ejemplo de aplicación al mapeo logístico, utilizando un modelo de redes neuronales recurrentes.

Palabras Claves: Redes Neuronales, Entrenamiento, Series Temporales.

1. INTRODUCCIÓN

Con el desarrollo de máquinas de cómputo cada vez más sofisticadas, tareas que parecían difíciles de realizar, se volvieron más asequibles al crear algoritmos que permitieron en un número finito de pasos encontrar una respuesta satisfactoria en una medida de tiempo prudente y con un consumo aceptable de recursos (memoria, energía, etc). Estos tipos de sistemas poseen un alto grado de sofisticación debido a que se contemplan todos los posibles escenarios o casos particulares que pueden llegar a presentarse. Dichos sistemas son conocidos como *Expertos* y son un intento de emular la inferencia lógica que se ejerce mediante el razonamiento. Los algoritmos entonces son un conjunto de reglas que se ejecutan dependiendo de las condiciones que se deben cumplirse[5].

Sin embargo, no todos los problemas son de esta índole, es decir, problemas que requieran encontrar valores específicos, por citar algunos ejemplos: el cálculo de las raíces de un polinomio, los autovalores de una matriz, la solución numérica de una ecuación diferencial o la multiplicación de números con una gran cantidad de cifras.

Problemas como la categorización de objetos, de acuerdo a características comunes o el reconocimiento de patrones en un conjunto de datos de naturaleza aleatoria resultan ser casi imposibles de abordar mediante esta estructura de sentencias lógicas,

son problemas cuya esencia es distinta a la de aquellos que requieren una gran capacidad de cómputo [5]. Es necesario entonces definir una nueva perspectiva que permita ser cifrada mediante el lenguaje matemático y que genere una solución a dichas cuestiones [3,4].

Los algoritmos representan órdenes concretas considerando todos los posibles casos particulares para la solución de un problema. En la *Perspectiva Sub-simbólica* se diseñan sistemas con la capacidad de realizar modificaciones a su estructura, que emule el proceso de razonamiento de organismos inteligentes para resolver un problema. Dicha perspectiva también se identifica como el estudio de los mecanismos por el cual el razonamiento humano genera soluciones; esto implica a su vez que un conocimiento más profundo de dichos procesos conlleva a unos esquemas de aprendizaje automatizado más complejos y potentes [5].

Tales sistemas se conocen como *Redes Neuronales*. Su arquitectura tomó como referencia el funcionamiento de los sistemas nerviosos de los organismos que captan estímulos del medio en el que se encuentran, procesan la información obtenida y generan una respuesta [5]. En el presente trabajo, se hace una revisión histórica de los inicios de las redes neuronales, destacando sus principales rasgos, además de mencionar aplicaciones para las cuales han tenido un impacto significativo en su ulterior desarrollo.

2. JUSTIFICACIÓN

De acuerdo a las líneas de investigación propuestas por la Universidad Nacional Autónoma de Honduras, este trabajo se encuentra en el *Eje de investigación : Población y Condiciones de Vida*, apartado ciencia. Conforme a las líneas de investigación del programa de Maestría en Matemática con Orientación en Estadística, se encasilla en el tópico : *Series de Tiempo*.

La importancia del estudio en redes neuronales está en que permite reconocer patrones y clasificar una inmensa cantidad de datos, difíciles de abordar mediante algoritmos clásicos, que a su vez, abre paso para que se realicen inferencias más precisas de las características de los objetos de estudio. Datos relevantes como las medidas de humedad, radiación solar [6] o temperatura promedio global por año [10], son útiles para verificar cambios considerables en la dinámica del comportamiento climático del planeta tierra, sin embargo, su comportamiento a futuro requiere de modelos sofisticados, que solo pueden construirse mediante redes neuronales.

También identificar patrones en los tejidos del cuerpo humano para encontrar potenciales amenazas a la salud, como tumores, además de inferir en que estado de desarrollo se encuentran ameritan estos tipos de análisis [8].

Las redes neuronales abren nuevos campos de investigación en el área computacional y de modelamiento, dado que permite generar modelos a partir de muestras de datos considerables, además de saber organizarlos y extraer información relevante de ellos, que puede ser útil en la toma de decisiones.

3. ANTECEDENTES

El primer modelo computacional de una neurona artificial fue propuesto por McCulloch y Pitts [1], basado en el conocimiento que se tenía hasta ese momento de la estructura fisiológica de una neurona y su funcionamiento a principios de la década de 1940.

Su trabajo se resume en cinco conclusiones :

- (1) El proceso de activación de una neurona es *todo o nada*, esto significa que un número mínimo de conexiones deben estar presentes para que la neurona pueda excitarse y permitir el paso de una señal generada por catalizadores químicos. Además que dicho número es independiente de la actividad previa a la activación de la neurona y su posición.
- (2) El único retraso significativo en el funcionamiento del sistema nervioso es el sináptico.
- (3) La actividad de cualquier inhibidor sináptico, evita la excitación de la neurona.
- (4) La estructura de la red neuronal no cambia conforme el tiempo transcurre.
- (5) El período de *adición latente* (el tiempo que le toma a la neurona detectar los estímulos externos) es menor a los 0.25 milisegundos. El *retraso sináptico* (el tiempo necesario para que una neurona emita una respuesta ante los estímulos externos) es del orden de los 0.5 milisegundos.

A pesar del importante avance que supuso el trabajo de McCulloch y Pitts en la comprensión del funcionamiento neuronal, no todos sus conceptos son utilizados en la implementación computacional de las redes neuronales actualmente. El modelo de McCulloch y Pitts le confiere una característica binaria a la neurona, definiendo sus posibles salidas en 0 y 1 (en referencia a su rasgo todo o nada). Este tipo de neuronas son utilizadas en *Máquinas de Boltzmann*, sin embargo en la actualidad se sabe que la respuesta o valor que genera una neurona artificial se encuentra en un espacio continuo.

Esto se debe a que las entradas de una neurona artificial también son continuas, además que el número de excitaciones externas es arbitrario. Otro de los motivos por los cuales el modelo de McCulloch y Pitts no sigue vigente en la actualidad, es que en la práctica no existe un tiempo de retardo, las señales fluyen sin obstáculo alguno, como también la configuración de la red neuronal está dividida por capas, lo que implica que un grupo de neuronas se activen al mismo tiempo, en comparación a las máquinas de Boltzmann cuyas neuronas pueden hacerlo asincrónicamente. En las redes neuronales actuales los inhibidores no desactivan su funcionamiento, en cambio, mediante una suma ponderada, el efecto que tiene es negativo en comparación a los valores de entrada provenientes de las distintas conexiones que posee una neurona en particular.

En general el proceso de entrenamiento no exige que las conexiones entre las neuronas cambien, de acuerdo a la quinta condición del modelo de McCulloch y Pitts, sin embargo, cada conexión posee un peso que se modifica conforme la red neuronal analiza distintos casos.

La ventaja del modelo de McCulloch y Pitts que supuso en un principio está en que las neuronas puede representar cualquier expresión lógica finita; segundo, es útil para implementar una arquitectura paralela de procesamiento, y tercero, sentaron las bases para el desarrollo posterior de paradigmas de aprendizaje[1].

En 1949, Donald, O. Hebb fue el primero en establecer el concepto de *actualización sináptica de los pesos* y de utilizar el término *conexionismo*, él establece según [2]: Cuando un axón de la neurona A, está lo suficientemente cerca para excitar a la neurona B de manera persistente y repetitiva, entonces surge un cambio metabólico en ambas neuronas que permite aumentar la eficiencia en que la neurona A excita a la neurona B.

Hebb realiza cuatro contribuciones fundamentales, a saber [2]:

- (1) Establece que la información es almacenada mediante las distintas conexiones que se dan en la red neuronal o en los pesos de los espacios sinápticos.
- (2) El peso de conexión entre neuronas para la tasa de aprendizaje, es proporcional al producto de los valores de activación de las neuronas artificiales, este postulado asume que los valores de activación son positivos y que pueden incrementarse indefinidamente.
- (3) Se asume que el valor de los pesos sinápticos es simétrico, es decir, que el efecto de la neurona A para excitar la neurona B, es el mismo que el de la neurona B sobre la neurona A.
- (4) Se postula la *teoría del ensamble*, la cual dice que el aprendizaje ocurre cuando los pesos y patrones de conexión cambian y los ensamblajes son afectados considerablemente. Además, si un grupo de neuronas conectadas entre sí, con vínculos débiles, tiende a activarse en conjunto, los pesos de las conexiones que mantienen entre ellas se modifican, para que el ensamble en general se fortalezca.

Las contribuciones de Hebb son consideradas en la actualidad para la implementación computacional de las redes neuronales.

En 1958, en su artículo [9] Frank Rosenblatt presenta un nuevo modelo para la estructura de una red neuronal artificial conocida como *perceptrón*. Los perceptrones reciben ese nombre por el interés de Rosenblatt de aplicar modelos de redes neuronales artificiales a problemas relativos a la percepción, como el reconocimiento de voz o el reconocimiento óptico de caracteres. Fue la primera red neuronal artificial que se implementó computacionalmente, que fue simulada en una computadora IBM 704 en el laboratorio de Aeronáutica en Cornell. El perceptrón es capaz de identificar patrones en un conjunto de objetos modificando sus conexiones, fue descrita como una *máquina de aprendizaje*. En su modelo, Rosenblatt define dos conjuntos de neuronas, las *unidades singulares de asociación (AU)* que están conectadas a las *unidades de respuesta (RU)*. Ideó dos maneras en que las AU pudieran conectarse con las RU. En la primera, la activación de las RU depende de que las AU se exciten primero. En la segunda manera, los inhibidores de activación existen entre las AU y RU, permitiendo controlar los tiempos de excitación de las AU. Esta segunda opción fue más usada para la implementación computacional [1,3].

El ajuste de los parámetros en este modelo se realiza en una única capa, en las

sinápsis de las RU, ya que las AU se limitan a proporcionar la entrada de la información a la red. De ahí que normalmente, se interprete el perceptrón como una red neuronal de una sola capa [4].

Se pretendía ilustrar con el perceptrón algunas propiedades de los sistemas inteligentes. A diferencia del enfoque simbólico de McCulloch y Pitts, basado en lógica propocional, Rosenblatt optó por un enfoque probabilístico para analizar el modelo idealizado de una red biológica neuronal. El modelo de perceptrón permitió establecer dos interrogantes:

- (1) La forma en que la información es almacenada.
- (2) El uso de la información para reconocer patrones.

La respuesta fue que la información está contenida en conexiones ubicadas en lugares específicos de la red neuronal. La nueva información toma forma de nuevas conexiones o canales de transmisión (esto en el sistema nervioso) que se ve reflejado en el cambio de los valores de los pesos sinápticos en la red neuronal artificial. De esto se deduce que nuevos estímulos harán uso de estas vías sin la necesidad de que la información sea previamente identificada [3].

El principal mecanismo de aprendizaje del perceptrón es *auto-organizado*, esto consiste en que la respuesta de salida, en un principio tiene una estructura aleatoria [1]. En su configuración inicial, el perceptrón era incapaz de distinguir unos patrones de otros. Sin embargo, mediante un proceso de aprendizaje conseguía diferenciar patrones de interés. Este proceso de aprendizaje se realizaba por esfuerzo: la actividad neuronal que contribuye a una respuesta correcta de la red neuronal se aumenta, mientras que la que no contribuye se disminuye [4].

Rosenblatt consiguió demostrar matemáticamente el teorema de convergencia del perceptrón. Si un perceptrón es capaz de discriminar correctamente entre dos clases, el algoritmo de aprendizaje garantiza que el perceptrón aprenderá a hacerlo en un número finito de pasos de entrenamiento [4].

Una vez dotadas de un algoritmo de aprendizaje, las redes son capaces de ajustar dinámicamente las conexiones sinápticas entre neuronas. Usualmente este cambio se realiza de forma progresiva en función de la realimentación obtenida del éxito o el fracaso de la red durante su funcionamiento. Una red neuronal se ajusta automáticamente para producir una salida deseada, sin que su diseñador tenga que conocer cómo se puede resolver un problema, ni la secuencia de pasos necesaria para llegar a una solución desde un punto de vista algorítmico. De hecho esa es la característica clave para los sistemas basados en redes neuronales artificiales: su capacidad para resolver problemas sin necesidad de conocer cómo se resuelven [4].

En cierta medida, las redes neuronales alcanzan la solución de un problema mediante un mecanismo de prueba y error. Este mecanismo les permite generar soluciones gradualmente mejores, reteniendo la arquitectura de las conexiones entre neuronas que conducen a un rendimiento mejor. La arquitectura final de la red neuronal artificial es el resultado de su propia experiencia, de forma análoga a como se cree que se forman y evolucionan los circuitos neuronales del cerebro [4,8].

Los perceptrones simples permiten reconocer patrones. Desde el punto de vista de aprendizaje automático, son clasificadores binarios: funciones que permiten decidir si una entrada dada pertenece a una clase específica o no. Para ser más precisos, los perceptrones son un tipo de clasificador lineal: clasificadores cuya frontera de decisión, la que separa una clase de los de otra, viene dada en forma de línea recta en dos dimensiones, de plano en tres dimensiones o de hiperplano en general. Matemáticamente, la frontera de decisión definida por un clasificador lineal es de la forma:

$$(3.1) \quad b + \sum_i x_i \omega_i = 0.$$

Si se considera el sesgo b como un peso más, asociado a una entrada fija:

$$(3.2) \quad \sum_i x_i \omega_i = 0.$$

Por tanto, un perceptrón sólo será capaz de clasificar correctamente clases que son linealmente separables.

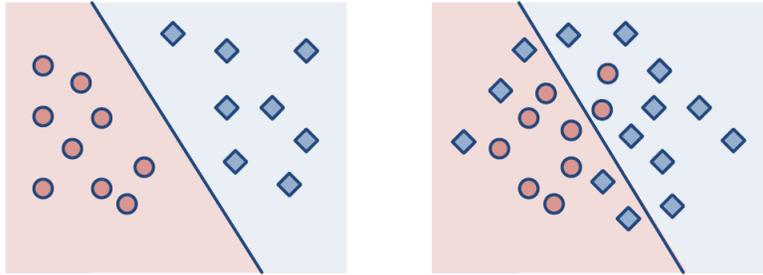


FIGURE 1. Separación gradual de los casos por el perceptrón [4].

Un perceptrón para poder clasificar convierte los datos de entrada en un vector de características x . A continuación, se utilizan los pesos ω asociados a cada una de estas características para obtener un valor escalar a partir del vector de entrada.

Dicho valor viene dado por:

$$(3.3) \quad z = \sum_i x_i \omega_i.$$

Finalmente este valor z se emplea para clasificar el ejemplo x . Si z se halla por encima del umbral de activación de la neurona, ésta se activa y se decide que el vector de entrada corresponde a un ejemplo de la clase objetivo, habitualmente llamada clase positiva. Si no es así, el vector de entrada no provoca la activación de la neurona y el ejemplo se asocia con la clase negativa.

Mediante el procesamiento de todos los posibles casos o ejemplos que se le presenten a la red neuronal, el perceptrón actualiza los pesos de las conexiones [4,5].

