



Multimodelos empleando Mapas Auto-Organizados de Kohonen para la predicción de series de tiempo caóticas

TESIS

que para obtener el grado de
Maestro en Ingeniería Eléctrica

presenta

José Luis Guillén Aguirre

Dr. J. Jesús Rico Melgoza

Director de tesis

Universidad Michoacana de San Nicolás de Hidalgo
Facultad de Ingeniería Eléctrica
División de Estudios de Posgrado

Morelia, Michoacán, Octubre 2008.

MULTIMODELOS EMPLEANDO MAPAS AUTO-ORGANIZADOS DE
KOHONEN PARA LA PREDICCIÓN DE SERIES DE TIEMPO
CAÓTICAS

Los miembros del Jurado de Examen de Grado aprueban
la tesis de maestría de *José Luis Guillén Aguirre*

Dr. J. Jesús Rico Melgoza
Asesor

J. Jesús Rico Melgoza

Dr. Juan José Flores Romero



Dr. Félix Calderón Solorio




Dr. J. Aurelio Medina Ríos



Dr. Manuel Madrigal Martínez



Dr. J. Aurelio Medina Ríos
Jefe de la División de Estudios de Posgrado
de la Facultad de Ingeniería Eléctrica





Índice general

Índice de figuras	4
1. Introducción.....	7
1.1. Presentación del problema	8
1.2. Objetivos	8
1.3. Metodología	8
1.4. Descripción de la tesis.....	12
1.5. Contribución	14
2. Respuesta en el tiempo de sistemas caóticos clásicos	15
2.1. Introducción	15
2.2. Sistema de Lorenz.....	16
2.3. Sistema de Chua.....	19
2.4. Sistema de Chen.....	23
2.5. Sistema de Rössler	25
2.6. Conclusiones	27
3. Características de los sistemas no lineales.....	28
3.1. Introducción	28
3.2. Mapas iterados	29
3.3. Mapas de Poincaré	36
3.4. Ciclos límite	41
3.5. Atractores extraños	44
3.6. Gráficos de recurrencia	46
3.7. Información mutua promedio.....	48
3.8. Algoritmo de vecinos falsos	50
3.9. Conclusiones	51
4. Mínimos cuadrados	52
4.1. Introducción	52
4.2. Mínimos cuadrados convencionales.....	52
4.3. Mínimos cuadrados ponderados.....	55

4.4.	Comparación entre mínimos cuadrados convencionales y mínimos cuadrados ponderados	57
4.5.	Conclusiones	60
5.	Clustering con Kohonen.....	61
5.1.	Introducción	61
5.2.	Algoritmo de los mapas auto-organizados	62
5.3.	Ejemplos de clasificación con mapas auto-organizados.....	68
5.4.	Conclusiones	82
6.	Predicción de series de tiempo caóticas	83
6.1.	Introducción	83
6.2.	Teorema de Takens	83
6.3.	Modelado dinámico: Multi-modelos	84
6.4.	Casos de estudio.....	86
6.5.	Estimación del error	86
6.6.	Ejemplos de predicción de series de tiempo	87
6.7.	Comportamiento del error medio cuadrático normalizado en el sistema	108
6.8.	Conclusiones	110
7.	Conclusiones y trabajos futuros.	111
	Bibliografía	114

Índice de figuras

Figura 1.1. Relación entre regímenes de comportamiento de la serie y la red neuronal.....	11
Figura 1.2. Esquema de reconstrucción del modelo global del sistema.....	11
Figura 1.3 Diagrama del esquema de predicción.....	12
Figura 2.1. Gráfico de la variable $x(t)$ de la serie de tiempo de Lorenz.....	17
Figura 2.2. Plano de fase del Sistema de Lorenz, mostrando la trayectoria de x vs. z y la estructura subyacente, conocida como <i>Mariposa</i>	18
Figura 2.3. Gráfico tridimensional de la trayectoria del sistema de Lorenz.....	19
Figura 2.4. Circuito elemental de Chua.....	20
Figura 2.5. Función de transferencia del resistor no lineal del circuito de Chua.....	20
Figura 2.6. Nacimiento de la doble espiral.....	21
Figura 2.7 Atractor de Chua.....	21
Figura 2.8. Atractor de Chua visto en tres dimensiones.....	22
Figura 2.9. Atractor de Chua que muestra los semiplanos en que se desarrollan las espirales.....	22
Figura 2.10. Proyección en el plano xy del atractor de Chen.....	24
Figura 2.11. Vista en tres dimensiones del atractor de Chen.....	24
Figura 2.12. Serie de tiempo para la variable x del Sistema de Rössler.....	25
Figura 2.13. Atractor extraño del Sistema de Rössler.....	26
Figura 2.14. Proceso de expansión, doblamiento y re-inyección de un sistema caótico.....	27
Figura 3.1. Gráfico de la ecuación logística.....	30
Figura 3.2. Gráfico iterado de la ecuación logística con $r = 2.9$	31
Figura 3.3. Gráfico de “telaraña” para la respuesta del sistema.....	32
Figura 3.4. Gráfico de la ecuación logística para un valor de $r = 0.0$	33
Figura 3.5. Respuesta oscilatoria del sistema cuando $r = 3.2$	34
Figura 3.6. Gráfico de telaraña que muestra el ciclo de óctuple periodo del sistema en $r = 3.56$. No se muestra el transitorio inicial.....	35
Figura 3.7. Diagrama orbital para la ecuación logística.....	35
Figura 3.8. Gráfico de telaraña para $r = 3.7$. No se muestra el transitorio inicial.....	36
Figura 3.9. Diagrama de fase de un péndulo.....	37
Figura 3.10. Mapa de Poincaré.....	38
Figura 3.11. Serie de tiempo para el oscilador dado por la ecuación (3.4).....	39
Figura 3.12. Divergencia exponencial en un sistema ante condiciones iniciales ligeramente distintas.....	40
Figura 3.13. Diagrama de fase que muestra la divergencia exponencial del sistema ante condiciones iniciales ligeramente distinta.....	40
Figura 3.14. Mapa de Poincaré para el oscilador de la ecuación (3.4).....	41
Figura 3.15. Gráfico del espacio de fase para la ecuación de Van der Pol con $\mu = 1.5$	43

Figura 3.16. Gráfico del espacio de fase para la ecuación de Van der Pol con $\mu = 5$.	43
Figura 3.17. Atractor caótico de Hénon.	45
Figura 3.18. Ampliación sobre una región del atractor de Hénon de la Figura 3.17.	46
Figura 3.19. Gráficos de recurrencia para diferentes series de tiempo.	47
Figura 4.1. Ajuste de curvas.	53
Figura 4.2. Gráfico de la función (4.13) y de los datos generados.	58
Figura 4.3. Gráfico de las funciones obtenidas con los métodos de mínimos cuadrados convencionales y mínimos cuadrados ponderados.	59
Figura 5.1. Arreglo de celdas para un mapa.	62
Figura 5.2. Arreglo de neuronas: (a) hexagonal. (b) rectangular.	63
Figura 5.3. Conexión en paralelo del vector de entrada con las neuronas de la red.	63
Figura 5.4. Subconjunto de relación de la neurona ganadora.	64
Figura 5.5. Funciones de relaciones laterales posibles para los subconjuntos topológicos.	65
Figura 5.6. Ejemplo de vecindarios topológicos $N_c(t)$, donde $t_1 < t_2 < t_3$.	66
Figura 5.7. Efecto de una baja y una alta resolución.	67
Figura 5.8. Puntos de entrada al mapa auto-organizado.	69
Figura 5.9. Mapa inicial.	70
Figura 5.10. Mapa final.	70
Figura 5.11. Mapa inicial de una sola dimensión.	71
Figura 5.12. Mapa final de una sola dimensión.	72
Figura 5.13. Puntos empleados para mostrar la característica de clasificador del mapa auto-organizado.	72
Figura 5.14. Mapa final.	73
Figura 5.15. Matriz de distancia unificada de la red.(U-Matrix)	76
Figura 5.16. Planos componentes del mapa auto organizado del ejemplo.	77
Figura 5.17. Mapa con etiquetas.	78
Figura 5.18. Proyección de los datos en el mapa.	79
Figura 5.19. U-matrix para la serie de tiempo de Lorenz.	80
Figura 5.20. Planos componentes para la serie de tiempo de Lorenz	80
Figura 5.21. Proyección de los datos del mapa para la serie de tiempo de Lorenz.	81
Figura 5.22. Atractor de Lorenz, proyección en dos dimensiones.	81
Figura 6.1. Esquema del proceso de reconstrucción de sistemas.	88
Figura 6.2. Primeros 1000 puntos de la serie de tiempo de Lorenz.	91
Figura 6.3. Gráfico de la predicción de la serie de Lorenz, modalidad de un paso.	92
Figura 6.4. Gráfico de error para la predicción de un paso del sistema de Lorenz.	93
Figura 6.5. Gráfico de la predicción de la serie de Lorenz, modalidad iterada.	94
Figura 6.6. Gráfico de error para la predicción iterada del sistema de Lorenz.	94

Figura 6.7. Gráfico de la predicción del sistema de Chen, modalidad de un paso.....	96
Figura 6.8. Gráfico de error para la predicción de un paso del sistema de Chen.....	96
Figura 6.9. Gráfico de la predicción del sistema de Chen, modalidad iterada.....	97
Figura 6.10. Gráfico de error para la predicción iterada del sistema de Chen.....	97
Figura 6.11. Primeros 1500 puntos para la serie de tiempo del sistema de Chua.....	99
Figura 6.12. Gráfico de la predicción del sistema de Chua, modalidad de un paso.....	100
Figura 6.13. Gráfico de error para la predicción de un paso del sistema de Chen.....	100
Figura 6.14. Gráfico de la predicción del sistema de Chua, modalidad iterada.....	101
Figura 6.15. Gráfico de error para la predicción iterada del sistema de Chua.....	101
Figura 6.16. Serie de tiempo para la variables x del sistema de Rössler.....	102
Figura 6.17. Gráfico de la predicción del sistema de Rössler, modalidad de un paso.....	103
Figura 6.18. Gráfico de error para la predicción de un paso del sistema de Rössler.....	104
Figura 6.19. Gráfico de la predicción del sistema de Rössler, modalidad iterada.....	104
Figura 6.20. Gráfico de error para la predicción iterada del sistema de Rössler.....	105
Figura 6.21. Gráfico de la predicción del sistema de Duffing, modalidad de un paso.....	106
Figura 6.22. Gráfico de error para la predicción de un paso del sistema de Duffing.....	107
Figura 6.23. Gráfico de la predicción del sistema de Duffing, modalidad iterada.....	107
Figura 6.24. Gráfico de error para la predicción iterada del sistema de Duffing.....	108
Figura 6.25. Gráfico de comportamiento del error medio cuadrático.....	109
Figura 6.26. Gráfico de comportamiento del error medio cuadrático, para pronóstico iterado.....	110

1. Introducción.

El estudio de los sistemas caóticos ha cobrado una gran importancia desde el descubrimiento del primer sistema caótico en 1963, con las investigaciones de Lorenz sobre modelos que explicaran los fenómenos climáticos [Lorenz, 1963]. A partir de entonces, las metodologías para el análisis de tales sistemas es un tema de gran importancia, a tal grado que se han creado asociaciones dedicadas a la investigación en las áreas de sistemas caóticos. Otro aspecto de la importancia de los sistemas caóticos es el hecho de que la mayoría de los sistemas reales presentan un comportamiento no lineal, desde el punto de vista de modelado matemático, y por tanto una gran posibilidad de regímenes caóticos. Debe aclararse que una condición necesaria para que un sistema pueda presentar un comportamiento caótico es el hecho de que debe ser un sistema no lineal, y que tales sistemas, pueden o no, encontrarse en un momento dado en un régimen caótico. Entre los sistemas caóticos podemos mencionar, como ya se dijo anteriormente, el clima o fenómenos sísmicos [van Zyl y Omlin, 2001], sistemas desarrollados por el hombre, como lo son los sistemas eléctricos de potencia, y comportamientos biológicos de algunos trastornos de salud, como puede ser la metástasis de células cancerígenas [Varsta et al, 1997; Kaplan y Glass, 1995], por mencionar sólo algunos. Todos estos sistemas tienen en común la gran importancia que representa el poder entenderlos, principalmente para predecir los primeros, que son fenómenos naturales y predecir, e incluso controlar los últimos.

Los fenómenos caóticos se consideraron siempre como fenómenos estocásticos, es decir fenómenos de naturaleza aleatoria, y por tanto, impredecibles, en oposición a los sistemas deterministas, cuyos modelos permiten conocer el estado del sistema en todo instante de tiempo. Sin embargo, en la actualidad es bien conocida la aparición de patrones en el comportamiento de tales sistemas, lo cual trae consigo la posibilidad de predicción [Principe et al, 1998]. Más aún, es posible especificar el grado en que es predecible el sistema en cuestión [Smith et al, 1999].

1.1. Presentación del problema

Se pretende caracterizar series de tiempo de sistemas no lineales en regímenes caóticos, para su predicción. Para tal propósito se emplearán métodos analíticos de obtención de características importantes del sistema, como lo son el método de *información mutua promedio*, y el *algoritmo de vecinos falsos*; redes neuronales en conjunto con modelos locales o multi-modelos, para su clasificación y el método de mínimos cuadrados ponderados para la predicción.

1.2. Objetivos

El objetivo principal de la tesis es presentar un esquema de predicción para sistemas caóticos basado en clasificación de modelos locales. La clasificación se llevará a cabo mediante el uso de *mapas auto-organizados de Kohonen*. Los modelos obtenidos deben ser suficientemente simples para ser tratados por métodos bien conocidos como lo son los *mínimos cuadrados*.

1.3. Metodología

A continuación se describe la metodología con la cual se llevará a cabo el proceso de predicción.

El esquema que se propone en esta tesis, consiste a grandes rasgos de cuatro etapas:

- Pre-procesamiento de la señal
- Obtención de parámetros

-
- Clasificación de modelos
 - Reconstrucción del sistema

La primera etapa consiste en el acondicionamiento de la señal o serie de tiempo, ya sea mediante filtrado, eliminación de tendencias que pudiera presentar tal señal, o normalización de los datos, y depende grandemente de los datos mismos, puesto que puede ser posible que para una señal se aplique alguno de los procesos mencionado, todos o ninguno. Es por tanto un proceso relativamente heurístico, que la respuesta predictiva final obtenida permite refinar en procesos posteriores.

Una serie de tiempo se puede definir como *aquella sucesión de observaciones de un fenómeno que es variable en el tiempo* [Chao, 1993]. Cabe señalar que éste término se emplea a lo largo de la tesis de una manera más bien generalizada, dado que no únicamente se utilizarán variables temporales, sino cualquier dato que forme parte de una sucesión y que sea susceptible de predicción bajo las condiciones que se exponen en el capítulo 6.

Una vez que la señal está acondicionada, el esquema requiere caracterizar adecuadamente los datos, y para ello es necesario tener conocimientos sobre su dinámica. Una de las características más importantes a conocer es probablemente la dimensión o espacio de fase. Sin embargo, es poco común que en sistemas complejos se cuente con observaciones de todas y cada una de las variables que intervienen o determinan la dinámica del sistema, siendo más común contar con una sola variable observable. Sin embargo, se ha demostrado [Takens, 1981] que una variable observable $x(t)$ de un sistema dinámico y sus versiones atrasadas

$$x(n) = [x(n), x(n - \tau), \dots, x(n - (N - 1)\tau)] \quad (1.1)$$

se pueden usar para recrear una trayectoria que preserve las invariantes dinámicas del sistema original. Para la ecuación (1.1), N es el doble de la dimensión espacial a recrearse y τ es un atraso normalizado de la variable. Tanto el parámetro τ como N son factibles de calcularse a través de métodos analíticos como lo es la Información Mutua Promedio [Fraser y Swinney, 1986] y el método de Vecinos Falsos [Kennel et al, 1992]. Estos métodos se introducen en el capítulo 3.

La arquitectura de redes neuronales y procesos de señales usados para modelar sistemas nerviosos se puede clasificar de manera muy general en tres categorías, basadas en diferentes filosofías: Las *redes unidireccionales*, que transforman un conjunto de señales de entrada en un conjunto de señales de salida; la transformación entrada-salida deseada se determina mediante ajustes supervisados de los parámetros desde el exterior del sistema. En las *redes recurrentes*, la información de entrada define el estado inicial de actividad de un sistema recurrente, y después de algunas transiciones del sistema, el estado asintótico final se identifica como la salida del cálculo. En la tercera categoría, células vecinas en una red compiten en su actividad por medio de una interacción lateral mutua y se desarrollan, adaptándose como detectores específicos de diferentes patrones de la señal de entrada. Esta categoría es denominada *competitiva* o *auto-organizada* [Kohonen, 1987].

Los mapas auto-organizados tienen la propiedad especial de crear “representaciones internas” espacialmente organizadas, de diversas características de señales de entrada y sus abstracciones. Los mapas auto-organizados pueden aprender a detectar irregularidades y correlaciones en sus entradas y adaptar sus respuestas futuras de manera consecuente [Kohonen, 1989]. Esta característica resulta de gran valor cuando la intención es aprender el comportamiento de sistemas que presentan una gran riqueza en cuanto a comportamientos, como lo son los sistemas caóticos. En una escala de complejidad, cuando el concepto de complejo está asociado a no linealidad u orden del sistema, los sistemas caóticos se encuentran entre las señales periódicas completamente definidas y el ruido aleatorio [Principe et al, 1998]. Como en toda red neuronal, es necesaria una etapa de entrenamiento, que normalmente se realiza con datos ya observados del sistema, lo que permite a la red aprender el comportamiento del sistema. El mapa auto-organizado en su aprendizaje, genera asociaciones entre los regímenes locales del sistema y cada una de sus neuronas. El número de asociaciones generado depende entonces del número de neuronas con que inicialmente fue creada la red. Cada asociación régimen-neurona implica la creación de un modelo local del sistema, lo suficientemente simple para ser tratado como un modelo lineal. De aquí se desprende el hecho de que el esquema de predicción utilice un paradigma de multi-modelos. La Figura 1 ilustra lo anterior. El capítulo 5 hace una revisión del proceso de competencia y aprendizaje de los mapas auto-organizados y ejemplifica su uso mediante problemas sencillos de clasificación.

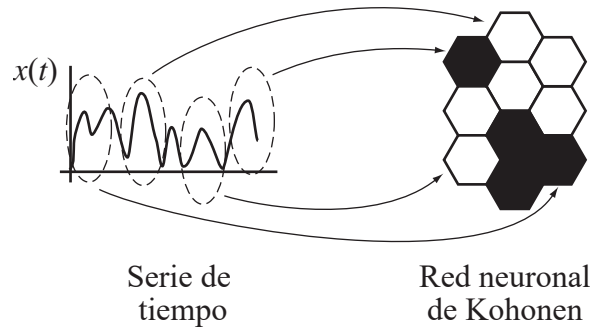


Figura 1. Relación entre regímenes de comportamiento de la serie y la red neuronal.

Finalmente, una vez realizada la clasificación de los regímenes de comportamiento del sistema en el entrenamiento de la red, en el paso anterior, es necesario realizar una reconstrucción del sistema. Esto se lleva a cabo mediante el uso del método de mínimos cuadrados ponderados, los cuales se revisan en el capítulo 4. La ventaja que representa el uso de ponderaciones reside en el hecho de que el paradigma de multi-modelos es discontinuo [McNames, 1999], y el efecto de los pesos es el de suavizar las transiciones de un modelo a otro. De esta manera, el comportamiento global del sistema es aproximado mediante un conjunto de modelos locales lineales. La Figura 2 ilustra este concepto.

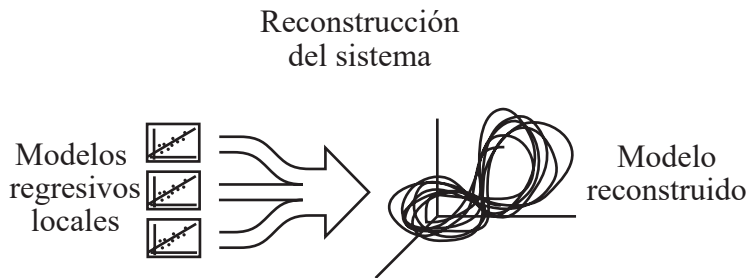


Figura 2. Esquema de reconstrucción del modelo global del sistema.

En base a cada una de las etapas previamente descritas, es posible formar un diagrama que ilustre el proceso completo de predicción, como lo muestra la Figura 3.

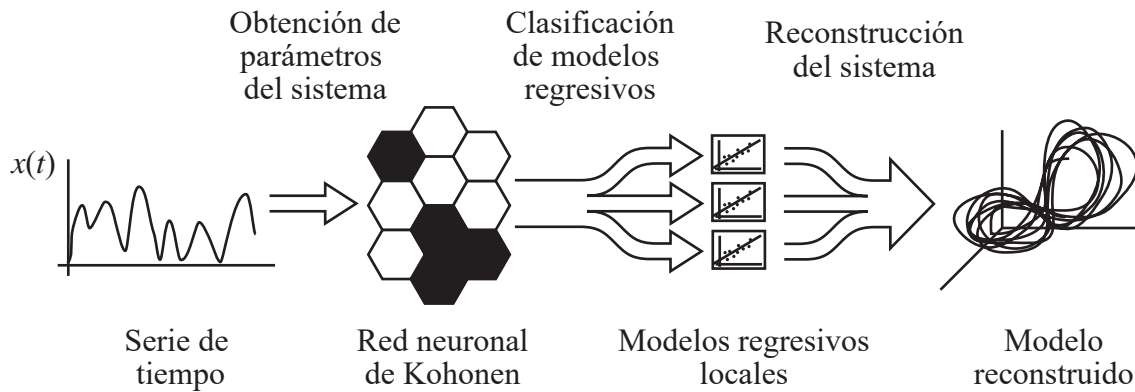


Figura 3 Diagrama del esquema de predicción.

En el proceso predictivo, una vez que se inicia la predicción, es posible realizar dos tipos de predicción:

- Predicción de un solo paso.
- Predicción iterada.

El primer tipo de predicción consiste en introducir un vector de datos ya observado de la serie de tiempo del sistema, denominado *interrogante*, y predecir el siguiente dato que aún no se ha observado. Esto nos permite estar un paso adelante en la respuesta del sistema. A primera vista, este resultado parece poco importante, pero puede significar anticipar, por ejemplo, la demanda del sistema eléctrico en una hora, si la serie de tiempo es una serie de lecturas horarias.

La predicción iterada, al igual que la anterior, requiere de un vector de datos observados de la serie de tiempo del sistema, pero el sistema puede predecir un número indefinido de datos posteriores. Cabe señalar que este método es menos preciso que el anterior, dada la característica de divergencia de todo sistema caótico.

1.4. Descripción de la tesis

Se describen a continuación, de manera breve, los conceptos y temas a abordarse en cada uno de los siguientes capítulos que conforman la tesis.

El capítulo 1 presenta una introducción general de la tesis, así como su objetivo, contenido y metodología. Dentro de esta última se reseña de manera breve los pasos para alcanzar el objetivo.

En el capítulo 2 se presentan varios modelos de sistemas caóticos clásicos, sus propiedades y su respuesta en el tiempo. Se hace una breve reseña de la forma en que surgen en la literatura, y aquellos valores paramétricos para los cuales presentan un claro comportamiento caótico. Se muestran gráficos de series de tiempo típicas de cada sistema y de sus atractores proyectados en dos dimensiones y aquellos tridimensionales donde es aplicable.

En el capítulo 3 se describen algunas herramientas básicas para el análisis de sistemas caóticos, así como algunas características que son esenciales para que un sistema pueda ser catalogado como tal. Se presentan conceptos típicos en la literatura de sistemas dinámicos no lineales: *atractores*, *bifurcaciones*, etc. y métodos de análisis comunes, como lo son los *Mapas de Poincaré*, con la finalidad de simplificar la introducción al campo.

El capítulo 4 introduce el método de *mínimos cuadrados ponderados*, y su uso en la predicción de series de tiempo. Se muestran las ecuaciones que definen tal método, y se presentan ejemplos de predicción usando esta metodología.

El capítulo 5 describe los Mapas auto-organizados de Kohonen, red neuronal no supervisada, caracterizada por su capacidad de representación gráfica, la cual se emplea en el método de predicción como la base de clasificación de los modelos locales o multi-modelos. Se detalla el principio de funcionamiento de la red y se ejemplifica su capacidad de clasificación.

Los casos de estudio realizados se muestran en el capítulo 6. Se hace uso de diversos sistemas bien conocidos como lo son los sistemas de *Lorenz*, *Chen*, *Chua*, etc. para ejemplificar el proceso de predicción. Se describe así mismo la teoría detrás de los modelos locales y de predicción de series de tiempo caóticos.

1.5. Contribución

La principal contribución de esta tesis es establecer un método de predicción de sistemas dinámicos no lineales en régimen caótico. El esquema que se presenta es una metodología novedosa, fundamentada en el uso de modelos locales lineales o multi-modelos lineales, los cuales se crean a partir de la clasificación de una red neuronal de Kohonen o mapa auto-organizado.

2. Respuesta en el tiempo de sistemas caóticos clásicos

2.1. *Introducción*

En este capítulo se presentan algunos de los sistemas que han significado la base del estudio de sistemas caóticos, desde sus inicios, con los estudios meteorológicos de Lorenz, hasta los sistemas caóticos “artificiales” usando redes neuronales, como lo son los circuitos de Chen y Chua [Ueta y Chen, 2000].

El análisis de la dinámica no lineal se aborda de una manera más intuitiva desde el punto de vista gráfico [Strogatz, 2000]. Es por ello que este capítulo, y en general, en toda la tesis, se presentan un gran número de gráficos, tanto de las series de tiempo, como de los atractores de cada sistema. Estos atractores, aparte de proporcionar una gran cantidad de información acerca del comportamiento del sistema, conllevan una estética visual inherente, la misma de la gran mayoría de los procesos naturales, como por ejemplo, la ramificación de un árbol, o la geometría de un copo de nieve.

2.2. Sistema de Lorenz

En 1963, Edward N. Lorenz publicó en la *New York Academy of Sciences* un artículo en el cual describía el comportamiento de la solución del sistema de ecuaciones diferenciales siguiente [Lorenz, 1963]:

$$\begin{aligned}\dot{x} &= \sigma(x - y) \\ \dot{y} &= rx - y - xz \\ \dot{z} &= xy - bz\end{aligned}\tag{2.1}$$

donde x , y y z son las variables de estado y σ , r y b , son los parámetros del sistema. σ es conocido como *número de Prandtl*, y r es el *número de Rayleigh*. Este conjunto de ecuaciones es la simplificación de un sistema que describe el flujo en dos dimensiones de un fluido en una caja rectangular, la cual es calentada en su parte inferior. Inicialmente era empleado para simular el efecto de la convección atmosférica a mediana escala. Este sistema tiende a presentar una divergencia exponencial ante cambios mínimos en su condición inicial, interpretándose esto como evidencia de la dificultad fundamental en la predicción de la evolución de flujos en régimen turbulento [Strogatz, 2000]. Lorenz descubrió que aún cuando se presentaba una respuesta dinámica extremadamente errática, siempre estaba dentro de una región del *espacio de fase*. El espacio de fase es un concepto interesante, empleado por connotados científicos como Poincaré (1881), en su tratado de solución de ecuaciones diferenciales, Gibbs (1902), en sus desarrollos de mecánica estadística y Birkhoff (1927), en su tratado de sistemas dinámicos. La definición de espacio de fase es la siguiente:

Se denomina espacio de fase a un espacio euclidiano M -dimensional, que se denota por Γ , cuyas coordenadas son X_1, X_2, \dots, X_M , y cada punto en él representa un posible estado instantáneo de algún sistema.

El espacio de fase es una generalización en M dimensiones de un sistema, y es común denominar como *plano de fase* a aquellos espacios de dos dimensiones. Graficando cualquiera de las variables de estado en el dominio del tiempo, con valores paramétricos de $\sigma = 10$, $b = 8/3$ y $r = 28$, se obtiene una gráfica para cualquiera de las variables como la que muestra la Figura 4.

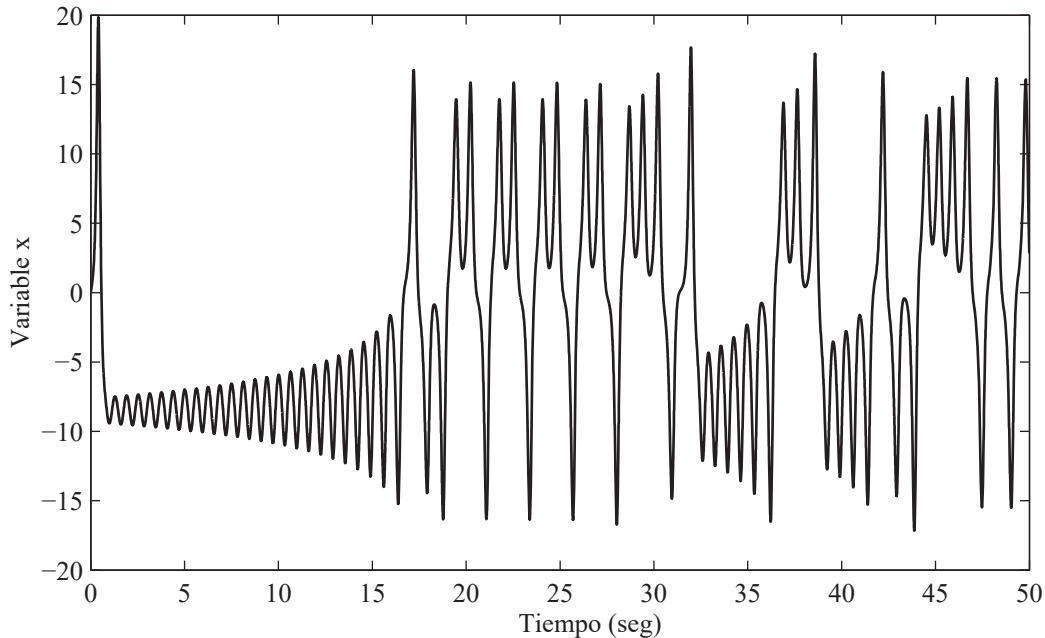


Figura 4. Gráfico de la variable $x(t)$ de la serie de tiempo de Lorenz.

Graficando el estado del sistema en el plano de fase es posible visualizar la interesante estructura llamada *Atractor de Lorenz*, la cual dio paso a la conocida frase de *Efecto Mariposa*¹, por la forma de alas que presenta.

La trayectoria que se crea conforme el sistema evoluciona, pareciera que se entrecruza repetidamente, pero no es más que el efecto de proyectar las tres dimensiones en un plano. En el espacio tridimensional nunca ocurre un cruce de la trayectoria, como lo demuestra el *teorema de unicidad*. La trayectoria en el plano de fase para este sistema se muestra en la Figura 5.

¹ De acuerdo al título de la conferencia que Lorenz dictó en 1972 en la Reunión de la Asociación americana para el avance de las ciencias. (*American Association for the Advancement of Science*) y que llevó como título “Predictibilidad: ¿Puede el aleteo de una mariposa en Brasil, provocar un tornado en Texas? (*Predictability: Does the Flap of a Butterfly’s Wings in Brazil set off a Tornado in Texas?*)

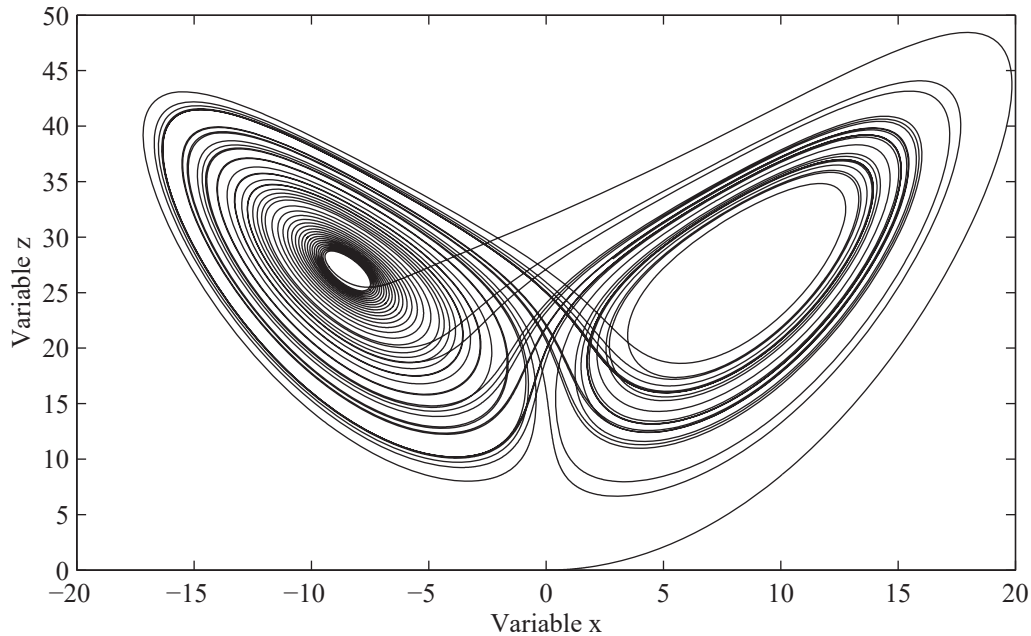


Figura 5. Plano de fase del Sistema de Lorenz, mostrando la trayectoria de x vs. z y la estructura subyacente, conocida como *Mariposa*.

A primera vista, la trayectoria parece que está establecida sobre una superficie definida por las dos espirales, pero conforme se realiza un acercamiento sobre ellas, es posible darse cuenta de la “rugosidad” de las mismas, de tal manera que es realmente más que una superficie la que ocupa la trayectoria. Tal superficie se dice, tiene un *incremento* en su dimensión, sin llegar a ser tridimensional, con lo cual se puede definir entonces como un *fractal* [Mandelbrot, 1982]. Es posible definir un fractal como el conjunto de puntos con volumen cero y área superficial infinita. Una gráfica tridimensional de la trayectoria del sistema se muestra en la Figura 6.

Cabe destacar que Lorenz acuñó el término *efecto mariposa* para describir la enorme sensibilidad de un sistema caótico a las condiciones iniciales, es decir, un cambio muy sutil de tales condiciones, generará un comportamiento completamente distinto posteriormente. En el caso del clima, que estudiaba Lorenz, el aleteo de la mariposa es el cambio de condiciones iniciales, que al paso de los días puede provocar comportamientos climáticos extremos.

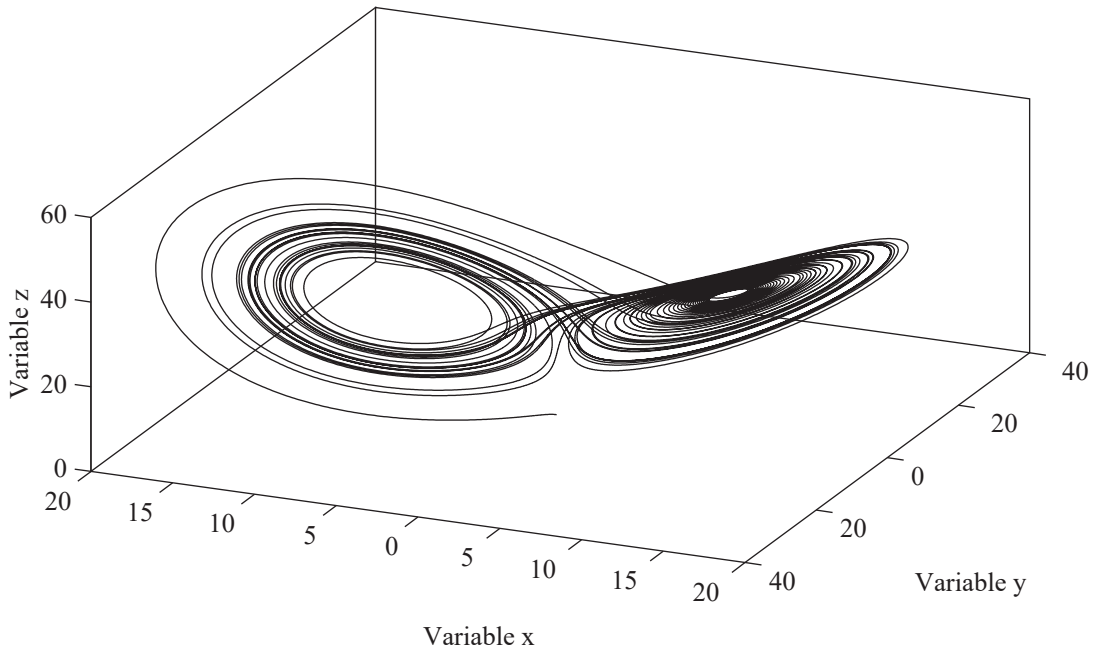


Figura 6. Gráfico tridimensional de la trayectoria del sistema de Lorenz.

2.3. Sistema de Chua

El sistema de Chua o de doble espiral, es un atractor caótico bien conocido que puede observarse como salida del circuito de Chua [Chua et al, 1986], un circuito eléctrico autónomo simple, compuesto de cuatro elementos lineales y uno no lineal. La dinámica del circuito está descrita por el siguiente conjunto de ecuaciones:

$$\begin{aligned}
 \dot{x} &= \alpha(y - m_0x - \frac{1}{3}m_1x^3) \\
 \dot{y} &= x - y + z \\
 \dot{z} &= -\beta y
 \end{aligned}
 \tag{2.2}$$

el comportamiento del sistema está gobernado por los cuatro parámetros α , β , m_0 y m_1 , y es el circuito autónomo más simple capaz de presentar un comportamiento caótico.

