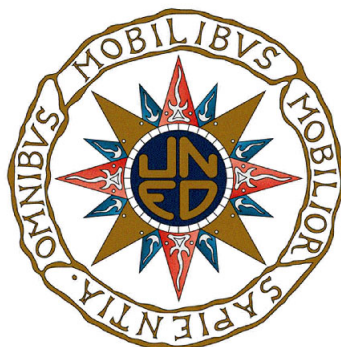


**UNIVERSIDAD NACIONAL DE EDUCACIÓN A DISTANCIA  
FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES  
DEPARTAMENTO DE ECONOMÍA APLICADA CUANTITATIVA I**



**ESTIMACIÓN DEL RETARDO TEMPORAL PARA CONTRASTAR  
SERIES DE TIEMPO DE ACTIVOS FINANCIEROS**

**TESIS DOCTORAL**

**RINA BETZABETH OJEDA CASTAÑEDA  
MAESTRA EN CIENCIAS MATEMÁTICAS**

**MADRID**

**2013**



**Universidad Nacional de Educación a Distancia**  
**Departamento Economía Aplicada Cuantitativa I**  
**Facultad de Ciencias Económicas y Empresariales**

**ESTIMACIÓN DEL RETARDO TEMPORAL PARA CONTRASTAR  
SERIES DE TIEMPO DE ACTIVOS FINANCIEROS**

**Tesis presentada por**  
**Rina Betzabeth Ojeda Castañeda**  
**Maestra en Ciencias Matemáticas**

**Director: Dr. Mariano Matilla García**  
**Codirector: Dr. Manuel Ruiz Marín**

## **DEDICATORIAS**

*A la memoria de mis padres Emma y Adelaido.*

*A la memoria de mis hermanos Francisco Xavier y Renato Hernán.*

*A mis hermanos Emma Luz, Hugo, Jorge y Gerardo.*

*Gracias por su apoyo y ejemplo de dedicación, constancia y congruencia en todos los actos de su vida personal y profesional.*

## **AGRADECIMIENTOS**

Mi más sincero agradecimiento a mis directores de tesis; al Dr. Mariano Matilla García, por su dirección y apoyo incondicional durante el desarrollo de esta investigación, proporcionándome en todo momento una guía para realizarla; al Dr. Manuel Ruiz Marín por la revisión, comentarios y sugerencias a este trabajo.

Quiero dedicar un especial agradecimiento a la Dra. María del Pilar Alonso Reyes y al M.C. José Antonio Flores Díaz, mis queridos colegas y amigos por sus comentarios y sugerencias a la parte estadística del presente trabajo.

Agradezco también al M.C. Francisco Cepeda Flores, director de mi centro de trabajo (Centro de Investigación en Matemáticas Aplicadas, CIMA, de la Universidad Autónoma de Coahuila), por su apoyo moral y gestión en la obtención de recursos para realizar algunas estancias cortas en la UNED, con el fin de llevar a cabo este trabajo de investigación. Gracias a mis colegas del Cuerpo Académico de Análisis Numérico y Optimización del CIMA, Humberto, Irma y Josefina, por la motivación para continuar este proceso de investigación, y en especial a José Luis Fraga, colaborador de nuestro cuerpo académico y del CIMA, por su apoyo en la edición de esta tesis en Latex.

## RESUMEN

El primer paso crucial en la descripción de un modelo general para una serie de tiempo financiera, es la selección de los retardos que deben ser incluidos, ya que es esencial en la predicción y desarrollo de los modelos económicos. Tradicionalmente, para la especificación de los retardos apropiados en un modelo se utilizan los coeficientes de autocorrelación y de autocorrelación parcial. Sin embargo, es conocido que procesos con autocorrelación cero pueden exhibir todavía un alto orden de dependencia o de dependencia no lineal, como sería el caso de algunos procesos bilineales e incluso con modelos deterministas caóticos puros, entre otros. En general, los procedimientos basados en la autocorrelación pueden arrojar resultados engañosos cuando se analizan modelos no lineales, por lo que no es posible detectar importantes relaciones no lineales en los datos, limitando su utilidad en la detección de retardos de tiempo apropiados (retrasos), especialmente en aquellos escenarios donde los fenómenos no lineales son más la regla que la excepción. En la actualidad es un hecho generalmente aceptado el que las variables económicas más importantes son de naturaleza no lineal y que entre ellas abundan relaciones no lineales.

Desde un punto de vista econométrico y estadístico, esta situación ha motivado el desarrollo de pruebas de dependencia serial, más aún, cuando las herramientas estadísticas que existen, tienen muchas limitaciones para llevar a cabo la selección de retardos significativos, lo que conduce a que probablemente se hagan elecciones de valores inapropiados. Por tal motivo en la presente investigación, tomando un camino diferente al de las autocorrelaciones y autocorrelación parcial, se propone y desarrolla un nuevo procedimiento estadístico no paramétrico que permite identificar los retardos significativos que deben estar presentes en la descripción y el desarrollo de modelos de series de tiempo de naturaleza desconocida, sean éstos lineales o no lineales. Adicionalmente se diseña y aplica un procedimiento bootstrap basado en una prueba estadística para seleccionar entre los posibles desfases pertinentes, logrando que esta contribución marque una diferencia interesante con respecto a otros procedimientos disponibles, que no proporcionan un marco de prueba de hipótesis.

## SUMMARY

In the general model description of a financial time series, the selection of the lags to be included is certainly a crucial first step, which is essential in forecasting and economic model building. Traditionally, autocorrelation and partial autocorrelation coefficients have been utilized in specifying the appropriate lags. However, it is known that zero autocorrelation processes may still exhibit a high order or nonlinear dependence, as in the case of some bilinear processes and even for pure deterministic chaotic models, among others. In general, the procedures based on the autocorrelation may yield misleading results when analyzing nonlinear models, and so might fail to detect important nonlinear relationships present in the data, limiting its usefulness in detecting appropriate delays especially in those scenarios where nonlinear phenomena are more the rule than the exception. Today it is generally accepted that the most important economic variables are nonlinear nature and that nonlinear relationships abound. From an econometric and statistical point of view, this situation has motivated the development of serial dependence tests, especially when there are statistical tools with many limitations to perform the selection of significant delays, which probably leads to choices made inappropriate values. Therefore in the present investigation, taking a different way than the autocorrelations and partial autocorrelation is proposed and developed a new nonparametric statistical procedure that identifies the significant delays that must be present in the description and development of models of series time of unknown nature, whether linear or nonlinear. It designs a bootstrap based statistical test for selecting among potential relevant lags. This later contribution makes an interesting difference with other available procedures, which do not provide a hypothesis testing framework.

---

# ÍNDICE GENERAL

<b>1. INTRODUCCIÓN</b>	<b>1</b>
1.1. Objetivos de la Tesis . . . . .	6
1.1.1. Objetivo general . . . . .	6
1.1.2. Objetivos específicos . . . . .	7
1.2. Metodología de Trabajo . . . . .	7
1.3. Justificación . . . . .	9
1.4. Alcance y Limitaciones de la Tesis . . . . .	10
1.5. Organización de la Tesis . . . . .	11
<b>2. SISTEMAS DINÁMICOS Y TEORÍA DEL CAOS</b>	<b>13</b>
2.1. Sistemas Dinámicos Deterministas . . . . .	14
2.1.1. Sistemas dinámicos deterministas no lineales . . . . .	19
2.2. Reconstrucción del Atractor y la Integral de Correlación. . . . .	22

2.2.1.	Integral de correlación . . . . .	25
2.2.2.	Procedimiento para la estimación de la integral de correlación . . . . .	27
2.3.	Sistemas Dinámicos Estocásticos . . . . .	29
2.3.1.	Distribución conjunta y marginal del proceso . . . . .	31
2.3.2.	Funciones de distribución condicionadas . . . . .	34
2.3.3.	Procesos estacionarios . . . . .	36
2.3.4.	Proceso de ruido blanco . . . . .	37
2.3.5.	Proceso caminata aleatoria . . . . .	38
2.3.6.	Modelos de procesos estocásticos no lineales . . . . .	38

**3. LA INTEGRAL DE CORRELACIÓN COMO IDENTIFICADOR DEL RETARDO SIGNIFICATIVO DE UNA SERIE DE TIEMPO 42**

3.1.	Generación de Vectores Retardados . . . . .	44
3.2.	Estimador de la Integral de Correlación . . . . .	46
3.3.	Identificación del Retardo Temporal Óptimo en Series de Tiempo Estacionarias . . . . .	50
3.4.	Identificación del Retardo Óptimo en los Residuales . . . . .	61
3.4.1.	Convergencia en probabilidad de la serie de los residuales al cuadrado . . . . .	63
3.4.2.	Convergencia en probabilidad de la integral de correlación. . . . .	64
3.4.3.	Procesos generados por modelos ARMA-GARCH . . . . .	67

**4. METODOLOGIA PARA LA VALIDACION EMPIRICA 70**

4.1.	Simulación Monte Carlo y Bootstrap . . . . .	71
------	--	----



4.1.1.	Método Monte Carlo. . . . .	72
4.1.2.	Método Bootstrap . . . . .	75
4.2.	Procedimientos para la Validación Empírica . . . . .	76
4.2.1.	Primer procedimiento de validación: Promedio de la $\hat{C}_{m,n}^{\tau}(\varepsilon)$ . . . . .	79
4.2.2.	Segundo procedimiento de validación: Prueba para identificar retardos en modelos lineales y no lineales. . . . .	80
4.2.3.	Procedimiento automático para seleccionar retardos . . . . .	82
4.2.4.	Procedimiento para seleccionar retardo con independencia del radio . . . . .	83
4.3.	Procedimiento para el Uso de las Técnicas Propuestas a los Residuos Estimados de un Modelo Ajustado. . . . .	84
4.4.	Software Utilizado en las Simulaciones Monte Carlo y Bootstrap . . . . .	87
<b>5.</b>	<b>RESULTADOS EMPÍRICOS DE LA APLICACIÓN DE LA TÉCNICA</b>	<b>89</b>
5.1.	Resultados de la Aplicación del Procedimiento 1: Identificación del Retardo a Partir del Valor Máximo del Promedio de $\hat{C}_{m,n}^{\tau}(\varepsilon)$ . . . . .	90
5.2.	Resultados Numéricos de la Identificación del Retardo a Partir del Valor Máximo del Promedio de $\hat{C}_{m,n}^{\tau}(\varepsilon)$ , Modificando el Valor del Parámetro $\varepsilon$ . . . . .	93
5.3.	Resultados de la Aplicación del Procedimiento 2: . . . . .	96
5.3.1.	Experimentos para verificar el tamaño de la prueba . . . . .	96
5.3.2.	Experimentos para verificar la potencia de la prueba . . . . .	99
5.3.3.	Resultados numéricos de la aplicación del procedimiento automático para seleccionar retardos . . . . .	101

5.4.	Resultados Numéricos de la Aplicación de las Técnicas Propuestas a los Residuos Estimados de un Modelo Ajustado. . . . .	104
5.5.	Otros Resultados de las Simulaciones . . . . .	108
<b>6.</b>	<b>APLICACIÓN DE LA TÉCNICA A SERIES NO ESTACIONARIAS</b>	<b>110</b>
6.1.	Procesos Lineales No Estacionarios . . . . .	111
6.1.1.	Vectores retardados en los procesos de caminata aleatoria . . . . .	114
6.1.2.	Vectores retardados en otros procesos lineales no estacionarios . . . . .	116
6.2.	Procesos No Lineales con Tendencia . . . . .	117
6.3.	La Estacionariedad y las Medidas de Reconstrucción. . . . .	118
6.4.	Aplicación Empírica de la Técnica Propuesta a Series de Tiempo Lineales No Estacionarias . . . . .	119
6.4.1.	Resultados de la aplicación del Procedimiento 1: Identificación del retardo a partir del valor máximo del promedio de $\hat{C}_{m,n}^{\tau}(\varepsilon)$ . . . . .	120
6.4.2.	Resultados de la aplicación del procedimiento 2: Prueba para identificar retardos en modelos no estacionarios lineales y no lineales. . . . .	123
6.5.	Futuras aplicaciones de la técnica propuesta a series de tiempo no estacionarias . . . . .	126
<b>7.</b>	<b>CONCLUSIONES Y CONSIDERACIONES FINALES</b>	<b>127</b>
7.1.	Identificación Correcta del Retardo Temporal Significativo . . . . .	129
7.2.	Identificación Correcta del Retardo Temporal Significativo con Distintos Valores del Parámetro $\varepsilon$ . . . . .	130
7.3.	Tamaño y Potencia de la Prueba Estadística Propuesta . . . . .	130

7.4.	Aplicación Automática del Procedimiento de Identificación . . . . .	132
7.5.	Uso de la Prueba Propuesta como Herramienta de Diagnóstico. . . . .	132
7.6.	Aplicación del Procedimiento a Series de Tiempo Lineales No Estacionarias y No Explosivas. . . . .	133
<b>A.</b>	<b>APENDICE DEL CAPITULO 5:</b>	
	<b>RESULTADOS TAMAÑOS DE MUESTRA PEQUEÑOS</b>	<b>135</b>
A.1.	Introducción . . . . .	135
A.2.	Resultados de la Aplicación del Procedimiento 1: Identificación del Retardo a Partir del Valor Máximo del Promedio de $\hat{C}_{m,n}^{\tau}(\epsilon)$ . . . . .	136
A.3.	Resultados de la Aplicación del Procedimiento 2: . . . . .	141
A.4.	Resultados Numéricos de la Aplicación del Procedimiento Automático para Seleccionar Retardos. . . . .	146
A.4.1.	Resultados de la identificación del retardo significativo a partir del valor máximo del promedio $\hat{C}_{m,n}^{\tau}(\epsilon)$ . . . . .	146
A.4.2.	Resultados del rendimiento de la prueba con muestras pequeñas . . . . .	147
A.5.	Resultados Numéricos de la Aplicación de las Técnicas Propuestas a los Residuos Estimados de un Modelo Ajustado. . . . .	148
.	<b>Bibliografía</b>	<b>153</b>

---

## ÍNDICE DE FIGURAS

2.1.	Evolución temporal de $x_{n+1} = x_n$ ; $\mu = -0.8$ , $x_1 = 0.08$ . . . . .	16
2.2.	Dada la condición inicial $X_0 = 0.25$ , a) muestra la órbita para $c = 0.5$ , b) muestra la órbita para $c = 2.0$ , c) muestra la órbita para $c = 3.3$ , d) muestra la órbita aperiódica para $c = 3.9$ .	21
2.3.	Gráficas de la serie de tiempo caótica generada con el modelo Logístico con $c = 4$ . . . . .	21
2.4.	Gráficas de la serie de tiempo de ruido blanco. . . . .	32
2.5.	Gráficas de la serie de tiempo camino aleatorio. . . . .	33
2.6.	Gráficas de la serie de tiempo camino aleatorio. (a) Evolución de la serie estacionaria, (b) Gráfica plana de dispersión $(X_{t-1}, X_t)$ serie estacionaria, (c) Evolución de la serie no estacionaria, (d) Gráfica plana de dispersión $(X_{t-1}, X_t)$ serie no estacionaria . . . . .	39
3.1.	Puntos finitos $X_{i,1}^\tau(\varepsilon)$ ; $m = 1$ y $X_{i,2}^\tau(\varepsilon)$ ; $m = 2$ , cercanos en dos vecindades de radio $\varepsilon$ distintas <sup>1</sup> . . . . .	47
3.2.	Vectores retardados obtenidos de la serie de tiempo generada con el modelo, $X_t = e_t \sim N(0, 1)$ , $m = 2$ , $\tau = 1, 2, 3, 4, 5, 6$ . No existe retardo del modelo. . . . .	52

3.3.	Vectores retardados obtenidos de la serie de tiempo generada con el modelo, $X_t = e_t + 0.8e_{t-3}^2; e_t \sim N(0, 1), m = 2, \tau = 1, 2, 3, 4, 5, 6, \tau_o = 3$ . . . . .	53
3.4.	Vectores retardados obtenidos de la serie de tiempo generada con el modelo, $X_t = e_t + 0.8X_{t-2}; e_t \sim N(0, 1), m = 2, \tau = 1, 2, 3, 4, 5, 6, \tau_o = 2.$ . . . . .	53
3.5.	Vectores retardados obtenidos de la serie de tiempo generada con el modelo, $X_t = 4X_{t-1}(1 - X_{t-1}), m = 2, \tau = 1, 2, 3, 4, 5, 6, \tau_o = 1.$ . . . . .	54
3.6.	Vectores retardados obtenidos de la serie de tiempo generada con el modelo, $X_t = e_t +  X_{t-1} ^{0.8}; e_t \sim N(0, 1), m = 2, \tau = 1, 2, 3, 4, 5, 6, \tau_o = 1.$ . . . . .	54
3.7.	Gráficas de la evolución de (a) serie generada con un modelo AR(1)+GARCH(1,1), (b) serie de los residuales al cuadrado. . . . .	68
6.1.	Indicadores Económicos Coyunturales de México <sup>3</sup> : a) PIB (precios corrientes), b) índice actividad industrial, base 2003, c) tasa de desempleo por nivel de instrucción, secundaria completa, d) inflación mensual. . . . .	112
6.2.	Vectores retardados obtenidos de la serie de tiempo generada con el modelo: $X_t = X_{t-1} + e_t; X_0 = 1.0, e_t \sim N(0, 1); m = 2, \tau = 1, 2, \dots, 6; \tau_0 = 1$ . . . . .	115
6.3.	Vectores retardados obtenidos de la serie de tiempo generada con el modelo: $X_t = X_{t-2} + e_t; X_0 = 1.0, X_1 = 2.266, e_t \sim N(0, 1),$ $m = 2, \tau = 1, 2, \dots, 6; \tau_0 = 2$ . . . . .	115
6.4.	Vectores retardados obtenidos de la serie de tiempo generada con el modelo: $X_t = X_{t-3} + e_t; X_0 = 1.0, X_1 = 2.266, X_2 = 1.249; e_t \sim N(0, 1), m = 2, \tau =$ $1, 2, \dots, 6; \tau_0 = 3$ . . . . .	116
6.5.	Vectores retardados obtenidos de la serie de tiempo generada con el modelo: $X_t = 0.5 + t + X_{t-1} + e_t; X_0 = 1.0 e_t \sim N(0, 1), m = 2, \tau = 1, 2, \dots, 6; \tau_0 = 1$ . .	117

6.6. Vectores retardados obtenidos de la serie de tiempo generada con el modelo:

$$X_t = 0.5 + t^2 + X_{t-1} + e_t; X_0 = 1.0 \quad e_t \sim N(0, 1); m = 2, \tau = 1, 2, \dots, 6; \tau_0 = 1 \dots 118$$

---

## ÍNDICE DE TABLAS

4.1. Modelos generadores de series de tiempo estacionarias . . . . .	77
4.2. Modelos generadores de series de tiempo no estacionarias . . . . .	78
4.3. Modelos para generar $\{X_t\}$ y filtro para los residuales $\{\hat{e}_t\}$ . . . . .	86
5.1. Promedios de $\hat{C}_{m,n}^\tau$ para procesos estacionarios . . . . .	91
5.2. Promedios de $\hat{C}_{m,n}^\tau$ para procesos estacionarios, $m = 3, T_1 = 50$ , variando el valor de $\epsilon = a\sigma_x$ . . . . .	93
5.3. Promedios de $\hat{C}_{m,n}^\tau$ para procesos estacionarios, $m = 3, T_2 = 100$ , variando el valor de $\epsilon = a\sigma_x$ . . . . .	94
5.4. Promedios de $\hat{C}_{m,n}^\tau$ para procesos estacionarios, $m = 3, T_3 = 200$ , variando el valor de $\epsilon = a\sigma_x$ . . . . .	94
5.5. Promedios de $\hat{C}_{m,n}^\tau$ para procesos estacionarios, $m = 3, T_4 = 500$ , variando el valor de $\epsilon = a\sigma_x$ . . . . .	95
5.6. Tamaño de la prueba para el Modelo 1 con $T = 500$ y $1000$ . . . . .	97

5.7. Tamaño de la prueba para el Modelo 2 con $T = 500$ y $1000$ . . . . .	98
5.8. Tamaño de la prueba para el Modelo 4 con $T = 500$ y $1000$ . . . . .	98
5.9. Potencia de la prueba para series generadas de tamaño $T_1 = 500$ . . . . .	99
5.10. Potencia de la prueba para series generadas de tamaño $T_2 = 1000$ . . . . .	100
5.11. Promedios de la $\hat{C}_{m,n}^\tau$ para el Modelo 11; $T = 500$ y $1000$ $\tau = 6, 10$ . . . . .	102
5.12. Resultados Tamaño de la Prueba, para el Modelo 11, $AR(10)$ . . . . .	103
5.13. Proporción de veces que algunas posibles estructuras de retardos son seleccionadas.	104
5.14. Tamaño de la prueba de la serie residual con el modelo $ARMA(1,1)-N(0,1)$ , $T_1 = 500$ . . . . .	105
5.15. Tamaño de la prueba de la serie residual con el modelo $ARMA(1,1)-N(0,1)$ . $T_2 = 1000$ . . . . .	105
5.16. Tamaño de la prueba de la serie residual con el modelo $ARMA(1,1)-N(0,1)$ . $T_3 = 2000$ . . . . .	106
5.17. Potencia de la prueba en la serie de residuales del modelo: $ARMA(1,0)-GARCH(1,1)$ . . . . .	107
5.18. Potencia de la prueba en la serie de residuales del modelo: $ARMA(0,1)-GARCH(1,1)$ . . . . .	107
5.19. Potencia de la prueba en la serie de residuales del modelo: $ARMA(1,1)-GARCH(1,1)$ . . . . .	107
5.20. Tiempo de CPU(segundos) consumido en las simulaciones de los procedimientos .	109
6.1. Promedios de $\hat{C}_{m,n}^\tau$ para procesos no estacionarios, $T = 500$ . . . . .	121
6.2. Promedios de $\hat{C}_{m,n}^\tau$ para procesos no estacionarios, $T_1 = 50$ . . . . .	122