Para cada ejemplo del conjunto de entrenamiento [5]:

- (1) Si la salida del perceptrón es correcta, se dejan los pesos tal cual.

- (2) Si la unidad de salida incorrectamente da un cero (o falso negativo) se añade el vector de entrada al vector de pesos.
- (3) Si la unidad de salida incorrectamente da un uno (un falso positivo) se resta el vector de entrada del vector de pesos.

Habitualmente se recorre varias veces el conjunto de entrenamiento hasta que el algoritmo converja y la tasa de error sea menor a una cota previamente establecida. La versión más sencilla del aprendizaje consistiría simplemente en recorrer el conjunto de datos de entrenamiento un número fijo de veces [4].

El modelo propuesto por Rosenblatt utilizaba entradas y salidas binarias, sin umbral. No obstante es habitual utilizar una codificación bipolar de entradas, $-1,+1$, ya que esta codificación tiene características que resultan útiles en la práctica. Por un lado permite diferenciar situaciones en las que se tiene que codificar utilizando un tercer valor $x = 0$. Por otro lado, la codificación bipolar usando -1 facilita el ajuste de los pesos durante el entrenamiento del perceptrón. Si se usa la codificación binaria, solo será necesario modificar los pesos asociados a las entradas distintas de cero [5].

Los pesos del algoritmo se inicializan en cero. Para cada valor de x se calcula la salida real del perceptrón y y se compara con una salida de referencia o deseada d . Cuando ambas salidas difieren, se actualizan los pesos usando la ecuación:

$$(3.4) \quad \omega(t + 1) = \omega(t) + (d - y)x.$$

Es importante señalar que los pesos se actualizan tras mostrarle al perceptrón cada ejemplo de entrenamiento.

Mediante el *Error Cuadrático Medio* (MSE) se mide el ajuste del modelo a los datos de referenica o verificación. El MSE se define como sigue [4]:

$$(3.5) \quad MSE = \frac{1}{n} \sum_{j=1}^n (d_j - y_j)^2 = \frac{1}{n} \| \mathbf{d} - \mathbf{y} \|_2^2$$

El algoritmo de aprendizaje del perceptrón converge si las clases son linealmente separables, lo que a su vez hará que la tasa de error MSE converja a cero [5].

Las redes neuronales pueden aplicarse a la predicción de datos en series de tiempo.

Los conceptos que a continuación se presentan, están basados en [5].

4. SERIES TEMPORALES

Una serie temporal se puede definir como una colección de datos o valores de un suceso determinado a lo largo del tiempo, $x(t)_{t \in R^+}$, donde t representa la variable tiempo y $x(t)$ el valor de la serie en dicho instante de tiempo. Las series temporales se caracterizan porque su evolución temporal no depende explícitamente de la variable tiempo, sino de los valores de la serie en instantes anteriores de tiempo

o incluso de otras variables temporales que pudieran afectar a la evolución de la serie.

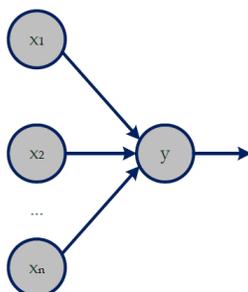


FIGURE 2. Arquitectura sencilla de un perceptrón [4].

La predicción del comportamiento en el futuro de una serie temporal es un objetivo importante en muchas áreas de aplicación, como puede ser en la física, la biología, negocios, etc. Así, por ejemplo, la predicción de la serie que describe la evolución del número de manchas solares observadas anualmente, permite estudiar la actividad solar, que a su vez influye en las transmisiones radioeléctricas o en fenómenos como el descenso térmico global de la Tierra.

En ocasiones, es posible construir modelos mediante la aplicación de ciertas leyes físicas, los cuales permiten describir el comportamiento o evolución temporal del fenómeno estudiado. Dichos modelos consisten generalmente de un conjunto de ecuaciones diferenciales o ecuaciones en diferencias que expresan el valor de la serie temporal en un instante de tiempo como una función de sus valores anteriores. Estos modelos pueden utilizarse para predecir el comportamiento de la serie en el futuro. Sin embargo, para la mayor parte de las series temporales que describen fenómenos reales, no es fácil ni inmediato conocer las ecuaciones que describen la relación explícita entre el valor de la serie en un instante de tiempo y sus valores anteriores, sino que sólo se dispone de un conjunto de datos observados. Estos datos requieren una interpretación para poder construir un modelo que aproxime la evolución de la serie temporal y permita predecir su comportamiento en el futuro.

Las redes neuronales artificiales han sido ampliamente utilizadas en el marco de predicción de series temporales. Esto es debido a características propias de las redes de neuronas, como son:

- (1) La capacidad de las redes de neuronas para aproximar y capturar relaciones a partir de un conjunto de ejemplos.
- (2) La capacidad de las redes de neuronas para construir relaciones no lineales.
- (3) La capacidad de las redes de neuronas para construir relaciones a partir de información incompleta o información con ruido.
- (4) Los modelos basados en redes de neuronas son fáciles de construir y utilizar.

De manera genérica, el problema de predicción de series temporales se puede formular del siguiente modo: a partir de un conjunto de muestras de la serie, conocer los valores en el futuro, es decir, su evolución o comportamiento a lo largo del tiempo.

En este caso se suelen distinguir dos casos:

- (1) *Predicción en un paso de tiempo* : la predicción en un paso de tiempo consiste en predecir el valor de la serie en el instante de tiempo inmediatamente siguiente al instante actual t , a partir de las muestras disponibles hasta dicho instante de tiempo. Es decir, predecir el valor $x(t + 1)$ utilizando un cierto número de muestras anteriores $x(t)$, $x(t - 1)$, $x(t - 2)$,..., número que depende de la serie temporal.
- (2) *Predicción en múltiples pasos de tiempo* : consiste en predecir el comportamiento de la serie, no únicamente en el instante inmediatamente siguiente al instante actual t , sino en un futuro más lejano, concretamente en el llamado intervalo de predicción $[t + 1, t + h + 1]$, siendo h un número natural que representa el horizonte de predicción. Es decir, consiste en predecir los valores de la serie $x(t + 1)$, $x(t + 2)$,... y $x(t + h + 1)$ a partir de la información disponible en el instante de tiempo actual t .

Cuando no se dispone de las ecuaciones que describen el comportamiento de la serie temporal, es necesario recurrir a métodos de aproximación para construir modelos que permitan resolver el problema de predicción, bien se trate de predicción en un paso de tiempo o en múltiples pasos de tiempo. La dificultad de predicción de las series temporales depende del comportamiento dinámico de dicha serie. Existen series que poseen comportamientos periódicos y estables que pueden ser explicados y predichos utilizando modelos basados en técnicas clásicas y lineales. Sin embargo, existen otras series temporales más complejas, para las cuales las técnicas clásicas podrían resultar deficientes.

Este trabajo se centra en series temporales cuyo comportamiento puede ser descrito por modelos no lineales regresivos. Estos modelos se caracterizan porque recogen el comportamiento temporal de la serie expresando el valor de la serie en el instante de tiempo $t + 1$ como una función no lineal de $r + 1$ valores de la serie temporal en instantes anteriores de tiempo, es decir:

$$(4.1) \quad x(t + 1) = F(x(t), x(t - 1), \dots, x(t - r)) + o(t),$$

donde t es la variable discreta tiempo; $o(t)$ es un error residual que se asume ruido blanco gaussiano; F es una función no lineal desconocida y que, por tanto, debe ser estimada o aproximada a partir de un conjunto de datos observados de la serie temporal, $x(t)_{t=0, \dots, N}$. Por esta razón, la construcción de estos modelos involucra la determinación de la función F , a partir del conjunto de muestras disponibles, mediante técnicas de aproximación, entre las cuales se encuentran las redes de neuronas artificiales.

El perceptrón es un modelo de red neuronal considerado estático, ya que su arquitectura está definida para encontrar relaciones entre patrones de entrada y salida independientes de la variable tiempo. Sin embargo, esto no impide que puedan utilizarse para tratar información temporal y abordar así el problema de predicción de series temporales. Para ello, basta que la entrada de la red esté compuesta no sólo del valor del patrón en un determinado instante de tiempo, sino también de

una cierta historia de dicho patrón, la cual se representa utilizando una secuencia finita temporal en el pasado.

El problema de predicción de series temporales puede abordarse también con estructuras dinámicas de redes de neuronas recurrentes. Estas se caracterizan porque la propia arquitectura de la red está diseñada para el procesamiento de información temporal, permitiendo que la salida de la red no sólo dependa de la entrada, sino de estados anteriores de la propia red. Esto hace que las redes recurrentes estén preparadas para capturar comportamientos dinámicos, sin necesidad de presentar a la red una secuencia de valores anteriores de la serie.

Considerando el vector $(x(t), x(t-1), \dots, x(t-r))$ como patrón de entrada, las redes neuronales estáticas pueden utilizarse para aproximar la función F en la ecuación (2.6), obteniendo el siguiente modelo de predicción :

$$(4.2) \quad \tilde{x}(t+1) = \tilde{F}(x(t), x(t-1), \dots, x(t-r)),$$

donde \tilde{F} representa la aproximación neuronal de la función F y $\tilde{x}(t+1)$ es la predicción proporcionada por el modelo neuronal en el instante de tiempo $t+1$. Se considera que la red neuronal tiene $r+1$ neuronas de entrada que reciben los valores de la serie temporal en los $r+1$ instantes anteriores, y una neurona de salida representa la predicción en un paso de tiempo de la serie temporal.

Los patrones para el entrenamiento de la red de neuronas se obtienen del conjunto de muestras disponibles de la serie temporal del siguiente modo: dado el conjunto $x(t)_{t=0, \dots, N}$ y dado el valor de r , cada patrón se construye utilizando los $r+1$ valores anteriores de la serie. De este modo, cada patrón de entrada recoge la información temporal necesaria para predecir el valor de la serie en el instante de tiempo $t+1$, para todo $t = r, \dots, N-1$.

Al realizar el entrenamiento, se pretende que la red capture la relación entre el valor de la serie en el instante de tiempo y los r valores pasados. El ajuste de los pesos se realiza para minimizar el error cuadrático medio en la salida, es decir:

$$(4.3) \quad e(t+1) = \frac{1}{2} (x(t+1) - \tilde{x}(t+1))^2 \quad \forall t = r, \dots, N-1.$$

Una vez realizado el entrenamiento de la red, el modelo puede utilizarse con el propósito de predecir en un paso de tiempo; hasta representar a la red los valores de la serie en los $r+1$ instantes anteriores de tiempo. Uno de los problemas del uso de estos modelos estocásticos para la predicción de series temporales es la definición de la longitud de la serie temporal que hay que presentar como patrón de entrada a la red, es decir, la definición del valor de r . Dicho valor dependerá de cada serie temporal y será necesario realizar una estimación. Esta estimación puede llevarse a cabo realizando un análisis previo de la serie temporal o de modelos existentes que expliquen parcialmente el comportamiento de la serie. En último caso, y ante la falta de información, mediante prueba y error, lo cual implica entrenar diferentes arquitecturas de redes de neuronas variando el número de entradas a la red, hasta

conseguir una arquitectura capaz de resolver el problema de predicción.

Sea $h > 1$ el horizonte de predicción. Como se ha comentado anteriormente, la predicción en múltiples pasos de tiempo consiste en predecir el valor de la serie en el instante de tiempo $t + h + 1$ o en el intervalo de tiempo $[t + 1, t + h + 1]$ a partir de los valores de la serie hasta el instante de tiempo actual t . Los esquemas para la predicción en series temporales en múltiples pasos de tiempo son los siguientes:

- (1) *Esquema de predicción 1* : consiste en utilizar de manera recurrente el modelo neuronal para la predicción en un paso de tiempo. Dicho modelo necesita como entrada los valores de la serie en los $r + 1$ instantes de tiempo inmediatamente anteriores. Por tanto, si el objetivo es predecir el valor de la serie en el instante de tiempo $t + h + 1$, dicha predicción viene dada por la siguiente expresión:

$$(4.4) \quad \tilde{x}(t + h + 1) = \tilde{F}(x(t + h), x(t + h - 1), \dots, x(t + h - r)).$$

Sin embargo, en el instante actual de tiempo t no toda la información de entrada a la red está disponible, pues los valores de la serie $x(t + h)$, $x(t + h - 1)$, ..., $x(t + 1)$ para $h > 1$, no se conocen. Para afrontar este problema se utilizan como entradas del modelo neuronal los valores predichos por la red neuronal en instantes anteriores de tiempo $\tilde{x}(t + h)$, $\tilde{x}(t + h - 1)$, ..., $\tilde{x}(t + 1)$. De este modo el modelo de predicción de un paso puede utilizarse con el propósito de predicción en múltiples pasos de tiempo retroalimentando la salida de la red hacia la entrada.

Durante el entrenamiento, la red ha capturado las relaciones entre la muestra actual y las muestras observadas en instantes anteriores de tiempo. Sin embargo, cuando el modelo se utiliza para predecir en múltiples pasos de tiempo, un grupo de neuronas de entrada reciben los valores de la serie temporal predichos por la red en instantes anteriores de tiempo, de manera que los errores cometidos en la predicción de dichos valores son propagados hacia las predicciones futuras, pudiendo influir negativamente en la calidad de dichas predicciones.

- (2) *Esquema de predicción 2* : Otra forma de abordar el problema de predicción en múltiples pasos de tiempo utilizando redes de neuronas estáticas consiste en construir un modelo regresivo para predecir, directamente, el valor de la serie en el instante $t + h + 1$ a partir de la información disponible en el instante actual t :

$$(4.5) \quad x(t + h + 1) = G(x(t), x(t - 1), \dots, x(t - d)),$$

donde G es una función no lineal que relaciona el valor de la serie en el instante de tiempo $t + h + 1$ con una secuencia finita de valores de la serie temporal disponibles en el instante de tiempo actual t . A diferencia del

modelo para predecir en un paso de tiempo, en este caso, el valor de la serie en el instante de tiempo $t + h + 1$ no se predice utilizando los $r + 1$ valores inmediatamente anteriores, sino información disponible en el instante actual.

Utilizando como entrada el vector $(x(t), x(t - 1), \dots, x(t - d))$ las redes de neuronas estáticas se pueden utilizar para aproximar la relación G , obteniendo el siguiente modelo de predicción:

$$(4.6) \quad \tilde{x}(t + h + 1) = \tilde{G}(x(t), x(t - 1), \dots, x(t - d)).$$

En este caso, la salida de la red neuronal representa la predicción de la serie en el instante $t + h + 1$.

Los patrones para el entrenamiento de la red se obtiene de las muestras disponibles de la serie temporal, $x(t)_{t=0, \dots, N}$ y considerando como salida deseada para el patrón en el instante de tiempo t , el valor de la serie en el tiempo $t + h + 1$.

La formulación de este modelo de predicción tiene sentido siempre que exista una relación, aunque desconocida a priori, entre el valor de la serie en el instante de tiempo $t + h + 1$ y la secuencia $(x(t), x(t - 1), \dots, x(t - d))$.