El elemento no lineal del sistema es un resistor *lineal a segmentos* de tres regiones, conocido como *diodo de Chua*, y que presenta una simetría respecto al origen. La Figura 7 muestra el circuito de Chua y la Figura 8 muestra la gráfica de función de transferencia para el resistor no lineal R_N del circuito.

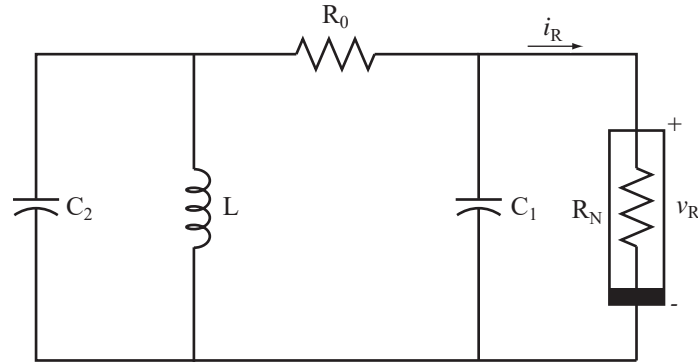


Figura 7. Circuito elemental de Chua.

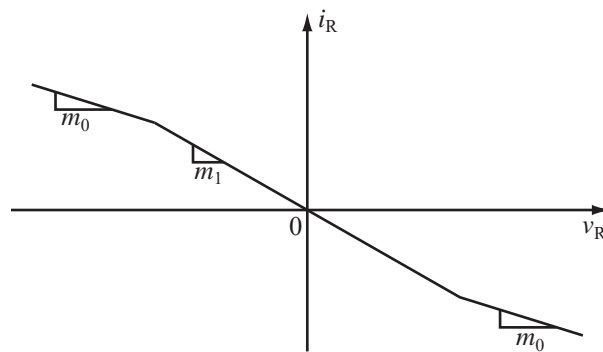


Figura 8. Función de transferencia del resistor no lineal del circuito de Chua .

Como puede verse en la Figura 8, los parámetros m_0 y m_1 de (2.2) son las pendientes que presentan los segmentos lineales de la función de transferencia del resistor R_N . Esta característica le proporciona un comportamiento topológicamente equivalente al de un sistema de ecuaciones no lineales. Por este motivo y por el hecho de que el sistema presenta diferentes fenómenos no lineales, el circuito presenta una gran variedad de atractores extraños estructuralmente inestables. Es común que en serie con el inductor L se conecte otro resistor, lo cuál confiere al circuito una mayor diversidad de comportamientos. Entre los comportamientos interesantes del sistema se puede citar el denominado *nacimiento de la doble espiral*, proceso que consiste en el paso de una dinámica asimétrica del sistema a una simétrica. La Figura 9 muestra este proceso, el cual es relativamente simple. En (a) el sistema se comporta como un oscilador simple, pero una vez que se alcanzan ciertos valores críticos, la trayectoria de estado del sistema se altera completamente (b),

estableciéndose finalmente en otra espiral alejada de la primera (c). El comportamiento estable del sistema oscila repitiendo estos saltos indefinidamente.

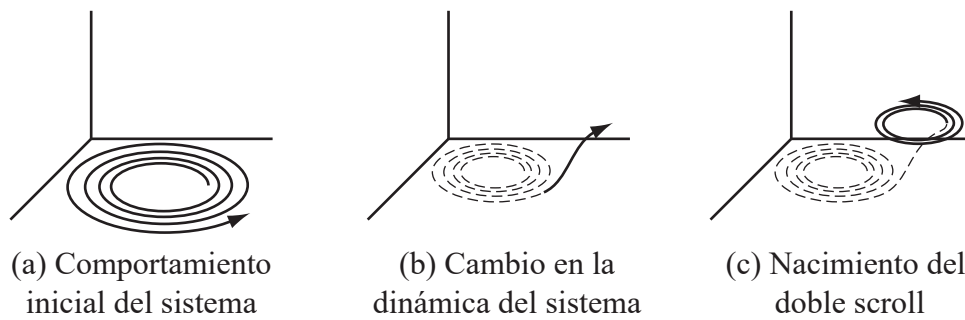


Figura 9. Nacimiento de la doble espiral.

La Figura 10 muestra una proyección en el plano xy del atractor en estado caótico y con la doble espiral claramente definida.

Los valores paramétricos empleados para obtener esta representación son: $\alpha = 9$, $\beta = 14.286$, $m_0 = 2/7$, $m_1 = -1/7$.

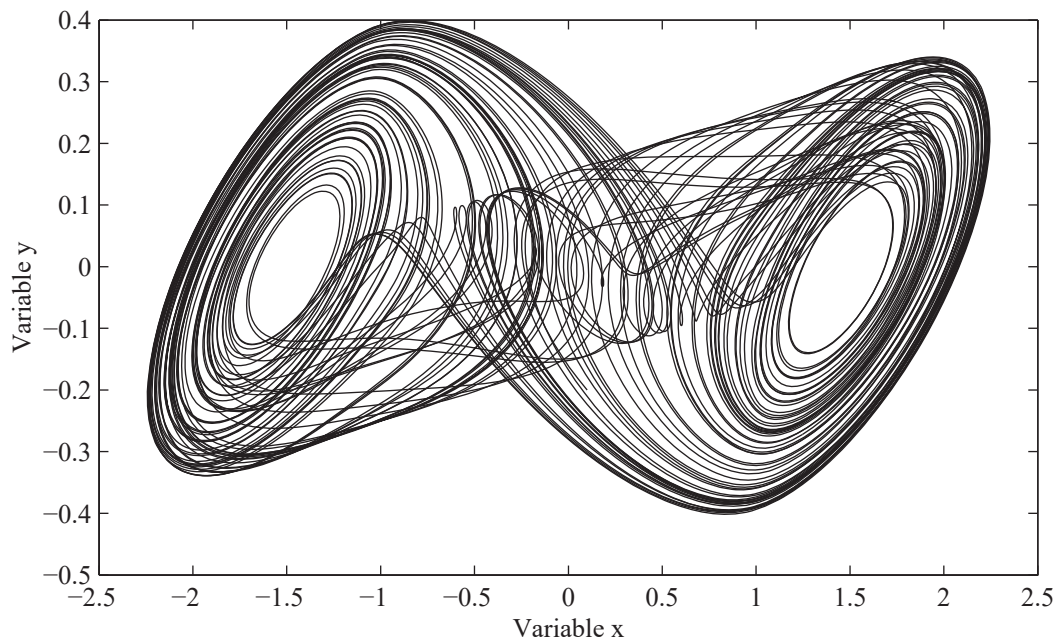


Figura 10 Atractor de Chua.

La Figura 11 muestra el mismo atractor en el espacio de tres dimensiones.

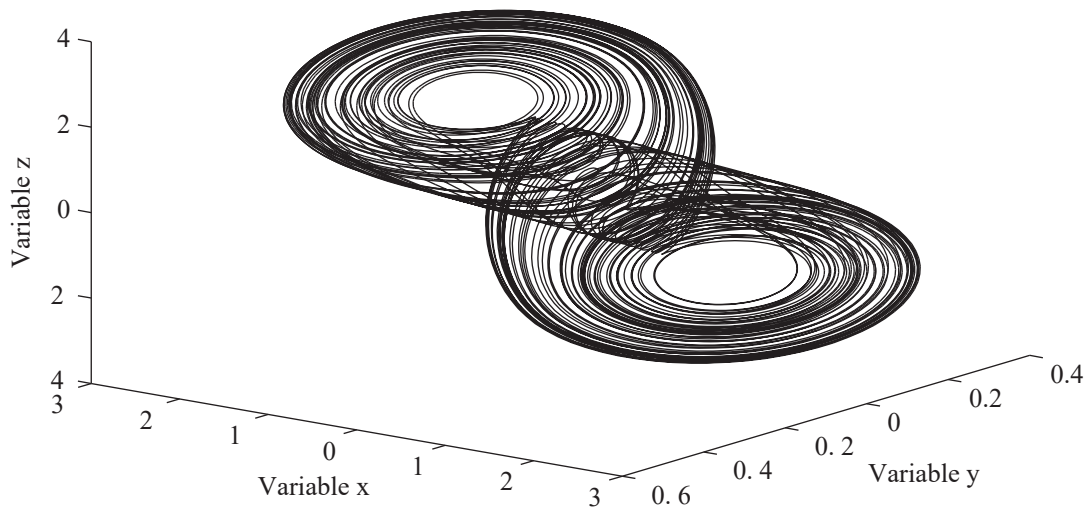


Figura 11. Atractor de Chua visto en tres dimensiones.

La Figura 12 muestra como cada espiral se desarrolla en un semiplano.

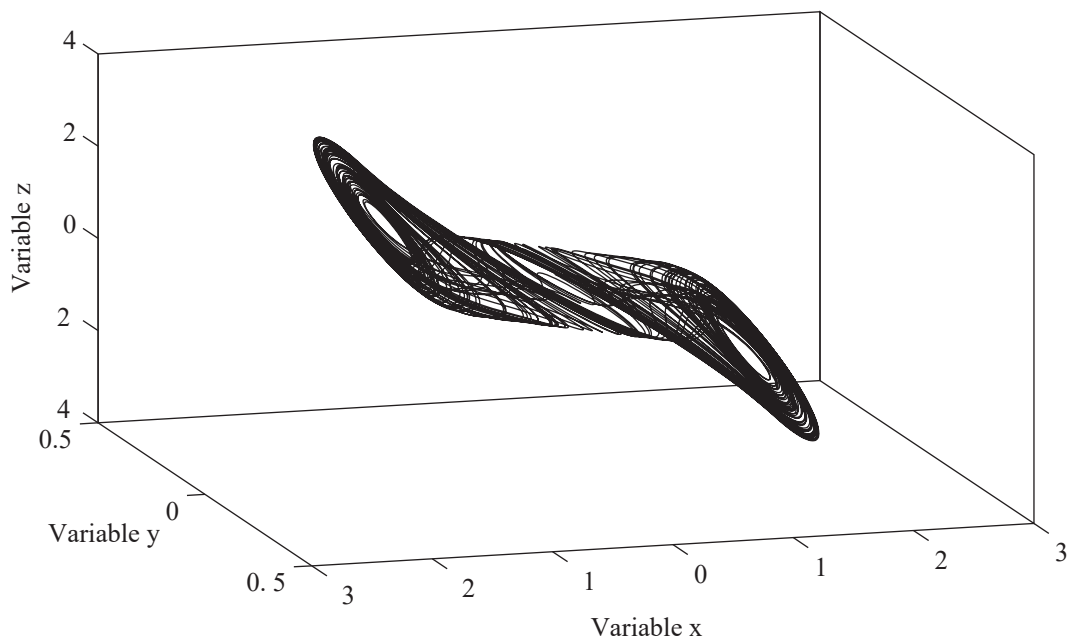


Figura 12. Atractor de Chua que muestra los semiplanos en que se desarrollan las espirales.

El circuito puede ser alterado de tal manera que es capaz de producir atractores multi-esprial, ya sea de número par o impar, mediante la modificación de los segmentos lineales del resistor R_N [Aziz Alaoui, 2000].

2.4. Sistema de Chen

El modelo matemático del atractor caótico de Chen está descrito por el siguiente conjunto de ecuaciones de estado:

$$\begin{aligned}\dot{x} &= a(y - x) \\ \dot{y} &= (c - a)x - xz + cy \\ \dot{z} &= xy - bz\end{aligned}\tag{2.3}$$

este sistema inicialmente recuerda al de Lorenz, dado que presenta el mismo grado de complejidad y dos términos no lineales, pero se ha demostrado que no son equivalentes topológicamente.

Presentado en [Ueta y Chen, 2000], el principal interés para este sistema era la de anti-control de caos, o *caotificación*, que consiste en generar comportamientos caóticos en sistemas que originalmente son no-caóticos o incluso estables, mediante controles simples e implementables. El sistema de ecuaciones de Chen surge del anti-control del sistema de Lorenz, de ahí su similitud.

La Figura 13 muestra la proyección bidimensional del atractor. En ella se puede ver la simetría con respecto a $x=0$, que es la misma que presenta el sistema de Lorenz, pero una diferencia es el hecho que este sistema presenta una bifurcación adicional hacia una trayectoria fuera de la simetría y de los semiplanos en los cuales residen las espirales, creando una estructura en forma de semicono por encima de ellos.

Nuevamente, la proyección en un plano crea la ilusión que las trayectorias se entrecruzan, pero en la proyección tridimensional de la Figura 14, puede verse que la trayectoria está fuera de los semi-planos que forman las espirales simétricas. Los parámetros empleados para obtener estas imágenes son: $a=38$, $b=3$ y $c=28$.

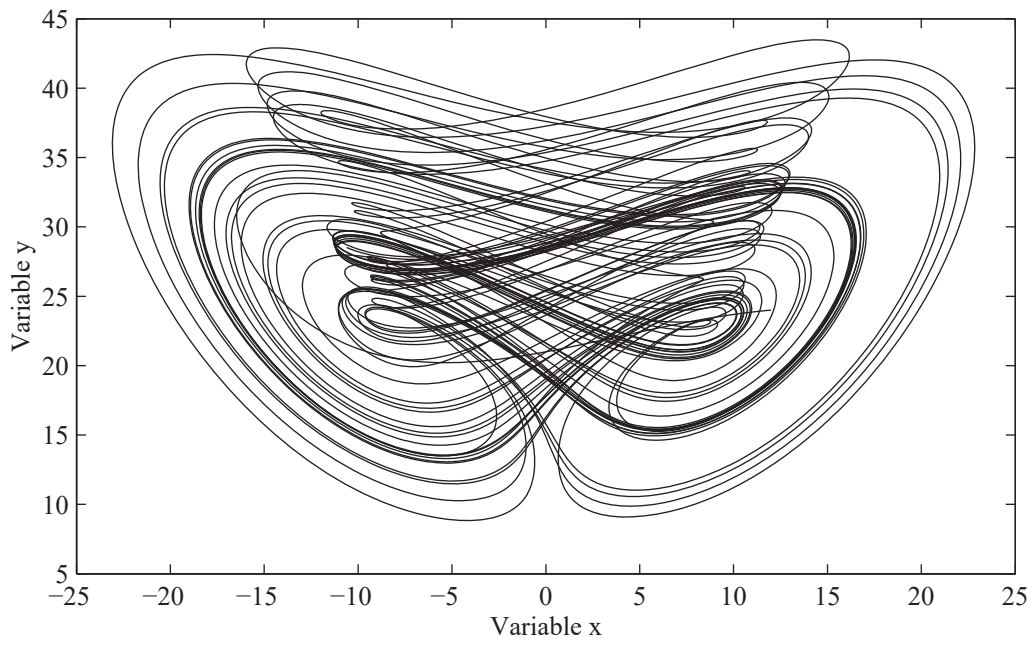


Figura 13. Proyección en el plano xy del atractor de Chen.

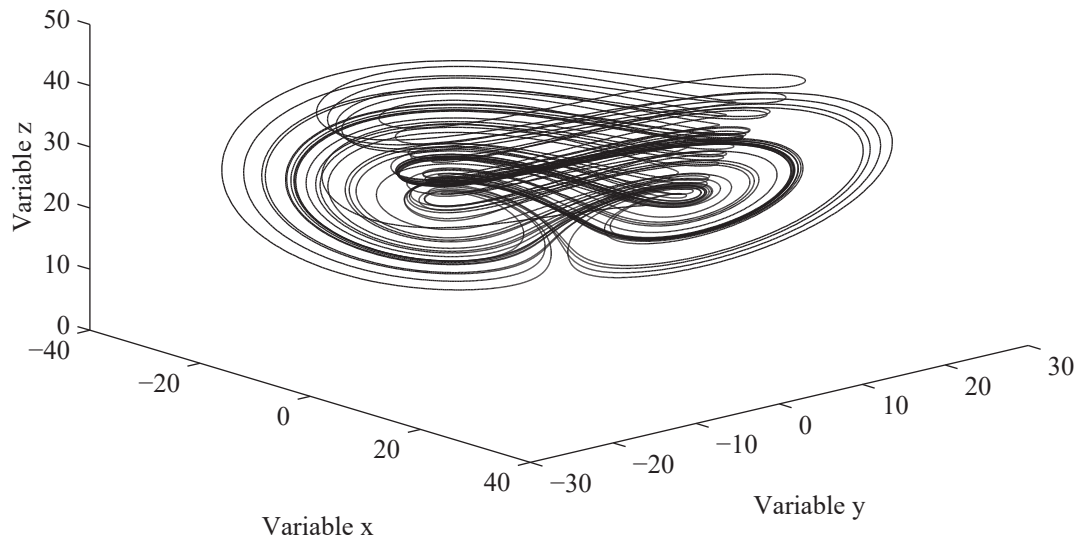


Figura 14. Vista en tres dimensiones del atractor de Chen.

2.5. Sistema de Rössler

El sistema de Rössler es probablemente el sistema tridimensional más simple que presenta comportamiento caótico [Rössler, 1976]. Está definido por el siguiente conjunto de ecuaciones diferenciales acopladas:

$$\begin{aligned}\dot{x} &= -(z + y) \\ \dot{y} &= x + ay \\ \dot{z} &= b + xz - cz\end{aligned}\tag{2.4}$$

el único factor no lineal del sistema es el término xz en la última ecuación. Este es un sistema artificial creado para explicar la cinética de reacciones químicas. El conjunto de ecuaciones resulta bastante simple, pero aún así, el comportamiento de sus soluciones presenta dinámicas de gran riqueza.

Es común fijar los parámetros a y b , dejando a c como parámetro de control con la finalidad de observar las bifurcaciones que suceden al sistema conforme c aumenta, hasta llegar a un comportamiento caótico.

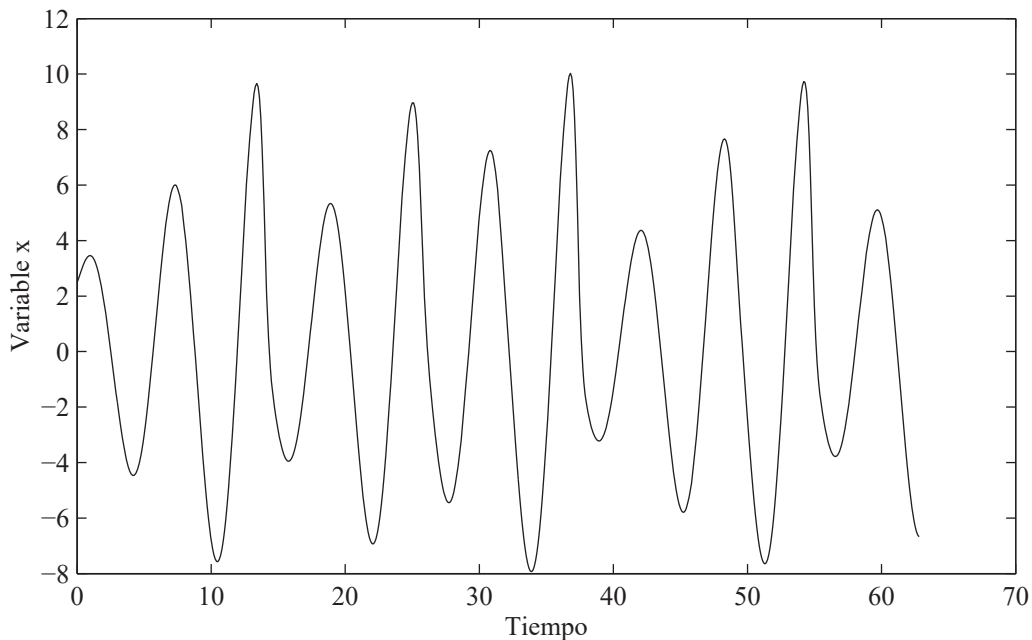


Figura 15. Serie de tiempo para la variable x del Sistema de Rössler.

La Figura 15 ilustra una serie de tiempo para la variable x del sistema, y muestra un comportamiento oscilatorio relativamente simple, de tal manera que para valores de $a = b = 0.2$ y para $c=2.5$, el atractor del sistema es un simple ciclo límite muy similar al mostrado por un péndulo. Conforme c aumenta, el sistema pasa a ser de doble periodo, lo cual obliga a que, con la finalidad de que la trayectoria del sistema no se interseque consigo misma, el atractor deba ser un atractor tridimensional.

Este sistema muestra de una manera sencilla el proceso de *expansión, doblamiento y re-inyección* de las trayectorias que todo sistema caótico requiere a fin de que las trayectorias puedan divergir y sin embargo permanecer acotadas.

La Figura 16 muestra el atractor del sistema cuando c ha sobrepasado el valor de 5. Es posible observar que el sistema se mantiene parcialmente en una trayectoria bidimensional, mediante una espiral creciente (*expansión*), que Rössler denominó *espiral caótica*, hasta alcanzar un punto crítico, en el cual la trayectoria salta hacia la tercera dimensión (*doblamiento*), evitando así el cruce de trayectorias y reingresando nuevamente en el plano bidimensional muy cerca de su punto de arranque inicial (*re-inyección*).

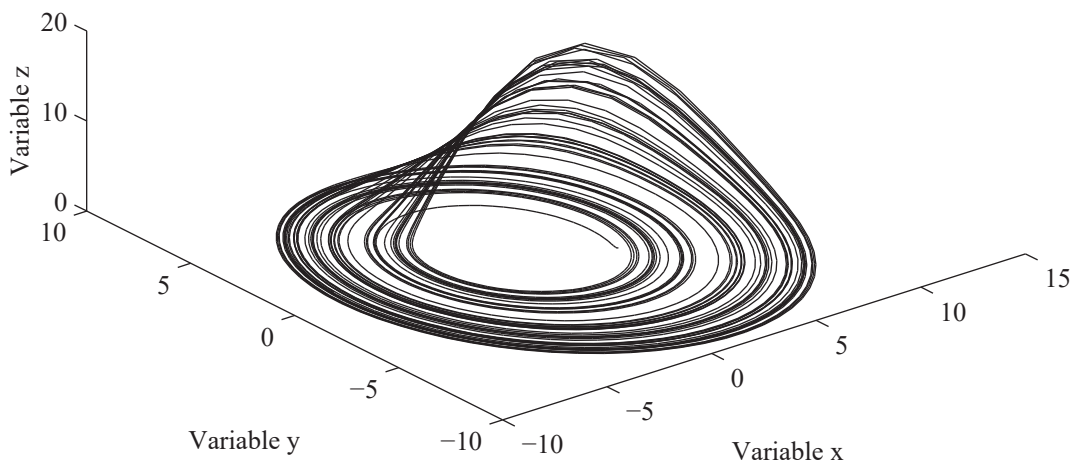


Figura 16. Atractor extraño del Sistema de Rössler.

La Figura 17 muestra de manera esquemática este proceso para su mejor comprensión.

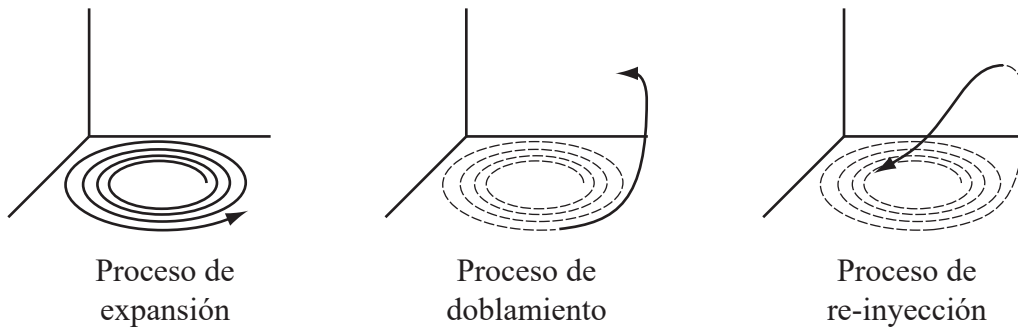


Figura 17. Proceso de expansión, doblamiento y re-inyección de un sistema caótico.

Posteriormente a su descubrimiento inicial, Rössler descubrió dinámicas caóticas más complejas provocadas por la interacción de diversos puntos de bifurcación del sistema, a la cual llamó *hélice de caos* [Rössler, 1979].

2.6. Conclusiones

En este capítulo se introdujeron varios de los sistemas caóticos más comunes en la literatura: Lorenz, Chen, Chua y Rössler, las ecuaciones que los definen, valores paramétricos para los cuales presentan comportamientos caóticos y gráficos característicos de sus atractores y series de tiempo.

Se usaron los sistemas para introducir conceptos propios de los estudios de la dinámica caótica, como son *espacio de fase*, el proceso de *expansión*, *doblamiento*, y *re-inyección* de una trayectoria, aparición de múltiples scrolls, etc. Basados en estos conceptos, se está en posición de comprender la base de la predicción, formada por el teorema de Takens, así como de los algoritmos de vecinos falsos e información mutua promedio los cuales se presentarán de una manera más formal en el capítulo 3.

3. Características de los sistemas no lineales.

3.1. Introducción

En este capítulo se describen algunos mapas y procedimientos empleados para el análisis de sistemas caóticos. Se consideran únicamente aquellos que posteriormente se emplearán en el análisis de los sistemas bajo estudio como objetivo de la tesis. Entre los mapas útiles se pueden mencionar los *mapas iterados*, *mapas de Poincaré*, *gráficos de telaraña*, *mapas de bifurcación*, etc. y algoritmos de gran importancia en las técnicas de análisis de datos, tales como el *algoritmo de información mutua promedio* y el de *vecinos falsos*. Se utilizan sistemas típicos como ejemplo para ilustrar el uso de cada una de estas herramientas, y se describe de manera breve de dónde surgen y el análisis del mismo a la luz de los resultados que arrojan.

Los sistemas ejemplo, aunque simples, muestran los mismos comportamientos de aquellos sistemas de mayor tamaño y complejidad, como lo son, por ejemplo los sistemas eléctricos de potencia, y permiten un mejor entendimiento del análisis básico de los sistemas caóticos.

Cabe señalar que existe una gran cantidad de herramientas adicionales, pero que salen del objetivo de la tesis. Entre éstas se pueden mencionar los exponentes de Lyapunov [Eckmann y Ruelle, 1985], la dimensión de correlación [Grassberger y Procaccia, 1983], por citar algunas.

Una vez que se ha visto en el capítulo anterior la riqueza de comportamientos de los sistemas caóticos, se definen en este capítulo algunos de los términos empleados.

3.2. Mapas iterados

Los mapas iterados son la manera más simple de describir un sistema no lineal [University of Western Ontario, 2003]. Usados ya en el método de Newton para la determinación de raíces de polinomios a finales del siglo XVII, cayeron en desuso y fueron retomados hasta finales de siglo XIX y principios del siglo XX, para resolver problemas de convergencia de funciones, especialmente funciones racionales complejas [Wolfram, 2002]. Tales mapas se crean a partir de un número inicial cualquiera que se pasa como argumento a alguna función matemática, el resultado que se obtiene se pasa como argumento nuevamente de la misma función, repitiéndose tal proceso una infinidad de veces. Los valores obtenidos a lo largo del proceso forman una *solución*, *trayectoria* u *órbita* del sistema.

Un mapa iterado interesante lo constituye el mapa logístico o ecuación logística, conocido también como modelo de Verhulst [Quetelet y Verhulst, 1850], por Pierre-Francoise Verhulst, matemático belga, quien la dedujo en 1838. La ecuación logística es un modelo simplificado de tiempo continuo del crecimiento poblacional de especies sin depredadores y con una fuente alimenticia limitada. Está dado por la ecuación diferencial siguiente:

$$\frac{dN}{dt} = \frac{rN(K - N)}{K} \quad (3.1)$$

donde r es conocido como parámetro de Malthus, y representa la razón máxima de crecimiento poblacional, K es la constante conocida como capacidad de portación, es decir, la máxima población sostenible, en tanto que N es el número de individuos presentes.

Dividiendo ambos lados de la ecuación por K y definiendo $x = N/K$, nos lleva a la ecuación diferencial:

$$\frac{dx}{dt} = rx(1-x) \quad (3.2)$$

Con la llegada de las computadoras digitales la Ecuación (3.2) se puede sustituir por una ecuación de *recurrencia cuadrática*:

$$x_{n+1} = rx_n + rx_n^2 \quad (3.3)$$

la cual puede reexpresarse en la siguiente forma:

$$x_{n+1} = rx_n(1-x_n) \quad (3.4)$$

donde r , (algunas veces se emplea μ), es la capacidad natural de un organismo para sobrevivir y reproducirse y se conoce como potencial biótico. La gráfica de esta función es una parábola con su valor máximo de $r/4$, situado en $x = 0.5$ y se muestra en la Figura 18.

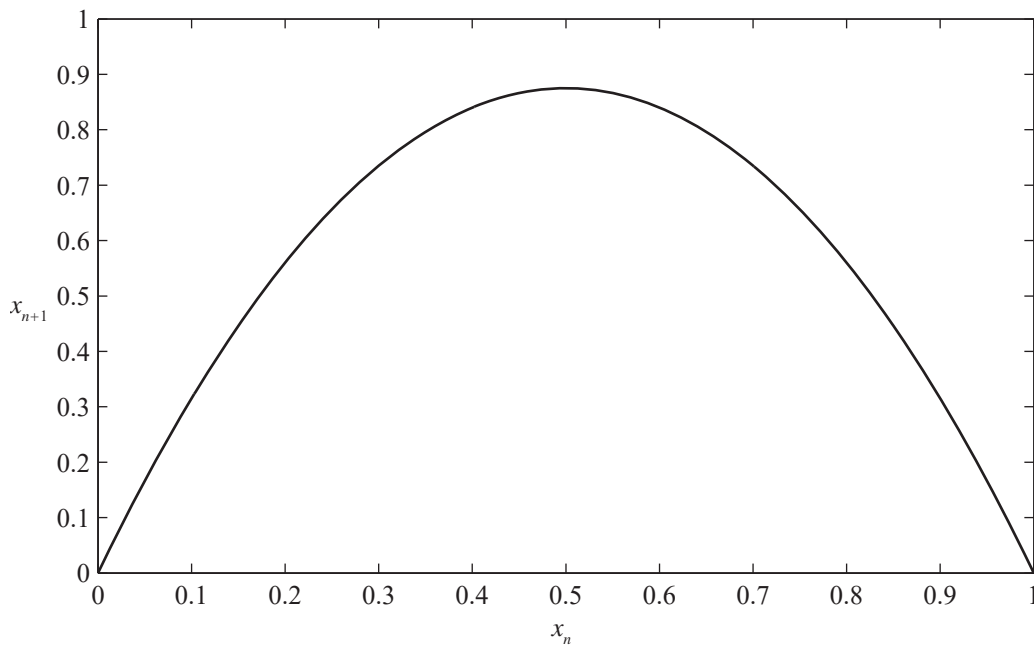


Figura 18. Gráfico de la ecuación logística.

Si se itera sobre la Ecuación (3.4) con valores de $r = 2.9$ y $x_0 = 0.01$, se obtiene la gráfica de la Figura 19. Como puede verse en esta figura, después de un tiempo el sistema tiende a establecerse en un valor determinado, denominado *atractor* del sistema, lo cual no es sino el estado hacia el cual el sistema evoluciona después de haber arrancado de un cierto valor inicial. Posteriormente se verá que este atractor no necesariamente es un valor único, sino que puede tomar un número infinito de valores, y sin embargo seguir siendo un atractor. En este caso en particular se trata de un atractor de *punto fijo*, puesto que el valor de estado estable es una constante. Los puntos fijos se denotan por x^* y deben satisfacer el criterio $x^* = f(x^*)$ o $x_{t+1} = x_t$.

Para el intervalo $0 < r < 1$, la iteración sobre el sistema siempre decae después de un tiempo. El significado de ésto es que un valor de máximo crecimiento poblacional demasiado pequeño condena a la población a extinguirse. Inmediatamente después de sobrepasar el valor de 1, el sistema tiende a estabilizarse en valores continuamente crecientes, hasta llegar a la vecindad de $r = 3.0$, pero siempre en un punto fijo estable. La Figura 19 muestra el comportamiento del sistema cuando $r = 2.9$.

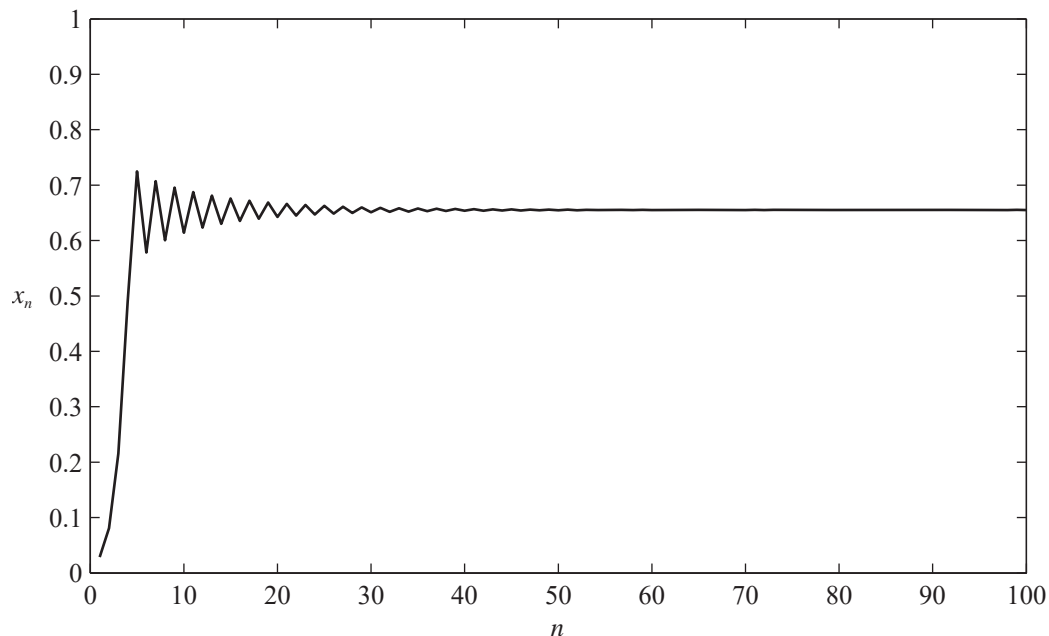


Figura 19. Gráfico iterado de la ecuación logística con $r = 2.9$.

El punto fijo de un sistema es un estado de equilibrio del mismo, y tal equilibrio puede ser estable o inestable. Una gráfica informativa tanto de los puntos fijos estables como de los *ciclos estables* del sistema conforme r varía se obtiene de la *gráfica de telaraña*, mostrada en la Figura 20.

Este tipo de gráficas muestra como los valores iterados varían con respecto al equilibrio, cuando $x_n = x_{n+1}$, lo cual se puede representar por una línea diagonal. Es claro ver que después de cada iteración, el sistema tiende a un punto sobre la curva, pasando por una etapa transitoria hasta su establecimiento, que será su respuesta final. Cada punto en que la línea h toca la parábola representa un punto máximo o mínimo local de la oscilación del sistema mostrado en la Figura 19, por ello, esta forma de visualización tiene la capacidad de mostrar dinámicas interesantes ocultas en el sistema, entre ellos puntos fijos o ciclos estables que de otra manera resultaría difícil encontrar.

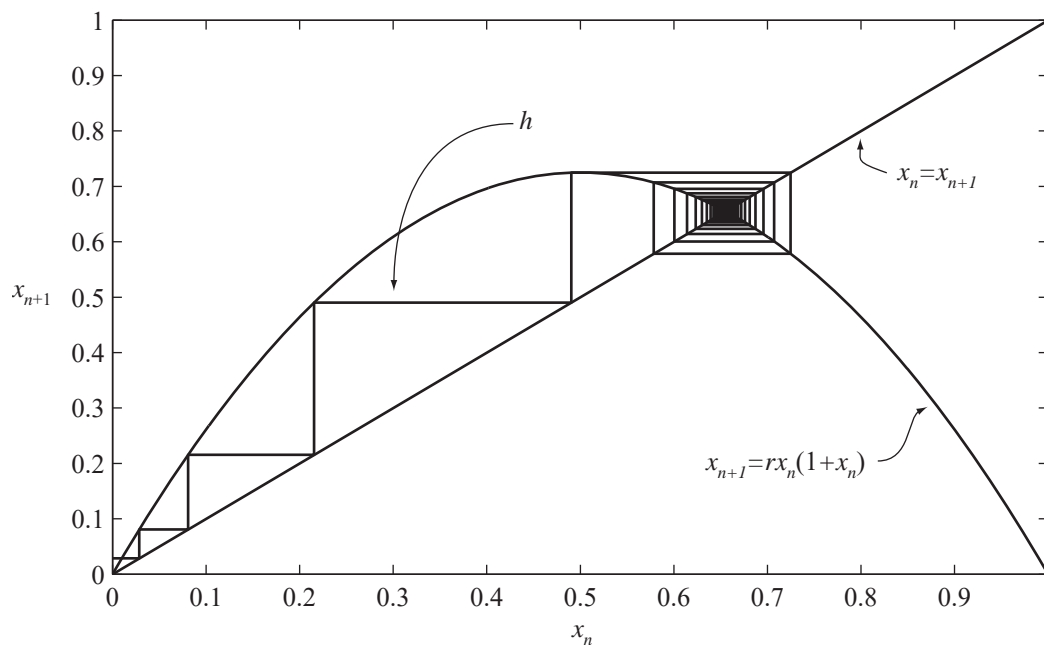


Figura 20. Gráfico de “telaraña” para la respuesta del sistema.