6.3.	Promedios de $\hat{C}_{m,n}^\tau$ para procesos no estacionarios, $T_2 = 100$ . . . . .	122
6.4.	Promedios de $\hat{C}_{m,n}^\tau$ para procesos no estacionarios, $T_3 = 200$ . . . . .	123
6.5.	Potencia de la prueba para series no estacionarias generadas de tamaño, $T_1 = 500$ . . . . .	124
6.6.	Potencia de la prueba para series no estacionarias generadas de tamaño $T_2 = 1000$ . . . . .	124
6.7.	Potencia de la prueba para muestras de tamaño pequeño modelos no estacionarios. . . . .	125
A.1.	Promedios de la Integral de Correlación para procesos estacionarios, $T_1 = 50$ . . . . .	138
A.2.	Promedios de la Integral de Correlación para procesos estacionarios, $T_2 = 100$ . . . . .	139
A.3.	Promedios de la Integral de Correlación para procesos estacionarios, $T_3 = 200$ . . . . .	140
A.4.	Tamaño de la prueba para muestras de tamaño pequeño Modelo 1. . . . .	142
A.5.	Tamaño de la prueba para muestras de tamaño pequeño Modelo 2. . . . .	143
A.6.	Tamaño de la prueba para muestras de tamaño pequeño Modelo 4. . . . .	144
A.7.	Potencia de la prueba para muestras de tamaño pequeño . . . . .	145
A.8.	Promedios de la Integral de Correlación para el Modelo 11. . . . .	146
A.9.	Resultados del Tamaño de la Prueba, para el Modelo 11, AR(10) . . . . .	147
A.10.	Tamaño de la prueba de la serie residual con el modelo ARMA(1,1)-N(0,1) . . . . .	149
A.11.	Tamaño de la prueba de la serie residual con el modelo ARMA(1,1)-N(0,1) . . . . .	149
A.12.	Tamaño de la prueba de la serie residual con el modelo ARMA(1,1)-N(0,1) . . . . .	150
A.13.	Potencia de la prueba en la serie de residuales del modelo ARMA(1,0)-GARCH(1,1) . . . . .	151
A.14.	Potencia de la prueba en la serie de residuales del modelo ARMA(0,1)-GARCH(1,1) . . . . .	151
A.15.	Potencia de la prueba en la serie de residuales del modelo ARMA(1,1)-GARCH(1,1) . . . . .	151

# GLOSARIO

## Nomenclatura

- AR( $p$ ) — Proceso Estocástico Autorregresivo de orden  $p$ .
- ARMA( $p,q$ ) — Proceso Estocástico Autorregresivo de orden  $p$  con Media Móvil de orden  $q$ .
- ARCH( $p$ ) — Proceso Estocástico Autorregresivo Condicionalmente Heteroscedástico de orden  $p$ .
- BDS — Prueba estadística no paramétrica basada en la estimación de la integral de correlación en varias dimensiones, para probar la independencia de una serie de tiempo. Prueba desarrollada por Brock, Dechert y Scheinkman en 1986.
- BL( $p,q,m,h$ ) — Proceso Estocástico Bilinial de orden  $(p, q)$  en la parte lineal y  $(m, h)$  en los productos bilineales.
- CPU — Unidad de Procesamiento Central.
- DGP — Proceso Generador de Datos.
- GARCH( $p,q$ ) — Proceso Estocástico Generalizado Autorregresivo Condicionalmente Heteroscedástico de orden  $(p, q)$ .
- G-P — Grassberger y Procaccia, autores del estadístico insesgado para estimar la integral de correlación.
- i.i.d. — Proceso Estocástico Independiente e Igualmente Distribuido.
- MA( $q$ ) — Proceso Estocástico de Media Móvil de orden  $q$
- $N(\mu, \sigma^2)$  — Proceso Estocástico Normal con media  $\mu$  y varianza  $\sigma^2$
- NAR( $p$ ) — Proceso Estocástico No Lineal Autorregresivo de orden  $p$
- NARMA( $p,q$ ) — Proceso Estocástico No Lineal Autorregresivo de orden  $p$  con Media Móvil de orden  $q$

NMA( $q$ )	—	Proceso Estocástico No Lineal de Media Móvil de orden $q$
PIB	—	Producto Interno Bruto.
Residuales	—	Estimación de los errores verdaderos (Residuos, $e$ )

## Símbolos

$B$	—	Número de realizaciones Bootstrap.
$C_m^\tau(\varepsilon)$	—	Integral de Correlación en una distancia $\varepsilon$
$\hat{C}_{m,n}^\tau(\varepsilon)$	—	Estimador de la Integral de Correlación en una distancia $\varepsilon$
$C_{m,n}^\tau(e_t^{2,m,\tau}; \varepsilon)$	—	Integral de Correlación de la serie de residuales al cuadrado en una distancia $\varepsilon$ .
$\hat{C}_{m,n}^\tau(\hat{e}_t^{2,m,\tau}; \varepsilon)$	—	Estimador de la Integral de Correlación de la serie de residuales estimados. al cuadrado en una distancia $\varepsilon$
$\{e_t\}$	—	Serie de residuales exactos.
$\{\hat{e}_t\}$	—	Serie de residuales estimados.
$\varepsilon$	—	Distancia entre dos puntos del espacio de reconstrucción en $\mathbb{R}^m$ .
$H(\bullet)$	—	Función Heavside que toma el valor de 1 si dos observaciones están dentro de una distancia $\varepsilon$ , y 0 en otro caso.
$N^{m,\tau}$	—	Número de vectores retardados que cumplen estar a una distancia $\varepsilon$ .
$N_r$	—	Número de replicas para la simulación Monte Carlo.
$\ \bullet\ _\infty$	—	Norma infinito.
$\sigma_x$	—	Desviación estándar de la serie de tiempo $\{x_t\}$ .
$T$	—	Número total de observaciones de la serie de tiempo.
$\tau$	—	Retardo Temporal de una Serie de Tiempo.
$\tau_o$	—	Retardo Temporal Óptimo (significativo) de una Serie de Tiempo.
$\theta$	—	Vector de parámetros del modelo estocástico.
$\hat{\theta}$	—	Vector de parámetros estimados del modelo estocástico.
$V_\varepsilon(X_i^{m,\tau})$	—	Vecindad de un punto del espacio de reconstrucción en $\mathbb{R}^m$ de radio $\varepsilon$ .

- $\{x_t\}_{t=1}^T$  — Serie de tiempo escalar con  $T$  número de observaciones
- $X_i^{m,\tau}$  —  $i$ -ésimo vector de retardo ( $m$ -historia). Punto del espacio de reconstrucción en  $\mathbb{R}^m$ .

---

---

# CAPÍTULO 1

---

## INTRODUCCIÓN

Desde la última década, del siglo XX a la fecha, una gran parte de la investigación económica y financiera ha estado relacionada con el desarrollo y aplicación de técnicas para el análisis de series de tiempo no lineales, con el fin de determinar su naturaleza y de proponer modelos que reproduzcan el comportamiento de la dinámica subyacente.

El interés por este tipo de investigación se ha visto incrementado porque, al igual que como ocurre con las series de tiempo que surgen en otras áreas de la ciencia, las series económicas y financieras son generadas por sistemas dinámicos complejos, algunos plenamente identificados como no lineales y otros, la gran mayoría, cuya naturaleza es casi imposible de determinar a priori. En muchos casos, las series de tiempo experimentales que se obtienen son de tamaño pequeño y de precisión finita, por lo que no es tan fácil distinguir la presencia de una dinámica no lineal conteniendo componentes estocásticas.

El objetivo más importante de llevar a cabo el análisis de una serie de tiempo es conocer la ley de la dinámica que gobernó su generación, por lo cual, es necesario entender por completo al sistema dinámico físico, químico, biológico o económico subyacente del cual se derivó la serie. Cuando

no existe una teoría en la cual apoyarse para lograr ese objetivo o dicha teoría se encuentra lejos de estar completa, y con lo único que se cuenta son los propios datos de la serie, esto es, las observaciones o mediciones obtenidas para llevar a cabo el análisis de éstas, se aplica comúnmente el siguiente procedimiento:

- 1). Reconocer las características o hechos importantes de los datos observados;
- 2). Construir un modelo empírico de la serie de tiempo, incorporando lo más que se pueda la teoría fundamental disponible;
- 3). Verificar que el modelo construido sea capaz de capturar las características o hechos identificados en el paso 1).
- 4). Buscar mejorar el modelo aún más, si fuera necesario.

Fundamentalmente, un modelo empírico de series de tiempo representa una hipótesis relacionada con la probabilidad de su transición en el tiempo, esto es, en la dinámica. El paso 1) del procedimiento presentado, proporcionaría el tipo y la estructura de los datos para seleccionar el modelo que se va a requerir, y el paso 3) probaría la bondad de ajuste del modelo seleccionado. El paso 2) permitiría facilitar la tarea de especificación de una clase bastante amplia de modelos, de los cuales se deberá seleccionar el mejor. Un requerimiento obvio es que esa clase debe ser lo suficientemente amplia para que incluya los modelos capaces de capturar las características identificadas en el paso 1); amplitud que por supuesto se puede ver restringida por los recursos computacionales y, herramientas numéricas y estadísticas que se dispongan para la investigación.

Otra de las funciones importantes de un modelo empírico que no debe pasarse por alto es que permite agudizar la percepción del conocimiento del fenómeno o proceso que se quiere estudiar. Si bien es cierto que un modelo empírico nunca puede sustituir a la teoría subyacente, si puede contribuir al desarrollo de esta última, y al mismo tiempo, cada avance que se lleve a cabo en la ampliación de dicha teoría puede contribuir al desarrollo de un modelo empírico más adecuado.

En la descripción de un modelo empírico general de una serie de tiempo que representa el comportamiento de fenómenos económicos, el primer paso crucial para hacer una buena selección del modelo experimental, es ciertamente la selección de los retardos temporales que deben ser incluidos en el modelo. Tradicionalmente, los coeficientes de autocorrelación y de autocorrelación parcial han sido utilizados para la especificación de los retardos apropiados. Sin embargo, está bien establecido en [Granger, 1983], que procesos con autocorrelación cero pueden exhibir un alto orden de dependencia o dependencia no lineal. Este es el caso de algunos procesos bilineales e incluso con modelos deterministas caóticos puros, entre otros. Un sistema determinista no lineal y caótico puede generar una serie de tiempo que tiene la apariencia de aleatoria y que por tanto es imposible de predecir.

En general, los procedimientos basados en la autocorrelación pueden ser engañosos para los modelos no lineales, y por lo tanto pueden no detectar importantes relaciones no lineales en los datos, siendo, por tanto, de limitada utilidad en la detección de retardos de tiempo apropiados, especialmente en aquellos escenarios donde los fenómenos no lineales son más la regla que la excepción. Que las variables económicas más importantes sean de naturaleza no lineal y que además abunden relaciones no lineales entre ellas son actualmente hechos generalmente aceptados. Desde un punto de vista econométrico y estadístico, esta situación ha motivado el desarrollo de pruebas de dependencia serial no lineal. Para mayor detalles ver: [Granger and Anderson, 1978], [McLeod and Li, 1983], [Engle, 1982], [Bollerslev, 1986], [Subba Rao, 1980], [Hinich, 1982], [Ashley, Altug and Patterson, 1997], [Ashley and Patterson, 1989], y [Tjostheim, 1996].

Una de las aportaciones más importantes en el desarrollo de pruebas estadísticas para probar no linealidad estocástica y caos en los datos económicos se debe a Grassberger y Procaccia [G-P, 1983], quienes introdujeron el concepto de Integral de Correlación y desarrollaron un procedimiento para organizar los datos de la serie en  $m$ -historias (información pasada). Este procedimiento se basa en el conocido *método de vectores de retardo* propuesto por [Takens, 1981], para llevar a cabo la reconstrucción del espacio fase de un sistema dinámico desconocido.

La construcción a partir de una serie temporal de un espacio de estados equivalente al espacio

de estados del sistema original, denominado espacio reconstruido, es la base de muchas de las pruebas estadísticas que se aplican en el análisis de series temporales no lineales. Sin embargo estas pruebas, por construcción, no están diseñadas para seleccionar retardos relevantes, y esto explica en parte la escasez de técnicas no paramétricas para investigar la dependencia del retardo, sin importar la naturaleza del proceso lineal o no lineal, lo cual es un aspecto que es desconocido en la mayoría de los casos prácticos. Algunas notables excepciones son: [Granger and Lin, 1994], [Tjostheim and Auestad,1994], [Granger et al, 2003], y [Matilla-García and Ruiz,2009].

Una característica común de la mayoría de las técnicas desarrolladas es el uso de las medidas basadas en la entropía para identificar el retardo correcto. Sin embargo, en la actualidad no existe un método riguroso y definitivo para determinar el valor óptimo del retardo temporal. Dos de los métodos más utilizados por los analistas, pero que no producen buenos resultados para series de tiempo no lineales de las cuales se desconoce el sistema dinámico que las generó, son: el método que identifica como retardo óptimo el correspondiente al primer cero de la función de autocorrelación ([Barnett and Chen,1988]; [Ramsey et al,1990]), con lo que se pretende conseguir que los elementos de la serie que constituyen el  $m$ -vector de retardos no presenten dependencia lineal; y el método que utiliza el retardo donde se obtiene el primer mínimo de la función media de información mutua (AMI, Averaged Mutual Information), ([Fraser and Swinney,1986]), que mide cualquier correlación entre los datos de la serie temporal, pero que requiere del conocimiento de la distribución de probabilidad de los datos. Existen otros trabajos de investigación donde se proponen otros tipos de métodos, como el de “factor de sensibilidad” utilizado por [Buzug et al,1999] y los relacionados con la “descomposición de valores singulares”, aplicada al algoritmo G-P por [Albano et al, 1988].

En el trabajo de investigación que se presenta en esta tesis, se toma un camino diferente a los trabajos antes mencionados, y con base en la ya muy conocida integral de correlación (en [G-P, 1983], se propone y desarrolla un nuevo procedimiento estadístico no paramétrico para identificar los retardos temporales que deben ser usados en los modelos de naturaleza desconocida (lineales o no lineales). En adición a esta técnica, se diseña un procedimiento bootstrap basado en una prue-



ba estadística para hacer una selección adecuada entre desfases potencialmente relevantes. Esta última contribución hace una diferencia interesante con otros procedimientos disponibles, que no proporcionan un marco de prueba de hipótesis.

La integral de correlación inicialmente fue diseñada para medir la dimensión fractal de datos caóticos, convirtiéndose posteriormente en una herramienta importante para analizar algunas propiedades de las series de tiempo generadas en las ciencias físicas y naturales (ver por ejemplo [Figueiredo et al,2011] y [Mathew and Picu,2011], como últimas contribuciones). Ha sido ampliamente utilizada en las finanzas y la macroeconomía, básicamente para probar independencia serial. La contribución más relevante es la muy conocida prueba BDS [Brock et al,1990], la cual se basa precisamente en el hecho de que la integral de correlación factoriza cuando los elementos de una serie de tiempo son iid (independientes e idénticamente distribuidos). Curiosamente, la integral de correlación también ha sido utilizada [Hiemstra and Jones,1994] para detectar posibles relaciones no lineales de causalidad entre series de tiempo. Sin embargo, ha habido algunos comentarios seriamente críticos a la prueba, ([Tjostheim,1996], [De Lima,1996], [Barnett et al, 1997] entre otros), en cuanto a una sobreestimación mayor del nivel correcto de ella, por lo tanto tiende a reportar una tasa alta de rechazos falsos. Esto puede explicarse, en parte, debido al impacto no trivial de algunas de las decisiones que se llevan a cabo con anterioridad sobre la selección de diversos parámetros decisivos que deben ser considerados para poder utilizar la prueba. Desafortunadamente, existen herramientas estadísticas muy limitadas (la mayoría de ellas basadas en experimentos Monte Carlo masivos) acerca de su selección, lo que hace que probablemente se elijan valores inapropiados.

Como se mostrará a lo largo de los siguientes capítulos de esta tesis, la prueba que se propone y desarrolla, es robusta para tomar aquellas decisiones previas, particularmente minimizando el papel que desempeñan cuando se elije un retardo adecuado en el modelo seleccionado. Estrechamente relacionada, otra propiedad especialmente relevante, es que el procedimiento no se ve afectado por el problema de “la maldición de la dimensionalidad”, ya que no se basa en la estimación no paramétrica de diferentes momentos. Más bien, se basa únicamente en la integral de correlación, que originalmente requiere la incorporación de un parámetro de retardo de tiempo que describa

mejor las propiedades dinámicas básicas del proceso. Esta característica, que no fue utilizada en las pruebas de tipo BDS, ya que el tiempo de retraso se fija en uno, permite considerar escenarios donde la dependencia no necesariamente está contenida en las clásicas  $m$ -historias formadas consecutivamente, sino que ahora se pueden retrasar estos vectores adecuadamente y por lo tanto, hacer una búsqueda a través de varios retardos, hasta encontrar el mejor.

Otra contribución de esta investigación es que la técnica propuesta puede ser usada para examinar la estructura de retardos en los modelos residuales, y por lo tanto, parece ser útil como herramienta potencial de diagnóstico. Estos dos últimos atributos hacen que el enfoque sea ciertamente general, tanto para los macroeconometristas y los analizadores financieros.

## **1.1. Objetivos de la Tesis**

Tomando como base los resultados que se tienen de varios trabajos de investigación sobre el concepto, interpretación estadística, cálculo y uso de la integral de correlación, así como la hipótesis de la existencia de un retardo temporal significativo en una serie de tiempo observada, de acuerdo con los resultados empíricos de [Matilla-García et al,2004], en este trabajo de tesis se presenta el desarrollo y validación teórica y empírica de una nueva técnica que permite identificar el retardo temporal significativo de una serie de tiempo generada por un proceso estocástico o determinista, lineal o no lineal.

### **1.1.1. Objetivo general**

Proponer, desarrollar y validar teórica y empíricamente, un procedimiento para la Estimación de la Integral de Correlación  $C_{m,n}^{\tau}(\varepsilon)$ , como la herramienta principal para identificar el retardo temporal óptimo,  $\tau_0$ , de una serie de tiempo observada.

### 1.1.2. Objetivos específicos

Con el fin de obtener el objetivo general se proponen los siguientes objetivos específicos:

- Revisar y estudiar los conceptos teóricos necesarios para introducir al estimador de la Integral de Correlación,  $\hat{C}_{m,n}^\tau(\varepsilon)$ , como la herramienta principal que permite, identificar el retardo temporal óptimo,  $\tau_0$ , de una serie de tiempo observada.
- Demostrar teóricamente que dada una serie de tiempo  $\{x_t\}$ , el valor máximo del estimador de la Integral de Correlación,  $\hat{C}_{m,n}^\tau(\varepsilon)$ , es obtenido cuando para una  $m$  y  $\varepsilon$ , fijas dadas, se calcula  $\hat{C}_{m,n}^\tau(\varepsilon)$  con el retardo óptimo,  $\tau_0$ .
- Demostrar empíricamente que se obtiene el resultado teórico, al aplicarse la técnica a series de tiempo generadas por diferentes modelos de procesos estocásticos lineales o no lineales, y procesos puramente deterministas caóticos, con la característica de ser estacionarios.
- Demostrar teórica y empíricamente que la técnica también puede ser usada para identificar el retardo temporal óptimo de la serie de tiempo de los residuos al cuadrado, obtenidos al filtrar con un proceso  $ARMA(p, q)$ , una serie generada con un modelo  $ARMA - GARCH$  dado.

## 1.2. Metodología de Trabajo

Para llevar a cabo las actividades relacionadas con este trabajo de investigación y alcanzar los objetivos mencionados en la sección anterior, se sigue la metodología que se detalla a continuación:

1. La revisión y estudio de los conceptos teóricos necesarios para conocer las características y propiedades de la integral de correlación, así como el manejo y comportamiento de su estimador,  $\hat{C}_{m,n}^\tau(\varepsilon)$ , se lleva a cabo mediante una búsqueda y selección de material bibliográfico relacionado con los temas de desarrollo y uso de herramientas de análisis para series de

tiempo dinámicas y estocásticas no lineales, particularmente de investigaciones de reciente publicación.