5. MAPEO LOGÍSTICO

En 1976 Robert May establecía el *Mapeo Logístico* [11], un modelo poblacional discreto que expresa la relación dinámica entra la población existente y la población máxima posible de una especie sujeta a determinadas condiciones. El modelo de mapeo logístico viene dado por:

$$(5.1) \quad x_{n+1} = rx_n(1 - x_n)$$

Donde x_n representa la fracción de población máxima en el tiempo con $x_n \in [0, 1]$ y r es la tasa de crecimiento poblacional cuyo valor se encuentra en el intervalo $[0, 4]$.

Naturalmente el Mapeo Logístico genera una serie de tiempo, lo que permite a su vez que pueda ser analizado mediante redes neuronales.

Mediante una simulación de la ecuación (5.1) con las condiciones iniciales $x_0 = 0.1$ y $r = 3.95$ se generan mil valores. De los cuales 900 serán utilizados para entrenar la red y los restantes 100 para verificar el aprendizaje de la misma. Se entrena una red neuronal recurrente con dos capas ocultas, una de tres neuronas y otra de una y se realizan 5,000 iteraciones. El error de la red neuronal en cada iteración se muestra en la figura 3. Lo que muestra un rápida convergencia del mismo a cero.

El comportamiento de los datos para la verificación del aprendizaje de la red neuronal se muestra en la figura 4. La predicción realizada por la red neuronal se muestra en la figura 5.

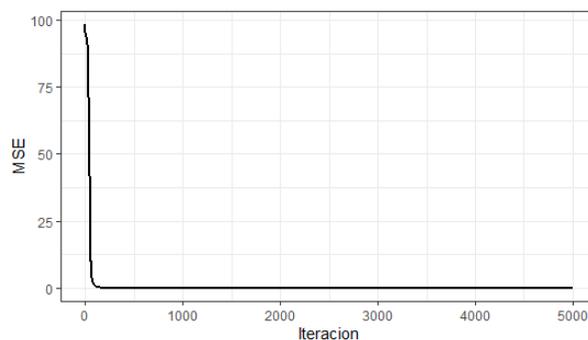


FIGURE 3. Error de Itación de la Red Neuronal

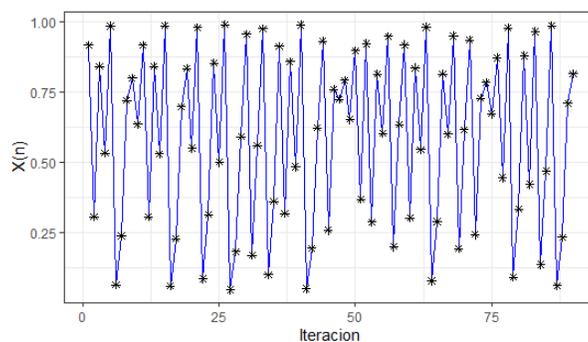


FIGURE 4. Comportamiento de los datos de Verificación.

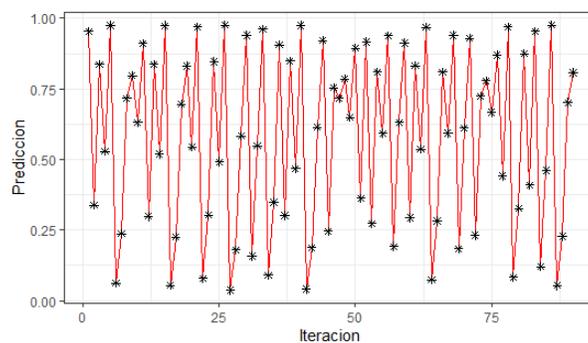


FIGURE 5. Predicción realizada por la red neuronal

Al realizar un gráfico de los datos de entrenamiento junto con los datos de predicción, se observa una clara tendencia lineal (Figura 6).

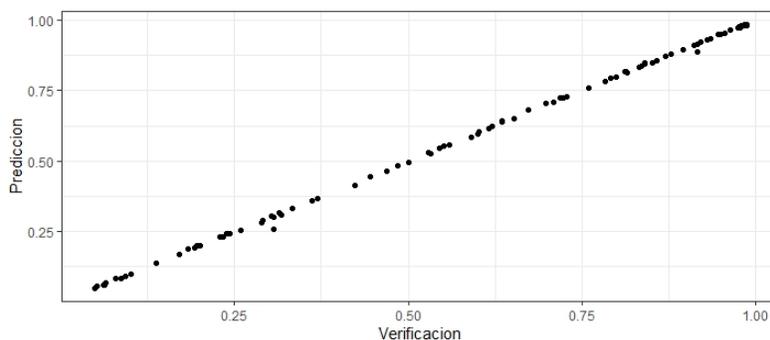


FIGURE 6. Datos de Predicción y Entrenamiento

La correlación entre ambos conjuntos de datos, resultó ser de 0.9954 y la pendiente de la regresión lineal fue de 0.997901 y su correspondiente intercepto de 0.002120. Estos valores demuestran que los datos de entrenamiento fueron predichos satisfactoriamente.

6. CONCLUSIONES

- (1) La estructura de una red neuronal artificial, se basa en la estructura morfológica de las células neuronales, la respuesta a los estímulos externos, se simulan mediante funciones de activación y el ajuste de pesos en las conexiones.
- (2) Son los datos los que determinan el proceso de aprendizaje de la red, un número considerable de ellos permite que los parámetros sean capaces de sintetizar un mayor número de posibilidades a las que una red neuronal puede llegar a estar expuesta, y que la respuesta que genere sea satisfactoria, esto se cuantifica mediante el valor del *Error Cuadrático Medio*.
- (3) Una de las aplicaciones usuales de las redes neuronales es la clasificación de objetos en categorías y también la predicción en series de tiempo. El horizonte de predicción depende fuertemente del comportamiento de la serie, pues estos datos suponen el conjunto de entrenamiento de la red y a su vez determinan la sensibilidad de la estructura de la red neuronal, lo que lleva a establecer *diferentes esquemas de predicción*.
- (4) En el mapeo logístico la gráfica para el Error Cuadrático Medio converge rápidamente a cero, cerca de la iteración 100, el valor del error decae drásticamente. Esto demuestra la rapidez de convergencia de la red para el ajuste de los datos de entrenamiento.
- (5) El factor de correlación lineal en el modelo de Mapeo Logístico fue de 0.997901 que junto con los parámetros de la regresión lineal, se permite decir que la relación entre los datos de verificación y los de predicción están próximos a la función identidad. Esto muestra que la red neuronal predijo los datos satisfactoriamente.

REFERENCES

- [1] R.C. EBERHART, R.W. DOBBINS (1990) *Early Neural Network Development History: The Age of Camelot*. IEEE Engineering in Medicine and Biology.
- [2] D.O. HEBB (1949) *The Organization of Behavior: A Neuropsychological Theory* New York: Wiley.
- [3] MARTA VIDAL GONZÁLEZ (2015) *El uso del Perceptrón Multicapa para la clasificación de patrones en conductas adictivas*. Disertación de grado. Universidad de las Islas Baleares.
- [4] FERNANDO BERZAL (2018) *Redes Neuronales y Deep Learning* Granada: <http://deep-learning.ikor.org>
- [5] PEDRO I. VIÑUELA, INÉS M. GALVÁN (2004) *Redes de Neuronas Artificiales: Un Enfoque Práctico* Universidad Carlos III de Madrid, Departamento de Informática. España: Prentice Hall.
- [6] LUIS MAZORRA AGUILAR (2015) *Modelo Predictivo de Radiación Solar mediante técnicas de Machine Learning. Aplicación a la Isla de Gran Canaria* Tesis Doctoral, Departamento de Ingeniería Eléctrica, Universidad de Las Palmas, Islas de Gran Canaria.
- [7] CHRISTOPHER M. BISHOP (2006) *Pattern Recognition and Machine Learning* Cambridge: Springer.
- [8] TON J. CLEOPHAS, AEIKO H. ZWINDERMAN (2015) *Machine Learning in Medicine- a Complete Overview* New York: Springer.
- [9] F. ROSENBLATT (1958) *The Perceptron : A Probabilistic Model for Information Storage and Organization in the Brain*. Cornell Aeronautical Laboratory, Psychological Review Vol. 6, No 6.
- [10] AGENCIA DE PROTECCIÓN AMBIENTAL DE LOS ESTADOS UNIDOS. *Métodos, Herramientas y Bases de Datos*. <https://espanol.epa.gov/espanol/metodos-modelos-herramientas-y-base-de-datos>.
- [11] MITCHELL J. FEIGENBAUM *Quantitative Universality for a Class of Nonlinear Transformation* Theoretical Division, Los Alamos Scientific Laboratory, Los Alamos, New Mexico. 1978 Plenum Publishing Corporation.

APLICACIÓN DE MODELOS VAR AL ANÁLISIS DE RIESGO FINANCIERO (HONDURAS)

ANGEL MATUTE

ABSTRACT. La dependencia de la estabilidad del sistema financiero del entorno económico ha sido un tema muy estudiado durante los últimos años, especialmente desde la crisis financiera en 2008 en Estados Unidos, lo cual repercutió a nivel internacional. Con el objetivo de identificar las variables económicas a considerar en Honduras para prevenir efectos negativos en el sistema financiero, en este artículo se analiza la relación entre las variables económicas, índice mensual de actividad económica (IMAE), índice de precios al consumidor (IPC), tasas de interés activa (variables consideradas de vital importancia para la estabilidad del sistema financiero) y el índice de Mora del sistema bancario (IM), esto mediante modelos de vectores autoregresivos (VAR), obteniendo mediante la estimación de dicho modelo, un ajuste considerablemente bueno, excepto por algunos fallos en los test a los que es sometido el modelo para verificar el ajuste. Este modelo mediante los coeficientes encontrados muestra relaciones esperadas (según la teoría revisada) entre estas variables, especialmente del IMAE y el IPC con el índice de mora.

1. INTRODUCCIÓN

Las mediciones de la exposición al riesgo de un grupo de instituciones financieras a un escenario macroeconómico o microfinanciero particular (denominadas pruebas de estrés o de resistencia), se dan con frecuencia por las instituciones financieras para evaluar sus exposiciones al crédito y otros riesgos. Las pruebas de resistencia también pueden ayudar a los formuladores de políticas en instituciones a evaluar las posibles implicaciones de los diferentes riesgos para la estabilidad del sistema financiero en su conjunto [1]. De acuerdo a esta investigación realizada en Inglaterra, la aplicación de estas pruebas tienen un gran valor para la estabilidad financiera.

La crisis financiera mundial sucedida en el año 2009 originada principalmente en el sistema financiero estadounidense nos recuerda lo difícil que es analizar de forma adecuada la vulnerabilidad de las entidades de crédito [2]. Por ello en un sistema financiero siempre es necesario conocer la calidad de la cartera crediticia, generalmente representada por el índice de mora, esto permite en gran manera conocer el desempeño del sistema financiero en una economía, pues impactos fuertes en este indicador generan efectos adversos para los entes financieros principalmente en la rentabilidad, lo cual puede generar otros efectos negativos que los puede llevar incluso a situación de quiebra y de este modo se puede provocar una crisis financiera

Palabras y frases clave. Índice de Mora, Índice Mensual de Actividad Económica (IMAE), Tasas de Interés.

nacional.

Por ello es deseable utilizar un modelo econométrico que mida la relación de causalidad entre las variables macroeconómicas e índice de mora [1], pues desde el punto de vista económico los indicadores financieros están muy relacionados con la actividad económica, en esta investigación se pretende pronosticar el indicador índice de mora dadas las variables macroeconómicas en Honduras:

Mediante un modelo de vectores autorregresivos (VAR) que según Gujarati [3] permite hacer pronósticos basados en rezagos de las variables tanto macroeconómicas como de el mismo índice de mora, donde el modelo se elabora mediante una serie de ecuaciones simultáneas y el conjunto de variables (macroeconómicas e índice de mora) debe tratarse en igualdad de condiciones, es decir no debe haber ninguna distinción a priori entre variables endógenas y exógenas, y es en este contexto que se desarrolla el modelo VAR, tomando como base la investigación [4].

La elaboración de estos modelos con cifras mensuales en el periodo de junio 2011 a enero 2021 (116 meses en total) permite tener una perspectiva más clara, basada en los resultados de los dos modelos que predicen como se comportará el índice de mora dada la situación macroeconómica actual y conociendo sus perspectivas a futuro en el país.

Se tiene como objetivo principal, pronosticar a futuro la evolución del índice de mora mediante la elaboración de un modelo econométrico (VAR) utilizando las técnicas estadísticas necesarias. También con el objetivo de observar cual modelo se ajusta de forma más adecuada a la realidad económica y detectar cuales son las variables macroeconómicas que influyen más sobre la calidad de la cartera de crédito del sistema financiero hondureño.

2. JUSTIFICACIÓN

La Universidad Nacional Autónoma de Honduras (UNAH) establece 4 ejes principales en los cuales se deben enfocar las investigaciones realizadas, cada eje tiene sus temas prioritarios a tomar en cuenta a la hora de realizar una investigación, Este trabajo que trata temas de econometría y series de tiempo aplicados al riesgo crediticio del sistema financiero, sigue la línea del eje de investigación 1 (Desarrollo Económico y Social) específicamente en el tema prioritario Globalización, productividad y competitividad.

La Maestría en Matemáticas con Orientación en Estadística Matemática tiene entre sus líneas de investigación las áreas siguientes: Series de tiempo y Econometría y actuaría en las cuales este trabajo esta vinculado dados los temas a tratar.

El sistema financiero de un país tiene un rol determinante en el crecimiento económico, ya que proporciona un canal eficiente en el que trazan las rutas de fondos de sectores institucionales con excedentes hacia sectores deficitarios [7].

La oportuna detección de la fragilidad financiera de las entidades financieras, que permita implementar las medidas correctivas necesarias para restaurar su solidez

y limitar las pérdidas potenciales, es de vital importancia [8] y dado que la mora puede constituir un indicador del efecto que tiene el ciclo económico sobre el valor de los activos de las entidades financieras (los créditos), pues este, al afectar conjuntamente a empresas y familias, incide necesariamente, a través del aumento de los volúmenes de créditos morosos dudosos y fallidos, en el índice de mora, lo que puede constituir un indicador de su riesgo de quiebra [6].

Además en [11] se demostró que el complejo análisis de series de tiempo univariantes sugerido por Box y Jenkins, en la actualidad es de gran utilidad en el campo estadístico para la determinación de proyecciones de una gran cantidad de variables de tipo económico social demográfico y de diferentes campos, los cuales son susceptibles a cambios en función del tiempo, y a su vez detectar los efectos de dichos cambios.

El Modelo VAR también es muy útil cuando existe evidencia de simultaneidad entre un grupo de variables y que sus relaciones se transmiten a lo largo de un determinado número de períodos. Al no imponer ninguna restricción sobre la versión estructural del modelo, no se incurre en los errores de especificación que dichas restricciones pudieran causar al ejercicio empírico. De hecho, la principal motivación detrás de los modelos VAR es la dificultad en identificar variables como exógenas, como es preciso hacer para identificar un modelo de ecuaciones simultáneas [10].

Por lo tanto, la información generada a partir de estos modelos econométricos desempeñaría un papel preponderante en la toma de decisiones en los mercados de crédito, proporcionando al sector financiero una herramienta estadística para poder predecir a futuro la evolución conjunta de las variables macroeconómicas con la calidad de la cartera crediticia.