Es claro, de acuerdo a lo dicho en el párrafo anterior, que el valor $x_t = 0$ es un punto fijo del sistema, como lo muestra también el mapa iterado de la Figura 21. Desde el punto de vista poblacional, es posible decir que una población de cero individuos no producirá

descendencia. La Figura 21 es interesante además porque representa el único punto fijo del sistema para valores de $r < 1$, pero sus características cambian conforme r aumenta. Y cuando r alcanza el valor exacto de 1, entonces, todo valor de x_t es un punto fijo. Cuando r sobrepasa el valor de 1, $x^* = 0$ se vuelve un punto fijo inestable, ya que únicamente un valor inicial exacto de $x_0 = 0$ conduce a esta respuesta estable, y cualquier alteración, por pequeña que sea, repele la trayectoria del sistema, alejándose de este punto fijo. Es común que a tales puntos fijos se les denomine *repulsores*.

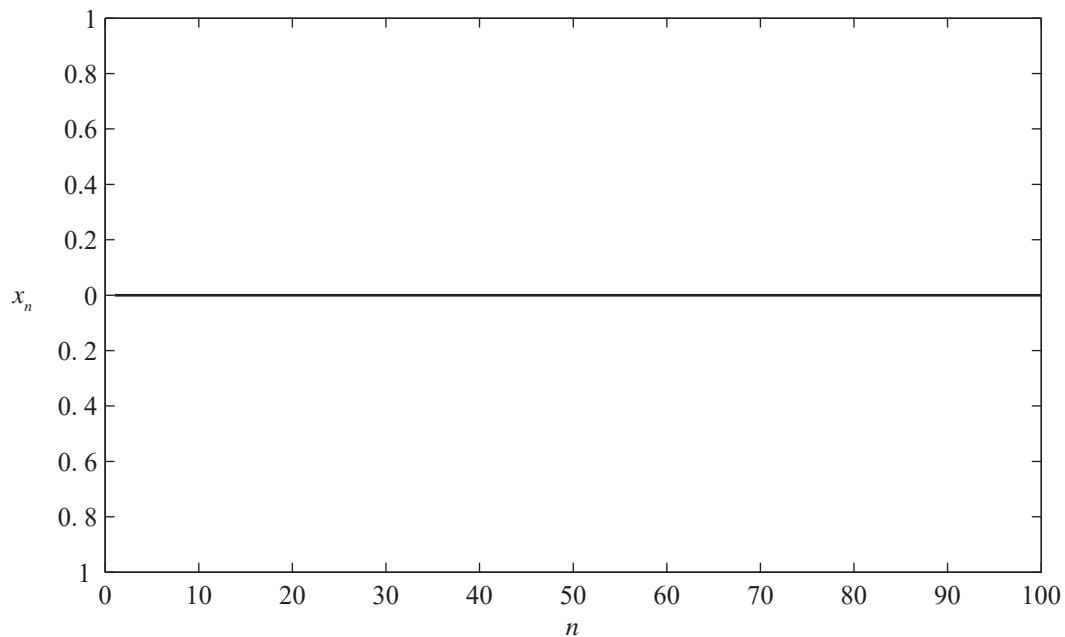


Figura 21. Gráfico de la ecuación logística para un valor de $r = 0.0$.

Cuando el valor de r alcanza o sobrepasa el valor de 3.0, el sistema entonces no tiende hacia un punto fijo, sino que se mantiene oscilando. La oscilación puede entenderse desde el punto de vista de una generación que tiene suficiente alimento, lo cual provocará un crecimiento en su número, pero en la siguiente generación el número de individuos sobrepasará el nivel de mantenimiento, produciendo un subsecuente decaimiento. Este es en sí un comportamiento de estado estable que puede verse en cultivos de bacterias. La Figura 22 muestra este comportamiento oscilatorio. En este tipo de oscilación, cuando x_n se repite cada dos iteraciones, se denomina ciclo de doble periodo. Al incrementar más el valor de r , el ciclo se vuelve un ciclo de cuádruple periodo, luego óctuple, y así

sucesivamente. Este cambio se denomina *bifurcación*. La Figura 23 muestra el ciclo óctuple para la ecuación logística.

Para observar todas las posibles bifurcaciones que surgen en el sistema sería necesaria una gráfica de telaraña para cada valor de r , pero este proceso se puede sintetizar mediante un *diagrama orbital*. En él se grafica el estado estable del sistema como una función de r . la Figura 24 muestra un diagrama orbital para la ecuación logística. Cuando r alcanza un valor aproximado de 3.57, el mapa muestra cómo el sistema se vuelve *caótico*, es decir, su periodo se vuelve infinito. Más allá de este valor, el comportamiento del sistema es una mezcla de caos y estabilidad, denominados *ventanas periódicas*.

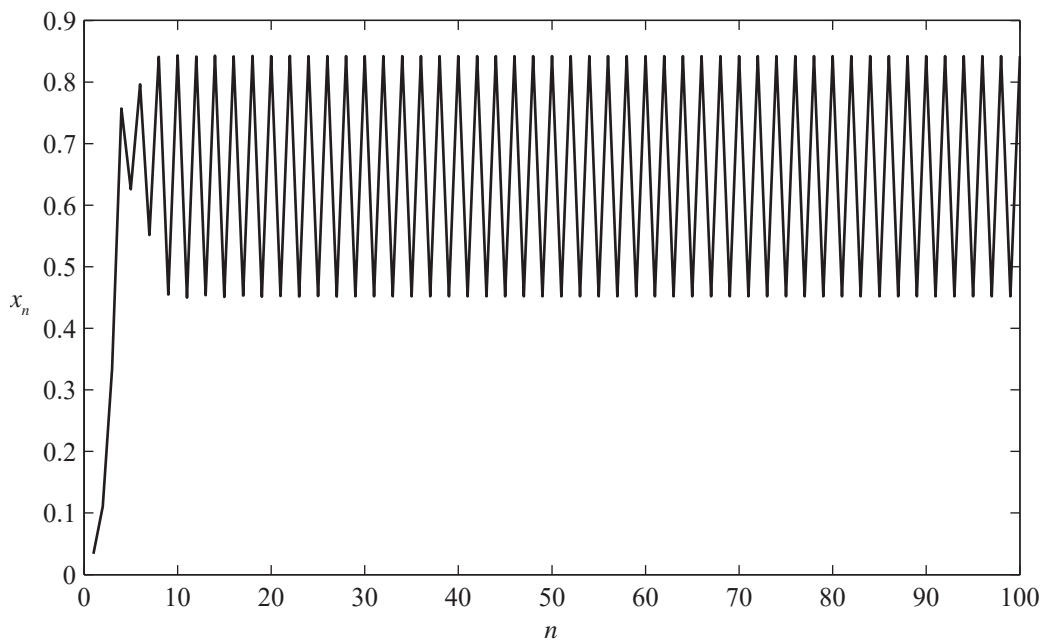


Figura 22. Respuesta oscilatoria del sistema cuando $r = 3.2$.

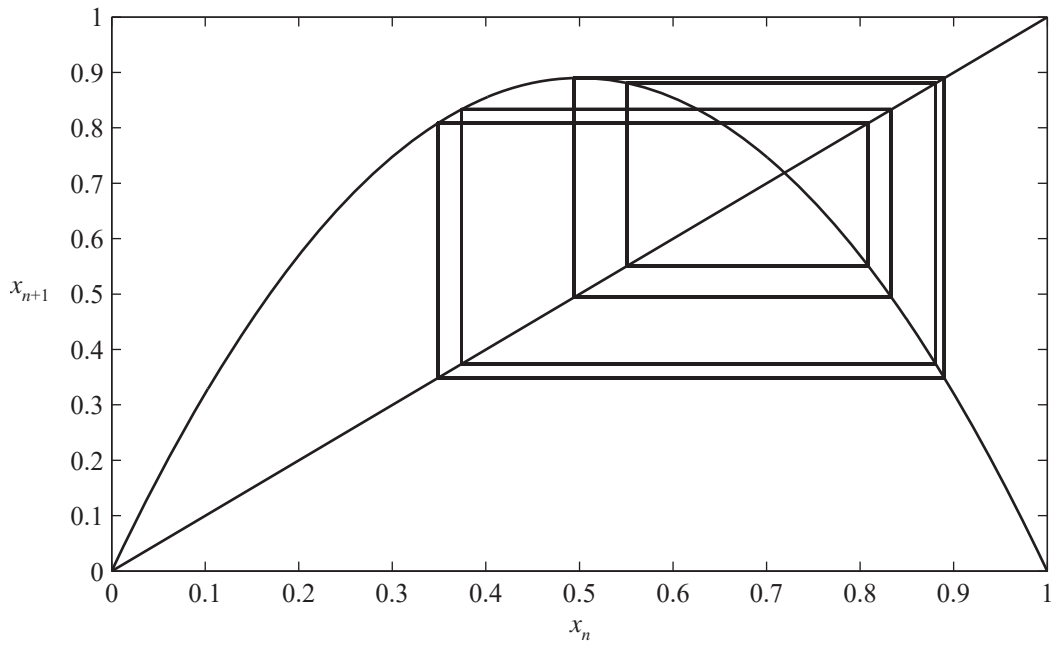


Figura 23. Gráfico de telaraña que muestra el ciclo de óctuple periodo del sistema en $r = 3.56$. No se muestra el transitorio inicial.

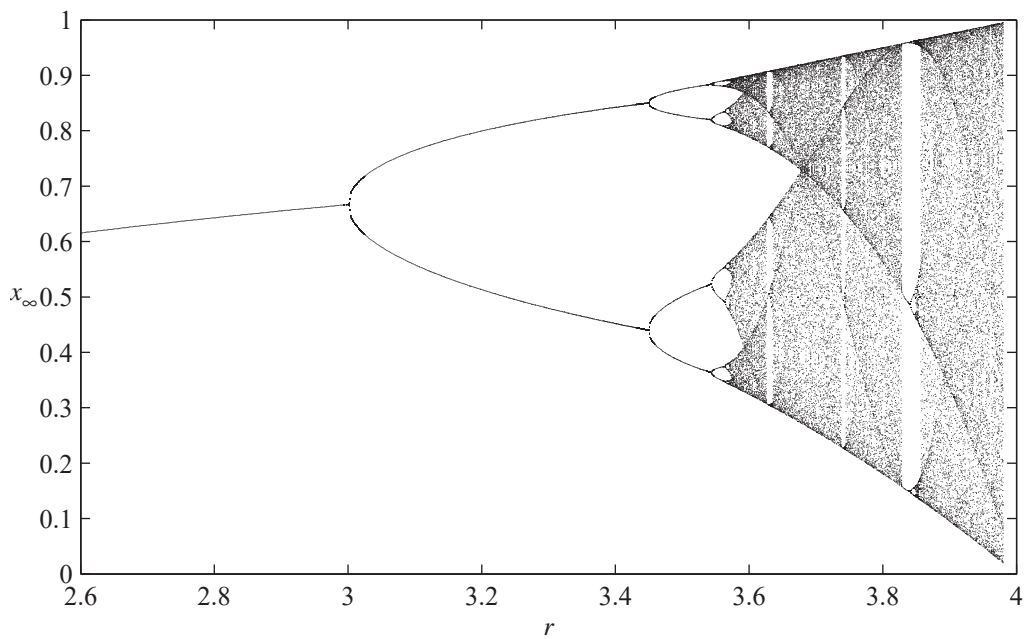


Figura 24. Diagrama orbital para la ecuación logística.

La Figura 25 muestra un gráfico de telaraña cuando el sistema ha entrado a un régimen caótico, cuando r alcanza valores superiores a 3.57.

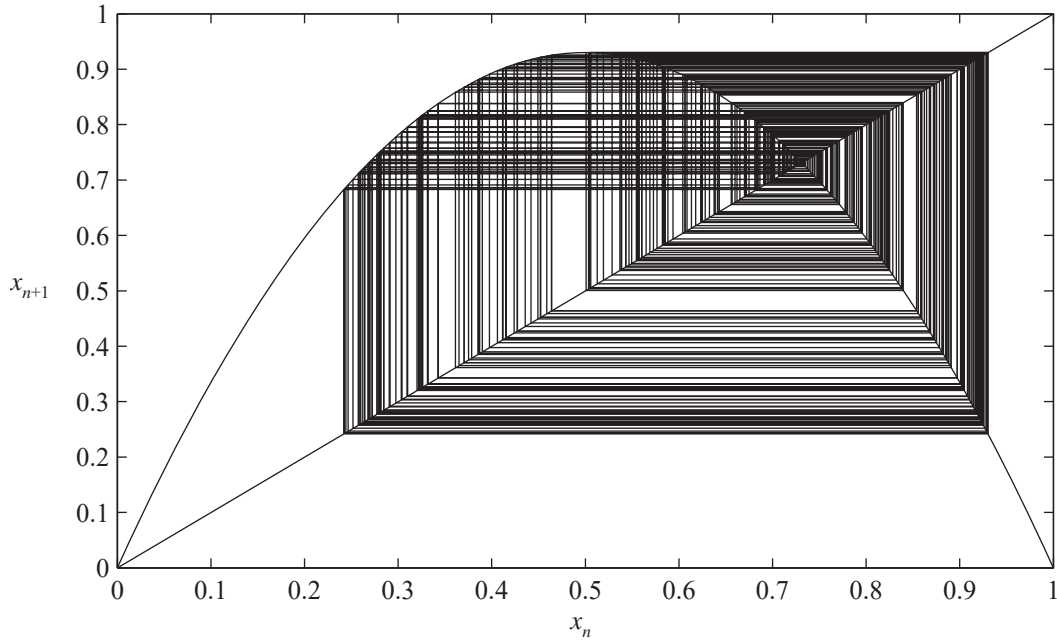


Figura 25. Gráfico de telaraña para $r = 3.7$. No se muestra el transitorio inicial.

Se tienen entonces diversas representaciones gráficas del sistema, y cada una permite un análisis distinto del sistema. Podemos concluir, por tanto, que su uso depende de la característica que se requiera visualizar, así, si es necesario determinar el valor estable al que tiende el sistema, se puede emplear el mapa iterado, pero si la intención es visualizar el régimen que presentará el sistema conforme varían sus parámetros, es preferible una gráfica de telaraña o un diagrama orbital.

3.3. Mapas de Poincaré

Frecuentemente se utiliza la construcción de una gráfica de espacio de fase para entender mejor el comportamiento de un sistema. Un mapa de fase espacial es una muestra de los parámetros que describen el comportamiento del sistema. Su utilidad reside en que permite un análisis gráfico del comportamiento del sistema. La Figura 26 muestra un ejemplo de una gráfica de espacio de fase para un péndulo simple. En el punto A de la Figura 26, el péndulo está a la máxima distancia positiva de su punto neutral pero su velocidad es cero. Esto se muestra como punto A en el diagrama de fase espacial. En B, la distancia del péndulo de su posición neutral es cero, pero está a su máxima velocidad (en un sentido negativo). Los otros puntos de la gráfica de espacio de fase muestran la relación entre la

velocidad y la posición de las otras posiciones del péndulo. En este caso, cuando no hay fricción, el movimiento del péndulo está limitado a permanecer dentro del patrón elíptico, el cual es el *atractor* del sistema. Se puede apreciar que este atractor es periódico porque el patrón que sigue el sistema se repite exactamente en cada órbita alrededor de su origen.

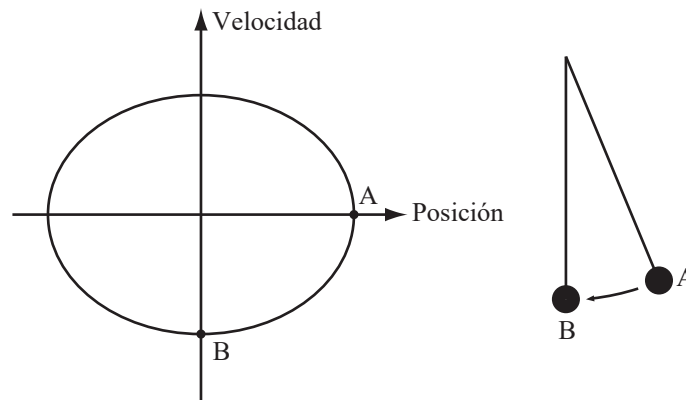


Figura 26. Diagrama de fase de un péndulo.

Es posible simplificar la presentación del atractor tomando un corte bidimensional del espacio de fase que cruce el atractor, esto también hace más evidente su estructura. A tal corte se le denomina *Mapa de Poincaré*. Los mapas de Poincaré son útiles en el estudio de flujos giratorios, como los que se presentan en la vecindad de una trayectoria periódica, o en los sistemas caóticos, dada su sensibilidad a condiciones iniciales. Si se considera un sistema n -dimensional $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ donde $\dot{\mathbf{x}}$ representa la primera derivada de \mathbf{x} y S una *sección de superficie* de $n-1$ dimensiones en el espacio de fase, no necesariamente plana, pero transversal al flujo, es decir, todas las trayectorias que comienzan en S deben cruzarlo. El mapa de Poincaré, denotado por \mathbf{p} , es un mapeo discreto de S sobre ella misma, es decir $\mathbf{p} : S \rightarrow S$, que recibe muchas de las propiedades del sistema original, por ejemplo, su periodicidad o cuasi-periodicidad [Strogatz, 2000;Löffelmann, 1998].

Los mapas de Poincaré representan la solución del sistema a través del tiempo. Si, por ejemplo, el sistema está siendo atraído hacia un ciclo límite, en el mapa se pueden ver puntos convergentes hacia un punto o puntos estacionarios, dependiendo del periodo del sistema. La Figura 27 (a) muestra el mapa que crearía un sistema cuyo estado estable es un punto fijo, y en (b) se puede ver distintos cruces de la trayectoria cuando el estado estable

del sistema es de doble periodo. Recordando lo visto en la sección anterior, un sistema caótico tiene un periodo infinito, lo cual se reflejará en el mapa de Poincaré como un gran número de puntos que crean un objeto.

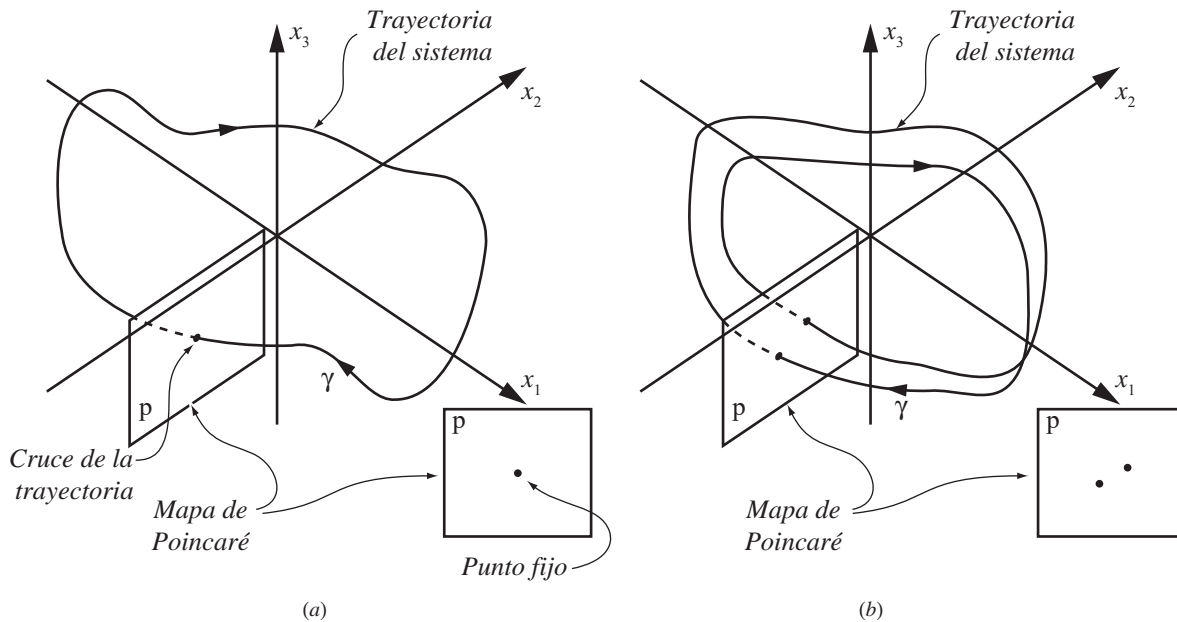


Figura 27. Mapa de Poincaré.

Un mapa de Poincaré muestra el atractor del sistema, ya sea un punto fijo, un ciclo límite o un atractor caótico. Un ejemplo interesante para la obtención de mapas de Poincaré lo constituye el comportamiento de un oscilador sinusoidal forzado, como el que representa la siguiente ecuación:

$$\ddot{x} + 0.05\dot{x} + x^3 = 7.5 \cos t \tag{3.5}$$

La ecuación (3.5) modela, por ejemplo, el movimiento de una estructura sometida a deflexiones elásticas de gran magnitud, donde \dot{x} y \ddot{x} son la primera y segunda derivada de x , respectivamente. Una serie de tiempo obtenida de la solución de esta ecuación se muestra en la Figura 28, para los valores iniciales de $x = 3$, $\dot{x} = 4$.

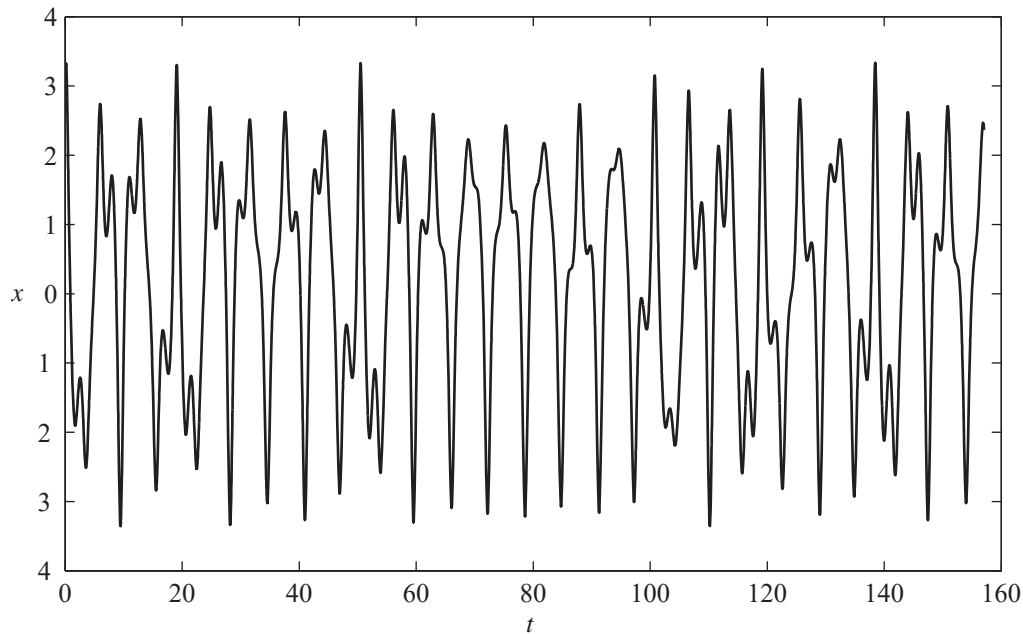


Figura 28. Serie de tiempo para el oscilador dado por la ecuación (3.5).

En la Figura 28 es posible ver una cierta recurrencia del sistema, dada la aparición de patrones a intervalos irregulares, sin embargo, no es una repetición exacta, y por tanto, el sistema es no-periódico.

Si dos sistemas se inician de valores ligeramente distintos, sucede una divergencia de la respuesta del sistema en el tiempo que en promedio se incrementa en un múltiplo fijo, y rápidamente las respuestas se vuelven no correlacionadas. Tal situación es inevitable en los sistemas físicos reales, dada la imposibilidad de obtener condiciones de inicio idénticas. La Figura 29 ilustra lo anterior. Esta característica es de enorme importancia en los sistemas caóticos. Los valores iniciales en este caso son de $x = 3$, $\dot{x} = 4$ para la línea continua y de $x = 3.01$, $\dot{x} = 4.01$ para la discontinua.

Un mapa de plano de fase muestra esta misma divergencia de una manera más clara (Figura 30). Nuevamente, los dos sistemas evolucionan de una manera muy similar durante un tiempo, pero al final, divergen y se vuelven no correlacionados. Sin embargo, estas trayectorias, aunque divergentes, representan versiones no correlacionadas del mismo patrón: un atractor caótico. A fin de encontrar este patrón, es necesario entonces el uso de mapas de Poincaré.

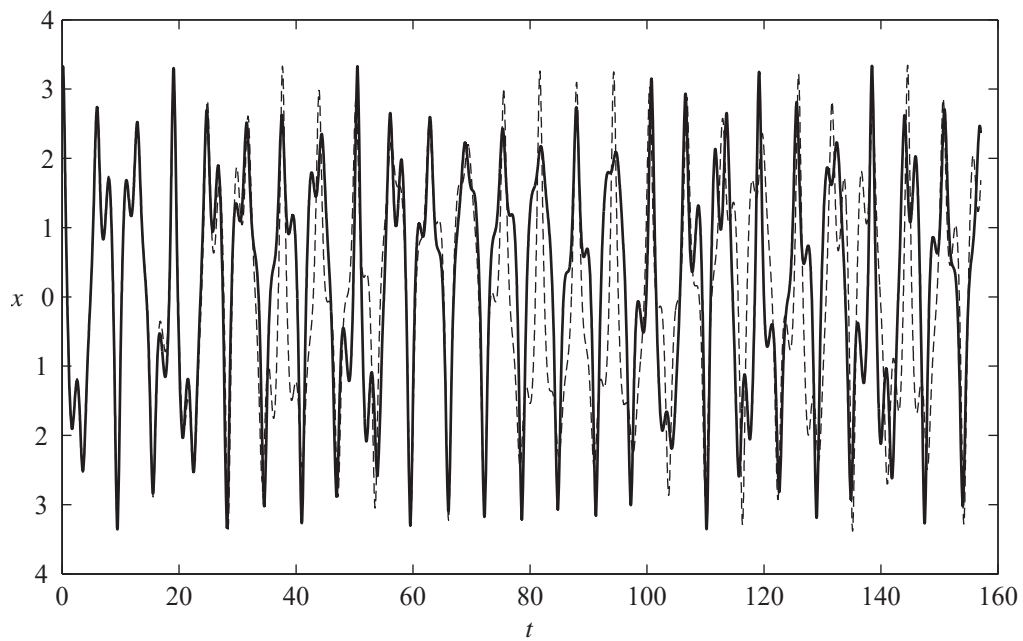


Figura 29. Divergencia exponencial en un sistema ante condiciones iniciales ligeramente distintas.

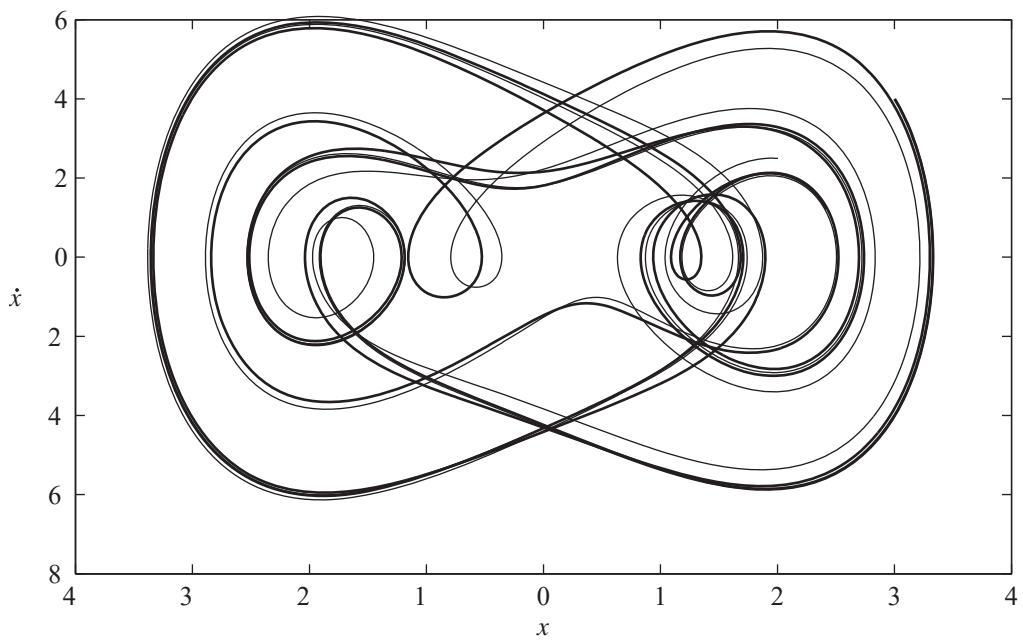


Figura 30. Diagrama de fase que muestra la divergencia exponencial del sistema ante condiciones iniciales ligeramente distinta.

La Figura 31 muestra el mapa de Poincaré para el oscilador del ejemplo. Cada punto en el mapa es una especie de “fotografía” instantánea, realizada en sincronía con la función

forzante del sistema, $7.5\cos(t)$, que para el caso de este oscilador es de 2π segundos, y delimita la región espacial que ocupa el atractor para ese valor. El sistema ejemplo presenta un comportamiento caótico, y por tanto su atractor es el lugar geométrico mostrado en la Figura 31. Si el sistema no fuera caótico, el mapa presentaría un solo punto para el caso de un atractor de punto fijo, o un par de puntos en el caso de un sistema de doble periodo.

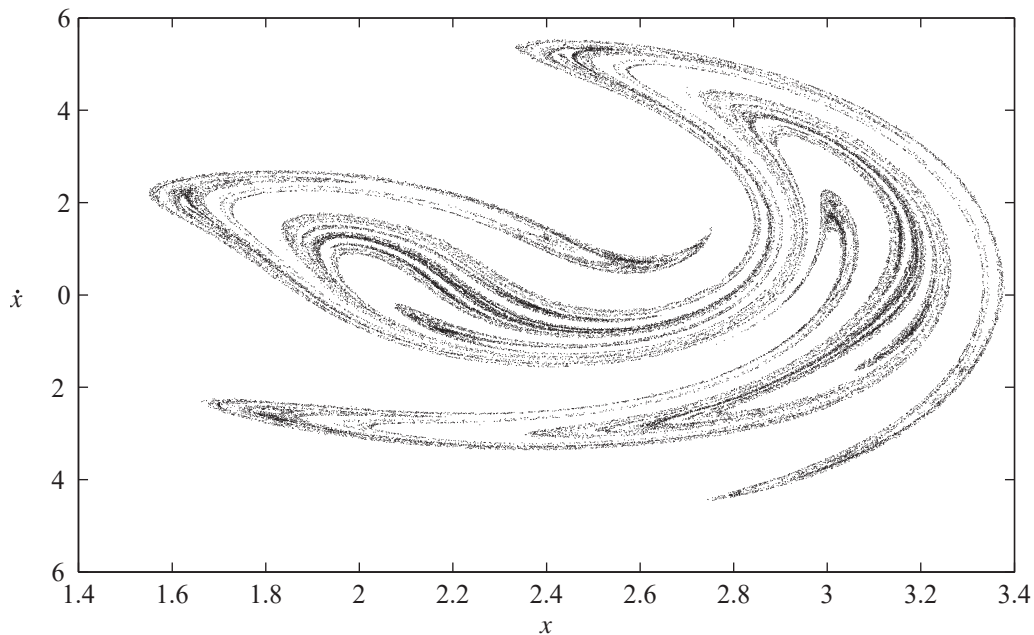


Figura 31. Mapa de Poincaré para el oscilador de la ecuación (3.5).

3.4. Ciclos límite

La respuesta de un sistema dinámico no lineal puede presentar estados de equilibrio asintóticamente estables, pero también pueden aparecer oscilaciones que atraen todas las trayectorias cercanas. A tales estados de equilibrio se les denomina *ciclos límite*. Tales oscilaciones son cerradas y *aisladas* en el sentido de que todas las trayectorias vecinas son abiertas. Para asegurar la existencia de un solo ciclo límite estable, es necesario que el *origen* del sistema sea inestable, es decir, para un sistema de dos variables $(0,0)$, de tal manera que las trayectorias de magnitud muy pequeña se muevan hacia fuera, y las de gran magnitud decaigan. También es posible encontrar ciclos límite inestables, o semi-estables [Strogatz, 2000; Kaplan y Glass, 1995; Thompson y Stewart, 1993].

Los ciclos límite son de gran importancia científica, ya que modelan los sistemas que exhiben oscilaciones auto-sostenidas, tales como el latido de un corazón, reacciones químicas que oscilan espontáneamente o vibraciones auto-excitadas en puentes y alas de aviones. En cada caso hay una oscilación estándar de algún periodo, forma de onda y amplitud preferidos, de tal manera que si el sistema se perturba ligeramente, siempre volverá a tal oscilación estándar.

Los ciclos límite son fenómenos inherentemente no lineales: No pueden suceder en sistemas lineales. Existen sistemas lineales que pueden tener órbitas cerradas, pero no aisladas. Dado un sistema lineal $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$, sea \mathbf{x} una solución periódica, entonces también lo será $c\mathbf{x}(t)$ para cualquier constante $c \neq 0$. Por tanto, \mathbf{x} estará rodeada por una familia de soluciones uniparamétricas de curvas cerradas. En consecuencia, la amplitud de la oscilación lineal queda establecida completamente por la condición inicial. Cualquier perturbación a la amplitud persistirá por siempre, a diferencia de los ciclos límite, que quedan determinados por la estructura misma del sistema.

Una ecuación que ha tenido un importantísimo papel en el desarrollo de los estudios acerca de la dinámica no lineal es la ecuación de *Van der Pol*:

$$\ddot{x} + \mu(x^2 - 1)\dot{x} + x = 0 \quad (3.6)$$

donde $\mu \geq 0$ es un parámetro que determina la no-linealidad del sistema. Esta ecuación es un modelo matemático de la oscilación auto-sostenida de un circuito triodo con una característica de voltaje-corriente cúbica. Tiene un interés relevante, sobre todo para los casos extremos de μ pequeña o muy grande. Cuando μ es pequeña, se obtienen oscilaciones no lineales muy débiles, es decir, oscilaciones que difieren ligeramente de un movimiento armónico simple, mientras que para μ muy grande, se obtienen oscilaciones de relajamiento, es decir, oscilaciones altamente no lineales con saltos bruscos periódicos. En la Ecuación (3.6), el segundo término actúa como un amortiguamiento positivo ordinario para $|x| > 1$, pero como un amortiguamiento *negativo* para $|x| < 1$. El resultado de ello es que las oscilaciones de gran amplitud decaen, pero las impulsa a crecer si se vuelven demasiado pequeñas. Tal comportamiento lleva al sistema a establecerse en una oscilación auto-sostenida. La solución numérica del sistema para un valor de $\mu=1.5$, y condiciones iniciales

de $x(0) = 0.5$ y $\dot{x}(0) = 0$ se muestra en la Figura 32 y en la Figura 33 la solución para los mismos valores iniciales, pero con $\mu=5$.

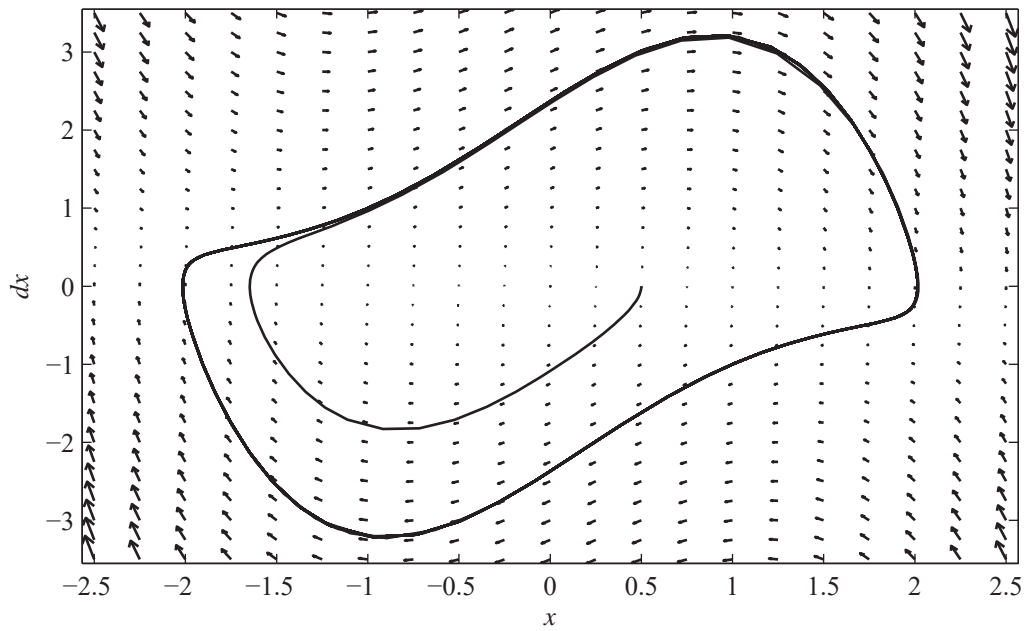


Figura 32. Gráfico del espacio de fase para la ecuación de Van der Pol con $\mu = 1.5$.