2. La verificación y evaluación del comportamiento de las técnicas propuestas en la identificación del retardo temporal significativo de series de tiempo estacionarias, lineales, no lineales y caóticas, se lleva a cabo a través de simulaciones. En específico utilizando las técnicas de simulación Monte Carlo y Bootstrap, en las cuales se consideran los siguientes puntos:

- a) Selección de un conjunto de modelos de procesos estocásticos y/o determinísticos estacionarios, lineales o no lineales, y algunos lineales no estacionarios, que generan las series de tiempo a estudiar.
- b) Definición del tamaño de las series (número de observaciones a generar) a estudiar y del conjunto de parámetros de inmersión  $(m, \tau, \varepsilon)$  para llevar a cabo la construcción de los vectores de retardo requeridos.
- c) Definición del número de replicas para la simulación Monte Carlo y el número de muestras Bootstrap para ejecutar la prueba estadística que se propone, y obtener el tamaño y la potencia de la prueba.

3. Por lo que respecta a la verificación y evaluación del comportamiento de las técnicas propuestas al aplicarlas a series de tiempo de los residuos al cuadrado, obtenidos al filtrar con un proceso  $ARMA(p, q)$ , una serie generada con un modelo  $ARMA - GARCH$ , se incluye dentro de las simulaciones mencionadas en el punto anterior, la técnica de ajuste para estimar los parámetros del modelo propuesto, así como el cálculo de la series de residuales al cuadrado.

4. La implantación computacional de los algoritmos de simulación correspondientes, se realiza dentro del entorno de trabajo del software MATLAB.

### 1.3. Justificación

El mercado financiero ha sido y es objeto de estudio por parte de numerosos investigadores y profesionales de las finanzas que intentan encontrar pautas de regularidad que les permita entender el comportamiento de los activos y, si fuera posible, incluso realizar predicciones. Esto ha llevado a proponer y establecer una enorme cantidad de modelos que, en cierta medida, imitan el comportamiento de las series financieras.

Desde hace años se utilizan modelos estocásticos como los GARCH, que consideran la información pasada de la serie y su volatilidad como factores explicativos del comportamiento presente. De igual forma se están utilizando otro tipo de modelos estocásticos no lineales como los NAR, NMA y los Biliniales que capturan una buena parte de la dinámica no lineal. Por otra parte y desde la perspectiva de los sistemas dinámicos deterministas no lineales, también se ha venido trabajando con estos, con el fin de detectar la existencia de caos en las series financieras, aunque en la actualidad se usan en un sentido más amplio para capturar alguna componente determinista no lineal en los datos de la serie.

Como se señalaba anteriormente, existen muy pocas técnicas para analizar series de tiempo no lineales de tipo continuo y discreto observadas, sean éstas producto de procesos o modelos deterministas estocásticos y, en una buena parte de las existentes, su procedimiento está basado de una u otra forma, en la generación de los vectores de retardo y el cálculo de la dimensión de correlación, que en el contexto de la teoría del caos es una medida asociada con el atractor de un sistema dinámico determinista, que es invariante bajo el operador evolución de dicho sistema.

Una de las razones por la que la dimensión de correlación es comúnmente aplicada, además de su propiedad de invarianza, es porque puede ser obtenida fácilmente a partir de la estimación de la integral de correlación,  $C_m(\varepsilon)$ , la cual es definida como la probabilidad de que dos puntos del atractor, tomados aleatoriamente según una medida de probabilidad  $\mu$  definida sobre el atractor, estén a una distancia menor que un valor  $\varepsilon$  dado.

Para hacer una buena estimación de la integral de correlación  $C_m(\varepsilon)$ , es necesario conocer a priori el valor de tres parámetros,  $(m, \tau, \varepsilon)$  que corresponden a la dimensión de inmersión, el retardo temporal y la distancia entre dos vectores del espacio fase de reconstrucción respectivamente. Aun cuando, en la teoría de reconstrucción de la dinámica de un sistema, la dimensión de correlación,  $m$ , es independiente de la selección del parámetro de retardo temporal  $\tau$ , en la práctica, la estimación de este para series de tiempo finitas sí depende de la selección. Además de que un modelo que genera series de tiempo que tienen una dependencia no lineal, debe pensarse como un mecanismo que consta de dos partes: una relación funcional no lineal y un retardo, en este caso fijo.

En el análisis de series temporales económicas observadas, la mayoría de las veces, el segundo elemento, el retardo temporal, se olvida o desprecia en la investigación, siendo precisamente que, tanto para la reconstrucción del sistema dinámico subyacente a la serie, como para el propio modelo en el caso de la predicción, existe un retardo significativo que es el que aporta mayor información para lograr la identificación del patrón de comportamiento del sistema.

Dada la importancia de que los interesados en el análisis de series financieras, cuenten con una técnica que les permita identificar el retardo temporal óptimo, en este trabajo de tesis, considerando la hipótesis de la existencia de un retardo significativo, se propone y desarrolla una nueva técnica sencilla de implantar y utilizar, para que a partir de una serie generada por diferentes modelos de procesos dinámicos estocásticos y deterministas, lineales y no lineales, así como para las series de residuales al cuadrado resultantes del ajuste de un modelo a una serie dada, se identifique el retardo significativo que proporciona mayor información del comportamiento de la dinámica.

## **1.4. Alcance y Limitaciones de la Tesis**

El alcance de este trabajo comprende la propuesta, desarrollo y validación teórica y empírica de una nueva técnica para identificar el retardo temporal óptimo de una serie de tiempo observada o generada por procesos de tipo estocástico o determinista, lineales o no-lineales.

La validación de la nueva técnica se obtendrá a partir de su aplicación a series de tiempo generadas por procesos estacionarios, lineales y no lineales, no estacionarios lineales, y a las series de tiempo de los residuales al cuadrado, obtenidos al filtrar con un proceso  $ARMA(p, q)$  a una serie generada con modelos del tipo  $ARMA - GARCH$ .

## 1.5. Organización de la Tesis

La presentación del desarrollo del presente trabajo de investigación se organiza como sigue: en el Capítulo 2 se presenta el marco teórico necesario para sustentar el enfoque de esta investigación, esto es, se introducen los conceptos básicos de la teoría de Sistemas Dinámicos tanto deterministas como estocásticos, así como teoremas y resultados que permiten llevar a cabo la reconstrucción del atractor de una serie de tiempo determinista; la definición e interpretación de la Integral de Correlación,  $C_m(\varepsilon)$ , y el estadístico que la estima,  $\hat{C}_{m,n}^\tau(\varepsilon)$ .

En el Capítulo 3 se presentan los fundamentos teóricos que permiten establecer formalmente la hipótesis de que dada una serie de tiempo  $\{x_t\}$  estacionaria, el valor máximo de la estimación de la integral de correlación se obtiene cuando ésta es calculada para una  $m$  y  $\varepsilon$  fijas, con el retardo óptimo  $\tau_0$ . Esto es, que existe un retardo temporal significativo que aporta la mayor información de la dinámica subyacente a la serie de tiempo. Adicionalmente se presenta la demostración teórica, de que la técnica propuesta también se puede aplicar a la serie de tiempo de los residuales al cuadrado, obtenidos al filtrar con un procedimiento  $ARMA(p, q)$  la serie de tiempo generada por un proceso  $ARMA - GARCH$ .

En el Capítulo 4 se presentan los procedimientos y algoritmos de simulación con los métodos Monte Carlo y Bootstrap aplicados para llevar a cabo la validación empírica de la técnica y prueba estadística que se propone en la investigación. En el Capítulo 5 y su apéndice, se presentan los resultados de las simulaciones Monte Carlo y Bootstrap llevadas a cabo con modelos generadores de series de tiempo estacionarias estocásticas y deterministas, lineales o no lineales, y de series de tiempo residuales, de tamaño grande (500 y 1000 observaciones) y tamaño pequeño (50, 100 y 200

observaciones) respectivamente.

En el capítulo 6 se discute el comportamiento que la técnica propuesta pudiera tener al ser aplicada directamente a series de tiempo generadas por procesos no estacionarios, lineales o no lineales. Se introduce y valida empíricamente la hipótesis de que el estadístico que estima la integral de correlación,  $\hat{C}_{m,n}^{\tau}(\varepsilon)$ , sería una buena herramienta para identificar el retardo temporal significativo de una serie de tiempo generada por modelos lineales no estacionarios que representan procesos no explosivos, como por ejemplo el de caminata aleatoria, y se deja abierta, para próximas investigaciones, la posibilidad de que dicha herramienta pudiera ser utilizada para series de tiempo no estacionarias, con tendencia lineal o no lineal explosiva, en donde el estimador de la integral de correlación no puede obtenerse numéricamente debido a problemas de desbordamiento en el aritmética de punto flotante de la computadora.

Por último en el Capítulo 7 se presentan las conclusiones y consideraciones finales de esta investigación, así como de futuras líneas de investigación.



---

---

## CAPÍTULO 2

---

### SISTEMAS DINÁMICOS Y TEORÍA DEL CAOS

Desde 1990 a la fecha, la Economía Financiera avanza aceleradamente gracias al empleo de procedimientos computacionales intensivos en el análisis de fenómenos financieros, los cuales eran anteriormente realizados mediante técnicas analíticas. Una importante aportación cuantitativa a la Economía Financiera ha sido la Teoría del Caos, la cual postula en términos informales que ciertos procesos no lineales sencillos, puramente deterministas, exhiben comportamientos dinámicos extraordinariamente complejos y aparentemente erráticos.

Cuando se trata de investigar la evolución en el tiempo del sistema dinámico que representa un fenómeno financiero, en lugar de un conjunto de ecuaciones, lo que en realidad se tiene es una serie de datos en distintos momentos del tiempo. En la literatura financiera se conocen como *series de tiempo*. Estos datos aparentan ser aleatorios, sin embargo, es posible que la apariencia de aleatoriedad se trate del resultado de la evolución de un cierto número de variables, que responden a ciertas leyes teóricas.

El interés que surge en este caso es saber, si ciertas series aparentemente aleatorias, como las bursátiles, pudieran estar generadas por procesos deterministas. De manera que una vez que se

especifique el proceso que las generó, este pudiera ser utilizado para realizar predicciones.

La aplicación de la Teoría del Caos en el contexto de la predicción bursátil es una cuestión compleja y en ocasiones bastante polémica. En este capítulo 2 de la tesis se plantean de manera introductoria, algunos conceptos básicos que permiten sustentar el enfoque de esta investigación, el cual se basa en la reconstrucción de la dinámica asintótica de un sistema dinámico de las realizaciones de un proceso de series de tiempo observadas. Para ampliar la información relacionada con las definiciones y conceptos de la Teoría del Caos aquí presentados consultar: [Broer,Takens,2009], [Kantz and Schreiber,2005], [Solé and Manrubia,2001], [Clinton,2003], [Small,2005] y las relacionadas con procesos y modelos estocásticos [Ross,1996] [Franses and van Dijk, 2004], [Peña, 2005].

En la sección 2.1 se presentan conceptos teóricos básicos de los sistemas dinámicos deterministas, el espacio fase, los atractores y su dimensión, así como otros conceptos relacionados con el comportamiento errático o caótico de las trayectorias de evolución de una serie temporal. En la sección 2.2 se introducen los conceptos de reconstrucción del espacio fase y el de la integral de correlación, desde el punto de vista de sistemas dinámicos determinísticos discretos. De igual forma se discute el uso de ambos conceptos en el análisis de series de tiempo no lineales o caóticas. En la sección 2.3 se presentan los conceptos teóricos básicos de los sistemas dinámicos estocásticos y se discuten las *series de tiempo* que son generadas por procesos puramente aleatorios o combinados con una parte determinista.

## 2.1. Sistemas Dinámicos Deterministas

Los sistemas dinámicos son sistemas que presentan una variación de las magnitudes que los definen en el tiempo, esto es,

**Definición 2.1** *Un sistema dinámico  $\vec{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , es aquel que describe su evolución sobre un espacio de estado dado, mediante un conjunto de  $n$  ecuaciones diferenciales vectoriales de la*

forma:

$$\frac{d\vec{X}(t)}{dt} = \vec{F}(\vec{X}(t); \beta). \quad (2.1)$$

Equivalentemente cuando la evolución temporal no es continua y se asumen únicamente valores discretos, existen  $n$  ecuaciones en diferencias de la forma:

$$\vec{X}_{n+1} = F(\vec{X}_n; \beta). \quad (2.2)$$

El conjunto de variables del sistema está representado por  $\vec{X}$ , que es un punto de un conjunto abierto  $U \subset \mathbb{R}^n$ , la evolución temporal está representada por  $t \in \mathbb{R}$  y es la variable que parametriza las trayectorias (o soluciones) del vector  $\vec{X}$ . El vector  $\beta$  es una colección de  $p$  parámetros,  $\beta \in \mathbb{R}^p$ , que se fijan en cada evolución del sistema, y  $F \in \mathbb{R}^n$ , es un vector de funciones. El espacio  $\mathbb{R}^n$ , donde no aparece el tiempo, sino las variables  $\vec{X}$  dependientes se denomina *espacio de fases* (o *espacio de estados*). La geometría de las soluciones en el espacio fase proporciona toda la información sobre la dinámica del sistema.

La relación (2.2) indica que el punto  $\vec{X}_{t+1}$  depende de la localización del sistema en el tiempo anterior y del conjunto de parámetros  $\beta$ , esto es, que si se conociera el verdadero estado de  $\vec{X}_t$  y a  $F(\bullet)$ , entonces se podría predecir en forma exacta el valor futuro de  $\vec{X}_{t+1}$ . En general, en un modelo discreto se utilizan períodos: primer periodo, segundo período,... . Si se considera el tiempo discreto  $t = 1, 2, \dots$ , entonces una función de esta variable  $X : \mathbb{N} \rightarrow \mathbb{R}^n$  es simplemente una sucesión de vectores,  $X_1, X_2, X_3, \dots$

Un ejemplo sencillo de un sistema dinámico unidimensional discreto es:

$$X_{n+1} = \mu X_n. \quad (2.3)$$

Si se parte de un cierto punto inicial  $x_1$ , la trayectoria u órbita del sistema estaría dada por  $orb(x_1) = (x_1, x_2, x_3, x_4, \dots)$  que se obtendría de iterar (2.3) considerando el punto inicial, esto es,  $x_2 = \mu x_1$ ;  $x_3 = \mu x_2 = \mu^2 x_1$ ;  $x_4 = \mu x_3 = \mu^3 x_1$ ;  $\dots$ . Por lo que

$$x_k = \mu^k x_1. \quad (2.4)$$

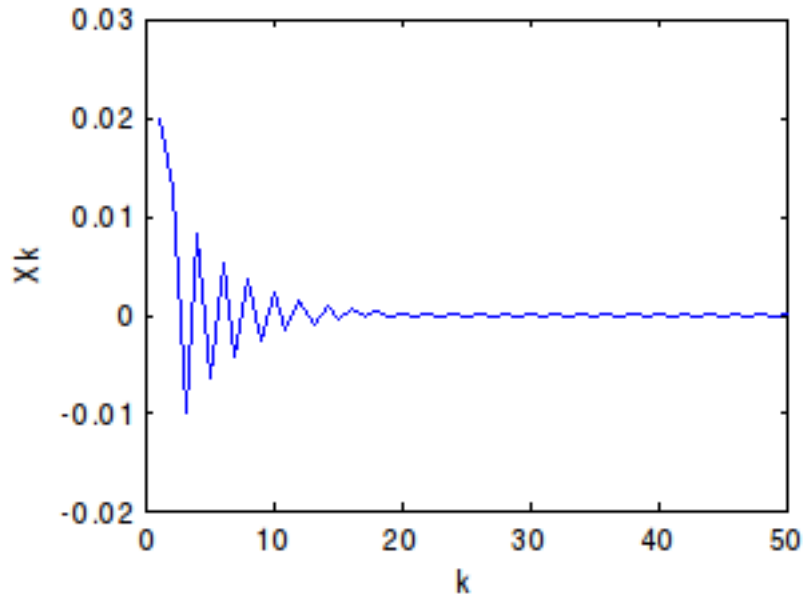


Figura 2.1: Evolución temporal de  $x_{n+1} = x_n$ ;  $\mu = -0.8$ ,  $x_1 = 0.08$

es la solución analítica de (2.3) para cualquier iteración conociendo el punto inicial.

Además para este caso sencillo, en función del valor de  $\mu$  se puede describir de manera fácil la dinámica del sistema:

- Si  $|\mu| < 1$  el sistema tiende al punto final  $x_f = 0$ , de forma monótona decreciente si  $\mu > 0$ , y de forma oscilante si  $\mu \leq 0$
- Si  $|\mu| > 1$  el sistema diverge a infinito cuando  $k \rightarrow \infty$ , de forma monótona creciente u oscilante dependiendo del signo de  $\mu$ , como en el caso anterior.
- Si  $|\mu| = 1$  el sistema permanece en el punto inicial  $x_0$  para  $\mu > 0$ , y oscila entre  $x_0$  y  $-x_0$  si  $\mu < 0$

La figura 2.1 representa en coordenadas cartesianas  $(k, x_k)$  la evolución temporal del sistema (2.4) para  $\mu = -0.8$  y  $x_1 = 0.08$ . El espacio de fases en este caso es unidimensional, y F se reduce a una sola función denotada por  $f$ .

Cuando la función  $f$  es invertible, es decir, cuando existe la función inversa de  $f(\vec{x}; \beta)$ ,  $f^{-1}(\vec{x}; \beta)$ , se puede estudiar la secuencia doblemente infinita

$$\{\dots, f^{-n}(x_1; \beta), \dots, f^{-1}(x_1; \beta), x_1, f(x_1; \beta), \dots, f^n(x_1; \beta), \dots\}. \quad (2.5)$$

Si  $f$  no es invertible, entonces sólo es posible estudiar la secuencia infinita

$$\{x_1, f(x_1; \beta), \dots, f^n(x_1; \beta), \dots\}. \quad (2.6)$$

Tradicionalmente, los economistas han utilizado los modelos discretos directamente. Se ha mostrado que aun los modelos más sencillos poseen gran riqueza y complejidad, y son capaces de representar más acertadamente la realidad económica. Además, al discretizar el tiempo es posible introducir elementos estocásticos de manera relativamente sencilla.

**Definición 2.2** *Un sistema dinámico discreto de primer orden es una ecuación de la forma  $X_{t+1} = f(X_t, t, \beta)$ . Si la variable  $t$  no aparece explícitamente, se dice que el sistema es **autónomo**. El sistema es de **orden**  $m$  si es de la forma:*

$$X_{t+m} = f(X_{t+m-1}, X_{t+m-2}, \dots, X_t, t, \beta). \quad (2.7)$$

Es decir, el estado de la variable depende no sólo del estado en el período anterior, sino de estados en los  $m$  períodos anteriores. Es común denotar la función **no autónoma**

$$f(X_{t+m-1}, X_{t+m-2}, \dots, X_t, t; \beta). \quad (2.8)$$

por

$$f(X_{t+m-1}, X_{t+m-2}, \dots, X_t; \beta). \quad (2.9)$$

Si la función  $f$  es *lineal*, se dice que el sistema es **lineal**.

Cuando se trata con sistemas dinámicos complejos que presentan comportamientos complicados en el espacio de fase, no se puede encontrar una solución analítica para la evolución del sistema. En estos casos se utilizan técnicas de estudio del sistema a nivel local, esto es, cerca de los denominados *puntos fijos* (o *puntos de equilibrio*).

El primer paso en el estudio del sistema dinámico es entonces el análisis de sus puntos fijos y el tipo de estabilidad que presentan, por lo que es importante definir lo que es un punto fijo.

**Definición 2.3** *Dado el sistema dinámico descrito por  $X_{t+1} = f(X_t)$ , se dice que  $X^*$  es un **punto fijo o punto de equilibrio** si se cumple  $X^* = f(X^*)$*

El análisis cualitativo de los puntos fijos da una idea de la evolución del sistema y de la estabilidad de éstos.

**Definición 2.4** *Supóngase que el punto fijo  $X^*$  de un sistema dinámico discreto dado por  $X_{t+1} = f(X_t)$  satisface la siguiente condición: existe un número  $\delta > 0$  tal que se cumple*

$$\|X_0 - X^*\| < \delta \Rightarrow \lim_{t \rightarrow \infty} X_t = X^*, \quad (2.10)$$

*es decir, la solución converge al punto fijo (equilibrio) cuando  $t \rightarrow \infty$ , siempre y cuando el valor inicial esté suficientemente cerca de  $X^*$ . Se dice entonces que el punto fijo es asintóticamente estable.*

De manera análoga se puede definir inestabilidad en un punto fijo.

Se dice que un punto  $X^*$  atrae al movimiento  $X_t$  si  $\lim_{t \rightarrow \infty} X_t = X^*$ , y que es un atractor (o que es asintóticamente estable), si atrae a todos los movimientos que se inician en su entorno. Se distinguen dos tipos de atractores simples: los puntos fijos y los ciclos límites, que representan el comportamiento a que llega un estado estable o que se repite continuamente.