3. ANTECEDENTES

3.1. Descripción de modelos.

3.1.1. *Modelos ARIMA*. Según [12], los modelos ARIMA son matemáticamente muy intuitivos y permiten estudiar el comportamiento de una gran variedad de series temporales y modelar algunas de las más importantes fuentes de error que en ellas se encuentran. Las correcciones de tendencias (invariabilidad en el tiempo), estimaciones de estacionalidad o la visualización de errores aleatorios hacen de estos modelos de los más versátiles y utilizados para realizar pronóstico y estimación futura.

Los modelos ARIMA desde los años setenta han ganado un papel fundamental en la literatura sobre predicción. Siendo introducidos originalmente por Box y Jenkins, esta generalización de los modelos autorregresivos y de medias móviles (ARMA) probó ser bastante efectiva a pesar de su sencillo cálculo mediante herramientas computacionales, y se instauró de ese modo como una alternativa relevante a los modelos estructurales, por entonces muy extendidos en la práctica de predicción. Al pasar de modelos Autoregresivos (AR) univariados a el caso multivariado se llega a lo que son los modelos de Vectores Autoregresivos (VAR), que serán la base

principal para el estudio en este artículo.

3.1.2. *Modelos VAR.* Según [13], dado los resultados de Engle y Granger en 1987, se popularizó la utilización de los modelos VAR desarrollados por Sims en 1980 en los cuales había enfatizado el uso de modelos VAR irrestrictos como un medio para modelar las relaciones económicas sin hacer supuestos innecesarios. Por otra parte, un modelo VAR con cointegración a menudo se basa en la idea de un “largo plazo”, o equilibrio móvil, definido por la teoría económica y caracterizado por un vector.

Para la elaboración de un modelo VAR, es necesario que las variables incluidas en el modelo sean estacionarias, es decir, todas las variables deben tener varianza y media constante en el tiempo [17].

Como se explica en [16], el modelo VAR puede ser escrito en su forma más general como.

$$Z_{t+1} = \Gamma + \sum_{j=1}^p \phi_j Z_{t+1-j} + \varepsilon_{t+1}$$

donde

Γ : es un vector constante,

ϕ_j : son matrices de coeficientes,

ε_{t+1} : es un vector de residuos,

Z_{t+1} : es el vector de n variables endógenas que incluye el modelo y

p : es el número de retardos de cada variable en cada ecuación.

Esta ecuación se puede reescribir como:

$$(3.1) \quad (I_n - \phi_1 L^1 - \phi_2 L^2 - \dots - \phi_p L^p) Z_{t+1} = \Phi(L) Z_{t+1} = \Gamma + \varepsilon_t$$

con $\Phi(L)$ matriz polinomial en el operador de retardos (L) formada por la matriz identidad I_n , y las matrices ϕ_j de $(n \times n)$.

Según [17], este proceso es estable si

$$(3.2) \quad \det(\Phi(L)) \neq 0 \text{ para } |L| \leq 1$$

La matriz identidad I_n evidencia que en este modelo VAR las variables se relacionan con los valores retardados de todas las variables en el vector, pero no existe relación entre variables contemporáneas; es decir, cada variable no se relaciona con los valores de las demás en el mismo periodo, a esta ecuación se le conoce como la forma reducida-autoregresiva, que se obtiene a partir de un modelo Estructural de Vectores Autoregresivos (*SVAR*) el cual se puede escribir como sigue:

$$(3.3) \quad (B_0 - B_1 L^1 - B_2 L^2 - \dots - B_p L^p) Z_{t+1} = B(L) Z_{t+1} = k + u_t$$

Donde k es un vector $(n \times 1)$, u_t un vector $(n \times 1)$ de perturbaciones, que en este modelo estructural se denominan *shocks estructurales* y $B(L)$ un polinomio de matrices en el operador de retardos formado por matrices $B_j (n \times n)$.

En este modelo la matriz B_0 refleja las relaciones contemporáneas entre las variables endógenas que constituyen Z_{t+1} , que en caso de ser B_0 igual a la matriz identidad (I_n) obtenemos la forma reducida-autoregresiva, lo cual se da cuando

suponemos que no existe relación contemporánea entre las variables en Z_{t+1} , pero de no ser así, al conocer la matriz B_0 si multiplicamos B_0^{-1} a ambos lados de la ecuación se obtiene la forma reducida-autoregresiva y viceversa.

3.2. Trabajos realizados. En Perú en 2003 pensando en factores macroeconómicos como desencadenantes de la quiebra de una entidad bancaria en [8] Moron y Loo-Kung elaboraron un sistema de alerta temprana para la detección temprana de fragilidad de las entidades bancarias para restaurar la solidez y minimizar las pérdidas potenciales. Se basaron en la utilización de un panel de datos logit de efectos fijos. Después de descartar algunas variables macroeconómicas bajo ciertos criterios estadísticos establecidos, se trabajó con las variables (1) indicador de vulnerabilidad externa, (2) la devaluación y (3) la inflación, además de otras variables microeconómicas.

En Inglaterra en el año 2005 se realizaron dos investigaciones que aportan información relevante para este trabajo, en [1] elaboraron un modelo de prueba de estrés considerando el impacto en el sistema financiero de un choque en la evolución futura del entorno macroeconómico, el cual es importante ya que impone un modelo explícito de correlaciones de los determinantes macroeconómicos del riesgo de crédito, como el desempleo y los precios de las propiedades, entre sí y a lo largo del tiempo. También se menciona que el entorno macroeconómico cambiante afectará los balances de los prestatarios y por lo tanto, la incidencia de incumplimiento.

Por otro lado, en [4] intentaron dar cuenta de la dinámica entre la tasa de amortización de préstamos de los bancos (Eliminación de préstamos por cobrar del inventario debido al incumplimiento del pago de estos por parte de los clientes) y las variables macroeconómicas clave, utilizando un modelo de autorregresión vectorial parsimoniosa (VAR). A pesar de que ningún modelo macroeconómico teórico convencional describe la relación de las variables macroeconómicas con las amortizaciones bancarias y viceversa, intentaron resolver este problema en dos pasos. Primero, el documento discute un modelo macroeconómico para ayudar a guiar la elección de un número parsimonioso de variables macroeconómicas para incluir en la especificación. En segundo lugar, aumentaron este vector de variables macroeconómicas con los datos de amortización y dejaron que se afecten entre sí de forma autorregresiva sin restringir la dinámica.

Este documento lleva a cabo uno de los primeros análisis multivariados de cómo el desarrollo macroeconómico afecta las amortizaciones de préstamos de los bancos del Reino Unido, tanto en conjunto como a nivel sectorial. Tomando como variables macroeconómicas la brecha del producto (diferencia entre lo que la economía está produciendo y lo que esta puede producir), la tasa de interés nominal a corto plazo, la inflación anual RPIX y la tasa de cambio real.

Concluyeron que el sistema bancario del Reino Unido en su conjunto parece ser robusto frente a grandes impactos macroeconómicos adversos.

En Colombia en 2012 por medio de la utilización de modelos VAR y el cálculo de las funciones impulso respuesta resultantes, se encuentra que si bien no parece

haber una relación directa significativa entre flujos de capital y estabilidad financiera, existe una relación indirecta entre estas dos variables, intermediada por el indicador cartera/PIB. Los incrementos en los flujos de capital tienden a propiciar el aumento del crédito de las entidades financieras. [18]

En Argentina en 2017, Betti [19] se enfoca en la elaboración de un modelo macroeconómico, confrontándose con los modelos estructurales y con los VAR. Se concluye la preferencia por los VAR dada su estructura más versátil. Permiten elegir el conjunto de variables endógenas y exógenas más pertinentes en función del objetivo buscado. Asimismo, son estructuras que no requieren una gran cantidad de herramientas computacionales sofisticadas, ni excesivos conocimientos de programación, tornándolos más accesibles para los profesionales que se desempeñan en las entidades financieras.

4. MODELO

Basados en la revisión teórica, se consideran para la elaboración de un modelo de Vectores autoregresivos (VAR) las variables:

- IM Índice de Mora
- IPC Índice de precios al consumidor
- IMAE Índice Mensual de Actividad Económica
- TI Tasa de Interés Activa

4.1. Requisito de estacionariedad. Es necesario primero evaluar la estacionariedad, es decir, comprobar si cada una de las variables tienen media y varianza constante en el tiempo, puesto que es requisito para elaborar el modelo tal como se menciono en la descripción del modelo. En primer lugar para facilitar el análisis se hace una transformación de las variables usando logaritmos, sus gráficas se muestran en la figura 1, donde:

- LTIM : Logaritmo del Índice de Mora
- LTIPC : Logaritmo del Índice de precios al consumidor
- LTIMAE : Logaritmo del Índice Mensual de Actividad Económica
- LTTI : Logaritmo de la Tasa de Interés Activa

Usando la librería VARS del Paquete estadístico R, se prueba en cada una de las variables, cuantas diferencias son necesarias para que estas sean estacionarias, obteniendo que estas cuatro variables necesitan solo una diferenciación para ser estacionarias.

Una vez extraídas las variables en diferencias podemos observar en la figura 2 que presentan un comportamiento estacionario, es decir que con una diferenciación estas se vuelven estacionarias, donde:

- DLTIM : Primera diferencia del logaritmo del Índice de Mora
- DLTIPC : Primera diferencia del logaritmo del Índice de precios al consumidor
- DLTIMAE : Primera diferencia del logaritmo del Índice Mensual de Actividad Económica
- DLTTI : Primera diferencia del logaritmo de la Tasa de Interés Activa

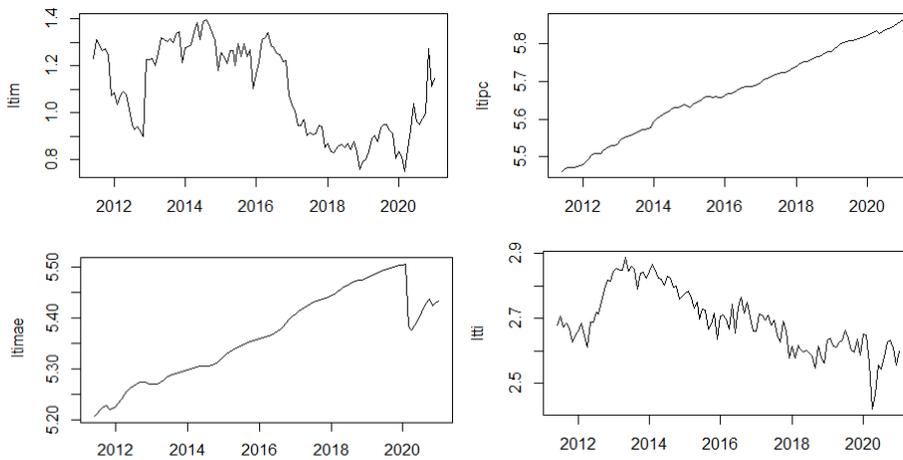


FIGURA 1. Logaritmo del Índice de Mora, IPC, IMAE y la Tasa de Interés Activa

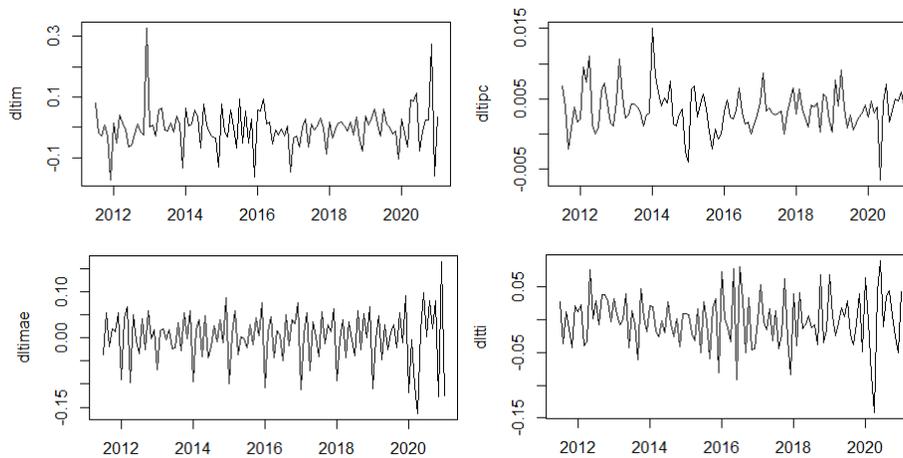


FIGURA 2. Primera diferencia del logaritmo del Índice de Mora, IPC, IMAE y la Tasa de Interés Activa

4.2. Estimación del Modelo VAR. Con variables estacionarias se puede elaborar el modelo VAR. Primero utilizando los criterios de información de Akaike (AIC), Hannan y Quinn (HQ), Bayesiana de Schwarz (SC) y el Error de Predicción Final (FPE) (estadísticos utilizados para elegir el rezago óptimo a estimar en un modelo) se considera que un modelo con dos rezagos es el más indicado para estas variables, como se puede observar en la tabla 1, ya que dos de los cuatro estadísticos indican que dos es el rezago óptimo.

Conociendo el número de rezagos óptimo, se procede a hacer la estimación, puesto que en la figura 2 se observa que los datos tienen media cero, y no se observa

CUADRO 1

| | <i>Rezago Óptimo</i> |
|------------|----------------------|
| <i>AIC</i> | 2 |
| <i>HQ</i> | 1 |
| <i>SC</i> | 1 |
| <i>FPE</i> | 2 |

tendencia, se hace la estimación sin intercepto ni tendencia, con lo cual se obtiene la ecuación 4.1 en su forma de matricial.

$$(4.1) \quad \begin{pmatrix} dltim \\ dltimeae \\ dltti \\ dltipc \end{pmatrix} = \begin{pmatrix} -\mathbf{0.17} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \mathbf{1.44} & -\mathbf{0.40} & 0 \\ 0 & 0 & 0 & \mathbf{0.61} \end{pmatrix} \begin{pmatrix} dltim(-1) \\ dltimeae(-1) \\ dltti(-1) \\ dltipc(-1) \end{pmatrix} \\ + \begin{pmatrix} 0 & -\mathbf{1.01} & 0 & \mathbf{3.65} \\ 0 & 0 & -\mathbf{0.06} & 0 \\ 0 & 0 & -\mathbf{0.28} & 0 \\ 0 & \mathbf{0.08} & 0 & 0 \end{pmatrix} \begin{pmatrix} dltim(-2) \\ dltimeae(-2) \\ dltti(-2) \\ dltipc(-2) \end{pmatrix}$$

Estos coeficientes son significativos al 10%. Existen tres test estadísticos mediante los cuales se deben evaluar los supuestos del modelo, estos tests se hacen sobre los residuos del modelo, evaluando:

- Autocorrelación: se desea que no haya autocorrelación entre los residuos,
- Normalidad: se desea que los residuos tengan una distribución normal,
- Heterocedasticidad: se desea que no exista heterocedasticidad, es decir, que la varianza de los residuos se mantenga constante.

4.3. Pruebas sobre el modelo. Las raíces del polinomio característico del modelo son 0.8134, 0.4353, 0.4353, 0.3743, 0.3743, 0.3173, 0.3173 y 0.3152, las cuales son todas menores que uno, y por tanto el modelo es estable.