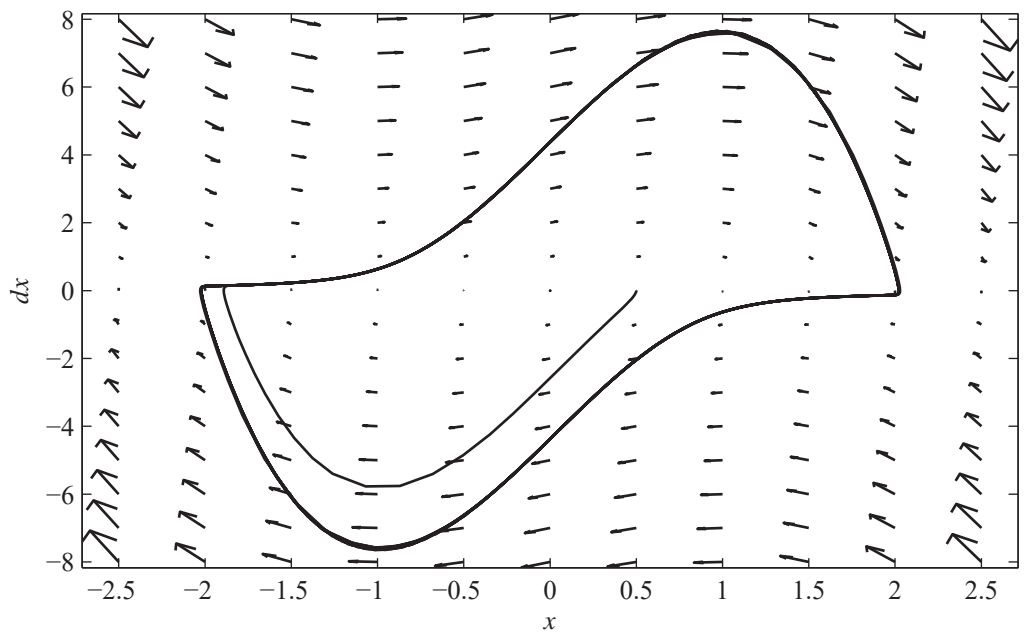


Figura 33. Gráfico del espacio de fase para la ecuación de Van der Pol con $\mu = 5$.

El campo vectorial que se muestra en las figuras anteriores ilustra el hecho de que la ecuación de Van der Pol tiene un único ciclo límite estable para cada $\mu > 0$. Posteriormente, esta ecuación se generalizó en lo que ahora es conocido como *ecuación de Liénard*, la cual es una familia de ecuaciones diferenciales de segundo orden de la forma:

$$\ddot{x} + f(x)\dot{x} + g(x) = 0 \quad (3.7)$$

y mediante el Teorema de Liénard es posible encontrar tales ciclos límite [Liénard, 1928].

3.5. Atractores extraños

A diferencia de sistemas cuyo estado estable es un punto fijo o un ciclo límite, aquellos que presentan un régimen caótico una vez que el transitorio del sistema ha decaído completamente, se dice que están gobernados por un *atractor caótico* o *atractor extraño* (término acuñado por Ruelle y Takens). Bajo esta perspectiva, un atractor extraño es el conjunto límite de una trayectoria caótica. Prácticamente la mayoría de los sistemas no lineales son impredecibles, pero aún así, presentan patrones: Tales patrones son en realidad formas geométricas fractales, por tanto, también es posible decir que son atractores fractales [Strogatz, 2000;Thompson y Stewart, 1993;Kaplan y Glass, 1995].

De una manera informal, un atractor es un conjunto al cual convergen todas las trayectorias cercanas. Un atractor extraño, denotado por A , cumple con las siguientes propiedades:

- A es un conjunto invariante. Cualquier trayectoria que inicia en A , permanece en A para cualquier tiempo.
- A atrae un conjunto de condiciones iniciales abierto. Dado un conjunto abierto U que contiene a A , y para cualquier condicional inicial que pertenezca a U , entonces A atrae todas las trayectorias que inician lo suficientemente cercanas a él. Al mayor conjunto U se le denomina *cuenca de atracción* de A .
- A es mínimo. No existen subconjuntos propios de A .
- A es sensible a las condiciones iniciales.

Inicialmente estos atractores se denominaron extraños por su propiedad fractal, pero en la actualidad esta propiedad geométrica se considera de menor importancia que la propiedad dinámica de sensibilidad a las condiciones iniciales.

Los atractores caóticos pueden coexistir con otros tipos de estados estables periódicos, cada cual con su dominio adecuado de atracción.

Un atractor caótico típico es el que muestra el sistema de Hénon [Hénon, 1976], definido por el sistema de ecuaciones siguiente:

$$\begin{aligned}x_{n+1} &= y_n + 1 - ax_n^2 \\ y_{n+1} &= bx_n\end{aligned}\tag{3.8}$$

estas ecuaciones se pueden arreglar en una ecuación recurrente:

$$x_{n+1} = 1 - ax_n^2 + bx_{n-1}\tag{3.9}$$

la cual, con valores para $a = 1.4$ y $b = 0.3$, proporciona la Figura 34.

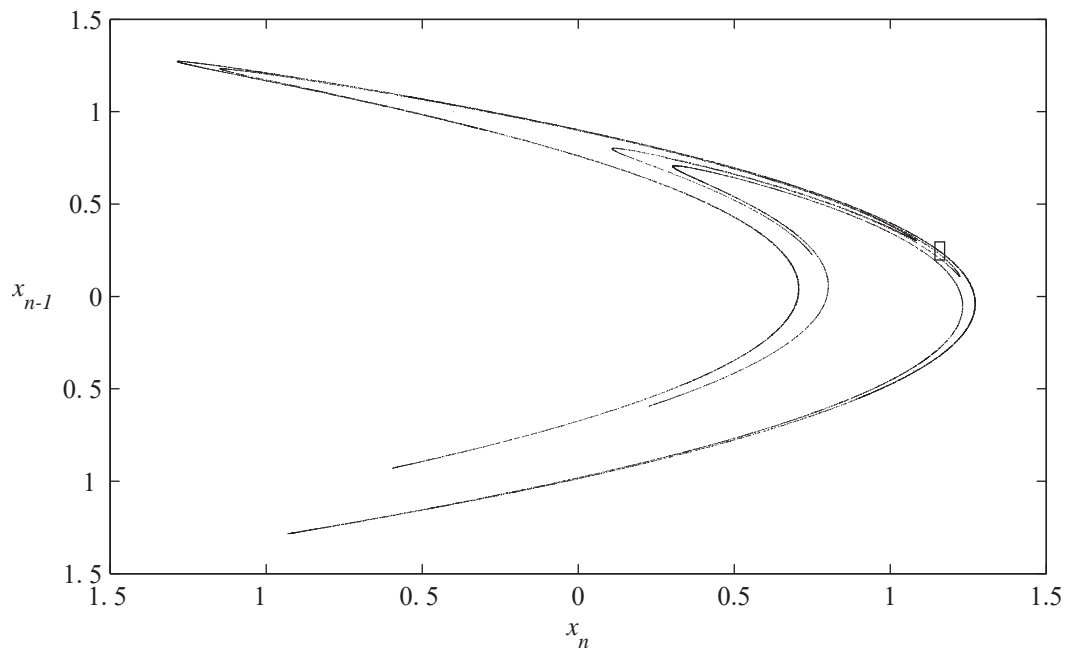


Figura 34. Atractor caótico de Hénon.

Este atractor presenta características interesantes, por ejemplo, lo que aparentan ser líneas simples, al ampliar la imagen es posible darse cuenta que son líneas dobles y triples, sucesivamente, conforme se amplía sobre algún área en particular del atractor. En la Figura 34 se marca un recuadro en la parte derecha del atractor, haciendo una ampliación de esa área, se obtiene la Figura 35, que muestra tal multiplicidad de líneas. Ésto revela la característica fractal del atractor.

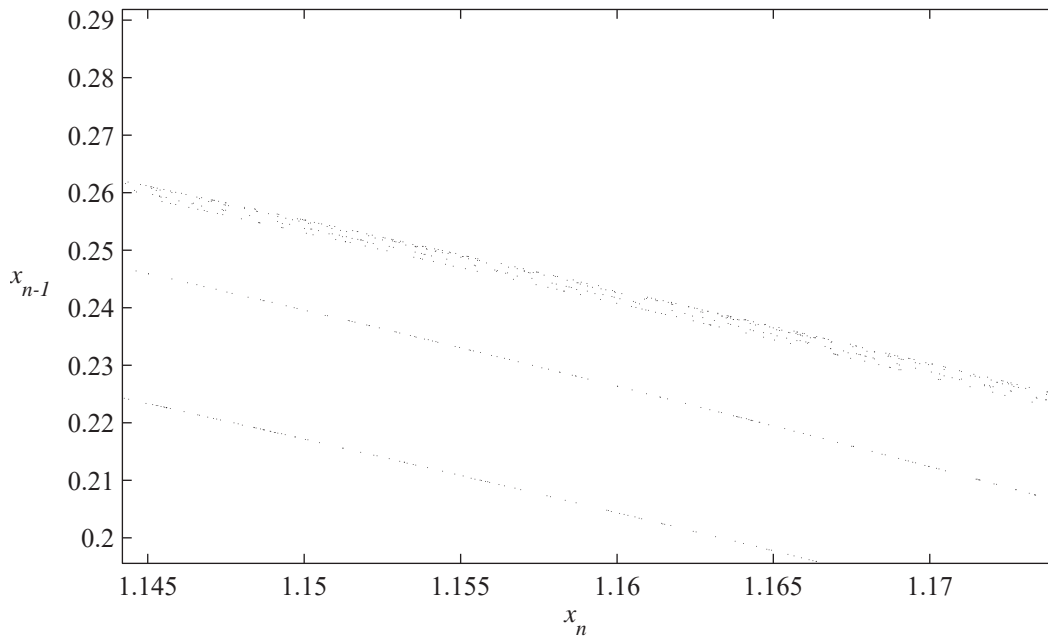


Figura 35. Ampliación sobre una región del atractor de Hénon de la Figura 34.

3.6. Gráficos de recurrencia

Los gráficos de recurrencia son utilizados para revelar características de no-estacionariedad de una serie de tiempo, así como para indicar su grado de aperiodicidad. Inicialmente propuestos por [Eckmann et al, 1987] para el estudio de órbitas en el espacio de fase. Con los gráficos de recurrencia es posible detectar, asimismo, patrones ocultos y cambios estructurales en datos, así como similitudes en patrones en una serie de tiempo.

La suposición fundamental que subyace en los gráficos de recurrencia es el hecho que una serie de tiempo observable es la concretización de algún proceso dinámico. Procesos altamente complejos tales como el clima o el mercado cambiario, con un gran número de factores que influye sobre su dinámica, pero cuya concretización es una única medición: en

el caso del clima, las lecturas de temperatura; y en el mercado cambiario, el precio promedio de cierre de valores accionarios. En ambos casos se tienen sistemas dinámicos multidimensionales y una salida uni-dimensional. Matemáticamente se ha probado que es posible recrear una “fotografía” topológicamente equivalente del sistema multidimensional original mediante el uso de una sola serie de tiempo observable [Takens, 1981]. La idea básica, es que el efecto de todas las otras variables del sistema (no observadas) está ya reflejado en la serie de salida observada. Además, las reglas que gobiernan el comportamiento del sistema original se pueden recuperar de su salida.

Un gráfico de recurrencia de una secuencia de N puntos x_1, x_2, \dots, x_N , es la visualización de la matriz de recurrencia de esa secuencia. El color del elemento (i,j) y (j,i) en el gráfico de recurrencia es negro si la distancia entre el i -ésimo y el j -ésimo puntos en la serie de tiempo yace entre algún *corredor de umbral*:

$$\delta_l < \|x_i - x_j\| < \delta_h \quad (3.10)$$

para alguna norma adecuada, donde δ_l es el valor mínimo y δ_h el valor máximo de umbral y es de color blanco para cualquier otra. Esta técnica, aplicada a una serie de tiempo, permite visualizar las correlaciones temporales existentes dentro de la serie. La Figura 36 muestra los gráficos de recurrencia de distintas series de tiempo.

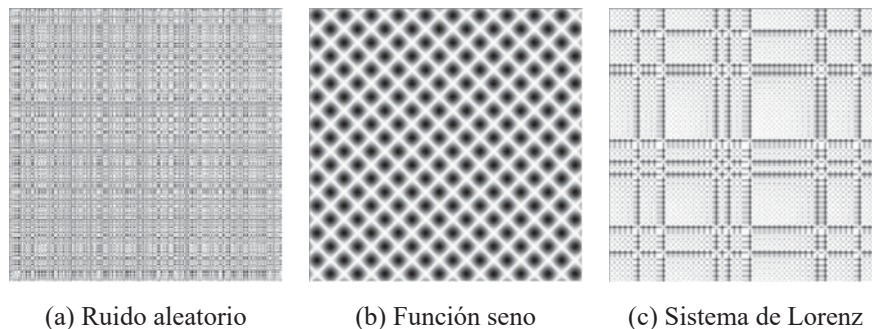


Figura 36. Gráficos de recurrencia para diferentes series de tiempo.

En (a) se tiene una serie de datos aleatorios obtenidos de una distribución uniforme que no muestra ningún patrón definido, como es de esperarse, pues no existe correlación entre un valor dado y valores previos o posteriores. En (b) se grafica la serie de datos para

una función seno. Esta gráfica muestra un patrón perfectamente definido dado que se trata de una función periódica. En (c) se tiene el gráfico para una serie de tiempo del sistema de Lorenz. Precisamente en esta última serie es donde cobra valor el gráfico de recurrencia, pues permite visualizar el patrón que emerge de una serie caótica. Aún cuando no es un patrón perfectamente definido, como aquel para la función senoidal en (b), sí lo es mucho más definido que para la serie de ruido de (a). Esto también muestra que existe una cierta correlación entre valores de la serie de tiempo de Lorenz. Estos gráficos fueron obtenidos con el paquete Análisis visual de recurrencias (Visual Recurrence Analysis) [Kononov, 2003]. El uso de esta herramienta nos permite determinar de una manera simple parámetros importantes en el proceso de predicción del esquema que se propone en la tesis. Por un lado, es posible visualizar si una serie dada presenta patrones en su comportamiento, y por otro, calcula otros importantes parámetros, como lo son la información mutua promedio y la cantidad de vecinos falsos, que se revisan en las secciones posteriores.

Entre las características importantes del gráfico de recurrencia se puede mencionar el hecho que, a diferencia de la transformada rápida de Fourier, permite ver no sólo las frecuencias presentes, sino dónde ocurren las señales correspondientes; es bastante robusto al ruido y preserva las invariantes del sistema dinámico del cual provenga la serie de tiempo.

Los gráficos de recurrencia de series de tiempo caóticas presentan una estructura rica y altamente característica. El primer esfuerzo en la literatura por definir una métrica cuantitativa para esta estructura se denominó *Análisis de cuantificación de recurrencia*, creado por [Zbilut y Webber, 1992]. Este análisis define diversas medidas estadísticas de los puntos negros en el gráfico. Estas cantidades miden muchas de las mismas propiedades dinámicas que hacen métodos como el *exponente de Lyapunov* y la dimensión de correlación.

3.7. Información mutua promedio.

La información mutua promedio, denotada por I , proporciona una medida general para la dependencia de dos conjuntos de variables dinámicas. Propuesta en [Fraser y Swinney,

1986], como un estimado para determinar la dimensión de τ_s , la coordenada de retraso para la graficación de atractores.

Cuando es posible observar solamente una variable, $x(t)$, de algún sistema, y se requiere construir un atractor multi-dimensional, es factible utilizar $(x(t), x(t+\tau), \dots, x(t+n\tau))$ para crear un atractor de dimensiones $(n+1)$, donde τ es un valor entero de retraso. Es necesario elegir un valor adecuado de retraso τ que proporcione la máxima información posible acerca de la serie. Un retraso pequeño tiene como efecto que $x(t)$ sea muy similar a $x(t+\tau)$. Si se grafica $x(t)$ contra $x(t+\tau)$, todos los datos permanecerán muy cerca de la línea $x(t) = x(t+\tau)$. Si el retraso es grande, ambas series serán forzosamente independientes y no surgirá una verdadera ganancia de información acerca del sistema.

Dado que la información mutua es una medida de dependencia, un valor que tienda a cero será un valor adecuado, siendo de uso común emplear el primer mínimo que se alcance en el cálculo de I .

El algoritmo para el cálculo de I evalúa el histograma normalizado de $x(t)$, y de su probabilidad $P(x(t))$, y el histograma normalizado de $x(t+\tau)$, y su probabilidad $P(x(t+\tau))$, así como el histograma conjunto normalizado de las variables $x(t)$ y $x(t+\tau)$, $P(x(t), x(t+\tau))$. Promediando sobre la información mutua obtenida de:

$$\log_2 \left[\frac{P(x(t))P(x(t+\tau))}{P(x(t), x(t+\tau))} \right] \quad (3.11)$$

sobre todas las mediciones, se llega a:

$$I = \sum_{x(t), x(t+\tau)} P(x(t), x(t+\tau)) \log_2 \left[\frac{P(x(t))P(x(t+\tau))}{P(x(t), x(t+\tau))} \right] \quad (3.12)$$

Lo que esta ecuación determina es que dado un valor de $x(t)$, cuál es la probabilidad de que τ pasos adelante el valor de $x(t+\tau)$ nos proporcione la máxima información sobre el sistema.

La información mutua promedio nos permite entonces definir un parámetro importante dentro del esquema de predicción, objetivo de la tesis: el número de pasos que la serie de

tiempo debe adelantarse, para crear cada vector de entrenamiento de la red neuronal de Kohonen. Una serie de tiempo de n elementos está definida por (x_1, x_2, \dots, x_n) . Si para esa serie se obtiene un valor de información mutua promedio de $\tau=4$, entonces, para recrear un atractor tridimensional, se generará el siguiente arreglo:

$$\begin{array}{ccc} x_1 & x_5 & x_9 \\ x_2 & x_6 & x_{10} \\ \vdots & \vdots & \vdots \\ x_{n-9} & x_{n-5} & x_n \end{array}$$

El cálculo de este parámetro es parte de la etapa de parametrización de la serie de tiempo dentro del esquema de predicción que se explicó en la sección 1.3.

3.8. Algoritmo de vecinos falsos.

El algoritmo de vecinos falsos proporciona la dimensión envolvente, m , del atractor de un sistema. La dimensión envolvente es la dimensión del espacio de estado de un sistema. Un sistema de dos variables tiene, por ejemplo, una dimensión envolvente de $m=2$. El teorema de Takens establece que un atractor se *desenvolverá* completamente, sin cruces de flujos, si la dimensión envolvente se elige de tal manera que sea mayor a $2m$.

En [Kennel et al, 1992] se propone el algoritmo que proporciona un valor de m adecuado, basado en el hecho de que dos puntos cercanos entre sí en una dimensión m , podrían estar muy alejados en una dimensión $m+1$. Esto sucede cuando la menor dimensión es simplemente una proyección de dimensiones superiores. Así, el algoritmo establece primeramente vecinos candidatos, incrementa la dimensión m y entonces inspecciona en búsqueda de vecinos falsos. Cuando no se logra identificar vecinos falsos, se puede decir que el atractor está completamente desenvuelto, y m se toma como la dimensión envolvente.

Este es otro parámetro importante en el esquema de predicción que se propone en la tesis. Se emplea en la etapa de parametrización definida en la sección 1.3 y define la dimensión del arreglo que se crea para entrenar la red de Kohonen, en conjunto con el parámetro τ obtenido en la sección anterior.

Como ejemplo, si se tiene un $\tau = 10$ y $m=3$, para una serie de tiempo definida por (x_1, x_2, \dots, x_n) , se creará un arreglo como el siguiente:

$$\begin{array}{ccc} x_1 & x_{11} & x_{21} \\ x_2 & x_{12} & x_{22} \\ \vdots & \vdots & \vdots \\ x_{n-21} & x_{n-11} & x_n \end{array}$$

La obtención de estos dos parámetros completa la etapa de preprocesamiento de la serie de tiempo. Se está entonces en la posibilidad de entrenar la red de Kohonen, que se revisa en el capítulo 5.

3.9. Conclusiones

En este capítulo se reseñaron diversos métodos de análisis de caos. Tal análisis ha venido cobrando una enorme importancia desde el “descubrimiento” del caos en 1962, con las investigaciones de Edward N. Lorenz sobre modelos climáticos, dada su presencia en muchas series de tiempo de interés científico.

Se definieron conceptos básicos de sistemas caóticos y características que deben presentar los mismos a fin de que se puedan catalogar como caóticos, entre ellas, la más importante es probablemente la sensibilidad a las condiciones iniciales.

Los sistemas caóticos, dada su complejidad, resultan más fáciles de analizar desde un punto de vista gráfico, es por ello que un gran número de métodos son métodos gráficos. Tal es el caso de los mapas iterados, los mapas de Poincaré y los mapas de recurrencia.

4. Mínimos cuadrados

4.1. Introducción

La técnica de aproximación por el criterio de *mínimos cuadrados* es la forma más común de regresión lineal y proporciona una solución al problema de encontrar la mejor función suave que se aproxime a un conjunto dado de puntos, por tanto, también es conocida como *aproximación funcional*.

Antes de aplicar este criterio, debe escogerse la forma de la curva que deba ajustarse al conjunto de puntos. La ecuación de esa curva puede obtenerse por conocimiento previo del problema, o de una manera arbitraria, observando qué ecuación conocida describe aproximadamente a tal conjunto.

En este capítulo se introducen los conceptos de mínimos cuadrados convencionales, y de mínimos cuadrados ponderados.

4.2. Mínimos cuadrados convencionales

El procedimiento de mínimos cuadrados para determinar un *polinomio de regresión* parte de que cada punto y (observación) del conjunto dado, es de la forma:

$$y = a_0 + a_1x + a_2x^2 + \cdots + a_mx^m \quad (4.1)$$

donde a_0, a_1, \dots, a_m , son los coeficientes de la función a determinar. La Figura 37 muestra una gráfica de la aproximación de la curva.

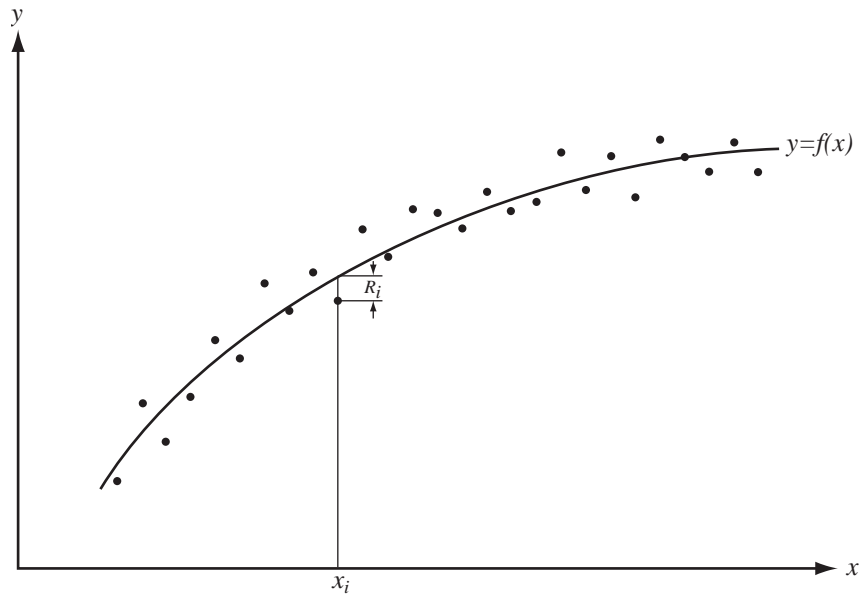


Figura 37. Ajuste de curvas.

La figura muestra la diferencia de la ordenada de la curva para $x = x_i$ menos la del punto (x_i, y_i) , denominada *residuo*, y representada como R_i . Se tiene entonces que:

$$R_i = f(x_i) - y_i \quad (4.2)$$

es decir:

$$R_i = a_0 + a_1 x_i + a_2 x_i^2 + a_m x_i^m - y_i \quad (4.3)$$

el método de los mínimos cuadrados consiste entonces en determinar los valores de los parámetros a_0, a_1, \dots, a_m de manera que hagan mínima la suma de los cuadrados de los residuos. Esta suma equivale a:

$$\sum_{i=1}^n R_i^2 = \sum_{i=1}^n (a_0 + a_1 x_i + a_2 x_i^2 + \dots + a_m x_i^m - y_i)^2 \quad (4.4)$$

se obtiene el mínimo de ésta igualando a cero sus primeras derivadas parciales con respecto a todos y cada uno de los parámetros. Derivando con respecto a a_j , donde $j = 0,1,2, \dots, m$, se obtiene:

$$\begin{aligned} \frac{\delta}{\delta a_j} \sum_{i=1}^n R_i^2 &= \frac{\delta}{\delta a_j} \sum_{i=1}^n (a_0 + a_1 x_i + a_2 x_i^2 + \dots + a_m x_i^m - y_i)^2 \\ &= \sum_{i=1}^n \frac{\delta}{\delta a_j} (a_0 + a_1 x_i + a_2 x_i^2 + \dots + a_m x_i^m - y_i)^2 \\ &= \sum_{i=1}^n 2(a_0 + a_1 x_i + a_2 x_i^2 + \dots + a_m x_i^m - y_i) x_i^j \end{aligned} \quad (4.5)$$

igualando con cero esta derivada se llega a:

$$a_0 \sum_{i=1}^n x_i^j + a_1 \sum_{i=1}^n x_i^{j+1} + a_2 \sum_{i=1}^n x_i^{j+2} + a_3 \sum_{i=1}^n x_i^{j+3} + \dots + a_m \sum_{i=1}^n x_i^{j+m} = \sum_{i=1}^n x_i^j y_i \quad (4.6)$$

finalmente, considerando $j = 0,1,2, \dots, m$, se obtiene el siguiente sistema de *ecuaciones normales*:

$$\left\{ \begin{array}{l} na_0 + a_1 \sum x + a_2 \sum x^2 + \dots + a_m \sum x^m = \sum y \\ a_0 \sum x + a_1 \sum x^2 + a_2 \sum x^3 + \dots + a_m \sum x^{m+1} = \sum xy \\ a_0 \sum x^2 + a_1 \sum x^3 + a_2 \sum x^4 + \dots + a_m \sum x^{m+2} = \sum x^2 y \\ \vdots \\ a_0 \sum x^m + a_1 \sum x^{m+1} + a_2 \sum x^{m+2} + \dots + a_m \sum x^{2m} = \sum x^m y \end{array} \right. \quad (4.7)$$

por simplicidad se omiten los subíndices de x y y , así como los rangos de las sumas.

Para definir las ecuaciones normales en notación matricial se definen las siguientes matrices:

$$\mathbf{X} = \begin{pmatrix} 1 & x_1^1 & x_1^2 & \dots & x_1^m \\ 1 & x_2^1 & x_2^2 & \dots & x_2^m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n^1 & x_n^2 & \dots & x_n^m \end{pmatrix} \quad (4.8)$$

$$\mathbf{Y} = (y_1 \quad y_2 \quad \cdots \quad y_n)^T \quad (4.9)$$

$$\mathbf{B} = (a_1 \quad a_2 \quad \cdots \quad a_m)^T \quad (4.10)$$

la primera matriz, \mathbf{X} , es una matriz $n \times (m+1)$ que consiste esencialmente de los valores dados de las x , donde se añade una columna de 1's para dar cabida a los términos constantes. \mathbf{Y} es un vector de $n \times 1$ que consiste de los valores observados de y , y \mathbf{B} es un vector $(m+1) \times 1$ que consiste de las estimaciones de mínimos cuadrados de los coeficientes de regresión. Al emplear estas matrices se puede escribir el sistema de ecuaciones normales de la siguiente manera:

$$\mathbf{B} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \quad (4.11)$$

4.3. Mínimos cuadrados ponderados

Una suposición común de los mínimos cuadrados convencionales es que cada punto proporciona un mismo nivel informativo, en lo que se refiere a la parte determinista de la variación total del proceso, sin embargo, ésta no es aplicable a todos los modelos. En los casos donde esta suposición no es razonable, los mínimos cuadrados ponderados pueden maximizar la eficiencia en la estimación paramétrica. Esto se lleva a cabo tratando de proporcionar a cada punto la cantidad adecuada de "influencia" sobre los parámetros a estimar [NIST/SEMATECH, 2003].

A diferencia de los mínimos cuadrados convencionales, los mínimos cuadrados ponderados no están asociados con un tipo particular de función que describa la relación entre las variables del proceso, sino que en su lugar, refleja el comportamiento de los errores aleatorios en el modelo. Los mínimos cuadrados ponderados se pueden usar con funciones lineales o no lineales en sus parámetros. Trabaja mediante la incorporación de una constante, o *ponderación*, asociada con cada dato dentro del criterio de ajuste de la curva. El tamaño de la ponderación indica la precisión de la información contenida en la observación asociada. Cuando el criterio de ponderación se optimiza para encontrar estimaciones de los parámetros, los *pesos* determinan la contribución de cada observación a la estimación final de los parámetros. Así, se puede concluir que los mínimos cuadrados

ponderados es un método eficiente y que hace buen uso de conjuntos pequeños de datos. Es, asimismo, capaz de proporcionar diferentes tipos de intervalos estadísticos fácilmente interpretables para estimación, predicción u optimización, pero su principal ventaja es la capacidad para manejar situaciones en las cuales los datos son de una calidad variable (entendiéndose calidad como la importancia o influencia de un dato sobre los parámetros). Si la desviación estándar del error en los datos no es constante sobre todos los niveles de las variables independientes, el uso de mínimos cuadrados ponderados cuyos pesos sean inversamente proporcionales a la varianza en cada nivel de las variables independientes, proporciona el estimado paramétrico más preciso posible.

Entre las desventajas de los mínimos cuadrados ponderados está el suponer que los pesos se conocen de manera exacta, y éste, desafortunadamente, casi nunca es el caso en aplicaciones reales, de ahí el uso de pesos estimados. El efecto de usar pesos estimados es difícil de valorar, pero la experiencia indica que aunque la mayoría de los resultados de análisis de regresión son insensibles a pequeñas variaciones en los pesos, pueden afectarse gravemente por los cambios relativamente grandes que se observan, en especial, cuando tales estimaciones se hacen de una manera inexperta y basada en pequeñas cantidades de datos. A pesar de ello, se puede obtener en gran medida los beneficios de los mínimos cuadrados ponderados, utilizando pesos cuidadosamente aproximados.

Los mínimos cuadrados ponderados, como cualquier otro método de mínimos cuadrados, es sensible a los efectos de datos atípicos. Si tales datos no son analizados y tratados adecuadamente, tendrán un impacto negativo en la estimación final de los parámetros.

En la estimación por mínimos cuadrados ponderados, los valores desconocidos de los parámetros a_0, a_1, \dots, a_m , en la función de regresión, se calculan mediante la minimización de la suma del cuadrado de la diferencia entre las respuestas observadas y estimadas, es decir los residuos. A diferencia de los métodos convencionales, cada término en el criterio de mínimos cuadrados incluye un peso adicional w_i , que determina qué tanto influye cada observación en el estimado final de los parámetros. El criterio de mínimos cuadrados ponderados que se minimiza para obtener la estimación es:

$$Q = \sum_{i=1}^n w_i [y_i - f(x_i; \beta)]^2 \quad (4.12)$$

4.4. Comparación entre mínimos cuadrados convencionales y mínimos cuadrados ponderados

A continuación se realiza un ejemplo comparativo del uso de los mínimos cuadrados convencionales y mínimos cuadrados ponderados.

Se puede emplear una función simple para generar datos a intervalos regulares, a los cuales se les agrega ruido uniformemente distribuido con la finalidad de ocultar la función generadora.

Una función útil puede ser la ecuación (4.13)

$$y = 10 - 2x \quad (4.13)$$

Los datos de la Tabla 1 se generaron a partir de la ecuación (4.13) más ruido. Adicionalmente los datos 4 y 10 se modificaron en mayor grado, de tal manera que aparezcan como datos atípicos dentro del proceso.

Dato	Valor
1	8.000
2	5.6821
3	5.0950
4	-3.000
5	-0.4282
6	-1.1044
7	-3.2690
8	-5.4221
9	-7.9597
10	1.000

Tabla 1. Datos generados.

El gráfico de la función y de los puntos se muestra en la Figura 38.

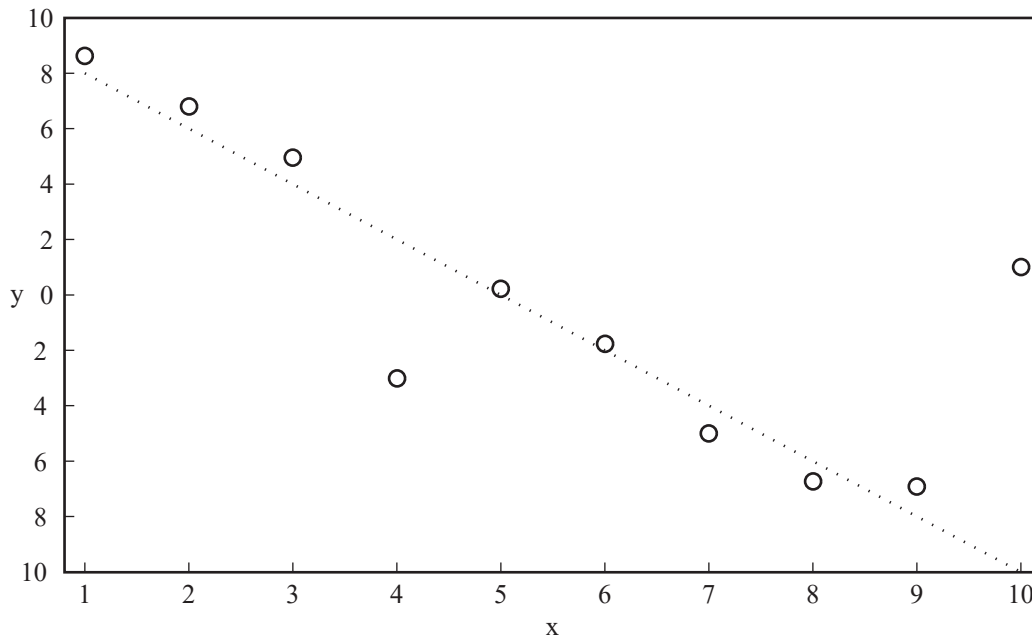


Figura 38. Gráfico de la función (4.13) y de los datos generados.

Aplicando los métodos de mínimos cuadrados para la estimación de los parámetros de la función de aproximación, se llega a:

$$y' = 7.5158 - 1.4002x \quad (4.14)$$

Con el método de mínimos cuadrados ponderados se llega a la siguiente ecuación:

$$y' = 10.8920 - 2.1245x \quad (4.15)$$

la aproximación que se obtiene en este caso es mucho mejor. Al graficar ambos resultados se puede ver de una manera más clara, como lo muestra la Figura 39.

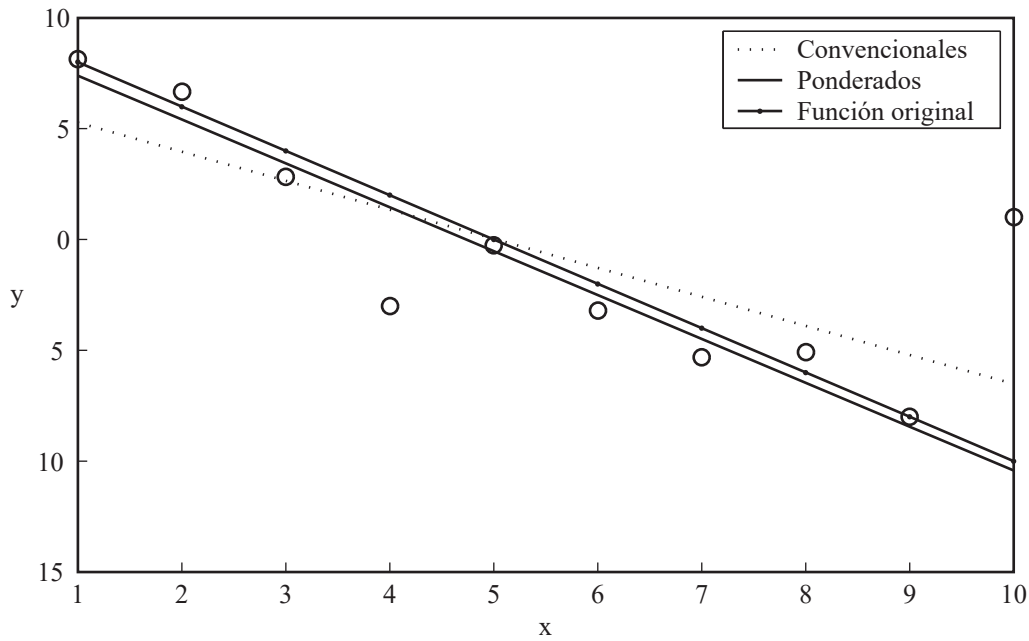


Figura 39. Gráfico de las funciones obtenidas con los métodos de mínimos cuadrados convencionales y mínimos cuadrados ponderados.

Los mínimos cuadrados ponderados presentan un mejor comportamiento ante datos atípicos, y esta característica se vuelve importante en el esquema de predicción que se propone en la tesis.

En el proceso predictivo que se explicó en la sección 1.3 se requiere de un vector interrogante para iniciar la predicción, el cual se pre-procesa y se pasa a la etapa de clasificación. Durante el entrenamiento de la red neuronal, que es la parte central del proceso de clasificación, un conjunto de vectores de entrenamiento quedaron asociados a cada neurona de la red. El vector interrogante, en el proceso de clasificación, es entonces comparado con todas las neuronas, resultando una sola neurona ganadora. Una nueva comparación, pero ahora con el conjunto de vectores de entrenamiento asociados con esa neurona, determina los pesos que se emplearán en el método de mínimos cuadrados. La intención es que aquellos vectores con los cuales el vector interrogante presente mayor diferencia, tengan una menor influencia en la predicción.