Se dice que el sistema es límite cuando es asintóticamente estable en el sentido convencional, o lo que es equivalente, al cabo de cierto tiempo, se encuentra atrapado en el ciclo, definiendo un estado estacionario no constante, sino oscilatorio. En estos sistemas es importante fijar la atención en el comportamiento a largo plazo, concentrándose en las propiedades asintóticas del sistema dinámico, el atractor del sistema, esto es, en el conjunto compacto de puntos al que convergen todas las trayectorias originadas en un rango de condiciones iniciales. En los sistemas simples, las

órbitas en el espacio de fase se aproximan a un punto fijo, y los atractores son puntuales, o bien son órbitas con ciclo límite y atractores periódicos. En ambos casos el sistema es predecible.

### 2.1.1. Sistemas dinámicos deterministas no lineales

La teoría de sistemas dinámicos no lineales discretos en general y las ecuaciones en diferencia no lineales en particular, aun cuando han estado experimentando un gran desarrollo, todavía se encuentran incompletas. En la etapa actual de su desarrollo, el enfoque pragmático de recurrir a la experimentación por medio de las computadoras digitales, ha desempeñado un papel muy importante en el proceso de aprendizaje. Sin embargo, es importante que, para llevar a cabo este tipo de experimentos, se conozcan algunas definiciones de trabajo que son análogas a su contraparte de los sistemas dinámicos de tiempo continuo.

Algunos sistemas dinámicos deterministas no lineales, esto es, de la forma (2.1) y (2.2) donde  $F$  está formado por funciones no lineales, presentan un comportamiento errático o caótico y frecuentemente tienen un atractor en el espacio de fase el cual es *extraño*. En un *atractor extraño*, las trayectorias vagan de manera errática y son altamente sensitivas a condiciones iniciales, se verifica, entre otras cosas, un conjunto innumerable de ciclos de periodos arbitrariamente grandes y un conjunto de trayectorias no periódicas.

La versión discreta del modelo logístico es el mejor ejemplo de un sistema discreto no lineal sencillo que posee un comportamiento caótico para ciertos valores del parámetro  $c$  en el intervalo  $[0, 4]$ . Dentro de estos valores el sistema lleva al intervalo  $[0, 1]$  en sí mismo.

$$X_{t+1} = f(X_t) = cx_t(1 - x_t). \quad (2.11)$$

los puntos fijos de (2.11) están dados al resolver,

$$x^* = f(x^*) = cx^*(1 - x^*). \quad (2.12)$$

Por lo que (2.11) tiene dos puntos fijos:

$$x_1^* = 0; \quad x_2^* = \frac{c-1}{c}. \quad (2.13)$$

Si  $\pm c \in (0, 1)$ , entonces  $x_1^*$  es el único punto fijo dentro de  $[0, 1]$  que además es asintóticamente estable. Se puede verificar fácilmente que dada cualquier condición inicial  $x_0 \in [0, 1]$  su órbita converge a 0.

Si  $1 < c < 3$ , entonces se tienen los dos punto fijos (2.13) en  $[0, 1]$ ; y en este caso  $x_1^* = 0$  es inestable y  $x_2^* = \frac{c-1}{c}$  es asintóticamente estable, por lo que dada  $x_0 \in [0, 1]$  su órbita converge a  $\frac{c-1}{c}$ .

Si  $3 < c < 4$ , entonces existen 2-ciclos en el sistema (2.11), por lo que es necesario encontrar los puntos de período 2, esto es,  $x$  puntos que satisfacen  $f^2(x) = x$ , para lo cual se debe resolver la ecuación  $f(f(x)) - x = 0$ , lo que lleva al siguiente polinomio de orden 4, del cual ya se conocen dos raíces  $x_1^* = 0$ ;  $x_2^* = \frac{c-1}{c}$  ya que estos puntos están en el intervalo  $[0, 1]$  y necesariamente son raíces dado que  $f(x) = x \Rightarrow f(f(x)) = x$

$$c^3x^4 - 2c^3x^3 + (c^3 + c^2)x^2 - (c^2 - 1)x = 0. \quad (2.14)$$

Factorizando ambas raíces del polinomio utilizando el algoritmo de la división, se obtienen los puntos de periodo 2:

$$x_3^* = \frac{(c+1) + \sqrt{(c-3)(c+1)}}{2c}; \quad x_4^* = \frac{(c+1) - \sqrt{(c-3)(c+1)}}{2c}. \quad (2.15)$$

Estos dos puntos no existían cuando  $c < 3$ , ya que en este caso las raíces no son reales; para  $c = 3 \Rightarrow x_2^* = x_3^* = x_4^* = \frac{2}{3}$  y, si  $c > 3$  se tiene que  $x_3^*$  y  $x_4^*$  son puntos de periodo 2 y por lo tanto existe un 2-ciclo. Analizando la estabilidad de estos dos puntos de período 2, se encuentra que la órbita es estable para valores de  $c$  que satisfacen  $|4 + 2c - c^2| < 1$ , o bien si  $3 < c < 1 + \sqrt{6}$ .

En la figura 2.2 se muestran las órbitas correspondientes (los puntos se unen para mayor claridad), tomando como punto inicial  $x_0 = 0.25$ , y los valores de  $c = 0.5$ ,  $c = 2.0$ ,  $c = 3.3$  y  $c = 3.9$ . Las tres primeras imágenes, a)-c) muestran órbitas con comportamiento estable o cíclico, sin embargo la imagen d) muestra que la órbita tiene un comportamiento aperiódico. Además, si se toman órbitas que se forman con condiciones iniciales muy cercanas, estas se separan con rapidez exponencial, por lo que se dice que el sistema exhibe un comportamiento caótico. La figura 2.3(a)



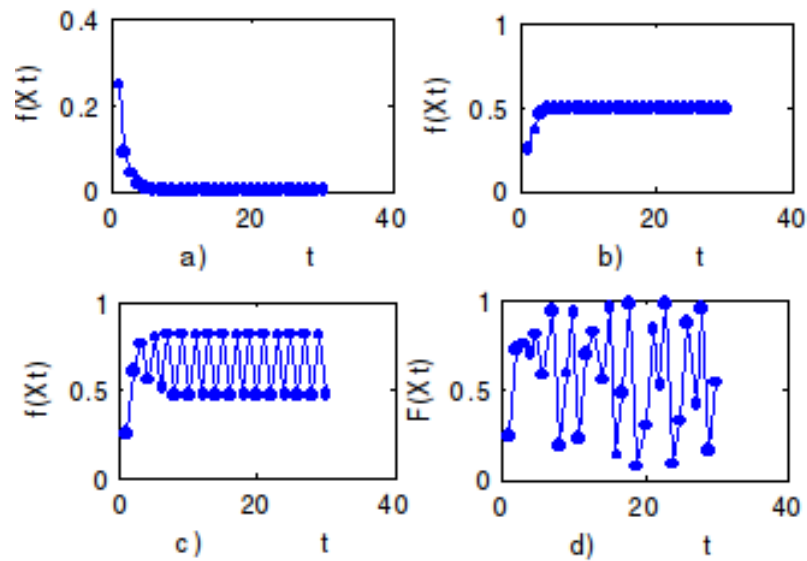
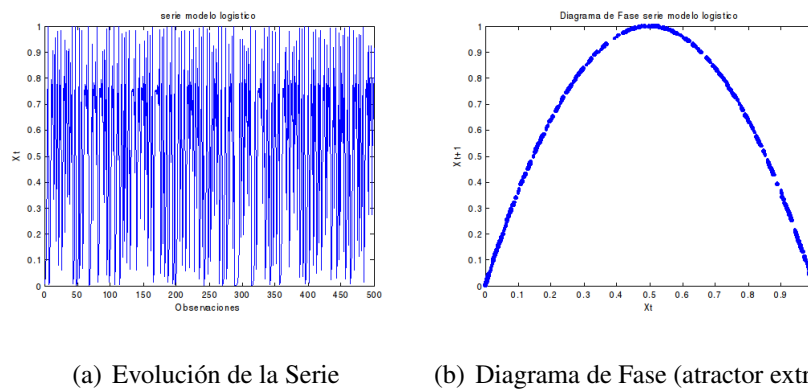


Figura 2.2: Dada la condición inicial  $X_0 = 0.25$ , a) muestra la órbita para  $c = 0.5$ , b) muestra la órbita para  $c = 2.0$ , c) muestra la órbita para  $c = 3.3$ , d) muestra la órbita aperiódica para  $c = 3.9$ .

presenta la gráfica que muestra la órbita para  $c = 4.0$  (caso caótico) y puede verse que aparentemente no existe ninguna pauta en la evolución dinámica de la serie, pareciera completamente aleatoria, aun cuando en la ecuación (2.11) no hay nada de aleatorio. La figura 2.3(b) corresponde al diagrama de fase (gráfica en el plano  $X_t, X_{t-1}$ ), y que corresponde a una parábola sencilla (atractor extraño). La dimensión de un atractor es una parte importante en el conocimiento de un



(a) Evolución de la Serie

(b) Diagrama de Fase (atractor extraño)

Figura 2.3: Gráficas de la serie de tiempo caótica generada con el modelo Logístico con  $c = 4$ .

sistema, ya que es una medida de información necesaria para especificar un punto en el atractor con una aproximación dada. Se dice que un sistema dinámico determinista está definido sobre un

atractor  $A$ , que tiene dimensión  $m$ , la cual es más pequeña que la dimensión del sistema,  $n$ . Más aún, la dimensión del atractor es una cota inferior del número de variables necesarias para modelar el sistema. Un *atractor extraño* típicamente tiene una dimensión no entera (fractal).

## 2.2. Reconstrucción del Atractor y la Integral de Correlación.

Uno de los problemas que surge en la práctica es como obtener conocimiento del comportamiento caótico de un sistema a partir de una serie de tiempo observada, ya que se desconoce el verdadero estado del sistema. Una de las herramientas más poderosas con las que cuenta la investigación experimental de sistemas no lineales observados a través de series de tiempo, que ha sido aplicada en muchos campos de la ciencia, es la inmersión y reconstrucción del espacio de estado del proceso subyacente. Existen teoremas rigurosos acerca de la posibilidad de reconstruir un espacio de estado a partir de una serie de tiempo escalar. Para una muestra de datos discreta, todas las aproximaciones a la reconstrucción del espacio de estado pueden ser consideradas como variantes a la técnica central de inmersión del retardo temporal.

Conectar el espacio fase o el vector del espacio de estado  $X_t$  de variables dinámicas del sistema físico con la serie de tiempo  $\{x_t\}$  medida en experimentos, es un problema abordado por primera vez por [Packard et al, 1980], quien demostró que es posible reconstruir un vector del espacio de estado multidimensional usando retardos o adelantos temporales (desplazamientos negativos o positivos), con las mediciones escalares de la serie de tiempo  $\{x_t\}$ , es decir, se puede construir un vector sustituto  $V_t = (x_t, x_{t\pm\tau}, x_{t\pm 2\tau}, \dots)$  con las mediciones escalares  $x_t$ , considerando una opción particular de coordenadas (en las cuales cada componente del vector sustituto es una versión cambiada del tiempo de las demás) con el cambio de tiempo entre coordenadas adyacentes al mismo.

Las mediciones  $x_t$  que forman la serie de tiempo, son el resultado de aplicar la función de medición  $\varphi$  a los estados verdaderos  $X_t$ ; el efecto de aplicar esta función se puede describir en general como una proyección con una distorsión concurrente. Es claro que una proyección causa una pérdida

de información, por lo que podría parecer imposible reconstruir la verdadera evolución de  $X_t$  únicamente conociendo  $x_t$ , sin embargo, también puede resultar que exista un flujo de información de las variables inobservadas a las variables observadas que compensarían la proyección. Como el efecto de las distorsiones no puede ser deshecho en general, es de gran importancia contar con valores que sean invariantes bajo distorsiones, por lo que se sugiere que de alguna forma se transformen las mediciones de la serie en vectores de desplazamiento reconstruidos  $V_t \in \mathbb{R}^m$  convenientes.

Desde el punto de vista de los verdaderos estados, esto correspondería a la acción de inmersión de un mapeo de coordenadas  $\varphi : M \rightarrow \mathbb{R}^m$ , el cual es por supuesto desconocido, junto con una función  $f : M \rightarrow \mathbb{R}$  suave. La dimensión de inmersión  $m$  tiene que ser seleccionada de manera adecuada, y  $M \subset \mathbb{R}^d$  es una variedad compacta, donde  $d$  es la dimensión del espacio de estado verdadero del sistema.

Esta estructura constituida como la forma general de modelos matemáticos de sistemas dinámicos deterministas que generan series de tiempo, está pensada en el siguiente sentido: las posibles evoluciones u órbitas del sistema dinámico son secuencias de la forma  $\{x_n = \varphi^n(x_0)\}$ ;  $n \in (\mathbb{Z})^+$ . Se asume que lo que se observa o mide, en cada tiempo  $n$  no es todo el estado completo  $x_n$ , sino un número real  $y_n = f(x_n)$ , donde  $f$  es la función que asigna a cada estado el valor de su medición u observación cuando el sistema está en ese estado. De esta forma para cada evolución  $\{x_n = \varphi^n(x_0)\}$  existe una serie de tiempo  $Y$  de mediciones (observaciones) sucesivas  $\{y_n = f(\varphi^n(x_0)) = f(x_n)\}$ .

A  $\varphi$  se le piden las siguientes condiciones: i) que sea uno a uno lo que implica que  $\varphi(M)$  contiene la evolución del sistema y no tiene ninguna auto-intersección; ii) que se preserve la estructura diferencial de la variedad  $M$ . Una  $\varphi$  que cumple estas dos condiciones es llamada una inmersión. Esto es equivalente a que  $\varphi$  sea un difeomorfismo, esto es, un mapeo suave para el cual también existe una inversa suave.

Actualmente existen métodos que provienen de la teoría de sistemas dinámicos, que permiten a

través de la reconstrucción de los estados del sistema, reconstruir el atractor en el espacio desfasado (un espacio de menor dimensión al original), utilizando las ideas de la topología diferencial de [Takens,1981].

**Teorema 2.1** [Inmersión de Takens] Si  $M$  es una variedad compacta y  $\Phi : M \rightarrow \mathbb{R}^{2n}$  tal que  $\Phi \in C^2$ , entonces  $\Phi$  genéricamente, es una inmersión, esto es, un difeomorfismo entre el espacio de fase reconstruido y el original, si  $m > 2n + 1$

Aquí  $m$  es la dimensión de inmersión y  $n$  es la dimensión de la variedad compacta que contiene el atractor real del sistema, lo que implica que la dimensión del atractor reconstruido es la misma que aquella del atractor real, toda vez que es invariante bajo un cambio suave de coordenadas. Sin embargo, lo anterior supone que hay un número infinito de medidas infinitamente exactas de la única componente.

El significado del teorema de reconstrucción es el siguiente: para una serie de tiempo  $X = \{x_n\}$  se considera la secuencia de sus  $m$ -dimensional *vectores de reconstrucción*

$$\{x_n, x_{n\pm\tau}, \dots, x_{n\pm(m-1)\tau}\}_n \subset \mathbb{R}^m. \quad (2.16)$$

Esta secuencia es difeomórfica a la evolución del sistema determinista que generó la serie de tiempo  $X$ , si las siguientes condiciones se satisfacen:

- La dimensión  $m$  es suficientemente grande, por ejemplo, mayor que dos veces la dimensión del espacio de estado  $M$ .
- La pareja  $(\varphi, f)$ , es genérica en el sentido de que pertenecen al subconjunto abierto y denso que se menciona en el teorema de Inmersión de Takens.

Los resultados del teorema de reconstrucción consideran series de tiempo infinitamente grandes, por lo que el tiempo de muestreo,  $t$  y el retardo temporal  $\tau$ , pueden ser seleccionados casi arbitrariamente, sin embargo, en la práctica las series de tiempo que se analizan son finitas y de tamaño pequeño, y la selección óptima del retardo temporal es un problema difícil de resolver.

Parece intuitivamente claro que el valor de  $x_{t\pm k\tau}$  y el correspondiente  $x_{t\pm(k+1)\tau}$ , con  $k = 1, 2, \dots, m - 1$ , estarán correlacionados, ya que ambos se han generado a través de cierta ley determinista. También es claro que el valor de  $\tau$  no puede ser en la práctica totalmente arbitrario, ya que si el valor de  $\tau$  es muy reducido, los puntos  $x_{t\pm k\tau}$  y  $x_{t\pm(k+1)\tau}$  se encontrarán trivialmente correlacionados (con una clara dependencia lineal), y si el valor de  $\tau$  es excesivamente grande, la correlación entre ambos puntos, puede ser de hecho nula.

### 2.2.1. Integral de correlación

Desde un punto de vista matemático, se requiere que las series de tiempo  $\{x_t\}$  estén definidas para toda  $t$ , asumiendo que el tiempo es discreto, ya que cuando se obtienen las observaciones, éstas se registran en un tiempo discretizado de cualquier manera. De esta forma, una serie de tiempo es una función  $X$  que asigna a cada tiempo  $t \in \mathbb{Z}^+$  un valor  $x_t$ , el cual se asume es un número real.

Una propiedad importante que una serie de tiempo puede o no tener, es la de *estacionariedad*. Se dice que una serie de tiempo es estacionaria, en sentido estricto, sí, para cada  $m$ , existe una medida de probabilidad  $\mu_m \in \mathbb{R}^m$ , la cual describe la densidad de los vectores de reconstrucción  $m$ -dimensionales. De manera más formal lo anterior se puede traducir como que una serie de tiempo es estacionaria en sentido estricto, si para cada función continua no negativa  $g : \mathbb{R}^m \rightarrow \mathbb{R}$  se tiene

$$\int_{\mathbb{R}^m} g d\mu_m = \lim_{t \rightarrow \infty} \frac{1}{t} \sum_{i=0}^{t-1} g(x_i, \dots, x_{i\pm(m-1)\tau}). \quad (2.17)$$

Donde  $\mu_m$  es llamada la medida de reconstrucción  $m$ -dimensional. Los promedios, como el que aparece en el lado derecho de la ecuación (2.17), también pueden denotarse como  $E_i[g(x_i, \dots, x_{i\pm(m-1)\tau})]$ . Ya que como se mencionó anteriormente, las series de tiempo observadas nunca tienen una longitud infinita, los vectores de reconstrucción se pueden interpretar (de manera aleatoria) como intervalos  $\{x_t\}_{t=t_0}^{t=t_1}$  de estas series de tiempo. Tal y como se pudiera concluir desde un punto de vista estadístico, que los elementos de la población A son en promedio mayores que los de la población B, a partir de muestras lo suficientemente grandes de estas pobla-

ciones, se podría concluir de ese mismo modo de dos segmentos, lo suficientemente grandes, de una serie de tiempo estacionaria cuyas medidas de reconstrucción son significativamente diferentes, [Takens,2003].

Para una serie de tiempo estacionaria con medida de reconstrucción  $\mu_m$ , la integral de correlación en una distancia  $\varepsilon$ ,  $C_m^\tau(\varepsilon)$ , se define como la medida  $(\mu_m \times \mu_m)$  del conjunto  $\{(x, y) \in \mathbb{R}^m \times \mathbb{R}^m | dist(x, y) \leq \varepsilon\}$ . La  $dist(x, y)$  es considerada como la norma infinito,  $\|\bullet\|_\infty$ . La integral de correlación es una propiedad de la distribución espacial del conjunto de variables de estado de un sistema dinámico.

**Definición 2.5** *Dado un sistema dinámico  $(X, f)$  y un  $\varepsilon > 0$ , la integral de correlación, se define como la probabilidad de que dos puntos del atractor,  $x, y$ , tomados aleatoriamente, según una función de probabilidad  $\mu_m$  definida sobre este atractor, estén a una distancia menor que  $\varepsilon$ , esto es,*

$$C_m^\tau(\varepsilon) = \iint_{xy} H(\varepsilon - \|x - y\|_\infty) d\mu_m(x) d\mu_m(y) = P[\|x - y\|_\infty \leq \varepsilon]. \quad (2.18)$$

donde  $\mu_m(x)$  y  $\mu_m(y)$  son la distribución incondicional de  $x$  y  $y$  respectivamente.  $H(\bullet)$  denota la función Heavside, la cual toma el valor de 1 si ambas observaciones están dentro de una distancia  $\varepsilon$ , y 0 si esto no ocurre.

Considerando que

$$\begin{aligned} C_m^\tau(\varepsilon) &= P[\|x - y\|_\infty \leq \varepsilon] \\ &= P[|x_1 - y_1| \leq \varepsilon, \dots, |x_m - y_m| \leq \varepsilon] \\ &= C_{m|m-1}^\tau(\varepsilon) C_{m-1}^\tau(\varepsilon) \\ &= C_{m|m-1}^\tau(\varepsilon) C_{m-1|m-2}^\tau(\varepsilon) \cdots C_{2|1}^\tau(\varepsilon) C_1^\tau(\varepsilon), \end{aligned} \quad (2.19)$$

donde  $C_{m|m-1}^\tau(\varepsilon) = P[|x_m - y_m| \leq \varepsilon / |x_{m-1} - y_{m-1}| \leq \varepsilon]$ .