4.3.1. Autocorrelación. Para evaluar la autocorrelación, se utiliza la prueba multivariante de Portmanteau, la cual es utilizada para probar la ausencia de perturbaciones correlacionadas en serie hasta el orden h en un VAR estable (p), la hipótesis nula en este test es que no existe autocorrelación. si se realiza el test hasta el rezago 12, con un valor p de 0.49, no se rechaza la hipótesis nula, por tanto podemos concluir que no existe evidencia de autocorrelación en los residuos, justo lo que se desea.

4.3.2. Normalidad. Las versiones multivariadas y univariadas de la prueba de Jarque-Bera se aplican a los residuos de un VAR. La versión multivariante de esta prueba se calcula utilizando los residuos que son estandarizado por una descomposición de Choleski de la matriz de varianza-covarianza para los residuos centrados. se debe tener en cuenta que, en este caso, el resultado de la prueba depende del orden de

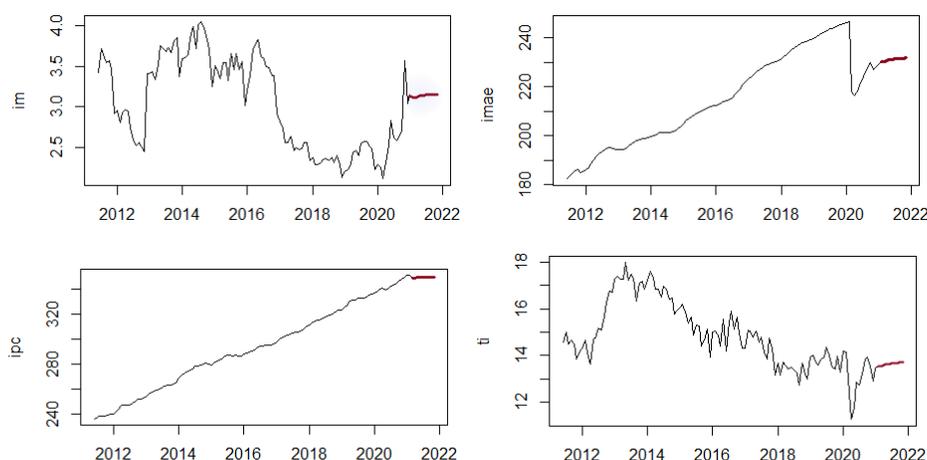


FIGURA 3. Variables junto con su pronóstico diez periodos adelante en rojo

las variables. siendo que en esta prueba la hipótesis nula es que los residuos se distribuyen de forma normal, al obtener para este test un valor $p \leq 0,05$, se rechaza la hipótesis nula, por lo cual se concluye que los residuos no se distribuyen de forma normal.

4.3.3. *Heterocedasticidad.* Este test prueba mediante el multiplicador de Lagrange (LM) para heterocedasticidad condicional autorregresiva (ARCH) (ARCH-LM), la existencia de heterocedasticidad hasta el rezago h . Al hacer el test hasta un rezago 11, este da un p valor de 0.95, y dado que la hipótesis nula en este test es que la varianza de los residuos es constante, entonces no se rechaza la hipótesis nula y se concluye que no existe heterocedasticidad en los residuos.

4.4. **Pronósticos.** Aun quedando pendiente la eliminación de algunos coeficientes, además de el problema de la no normalidad en los residuos, se elaboro el pronostico de las variables mediante el modelo VAR estimado, consiguiendo los resultados observados en la figura 3.

Los pronósticos revelan una evolución a futuro muy estable, lo cual no parece estar muy de acorde con la realidad, especialmente en las variables IMAE e IPC, que en condiciones normales tienen tendencia creciente, esto puede ser provocado por los problemas que aun quedan sin resolver en el modelo (coeficientes no significativos y la no normalidad en los residuos) que distorsionan el pronósticos.

Otro punto que podría estar influyendo en los pronósticos poco alejados de lo que se esperaría, es la actual crisis sanitaria provocada por el COVID-19 que podría haber generado distorsiones significativas en las variables y por tanto en sus pronósticos, lo que genera poca confianza en los pronósticos generados por este modelo.

5. CONCLUSIONES

- (1) A pesar de utilizar para el modelo un numero de rezagos indicado por los test HQ y SD solo se encontraron nueve coeficientes estadísticamente

significativos en el modelo, lo cual indica que es necesario eliminar gran cantidad de coeficientes.

- (2) Una causa de la poca cantidad de coeficientes significativos podría ser el problema de no normalidad en la distribución de los residuos del modelo, por lo tanto es otro asunto a corregir en el modelo.
- (3) Tanto el problema de los coeficientes no significativos, junto con la distribución no normal en los residuos, pueden ser la causa de los pronósticos en las variables tan poco realistas considerando la evolución histórica de las variables.
- (4) Para corregir los errores antes mencionados se pueden utilizar variables exógenas, ya sea para captar estacionalidad en las variables que al ser económicas tienen alta probabilidad de tenerla, o para mitigar el efecto de algunos valores extremos que podrían estar alterando la estimación.
- (5) Particularmente el índice de Mora, que es el punto central de esta investigación, solo es influenciado significativamente por tres variables:
 - El Índice de Mora rezagado un periodo de forma negativa,
 - El IMAE rezagado dos periodos de forma negativa y con mayor proporción que el Índice de Mora y,
 - El IPC de forma positiva y en mayor magnitud que cualquier otra variable.
- (6) Dadas las condiciones actuales, a pesar de los pronósticos diferentes de lo que se esperaría, no se puede descartar el modelo ni sus pronósticos, pues no estamos en un periodo de tiempo en el que la economía este funcionando en condiciones normales.

REFERENCIAS

1. P. Bunn, A. Cunningham y M. Drehmann, *Stress testing as a tool for assessing systemic risks*. Bank of England Financial Stability Review (2005), 116–26.
2. J. Gomez, I. Orozco y M. Drehmann, *Un Modelo de Alerta Temprana para el Sistema Financiero Colombiano*, Banco de la república, Colombia (2009).
3. G. Damodar, *Econometría*, McGraw-Hill/Irwin, Inc., (2010).
4. G. Hoggarth, S. Sorensen y L. Zicchino *Stress tests of UK banks using a VAR approach*. Bank of England (2005).
5. V. Romero, *Implementación de modelo para pronosticar variables en Cooperativas: caso de cobelén con cartera de créditos*. Universidad EAFIT (2015).
6. E. Guillen y L. Peñafiel, *Modelos predictor de la morosidad con variables macroeconómicas*, Revista Ciencia UNEMI., vol. 11, 2017.
7. D. Pineda, *Modelos predictor de la morosidad con variables macroeconómicas*, BCH (1993).
8. E. Moron y R. Loo'kung , *Sistema de alerta temprana de fragilidad Financiera*, Universidad del Pacífico, Peru (2003).
9. A. Andrade y C. Gonzalez, *Modelo Macro para Pruebas de Tensión de Riesgo de Crédito de Consumo en el Sistema Financiero Ecuatoriano*, Analitika, Revista de análisis estadístico., vol. 14(2), , 2017.
10. A. Novales, *Modelos vectoriales autoregresivos (VAR)*, Universidad Complutense., vol. 14(2), , 2016.
11. N. Quisbert, *Modelo ARIMA*. Revista Ciencia y Cultura (1997).
12. J. Gallego, A. Rodriguez, J. Minguez y F. Jimenes, *Modelos ARIMA para la predicción del gasto conjunto de oxígeno de vuelo y otros gases en el Ejército del Aire*, Sanidad mil., vol. 74(4), 1997.
13. C. Granger y R. Engle, *Econometría de las series de tiempo, cointegración y heteroscedasticidad condicional autoregresiva*, Cuestiones Económicas., vol. 20(2), 2004.
14. F. Camones, *CONTROL DE CALIDAD EN LOS PROCESOS ESTADISTICOS*, Centro de Investigación y Desarrollo del Instituto Nacional de Estadística e Informática (INEI) (2002).

15. P. Gonzalez, *Análisis de series temporales: Modelos ARIMA*, Universidad del País Vasco (UPV-EHU) (2009).
16. P. Alvarez, A. Crespo, F. Nuñez y C. Usabiaga, *Introducción de elementos autorregresivos en modelos de dinámica de sistemas*, Revista de dinámica de sistemas., vol. 2(1), 2006, pp. 37–66.
17. H. Lütkepohl y M. Krätzig, *Applied Time Series Econometrics*, Cambridge University Press (2004).
18. J. Gonzalez, *Flujos de capital, fragilidad financiera y desarrollo financiero en Colombia*, Ensayos sobre Política económica., vol. 30, 2006, pp. 69–109.
19. M. Betti, *Pruebas de estrés en entidades financieras. El modelo de vectores autorregresivos como metodología para la generación de escenarios macroeconómicos*, Revista De Investigación En Modelos Financieros., vol. I (2018-I), 2018.

ESCUELA DE MATEMÁTICAS, UNAH, TEGUCIGALPA, FRANCISCO MORAZAN
Dirección actual: Escuela de matemáticas, UNAH, Tegucigalpa, Francisco Morazan
Dirección de correo electrónico: angelavilez1@hotmail.com

MODELOS ESTADÍSTICOS DE ANÁLISIS DE FACTORES QUE AFECTAN EL RENDIMIENTO ESCOLAR EN MATEMÁTICA Y ESPAÑOL EN TERCER Y SEXTO GRADO EN HONDURAS

LESKY IBETH RIVAS

RESUMEN. En este artículo se hace una breve descripción de los modelos estadísticos utilizados en el análisis de factores que afectan el rendimiento escolar, se presenta un análisis de los datos utilizando el modelo de regresión múltiple, así como los resultados de las pruebas para verificar la existencia de multicolinealidad, heterocedasticidad y normalidad en el modelo.

ABSTRACT. In this article a description of the statistical models used in the analysis of factors that affect school performance is made, an analysis of the data using the multiple regression model is presented, as well as the results of the tests to verify the existence of multicollinearity, heteroscedasticity and normality in the model.

1. INTRODUCCIÓN

El rendimiento escolar se define como “los resultados obtenidos en pruebas o exámenes normalizados, que miden los conocimientos o competencias en una materia específica” [1]. Muchos países efectúan periódica y sistemáticamente mediciones destinadas a determinar lo que los estudiantes han aprendido durante su experiencia escolar, ya sea en exámenes por clase, evaluaciones nacionales o evaluaciones internacionales.

“El rendimiento escolar es medido, en su mayor parte por tests especialmente diseñados. Pero en algunos casos, el rendimiento escolar es medido a través de exámenes nacionales en lengua, matemática o ciencia” [2]. Estas evaluaciones sirven como base de información para la toma de decisiones en el sistema educativo.

Los modelos estadísticos utilizados en el estudio del rendimiento escolar, permiten conocer cuáles son las variables y en qué medida están presentes en el rendimiento escolar. Existen diferentes modelos utilizados, entre ellos se destacan:

1. Teoría de Respuesta al Ítem.

La Teoría de Respuesta al ítem analiza el comportamiento de una prueba a partir del funcionamiento de todos los ítems que la conforman [22].

2. Modelo de Regresión Múltiple.

La regresión lineal múltiple trata de ajustar modelos lineales o linealizables

Fecha: May, 2021.

Palabras y frases clave. Modelos de rendimiento escolar, regresión múltiple, multicolinealidad, heterocedasticidad, normalidad .

entre una variable dependiente y más de una variables independientes [6].

3. Modelo Multinivel.

Los modelos multinivel son, en esencia, ampliaciones de los modelos de regresión lineal clásicos, de tal forma que en realidad son varios modelos lineales para cada nivel. Así, los modelos del primer nivel están relacionados con uno de segundo nivel en el que los coeficientes de regresión del nivel 1 se regresan en un segundo nivel de variables explicativas y así sucesivamente para los diferentes niveles [7].

4. Modelos de Análisis de Estructuras de Covarianza o Ecuaciones Estructurales.

Los modelos de ecuaciones estructurales (MES) son una familia de modelos estadísticos multivariantes que permiten estimar el efecto y las relaciones entre múltiples variables. Los modelos de ecuaciones estructurales nacieron de la necesidad de dotar de mayor flexibilidad a los modelos de regresión [8].

2. JUSTIFICACIÓN

La Universidad Nacional Autónoma de Honduras (UNAH) presenta prioridades de investigación que tienen que ser atendidas institucionalmente en los próximos años mediante trabajos de investigación, actualmente existen trece temas prioritarios que han sido agrupados en cuatro ejes de investigación. Este trabajo sigue la línea de investigación del eje número tres de la UNAH población y condiciones de vida, este eje de investigación está conformado por cuatro temas prioritarios de investigación, enmarcándose en el tema número tres: cultura, ciencia y educación, y dentro de las líneas de investigación en la orientación en estadística de la maestría en matemática este trabajo sigue la número seis: Estadística multivariada y modelos lineales generalizados.

Las agencias internacionales realizan evaluaciones que monitorean las tendencias en el rendimiento estudiantil en matemáticas, ciencias y lectura.

Utilizan el currículo de matemática y español, definido en términos generales, como el concepto organizativo principal al considerar cómo se brindan las oportunidades educativas a los estudiantes, los factores que influyen, cómo debe organizarse el sistema educativo para facilitar este aprendizaje; lo que realmente se enseña en las aulas, las características de quienes lo enseñan y cómo se enseña; y, finalmente, qué es lo que los estudiantes han aprendido y qué piensan acerca de estos temas.

También recopilan información importante sobre los contextos de los estudiantes para aprender matemáticas y español. “Gracias a una reciente serie de estudios nacionales e internacionales, se han adquirido nuevos conocimientos importantes sobre lo que saben y pueden hacer los alumnos de las escuelas primarias de los países en desarrollo” [4].

Algunas de las agencias internacionales son: Programa para la Evaluación Internacional de Estudiantes (PISA), Evaluación Latinoamericana de la Calidad de la Educación (LLECE), Tendencias en el Estudio Internacional de Matemáticas y Ciencias (TIMSS), Progreso en el estudio internacional de alfabetización en lectura (PIRLS), Organización de las Naciones Unidas para la Educación, la Ciencia y la

Cultura (UNESCO). Y entre las nacionales se tienen: El Proyecto Mejorando el Impacto al Desempeño Estudiantil de Honduras, por sus siglas MIDEH, Unidad Externa de Medición de la Calidad de la Educación, por sus siglas UMCE.

El uso de modelos en la investigación sobre eficacia escolar constituye una interesante propuesta al movimiento y es realizada por Samuel Stringfield en 1994. Este expresa tres razones que considera de interés para elaborar modelos. En primer lugar, un modelo ofrece a los investigadores del área una fuente de información que permite considerar, relacionar entre sí, y tender puentes sobre la investigación previa. Posteriormente, permite abrir una serie de caminos, o alternativas, para desarrollar investigaciones futuras que, a posteriori, guiarán en la toma de las decisiones políticas relacionadas con la eficacia escolar. Por último, los modelos proporcionan una visión panorámica de las diferentes rutas de investigación posibles [3].

Es por esto que las agencias dedicadas a realizar estudios sobre rendimiento escolar, utilizan diferentes modelos en el análisis de los factores asociados al rendimiento escolar, aportando herramientas para la formulación de políticas públicas pertinentes en cada país.