4.5. Conclusiones

En este capítulo se revisan los procesos de mínimos cuadrados y de mínimos cuadrados ponderados, ambos útiles herramientas en la predicción o estimación de parámetros de sistemas, tanto lineales, como no lineales. El interés de esta revisión es el hecho de que el proceso de mínimos cuadrados forma la parte central en la etapa de reconstrucción de un sistema, en el esquema que se propone en esta tesis.

Se señalaron ventajas y desventajas del proceso de mínimos cuadrados, y sus diferencias esenciales con los mínimos cuadrados ponderados, así como las ecuaciones que definen a cada uno de estos métodos. Se mostraron los aspectos ventajosos que presentan los mínimos cuadrados ponderados y que lo hacen más adecuado para su uso en el esquema propuesto.

5. Clustering con Kohonen

5.1. Introducción

En este capítulo se presenta el algoritmo de la red neuronal conocida como *Mapa auto-organizado de Kohonen*. Se presenta la metodología de aprendizaje, adaptación y los diversos esquemas que pueden emplearse en la red [Kohonen, 1989].

Dentro del esquema de predicción, el mapa auto-organizado de Kohonen cumple con la función de clasificar los distintos modelos o regímenes que se presentan en una serie de tiempo. Es por tanto la parte central de la etapa de clasificación de modelos.

Dado que el propósito de un mapa auto-organizado es la visualización de espacios ordenados, se hace énfasis en las características visuales que presentan cada una de los ejemplos de clasificación. De la misma manera, se ejemplifica el uso de esta red mediante algunos problemas típicos en el área de clasificación, como la clasificación de pétalos de Iris. Estos datos están disponibles en el sitio <ftp://ftp.ics.uci.edu/pub/machine-learning-databases/>.

Se hace uso del *toolbox* de MATLAB, denominado SOM, (Self-Organizing Map) de la Helsinki University of Technology, programa bajo los términos de la fundación para el software gratuito, disponible en la dirección <http://www.cis.hut.fi/projects/somtoolbox/>.

5.2. Algoritmo de los mapas auto-organizados

Un mapa auto-organizado consiste en un conjunto de neuronas ordenadas en un arreglo de baja dimensionalidad. El número de neuronas puede variar de unas pocas a varios miles.

El algoritmo de auto-organización consiste en dos pasos básicos [Kohonen, 1989]:

- Selección de la celda con mejor correspondencia.
- Adaptación de los vectores de peso.

Selección de la celda con mejor correspondencia.

Considere la red bidimensional mostrada en la Figura 40:

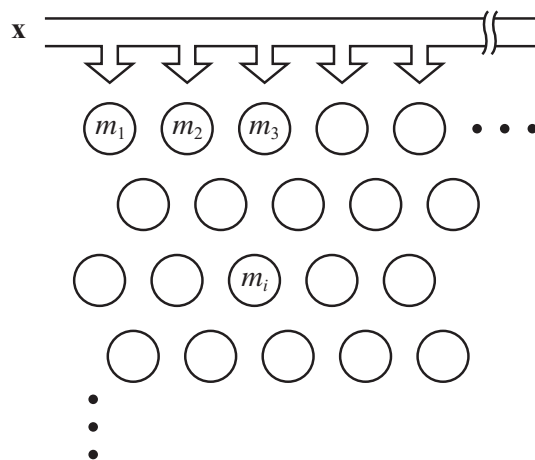


Figura 40. Arreglo de celdas para un mapa.

Donde cada m_i es una neurona de la red. El arreglo de tales neuronas puede ser rectangular o hexagonal, como muestra la Figura 41. En un arreglo hexagonal, una neurona presenta una conexión con sus seis vecinas. En la Figura 41 (a), la neurona encerrada en el hexágono pequeño tiene como vecinas a las neuronas encerradas en el hexágono mayor. En el arreglo rectangular cada neurona se conecta con sus cuatro vecinas más cercanas. En (b) se muestra el vecindario de una neurona en un arreglo rectangular.

A cada neurona se asocia un vector de peso que coincide en dimensiones con el vector de entrada del sistema. La intención es realizar una comparación entre el vector de entrada

y el vector de peso asociado a cada neurona, y llevar a cabo tal comparación con todas las neuronas de la red.

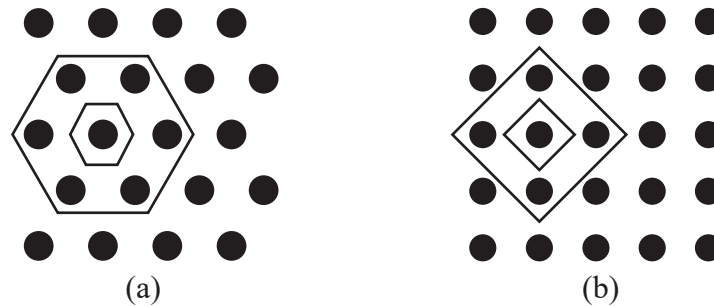


Figura 41. Arreglo de neuronas: (a) hexagonal. (b) rectangular.

El vector $\mathbf{x}=[x_1, x_2, \dots, x_n]^T \in \mathbf{R}^n$ es el vector de entrada que, por simplicidad y eficiencia computacional se considera conectado en paralelo a todas las neuronas m_i en la red, como muestra la Figura 42. El vector de peso de la celda i deberá, por tanto, estar denotado por $\mathbf{m}_i=[\mathbf{m}_{i1}, \mathbf{m}_{i2}, \dots, \mathbf{m}_{in}]^T \in \mathbf{R}^n$.

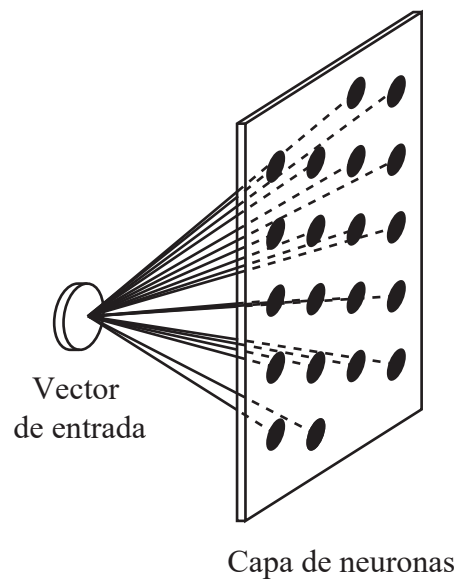


Figura 42. Conexión en paralelo del vector de entrada con las neuronas de la red.

Se requiere ahora encontrar una medida analítica, es decir, una medida de comparación entre el vector de entrada \mathbf{x} y el vector de pesos \mathbf{m}_i de las celdas. La medida analítica más simple para la correspondencia de \mathbf{x} con \mathbf{m}_i podría ser el producto interno $\mathbf{x}^T \mathbf{m}_i$, pero si el

algoritmo de auto-organización se empleará para relacionar patrones de señales naturales con espacios métricos vectoriales, entonces un criterio de correspondencia basado en distancias euclidianas entre \mathbf{x} y \mathbf{m}_i , resultará mejor y más conveniente. La distancia euclidiana d se define como la longitud de una línea recta entre dos puntos. En general, esta distancia queda definida por la ecuación

$$d(\mathbf{x}, \mathbf{y}) = \left(\sum_{i=1}^n (x_i - y_i)^2 \right)^{1/2} \quad (5.1)$$

donde n representa la dimensión del espacio. La distancia o diferencia mínima define a la celda ganadora, \mathbf{m}_c , que es la celda que aprenderá, o en otras palabras, adaptará sus pesos para parecerse más al vector de entrada.

Adaptación de los vectores de peso.

Una vez definida la celda ganadora, es crucial para la formación de los mapas ordenados en un arreglo que preserve la topología del sistema de entrada, que las celdas que ganan el aprendizaje, es decir, que presentan la diferencia mínima con respecto al vector de entrada, no se vean afectadas de manera independiente a todas las demás, sino que, como parte de un subconjunto relacionado topológicamente, Figura 43, se obtenga una corrección similar en cada celda de tal subconjunto.

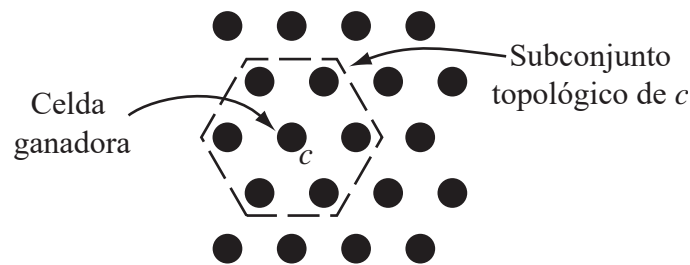


Figura 43. Subconjunto de relación de la neurona ganadora.

Durante el proceso de aprendizaje, cada subconjunto deberá abarcar varias celdas, y la corrección de la red en cada celda deberá reducirse en el largo plazo. Una característica importante de este tipo de aprendizaje es que los vectores de pesos tienden a obtener valores que se ordenan a lo largo de los ejes de la red.

En las redes neuronales el aprendizaje correlacionado en vecindarios espaciales de celdas se puede implementar usando varios tipos de *conexiones recurrentes laterales* y algunas otras interacciones laterales. Tales interacciones son en sí una función normalizada, que tiene como característica disminuir su valor conforme se aleja de un eje de simetría.

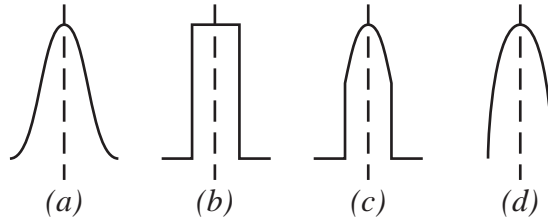


Figura 44. Funciones de relaciones laterales posibles para los subconjuntos topológicos.

El eje de la función se sitúa sobre la celda ganadora, y la magnitud de adaptación de las celdas vecinas se ve afectada por esta función. Puede verse que la celda ganadora se adapta en todos los casos completamente.

En esta red se desea forzar una interacción lateral directamente de manera general, para estructuras arbitrarias subyacentes de la red, mediante la definición de un *conjunto vecindario* N_c alrededor de c . Después de cada paso de aprendizaje, todas las celdas dentro de N_c se actualizan, mientras que celdas fuera del conjunto se dejan intactas. El vecindario se centra alrededor de la celda para la cual se obtuvo la mejor correspondencia con la entrada \mathbf{x} :

$$\|\mathbf{x} - \mathbf{m}_c\| = \min_i \{\|\mathbf{x} - \mathbf{m}_i\|\} \quad (5.2)$$

El ancho o radio de N_c puede variar en el tiempo. De hecho, para un buen ordenamiento global, ha mostrado ser ventajoso permitir que N_c sea ancho al principio y se contraiga monótonicamente con el tiempo, como lo muestra la Figura 45.

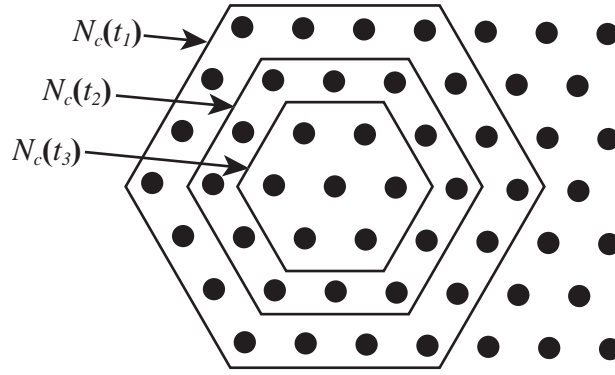


Figura 45. Ejemplo de vecindarios o subconjuntos topológicos $N_c(t)$, donde $t_1 < t_2 < t_3$.

La explicación para esto podría ser que un N_c inicialmente ancho correspondiente a una resolución espacial baja en el proceso de aprendizaje, induce primeramente un ordenamiento global “tosco” en los valores de \mathbf{m}_i , después de lo cual, reduciendo el tamaño de N_c , mejora la resolución espacial del mapa, sin destruir el ordenamiento global ya adquirido. Es posible incluso terminar el proceso con $N_c = 1$, es decir, actualizar únicamente la celda con mejor correspondencia, en cuyo caso, el proceso se reduce a un aprendizaje competitivo simple. Después de esto, sin embargo, se debe formar el “ordenamiento topológico” del mapa. Se puede explicar lo anterior empleando como símil el proceso de tramado de una foto, como lo muestra la Figura 46: al principio del proceso, la baja resolución significa un objeto irreconocible que en el caso del mapa sería una topología del sistema relativamente ajena a la del sistema de entrada, y conforme pasa el tiempo y aumenta la resolución, la imagen se vuelve más nítida, y en el mapa, la topología representa de manera más fiel aquella del sistema de entrada.

El proceso de actualización (en notación de tiempo discreto) puede ser:

$$\mathbf{m}_i(t+1) = \begin{cases} \mathbf{m}_i(t) + \alpha(t)[\mathbf{x}(t) - \mathbf{m}_i(t)] & \text{si } i \in N_c(t) \\ \mathbf{m}_i(t) & \text{si } i \notin N_c(t) \end{cases} \quad (5.3)$$

donde $\alpha(t)$ es la ganancia de adaptación, y $0 < \alpha(t) < 1$, y debe decrecer en el tiempo.

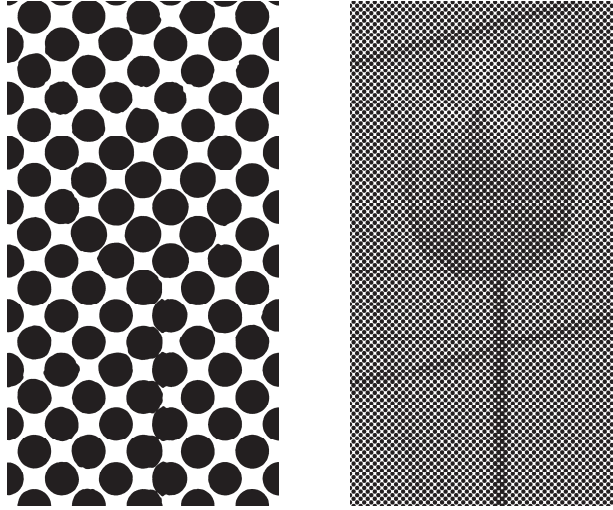


Figura 46. Efecto de una baja y una alta resolución.

Una notación alternativa es introducir un función *kernel* escalar $h_{ci}=h_{ci}(t)$

$$\mathbf{m}_i(t+1) = \mathbf{m}_i(t) + h_{ci}(t)[\mathbf{x}(t) - \mathbf{m}_i(t)] \quad (5.4)$$

por lo cual, $h_{ci}(t)=\alpha(t)$ dentro de N_c , y $h_{ci}(t)=0$ fuera de N_c . Por otro lado, la definición de h_{ci} , puede ser más general; una interacción lateral biológica tiene casi siempre la forma de una “curva de campana” como la mostrada en la Figura 44, inciso (a). Denotando las coordenadas de las celdas c e i por los vectores \mathbf{r}_c y \mathbf{r}_i , respectivamente, una forma apropiada para h_{ci} podría ser

$$h_{ci} = h_0 \exp(-\|\mathbf{r}_i - \mathbf{r}_c\|^2 / \sigma^2) \quad (5.5)$$

donde $h_0 = h_0(t)$ y $\sigma = \sigma(t)$ son funciones decrecientes en el tiempo adecuadas.

En ocasiones se sugiere que \mathbf{x} se normalice antes de emplearse en el algoritmo. La normalización, aunque no es necesaria en principio, puede mejorar la precisión numérica, dado que los vectores referencia resultantes tienden a tener el mismo rango dinámico.

Otro aspecto es la posibilidad de aplicar una medida general de distancia en la correspondencia, en cuyo caso, las reglas de correspondencia y de actualización deben ser mutuamente compatibles con respecto a la misma métrica. Por ejemplo, si se aplica la medida de producto interno de similitud, la ecuación de aprendizaje deberá ser

$$\mathbf{x}^T(t)\mathbf{m}_c(t) = \max_i \{\mathbf{x}^T(t)\mathbf{m}_i(t)\} \quad (5.6)$$

y

$$\mathbf{m}_i(t+1) = \begin{cases} \frac{\mathbf{m}_i(t) + \alpha'(t)\mathbf{x}(t)}{\|\mathbf{m}_i(t) + \alpha(t)\mathbf{x}(t)\|} & \text{si } i \in N_c(t) \\ \mathbf{m}_i(t) & \text{si } i \notin N_c(t) \end{cases} \quad (5.7)$$

con $0 < \alpha'(t) < \infty$, por ejemplo, $\alpha'(t) = 100/t$. Este proceso normaliza los vectores referencia a cada paso. El cálculo de la normalización vuelve significativamente lento el algoritmo de entrenamiento, pero por otro lado, el criterio de correspondencia lineal aplicado durante el reconocimiento es bastante simple y rápido, así como asequible a diversos tipos de cálculo computacional analógico.

Cuando el mapa auto-organizado se emplea como un *clasificador de patrones* en el cual las celdas o sus respuestas se agrupan en subconjuntos, cada uno correspondiente a una clase discreta de patrones, entonces el problema se vuelve un proceso de decidir a cuál subconjunto debe pertenecer el vector de entrada y debe tratarse de manera distinta. El mapa original está pensado principalmente para aproximar los valores de entrada, o su función de densidad probabilística, por medio de vectores cuantizados que se localizan en el espacio de entrada, para minimizar una función de error de cuantización. Por otro lado, si el conjunto de señales se clasificará en un número finito de categorías, entonces se requiere de diversos vectores código para representar cada clase, y su identidad dentro de la clase ya no es necesaria. De hecho, sólo decisiones realizadas en el borde de la clase tienen valor.

5.3. Ejemplos de clasificación con mapas auto-organizados.

Como primer ejemplo de clasificación con mapas auto-organizados se emplea una distribución uniforme de puntos aleatorios que definen un rectángulo.

Cada punto tiene como vector componente su coordenada, por tanto cada vector de peso del mapa será de dos dimensiones. La idea de estos ejemplos es ilustrar de forma simple en qué manera el mapa se adapta a las entradas.

Se generan 4320 puntos de manera aleatoria, y se define un mapa en un arreglo de 12×18 neuronas. Los puntos generados se muestran en la Figura 47. Los vectores de pesos del mapa se inicializan de manera aleatoria, quedando un mapa inicial como el que muestra la Figura 48. Las aristas representan la posición de la neurona.

Después del proceso de entrenamiento, el mapa debe haber alcanzado la topología del proceso de entrada, en este caso, una distribución regular sobre un área rectangular.

La Figura 49 muestra el mapa que se obtiene al final del entrenamiento. Puede observarse que el mapa se distribuye de manera uniforme y que prácticamente cubre toda el área que ocupan los puntos de entrada, por lo cual se concluye que el sistema ha aprendido el comportamiento del sistema y que de igual manera ha preservado la topología del mismo.

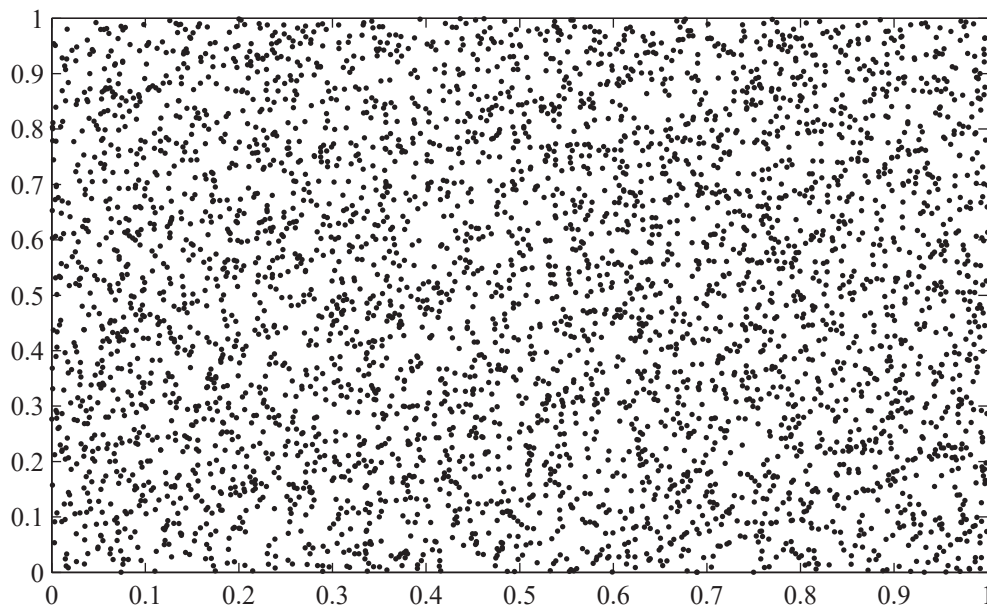


Figura 47. Puntos de entrada al mapa auto-organizado.

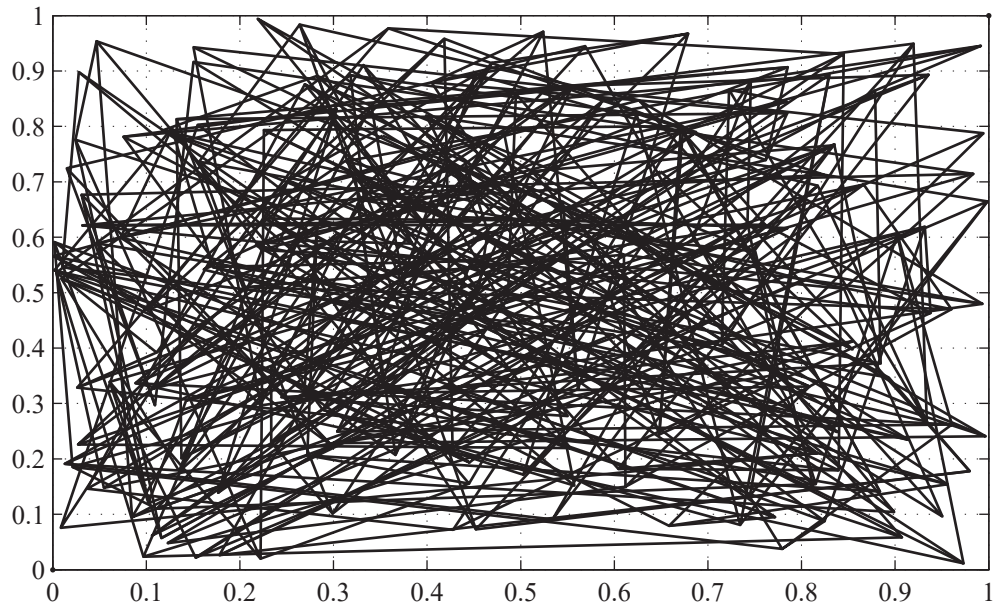


Figura 48. Mapa inicial.

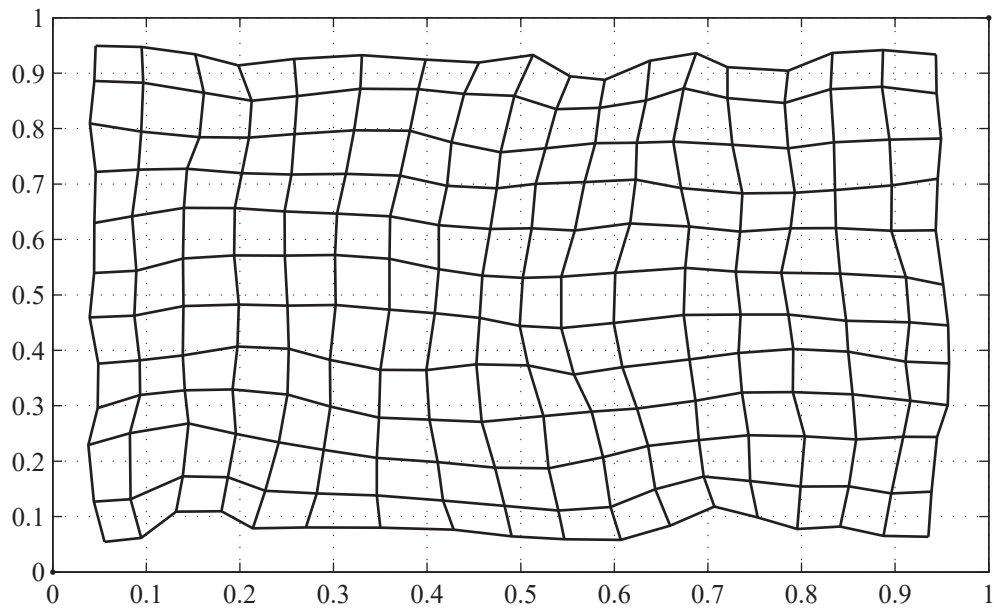


Figura 49. Mapa final.

Los mapas pueden ser de cualquier dimensión, y sin embargo, conservar la topología del sistema de entrada. Se empleará la misma distribución de puntos, pero empleando un mapa de una sola dimensión para ilustrar lo anterior.

Al igual que en el ejemplo anterior, la distribución de puntos es la que muestra la Figura 47.

El mapa se inicializa nuevamente con valores aleatorios, en un arreglo de 1×216 neuronas, el cual forma una curva abierta. Inicialmente las trayectorias del mapa se pueden cruzar, pero luego del entrenamiento, las neuronas tienden a separarse, evitando así cruces en la trayectoria que sigue la curva. Cada punto en la curva representa una neurona. Este mapa se muestra en la Figura 50.

Una vez realizado el entrenamiento se obtiene el mapa que se ilustra en la Figura 51. Al igual que en el caso del mapa bidimensional, este mapa trata de cubrir toda el área que está ocupada por puntos, preservando así la topología del sistema de entrada, a pesar de que el mapa es de menor dimensionalidad que el sistema de entrada.

Estos ejemplos simples ilustran de una manera clara la característica de preservación de topología que tienen los mapas auto-organizados.

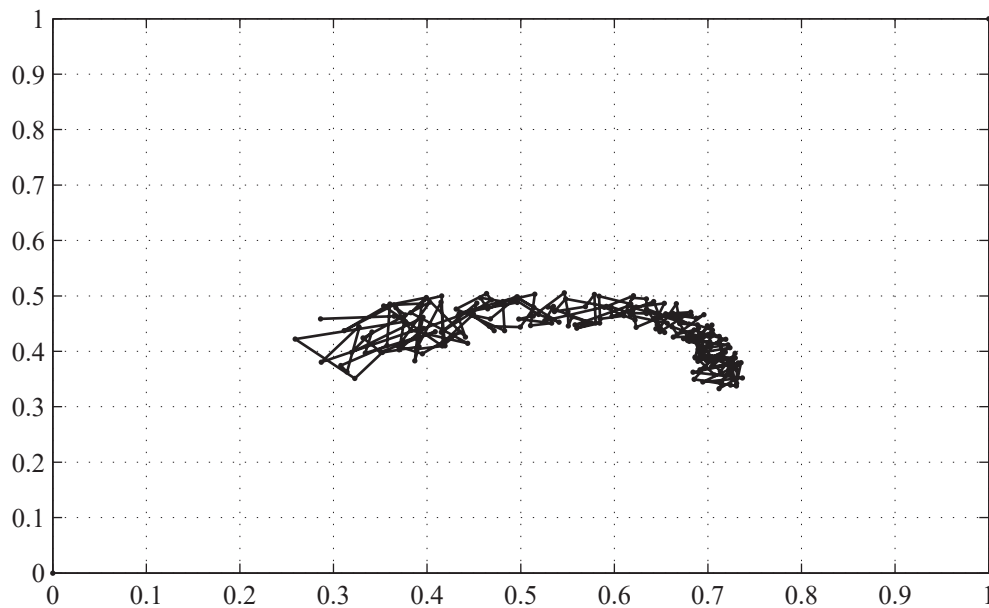


Figura 50. Mapa inicial de una sola dimensión.

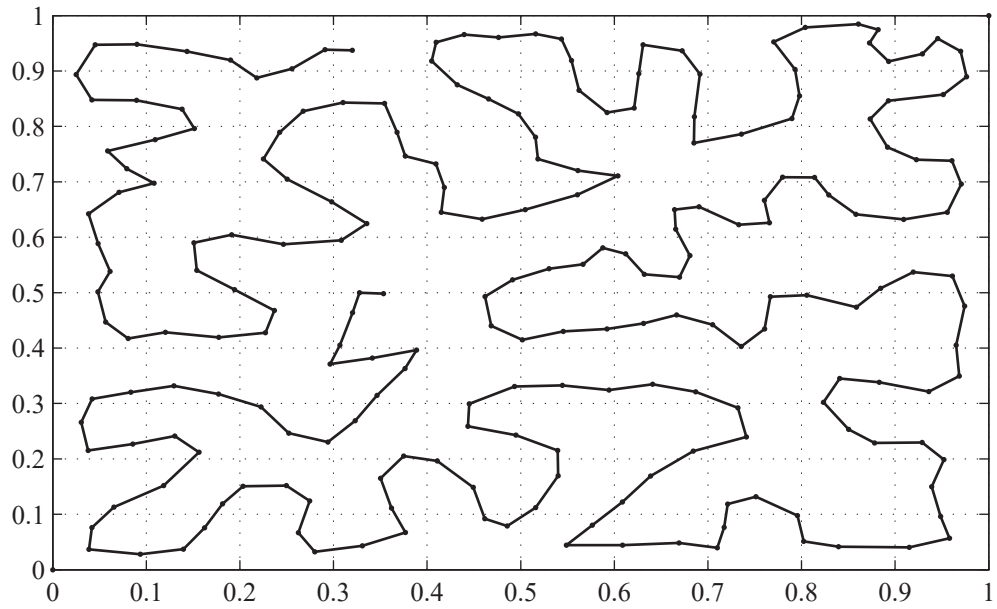


Figura 51. Mapa final de una sola dimensión.

Se ilustra ahora la característica de clasificación del mapa empleando un conjunto de datos con una clara separación entre ellos.

Los datos de entrenamiento son los que muestra la Figura 52.

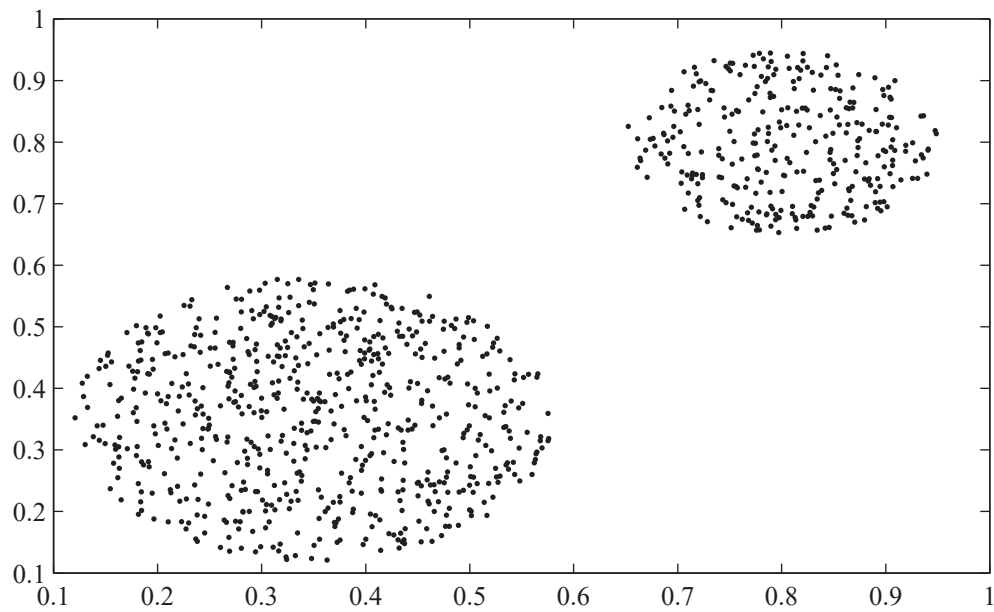


Figura 52. Puntos empleados para mostrar la característica de clasificador del mapa auto-organizado.

En la Figura 53 se muestra el mapa final obtenido del entrenamiento con los puntos del gráfico anterior. Cualquier punto que se presente al sistema posteriormente al entrenamiento puede fácilmente ser clasificado en uno u otro de los distintos cúmulos de datos. Nuevamente, los cruces de líneas representan neuronas del mapa.

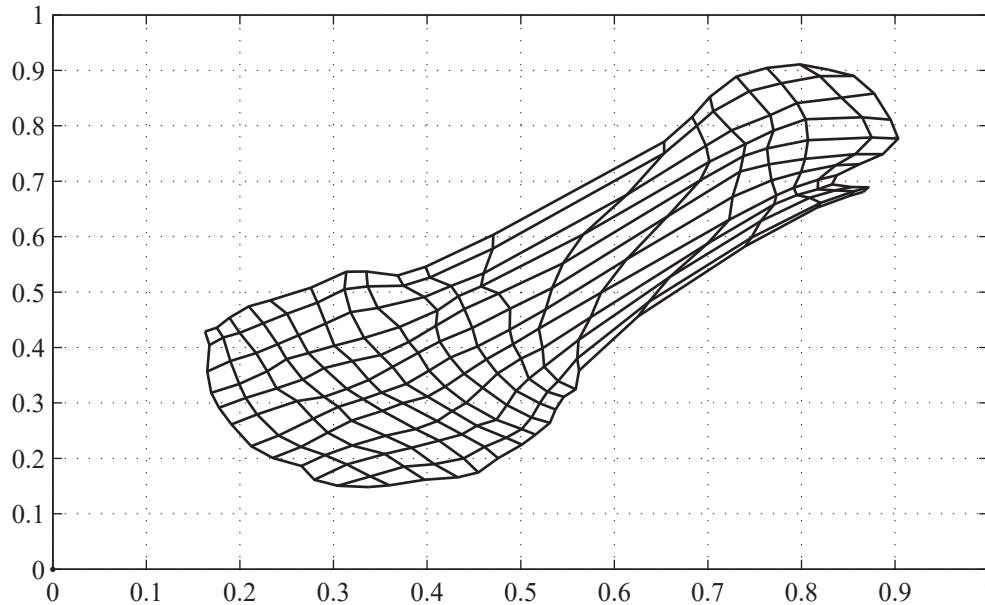


Figura 53. Mapa final.

Se pasa ahora a modelos más complejos y de mayor dimensión. Para ello se tiene un bien conocido conjunto de datos basados en diversas medidas de los pétalos y sépalos de Iris [Blake y Merz, 1998; Vesanto et al, 2000].

El conjunto de datos contiene tres clases, con cincuenta instancias cada una, y cada clase se refiere a un tipo específico de Iris.

Cada vector de entrada está formado por cuatro datos: largo y ancho de pétalos y largo y ancho de sépalos.

La siguiente tabla muestra los datos empleados:

Setosa				Versicolor				Virginica			
Sépalo		Pétalo		Sépalo		Pétalo		Sépalo		Pétalo	
Largo	Ancho	Largo	Ancho	Largo	Ancho	Largo	Ancho	Largo	Ancho	Largo	Ancho
5.1	3.5	1.4	0.2	7	3.2	4.7	1.4	6.3	3.3	6	2.5
4.9	3	1.4	0.2	6.4	3.2	4.5	1.5	5.8	2.7	5.1	1.9
4.7	3.2	1.3	0.2	6.9	3.1	4.9	1.5	7.1	3	5.9	2.1
4.6	3.1	1.5	0.2	5.5	2.3	4	1.3	6.3	2.9	5.6	1.8
5	3.6	1.4	0.2	6.5	2.8	4.6	1.5	6.5	3	5.8	2.2
5.4	3.9	1.7	0.4	5.7	2.8	4.5	1.3	7.6	3	6.6	2.1
4.6	3.4	1.4	0.3	6.3	3.3	4.7	1.6	4.9	2.5	4.5	1.7
5	3.4	1.5	0.2	4.9	2.4	3.3	1	7.3	2.9	6.3	1.8
4.4	2.9	1.4	0.2	6.6	2.9	4.6	1.3	6.7	2.5	5.8	1.8
4.9	3.1	1.5	0.1	5.2	2.7	3.9	1.4	7.2	3.6	6.1	2.5
5.4	3.7	1.5	0.2	5	2	3.5	1	6.5	3.2	5.1	2
4.8	3.4	1.6	0.2	5.9	3	4.2	1.5	6.4	2.7	5.3	1.9
4.8	3	1.4	0.1	6	2.2	4	1	6.8	3	5.5	2.1
4.3	3	1.1	0.1	6.1	2.9	4.7	1.4	5.7	2.5	5	2
5.8	4	1.2	0.2	5.6	2.9	3.6	1.3	5.8	2.8	5.1	2.4
5.7	4.4	1.5	0.4	6.7	3.1	4.4	1.4	6.4	3.2	5.3	2.3
5.4	3.9	1.3	0.4	5.6	3	4.5	1.5	6.5	3	5.5	1.8
5.1	3.5	1.4	0.3	5.8	2.7	4.1	1	7.7	3.8	6.7	2.2
5.7	3.8	1.7	0.3	6.2	2.2	4.5	1.5	7.7	2.6	6.9	2.3
5.1	3.8	1.5	0.3	5.6	2.5	3.9	1.1	6	2.2	5	1.5
5.4	3.4	1.7	0.2	5.9	3.2	4.8	1.8	6.9	3.2	5.7	2.3
5.1	3.7	1.5	0.4	6.1	2.8	4	1.3	5.6	2.8	4.9	2
4.6	3.6	1	0.2	6.3	2.5	4.9	1.5	7.7	2.8	6.7	2
5.1	3.3	1.7	0.5	6.1	2.8	4.7	1.2	6.3	2.7	4.9	1.8
4.8	3.4	1.9	0.2	6.4	2.9	4.3	1.3	6.7	3.3	5.7	2.1
5	3	1.6	0.2	6.6	3	4.4	1.4	7.2	3.2	6	1.8
5	3.4	1.6	0.4	6.8	2.8	4.8	1.4	6.2	2.8	4.8	1.8
5.2	3.5	1.5	0.2	6.7	3	5	1.7	6.1	3	4.9	1.8
5.2	3.4	1.4	0.2	6	2.9	4.5	1.5	6.4	2.8	5.6	2.1
4.7	3.2	1.6	0.2	5.7	2.6	3.5	1	7.2	3	5.8	1.6
4.8	3.1	1.6	0.2	5.5	2.4	3.8	1.1	7.4	2.8	6.1	1.9
5.4	3.4	1.5	0.4	5.5	2.4	3.7	1	7.9	3.8	6.4	2
5.2	4.1	1.5	0.1	5.8	2.7	3.9	1.2	6.4	2.8	5.6	2.2
5.5	4.2	1.4	0.2	6	2.7	5.1	1.6	6.3	2.8	5.1	1.5
4.9	3.1	1.5	0.1	5.4	3	4.5	1.5	6.1	2.6	5.6	1.4
5	3.2	1.2	0.2	6	3.4	4.5	1.6	7.7	3	6.1	2.3
5.5	3.5	1.3	0.2	6.7	3.1	4.7	1.5	6.3	3.4	5.6	2.4
4.9	3.1	1.5	0.1	6.3	2.3	4.4	1.3	6.4	3.1	5.5	1.8
4.4	3	1.3	0.2	5.6	3	4.1	1.3	6	3	4.8	1.8
5.1	3.4	1.5	0.2	5.5	2.5	4	1.3	6.9	3.1	5.4	2.1
5	3.5	1.3	0.3	5.5	2.6	4.4	1.2	6.7	3.1	5.6	2.4

4.5	2.3	1.3	0.3	6.1	3	4.6	1.4	6.9	3.1	5.1	2.3
4.4	3.2	1.3	0.2	5.8	2.6	4	1.2	5.8	2.7	5.1	1.9
5	3.5	1.6	0.6	5	2.3	3.3	1	6.8	3.2	5.9	2.3
5.1	3.8	1.9	0.4	5.6	2.7	4.2	1.3	6.7	3.3	5.7	2.5
4.8	3	1.4	0.3	5.7	3	4.2	1.2	6.7	3	5.2	2.3
5.1	3.8	1.6	0.2	5.7	2.9	4.2	1.3	6.3	2.5	5	1.9
4.6	3.2	1.4	0.2	6.2	2.9	4.3	1.3	6.5	3	5.2	2
5.3	3.7	1.5	0.2	5.1	2.5	3	1.1	6.2	3.4	5.4	2.3
5	3.3	1.4	0.2	5.7	2.8	4.1	1.3	5.9	3	5.1	1.8

Tabla 2. (continuación) Datos empleados en la clasificación de diferentes especies de Iris.