[G-P, 1983], propusieron un estadístico insesgado para estimar la integral de correlación, también

conocido como *suma de correlaciones*,

$$\begin{aligned}\hat{C}_{m,n}^\tau(\varepsilon) &= \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n H(\varepsilon - \|X_i^{m,\tau} - X_j^{m,\tau}\|_\infty), \\ H(Z) &= \begin{cases} 1 & \text{si } Z \geq 0 \\ 0 & \text{si } Z < 0 \end{cases} \quad (2.20) \\ X_i^{m,\tau} &= \{x_i, x_{i-\tau}, x_{i-2\tau}, \dots, x_{i-(m-1)\tau}\}, \text{ la } i\text{-ésima } m\text{-historia retardada de } x_i \\ \|X_i^{m,\tau} - X_j^{m,\tau}\|_\infty &= \max_{0 \leq k \leq m-1} \{|x_{i-k\tau} - x_{j-k\tau}|\}.\end{aligned}$$

$\hat{C}_{m,n}^\tau(\varepsilon)$  es un estadístico U generalizado ([Serfling, 1980], Capítulo 5); [Denker and Keller, 1983], que estima la integral de correlación (2.18)

$$C_m^\tau(\varepsilon) = \lim_{n \rightarrow \infty} \hat{C}_{m,n}^\tau(\varepsilon). \quad (2.21)$$

En el caso de que el límite de (2.18) exista,  $\hat{C}_{m,n}^\tau(\varepsilon)$  es interpretado como la fracción del número total de parejas de vectores de retardo  $(X_i^{m,\tau}, X_j^{m,\tau})$ , cuya distancia es menor que  $\varepsilon$  en una inmersión del espacio  $\mathbb{R}^m$ . Esto significa que el método de estimación está basado en tomar aleatoriamente un punto  $X_i^{m,\tau} \in \mathbb{R}^m$  (dentro del espacio fase reconstruido) y estudiar como evoluciona el número de puntos situados dentro de una vecindad de radio  $\varepsilon$  de él, esto es,  $V_\varepsilon(X_i^{m,\tau})$ .

### 2.2.2. Procedimiento para la estimación de la integral de correlación

Para estimar la integral de correlación de una serie de tiempo experimental, es necesario reconstruir la dinámica del sistema generador de dicha serie. Como se mencionó en la introducción, el cálculo de la integral de correlación propuesto en el algoritmo G-P está basado en el teorema de inmersión de Takens presentado en la Sección 2.2. El procedimiento que proponen G-P consta de los siguientes dos pasos:

**Paso 1.-** Establecer valores de los parámetros  $m$ , dimensión de inmersión y  $\tau$ , tiempo de retardo, para convertir la serie de tiempo escalar compuesta por un número finito de observaciones que siguen alguna función de distribución  $\mu_m, \{x_1, x_2, \dots, x_N\}$ , en un conjunto de  $m$ -historias

$$X_i^{m,\tau} = \{x_i, x_{i-\tau}, \dots, x_{i-(m-1)\tau}\}. \quad (2.22)$$

donde  $\tau$  es el tiempo de desfase entre elementos sucesivos del vector de retardos.

En (2.22) el tiempo  $(m-1)\tau$  generado para cada vector de inmersión es la distancia de desfase entre el primer y último elemento del vector de inmersión (también denominada en otras investigaciones como la longitud de ventana de la inmersión).

**Paso 2.-** Calcular la estimación de la integral de correlación

$$\hat{C}_{m,n}^{\tau}(\varepsilon) = \frac{1}{\binom{n}{2}} \sum_{i=1}^{n-1} \sum_{j=i+1}^n H(\varepsilon - \|X_i^{m,\tau} - X_j^{m,\tau}\|_{\infty}). \quad (2.23)$$

donde  $n = N - (m-1)\tau$  es el número de  $m$ -historias con  $\tau$  como desfase del tiempo, que pueden formarse con  $N$  observaciones. En (2.23),  $H$  es la función Heavside, tal que  $H(X_i^{m,\tau}, x_j^{m,\tau})$ ;  $i = 1 + (m-1)\tau, \dots, N-1$ ,  $j = i+1, \dots, N$  toma el valor de 1 si ambas observaciones están dentro de una distancia  $\varepsilon$  una de otra y el valor de 0 en caso contrario.

Dado que  $P[\|X_i^{m,\tau} - X_j^{m,\tau}\|_{\infty} \leq \varepsilon]$  es la probabilidad marginal de cada par de puntos, del conjunto  $\mathbf{X}^{m,\tau} = \{\mathbf{X}_1^{m,\tau}, \mathbf{X}_2^{m,\tau}, \dots, \mathbf{X}_n^{m,\tau}\}$ ,  $\hat{C}_{m,n}^{\tau}(\varepsilon)$  en la ecuación (2.23), es la probabilidad conjunta de todos los pares de puntos en el conjunto  $\mathbf{X}^{m,\tau}$  que satisfacen la condición de  $\varepsilon$ .

**Ejemplo 2.1** Sea  $\{x_t\}_{t=1}^N$ , si  $m = 2$  y  $\tau = 2$ , entonces se generan  $n = N - 2$  vectores retardados (2-historias) de la forma (2.22) y aplicando a (2.23) se tiene:

$$\begin{aligned} \mathbf{X}^{2,2} &= \{X_3^{2,2}, \dots, X_N^{2,2}\} \\ X_i^{2,2} &= \{x_i, x_{i-2}\}; \quad i = 3, \dots, N \\ \hat{C}_{2,N-2}^2(\varepsilon) &= \frac{1}{\binom{N-2}{2}} \sum_{i=3}^{N-1} \sum_{j=i+1}^N H(\varepsilon - \|X_i^{2,2} - X_j^{2,2}\|_{\infty}) \\ \|X_i^{2,2} - X_j^{2,2}\|_{\infty} &= \max_{0 \leq k \leq 1} \{|x_{i-2k} - x_{j-2k}|\} \\ &= \max\{|x_i - x_j|, |x_{i-2} - x_{j-2}|\} = d_{max_{ij}}. \end{aligned}$$

$$\begin{aligned} P[\|X_i^{2,2} - X_j^{2,2}\|_{\infty} \leq \varepsilon] &= P[d_{max_{ij}} \leq \varepsilon] \\ \hat{C}_{2,N-2}^2(\varepsilon) &= P[\|X_3^{2,2} - X_4^{2,2}\|_{\infty} \leq \varepsilon, \|X_3^{2,2} - X_5^{2,2}\|_{\infty} \leq \varepsilon, \\ &= \dots, \|X_{N-1}^{2,2} - X_N^{2,2}\|_{\infty} \leq \varepsilon]. \\ &= P[d_{max_{34}} \leq \varepsilon, d_{max_{35}} \leq \varepsilon, \dots, d_{max_{(N-1)N}} \leq \varepsilon]. \end{aligned}$$



si  $m = 3$  y  $\tau = 2$ , entonces se generan  $n = N - 4$  vectores retardados (3-historias) de la forma (2.22) y aplicando a (2.23) se tiene:

$$\begin{aligned}
\mathbf{X}^{3,2} &= \{X_5^{3,2}, \dots, X_N^{3,2}\} \\
X_i^{3,2} &= \{x_i, x_{i-2}, x_{i-4}\}; \quad i = 5, \dots, N \\
\hat{C}_{3,N-4}^2(\varepsilon) &= \frac{1}{\binom{N-4}{2}} \sum_{i=5}^{N-1} \sum_{j=i+1}^N H(\varepsilon - \|X_i^{3,2} - X_j^{3,2}\|_\infty) \\
\|X_i^{3,2} - X_j^{3,2}\|_\infty &= \max_{0 \leq k \leq 2} |x_{i-2k} - x_{j-2k}| \\
&= \max\{|x_i - x_j|, |x_{i-2} - x_{j-2}|, |x_{i-4} - x_{j-4}|\} \\
&= d_{max_{ij}}.
\end{aligned}$$

$$\begin{aligned}
P[\|X_i^{3,2} - X_j^{3,2}\|_\infty \leq \varepsilon] &= P[d_{max_{ij}} \leq \varepsilon] \\
\hat{C}_{3,N-4}^2(\varepsilon) &= P[\|X_5^{3,2} - X_6^{3,2}\|_\infty \leq \varepsilon, \|X_5^{3,2} - X_7^{3,2}\|_\infty \leq \varepsilon, \\
&= \dots, \|X_{N-1}^{3,2} - X_N^{3,2}\|_\infty \leq \varepsilon]. \\
&= P[d_{max_{56}} \leq \varepsilon, d_{max_{57}} \leq \varepsilon, \dots, d_{max_{(N-1)N}} \leq \varepsilon].
\end{aligned}$$

La estimación de la integral de correlación se utiliza en el cálculo de la dimensión de correlación y dimensión fractal que son medidas que permiten observar la dinámica de un atractor extraño y la geometría del objeto (ver [G-P, 1983]).

En este trabajo,  $\hat{C}_{m,n}^\tau(\varepsilon)$ , se propone y valida como una medida que permite identificar el retardo temporal significativo de una serie de tiempo generada por un proceso estocástico o determinista, lineal o no lineal.

### 2.3. Sistemas Dinámicos Estocásticos

Son sistemas dinámicos sometidos a los efectos del ruido. Presentan una variación de variables de estado aleatorias,  $X(t)$ , que evolucionan en función de otra variable, generalmente el tiempo,  $t$ . Esto significa que el estado subsecuente del sistema en el tiempo  $t + 1$  se determina tanto por las

acciones predecibles del sistema, como por elementos aleatorios. La mayoría de los sistemas de la vida real son estocásticos. Su comportamiento puede ser medido y aproximado con distribuciones y probabilidades, pero rara vez pueden ser determinados por un solo valor, (por ende son no deterministas). Por ejemplo, el tiempo que un cajero de un banco requiere para procesar el depósito de un cliente depende de varios factores, algunos de ellos pueden ser controlados, otros no; algunos son medibles, otros no, pero al final, realizando un conjunto de observaciones del tiempo de procesamiento de cada depósito del cajero, puede permitir ajustar los tiempos a una distribución y “predecir” cuál será el tiempo de proceso en un modelo de simulación para eventos discretos.

**Definición 2.6** Sea  $S$  un sistema dinámico discreto  $\vec{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  dado por

$$\vec{X}_{t+1} = \vec{F}(\vec{X}_t; \beta). \quad (2.24)$$

donde  $X_t$ ;  $t \in T$  es un proceso estocástico y  $T$  es un conjunto contable de índices del proceso, entonces se dice que (2.24) es un sistema dinámico estocástico discreto.

**Definición 2.7** Un proceso estocástico  $\{X_t, t \in T\}$  es una colección de variables aleatorias definidas sobre un espacio de probabilidad. Esto es,  $\{X_t : \Omega \rightarrow \mathbb{R}, t \in T\}$ ,  $\omega \rightarrow X_t(\omega) = X(\omega, t)$ , para cada,  $t \in T$ ,  $X_t(\omega)$  es una variable aleatoria.

El índice  $t$  se interpreta frecuentemente como tiempo, y por tanto, a  $X_t(\omega)$  se le refiere como el estado del proceso en el tiempo  $t$ . Por ejemplo  $X_t$  puede ser el precio de una acción en el tiempo  $t$ , o la cantidad total de ventas que se han registrado en el mercado en el tiempo  $t$ , o el número de clientes que van a depositar dinero en un banco al tiempo  $t$ .

El espacio de estados  $S$  de un proceso estocástico se define como el conjunto de todos los posibles valores que la variable aleatoria  $X_t$  puede tomar,  $S = \{X_t(\omega) | t \in T; \omega \in \Omega\}$ . Por tanto un proceso estocástico es una familia de variables aleatorias que describen la evolución a través del tiempo de algún proceso. Un ejemplo particularmente importante lo proporcionan las series de tiempo que registran observaciones de un proceso. En este caso el conjunto es ordenado y corresponde

a los instantes temporales (días, meses, años, etc). Para cada instante temporal está definida una variable aleatoria  $x_t$ , y los valores observados de las variables aleatorias en distintos instantes forman una serie temporal. Es decir, una serie de  $T$  datos,  $(x_1, x_2, \dots, x_T)$ , es una muestra de tamaño uno del vector de  $T$  variables aleatorias ordenadas en el tiempo correspondientes a los momentos  $t = 1, 2, \dots, T$ , por lo tanto la serie observada se considera una realización o trayectoria del proceso estocástico.

El proceso se caracteriza al definir la distribución de probabilidad conjunta de las variables aleatorias  $(x_1, x_2, \dots, x_T)$ , para cualquier valor de  $T$ . Estas distribuciones se denominan las *distribuciones finito-dimensionales* del proceso. Se dice que se conoce la estructura probabilística de un proceso cuando se conocen estas distribuciones, que determinan la distribución de cualquier subconjunto de variables y, en particular, las distribuciones marginales de cada variable.

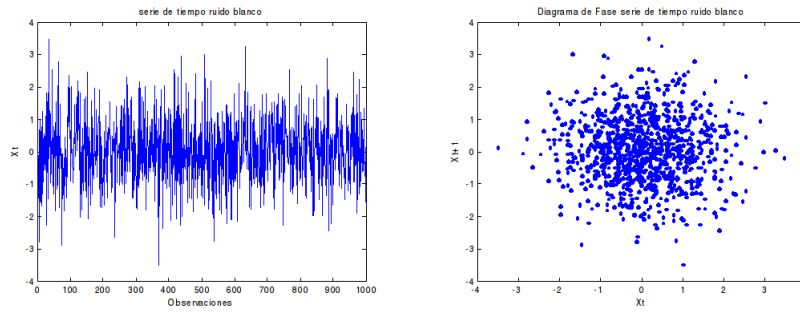
### 2.3.1. Distribución conjunta y marginal del proceso

En el análisis de series de tiempo generadas por procesos estocásticos se establece el supuesto de que las observaciones no provienen de variables aleatorias independientes. Se supone que existe una estructura de correlación entre las observaciones, por lo que no es fácil obtener la función de distribución conjunta del proceso. Para determinarla es necesario observar un gran número de realizaciones, para estimar la probabilidad de los distintos intervalos. Esta actividad se simplifica cuando se puede suponer que la distribución conjunta es normal multivariante, ya que quedaría determinada por el vector de medias y la matriz de varianzas y covarianzas entre variables.

**Ejemplo 2.2** *Considérese el proceso estocástico definido por*

$$X_t = e_t; \quad e_t \sim N(0, 1). \quad (2.25)$$

*Que se supone comienza en un instante  $t = 0$ , y donde las variables  $e_t$  son normales con media cero y varianza  $\sigma^2$ . Este proceso se conoce con el nombre de ruido blanco. En la figura 2.4(a) se presenta la evolución de un ruido blanco gaussiano de tamaño  $T = 1000$ , y en la figura 2.4(b) el*



(a) Evolución de la Serie

(b) Gráfica plana de dispersión

Figura 2.4: Gráficas de la serie de tiempo de ruido blanco.

diagrama de dispersión de dicho proceso, en donde se puede observar que la distribución de puntos  $(x_{t-1}, x_t)$  indica que sus valores en dos tiempos diferentes,  $t - 1$  y  $t$  no están correlacionados, por lo que al igual que en la gráfica (a) continua siendo aleatorio.

**Ejemplo 2.3** Considérese el proceso estocástico definido por

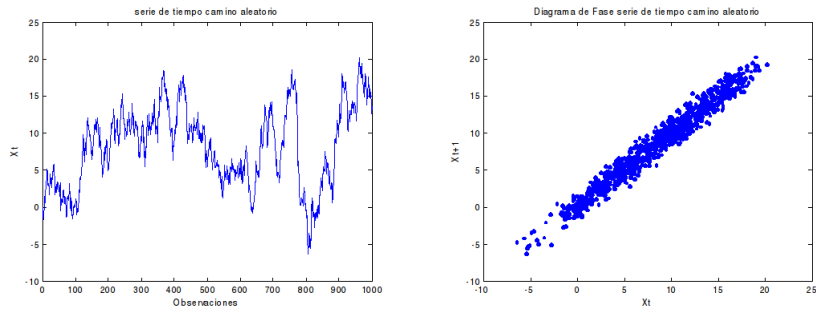
$$X_t = X_{t-1} + e_t; \quad e_t \sim N(0, 1). \quad (2.26)$$

Que se supone comienza en un instante  $t = 0$ ;  $X_0 = 0.0$ , y donde las variables  $e_t$  son normales con media cero y varianza  $\sigma^2$ . Este proceso se conoce con el nombre de caminata aleatoria y representa una variable cuyos cambios,  $e_t$ , son ruido blanco y por tanto imprevisibles. En este caso  $X_t$  depende más de su predecesor inmediato  $X_{t-1}$  y menos de sus predecesores antiguos.

La figura 2.5(a) presenta la evolución de la serie de tamaño  $T = 1000$ , en donde se puede apreciar la característica fundamental de este proceso que es la falta de afinidad de la serie a una media estable, empieza con valor cero y después va alejándose de ese valor a lo largo del tiempo. La figura 2.5(b) muestra el diagrama de fase (diagrama de dispersión) de dicho proceso.

**Definición 2.8** Se llama función de medias del proceso a una función del tiempo que proporciona el valor esperado de las distribuciones marginales  $x_t$  para cada  $t$ .

$$E[x_t] = \mu_t. \quad (2.27)$$



(a) Evolución de la Serie

(b) Gráfica plana de dispersión

Figura 2.5: Gráficas de la serie de tiempo camino aleatorio.

Cuando todas las variables tienen la misma media, entonces la función de medición es una constante,  $\mu$ , por lo que las realizaciones del proceso no muestran ninguna tendencia y se dice que el proceso es estable en la media. En caso contrario, las medias cambian con el tiempo, entonces las observaciones en distintos momentos muestran dicho cambio.

**Definición 2.9** Se llama función de varianzas del proceso a una función del tiempo que proporciona las varianzas en cada  $t$

$$Var[x_t] = \sigma_t^2. \quad (2.28)$$

Se dice que el proceso es estable en la varianza si esta es constante en el tiempo,  $\sigma_t^2 = \sigma^2; \forall t$ .

Un proceso puede ser estable en la media y no en la varianza y viceversa. Por ejemplo la caminata aleatoria, (2.26), tiene media constante, pero la varianza no es constante en el tiempo.

La estructura de dependencia lineal entre las variables aleatorias del proceso se representa por las funciones de covarianza y correlación.

**Definición 2.10** Se llama función de autocovarianza del proceso a la función de dos argumentos que describe las covarianzas entre dos variables del proceso en dos instantes cualesquiera

$$\gamma(t, t + j) = Cov(x_t, x_{t+j}) = E[(x_t - \mu_t)(x_{t+j} - \mu_{t+j})]. \quad (2.29)$$

En particular se tiene que

$$\gamma(t, t) = Var(x_t) = \sigma_t^2. \quad (2.30)$$

La función de medias y la de autocovarianzas cumplen en un proceso estocástico el mismo papel que la media y la varianza para una variable escalar. Las autocovarianzas tienen dimensiones, las del cuadrado de la serie, por lo que no son convenientes para comparar series medidas en unidades distintas. Se puede obtener una medida adimensional de la dependencia lineal generalizada, el coeficiente de correlación lineal entre dos variables.

**Definición 2.11** Se llama *coeficiente de autocorrelación de orden*  $(t, t + j)$  al *coeficiente de correlación entre las variables*  $x_t, x_{t+j}$ , y *función de autocorrelación* a la *función de dos argumentos que describe estos coeficientes para dos valores cualesquiera de las variables*;

$$\rho(t, t + j) = \frac{Cov(x_t, x_{t+j})}{\sigma_t \sigma_{t+j}} = \frac{\gamma(t, t + j)}{\sqrt{\gamma(t, t) \gamma(t + j, t + j)}}. \quad (2.31)$$

### 2.3.2. Funciones de distribución condicionadas

Además de las distribuciones marginales es de interés conocer lo que son las distribuciones condicionadas. Dentro de los procesos estocásticos uno de los más importantes es el denominado proceso de Markov. Intuitivamente un proceso es de Markov si conocido el valor actual del proceso, la distribución del valor futuro sólo depende del último valor, y no del camino recorrido hasta llegar a él. Por ello, se dice que un proceso de Markov tiene la propiedad de que

$$f(x_{t+1} | x_t, x_{t-1}, \dots, x_1) = f(x_{t+1} | x_t); t = 1, 2, \dots, \quad (2.32)$$

esto es, la distribución de la variable aleatoria en cualquier instante  $t$ , dados los valores previos del proceso solo depende del último valor observado. Un ejemplo de este tipo de proceso es la caminata aleatoria dada por (2.26),

$$f(x_{t+1} | x_t, x_{t-1}, \dots, x_1) = f(x_{t+1} | x_t) \sim N(x_t, \sigma^2). \quad (2.33)$$

Esto significa que la distribución de una variable generada por un paseo aleatorio en el instante  $t + 1$  condicionada a su valor anterior,  $x_t$ , es siempre normal con media del último valor observado y varianza constante, que no depende del instante considerado.

Una propiedad más débil que la markoviana es que la esperanza condicionada solo dependa del último valor observado, en particular cuando se cumple que

$$E(x_{t+1}|x_t, x_{t-1}, \dots, x_1) = E(x_{t+1}|x_t) = x_t. \quad (2.34)$$

al proceso que cumple con la propiedad anterior se le denomina *martingala*. La caminata aleatoria es una martingala.