3. ANTECEDENTES

La investigación que busca determinar los factores asociados con el rendimiento es tan antigua como la propia investigación empírica educativa; es decir, tiene más de un siglo de vida. Sin embargo, hay que fechar el nacimiento del movimiento de eficacia escolar en 1966, año en el que se publicó el “Informe Coleman”. A partir de los datos obtenidos, “se apuntó que la escuela desempeñaba un papel muy limitado en el rendimiento del alumno” [10]. Esta conclusión generó múltiples reacciones de todo tipo entre la comunidad educativa. Por una parte, algunos estudiosos que trabajaban sobre los factores explicativos del rendimiento asumieron como válidos los resultados del estudio de Coleman y, simplemente, obviaron el nivel escuela en sus indagaciones. “De esta manera, se potenció lo que se ha llamado la perspectiva individualista o psicológica, que busca identificar los factores personales, sociales y familiares asociados al rendimiento, pero no los factores escolares” [9].

Uno de los escépticos sobre estos resultados fue el inspector estadounidense George Weber. Durante sus años de trabajo como supervisor, encontró que había escuelas que, a pesar de encontrarse en ambientes desfavorecidos, conseguían enseñar a los niños a leer y a escribir, mientras que otras no lo hacían. Por lo tanto, “estudiar en una escuela u otra sí era importante para el futuro de los alumnos” [11], lo que contradecía el “Informe Coleman”. Así, elaboró un estudio cuya principal novedad fue “entrar en las aulas”. De esta forma, añadió los factores de proceso al modelo entrada-producto, con lo que obtuvo el modelo “input-process-product” (IPP). “En este trabajo encontró que, en efecto, había escuelas que marcaban una diferencia. Localizó algunos factores que caracterizaban a las buenas escuelas y ofreció así una primera lista de factores de eficacia escolar” [9].

Como ya se ha mencionado las agencias internacionales y nacionales dedicadas a realizar estudios sobre rendimiento escolar, utilizan diferentes modelos en el análisis de los factores asociados al rendimiento escolar. La agencia TIMSS informa el rendimiento de los estudiantes a través de puntajes de escala derivados de la escala de la teoría de respuesta al ítem, por sus siglas en inglés (IRT) [12]. Este modelo también es utilizado por la agencia PIRLS.

La UNESCO en los análisis del Primer Estudio Regional Comparativo y Explicativo (PERCE) y en el Segundo Estudio Regional Comparativo y Explicativo (SERCE) emplearon modelos jerárquicos lineales, los cuales corresponden a una particular técnica de regresión diseñada para utilizar y aprovechar la estructura jerárquica propia de los datos provenientes del ámbito educativo [13]. En el Tercer Estudio Regional Comparativo y Explicativo (TERCE) se utilizó modelos de teoría de respuesta al ítem, debido a las características y objetivos del TERCE.

La Unidad Externa de Medición de la Calidad Educativa (UMCE) en el análisis de factores asociados empleó modelos multinivel ya que se realiza un ajuste de datos para cada nivel de forma individual [14]. Y el Proyecto Mejorando el Impacto al Desempeño Estudiantil de Honduras (MIDEH) en el análisis de factores hace uso de la regresión lineal de mínimos cuadrados ordinarios, la cual permite establecer qué peso tiene cada factor sobre el desempeño [15].

4. MODELOS DE REGRESIÓN MÚLTIPLE

4.1. Características conceptuales. En el modelo de regresión lineal múltiple, el regresando (que puede ser la variable endógena o una transformación de las variables endógenas), es una función lineal de k regresores correspondientes a las variables explicativas (o a transformaciones de las mismas) y una perturbación aleatoria o error. El modelo también incluye un término independiente. Si designamos por y al regresando, por x_2, x_3, \dots, x_k a los regresores y por u al error o perturbación aleatoria, el modelo poblacional de regresión lineal múltiple vendrá dado por la siguiente expresión:

$$(4.1) \quad y = \beta_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_k x_k + u$$

Los parámetros $\beta_1, \beta_2, \beta_3, \dots, \beta_k$ son fijos y desconocidos.

En (4.1) se pueden distinguir dos componentes: un componente sistemático $\beta_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_k x_k$ y la perturbación aleatoria u . Llamando μ_y al componente sistemático, podemos escribir:

$$(4.2) \quad \mu_y = \beta_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_k x_k$$

Esta ecuación es conocida como función de regresión poblacional (FRP) o hiperplano poblacional. Cuando $k=2$, la FRP es específicamente una línea recta, cuando $k=3$, la FRP es específicamente un plano y, por último, cuando $k > 3$, la FRP es denominada genéricamente hiperplano, que no es susceptible de ser representado físicamente.

El parámetro μ_y es una función lineal en los parámetros $\beta_1 + \beta_2 + \beta_3 + \dots + \beta_k$. Ahora, supongamos que tenemos una muestra aleatoria de tamaño n , $(y_i, x_{2i}, x_{3i}, \dots, x_{ki})$ $i=1, 2, \dots, n$, extraída de la población estudiada. Si expresamos el modelo poblacional para todas las observaciones de la muestra, se obtiene el siguiente el sistema:

$$\begin{aligned} y_1 &= \beta_1 + \beta_2 x_{21} + \beta_3 x_{31} + \dots + \beta_k x_{k1} + u_1 \\ y_2 &= \beta_1 + \beta_2 x_{22} + \beta_3 x_{32} + \dots + \beta_k x_{k2} + u_2 \\ &\vdots \\ y_n &= \beta_1 + \beta_2 x_{2n} + \beta_3 x_{3n} + \dots + \beta_k x_{kn} + u_n \end{aligned}$$

El anterior sistema de ecuaciones puede expresarse de una forma más compacta usando la notación matricial. Así, vamos a denominar

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 & x_{21} & x_{31} & \dots & x_{k1} \\ 1 & x_{22} & x_{32} & \dots & x_{k2} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_{2n} & x_{3n} & \dots & x_{kn} \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_n \end{bmatrix} \quad u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

La matriz X es la matriz de regresores. Entre los regresores también se incluye el regresor correspondiente al término independiente. Este regresor, que a menudo se denomina regresor ficticio, toma el valor 1 para todas las observaciones. El modelo de regresión lineal múltiple, expresado en notación matricial es el siguiente:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{21} & x_{31} & \dots & x_{k1} \\ 1 & x_{22} & x_{32} & \dots & x_{k2} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_{2n} & x_{3n} & \dots & x_{kn} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_n \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

Si se tiene en cuenta las denominaciones dadas a vectores y matrices, el modelo de regresión lineal múltiple puede ser expresado de forma compacta de la siguiente manera:

$$(4.3) \quad y = X\beta + u$$

Donde, de acuerdo con la notación utilizada, y es un vector $n \times 1$, X es una matriz $n \times k$, β es un vector $k \times 1$ y u es un vector $n \times 1$ [16].

4.2. Multicolinealidad. El término multicolinealidad se atribuye a Ragnar Frisch. Originalmente, designaba una relación lineal “perfecta” o exacta entre algunas o todas las variables explicativas de un modelo de regresión. Para la regresión con k variables que incluye las variables explicativas X_1, X_2, \dots, X_k (donde $X_1 = 1$ para todas las observaciones de forma que den cabida al término del intercepto), se dice que existe una relación lineal exacta si se satisface la siguiente condición:

$$(4.4) \quad \lambda_1 X_1 + \lambda_2 X_2 + \dots + \lambda_k X_k = 0$$

Donde $\lambda_1, \lambda_2, \dots, \lambda_k$, son constantes tales que no todas son simultáneamente iguales a cero. Hoy en día, sin embargo, el término multicolinealidad incluye el caso de multicolinealidad perfecta, como lo indica (4.4) y también el caso en el cual hay X variables intercorrelacionadas pero no en forma perfecta, de la siguiente manera:

$$(4.5) \quad \lambda_1 X_1 + \lambda_2 X_2 + \dots + \lambda_k X_k + v_i = 0$$

Donde v_i es un término de error estocástico. [17]

La multicolinealidad significa que el comportamiento de buena parte de las variables guarda algún tipo de relación unas con otras y esa relación puede ser de menor o mayor grado. Solamente cuando dicha relación es de mayor grado podría ser un problema dentro de la modelación [18]. Se utilizan diferentes pruebas para la detección de la multicolinealidad entre ellas se tiene:

4.2.1. Matriz de correlación. La matriz de correlación es una matriz cuadrada $n \times n$ constituida por los coeficientes de correlación de cada pareja de variables; de manera que tendrá unos en su diagonal principal, y en los elementos no diagonales (i,j) los correspondientes coeficientes de correlación r_{ij} [18].

4.2.2. Prueba de Theil. Para la prueba de Theil se corren regresiones auxiliares de la variable dependiente contra los k regresores menos uno de ellos, al coeficiente de determinación de esas regresiones se le denomina R_i^2 . El efecto R^2 de Theil se obtiene con la siguiente expresión [18]:

$$(4.6) \quad R^2_{Theil} = R^2 - \left[\sum_{i=1}^n (R^2 - R_i^2) \right]$$

Donde R^2 es el coeficiente de determinación de la regresión original con todos los regresores y R_i^2 es el coeficiente de determinación de la regresión auxiliar i . Si el efecto de Theil fuera nulo no existiría multicolinealidad, entre mayor sea el efecto más grave es el problema [18].

4.2.3. Prueba de Klein. La regla de Klein es un principio práctico, propuesto por el premio Nobel Lawrence Klein. De acuerdo a dicho principio, la multicolinealidad es un problema a considerar si la R_i^2 de alguna regresión auxiliar es mayor que el coeficiente de determinación R^2 de la regresión original [18].

4.2.4. Factor de inflación de la varianza (VIF). El factor de inflación de la varianza (VIF) se utiliza como una medida del grado en que la varianza del estimador de mínimos cuadrados es incrementada por la colinealidad entre las variables, el VIF se define de la siguiente manera [18]:

$$(4.7) \quad VIF = \frac{1}{1 - R^2}$$

En donde R_i^2 es el coeficiente de determinación de la regresión auxiliar i . Por ejemplo, ante perfecta multicolinealidad $R_i^2=1$, lo cual hace que el VIF tienda a infinito, si la multicolinealidad es imperfecta y elevada, por ejemplo un $R_i^2=0.9$, el VIF será igual a 10. Es usual en la práctica si el VIF resulta mayor a 10 o incluso 5 sea considerado como evidencia de fuerte multicolinealidad [18].

4.3. Heterocedasticidad. Un supuesto importante del modelo clásico de regresión lineal es que la varianza de cada término de perturbación u_i , condicional a los valores seleccionados de las variables explicativas, es algún número constante igual a σ^2 . Éste es el supuesto de homocedasticidad, o igual (homo) dispersión (cedasticidad), es decir, igual varianza. Simbólicamente, [17]

$$(4.8) \quad E(u_i^2) = \sigma^2$$

Donde $i=1, 2, \dots, n$

Por tanto, si hay heterocedasticidad, simbólicamente,

$$(4.9) \quad E(u_i^2) = \sigma_i^2$$

Hay diversas razones por las cuales las varianzas de u_i pueden ser variables, una de ellas es por la presencia de datos atípicos o aberrantes. Otra fuente de heterocedasticidad surge de establecer que el modelo de regresión está correctamente especificado, con mucha frecuencia lo que parece heterocedasticidad puede deberse a que se omitan del modelo algunas variables importantes [17].

Se examinan algunos métodos para detectar la heterocedasticidad. La mayoría de estos métodos se basan en el examen de los residuos \hat{u}_i de los mínimos cuadrados ordinarios (MCO), pues son éstos los que se observan y no las perturbaciones u_i . Se espera que \hat{u}_i sean buenas estimaciones de u_i , esperanza que se cumple si el tamaño de la muestra es lo bastante grande [17].

4.3.1. Prueba de Breusch-Pagan. Esta prueba fue propuesta por Breusch y Pagan en 1979 y consiste en ajustar un modelo de regresión lineal con variable respuesta dada por residuales del modelo original al cuadrado u_i^2 y como covariables las variables del modelo original [19].

Por ejemplo, si se tienen $k=2$ covariables para explicar a y , entonces el modelo de regresión para estudiar la homocedasticidad es: $\hat{u}_i^2 = \delta_0 + \delta_1 X_1 + \delta_2 X_2 + \mu$ si se concluye que $\delta_1 = \delta_2 = 0$, significa que los residuales no son función de las covariables del modelo. El estadístico en esta prueba está dado por $n \times R^2$ y bajo la hipótesis nula verdadera, el estadístico tiene distribución χ_k^2 .

4.3.2. Prueba de Harvey. La prueba de Harvey implica ajustar un modelo de regresión auxiliar en el que la variable de respuesta es el logaritmo del vector de residuales al cuadrado del modelo original y la matriz de diseño X consta de una o más variables exógenas que se sospecha están relacionadas con la varianza del error. En ausencia de información previa sobre una posible elección de X , normalmente se utilizarían las variables explicativas del modelo original [17].

4.3.3. Prueba de Glejser. Después de obtener los residuales \hat{u}_i de la regresión los mínimos cuadrados ordinarios (MCO), Glejser sugiere una regresión sobre los valores absolutos de \hat{u}_i sobre la variable X que se cree muy asociada con σ_i^2 [17].

4.3.4. *Prueba Breusch-Pagan-Godfrey.* Es un test de autocorrelación en los errores y residuos estadísticos en un modelo de regresión. Hace uso de los errores generados en el modelo de regresión y un test de hipótesis derivado de éste. La hipótesis nula es que no exista correlación serial de cualquier orden sobre p [17].

Para ilustrar esta prueba, considere el modelo de regresión lineal con k variables

$$(4.10) \quad Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i$$

Suponga que la varianza del error σ_i^2 se describe como: $\sigma_i^2 = f(\alpha_1 + \alpha_2 Z_{2i} + \dots + \alpha_m Z_{mi})$ es decir, σ_i^2 es algún tipo de función de las variables Z no estocásticas; alguna de las X o todas ellas pueden servir como Z. Específicamente, suponga que,

$$(4.11) \quad \sigma_i^2 = \alpha_1 + \alpha_2 Z_{2i} + \dots + \alpha_m Z_{mi}$$

Es decir, σ_i^2 es una función lineal de las Z. Si $\alpha_2 = \alpha_3 = \dots = \alpha_m = 0$, $\sigma_i^2 = \alpha_1$, que es una constante. Por consiguiente, para probar si σ_i^2 es homoscedástica, se puede probar la hipótesis de que $\alpha_2 = \alpha_3 = \dots = \alpha_m = 0$.