Los datos anteriores están dados en centímetros.

El primer paso consiste en normalizar los valores. Para el presente caso, se realiza una normalización de la varianza de los datos, lo cual representa un proceso lineal. Una vez llevado a cabo, es necesario construir la topología de la red. Para este ejemplo se eligió una red con 78 neuronas hexagonales, (valor arbitrario) de “vecindario gaussiano”, e inicializadas aleatoriamente.

Después de un proceso de aprendizaje rápido, y un subsecuente refinamiento, se obtienen los mapas mostrados en la Figura 54 y la Figura 55.

La *matriz de distancia unificada* de la red (U-matrix) es una representación visual de distancia entre las neuronas. La tonalidad de las celdas indica distancias relativas, siendo un color claro, o valores altos, una distancia grande y por tanto una separación entre clases y un color oscuro o valores pequeños, una distancia pequeña, y por tanto, datos pertenecientes a la misma clase.

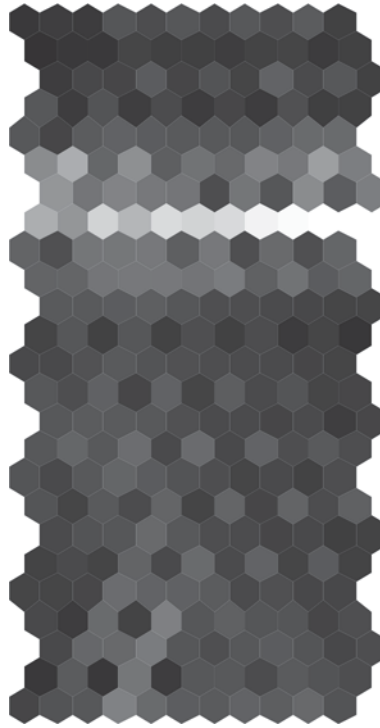


Figura 54. Matriz de distancia unificada de la red.(U-Matrix)

Este mapa facilita la inspección visual de las clases en ausencia de conocimientos previos de los datos, permitiendo el descubrimiento de estructuras subyacentes en el sistema. La Figura 54 muestra un cúmulo claramente definido en la parte superior del mapa. Se puede observar en la esquina superior izquierda un grupo de celdas oscuras, lo cual indica que hay un pequeño cúmulo en ese lugar y su significado es que existe una gran uniformidad en los datos para esa especie. Es precisamente este tipo de estructuras las que permite descubrir esta representación. Otra característica importante de la U-matrix es que tiene un mayor número de celdas que la topología de la red, ya que cada una de ellas muestra la distancia entre celdas vecinas.

La representación de *planos componentes* muestra la función de distribución relativa de los datos de entrada. Cada plano muestra la distribución relativa de un vector componente o variable de entrada. Mediante la comparación de los planos es relativamente simple determinar qué atributos están correlacionados si ellos son visualmente similares, y el grado de la similitud define la fuerza de tal correlación.

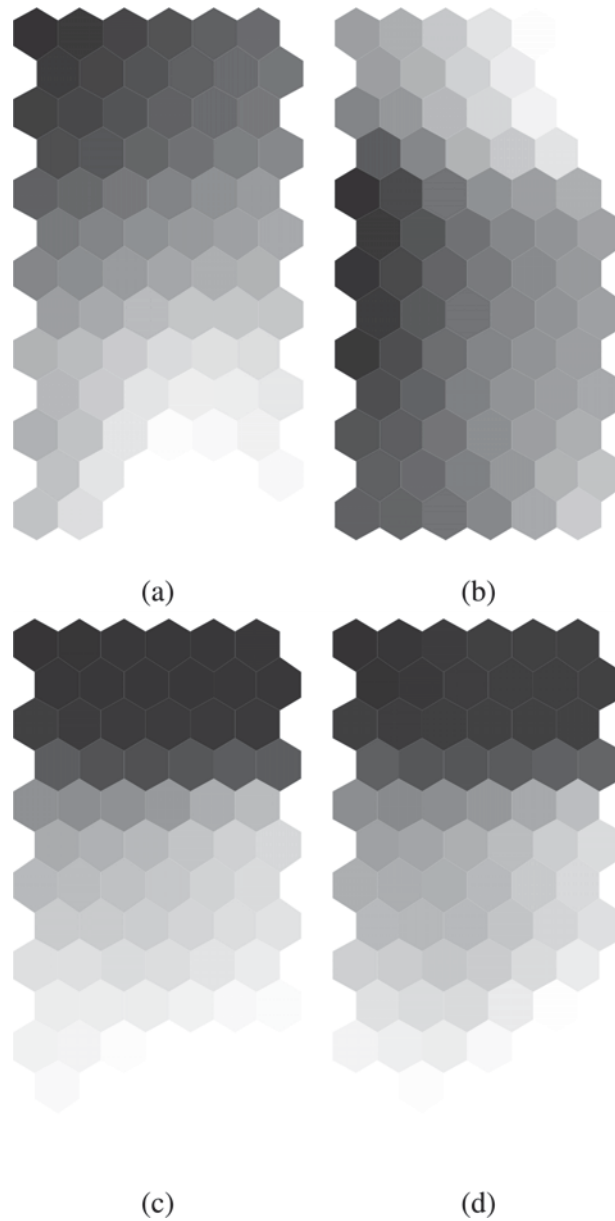


Figura 55. Planos componentes del mapa auto organizado del ejemplo

La Figura 55, incisos (a) y (b) representan los valores del largo y ancho del sépalo, y los incisos (c) y (d), el largo y ancho del pétalo. La similitud de estos últimos evidencia una clara correlación entre estas medidas, y una correlación mucho menor entre las primeras.

Existen muy variadas representaciones de los datos, por ejemplo, se pueden utilizar las etiquetas de entrada del sistema para determinar características que las otras representaciones no aclaran. La U-matrix mostrada en la Figura 54, muestra dos cúmulos:

el superior, perteneciente a la sub-especie *Setosa* y en la parte inferior indistintamente los otros dos. Utilizando la representación de etiquetas, Figura 56, es posible visualizar los cúmulos que no muestra la representación de U-matrix.

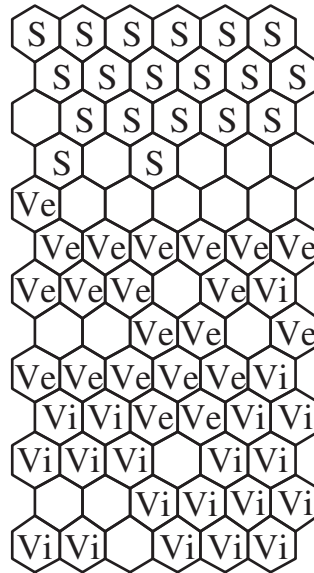


Figura 56. Mapa con etiquetas.

Los planos componentes son convenientes cuando es necesario visualizar muchos datos a la vez. Pero cuando el número de variables de interés es pequeño, una manera más conveniente de visualización la constituyen los gráficos de dispersidad. La Figura 57 muestra la proyección de los vectores de datos en el sub-espacio del mapa, así como el arreglo final del mapa.

Dado que la dimensión de entrada del sistema es mayor a tres, el gráfico de dispersidad es en realidad una proyección a una menor dimensión de los vectores de entrada. Se emplea el análisis de componentes principales para realizar la proyección de los datos. Este método está basado en el análisis de eigenvalues, también conocido como transformada de Karhunen-Loeve.

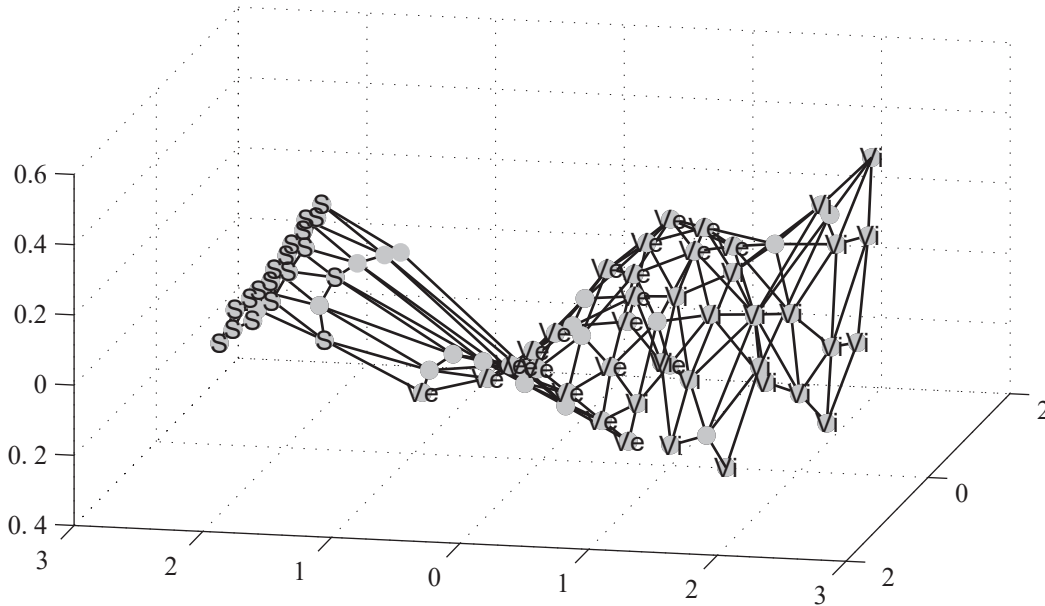


Figura 57. Proyección de los datos en el mapa.

Otro ejemplo de clasificación se realiza empleando las series de tiempo generadas por el sistema de ecuaciones de Lorenz (3.1). La serie se generó con los valores iniciales de $x=10$, $y=28$ y $z=8/3$, a lo largo de un tiempo de 10 segundos, a intervalos de centésimo de segundo.

La topología del mapa es hexagonal con 150 celdas, en un arreglo de 10×15 , con valores normalizados y una inicialización lineal del mapa.

Utilizando los valores obtenidos para las variables x , y y z , se obtiene la U-matrix de la Figura 58, y los planos componentes (que indican cada una de las tres variables) mostrados en la Figura 59. Ambas representaciones resultan en relativamente poca información sobre el sistema a primera vista. Sin embargo, la proyección del mapa muestra aspectos interesantes del comportamiento del sistema.

La Figura 60 muestra una proyección en dos dimensiones del mapa. En él se puede observar una estructura que recuerda la clásica proyección de “mariposa” mostrada en la Figura 5 y que se repite por comodidad en la Figura 61.

El mapa trata de recrear la trayectoria que sigue el sistema, concentrando neuronas en la región central, que es donde se dan los cruces de la trayectoria de un régimen a otro.

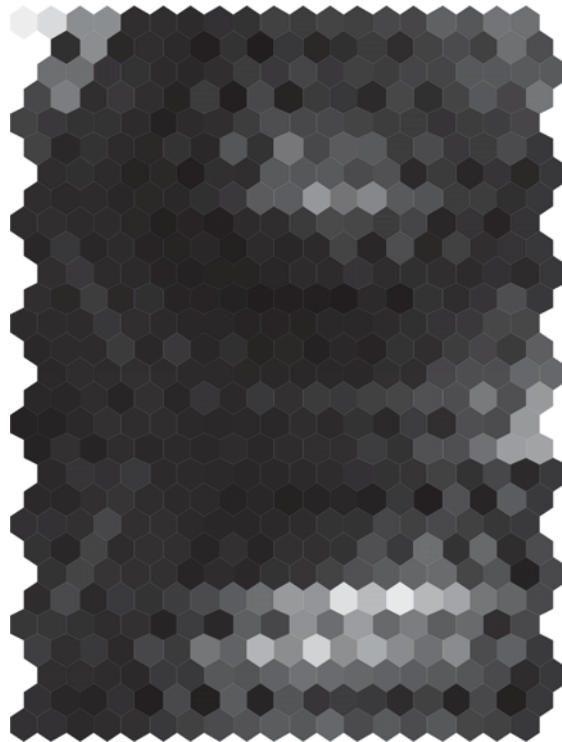


Figura 58. U-matrix para la serie de tiempo de Lorenz.

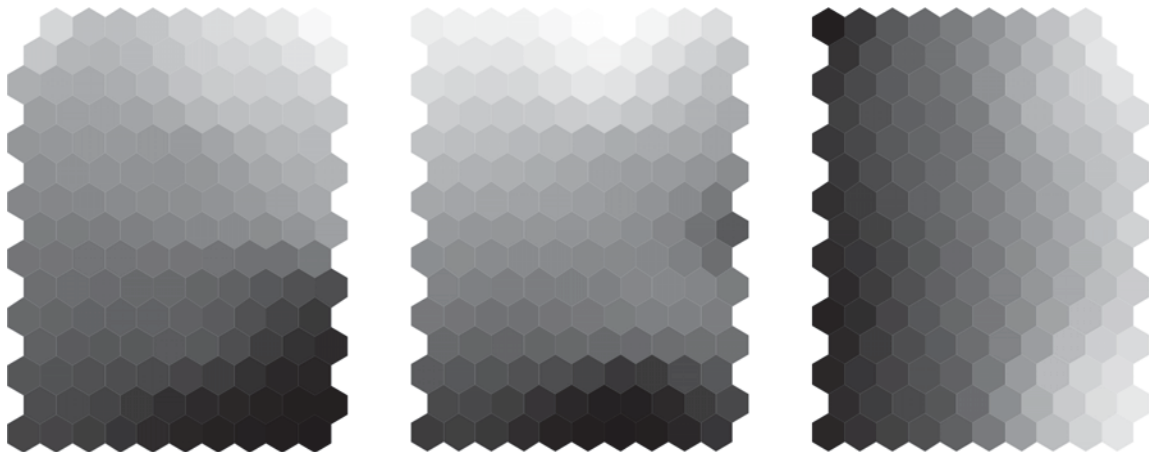


Figura 59. Planos componentes para la serie de tiempo de Lorenz

Se pueden observar “ventanas” o zonas de pocas neuronas hacia lo que sería el interior de las trayectorias que sigue cada régimen de oscilación. En base a esto, es posible observar que la U-Matrix presenta una característica similar, con regiones de celdas de color claro, que representan los bordes de los cúmulos, o las ventanas antes mencionadas, rodeadas de

celdas de color oscuro, que formarían parte del cúmulo, y por tanto, las regiones más concurridas de la trayectoria del sistema.

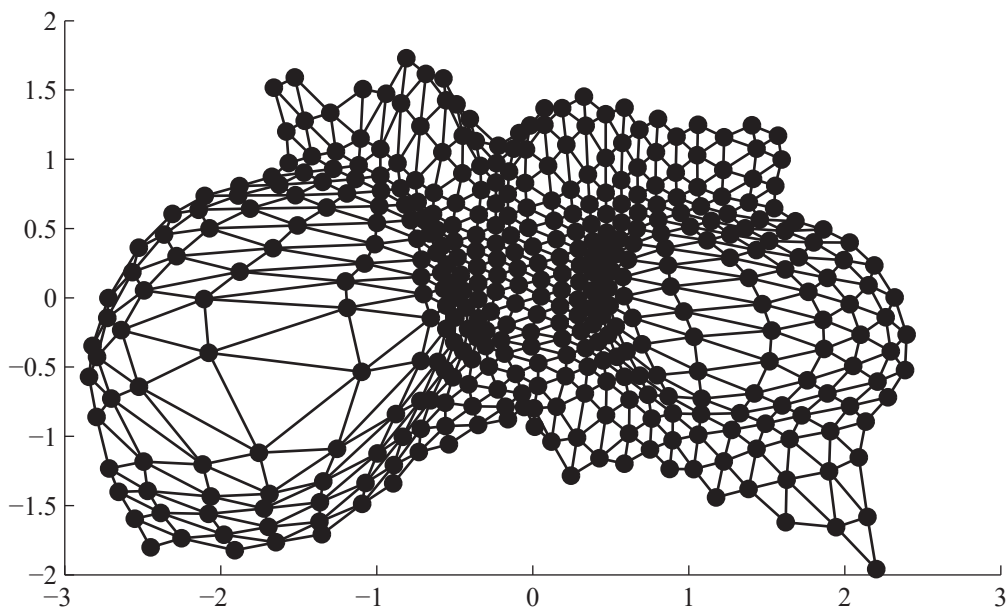


Figura 60. Proyección de los datos del mapa para la serie de tiempo de Lorenz.

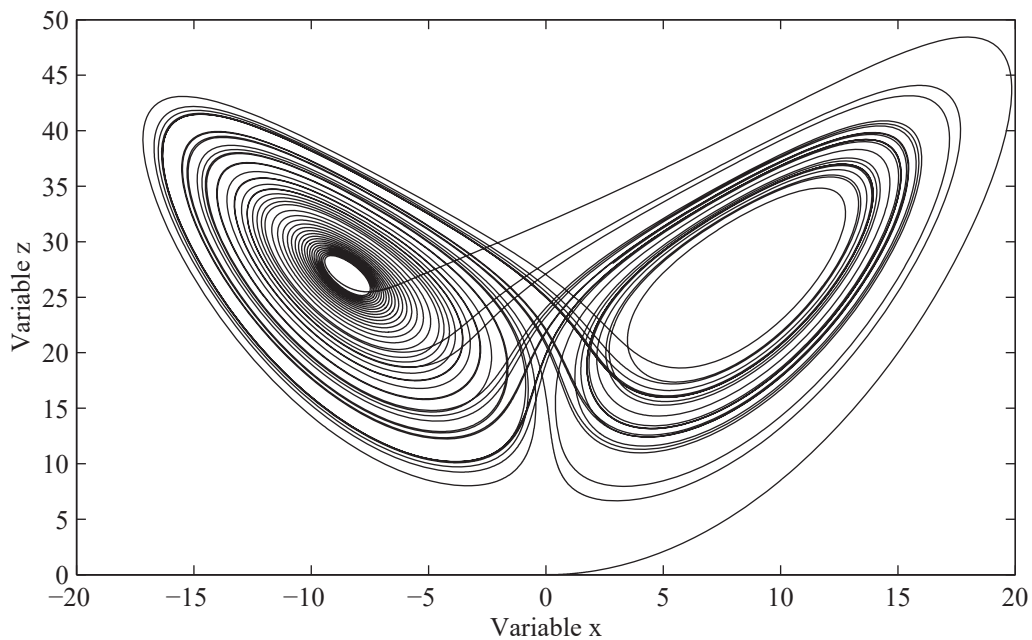


Figura 61. Atractor de Lorenz, proyección en dos dimensiones.

Para este caso, los planos componentes señalan una muy ligera correlación entre las variables, y por tanto representan buenas características para su uso en la predicción del sistema, de acuerdo a la metodología del esquema predictivo que se tiene como objetivo en esta tesis.

5.4. Conclusiones

Los mapas auto-organizados representan una excelente herramienta de análisis visual de sistemas tanto de baja dimensión como multidimensionales.

Resultan innumerables las características que se pueden obtener a partir de la simple inspección, y si tales observaciones son reforzadas por metodologías ya establecidas como el análisis de componentes principales o el agrupamiento de celdas mediante el algoritmo de los k vecinos cercanos, por citar algunos, el estudio que se logra hacer a partir de únicamente los resultados obtenidos del mapa, es suficiente para lograr una gran cantidad de información del comportamiento de cualquier sistema.

En este capítulo se revisó el algoritmo de aprendizaje y adaptación de los mapas auto-organizados. Se realizaron ejemplos que permiten observar sus principales características como clasificador de patrones.

La característica del mapa como preservador de la topología del sistema quedó bien establecida mediante ejemplos sencillos.

6. Predicción de series de tiempo caóticas

6.1. Introducción

En este capítulo se aborda el objetivo principal de la tesis: Presentar un esquema de predicción para sistemas caóticos, empleando el paradigma de modelos locales y mapas auto-organizados como clasificadores. Se introducen conceptos esenciales para lograr el objetivo, como lo es el teorema de Takens.

Se describen de una manera formal los pasos requeridos para llevar a cabo el proceso de predicción y se realizan pruebas para determinar las cualidades del esquema propuesto. Así mismo, se pretende definir lineamientos para un uso eficaz de tal esquema.

6.2. Teorema de Takens

Uno de los conceptos más importantes desarrollados en la teoría de sistemas dinámicos no lineales es el hecho que dado un sistema determinístico de la forma

$$x_{t+1} = F(x_t; \mu) \tag{6.1}$$

no es necesario observar el vector de estado x completo, sino únicamente aquellas observaciones libres de ruido de uno de sus componentes, la cual será suficiente para reconstruir el sistema dinámico entero.

El teorema de Takens [Takens, 1981] especifica que dado un sistema dinámico de tiempo continuo con un vecindario topológico diferenciable, compacto e invariante A , tal que A :

- Es de dimensión d_A
- Contiene sólo un número finito de puntos de equilibrio
- No contiene órbitas periódicas de periodo τ ó 2τ
- Contiene solamente un número finito de órbitas periódicas $p\tau$, $3 \leq p \leq m$

y si los Jacobianos del mapa de retorno de estas órbitas periódicas tienen distintos valores propios, entonces, con probabilidad de uno, una función de medida $h(x)$ proporcionará una función de coordenadas de atraso $\mathbf{H}=(h(x_\tau), h(x_{t-\tau}), \dots, h(x_{t-(m-1)\tau}))$, la cual es una envolvente diferenciable de A hacia $\mathbf{H}(A)$ para $m > 2d_A$. [Sauer et al, 1991]. El significado de esto, es que para casi cualquier función escalar $h(x)$, es decir, algún conjunto de observaciones, se pueden responder una amplia gama de interrogantes sobre la dinámica del sistema original, mediante la observación de un espacio definido por valores retrasados de $h(x)$. Es posible calcular entonces las invariantes del sistema, e incluso, mediante la interpolación del espacio de envolvente de atraso, se pueden realizar predicciones del comportamiento del sistema.

El objetivo de este teorema es obtener un resultado genérico, por tanto, proporciona únicamente modelos aproximados que pueden resultar útiles, y es en sí, una justificación del uso de para la envolvente de atraso.

6.3. Modelado dinámico: Multi-modelos

La tarea del modelado dinámico no lineal es aproximar el flujo reconstruido $F(x)$ en

$$x(n+1) = F(x(n)) \tag{6.2}$$

que involucra la trayectoria en el espacio reconstruido $x(n) \rightarrow x(n+1)$. Dado que no existe una manera algorítmica de obtener $F(\cdot)$, se emplea el marco de aproximación de funciones.

Se pueden tomar dos alternativas en la manera en que se realiza la aproximación del modelo:

- Modelado global
- Modelado local

Modelado Global

En el modelado global se emplea una sola función de aproximación, $\tilde{F}(\cdot)$, para aproximar todos los puntos en el espacio de reconstrucción, es decir $D=R^N$, donde D es algún dominio de aproximación.

Los modelos globales han sido investigados exhaustivamente en su utilización para modelado dinámico no lineal. Entre los modelos globales se pueden mencionar aquellos de aproximación polinomial, redes neuronales, y *funciones de base radial*.

Modelado Local

El modelado local o multimodelos, divide el espacio de estado en regiones locales y modela individualmente la dinámica de cada región. Una ventaja del modelado local sobre el global es la necesidad de un menor número de suposiciones estadísticas y geométricas acerca de los datos, lo cual lo hace a su vez, un modelo más general. Cada modelo local $\tilde{F}_r(\cdot)$, para $r=1, \dots, R$, es potencialmente más simple, siendo común el uso de modelos lineales, pero los parámetros deben cambiar a lo largo del espacio de estado.

Los modelos locales requieren de un conjunto de observaciones tanto de las variables observadas como de la salida del sistema. En este momento se puede construir el modelo y emplearlo para estimar la salida del sistema ante una nueva entrada q , denominada *interrogante (query)*.

La suposición que se hace en este modelo es que el sistema es localmente diferenciable.

Todo el proceso posterior del sistema se pospone hasta que el modelo es *interrogado*. En este punto, el proceso se puede realizar en tres pasos:

- Determinar k vecinos cercanos al punto de observación.

-
- Construir un modelo empleando únicamente los k vecinos cercanos.
 - Evaluar el modelo usando el vector de interrogantes como entrada, para estimar la salida del sistema.

El modelo global de la dinámica del sistema se obtiene al conjuntar todos los modelos locales de acuerdo a la siguiente ecuación:

$$\tilde{F}(x(n)) = \bigcup_{r=1, \dots, R} \tilde{F}_r(x(n)) \quad (6.3)$$

6.4. Casos de estudio

El uso del esquema de predicción que se propone en la tesis se ejemplifica utilizando los siguientes sistemas:

- Lorenz
- Chua
- Chen
- Rössler
- Duffing

De todos ellos se ha hecho una revisión en capítulos anteriores.

6.5. Estimación del error

Para la validación de los modelos empleados para las pruebas se empleó el *método de mantenimiento de validación cruzada*, que consiste en separar el conjunto de datos generados en dos subconjuntos, un conjunto inicial de *entrenamiento* y uno final de *prueba*. El error entre el conjunto de prueba y la predicción realizada por el sistema se mide utilizando la fórmula del *error medio cuadrático normalizado*:

$$EMCN \square \frac{1}{\sigma_y^2} \frac{1}{n_c} \sum_{i=1}^{n_c} (y_{v(i)+\tau} - \hat{g}_\gamma(\hat{x}_{v(i)+t-1}))^2 \quad (6.4)$$

donde σ_y^2 es la variancia de la serie de tiempo, $y_{v(i)+\tau}$ es el $(v(i)+\tau)$ -ésimo elemento de la serie de tiempo, $v(i)$ es el índice del primer elemento en el conjunto de prueba y $\hat{g}_y(\hat{x}_{v(i)+t-1})$ es la predicción del sistema.

Un estimador alternativo es la raíz del error medio cuadrático normalizado definido simplemente como:

$$REMCN = \sqrt{EMCN} \quad (6.5)$$

Dos características importantes de este error son su independencia a la escala de la serie y su fácil interpretación, ya que un valor de $REMCN=1$ significa que la aproximación no es mejor que calcular la media muestral de la serie, en promedio.

6.6. Ejemplos de predicción de series de tiempo

En esta sección se muestran ejemplos de la predicción realizada bajo el esquema que presenta esta tesis. Para cada ejemplo se muestran gráficos de la serie, una comparación del conjunto de prueba y la predicción del sistema y error medio cuadrático normalizado asociado con la predicción.

La Tabla 3 muestra la simbología empleada para los diversos parámetros del proceso de predicción.

Símbolo	Descripción
m	Dimensión envolvente.
d	Paso de retraso.
n_c	Número de neuronas de la red de Kohonen
n_d	Número de datos.
n_e	Número de datos para entrenamiento.
n_p	Número de datos de prueba
n_i	Número de pasos iterados

Tabla 3.

El proceso de predicción se puede esquematizar para facilitar su entendimiento. La Figura 62 muestra los pasos del proceso de reconstrucción del sistema.

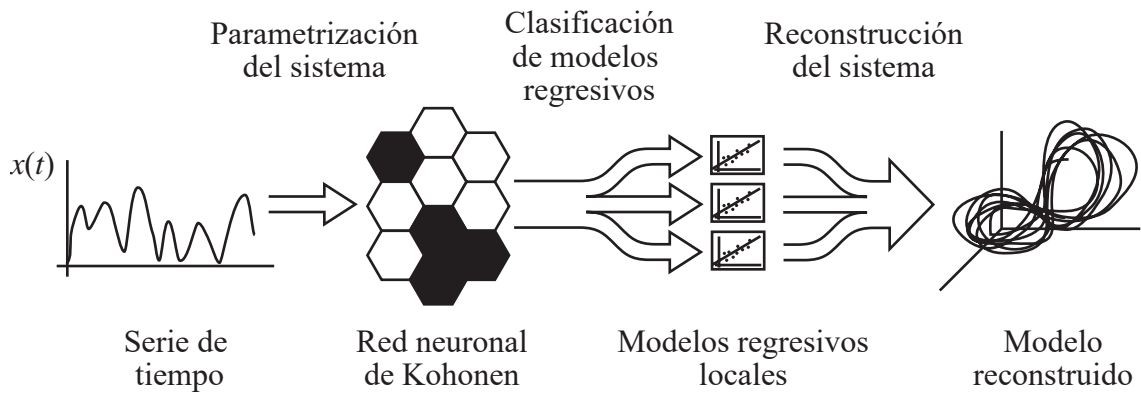


Figura 62. Esquema del proceso de reconstrucción de sistemas.

El primer paso consiste en *parametrizar* el sistema. Dada una serie de tiempo, se determinan los factores de envolvente, m , y el paso de retraso que permita obtener la máxima información del sistema, τ , mediante los métodos de información mutua promedio y de vecinos falsos, mencionados en el capítulo 3. Una vez obtenidos, la serie de tiempo se acondiciona para formar un arreglo que cumpla con las características que especifican tales parámetros. En el caso de un factor de envolvente $m=3$, y un paso de retraso $\tau=10$, se tendrá una matriz de datos de dimensiones $[n_d - (\tau \times m)] \times (m)$, ordenados de la siguiente manera:

$$\begin{array}{ccc}
 d_1 & d_{11} & d_{21} \\
 d_2 & d_{12} & d_{22} \\
 d_3 & d_{13} & d_{23} \\
 d_4 & d_{14} & d_{24} \\
 d_5 & d_{15} & d_{25} \\
 d_6 & d_{16} & d_{26} \\
 d_7 & d_{17} & d_{27} \\
 d_8 & d_{18} & d_{28} \\
 d_9 & d_{19} & d_{29} \\
 d_{10} & d_{20} & d_{30} \\
 d_{11} & d_{21} & d_{31} \\
 d_{12} & d_{22} & d_{32} \\
 d_{13} & d_{23} & d_{33} \\
 d_{14} & d_{24} & d_{34} \\
 \bullet & \bullet & \bullet \\
 \bullet & \bullet & \bullet \\
 \bullet & \bullet & \bullet
 \end{array}$$

En el siguiente paso, la matriz anterior formará el conjunto de datos de entrenamiento de la red de Kohonen, la cual crea una clasificación de las secuencias que le son

presentadas. Cada renglón es un vector de entrada, y queda asociado con aquella neurona que resulta ganadora ante esa entrada. Todos los vectores asociados al final del entrenamiento con una neurona, forman posteriormente la matriz \mathbf{X} en el proceso de mínimos cuadrados, que es el equivalente a definir los k vecinos cercanos del modelo local, para la determinación de los coeficientes de regresión que recrean la dinámica local del sistema.

De acuerdo con lo anteriormente expuesto, el número de celdas de la red de Kohonen determina asimismo el número de modelos locales del sistema, por tal motivo es un factor importante en el desempeño del sistema de predicción. Sin embargo, no hay una metodología que especifique un número adecuado de neuronas. Un número grande de neuronas significa un proceso de entrenamiento lento y un gran número de modelos locales, sin que ello signifique una verdadera ganancia de información del sistema a predecir. De igual manera, se requiere un mayor número de datos de entrenamiento, dado que un número muy pequeño de vectores asociados con cada modelo trae consigo la posibilidad de mal condicionamiento de la matriz \mathbf{X} , en el método de mínimos cuadrados.

El caso contrario, cuando se tienen pocas neuronas, se acerca al caso del modelado global, con la consecuente pérdida de la información en cuanto a regímenes de comportamiento del sistema, así como de generalidad del modelo.

Dada la inexistencia de metodología para optimizar el número de celdas de la red de Kohonen, se realizan pruebas con diversos números de celdas. Para todos los casos de estudio, se observó un valor de equilibrio alrededor de 100 celdas. Valores superiores a 200 celdas presentaron condicionamiento de las matrices de regresión en todos los casos.

Una vez determinados los k vecinos cercanos, se construyen los modelos regresivos empleando el método de mínimos cuadrados ponderados.

Una característica de los modelos locales es su discontinuidad. A fin de resolver esto, es necesaria una función de ponderación tal que decremente la influencia de los vecinos más alejados dentro de un mismo vecindario. La fórmula (4.12) define el método de mínimos cuadrados ponderados, donde \mathbf{w} representa la matriz de pesos. Y es precisamente el valor de \mathbf{w} lo que nos permite realizar el decremento en la influencia de un vecino lejano.

El modelado local es insensible a la función de ponderación que se emplee, en tanto ésta presente características de *no negatividad*, *monotonidad decreciente* y *suavidad*. Para el modelo regresivo utilizado en el método de predicción, se empleó la función ponderativa de la fórmula (6.6), denominada *bi-peso*.

$$BW(d_i, d_{k+1}) \square \begin{cases} \left(1 - \left(\frac{d_i^2}{d_{k+1}^2}\right)\right)^2 & d_i \leq d_{k+1} \\ 0 & \text{Cualquier otro caso} \end{cases} \quad (6.6)$$

donde d_i es la distancia euclidiana al i -ésimo vecino.

El tercer paso consiste en interrogar al sistema mediante los vectores de prueba, con la finalidad de formar el modelo completo del sistema, de acuerdo con la fórmula (6.3). En este paso se pueden definir diferentes modalidades: interrogar al modelo en un *paso*, actualizando los datos de interrogación con el valor real una vez que se ha obtenido, o interrogarlo iterativamente, empleando únicamente los valores predichos por el propio modelo.

En la modalidad de interrogación por paso, se predice un solo *paso adelante* (considerando un *paso adelante* de acuerdo al modelo regresivo que se emplee), y una vez que el valor real del sistema se obtiene, se integra al modelo regresivo para realizar la predicción sobre el siguiente valor. Esta modalidad es ventajosa en el hecho de que mantiene un error muy pequeño a lo largo de la predicción, sin embargo, su horizonte predictivo es de muy corto plazo.

En la modalidad de predicción iterada, el valor predicho se vuelve parte del modelo regresivo, de tal manera que cuando el número de predicciones es igual al paso de retraso, cada nueva predicción se realizará únicamente con valores predichos anteriormente y ningún valor real de la serie de tiempo. Es lógico suponer que dado que la predicción se realiza sobre una serie de comportamiento caótico, después de un cierto número de iteraciones la respuesta diverge exponencialmente de los valores reales de la serie de tiempo, de manera similar a la divergencia que presenta un sistema caótico ante cambios pequeños en sus condiciones iniciales. Aún cuando la divergencia se presentará con

seguridad, esta modalidad tiene un horizonte predictivo mayor que la modalidad de un solo paso.

Es interesante observar en la modalidad iterativa cómo la predicción, aún cuando ha divergido del valor real de la serie, presenta un comportamiento muy similar al del sistema, es decir, el modelo ha *aprendido* el comportamiento caótico de la serie de tiempo de entrada.