Se pueden resumir las diferencias entre las distribuciones condicionadas y las marginales como sigue:

- La distribución marginal de  $x_t$  representa lo que se puede conocer de esta variable sin saber nada de su trayectoria hasta el instante  $t$ .
- La distribución condicionada de  $x_t$  dada  $x_{t-1}, \dots, x_{t-k}$  representa lo que se puede decir de esta variable cuando se conocen los  $k$  valores anteriores del proceso.

En el paseo aleatorio (2.26) la media marginal es constante e igual a cero, mientras que la media condicionada a los valores anteriores, es igual al último valor observado y la varianza marginal crece con el tiempo, mientras que la varianza condicionada es constante.

Intuitivamente si se inicia una caminata aleatoria en  $t = 0$  y se pregunta por los valores futuros del proceso para un valor de  $t$  fijo, es decir, por su distribución marginal en ese instante, se puede decir que la distribución será normal, con media (promedio sobre todas las posibles realizaciones) igual a cero y varianza  $t\sigma^2$ . Esto significa que la probabilidad de que  $x_t$  tome valores alejados de su media crece con el tiempo, como puede verse en la figura 2.5. Sin embargo, si se calcula la distribución condicionada de  $x_t$  a un valor fijo de  $x_{t-1}$ , es decir, si se toman sólo aquellas realizaciones en el instante  $t - 1$ , tomando el valor  $x_{t-1}$  fijado, la distribución del valor siguiente  $x_t$  será normal con media  $x_{t-1}$  y varianza constante  $\sigma^2$ .

En las series de tiempo las distribuciones condicionadas van a tener mayor interés que las marginales, porque definen las predicciones que se pueden hacer del futuro, conocido el pasado y la incertidumbre de estas predicciones.

### 2.3.3. Procesos estacionarios

La obtención de las distribuciones de probabilidades del proceso no es una tarea fácil cuando se tienen variables económicas o sociales, porque en este caso solo se puede observar una realización del proceso, ya que no siempre es posible volver atrás en el tiempo para generar otras realizaciones. Es decir el proceso estocástico existe conceptualmente, pero no se pueden obtener muestras sucesivas o realizaciones independientes del mismo. Para poder estimar las características *transversales* del proceso, medias, varianzas, etc., a partir de su evolución *longitudinal*, es necesario suponer que las propiedades *transversales* (distribución de las variables en cada instante) son estables a lo largo del tiempo.

**Definición 2.12** *Un proceso estocástico o serie temporal, es estacionario en sentido estricto si:*

1. *Las distribuciones marginales de todas las variables son idénticas.*
2. *Las distribuciones finito-dimensionales de cualquier conjunto de variables sólo dependen de los estados entre ellas.*

La primera condición establece que tanto la media como la varianza de todas las variables son las mismas. También lo son los coeficientes de asimetría y curtosis de las distribuciones marginales, ya que estas distribuciones son las mismas para todos los retardos. La segunda condición impone que la dependencia entre las variables sólo dependa de sus retardos, es decir, la misma dependencia existe entre las variables  $x_t, x_{t+j}, x_{t+j+h}$ , que entre las variables  $x_{t+k}, x_{t+k+j}, x_{t+k+j+h}$ . Estas dos condiciones se pueden resumir estableciendo que la distribución conjunta de cualquier conjunto de variables no se modifica si se trasladan las variables en el tiempo, esto es:

$$F(x_i, x_j, \dots, x_k) = F(x_{i+h}, x_{j+h}, \dots, x_{k+h}). \quad (2.35)$$

Este tipo de estacionariedad es una condición muy fuerte, ya que para contrastarla es necesario tener las distribuciones conjuntas para cualquier selección de variables del proceso. Una propiedad más débil, pero más fácil de contrastar en la práctica, es la *estacionariedad en sentido débil*, que implica la estabilidad en la media, la varianza y la estructura de covarianzas a lo largo del tiempo.



**Definición 2.13** *Un proceso es estacionario en sentido débil si, para toda  $t$  :*

$$\begin{aligned}\mu_t &= \mu \\ \sigma_t^2 &= \sigma^2 \\ \gamma(t, t-k) &= E[(x_t - \mu)(x_{t-k} - \mu)] = \gamma_k; \quad k = 0, \pm 1, \pm 2, \dots ,\end{aligned}\tag{2.36}$$

Las dos primeras condiciones indican que la media y la varianza son constantes. La tercera, que la covarianza entre dos variables depende sólo de su separación. En un proceso estacionario las autocovarianzas y autocorrelaciones sólo dependen del retardo entre las observaciones y, en particular, la relación entre  $x_t$  y  $x_{t-k}$ , es siempre igual a la relación entre  $x_t$  y  $x_{t+k}$ . Por lo tanto, en los procesos estacionarios:

$$Cov(x_t, x_{t+k}) = Cov(x_{t+j}, x_{t+j+k}) = \gamma_k; \quad j = 0, \pm 1, \pm 2, \dots ,\tag{2.37}$$

Para las autocorrelaciones:

$$\rho_k = \frac{Cov(x_t, x_{t-k})}{\sqrt{var(x_t)var(x_{t-k})}} = \frac{\gamma_k}{\gamma_0},\tag{2.38}$$

en los procesos estacionarios

$$\gamma_0 = \sigma^2; \quad y \quad \gamma_k = \gamma_{-k}, \quad \rho_k = \rho_{-k}.\tag{2.39}$$

### 2.3.4. Proceso de ruido blanco

Es un proceso estacionario importante que representa una variable que oscila en torno a su media cero, con varianza constante y cuyo pasado no contiene información útil para predecir valores futuros, ya que las variables del proceso están incorrelacionadas para todos los retardos. Por esta razón se dice que el proceso *no tiene memoria*. La gráfica 2.4 es un ejemplo de ruido blanco.

**Definición 2.14** *Un proceso estacionario es de ruido blanco si*

$$\begin{aligned} E[x_t] &= 0, & t = 1, 2, \dots \\ \text{Var}(x_t) &= \sigma^2, & t = 1, 2, \dots \\ \text{Cov}(x_t, x_{t-k}) &= 0 & k = \pm 1, \pm 2, \dots, \end{aligned} \tag{2.40}$$

un proceso de ruido blanco no es necesariamente estacionario en sentido estricto, ni debe estar formado por variables independientes, ya que solamente se exige la incorrelación. Si adicionalmente se le impone la condición de que las variables del proceso sean independientes y no sólo incorrelacionadas, entonces se tendría un proceso de *ruido blanco estricto*. Si se supone que las variables tienen distribución normal, la incorrelación garantiza la independencia, y la normalidad la igualdad de las distribuciones marginales, por lo que el proceso será de ruido blanco estricto con variables normales, y se le denomina *proceso de ruido blanco normal o de ruido blanco gaussiano*, (2.25) del ejemplo 1 es un proceso de este tipo.

### 2.3.5. Proceso caminata aleatoria

Es un proceso autorregresivo (AR) definido por:

$$X_t = c + aX_{t-1} + e_t; \quad e_t \sim N(0, 1), \tag{2.41}$$

si  $|a| < 1$  el proceso es estacionario. Si  $|a| > 1$  se obtiene un proceso explosivo, donde los valores de la variable crecen sin límite hacia el infinito. Esto puede verse en la figura 2.6. El caso interesante es cuando  $|a| = 1$ , ya que el proceso no es estacionario pero tampoco es explosivo, la figura 2.5 presenta una realización de un proceso con  $|a| = 1$

### 2.3.6. Modelos de procesos estocásticos no lineales

En la práctica se encuentran series de tiempo generadas por modelos que representan procesos estocásticos no lineales. Estos modelos representan la relación entre el valor presente de la serie

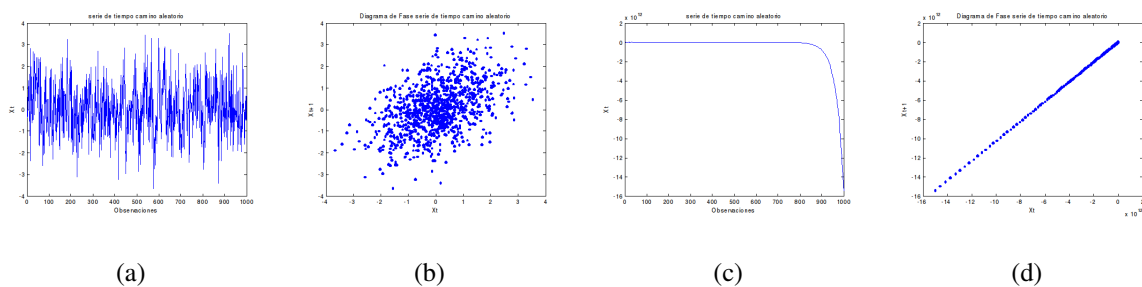


Figura 2.6: Gráficas de la serie de tiempo camino aleatorio. (a) Evolución de la serie estacionaria, (b) Gráfica plana de dispersión  $(X_{t-1}, X_t)$  serie estacionaria, (c) Evolución de la serie no estacionaria, (d) Gráfica plana de dispersión  $(X_{t-1}, X_t)$  serie no estacionaria

y su pasado con una función no lineal y las predicciones que minimizan el error cuadrático de predicción son funciones no lineales del pasado.

Una de las clases más importante de modelos que generan series de tiempo no lineales y, que están directamente motivadas por los sistemas dinámicos discretos deterministas no lineales, es la clase de autorregresión no lineal (NAR). En concreto, se dice que  $\{X_t\}$  es generada por un modelo autorregresivo no lineal de orden  $k$  con ruido, si existe una función  $f : \mathbb{R}^{k+1} \rightarrow \mathbb{R}$  no lineal, tal que

$$X_t = f(X_{t-1}, X_{t-2}, \dots, X_{t-k}, e_t); \quad t \in \mathbb{Z}, \quad (2.42)$$

donde  $e_t$  es una secuencia de variables aleatorias i.i.d. Normalmente,  $e_t$  es independiente de  $X_s$ ,  $s < t$ . Por el momento no se supone la existencia de los momentos de  $e_t$ . Claramente, en ausencia de  $e_t$  (ejemplo haciendo  $e_t = 0 \quad \forall t$ ) la ecuación (2.42) es justamente una ecuación en diferencias no lineal de orden  $k$ . En ocasiones resulta útil pensar en (2.42) como la vestidura estocástica de una estructura determinista ( $e_t = 0$ ), porque muchos modelos estocásticos se obtienen por perturbaciones aleatorias, de una manera u otra, de los modelos deterministas.

Algunas veces es más conveniente reescribir la ecuación (2.42) en su forma vectorial, introduciendo primero un vector de estado  $k$ -dimensional

$$\xi_t = (X_t, X_{t-1}, \dots, X_{t-k+1})', \quad (2.43)$$

y un vector de ruido  $k$ -dimensional

$$\eta_t = (e_t, 0, 0, \dots, 0)'. \quad (2.44)$$

Por lo que (2.42) puede reescribirse como

$$\xi_t = \phi(\xi_{t-1}, \eta_t), \quad (2.45)$$

donde  $\phi : \mathbb{R}^{2k} \rightarrow \mathbb{R}$  es definida por

$$\begin{aligned} \xi_t^{(1)} &= f\left(\xi_{t-1}^{(1)}, \dots, \xi_{t-1}^{(k)}, \eta_t^{(1)}\right) \\ \xi_t^{(2)} &= \xi_{t-1}^{(1)} \\ &\vdots \\ \xi_t^{(k)} &= \xi_{t-1}^{(k-1)}, \end{aligned} \quad (2.46)$$

normalmente  $\xi_t^{(j)}$  denota la  $j$ -ésima componente/coordenada de  $\xi$ .

Se puede obtener una primera clasificación de esta clase de modelos autorregresivos no lineales, suponiendo que se puede expresar la dependencia en función de un número finito de retardos, y escribir el modelo ARMA no lineal (NARMA) de orden  $(p, q)$  como:

$$X_t = f(X_{t-1}, \dots, X_{t-p}; e_{t-1}, \dots, e_{t-q}) + e_t, \quad (2.47)$$

donde  $f$  es una función no lineal arbitraria y la innovación  $e_t$  es un proceso de ruido blanco (2.25), independiente de  $(X_{t-k})$  para cualquier  $k > 0$ . Si  $q = 0$  se tiene un modelo AR no lineal (NAR) de orden  $p$ , y si  $p = 0$  se tiene un modelo MA no lineal (NMA) de orden  $q$ .

Suponiendo que  $f$  es una función suave con derivadas continuas, se puede aproximar mediante una serie, despreciando los términos de orden superior:

$$X_t = \alpha + \sum_{i=1}^p a_i X_{t-i} + \sum_{j=1}^r \sum_{k=1}^s b_{jk} X_{t-j} e_{t-k} + \sum_{i=1}^q c_i e_{t-i} + e_t. \quad (2.48)$$

Con lo que se obtiene un modelo de procesos bilineales. La introducción de procesos bilineales en el análisis de series temporales en tiempo discreto se debe a [Granger and Anderson,1978]. Se

representan como  $BL(p, q, m, h)$ , donde  $p, q$  se refieren a la parte lineal y  $m, h$  a los ordenes de los productos bilineales.

Analizando las propiedades de algunas formas bilineales sencillas caracterizadas como:

$$X_t = bX_{t-i}e_{i-j} + e_t. \quad (2.49)$$

Si  $i > j$  el modelo recibe el nombre de superdiagonal, si  $i = j$  diagonal, y si  $i < j$ , se llamará sub-diagonal.

En el caso determinista, un modelo bilinial es de la forma simple:

$$X_t = aX_{t-1} + bX_{t-1}u_{t-1}, \quad (2.50)$$

donde  $u_{t-1}$  es una entrada de control. La diferencia entre (2.50) modelo estocástico y (2.48) modelo determinista, bilineales ambos, radica en que la entrada de control de (2.49) es aleatoria e inobservable, por lo que se requiere una reformulación cuidadosa en términos probabilísticos antes de ser usados en uno u otro caso.

---

---

## CAPÍTULO 3

---

# LA INTEGRAL DE CORRELACIÓN COMO IDENTIFICADOR DEL RETARDO SIGNIFICATIVO DE UNA SERIE DE TIEMPO

En este capítulo, se presenta una propiedad intrínseca de las series de tiempo escalares que puede ser identificada y medida, la denominada Integral de Correlación,  $C_m^\tau(\varepsilon)$ , cuya estimación,  $\hat{C}_{m,n}^\tau(\varepsilon)$ , se propone en este trabajo como una herramienta estadística que permite identificar el retardo temporal óptimo (retardo significativo del modelo que proporciona una mayor información del comportamiento subyacente) de una serie de tiempo observada.

La estimación de la integral de correlación de [G-P, 1983],  $\hat{C}_{m,n}^\tau(\varepsilon)$ , es una medida de la frecuencia con la que se repiten patrones en los datos. Está basada en el hecho de que la dinámica caótica presenta regularidad en frecuencias altamente irregulares. Este hecho proviene de dos propiedades de la dinámica caótica que son: a) existe un gran número de puntos periódicos inestables en el espacio fase; b) las trayectorias caóticas son densas en el espacio fase. De esta forma, una trayectoria caótica puede pasar arbitrariamente cerca a una órbita periódica dada una infinidad de veces,

y trazar su patrón antes de ir a alguna otra parte del espacio fase, por lo que no debería causar ninguna sorpresa que cuando la integral de correlación captura un gran número de patrones que se repiten en un conjunto de datos, se rechaza la hipótesis de que los datos fueron generados por un proceso estocástico independiente e idénticamente distribuido, iid, ya que en este caso existiría una gran dependencia entre los datos.

De igual manera, y dadas las características de esta medida de frecuencia de repetición de patrones en los datos, se puede suponer que puede ser utilizada para identificar el retardo temporal significativo que proporciona mayor información del comportamiento subyacente de una serie de tiempo observada, al verificar que la estimación de la integral de correlación calculada con dicho retardo, es la que captura un mayor número de patrones que se repiten en los datos, en comparación con las estimaciones calculadas con retardos temporales distintos al significativo.

En general, cuando se estudia algún fenómeno financiero, con fines de reproducción o predicción de su comportamiento, lo que se observa en realidad no es el espacio fase del sistema dinámico que lo representa, que en la mayoría de los casos es desconocido, sino una secuencia de mediciones escalares, es decir, una serie de tiempo. Por esta razón, para llevar a cabo un análisis para especificar correctamente el retardo temporal que debe usarse en el modelo que reproduzca el comportamiento dinámico subyacente en esa secuencia de mediciones escalares, lo primero que debe hacerse es convertir estas mediciones en vectores de estado, para así poder calcular las integrales de correlación requeridas en la identificación del retardo significativo,  $\tau_o$ .

Con el fin de poder establecer posteriormente la metodología a seguir en la identificación del retardo significativo, en la sección 3.1 de este capítulo, se retoma el procedimiento para estimar la integral de correlación presentado en la sección 2.2.2 del Capítulo 2, y se presentan algunos elementos importantes de la formación de los vectores retardados ( $m$ -historias) requeridos para calcular el estimador,  $\hat{C}_{m,n}^{\tau}(\varepsilon)$  de la integral de correlación.

En la sección 3.2 se introduce a  $\hat{C}_{m,n}^{\tau}(\varepsilon)$  como un estadístico  $U$ , libre de parámetros de ruido (nuisance), que mide la probabilidad de que cualquier pareja particular en la serie de tiempo escalar

observada se encuentre dentro de una vecindad de radio  $\varepsilon$  dada. En la sección 3.3 se presentan las bases teóricas que permiten proponer a  $\hat{C}_{m,n}^\tau(\varepsilon)$ , como la medición que permite establecer un buen criterio para la selección del retardo temporal significativo,  $\tau_o$ , de una serie de tiempo estacionaria dada, es decir, el tiempo retardado de la variable, de la cual su valor actual,  $x_t$ , está dependiendo en mayor grado, y que le aportaría mayor información en relación a la dinámica que se quiere conocer. En la sección 3.4, se demuestra que este estimador también puede ser usado para identificar el retardo temporal significativo,  $\tau_o$ , de la serie de residuales al cuadrado, obtenida al filtrar con un proceso ARMA(p,q), una serie generada con un modelo ARMA-GARCH dado.

### 3.1. Generación de Vectores Retardados

Sea  $\{x_t\}_{t=1}^T$  una serie de tiempo de tamaño  $T$ , la *reconstrucción de vectores retardados* consiste en la formación de un vector de dimensión  $m$  para cada observación  $x_i$  de la forma

$$X_i^{m,\tau} = (x_{i-(m-1)\tau}, x_{i-(m-2)\tau}, \dots, x_{i-\tau}, x_i), \quad (3.1)$$

donde  $\tau$  es el retardo temporal seleccionado considerando la posible correlación que existe entre el valor actual de  $x$  y sus valores pasados. Cuando  $\tau > 1$ , la longitud de ventana de la inmersión se incrementa, esto es, se incrementa la distancia de desfase entre el primer y último elemento del vector de inmersión. El número total de vectores, (3.1) que se construyen es  $n = T - (m - 1)\tau$ , por lo que se requiere que el tamaño de la serie de tiempo observada sea  $T > 2(m - 1)\tau$ . En el ejemplo 2.1 presentado en el Capítulo 2, para el caso cuando  $m = 3$ ,  $\tau = 2$  y suponiendo que se tiene una serie de tiempo escalar de tamaño  $T = 50$ , el número de vectores retardados de la forma



3.1 que se pueden generar es  $n = 50 - 4 = 46$ , los que quedarían expresados como:

$$\begin{aligned}
 X_5^{3,2} &= \{x_1, x_3, x_5\} \\
 X_6^{3,2} &= \{x_2, x_4, x_6\} \\
 X_7^{3,2} &= \{x_3, x_5, x_7\} \\
 &\vdots \\
 X_{50}^{3,2} &= \{x_{46}, x_{48}, x_{50}\}
 \end{aligned}$$

Algunos teoremas relacionados con la inmersión mencionan que bajo ciertas circunstancias generales, el objeto geométrico generado por los vectores  $X_i^{m,\tau}$  es equivalente a la trayectoria original  $\{x_i\}$ , de hecho, ocurre que el atractor formado por  $X_i^{m,\tau}$  es equivalente al atractor en el espacio desconocido en el cual el sistema original se encuentra, si la dimensión  $m$  del espacio de coordenadas es suficientemente grande, ver Teorema 2.1 (Inmersión de Takens) en el Capítulo 2.

Como puede verse, cuando se inicia el análisis de una serie de tiempo escalar, es necesario, además de que el tamaño de la serie no sea muy pequeño, conocer un conjunto de parámetros que garanticen llevar a cabo de forma adecuada la construcción de los vectores de inmersión. Estos parámetros son: la dimensión de inmersión,  $m$  y el retardo temporal  $\tau$ . La forma en que se procede a seleccionarlos depende mucho de la dinámica subyacente en los datos, y del tipo de análisis que se quiere hacer. Sin embargo, en la mayoría de los casos no se tiene ninguna idea de como seleccionar el retardo temporal  $\tau$ , y como se mencionó en el Capítulo 1, al no existir procedimientos que tengan la finalidad de calcularlo, lo que se supone en la mayoría de los casos es que  $\tau = 1$ .