4.3.5. *Prueba general de White.* La prueba general de heterocedasticidad propuesta por White no se apoya en el supuesto de normalidad. Como ilustración de la idea básica, considere como ejemplo el siguiente modelo de regresión con tres variables [17]:

$$(4.12) \quad Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$$

Ahora para realizar la prueba de White se procede de la siguiente forma:

1. Estimar Y_i y obtener los residuos \hat{u}_i .
2. Efectuar la siguiente regresión (auxiliar):

$$\hat{u}_i^2 = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{3i} + \alpha_4 X_{2i}^2 + \alpha_5 X_{3i}^2 + \alpha_6 X_{2i} X_{3i} + v_i.$$
 Es decir, con el cuadrado de los residuos de la regresión original se hace la regresión sobre las variables o regresoras X originales, sobre sus valores al cuadrado y sobre el (los) producto(s) cruzado(s) de las regresoras. También pueden introducirse potencias más altas de las regresoras. Observe que hay un término constante en esta ecuación, aunque la regresión original puede o no contenerlo. Obtenga R^2 de esta regresión (auxiliar).
3. Según la hipótesis nula de que no hay heteroscedasticidad, puede demostrarse que el tamaño de la muestra (n) multiplicado por R^2 obtenido de la regresión auxiliar asintóticamente sigue la distribución ji cuadrada con gl igual al número de regresoras (sin el término constante) en la regresión auxiliar. Es decir,

$$n \cdot R^2 \sim \chi_{gl}^2$$
 Para este ejemplo, hay 5 gl porque hay 5 regresoras en la regresión auxiliar.
4. Si el valor ji cuadrada obtenido en el inciso 3 excede al valor ji cuadrada crítico en el nivel de significancia seleccionado, la conclusión es que hay heterocedasticidad. Si éste no excede el valor ji cuadrada crítico, no hay heterocedasticidad, lo cual quiere decir que en la regresión auxiliar \hat{u}_i^2 , $\alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = \alpha_6 = 0$.

Conviene hacer un comentario relacionado con la prueba de White. Si un modelo tiene muchas regresoras, la introducción de todas las regresoras, de sus términos elevados al cuadrado (o a potencias más elevadas) y de sus productos cruzados pueden consumir grados de libertad rápidamente. Por consiguiente, se debe tener cautela con esta prueba. En los casos en que el estadístico de prueba de White es significativo estadísticamente, la heterocedasticidad puede no necesariamente ser la causa, sino los errores de especificación [17].

4.3.6. *Errores estándar robustos de White.* El procedimiento generalizado de errores estándar de White para el modelo de regresión de k variables [17]:

$$(4.13) \quad Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i$$

La varianza de todo coeficiente de regresión parcial, $\hat{\beta}_j$, se obtiene así:

$$(4.14) \quad \text{var}(\hat{\beta}_j) = \frac{\sum \hat{w}_{ji}^2 \hat{u}_i^2}{(\sum \hat{w}_{ji}^2)^2}$$

Donde \hat{u}_i son los residuos obtenidos de la regresión original (4.13), y \hat{w}_j son los residuos proporcionados por la regresión (auxiliar) de la regresora X_j sobre las regresoras restantes en (4.13). Es un procedimiento tardado, pues se debe estimar (4.14) para cada variable X .

4.4. Normalidad. Los análisis de normalidad, también llamados contrastes de normalidad, tienen como objetivo analizar cuánto difiere la distribución de los datos observados respecto a lo esperado si, procediesen de una distribución normal con la misma media y desviación típica. Existen diferentes pruebas para realizar el análisis de normalidad, entre ellas Shapiro-Wilk y Jarque-Bera.

4.4.1. *Shapiro-Wilk.* La prueba de Shapiro-Wilk (SW), se restringió originalmente para tamaños de muestra pequeños ($n < 50$). Esta prueba fue la primera que fue capaz de detectar desviaciones de la normalidad, ya sea debido a la asimetría o curtosis, o ambos. Se ha convertido en la prueba preferida debido a sus buenas propiedades de potencia en comparación con una amplia gama de pruebas alternativas [20].

Las hipótesis a contrastar están establecidas de la siguiente manera.

H_0 : La muestra sigue una distribución normal

H_1 : La muestra no sigue una distribución normal

Shapiro y Wilk propusieron una prueba de bondad de ajuste basada en dos estimadores de desviación estándar: L_n , el mejor estimador (MEL) bajo la hipótesis de que esté presente una distribución normal estándar, y el estimador de máxima verosimilitud (EMV) bajo la misma hipótesis.

El estadístico de prueba definido por Shapiro y Wilk en 1965 es el siguiente.

$$(4.15) \quad W = \frac{(\sum_{i=1}^n a_i x_{(i)})^2}{[\sum_{i=1}^n (X_i - \bar{x})^2]}$$

Donde:

x_i : Es la i -ésima estadística de orden.

\bar{x} : la media de la muestra.

$$a_i = (a_1, \dots, a_n) = \frac{m^T V^{-1}}{(m^T V^{-1} m)^{1/2}}$$

$m = (m_1, \dots, m_n)^T$ son los valores esperados de las estadísticas de orden.

V : es la matriz de covarianza de las estadísticas de orden.

La prueba de Shapiro-Wilk fue modificada por Royston en 1982 para ampliar la restricción del tamaño de la muestra, cuyo propósito es habilitar el cálculo del estadístico W y su nivel de significancia para cualquier tamaño de muestra ($3 \leq n \leq 2000$) [20].

Los valores estadísticos para W recaen entre cero y uno. Para valores pequeños de W , hay evidencia de desviación de la normalidad, es decir se rechaza la normalidad; mientras que valor de uno indica la normalidad de los datos.

4.4.2. Jarque-Bera. Jarque y Bera formularon una prueba de normalidad que lleva su nombre, ellos plantean que existen distribuciones que pueden coincidir con la distribución normal, en media y varianza o sea, que su primer momento centrado en cero y su segundo alrededor de la media son los mismos, pero que no necesariamente el tercero y cuarto momentos centrados en la media son iguales. Esa es la razón que los conduce a plantear la prueba de normalidad basada en el sesgo, s , y la curtosis, k , de manera que las hipótesis a probar son [18]:

H_0 Errores normales vs H_1 Errores no normales o equivalentemente,

$H_0 : s = 0, k = 3$ vs $H_1 : s \neq 0$ y $k \neq 3$

Y el estadístico de prueba bajo H_0 es:

$$(4.16) \quad JB = \frac{\widehat{ks}^2}{6} + \frac{(\widehat{kk} - 3)^2}{24} \sim \chi^2_{(2)}$$

Donde \widehat{ks} es el coeficiente de sesgo y el \widehat{kk} coeficiente de curtosis dados por:

$$(4.17) \quad \widehat{ks} = \frac{\widehat{s}}{(\sqrt{\widehat{\sigma}^2})^3}, \widehat{kk} = \frac{\widehat{k}}{(\sqrt{\widehat{\sigma}^2})^4}$$

Y

$$(4.18) \quad \widehat{s} = \frac{1}{T} \sum_{t=1}^T \widehat{\epsilon}_t^3, \widehat{k} = \frac{1}{T} \sum_{t=1}^T \widehat{\epsilon}_t^4$$

Observe que si el estadístico de prueba JB es cercano a cero hay evidencia a favor de que los errores se distribuyen de manera normal, en caso contrario, es decir, cuando JB está alejado de cero, se rechaza la hipótesis nula.

5. ANÁLISIS Y RESULTADOS

5.1. Población estudiada y descripción de las variables. Desde su creación, el LLECE se constituyó como referente y marco regional de concertación y

cooperación entre los países en el ámbito de la evaluación en educación, y como apoyo técnico para la formación y capacitación de los equipos responsables de los sistemas de medición y evaluación nacionales, fuente de acceso a la información y bases de datos a disposición de los países para promover la elaboración de políticas educativas basadas en evidencia empírica.

Dentro de este marco de acción, el LLECE ha desarrollado tres versiones de su estudio regional comparativo y explicativo. En 1997 se aplicó el primer estudio, denominado PERCE, en el cual participaron 13 países. Este estudio evaluó matemática y lectura en tercer y cuarto grados de educación primaria. El segundo estudio (SERCE) se aplicó nueve años más tarde, en 2006, y evaluó las áreas de matemática, lectura en tercer y sexto grados, y en ciencias naturales en sexto grado. El tercer estudio (TERCE) se aplicó en 2013, evaluando las mismas áreas y grados que SERCE [22].

Debido a los esfuerzos realizados por la UNESCO y LLECE la secretaría de educación de Honduras ratificó un acuerdo en el año de 1999 que permite a LLECE la elaboración de informes de lenguaje, matemática y factores asociados a la calidad de la educación [21], esta es una de las razones por las cuales Honduras participó en el TERCE. Al no tener acceso público a las bases de datos de estudios realizados en Honduras sobre rendimiento escolar, se tomó la decisión de trabajar con las bases de datos públicas del TERCE, seleccionado únicamente los datos correspondientes a Honduras para realizar este estudio.

Las bases de datos de español y matemática tercer grado constan de 167 variables cada una, contienen preguntas relativas a aspectos educativos del hogar, procesos y estrategias de aprendizaje dentro de la sala de clases, actividades recreativas y disponibilidad de materiales. Mientras que las bases de datos de español y matemática sexto grado constan de 196 variables cada una, también contienen preguntas relativas a los aspectos que se tomaron para tercer grado [22].

5.2. Componentes Principales. El análisis de componentes principales ACP es particularmente útil para reducir la dimensionalidad de un grupo de datos. Un ACP se ocupa de explicar la estructura de varianza-covarianza de un conjunto de variables a través de algunas combinaciones lineales de estas. A menudo revela relaciones que no se sospechaban anteriormente y, por lo tanto permite interpretaciones que normalmente no resultarían [23].

Algebraicamente, los componentes principales son combinaciones lineales particulares de las variables aleatorias p , X_1, X_2, \dots, X_p . Geométricamente, estas combinaciones lineales representan la selección de un nuevo sistema de coordenadas obtenida mediante la rotación del sistema original, con X_1, X_2, \dots, X_p como ejes de coordenadas. Los nuevos ejes representan las direcciones con la máxima variabilidad y proporcionan una descripción más simple y parsimoniosa de la estructura de covarianza.

Se seleccionaron las variables que serían analizadas en los diferentes modelos utilizados, aplicando la técnica estadística de análisis de componentes principales

(ACP), el proceso consistió en realizar combinaciones con las variables seleccionadas, más la ayuda de estudios anteriores a fin de obtener un modelo explicativo, y finalmente se obtuvieron los modelos. Vale mencionar que toda la información de las cuatro bases de datos se procesó con el paquete estadístico RStudio.

VARIABLES UTILIZADAS PARA LOS MODELOS DE ESPAÑOL Y MATEMÁTICA TERCER GRADO:
Código escuela, código estudiante, género, edad, disponibilidad de libros y materiales en la escuela según el estudiante, trabajo infantil, clima en el aula escolar, prácticas de recreación, nivel educativo de los padres, ingresos en el hogar, servicios con los que cuenta el hogar, bienes con los que cuenta el hogar, alimentación en el hogar, servicios con los que cuenta la comunidad, clima negativo del barrio o comunidad hogar, clima positivo del barrio o comunidad hogar, gusto por la lectura, supervisión de estudios en el hogar, comunicación con la escuela, el director cuenta con 2 trabajos, instalaciones en la escuela, infraestructura del aula, servicios básicos de la escuela, disponibilidad de libros y materiales en la escuela según el director, programas escolares adicionales, satisfacción laboral según el director, evaluación desempeño docente, consejo docente, preparación docente, actualización docente, técnicas de evaluación, ambiente laboral según el docente, violencia en el aula según el docente, clima del aula según el docente, gestión directiva interna según el docente, acompañamiento docente, monitoreo de las prácticas docentes, puntaje estándar.

VARIABLES UTILIZADAS PARA LOS MODELOS DE ESPAÑOL Y MATEMÁTICA SEXTO GRADO:
Del modelo de tercer grado se quitaron las variables, disponibilidad de libros y materiales en la escuela según el estudiante, el director cuenta con 2 trabajos, disponibilidad de libros y materiales en la escuela según el director, consejo docente, violencia en el aula según el docente. Y se agregaron las variables; puntualidad del docente, actitud del docente en el salón de clases, preparación docente en la materia, metodología docente, liderazgo en el salón de clases, hostilidad en el salón de clases, razón por la que lee el niño, uso del computador, acceso a internet, uso del computador en el tiempo libre, nivel educativo de los padres según el estudiante, asistencia al centro educativo, gestión directiva interna.

5.3. Modelos y Resultados. Para los modelos se utilizaron las bases de datos públicas del TERCE, con los datos correspondientes a Honduras, se busca conocer los factores asociados al rendimiento escolar, utilizando regresión múltiple. Para ello, se cuenta con una variable puntaje estándar para el rendimiento en español y matemática en tercero y sexto grado, más las variables explicativas.

Modelo lineal de español tercer grado

El modelo con todas las variables tiene un R^2 ajustado de 0,2617, es capaz de explicar el 26,17%. En la tabla 1 se presentan las variables significativas del modelo lineal de español tercer grado y según los resultados obtenidos, es relevante examinar la posible existencia de multicolinealidad ya que de 35 variables en el modelo solo 17 son significativas.

Modelo lineal de matemática tercer grado

TABLA 1. Variables significativas del modelo lineal de español tercer grado

| Nombre de las variables | Estimación | Error Estándar | Valor t | pr(> t) |
|--|------------|----------------|---------|-----------|
| Género | -5,57 | 2,60 | -2,14 | 0,032 |
| Clima en el aula escolar | 5,80 | 0,79 | 7,29 | 0,000 |
| Prácticas de recreación | 1,87 | 0,31 | 6,07 | 0,000 |
| Ingresos en el hogar | 0,64 | 0,24 | 2,72 | 0,007 |
| Servicios con los que cuenta el hogar | 2,84 | 1,06 | 2,69 | 0,007 |
| Instalaciones en la escuela | 3,53 | 0,64 | 5,54 | 0,000 |
| Infraestructura del aula | -1,48 | 0,89 | -1,66 | 0,096 |
| Servicios básicos de la escuela | 10,18 | 1,52 | 6,69 | 0,000 |
| Libros en la escuela según el director | -11,53 | 1,61 | -7,17 | 0,000 |
| Evaluación desempeño docente | 1,64 | 0,66 | 2,49 | 0,013 |
| Consejo docente | 1,33 | 0,43 | 3,10 | 0,002 |
| Preparación docente | 0,83 | 0,43 | 1,94 | 0,052 |
| Actualización docente | 4,68 | 2,27 | 2,07 | 0,039 |
| Técnicas de evaluación | -0,94 | 0,42 | -2,22 | 0,027 |
| Ambiente laboral según el docente | -3,28 | 0,63 | -5,22 | 0,000 |
| Violencia en el aula según el docente | 1,42 | 0,73 | 1,96 | 0,050 |
| Acompañamiento docente | 0,96 | 0,35 | 2,70 | 0,007 |

TABLA 2. Variables significativas del modelo lineal de matemática tercer grado

| Nombre de las variables | Estimación | Error Estándar | Valor t | pr(> t) |
|--|------------|----------------|---------|-----------|
| Libros en la escuela según el estudiante | 3,81 | 1,52 | 2,52 | 0,012 |
| Trabajo infantil | 4,34 | 1,20 | 3,61 | 0,000 |
| Clima en el aula escolar | 8,22 | 0,78 | 10,59 | 0,000 |
| Nivel educativo de los padres | 1,11 | 0,42 | 2,62 | 0,009 |
| Servicios en el hogar | 2,61 | 1,03 | 2,53 | 0,011 |
| Servicios en la comunidad | -0,80 | 0,43 | -1,86 | 0,064 |
| Clima negativo del barrio | 0,44 | 0,21 | 2,09 | 0,037 |
| Gusto por la lectura | 1,55 | 0,78 | 2,00 | 0,046 |
| El director cuenta con 2 trabajos | -6,01 | 2,62 | -2,29 | 0,022 |
| Instalaciones en la escuela | 1,77 | 0,62 | 2,85 | 0,004 |
| Servicios básicos de la escuela | 10,40 | 1,48 | 7,04 | 0,000 |
| Libros en la escuela según el director | -11,11 | 1,59 | -7,00 | 0,000 |
| Programas escolares adicionales | -1,00 | 0,59 | -1,71 | 0,088 |
| Satisfacción laboral según el director | -1,05 | 0,64 | -1,66 | 0,098 |
| Evaluación desempeño docente | 3,67 | 0,66 | 5,59 | 0,000 |
| Preparación docente | 1,22 | 0,42 | 2,93 | 0,003 |
| Actualización docente | 3,69 | 2,21 | 1,67 | 0,094 |
| Ambiente laboral según el docente | -2,21 | 0,61 | -3,62 | 0,000 |
| Gestión directiva interna según el docente | 0,99 | 0,34 | 2,88 | 0,004 |
| Monitoreo de las prácticas docentes | 1,06 | 0,32 | 3,36 | 0,001 |

El modelo con todas las variables tiene un R^2 ajustado de 0,24, es capaz de explicar el 24,00%. En la tabla 2 se presentan las variables significativas del modelo lineal de matemática tercer grado y según los resultados obtenidos, es relevante examinar la posible existencia de multicolinealidad ya que de 35 variables en el modelo solo 20 resultaron significativas.