Lorenz

Este sistema ha jugado un papel principal en el estudio de sistemas caóticos, por tanto es uno de los más empleados en la determinación de exactitud de nuevos modelos predictivos. Se escriben abajo nuevamente las ecuaciones del sistema, que son las mismas de (2.1).

$$\begin{aligned}\dot{x} &= -\sigma(x - y) \\ \dot{y} &= rx - y - xz \\ \dot{z} &= xy - bz\end{aligned}\tag{6.7}$$

empleando los parámetros $\sigma=10$, $r=28$ y $b=8/3$, se obtiene la gráfica de la Figura 63.

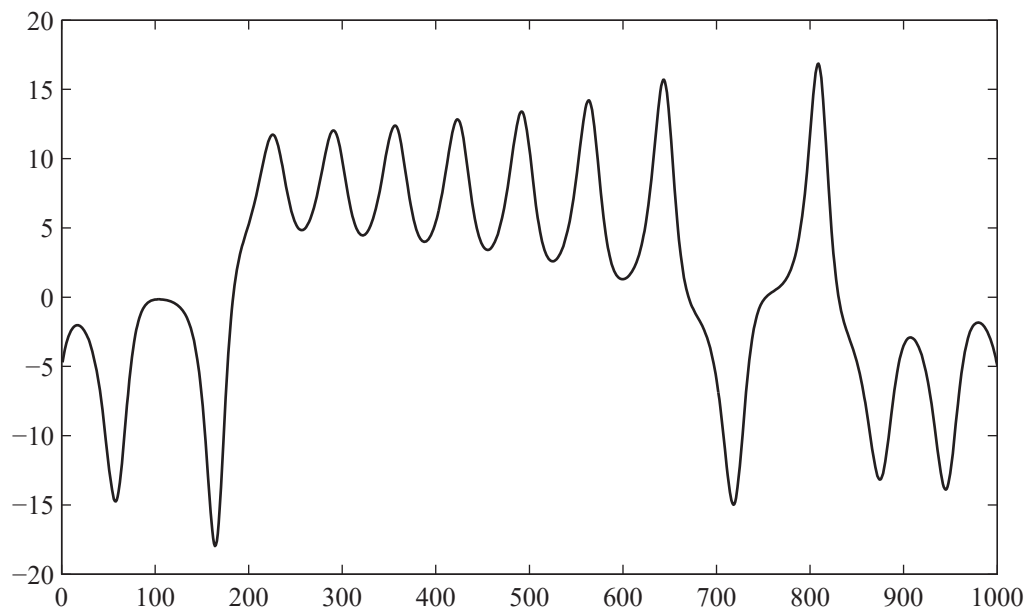


Figura 63. Primeros 1000 puntos de la serie de tiempo de Lorenz.

La Figura 63 muestra los primeros 1,000 puntos de la serie de Lorenz, muestreada cada 0.01 segundos, utilizando la función `ode45` de MATLAB. Como puede verse, la serie consiste de oscilaciones crecientes que al llegar a un nivel crítico, se colapsan y cambian a un estado distinto.

En el caso del sistema de Lorenz, los parámetros empleados son los que muestra la Tabla 4:

Variable	Valor	Variable	Valor
τ	10	n_e	10,000
m	3	n_p	500
n_c	100	n_i	200

Tabla 4

La Figura 64 muestra la predicción de un paso para el sistema de Lorenz. La línea discontinua representa la serie de tiempo original. Como puede verse, la predicción resulta de gran calidad.

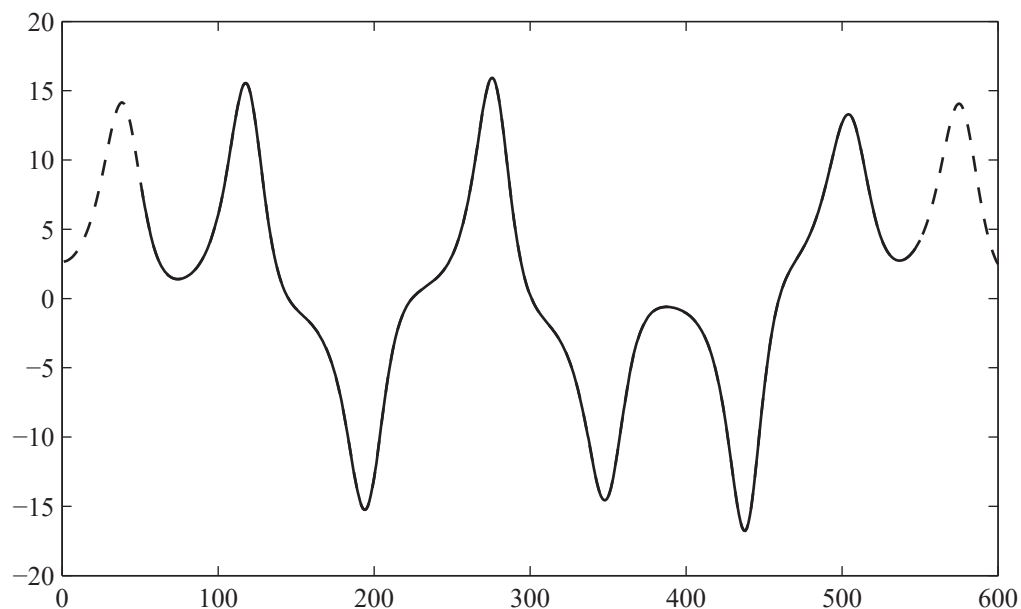


Figura 64. Gráfico de la predicción de la serie de Lorenz, modalidad de un paso.

Los estimadores de error obtenidos para esta serie son los que muestra la Tabla 5:

Predicción	Error medio cuadrático normalizado	Raíz del error medio cuadrático normalizado
Un paso	1.2353×10^{-6}	0.0011
Iterada	0.0221	0.1487

Tabla 5. Estimadores de error para las predicciones en el sistema de Lorenz.

La Figura 65 es un gráfico del error punto a punto entre la serie real del Sistema de Lorenz y la predicción realizada, mostrados en la Figura 64. El interés en este tipo de error es mostrar las tendencias del mismo conforme el sistema evoluciona en el tiempo.

De la Figura 64 y la Figura 65 es posible ver que el error es estable a lo largo de la predicción y muy cercano a cero, lo cual permite concluir que la predicción es de gran calidad para esta modalidad.

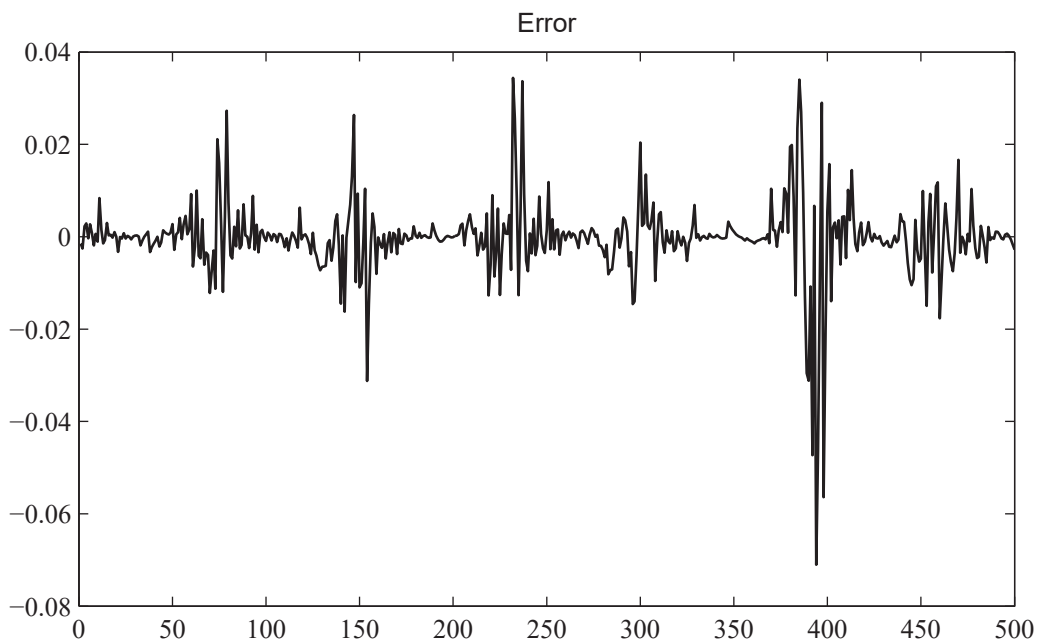


Figura 65. Gráfico de error para la predicción de un paso del sistema de Lorenz.

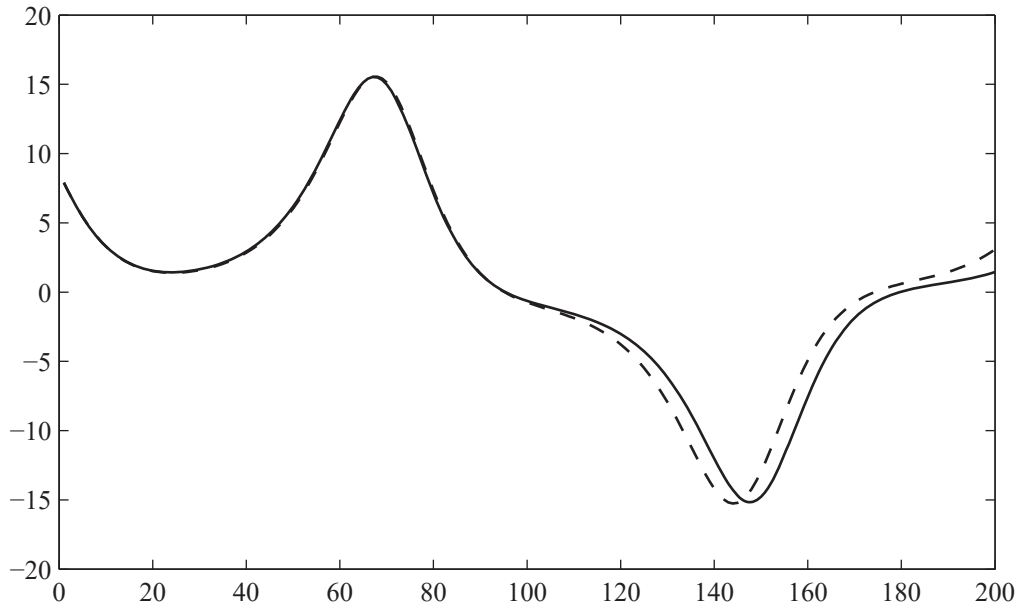


Figura 66. Gráfico de la predicción de la serie de Lorenz, modalidad iterada.

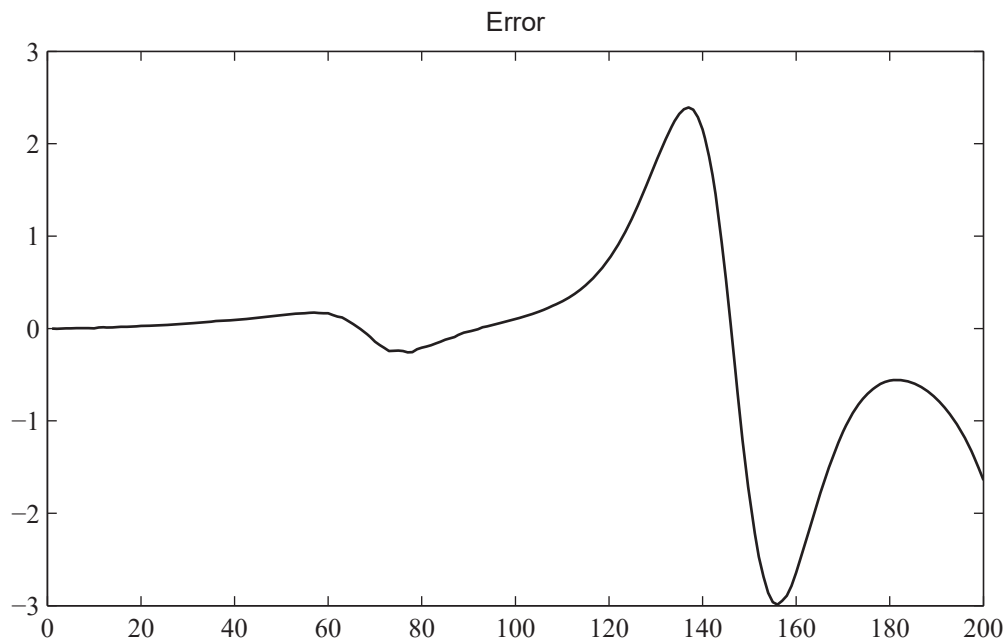


Figura 67. Gráfico de error para la predicción iterada del sistema de Lorenz.

La Figura 66 muestra la predicción iterada de la serie de tiempo de Lorenz. Nuevamente, la línea punteada representa la serie de tiempo original, en tanto que la Figura 67 muestra el error presente en la predicción. En este caso, el error tiende a crecer conforme

el sistema evoluciona, pero como se mencionó anteriormente, el sistema siempre tenderá a divergir. Más importante es el hecho de que la predicción comience con un error pequeño y éste se mantenga así dentro de un cierto “horizonte”, después del cual el crecimiento del error es inminente.

Chen

Para el sistema de Chen, el cual se analizó anteriormente, se obtuvieron los siguientes resultados.

Por comodidad, se repiten nuevamente las ecuaciones que definen a este sistema:

$$\begin{aligned}\dot{x} &= a(y - x) \\ \dot{y} &= (c - a)x - xz + cy \\ \dot{z} &= xy - bz\end{aligned}\tag{6.8}$$

los parámetros empleados para este caso son $a=35$, $b=3$ y $c=28$. El sistema se muestreo a intervalos de 0.01 segundos.

Los parámetros de entrada al método de predicción son los siguientes:

Variable	Valor	Variable	Valor
τ	2	n_e	10,000
m	3	n_p	500
n_c	100	n_i	200

Tabla 6

Nuevamente es posible ver una muy buena aproximación para la predicción de un paso, en la Figura 68. Los estimadores de error para este caso son los que muestra la Tabla 7.

Predicción	Error medio cuadrático normalizado	Raíz del error medio cuadrático normalizado
Un paso	2.3431×10^{-6}	0.0015
Iterada	1.4283	1.1951

Tabla 7. Estimadores de error para las predicciones en el sistema de Chen.

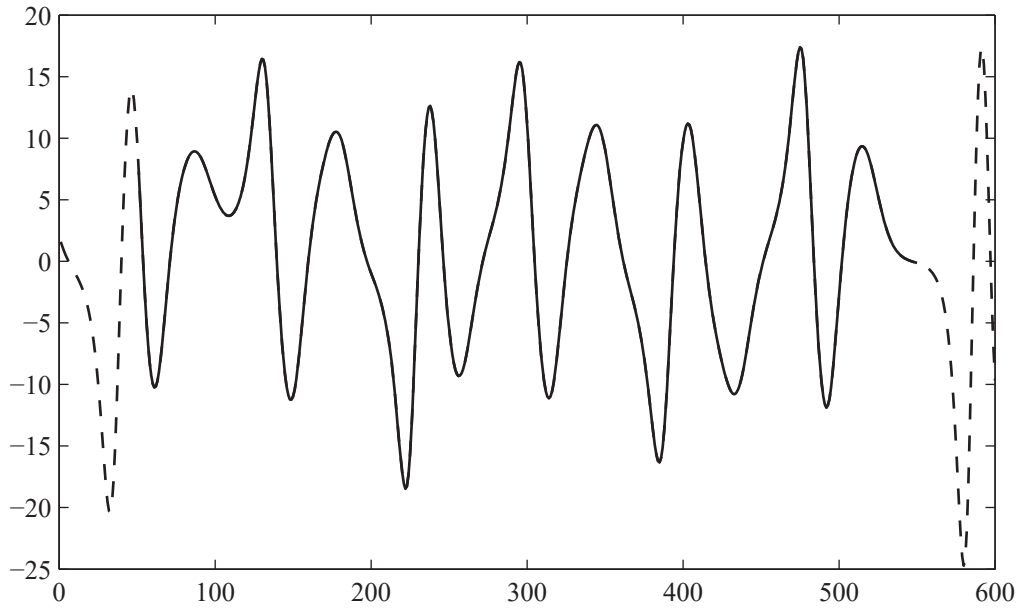


Figura 68. Gráfico de la predicción del sistema de Chen, modalidad de un paso.

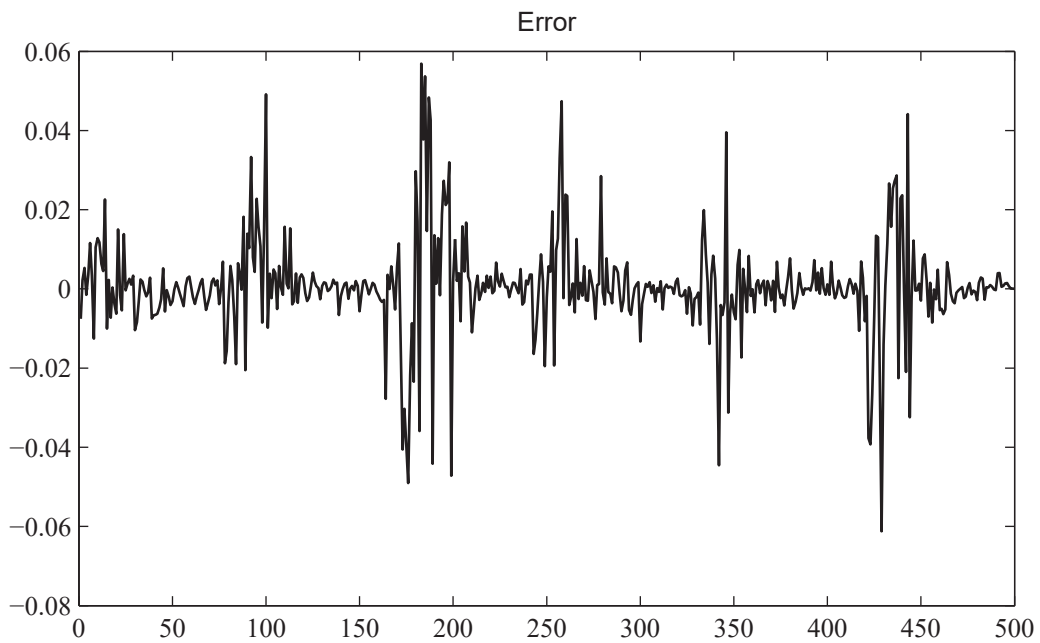


Figura 69. Gráfico de error para la predicción de un paso del sistema de Chen.

El error mostrado en la Figura 69 es el error involucrado en la predicción de un paso para el sistema de Chen. El error se mantiene estable a lo largo de la predicción y en

valores cercanos a cero. El estimado de error mostrado en la Tabla 7 permite concluir que la predicción es de calidad.

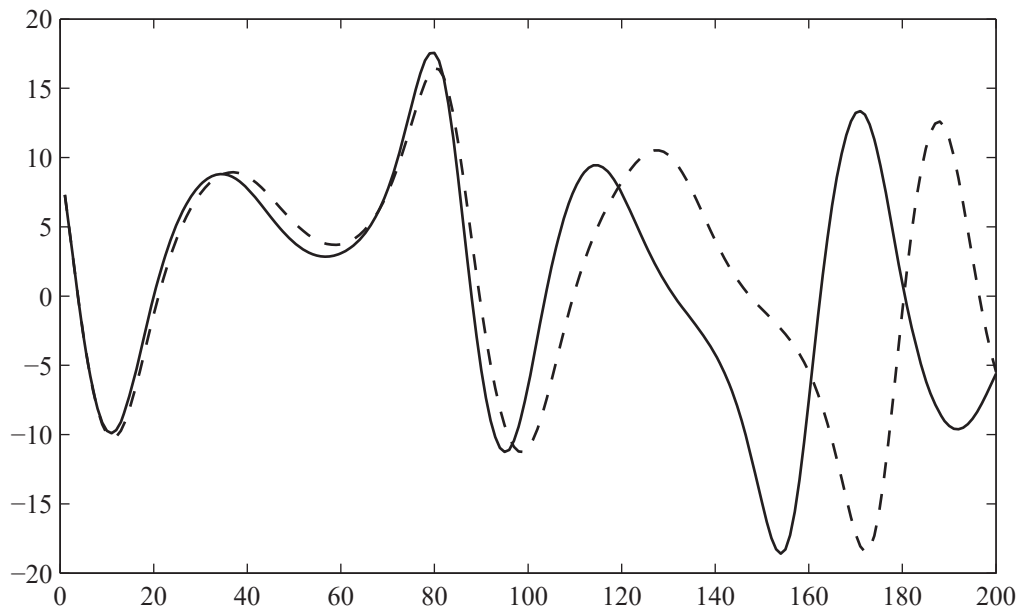


Figura 70. Gráfico de la predicción del sistema de Chen, modalidad iterada.

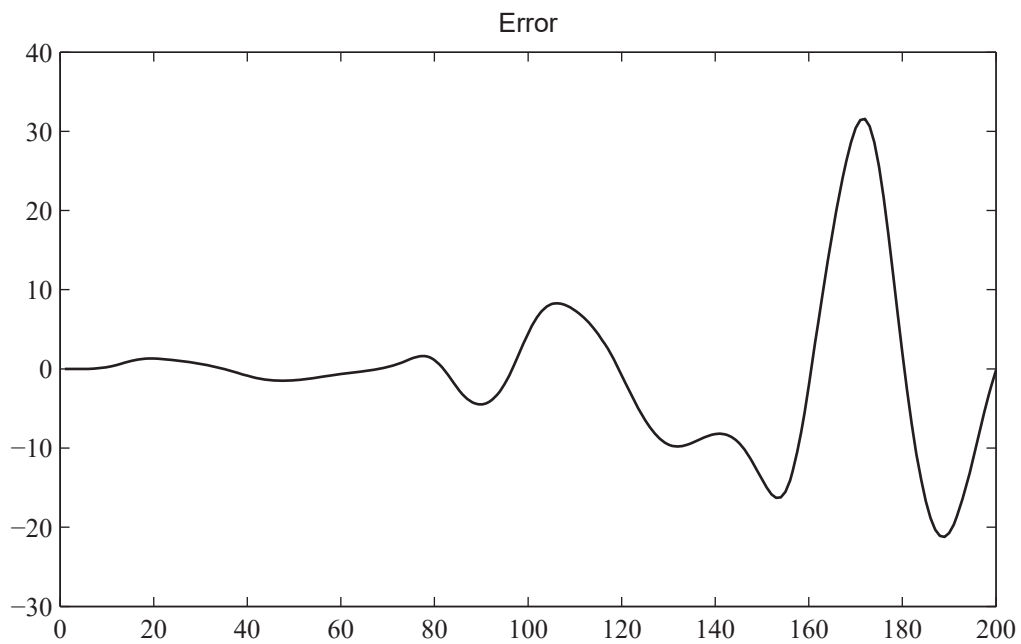


Figura 71. Gráfico de error para la predicción iterada del sistema de Chen.

El error presente en la predicción iterada para este caso es mucho mayor que el observado en el caso de Lorenz, sin embargo, de la Figura 70 es posible ver que aún cuando el sistema no “sigue” de una manera fiel al sistema, el comportamiento predicho por el método de predicción es muy similar al que muestra el sistema en la realidad. Aún así, el error en la predicción se mantiene dentro de un margen pequeño para un horizonte relativamente largo de tiempo, aproximadamente ochenta pasos adelante, de acuerdo con la Figura 71.

Chua

El circuito de Chua presenta prácticamente todos los fenómenos de bifurcación que se pueden encontrar en los sistemas caóticos. Lo anterior y el hecho de que es fácil de implementar físicamente lo convierten en un modelo para los estudios sobre caos, así como un ejemplo común para ilustrar éste fenómeno.

Las ecuaciones que describen este sistema son:

$$\begin{aligned}\dot{x} &= \alpha(y - m_0x - \frac{1}{3}m_1x^3) \\ \dot{y} &= x - y + z \\ \dot{z} &= -\beta y\end{aligned}\tag{6.9}$$

los valores paramétricos empleados son: $\alpha=9$, $\beta=14.286$, $m_0=-1/7$, $m_1=2/7$. Muestreado a 0.03 segundos, empleando la función `ode45` de MATLAB.

Una serie típica para una de las variables del sistema de Chua es la que muestra la Figura 72.

Este sistema es un oscilador con dos regímenes de oscilación. Cuando el sistema alcanza valores críticos, la oscilación “salta” de un régimen al otro.

Las ecuaciones del sistema pueden modificarse de tal manera que sea posible lograr un gran número de regímenes de oscilación. A este caso se le denomina *multi-scroll*, y se pueden ver ejemplos de tales atractores en [Aziz Alaoui, 2000].

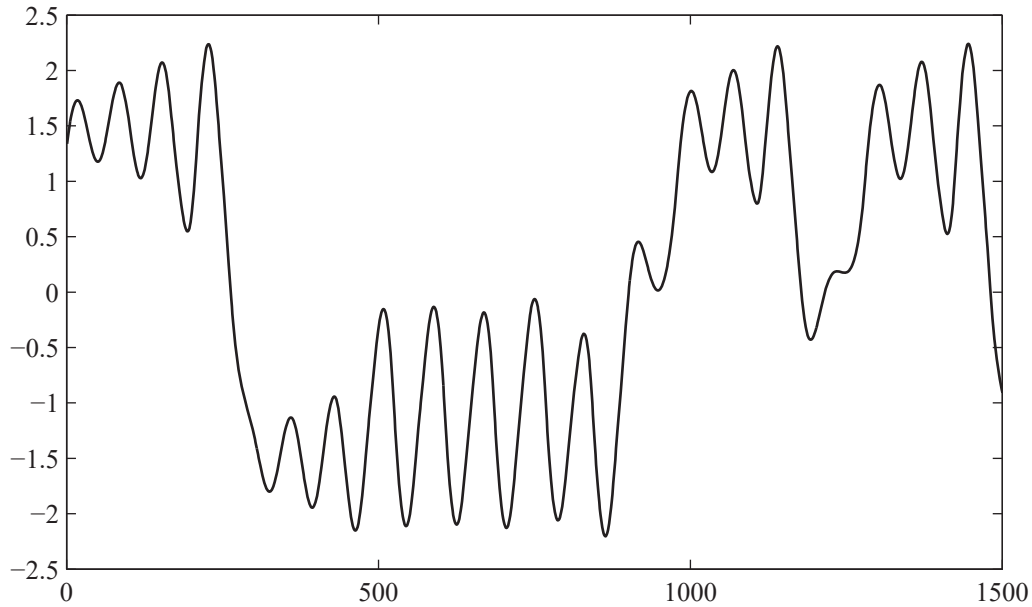


Figura 72. Primeros 1500 puntos para la serie de tiempo del sistema de Chua.

Para la predicción del Sistema de Chua se emplearon los siguientes parámetros:

Variable	Valor	Variable	Valor
τ	10	n_e	10,000
m	3	n_p	500
n_c	120	n_i	200

Tabla 8

El error involucrado en la predicción es el que muestra la Tabla 9, mostrada a continuación:

Predicción	Error medio cuadrático normalizado	Raíz del error medio cuadrático normalizado
Un paso	4.3014×10^{-7}	6.5585×10^{-4}
Iterada	0.5554	0.7452

Tabla 9. Estimadores de error para las predicciones en el sistema de Chua.

La Figura 73 muestra la predicción de un paso para este caso.

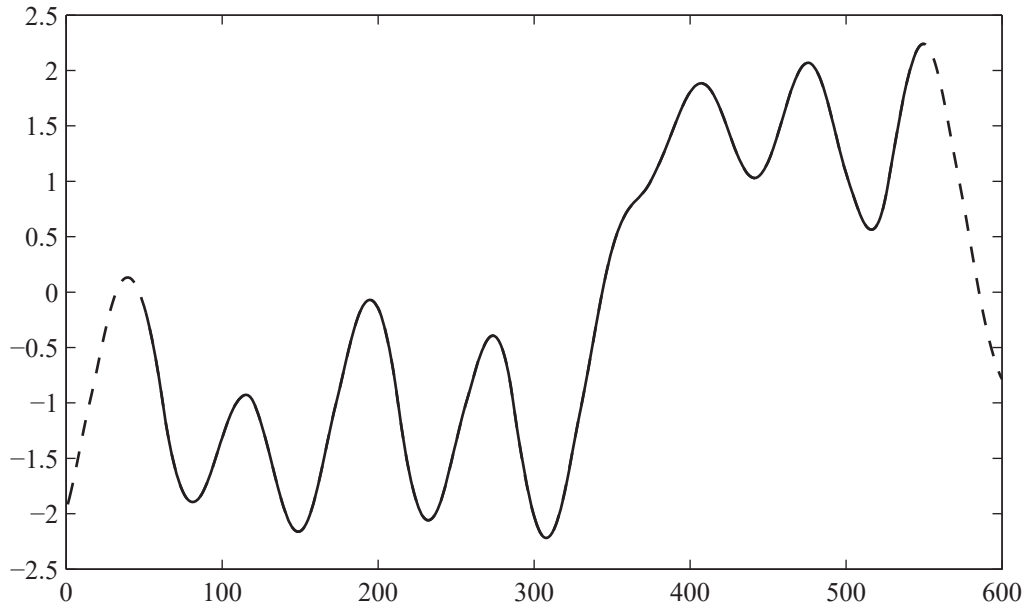


Figura 73. Gráfico de la predicción del sistema de Chua, modalidad de un paso.

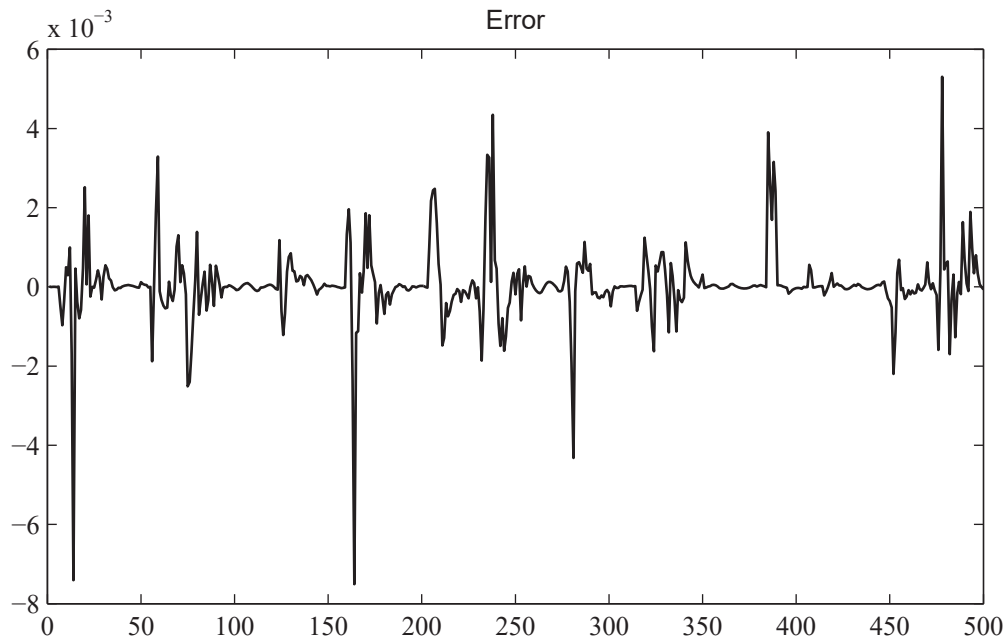


Figura 74. Gráfico de error para la predicción de un paso del sistema de Chen.

La Figura 74 muestra el error involucrado en la predicción de este sistema. La gráfica muestra un comportamiento estable del error a lo largo de la predicción.

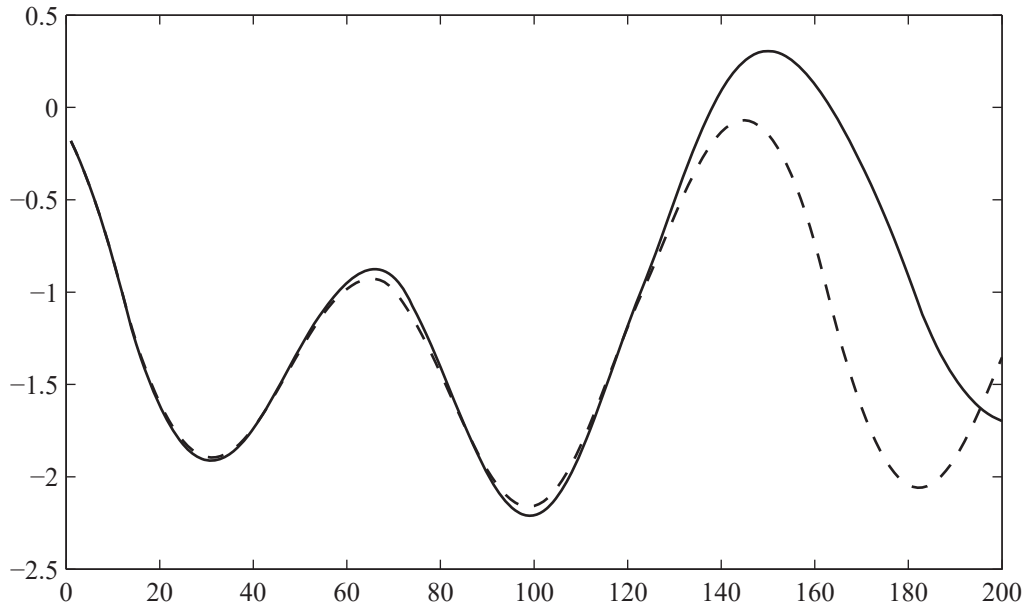


Figura 75. Gráfico de la predicción del sistema de Chua, modalidad iterada.

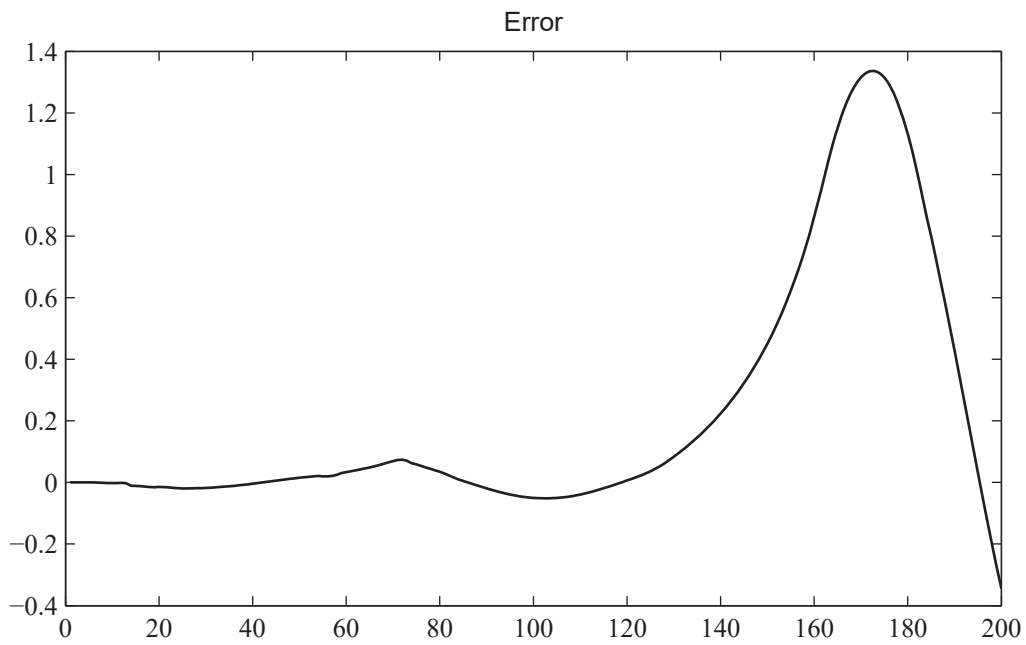


Figura 76. Gráfico de error para la predicción iterada del sistema de Chua.

La Figura 75 y la Figura 76 muestran el comportamiento de la predicción y el error observado para el caso iterado. Nuevamente se observa una gran divergencia del error al

final de la gráfica, pero el horizonte de predicción en el cual el error se mantiene pequeño es relativamente largo. Este es un caso muy similar al observado en el sistema de Chen.

Rössler

El sistema de Rössler se puede considerar como el sistema mínimo que exhibe caos. A pesar de esta simplicidad, resulta de una mayor complejidad analítica que sistemas como el de Lorenz.

La dinámica de este sistema está definida por el siguiente conjunto de ecuaciones:

$$\begin{aligned}\dot{x} &= -(z + y) \\ \dot{y} &= x + ay \\ \dot{z} &= b + xz - cz\end{aligned}\tag{6.10}$$

Un conjunto de parámetros comunes en los cuales el sistema exhibe caos es el siguiente son $a=0.2$, $b=0.2$ y $c=5$. Estos valores se emplearon para simular el sistema, muestréandolo a intervalos de 0.1 segundos. Una gráfica de la serie de tiempo es la que se muestra en la Figura 77.