Si se supone que la dinámica en el espacio fase está representada por un campo vectorial suave, entonces los “estados vecinos” deben estar sujetos al mismo tiempo de evolución. Esto significa que después de un intervalo corto de tiempo en el futuro, las dos trayectorias que emergen de estos estados vecinos deben seguir conservándose como vecinos, ya que si esto no ocurre, la evolución temporal futura va a ser muy diferente. Si adicionalmente se conociera la dimensión correcta de inmersión  $m$ , para algún conjunto de datos dados, ¿cuál sería el procedimiento adecuado para

obtener el retardo temporal significativo para modelar ese conjunto de datos?. En las siguientes secciones se discuten los fundamentos teóricos que permiten asegurar que el procedimiento adecuado sería calcular la estimación de la integral de correlación con diferentes retardos temporales, y seleccionar el retardo para el cual se obtuvo el máximo valor del estimador de la integral de correlación.

## 3.2. Estimador de la Integral de Correlación

Sea  $\{x_t\}_{t=1}^T$  una sucesión de observaciones escalares de longitud  $T$ . El estimador de la integral de correlación,  $\hat{C}_{m,n}^\tau(\varepsilon)$ , propuesto por [G-P, 1983], (ver sección 2.2.1 del Capítulo 2), es una medida de la correlación espacial de puntos dispersos o partículas en un espacio dimensional  $m$ . Esto es, dada una serie de tiempo  $\{x_t\}_{t=1}^T$ , sumergida en algún espacio de inmersión  $m$ , se pueden generar y graficar  $n = T - (m - 1)\tau$  puntos, de la forma  $X_i^{m,\tau} = (x_i, x_{i-\tau}, \dots, x_{i-(m-1)\tau}) \in \mathbb{R}^m$ ;  $i = 1 + (m - 1)\tau, \dots, T$ .

Si se selecciona uno de estos puntos,  $X_i^{m,\tau}$ , y se define a su alrededor una vecindad  $V_\varepsilon(X_i^{m,\tau})$  de radio  $\varepsilon$ , se puede obtener, mediante un conteo, el número de otros puntos que se encuentran dentro de esa vecindad. Adicionalmente si se definen otras vecindades correspondientes a los demás puntos  $X_j^{m,\tau}$ ,  $j = i + 1, \dots, T - 1$ , al número total de puntos normalizados, dentro de todas las vecindades, se le denomina, *el estimador de la integral de correlación*, cuya expresión matemática es:

$$\hat{C}_{m,n}^\tau(\varepsilon) = \frac{1}{\binom{n}{2}} \sum_{i=1}^{n-1} \sum_{j=i+1}^n H(\varepsilon - \|X_i^{m,\tau} - X_j^{m,\tau}\|). \quad (3.2)$$

En la figura 3.1, a modo de ejemplificar lo ya expuesto, se representan dos vecindades de radio  $\varepsilon$  para un número de puntos finitos correspondientes a una dimensión de inmersión  $m = 1$  (una recta) y  $m = 2$  (una superficie). Se puede ver en ambos casos, que cuando el radio se incrementa, entonces se incrementa también el número de puntos correlacionados espacialmente.

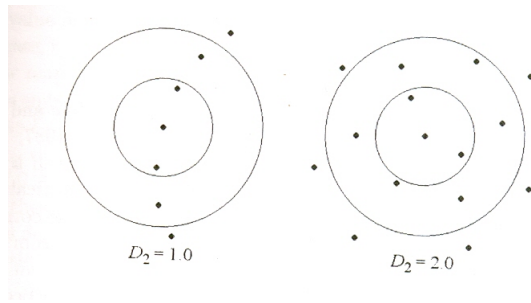


Figura 3.1: Puntos finitos  $X_{i,1}^T(\varepsilon)$ ;  $m = 1$  y  $X_{i,2}^T(\varepsilon)$ ;  $m = 2$ , cercanos en dos vecindades de radio  $\varepsilon$  distintas<sup>1</sup>.

Para los procesos generadores de datos de la forma:

$$x_t = G(\mathbf{X}_{t-1}, \theta) + e_t. \quad (3.3)$$

donde

$$\mathbf{X}_{t-1} = \{x_{t-1}, x_{t-2}, \dots\}$$

$G$  – una función medible de  $\mathbf{X}_{t-1}$  lineal o no lineal

$\theta$  – vector de parámetros

$\{e_t\}$  – es el residual, i.i.d,  $e_t$  independiente de  $\mathbf{X}_{t-1} \forall t$  con  $t = 2, \dots, T$

$T$  – número total de elementos de la serie

que cumplen los supuestos **A** a **D** presentados por [De Lima,1996], el estimador de la integral de correlación,  $\hat{C}_{m,n}^T(\varepsilon)$ , es un estadístico  $U$ , libre de parámetros de ruido (nuisance). Este tipo de estadísticos son de gran interés en la teoría formal de inferencia estadística porque:

- i. Tienen muchas de las propiedades deseables del promedio simple de tiempo  $U_t = \sum_{i=1}^T \frac{x_i}{T}$ , incluyendo teoremas del límite central y la ley de los grandes números.
- ii. En cierto contexto son estimadores de  $\theta$ , de varianza mínima dentro de la clase de estimadores insesgados de  $\theta$ , ([Serfling, 1980], p. 176).

<sup>1</sup>“Chaos and Time-Series Analysis”, Julián Clinton Sprott, Pág. 307, Oxford University Press

- iii. Convergen rápidamente a la normal ([Serfling, 1980], p.193, Teorema B).
- iv. Como es el caso de  $\hat{C}_{m,n}^\tau(\varepsilon)$ , muchos estadísticos de utilidad en la inferencia estadística, pueden escribirse en la forma de un estadístico  $U$ , ([Serfling, 1980], Capítulo 5).
- v. Geométricamente hablando,  $\hat{C}_{m,n}^\tau(\varepsilon)$  proporciona el volumen medio de un hipercubo de dimensión  $m$  con diámetro  $\varepsilon$ .

Considerando la sucesión de observaciones escalares  $\{x_t\}_{t=1}^T$  de longitud  $T$ , y analizando la expresión (3.2) se puede ver que el estimador permite medir la correlación espacial entre las  $n = T - (m - 1)\tau$ ,  $m$ -historias que se pueden generar a partir de esta serie. Si se parte del hecho de que se toman dos de las  $m$ -historias,  $X_i^{m,\tau}$ ,  $X_j^{m,\tau}$ , que pueden ser consideradas como dos conjuntos de datos consecutivos,

$$X_i^{m,\tau} = \{x_i, x_{i-\tau}, x_{i-2\tau}, \dots, x_{i-(m-1)\tau}\}; \quad (3.4a)$$

$$X_j^{m,\tau} = \{x_j, x_{j-\tau}, x_{j-2\tau}, \dots, x_{j-(m-1)\tau}\}; \quad j > i. \quad (3.4b)$$

ya que las componentes de (3.4a) y (3.4b) corresponden a las observaciones de la serie dada, si estas dos  $m$ -historias están cercanas, ambas son realizaciones de períodos en los cuales la variabilidad es relativamente pequeña comparada con la variabilidad media del proceso que generó la serie de tiempo. Esto debido a que los vectores retardados de períodos de baja variabilidad tienen una variación menor, y por lo tanto una mayor probabilidad de que sean  $\varepsilon$ -cercanos. Si se obtiene la distancia entre los elementos que ocupan la misma posición de ambos conjuntos

$$d_{ij}^k = |x_{i-k\tau} - x_{j-k\tau}|; \quad k = 0, \dots, m - 1; \quad j > i. \quad (3.5)$$

Se dice que las dos  $m$ -historias (3.4a) y (3.4b) están cercanas, si dado un  $\varepsilon > 0$

$$\|X_i^{m,\tau} - X_j^{m,\tau}\|_\infty = \max_{i \neq j, 0 \leq k \leq m-1} \{d_{ij}^k\} = d_{max_{ij}} \leq \varepsilon. \quad (3.6)$$

Si para cada una de las  $n$ ,  $m$ -historias se cuentan, cuántas de las restantes están próximas a la considerada, resumiendo el resultado de la comparación entre las dos  $m$ -historias de tamaño  $(m - 1)\tau$ ,

y, utilizando una función indicadora  $H$  (Heavside), se puede obtener el número total de  $m$ -historias que están próximas. La función  $H$  en si permite llevar a cabo la comparación (3.6), ya que toma el valor de uno si es verdadera y cero en otro caso. Por lo tanto de acuerdo a (3.2), todas las parejas de las  $m$ -historias son comparadas y aquellas dentro de  $\varepsilon$  una de la otra, son contadas para después normalizar ese número por el número de todas las posibles parejas.

**Notése** que, el estimador  $\hat{C}_{m,n}^\tau(\varepsilon)$ , en la ecuación (3.2), es un doble promedio de la función indicadora, que se espera que converja a  $C_m(\varphi)$  cuando  $T \rightarrow \infty$ . ([Denker and Keller,1983], teorema 1 y (3.9)) y [Brock and Dechert,1988a] demuestran que

$$\hat{C}_{m,n}^\tau(\varepsilon) \xrightarrow{d} C_m(\varepsilon), \quad (3.7)$$

además de que el estimador decrece monótonamente y tiende a cero cuando  $\varepsilon$  tiende a cero.

Sea  $A \in \mathbb{R}_{n \times n}$  la matriz de la función indicadora  $I(X_i^{m,\tau}, X_j^{m,\tau})$ , donde  $X_i^{m,\tau}$ ;  $i = 1 + (m-1)\tau, \dots, T-1$ ;  $X_j^{m,\tau}$ ;  $j = i+1, \dots, T$ , y tal que  $a_{ij}$  elemento  $ij$ -ésimo de la matriz  $A$  se define como:

$$a_{ij} = \begin{cases} 1 & \text{si } \|X_i^{m,\tau} - X_j^{m,\tau}\|_\infty \leq \varepsilon, \\ 0 & \text{en otro caso ;} \end{cases} \quad (3.8)$$

entonces esta matriz de unos y ceros está indicando si los vectores  $X_i^{m,\tau}$ ,  $X_j^{m,\tau}$  son cercanos en  $\mathbb{R}^m$ , en una vecindad de radio  $\varepsilon$ , por lo que la suma del número de unos de los renglones o columnas de la matriz  $A$ , denotada por  $N^{m,\tau}(\varepsilon)$ , permite obtener el número de vectores de retardo que están dentro de la vecindad dada, con lo cual la expresión (3.2), se puede reescribir como:

$$\hat{C}_{m,n}^\tau(\varepsilon) = \frac{2N^{m,\tau}(\varepsilon)}{n(n-1)} = \frac{N^{m,\tau}(\varepsilon)}{\binom{n}{2}}, \quad (3.9)$$

y en consecuencia la expresión anterior es la proporción del número de vectores de retardo que se encuentran dentro de la vecindad de radio  $\varepsilon$ , que implica por tanto que son cercanos, esto es, que son dependientes.

Por la ecuación (2.21) del Capítulo 2,

$$\hat{C}_{m,n}^\tau(\varepsilon) = \lim_{n \rightarrow \infty} \frac{2N^{m,\tau}(\varepsilon)}{n(n-1)}. \quad (3.10)$$

Por lo tanto,  $\hat{C}_{m,n}^\tau(\varepsilon)$ , que representa una medida de reconstrucción  $\mu_m$  asociada con los vectores de retardo de dimensión  $m$ , se define como *la probabilidad conjunta de todas las parejas de vectores de retardo que se encuentran a una distancia menor que  $\varepsilon$* , y por tanto se puede decir que mide la probabilidad de que cualquier pareja particular en la serie de tiempo escalar dada esté cerca.

$$\begin{aligned} \hat{C}_{m,n}^\tau(\varepsilon) = P[ & \|X_i^{m,\tau} - X_j^{m,\tau}\|_\infty \leq \varepsilon, \|X_i^{m,\tau} - X_{j+1}^{m,\tau}\|_\infty \leq \varepsilon, \\ & \dots, \|X_{i+1}^{m,\tau} - X_j^{m,\tau}\|_\infty \leq \varepsilon, \dots, \|X_{T-1}^{m,\tau} - X_T^{m,\tau}\|_\infty \leq \varepsilon], \end{aligned} \quad (3.11)$$

con  $i = 1 + (m-1)\tau, \dots, T-1$ ;  $j = i+1, \dots, T$ . Por (3.6)

$$\begin{aligned} \hat{C}_{m,n}^\tau(\varepsilon) = P[ & d_{max_{ij}} \leq \varepsilon, d_{max_{i(j+1)}} \leq \varepsilon, \dots, d_{max_{(i+1)j}} \leq \varepsilon, \\ & \dots, d_{max_{(T-1)T}} \leq \varepsilon], \end{aligned} \quad (3.12)$$

con  $i = 1 + (m-1)\tau, \dots, T-1$ ,  $j = i+1, \dots, T$

Se puede ver que (3.12) representa la probabilidad de que si los puntos de la serie de tiempo  $(x_i, x_j)$  están cercanos, también lo están  $(x_{i-\tau}, x_{j-\tau})$ ,  $(x_{i-2\tau}, x_{j-2\tau})$ ,  $\dots$ , y  $(x_{i-(m-1)\tau}, x_{j-(m-1)\tau})$ .

### 3.3. Identificación del Retardo Temporal Óptimo en Series de Tiempo Estacionarias

Dada una serie de  $t$  datos ordenados cronológicamente y correspondientes a una sola variable  $X_t = \{x_1, x_2, \dots, x_t\}$ , esta serie de puntos corresponde a una de las variables,  $j$ -ésimas, de un sistema  $m$ -dimensional subyacente. El comportamiento de los puntos del conjunto  $X_t$  es el resultado de la interacción con las restantes  $(m-1)$  variables, por lo que de alguna forma, la información contenida en  $X_t$  debe retener las propiedades del sistema  $m$ -dimensional.

De acuerdo al teorema de reconstrucción de Takens, cuando se construye el nuevo conjunto de vectores de retardo de dimensión  $m$ ,  $X_t^{m,\tau}$ , a partir de  $X_t$ , la elección tanto de  $\tau$  como de  $m$  es teóricamente arbitraria si los datos se encuentran libres de ruido y fueron medidos con precisión infinita, en cuyo caso únicamente se considera que la dimensión  $m$  sea lo “suficientemente grande” y cualquier retardo temporal. Sin embargo, estas consideraciones en la realidad no son factibles y se pueden presentar algunos problemas como:

- Series de Tiempo con datos que presentan ruido.
- Costo computacional caro e inestabilidad en el manejo de dimensiones de inmersión muy grandes. Para evitar la inestabilidad se requiere de series de tiempo de tamaño grande.
- Contaminación de ruido en las series de tiempo debido a la precisión finita de los cálculos en la computadora.

Por tal motivo, la calidad de la reconstrucción si depende de la elección de una mínima dimensión de inmersión  $m$ , y de un tiempo de retardo,  $\tau$ , óptimo, ya que la nueva serie de puntos  $m$ -dimensionales se obtiene, como ya se vió en la sección anterior, tomando como coordenadas de  $X_t^{m,\tau}$ , los valores  $x_t$  de la serie dada y los correspondientes  $\tau$ -retardos (desplazamientos). Es decir, que el rol que el retardo temporal óptimo tiene en la reconstrucción es de suma importancia, ya que si  $\tau$  fuera muy diferente al óptimo, se podrían presentar casos de redundancia, cuando  $x_t$  y  $x_{t-k\tau}$ ,  $k = 1, \dots, (m - 1)$ , están muy cercanos y el espacio de estado reconstruido se colapsa en una diagonal principal, o casos de irrelevancia, cuando  $x_t$  y  $x_{t-k\tau}$  están muy lejanos, por lo que no están relacionados, y cada vector reconstruido consiste de componentes irrelevantes.

En las figuras 3.2 a 3.6 se presentan las gráficas de los vectores retardados obtenidos a partir de series de tiempo generadas por tres modelos lineales estacionarios: un modelo completamente aleatorio,  $N(0,1)$ , un  $MA(3)$  y un  $AR(2)$ , así como por un modelo no lineal completamente determinista (modelo logístico) y uno autorregresivo no lineal estacionario,  $NAR(1)$ .

La construcción de los vectores retardados que se presentan en las gráficas se llevó a cabo con-

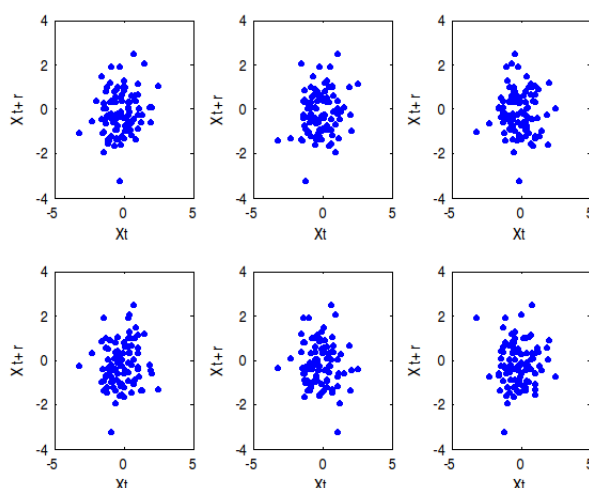


Figura 3.2: Vectores retardados obtenidos de la serie de tiempo generada con el modelo,

$X_t = e_t \sim N(0, 1)$ ,  $m = 2$ ,  $\tau = 1, 2, 3, 4, 5, 6$ . No existe retardo del modelo.

siderando una dimensión de inmersión  $m = 2$ , y seis diferentes valores de retardos temporales, entre los cuales se encuentra el retardo significativo al modelo generador de la serie de tiempo, que puede ser considerado como el retardo óptimo. El tamaño de las series generadas es de  $T = 100$ .

Puede verse en todas las figuras presentadas, a excepción de la figura 3.2, que la estructura que presenta la gráfica correspondiente a los vectores retardados generados con el retardo temporal significativo del modelo, es completamente diferente a las otras que corresponden a los vectores retardados generados con retardos diferentes al “óptimo”. En el caso de los vectores retardados obtenidos a partir del modelo determinista no lineal, figura 3.5, es claro que mientras que la gráfica de los vectores retardados obtenidos con el retardo  $\tau_o = 1$ , corresponde a la gráfica del atractor del modelo logístico, las estructuras de las otras gráficas con retardos temporales diferentes, muestran que al incrementarse el retardo, las coordenadas del vector reconstruido son irrelevantes.

Como lo mostraron [Matilla-García et al,2004], la importancia de la identificación del retardo temporal óptimo, radica en el hecho de que tanto en una serie de tiempo con datos observados, como en una generada por un proceso de la forma (3.3), existe un retardo temporal significativo que aporta una mayor información para reproducir la estructura dinámica correcta del proceso subyacente.



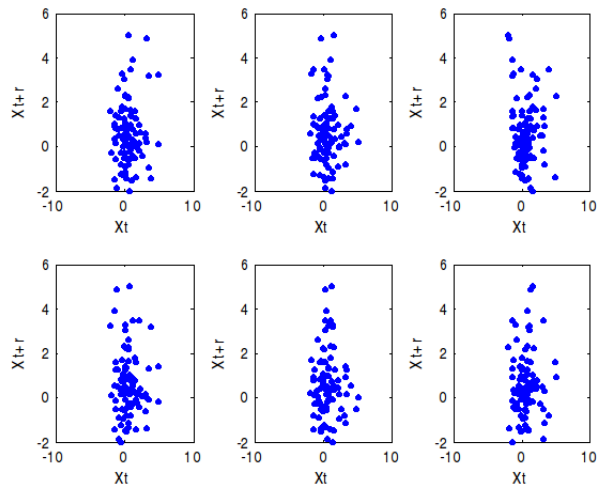


Figura 3.3: Vectores retardados obtenidos de la serie de tiempo generada con el modelo,  
 $X_t = e_t + 0.8e_{t-3}^2$ ;  $e_t \sim N(0, 1)$ ,  $m = 2$ ,  $\tau = 1, 2, 3, 4, 5, 6$ ,  $\tau_o = 3$

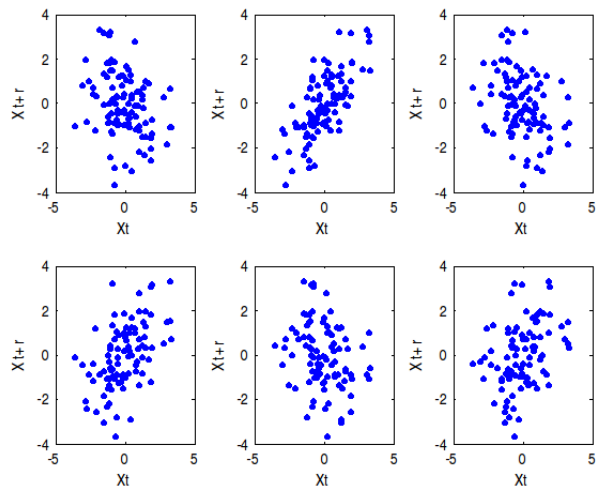


Figura 3.4: Vectores retardados obtenidos de la serie de tiempo generada con el modelo,  
 $X_t = e_t + 0.8X_{t-2}$ ;  $e_t \sim N(0, 1)$ ,  $m = 2$ ,  $\tau = 1, 2, 3, 4, 5, 6$ ,  $\tau_o = 2$ .