Modelo lineal de español sexto grado

TABLA 3. Variables significativas del modelo lineal de español sexto grado

| Nombre de las variables | Estimación | Error Estándar | Valor t | pr(> t) |
|--|------------|----------------|---------|-----------|
| Género | -6,77 | 2,42 | -2,79 | 0,005 |
| Nivel educativo de los padres | 1,63 | 0,37 | 4,42 | 0,000 |
| Servicios con los que cuenta el hogar | 3,56 | 0,86 | 4,16 | 0,000 |
| Bienes con los que cuenta el hogar | 2,00 | 0,29 | 6,96 | 0,000 |
| Clima en el salón de clases | -3,48 | 0,88 | -3,96 | 0,000 |
| Puntualidad del docente | 5,81 | 1,10 | 5,27 | 0,000 |
| Actitud del docente en clases | 1,19 | 0,54 | 2,20 | 0,028 |
| Preparación docente en la materia | 2,26 | 0,71 | 3,19 | 0,001 |
| Liderazgo en el salón de clases | -2,38 | 0,63 | -3,77 | 0,000 |
| Hostilidad en el salón de clases | 2,14 | 0,63 | 3,38 | 0,001 |
| Razón por la que lee el niño | 0,82 | 0,43 | 1,89 | 0,059 |
| Prácticas de recreación | -3,82 | 1,00 | -3,84 | 0,000 |
| Uso del computador | -11,08 | 2,63 | -4,21 | 0,000 |
| Acceso a internet | 10,94 | 2,46 | 4,44 | 0,000 |
| Uso del computador en el tiempo libre | 1,10 | 0,54 | 2,02 | 0,044 |
| Ingresos en el hogar | 0,53 | 0,19 | 2,78 | 0,006 |
| Servicios en la comunidad | 1,45 | 0,73 | 1,98 | 0,047 |
| Clima negativo del barrio | -0,75 | 0,28 | -2,68 | 0,007 |
| Supervisión de estudios en el hogar | 2,11 | 0,74 | 2,85 | 0,004 |
| Comunicación con la escuela | -1,51 | 0,41 | -3,71 | 0,000 |
| Servicios básicos de la escuela | 5,94 | 0,86 | 6,90 | 0,000 |
| Satisfacción laboral según el director | -0,90 | 0,53 | -1,71 | 0,088 |
| Evaluación desempeño docente | 0,13 | 0,06 | 2,23 | 0,026 |
| Preparación docente | -0,32 | 0,14 | -2,34 | 0,019 |

El modelo con todas las variables tiene un R^2 ajustado de 0,3251, es capaz de explicar el 32,51%. En la tabla 3 se presentan las variables significativas del modelo lineal de español sexto grado y según los resultados obtenidos, es relevante examinar la posible existencia de multicolinealidad ya que de 42 variables en el modelo solo 27 son significativas.

Modelo lineal de matemática sexto grado

TABLA 4. Variables significativas del modelo lineal de matemática sexto grado

| Nombre de las variables | Estimación | Error Estándar | Valor t | pr(> t) |
|---|------------|----------------|---------|-----------|
| Género | 6,05 | 2,66 | 2,28 | 0,023 |
| Edad | 7,35 | 2,78 | 2,64 | 0,008 |
| Nivel educativo de los padres | 1,64 | 0,50 | 3,30 | 0,001 |
| Bienes en el hogar | 1,13 | 0,32 | 3,57 | 0,000 |
| Puntualidad del docente | 3,05 | 1,18 | 2,60 | 0,009 |
| Actitud del docente en clases | 1,15 | 0,60 | 1,92 | 0,055 |
| Liderazgo en el salón de clases | -2,33 | 0,69 | -3,36 | 0,001 |
| Hostilidad en el salón de clases | 2,18 | 0,67 | 3,24 | 0,001 |
| Razón por la que lee el niño | 1,29 | 0,46 | 2,79 | 0,005 |
| Prácticas de recreación | -2,95 | 1,05 | -2,79 | 0,005 |
| Uso del computador | -8,37 | 2,98 | -2,81 | 0,005 |
| Acceso a internet | 5,63 | 2,92 | 1,93 | 0,054 |
| Uso del computador en el tiempo libre | 1,73 | 0,58 | 2,96 | 0,003 |
| Nivel educativo de los padres | 0,87 | 0,42 | 2,04 | 0,041 |
| Clima negativo del barrio | 0,55 | 0,23 | 2,40 | 0,017 |
| Clima positivo del barrio | -0,68 | 0,35 | -1,92 | 0,055 |
| Gusto por la lectura | 3,00 | 0,78 | 3,86 | 0,000 |
| Asistencia al centro educativo | 1,34 | 0,31 | 4,26 | 0,000 |
| Instalaciones en la escuela | 2,52 | 0,50 | 5,03 | 0,000 |
| Servicios básicos de la escuela | 2,00 | 0,95 | 2,10 | 0,036 |
| Preparación docente | 0,46 | 0,23 | 1,98 | 0,048 |
| Actualización docente | -3,52 | 1,67 | -2,10 | 0,036 |
| Clima del aula según docente | 2,49 | 0,71 | 3,49 | 0,000 |
| Acompañamiento docente | 1,78 | 0,61 | 2,93 | 0,003 |
| Gestión directiva interna según docente | -4,00 | 0,51 | -7,77 | 0,000 |
| Monitoreo de las prácticas docentes | 0,99 | 0,37 | 2,66 | 0,008 |

El modelo con todas las variables tiene un R^2 ajustado de 0,1999, es capaz de explicar el 19,99%. En la tabla 4 se presentan las variables significativas del modelo lineal de matemática sexto grado y según los resultados obtenidos, es relevante examinar la posible existencia de multicolinealidad ya que de 42 variables en el modelo solo 26 son significativas.

Una primera evidencia de posible elevada colinealidad entre las variables se deriva de la nula significancia de muchas de las variables, como se ha reflejado en los resultados de los cuatro modelos presentados. Para intentar confirmar esta evidencia es preciso realizar algunas exploraciones adicionales.

5.4. Verificación de los Supuestos. La información relevante en los modelos de regresión múltiple, está contenida en las variables seleccionadas. Los modelos operan bajo el supuesto de que el modelo contiene todas las variables relevantes para explicar el modelo. En este sentido la realización de pruebas de diagnóstico sobre la selección eficiente de las variables incluidas en los modelos es necesaria. La omisión de variables importantes en los modelos, es un problema relevante en la especificación de los mismos y en este sentido se pueden generar problemas de multicolinealidad y heterocedasticidad.

5.4.1. Supuesto de Colinealidad. Se realizó la matriz de correlación para cada uno de los modelos y mostró correlación entre algunas de las variables, rechazando así el supuesto de colinealidad. Para saber que tan grave es el problema de multicolinealidad en los modelos se realizaron las pruebas descritas anteriormente:

La prueba de Theil mostró para el modelo lineal de español tercer grado que el R^2 ajustada de 0.3812 la cual es superior a la de 0.2617 del modelo original, como el efecto es menor, esto indica un grado moderado de multicolinealidad, obteniendo resultados similares para los 3 modelos restantes.

La prueba de Klein para el modelo lineal de español tercer grado, en esta prueba se puede observar que hay una colinealidad muy alta en 4 de las variables como ser, servicios con los que cuenta el hogar, bienes con los que cuenta el hogar, supervisión de estudios en el hogar, servicios básicos de la escuela, ya que presentan un R^2 ajustada alto de 0,7159, 0,7119, 0,7313, 0,7589 respectivamente, en comparación con R^2 ajustada original de 0,2617, el resto de las variables presentó un grado moderado de colinealidad. Obteniendo resultados similares para los 3 modelos restantes.

En la prueba del factor de inflación de la varianza (VIF) los resultados indicaron que es posible establecer la existencia de problemas de colinealidad moderados en virtud de que las variables presentan una VIF por debajo de 5, esto para todos los cuatro modelos.

5.4.2. Supuesto de Homocedasticidad. Para el modelo lineal de español tercer grado se aplicaron seis pruebas, Prueba Breusch-Pagan, Prueba Harvey, Prueba Glesjer, Breusch-Pagan-godfrey, Breusch-Pagan-estudentizado, Prueba White, de las cuales cuatro rechazan la hipótesis nula y la Prueba Harvey, Breusch-Pagan-estudentizado no la rechazan. Debido a esto se decidió transformar la variable dependiente en logaritmo y realizar las pruebas nuevamente para comparar resultados. Y se obtuvieron resultados similares en este caso solo la prueba de Harvey sugiere que no hay heterocedasticidad.

5.4.3. Supuesto de Normalidad. La normalidad consiste en que todas las muestras, tanto de las variables independientes así como de la variable dependiente, tienen que estar distribuidos normalmente, es decir cumplir con la distribución normal.

Se realizaron las pruebas de Shapiro-Wilk y Jarquea-Bera, en ambas pruebas los cuatro modelos rechazaron la hipótesis nula, indicando que los modelos no se distribuyen normal.

6. CONCLUSIONES Y ESTUDIOS A FUTURO

1. La regresión múltiple se utiliza para examinar la relación entre varias variables independientes y una variable dependiente, con los supuestos de colinealidad, homocedasticidad, y normalidad.

2. En los cuatro modelos presentados y según las pruebas realizadas para verificar la existencia de problemas de colinealidad, se puede concluir que es un problema moderado y se pueden aceptar los modelos.
3. Para el problema de heterocedasticidad en los cuatro modelos presentados y según las pruebas realizadas, se puede concluir que existe heterocedasticidad en los modelos. Se realizó la prueba de errores estándar robustos de White, para saber que tan grave es el problema y según los resultados de los modelos lineales transformando la variable dependiente en logaritmo no es tan grave, así que se pueden aceptar los modelos.
4. Los cuatro modelos propuestos rechazan la hipótesis nula de normalidad.

Como trabajo a futuro utilizando los datos del TERCE se puede realizar modelaciones multinivel donde los datos están estructurados de manera jerárquica. En otras palabras, que algunas de las variables de análisis están anidadas o agrupadas en otras de mayor nivel y así comprender la varianza que produce cada nivel de agregación.

REFERENCIAS

1. UNESCO, *Informe de seguimiento de la educación para todos en el mundo*, Fortemps, 2007.
2. Eduardo Velez, *Factores que afectan el rendimiento académico en la educación primaria*, Revista Latinoamericana de Innovaciones Educativas, 1994.
3. Anita Magaly González Bustamante, *Dirección escolar exitosa en España*, Universidad Autónoma de Madrid, 2015.
4. UNESCO, *Situación y Tendencias 2000 Evaluación del Aprovechamiento Escolar*, Foro consultivo internacional sobre educación para todos, 2000.
5. Arturo de la Orden, Laura Oliveros, Joseph Mafokozí, Coral Gonzalez, *Modelos de investigación del bajo rendimiento*, Revista Complutense de Educación, Vol. 12 Núm. 1(2001): 159-178.
6. Montero Granados, *Modelos de regresión lineal múltiple*, Universidad de Granada, España, 2016.
7. F. Javier Murillo Torrecilla, *Los modelos multinivel como herramienta para la investigación educativa*, Magis. Revista Internacional de Investigación en Educación, vol. 1, núm. 1, julio-diciembre, 2008, pp. 45-62.
8. Miguel A. Ruiz, Antonio Pardo, Rafael San Martín, *Modelos de ecuaciones estructurales*, Papeles del Psicólogo, vol. 31, núm. 1, enero-abril, 2010, pp. 34-45.
9. Javier Murillo Torrecilla, *Investigación iberoamericana sobre eficacia escolar*, Bogotá: Convenio Andrés Bello, 2007.
10. James S. Coleman, and others, *Equality of Educational opportunity*, National Center for Educational, Statistics Number: OE-38001.
11. George Weber, *Inner-City Children Can Be Taught to Read: Four Successful Schools*, Council for Basic Education, 1971.
12. Michael O. Martin, Dana L. Kelly, *Technical Report, Volume III: Implementation and Analysis*, Library of Congress Catalog Card Number: 99-60717.
13. UNESCO, *Primer estudio internacional comparativo sobre lenguaje matemática y factores asociados, para alumnos del tercer y cuarto grado de la educación básica*, Publicado por el Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, 2000.
14. UMCE, *Factores Asociados al Rendimiento Académico 2017, Plan de Nación 2010-2022*, Universidad Pedagógica Nacional Francisco Morazán, 20018.
15. MIDEH, *Estudio de Factores Asociados al Desempeño Académico 2017*, USAID HONDURAS, 20017.
16. Ezequiel Uriel, *Regresión Lineal Múltiple*, Universidad de Valencia, versión 09-2013.
17. Danomar Gujarati, Dawn Porter, *Econometría*, McGraw-Hill/Irwin, 2010.
18. Luis Quintana, *Econometría aplicada usando R*, UNAM, 2016.
19. Freddy Hernández, Mauricio Mazo, *Modelos de Regresión con R*, Universidad Nacional de Colombia, 2020.

20. Mariana Carmona, Hanne Carrion, *Potencia de la prueba de estadística de normalidad*, UAEM, 2015.
21. SE, *Reporte Técnico Tercer Estudio Comparativo y Explicativo*, Secretaria de educación Honduras, 1999.
22. UNESCO, *Carta/acuerdo entre la secretaria de educación y la UNESCO*, Oficina Regional de Educación para América Latina y el Caribe, 2016.
23. Richard A. Johnson, Dean W. Wichern, *Applied Multivariate Statical analysis*, Pearson , 2007.

ESCUELA DE MATEMÁTICA, UNIVERSIDAD NACIONAL AUNTÓNOMA DE HONDURAS
Dirección actual: Departamento de Matemática Aplicada, Facultad de Ciencias, Universidad Nacional Autónoma de Honduras, Ciudad Universitaria
Dirección de correo electrónico: rivaslesky@gmail.com