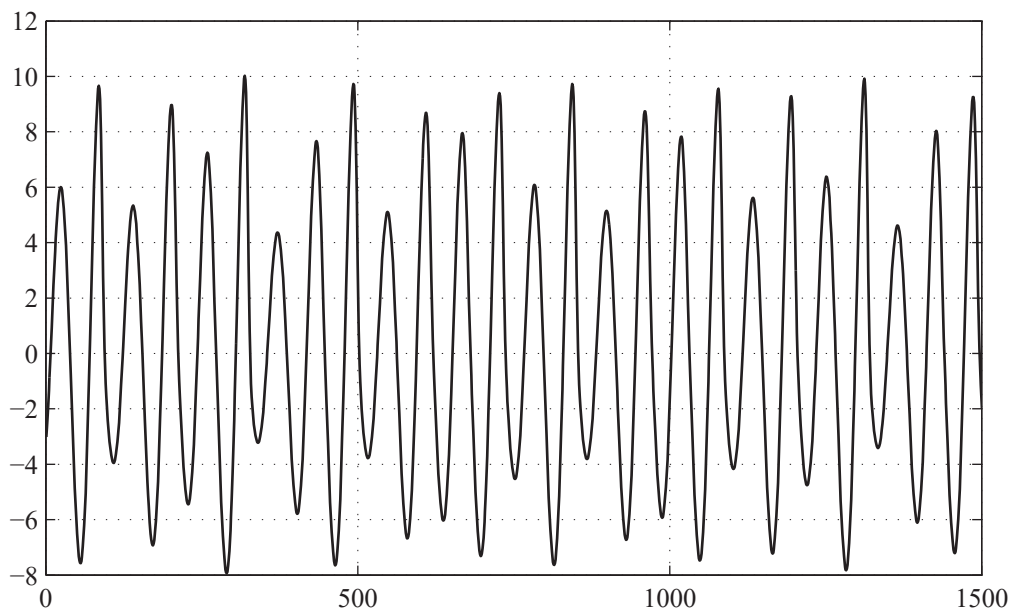


Figura 77. Serie de tiempo para la variables x del sistema de Rössler.

Los parámetros empleados para su predicción son:

Variable	Valor	Variable	Valor
τ	2	ne	10,000
m	3	np	500
n_c	100	ni	200

Tabla 10

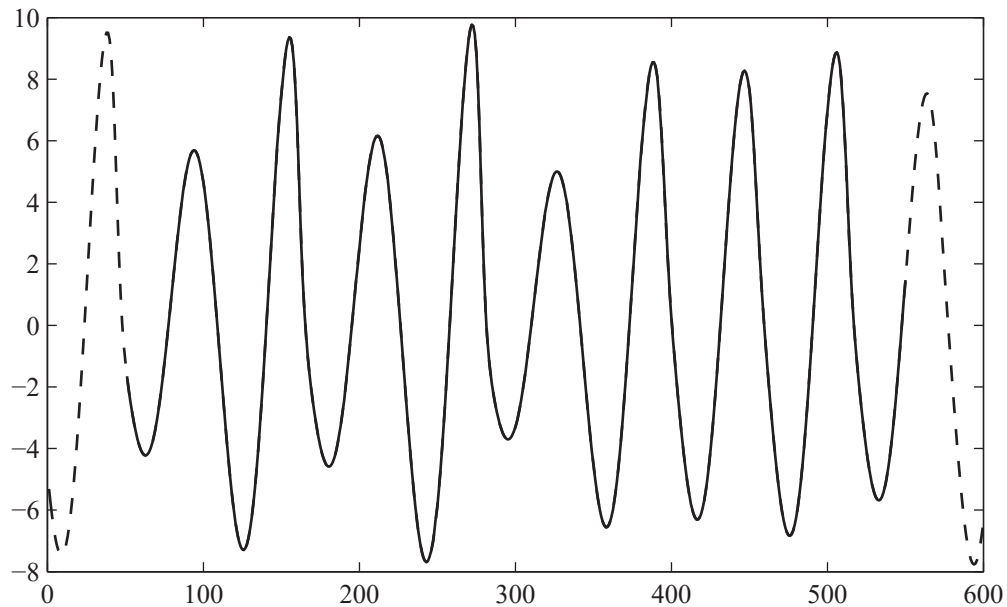


Figura 78. Gráfico de la predicción del sistema de Rössler, modalidad de un paso.

Los errores encontrados para este caso se muestran en la Tabla 11

Predicción	Error medio cuadrático normalizado	Raíz del error medio cuadrático normalizado
Un paso	2.1784×10^{-6}	0.0015
Iterada	0.0070	0.0839

Tabla 11. Estimadores de error para las predicciones en el sistema de Rössler.

En este caso, ambas modalidades presentan un error pequeño, y como puede apreciarse la Figura 78 y en la Figura 79, las predicciones son de muy buena calidad.

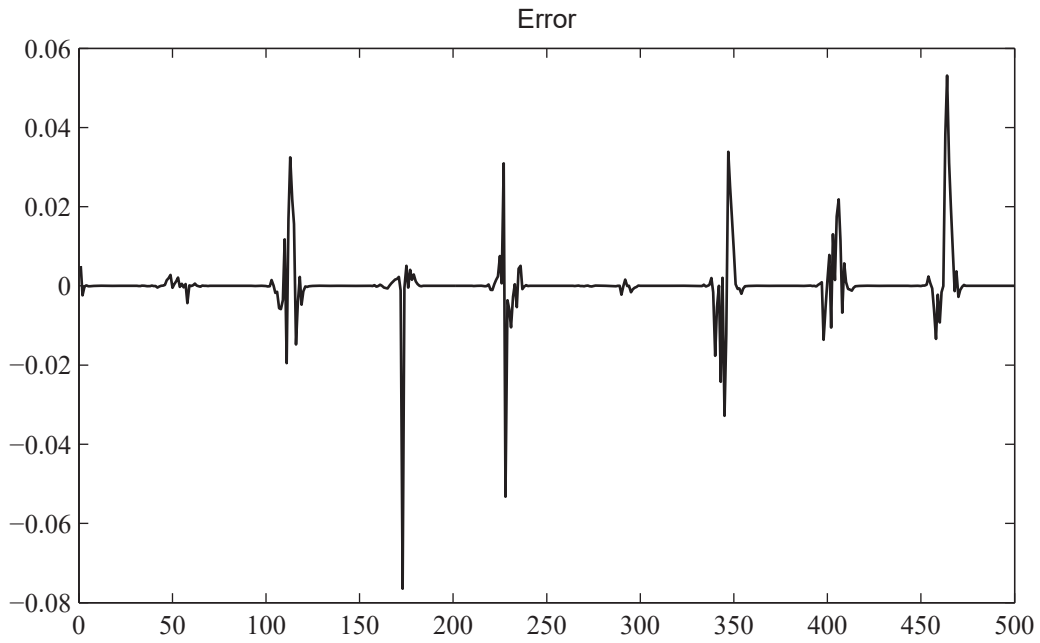


Figura 79. Gráfico de error para la predicción de un paso del sistema de Rössler.

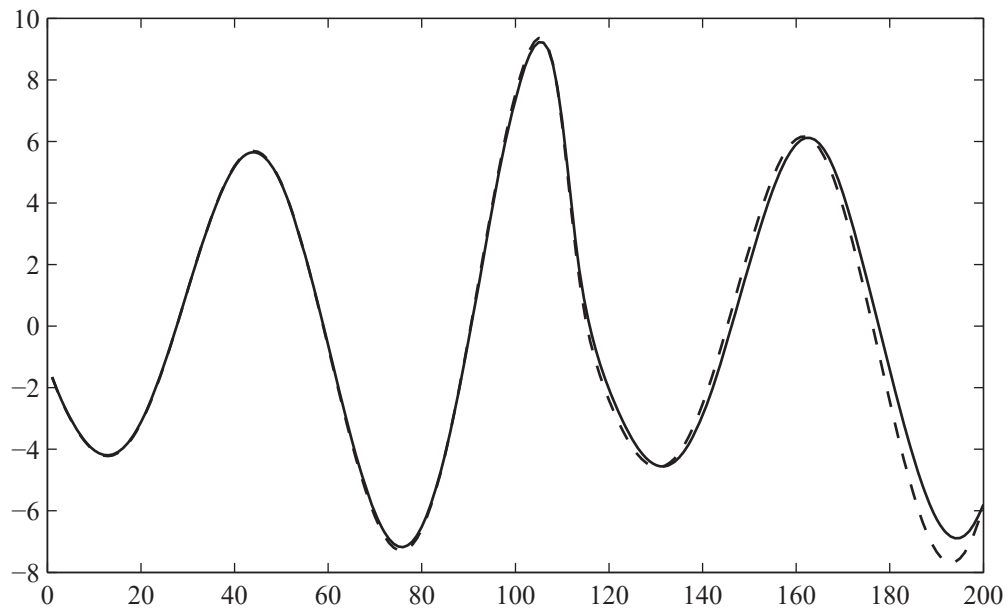


Figura 80. Gráfico de la predicción del sistema de Rössler, modalidad iterada.

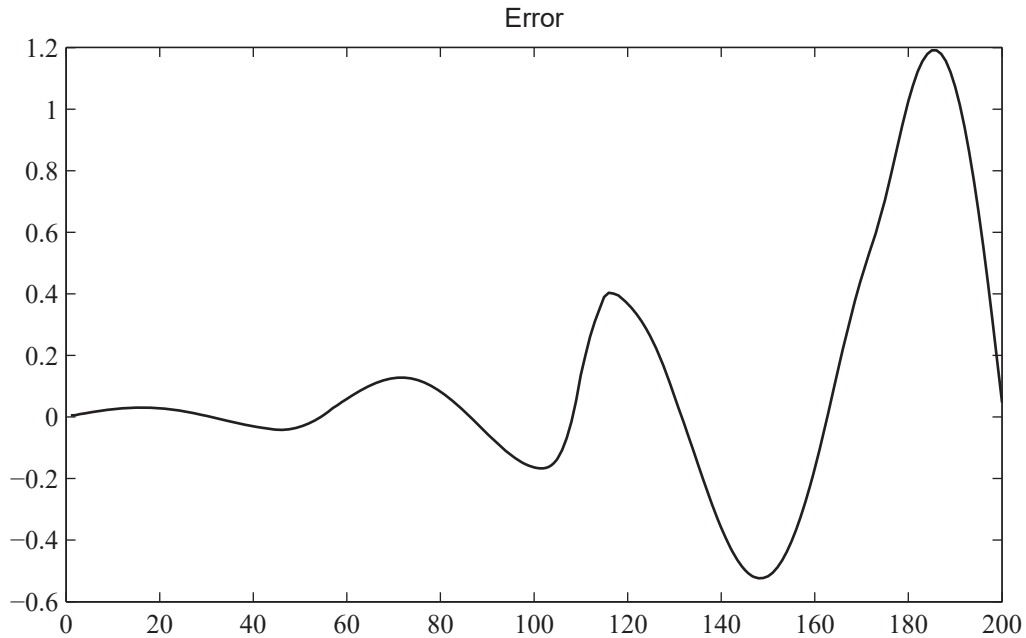


Figura 81. Gráfico de error para la predicción iterada del sistema de Rössler.

Duffing

La ecuación del oscilador de Duffing que se empleó para simular este sistema es la siguiente:

$$B \cos(t) = \ddot{x} + k\dot{x} + x^3 \quad (6.11)$$

utilizando los valores paramétricos de $k=0.05$ y $B=7.5$ a intervalos de 0.08 segundos.

Los parámetros empleados para su predicción son:

Variable	Valor	Variable	Valor
τ	2	ne	50,000
m	3	np	500
n_c	200	ni	200

Tabla 12

Los errores obtenidos en la predicción se muestran en la Tabla 13.

Predicción	Error medio cuadrático normalizado	Raíz del error medio cuadrático normalizado
Un paso	1.1454×10^{-5}	0.0034
Iterada	1.8385	1.3559

Tabla 13. Estimadores de error para las predicciones en el sistema de Duffing.

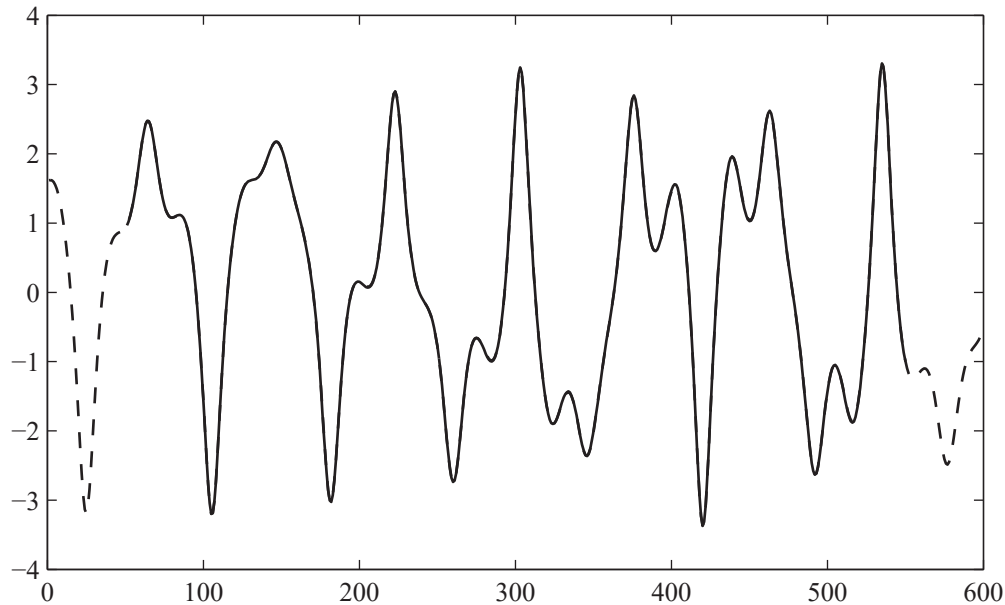


Figura 82. Gráfico de la predicción del sistema de Duffing, modalidad de un paso.

Este sistema presenta una dificultad relativamente mayor para su predicción, como lo demuestran los valores asociados al error. La predicción de un paso resulta suficientemente buena, aún así, es mayor en una orden de magnitud al de los sistemas previos. En cuanto al proceso iterado, su horizonte es relativamente corto y no presenta nunca una fidelidad hacia el sistema, pero logra imitar su comportamiento.

En este caso, el valor de información mutua (IM) que se obtuvo era relativamente alto, lo cual resultó discordante con la predicción, es decir, para el valor de IM, la predicción alcanzaba valores de error demasiado grandes, y recurriendo a un proceso iterativo de reducción del valor de IM, se llegó al valor de 2 empleado. Existe un gran número de factores que afecta la predicción de un sistema, entre ellos está la frecuencia de muestreo, el número de modelos locales empleados, la *dimensión fractal* del sistema, etc. Este último

concepto se refiere a la dimensión real del sistema, es decir, al aumento dimensional que sufre una longitud o una superficie al pasar a ser fractal, en oposición a la *dimensión topológica* o *euclidea* de los cuerpos bien definidos como son las líneas, las superficies bidimensionales o cuerpos tridimensionales.

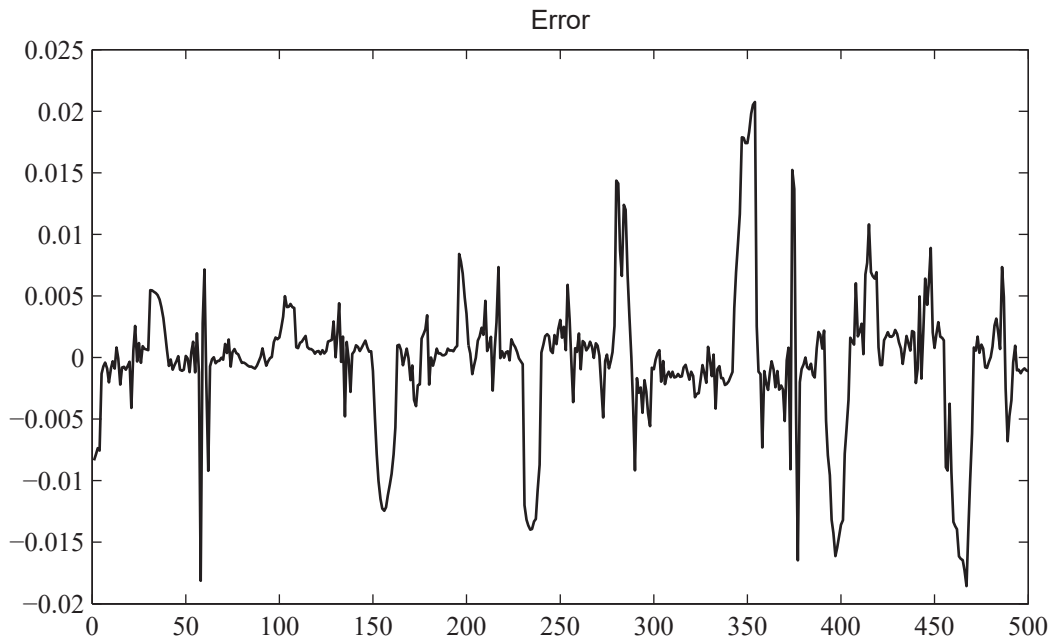


Figura 83. Gráfico de error para la predicción de un paso del sistema de Duffing.

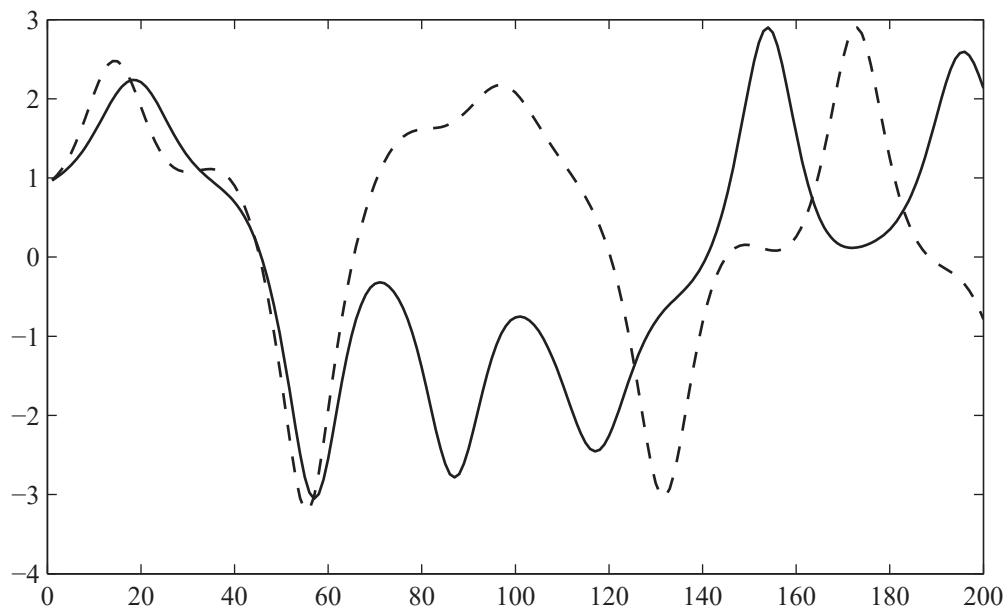


Figura 84. Gráfico de la predicción del sistema de Duffing, modalidad iterada.

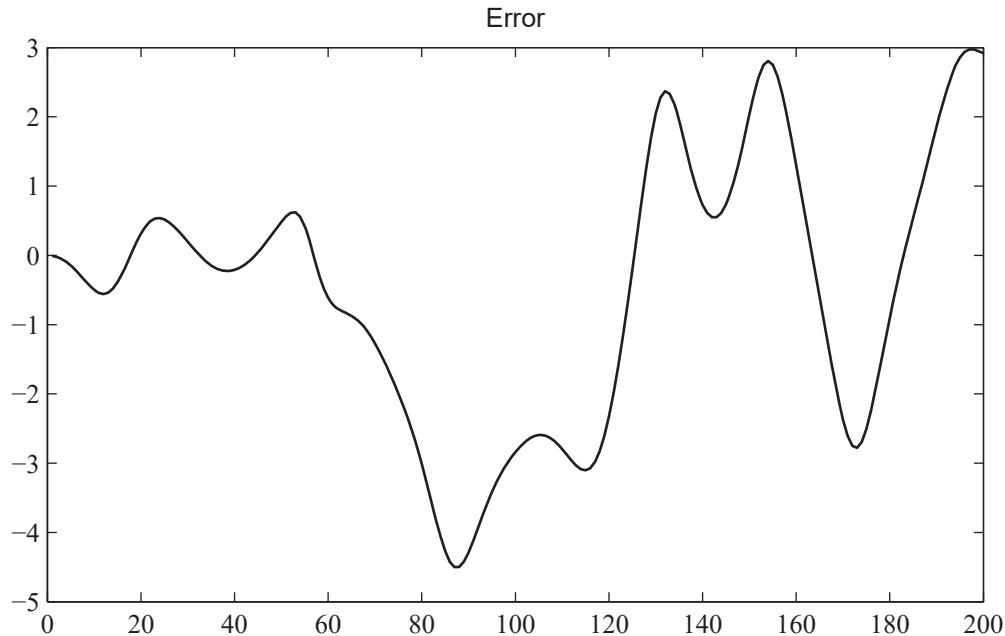


Figura 85. Gráfico de error para la predicción iterada del sistema de Duffing.

Es factible mejorar la predicción en este caso en particular mediante la optimización de los parámetros empleados, tanto en el muestreo del sistema como en el método de predicción.

6.7. Comportamiento del error medio cuadrático normalizado en el sistema

El error que muestran las tablas de la sección anterior es aquel que presenta el sistema ante una sola predicción, y está dado como un parámetro. En esta sección se presenta el comportamiento de tal error ante repetidas predicciones del sistema.

Se empleó el sistema de Lorenz para este estudio. Se realizaron 100 predicciones del sistema utilizando puntos aleatorios de la serie como valores iniciales de predicción. A partir de tal punto aleatorio se llevó a cabo la predicción de la serie, con quinientos puntos pronosticados mediante el esquema de un paso y doscientos mediante el método iterado. La Figura 86 muestra el comportamiento del error medio cuadrático normalizado para el caso de la predicción de un paso.

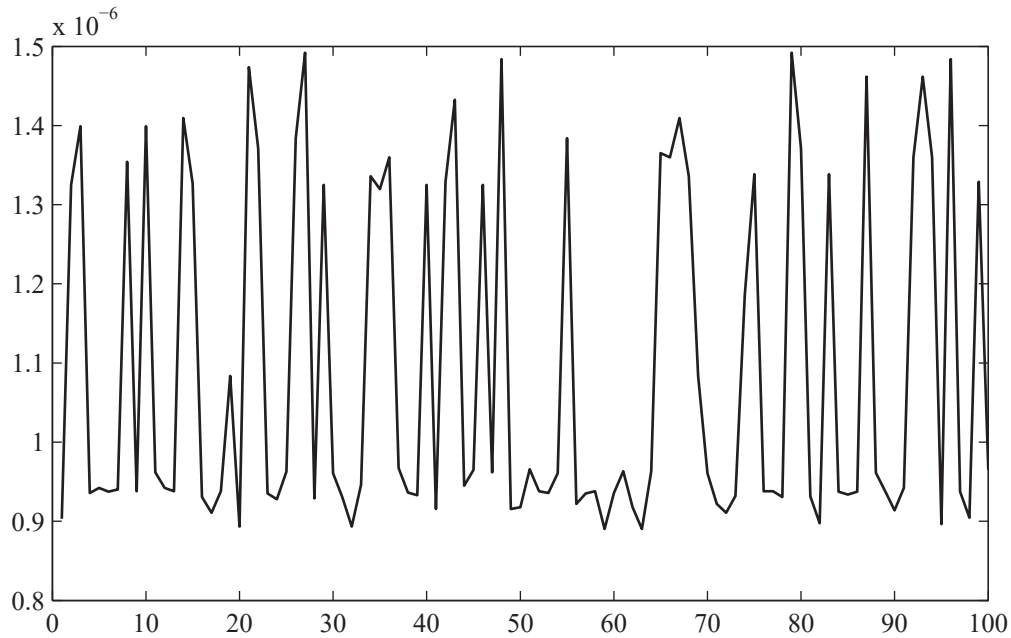


Figura 86. Gráfico de comportamiento del error medio cuadrático.

Es posible ver que el error se mantiene acotado dentro de un intervalo de valores, lo cual demuestra la estabilidad del sistema en cada predicción. Esto permite un cierto grado de confianza de que la predicción será adecuada, independientemente del punto de arranque de la misma.

La gráfica presentada resulta una gráfica genérica del error, dado que se obtienen resultados muy similares para cada uno de los otros sistemas revisados.

Un resultado interesante se puede observar en la Figura 87, que representa el comportamiento del error medio cuadrático para la predicción iterada. Como puede verse en cualquiera de las gráficas del error en la predicción iterada de la sección anterior, ésta tiende a divergir exponencialmente después de un cierto número de pasos, por lo cual podría esperarse que el error presentará una variación muy grande entre predicciones, sin embargo, la Figura 87 muestra un comportamiento acotado del error dentro de un rango relativamente pequeño de valores. Nuevamente, esto es señal de la estabilidad del esquema de predicción aún frente a tales condiciones.

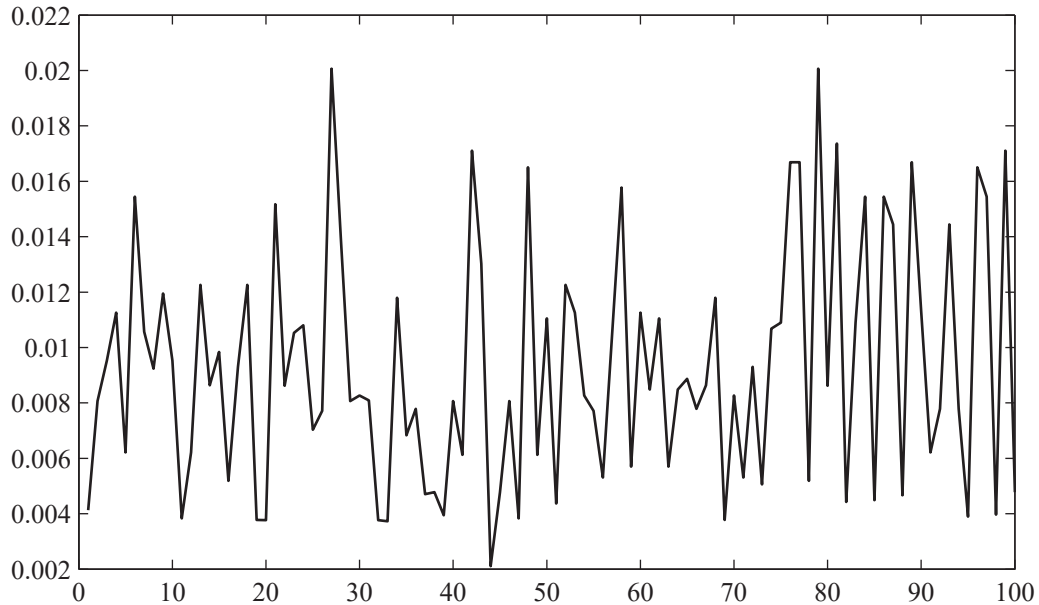


Figura 87. Gráfico de comportamiento del error medio cuadrático, para pronóstico iterado.

6.8. Conclusiones

El presente capítulo abordó el tema principal de la tesis: Definir un esquema de predicción basado en modelos locales que haga uso de un mapa auto-organizado como clasificador.

Se definieron conceptos importantes que definen los pasos a seguir para lograr el objetivo: El teorema de Takens, y el paradigma de modelos locales. Una vez definidos los anteriores conceptos se llevó a cabo la implementación del esquema, el cuál se probó con cinco distintos sistemas que presentan comportamientos caóticos conocidos.

En base a los errores observados, se puede concluir que el esquema presenta robustez y confiabilidad para la predicción de sistemas caóticos. Cabe señalar que las predicciones de tales sistemas se logran utilizando modelos lineales, lo cual confiere una gran simplicidad en su implementación. Otra característica importante que presenta este esquema es su modularidad, ya que es perfectamente posible sustituir en cualquiera de las etapas, lo que podríamos denominar el *núcleo*; por ejemplo, emplear un modelo de regresión no lineal en la etapa de reconstrucción del sistema, o alguna red neuronal distinta como clasificador, extendiéndose así las posibilidades del modelo a otro tipo de tarea.

7. Conclusiones y trabajos futuros.

El objetivo general de esta tesis fue investigar el uso de los modelos locales como método de predicción de series de tiempo, específicamente, aquellas series surgidas de sistemas caóticos. Para lograr tal objetivo, fue necesario el uso de diversas herramientas como lo son las redes neuronales; en este caso en particular, la llamada red neuronal de Kohonen, o mapa auto-organizado; y de métodos matemáticos bien establecidos como lo son los mínimos cuadrados, así como de métodos novedosos que proporcionan nueva información sobre tales sistemas: Información mutua promedio y el algoritmo de vecinos falsos, los cuales marcan la pauta que permite hacer uso del Teorema de Takens.

Los mapas auto-organizados proporcionaron la capacidad de “aprender” la dinámica de los sistemas, que para este caso en especial, es precisamente una de enorme riqueza en cuanto a comportamientos y posibilidades: la dinámica caótica. Los mapas cumplen con la tarea de *caracterizar* al sistema, es decir, determinar los atributos que son específicos a un sistema en particular, y que lo distinguen de los demás. Sin embargo, la caracterización va más allá de una caracterización generalizada del sistema, y se llega hasta el punto de hacerlo con los diferentes tipos de comportamiento, o regímenes que presenta. Así, tenemos todo un conjunto de caracterizaciones que representan *dinámicas locales* de un mismo sistema.

Las dinámicas locales observadas como salida de la red neuronal, proporcionan un conjunto de modelos lineales para la predicción del sistema. Se debe tener en cuenta que cualquier predicción realizada para un sistema en dinámica caótica no va más allá del corto plazo, y cuando mucho, mediano plazo, dado que cualquier sistema caótico presentará una divergencia exponencial en la trayectoria de su dinámica.

La incorporación de pesos al proceso de mínimos cuadrados, tiene como propósito incluir la información del tiempo que el sistema permanece en la dinámica de cierto modelo local. Decir esto es referirse no a un intervalo de tiempo continuo, sino en general a qué tanto el sistema tiende a estar en una cierta localidad de su trayectoria. Por otro lado, ayuda a limitar el efecto de valores atípicos que puedan presentarse. A la vez que permite eliminar el efecto de discontinuidad que presentan los modelos locales, al suavizar los bordes de los mismos, haciendo una transición suave de un modelo a otro.

Como se mencionó en el capítulo anterior, la modularidad del sistema permite una gran flexibilidad en cuanto al tipo de tarea que el esquema deba realizar, dándose la posibilidad de que funcione como un predictor, como es el caso que se presenta en esta tesis, o como un aproximador de funciones. El uso de redes neuronales de otro tipo, por ejemplo, redes recurrentes o redes regresivas, le confiere características distintas, lo cual puede resultar adecuado para algún cierto tipo de problema.

Una de las principales motivaciones para este trabajo es la posibilidad de predecir procesos reales, como la demanda de energía eléctrica, dada la enorme importancia que viene cobrando con los nuevos esquemas de mercado, pero no necesariamente está sujeto a este proceso. El análisis de series de tiempo como mediciones sísmicas o médicas, por ejemplo, también son de relevancia y afortunadamente, susceptibles de ser analizadas bajo el esquema propuesto. Se debe señalar que las mediciones reales pueden resultar increíblemente complejas, dada su característica “ruidosa”, ya que normalmente se obtienen a través de algún transductor, que por lo común es no lineal. El esquema propuesto en la tesis obtiene resultados relativamente buenos de predicción en este tipo de señales, sin embargo es susceptible de mejorarse mediante la optimización en la elección de los distintos parámetros involucrados en el proceso, y probablemente mediante el uso de

procesos regresivos menos sensibles al ruido. Otra alternativa es el uso de métodos de regularización para el pre-tratamiento de los datos de entrada del sistema

Bibliografía

[Aziz Alaoui, 2000]

Aziz Alaoui, M. A., "*Multispiral Chaos*", IEEE Proceedings Control of Oscillations and Chaos ed., IEEE, 2000, pp. 88-91.

Blake y Merz. "UCI Repository of machine learning databases." University of California, Department of Information and Computer Science. Irvine, CA. 1998. Anderson, E. (1935) "The Irises of the Gaspé Peninsula", Bulletin of the American Iris Society, 59, 2-5.

[Chao, 1993]

Chao, L. L., "*Estadística para las ciencias administrativas*", 3 ed., McGraw-Hill, 1993.

[Chua et al, 1986]

Chua, L. O., Komuro, M. y Matsumoto, T., "*Double Scroll Family*", IEEE Transaction on Circuits and Systems ed., 1986, pp. 1073-1118.

[Eckmann et al, 1987]

Eckmann, J. P., Kamphorst, S. O. y Ruelle, D., "*Recurrence Plots of Dynamical Systems*", Europhysics Letters, 1987, pp. 973-977.

[Eckmann y Ruelle, 1985]

Eckmann, J. P. y Ruelle, D., "*Ergodic Theory of Chaos and Strange Attractors*", Review of Modern Physics, 1985, pp. 617-659.

[Fraser y Swinney, 1986]

Fraser, A. M. y Swinney, H. L., "*Independent Coordinates for Strange Attractors from Mutual Information*", Physical Review A, 1986, pp. 1134-1140.

[Grassberger y Procaccia, 1983]

Grassberger, P. y Procaccia, I., "*Measuring the strangeness of a strange attractor*", Physica D, 1983, pp. 189-208.

-
- [Hénon, 1976]
Hénon, M., "A two-dimensional mapping with a strange attractor", Commun. Math. Phys., 1976, pp. 69-77.
- [Kaplan y Glass, 1995]
Kaplan, D. y Glass, H., "Understanding Nonlinear Dynamics", First ed., Springer-Verlag, 1995.
- [Kennel et al, 1992]
Kennel, M. B., Brown, R. y Abarbanel, H. D. I., "Determining Minimum Embedding Dimension using a Geometrical Construction", Phys. Rev.A, 1992, pp. 3403-3411.
- [Kohonen, 1987]
Kohonen, T., "Self-Organization and Associative Memory", 2nd ed., Berlin, Springer-Verlag, 1987.
- [Kohonen, 1989]
Kohonen, T., "The Self-Organizing Map", IEEE Publisher, 1989.
- [Kononov, 2003]
Kononov, E., "Visual Recurrence Analysis v.4.2",
<http://pw1.netcom.com/~eugenek/download.html>, 2003.
- [Liénard, 1928]
Liénard, A., "Etude des Oscillations Entretienue", Revue Générale de l'Electricité, 1928, pp. 901.
- [Löffelmann, 1998]
Löffelmann, H., "About Poincaré Maps", Technische Universität Wien, Institut für Computergraphik und Algorithmen., 1998.
- [Lorenz, 1963]
Lorenz, E., "Deterministic Nonperiodic Flow", Journal of Atmospheric Sciences, 1963, pp. 130-141.
- [Mandelbrot, 1982]
Mandelbrot, B. B., "The fractal geometry of nature", W. H. Freeman & Co., 1982.
- [McNames, 1999]
McNames, J., "Innovations in local modeling for time series prediction", Stanford University, 1999.
- [NIST/SEMATECH, 2003]
NIST/SEMATECH, "e-Handbook of Statistical Methods",
<http://www.itl.nist.gov/div898/handbook/>, 2003.

-
- [Principe et al, 1998]
Principe, J. C., Wang, L. y Motter, M. A., "*Local Dynamic Modeling with Self-Organizing Maps and Applications to Nonlinear System Identification and Control*", IEEE Proceedings, 1998, pp. 2240-2258.
- [Quetelet y Verhulst, 1850]
Quetelet, A. y Verhulst, P. F., "*Annuaire de l'Académie royale des sciences de Belgique*", 1850, pp. 97-124.
- [Rössler, 1976]
Rössler, O. E., "*An equation for continous chaos*", Physics Letters, 1976, pp. 397-398.
- [Rössler, 1979]
Rössler, O. E., "*Continuous Chaos: Four Prototypes Equations*", Annals Ney Yor Academy of Sciences, 1979, pp. 376-392.
- [Sauer et al, 1991]
Sauer, T., Yorke, J. A. y Casdagli, M., "*Embedology*", Journal of Statistical Physics, 1991, pp. 579-616.
- [Smith et al, 1999]
Smith, L. A., Ziehmann, C. y Fraedrich, K., "*Uncertainty dynamics and predictability in chaotic systems*", Q.J. Royal Meteorological Society, 1999, pp. 2855-2886.
- [Strogatz, 2000]
Strogatz, S. H., "*Nonlinear Dynamics and Chaos*", First ed., Westview Press, 2000.
- [Takens, 1981]
Takens, F., "*Detecting Strange Attractors in Turbulence*", 898 ed., Berlin, Springer, 1981, pp. 366-381.
- [Thompson y Stewart, 1993]
Thompson, J. M. T. y Stewart, H. B., "*Nonlinear Dynamics and Chaos*", First ed., John Wiley & Sons, 1993.
- [Ueta y Chen, 2000]
Ueta, T. y Chen, G., "*Bifurcation and Chaos of Chen's Equation*", IEEE International Symposium on Circuits and Systems ed., IEEE, 2000.
- [University of Western Ontario, 2003]
University of Western Ontario, "*Applied Mathematics*", <http://www.apmaths.uwo.ca/~bfraser/nll/version1/index.html>, 2003.

[van Zyl y Omlin, 2001]

van Zyl, J. y Omlin, C. W., "*Prediction of Seismic Events in Mines using Neural Networks*", International Joint Conference on Neural Networks, 2001, pp. 1410-1414.

[Varsta et al, 1997]

Varsta, M., Heikkonen, J. y Millan, J., "*Epileptic Activity Detection in EEG with Neural Networks*", 1997.

[Vesanto et al, 2000]

Vesanto, J., Himberg, J., Alhomieni, E. y Parhankangas, J., "*SOM Toolbox for Matlab 5*", 1 ed., Helsinki, Helsinki University of Technology, 2000.

[Wolfram, 2002]

Wolfram, S., "*A new Kind of Science*", Wolfram Media, 2002.

[Zbilut y Webber, 1992]

Zbilut, J. P. y Webber, C. L., "*Embeddings and delays as derived from quantification of recurrence plots*", Phys. Lett. A, 1992, pp. 199-203.