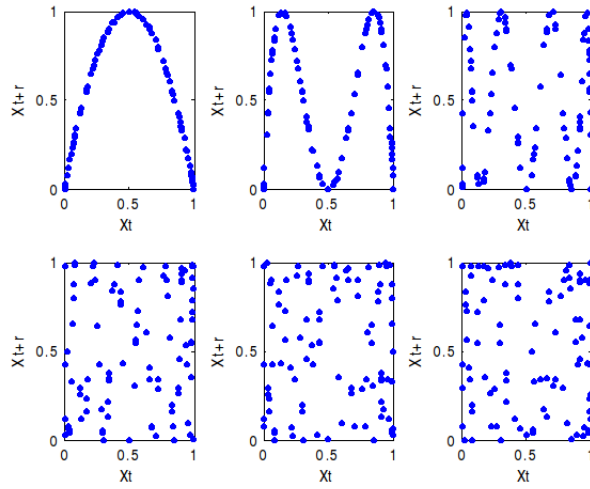


Figura 3.5: Vectores retardados obtenidos de la serie de tiempo generada con el modelo,  $X_t = 4X_{t-1}(1 - X_{t-1})$ ,  $m = 2$ ,  $\tau = 1, 2, 3, 4, 5, 6$ ,  $\tau_o = 1$ .

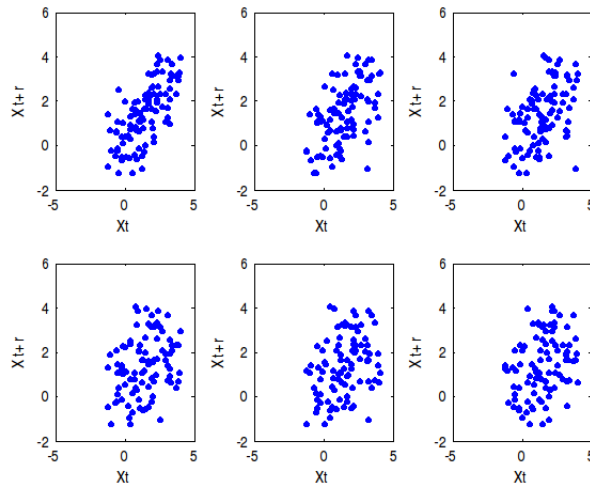


Figura 3.6: Vectores retardados obtenidos de la serie de tiempo generada con el modelo,  $X_t = e_t + |X_{t-1}|^{0.8}$ ;  $e_t \sim N(0, 1)$ ,  $m = 2$ ,  $\tau = 1, 2, 3, 4, 5, 6$ ,  $\tau_o = 1$ .

Considerando que la integral de correlación es una propiedad de la distribución espacial del conjunto de variables de estado, y que su estimación, por lo visto en la sección anterior, proporciona la fracción del número total de parejas de vectores de retardo  $(X_i^{m,\tau}, X_j^{m,\tau})$ , dentro de una vecindad de radio  $\varepsilon$  de él, esto es,  $V_\varepsilon(X_i^{m,\tau})$ , se propone utilizar este estimador para establecer una técnica que permita identificar el retardo óptimo de una serie de tiempo dada.

El procedimiento, en términos generales, de la aplicación de esta técnica consiste en: 1) estimar la integral de correlación, con los parámetros de dimensión de inmersión  $m$  y el radio  $\varepsilon$  fijos, haciendo variar el valor del parámetro  $\tau$  (por ejemplo:  $\tau = 1, 2, \dots, k; k \in \mathbb{Z}^+$ ); 2) identificar la  $\tau$ , para la cual se alcanzó el máximo valor de la estimación de la integral de correlación, esto es,  $\text{máx}\{\hat{C}_{m,n}^\tau(\varepsilon) | \tau \in \mathbb{N}\}$ .

En esta sección se demuestra que el estimador de la integral de correlación es una medición confiable que permite identificar el retardo temporal significativo de una serie de tiempo generada por un proceso de la forma (3.3).

Analizando las expresiones matemáticas (3.9) y (3.11) se puede establecer la hipótesis de que debería obtenerse un mayor número de parejas  $(X_i^{m,\tau}, X_j^{m,\tau})$ , tal que cada componente correspondiente de los vectores de retardo estén separadas entre sí en menos de un  $\varepsilon = 2\sigma$ , (donde  $\sigma$  es la desviación estándar de los datos de la serie de tiempo observada), al estimar la integral de correlación con el retardo temporal más significativo. En otras palabras, la probabilidad de que la distancia entre una pareja de vectores de retardo temporal óptimo seleccionados de forma aleatoria sea menor que  $\varepsilon$ , va a ser mayor que la probabilidad que se obtenga al considerar la distancia entre una pareja de vectores con un retardo temporal diferente al óptimo. Esto es, dada una serie de tiempo estacionaria de observaciones finitas  $\{X_t\}_{t=1}^T$ , una dimensión de inmersión  $m$  y un  $\varepsilon > 0$  fijos, al variar el tiempo de retardo  $\tau$ , dentro del conjunto  $1, 2, 3, \dots, r$ , en el cual se encuentra el retardo temporal significativo,  $\tau_o$ , por construcción de los vectores de retardo correspondientes a la inmersión, las  $m$ -historias, que se pueden generar con las observaciones de la serie de tiempo van

a ser diferentes, aún cuando ambas tengan en común el valor observado de la serie  $x_i$ :

$$X_i^{m,\tau_0} = (x_{i-(m-1)\tau_0}, \dots, x_{i-\tau_0}, x_i) \neq (x_{i-(m-1)\tau}, \dots, x_{i-\tau}, x_i) = \tilde{X}_i^{m,\tau}. \quad (3.13)$$

**Ejemplo 3.1** Dada una serie de tiempo  $\{x_t\} = x_1, x_2, \dots, x_T$  y la dimensión de inmersión  $m = 3$ , suponiendo que el retardo significativo del modelo es  $\tau_0 = 2$  y tomando  $\tau = 3$  retardo distinto al óptimo, se tendrían los siguientes conjuntos de vectores desfasados para  $\tau_0$  y  $\tau \neq \tau_0$  dados:

$$X_i^{3,2} = (x_{i-4}, x_{i-2}, x_i); \tau_0 = 2, i = 5, 6, \dots, T.$$

$$\tilde{X}_i^{3,3} = (x_{i-6}, x_{i-3}, x_i); \tau = 3, i = 7, 8, \dots, T.$$

Generando con el modelo  $x_t = 0.8x_{t-2} + e_t$ ;  $e_t \sim N(0, 1)$  una serie de tiempo de tamaño  $T = 10$  (tamaño muy pequeño con el fin de ejemplificar lo que se ha venido mencionando) se tendrían los siguientes resultados:

$$x_t = 0.9817, -2.291, 0.4315, -2.656, -1.232, -1.617, -0.7035, -1.260, -1.896, 0.1193.$$

como  $\sigma_x = 1.203$  se toma una  $\varepsilon = 1.96\sigma_x = 2.359$ . El conjunto de los vectores desfasados considerando la  $\tau_0 = 2$  está formado por  $n = 10 - (3 - 1)2 = 6$  vectores 3-dimensionales:

$$X_i^{3,2} = \begin{pmatrix} 0.9817 & 0.4315 & -1.232 \\ -2.291 & -2.656 & -1.617 \\ 0.4315 & -1.232 & -0.7035 \\ -2.656 & -1.617 & -1.260 \\ -1.232 & -0.7035 & -1.896 \\ -1.617 & -1.260 & 0.1193 \end{pmatrix}$$

para  $i = 5, \dots, 10$ . La matriz  $A$  (ver expresión 3.8), correspondiente al resultado de la comparación de las normas  $\|X_i^{3,2} - X_j^{3,2}\|_\infty$  contra la  $\varepsilon$  considerada es:

$$A = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ & 0 & 1 & 1 & 1 \\ & & 0 & 1 & 1 \\ & & & 1 & 1 \\ & & & & 1 \end{pmatrix}$$

por lo tanto  $N^{3,2}(\varepsilon) = 10$ , y

$$\hat{C}_{m,n}^{\tau}(\varepsilon) = \frac{2N^{3,2}(\varepsilon)}{6(5)} = \frac{20}{30} = .6667.$$

Por lo que respecta al conjunto de los vectores desfasados considerando la  $\tau = 3$ , está formado por  $n = 10 - (3 - 1)3 = 4$  vectores 3-dimensionales:

$$\tilde{X}_i^{3,3} = \begin{pmatrix} 0.9817 & -2.656 & -0.7035 \\ -2.291 & -1.232 & -1.260 \\ 0.4315 & -1.617 & -1.896 \\ -2.656 & -0.7035 & 0.1193 \end{pmatrix}$$

Y la matriz  $\tilde{A}$  que resulta de comparar las normas  $\|\tilde{X}_i^{3,3} - \tilde{X}_j^{3,3}\|_{\infty}$  contra la  $\varepsilon$  considerada es::

$$\tilde{A} = \begin{pmatrix} 0 & 1 & 0 \\ & 0 & 1 \\ & & 0 \end{pmatrix}$$

por lo tanto  $\tilde{N}^{3,3}(\varepsilon) = 2$ , y

$$\hat{C}_{m,n}^{\tau}(\varepsilon) = \frac{2\tilde{N}^{3,3}(\varepsilon)}{4(3)} = \frac{4}{12} = .3333.$$

Como se puede ver en el ejemplo, en general, si se toman dos parejas  $(i, j)$  en cada uno de los conjuntos  $X_{m,\tau_o}$ ,  $\tilde{X}_{m,\tau_o}$  que contienen las  $m$ -historias correspondientes, dado que los vectores generados con el retardo temporal significativo contienen una mayor información, las componentes que corresponden a valores de la serie van a generar trayectorias que se encuentran más cercanas entre sí, y por tanto

$$\|X_i^{m,\tau_o} - X_j^{m,\tau_o}\|_{\infty} < \|\tilde{X}_i^{m,\tau} - \tilde{X}_j^{m,\tau}\|_{\infty}, \quad (3.14)$$

por lo que la probabilidad marginal para cada pareja de puntos  $(i, j)$  de ambos conjuntos,  $X_{m,\tau_o}$ ,  $\tilde{X}_{m,\tau_o}$  tiene la siguiente relación:

$$P[\|X_i^{m,\tau_o} - X_j^{m,\tau_o}\|_{\infty} \leq \varepsilon] > P[\|\tilde{X}_i^{m,\tau} - \tilde{X}_j^{m,\tau}\|_{\infty} \leq \varepsilon], \quad (3.15)$$

relación que se conserva para la probabilidad conjunta (ver ecuación 3.11) de todas las parejas que cumplen con la condición de estar cercanas dentro de la distancia  $\varepsilon$  dada. Lo anterior se puede formalizar con el siguiente teorema.

**Teorema 3.1** *Sea  $\{x_t\}_{t=1}^T$  una serie de tiempo estacionaria generada por un modelo de la forma  $X_t = G(\mathbf{X}_{t-1}, \theta) + \mathbf{e}_t$ . De acuerdo con la definición (3.11), dados una dimensión de inmersión  $m \geq 2$ ;  $m \in \mathbb{N}$  y un radio  $\varepsilon > 0$ , si existe un retardo temporal óptimo, denotado por  $\tau_o$ , entonces*

$$\hat{C}_{m,n}^{\tau_o}(\varepsilon) = \text{máx}\{\hat{C}_{m,n}^{\tau}(\varepsilon) | \tau \in \mathbb{N}\}. \quad (3.16)$$

**Demostración:**

Sea  $\{X_t\}_{t=1}^T$  una serie de tiempo estacionaria, y sean  $X_{m,\tau_o}$ ,  $\tilde{X}_{m,\tau_o}$ , dos conjuntos de vectores de retardos de dimensión  $m$  (dimensión de inmersión) generados tomando a  $\tau_o$  como el retardo significativo del modelo generador de la serie, y  $\tau$  un retardo temporal distinto al significativo respectivamente. Por las ecuaciones (3.6) y (3.14) se tiene que:

$$d_{max_{ij}}^{\tau_o} < d_{max_{ij}}^{\tau}. \quad (3.17)$$

Si el radio de la vecindad de  $X_i^{m,\tau_o}$  y  $\tilde{X}_i^{m,\tau}$  es una  $\varepsilon$  pequeña,  $0 < \varepsilon < a\sigma$ , con  $a = 0.5, 1.0, 1.5, 2.0$  y  $\sigma$  la desviación estándar de la serie de tiempo, el mismo para ambos vectores, el número de parejas de vectores retardados que cumplen con el requerimiento de estar dentro del radio  $\varepsilon$  dado es mayor en el conjunto  $X_i^{m,\tau_o}$ , esto es,

$$N^{m,\tau_o}(\varepsilon) = \sum_{i=1+(m-1)\tau_o}^{n-1} \sum_{j=i+1}^n H(\varepsilon - d_{max_{ij}}^{\tau_o}) > \sum_{i=1+(m-1)\tau}^{n-1} \sum_{j=i+1}^n H(\varepsilon - d_{max_{ij}}^{\tau}) = N^{m,\tau}(\varepsilon). \quad (3.18)$$

Ya que el estimador de la integral de correlación se define como la fracción del número total de pares de puntos del conjunto de  $m$ -historias, cuya distancia es menor o igual a  $\varepsilon$ , ver ecuación (3.9), multiplicando (3.18) por  $\frac{2}{n(n-1)}$

$$\hat{C}_{m,n}^{\tau_o}(\varepsilon) > \hat{C}_{m,n}^{\tau}(\varepsilon) \quad \forall \tau \neq \tau_o. \quad (3.19)$$

Por otro lado, si el radio de la vecindad de  $X_i^{m,\tau_0}$  y  $\tilde{X}_i^{m,\tau}$  es una  $\varepsilon$  grande,  $0 \ll \varepsilon$ , todos los demás vectores tanto los generados con el retardo óptimo como los generados con un retardo diferente al óptimo, se encuentran dentro de la vecindad, ver Sección 3.2, figura 3.1, por lo que el número de parejas de vectores retardados que cumplen con el requerimiento de estar dentro del radio  $\varepsilon$  dado va a ser igual al número de parejas que se forman con los  $n$  vectores retardados,  $\binom{n}{2}$ , y por lo tanto

$$\hat{C}_{m,n}^{\tau_0}(\varepsilon) = \hat{C}_{m,n}^{\tau}(\varepsilon) \simeq 1.0 \quad \forall \tau \neq \tau_0 \quad (3.20)$$

Sea  $\hat{C}'_{m,n}(\varepsilon) = \max_{1 \leq \tau \leq s} \{\hat{C}_{m,n}^{\tau}(\varepsilon)\}$  por demostrar que  $\hat{C}'_{m,n}(\varepsilon) = \hat{C}_{m,n}^{\tau_0}(\varepsilon)$

Como

$$\hat{C}'_{m,n}(\varepsilon) = \max_{1 \leq \tau \leq s} \left\{ \frac{2N^{m,\tau}(\varepsilon)}{n(n-1)} \right\}, \quad (3.21)$$

por la ecuación (3.18)

$$\frac{2N^{m,\tau_0}(\varepsilon)}{n_{\tau_0}(n_{\tau_0}-1)} > \frac{2\tilde{N}^{m,\tau}(\varepsilon)}{n_{\tau}(n_{\tau}-1)} \quad \forall \tau \neq \tau_0, \quad (3.22)$$

entonces

$$\hat{C}'_{m,n}(\varepsilon) = \frac{2N^{m,\tau_0}(\varepsilon)}{n_{\tau_0}(n_{\tau_0}-1)} = \hat{C}_{m,n}^{\tau_0}(\varepsilon). \quad (3.23)$$

□

De acuerdo a lo anterior, la integral de correlación da su valor máximo en el intervalo de tiempo correcto, cuando se compara con retardos no óptimos. Por lo que la probabilidad de que cualquier par particular de los vectores de la serie de tiempo estén cerca, cuando se usa un tiempo de retardo diferente que el óptimo,  $\tau \neq \tau_0$ , implica que alguna información importante se está perdiendo, y entonces no será óptima para la detección de la estructura dinámica correcta de la serie de tiempo generada con por el proceso subyacente. Por consiguiente, este resultado es útil para seleccionar el intervalo de tiempo adecuado para el análisis de las propiedades del proceso, y es importante considerar las siguientes observaciones:

**Observación 1.** Nótese que el teorema 3.1 se cumple para toda  $\varepsilon$  y  $m$ . En ocasiones es difícil para el investigador elegir la distancia del radio  $\varepsilon$ . Algunos investigadores han realizado investigaciones en este sentido, como el clásico BDS ([Kanzler,1999] y [Kocenda and Briatka,2004]), pero no se

presenta un criterio claro de decisión. En nuestro caso, dicha selección no es tan importante, ya que sólo afecta al resultado para los casos extremos en (a) todas las parejas están dentro de  $\varepsilon$  o (b) el caso contrario. Desde un punto de vista práctico esta es una propiedad interesante ya que se reduce el riesgo de obtener conclusiones equivocadas debidas a una mala selección de la distancia del radio. Otras preocupaciones similares surgen cuando el investigador tiene que seleccionar el parámetro de inmersión  $m$ . No se dispone de una regla clara de decisión, y diferentes dimensiones de inmersión pueden conducir a conclusiones diferentes. Nuestro resultado tampoco se ve afectado por tal selección: la integral de correlación dará el valor máximo independientemente del parámetro  $m$  cuando se evalúa en el retardo correcto. Esta propiedad es potencialmente importante también para cuestiones prácticas, como la de permitir capturar el retardo correcto, aun cuando tal retardo no esté cercano en el tiempo.

**Observación 2.** La integral de correlación está estrechamente relacionada con la muy conocida medida de entropía. Existen otras técnicas no paramétricas para el estudio de la dependencia de los retardos en las que se han utilizado funcionales de entropía, que como bien se sabe, son medidas que están relacionadas con el uso de la integral de correlación, ya que hay una relación entre las cantidades teóricas de información y las integrales de correlación. La entropía de Renyi de orden 2 se define como

$$H = -\ln \int f_X^2(X) dX,$$

donde  $f_X(X)$  es la función de densidad de probabilidad y  $X$  un vector de variables aleatorias. Tomando logaritmos en la ecuación (2.18) del Capítulo 2, se obtiene

$$-\ln C_m^T(\varepsilon) = -\ln \left[ \iint H(\varepsilon - \|X_i^{m,\tau} - X_j^{m,\tau}\|_\infty) dF_m(X_i^{m,\tau}) dF_m(X_j^{m,\tau}) \right].$$

Existe una relación estrecha entre las dos últimas ecuaciones. La integral interna de la expresión (2.18),  $\left[ \int H(\varepsilon - \|X_i^{m,\tau} - X_j^{m,\tau}\|_\infty) dF_m(X_i^{m,\tau}) dF_m(X_j^{m,\tau}) \right]$  se comporta como  $\varepsilon^m f_X^2(X)$  para  $\varepsilon$  pequeñas (ver [Prichard and Theiler, 1995]), por lo tanto

$$H \simeq -\ln C_m^T(\varepsilon) + m \ln \varepsilon.$$

Por consiguiente, algunos procedimientos alternativos aprovechan el hecho de que la medida de



entropía va a ser mínima cuando se calcula utilizando el retardo correcto.

### 3.4. Identificación del Retardo Óptimo en los Residuales

Si  $\mathbf{x}_{t-1} = (x_{t-1}, \dots, x_1)$  es el vector de valores pasados de la serie  $\{x_t\}$ , una formulación de un modelo no lineal en la media está dada por:

$$x_t = f(\mathbf{x}_{t-1}) + e_t, \quad (3.24)$$

donde  $e_t$  es un proceso estacionario estricto de ruido blanco de media cero y varianza  $\sigma^2$ .

Según el teorema de Wold, si un proceso es estacionario, éste se puede aproximar mediante un modelo lineal con residuos sin autocorrelación serial, pero si el modelo generador es (3.24) los residuos no serían independientes.

Una serie de tiempo puede tener estructura no lineal no sólo en la media, sino también en la varianza. En series con datos financieros y climatológicos se observa con frecuencia que en las series estacionarias, la varianza de la serie en el instante  $t$  depende de la variabilidad de los datos anteriores, de manera que hay rachas de alta varianza, seguidas de datos con poca varianza. Entonces la expresión (3.24) se puede generalizar considerando este tipo de dependencias para obtener una formulación general para procesos no lineales, del tipo:

$$x_t = f(\mathbf{x}_{t-1}) + \mathbf{u}_t \mathbf{g}(\mathbf{x}_{t-1}), \quad (3.25)$$

donde  $u_t$  es un proceso de ruido blanco normal, iid  $N(0,1)$ , formado por variables independientes entre sí y de los valores  $\mathbf{x}_{t-1}$ , y las funciones  $f$  y  $g$  son no lineales en los valores pasados de la serie.

Los dos tipos de efectos no lineales, en la media y en la varianza pueden o no darse simultáneamente en una serie temporal. En esta sección para mostrar que la técnica de identificación del retardo óptimo es adecuada en el diagnóstico de series de residuales que surgen después de filtrar con un proceso ARMA(p,q) una serie dada, se van a considerar únicamente procesos generadores de

series que presentan efecto no lineal en la varianza. Como la varianza condicionada representa la incertidumbre de las predicciones, estos procesos van a reflejar un riesgo variable, entendiéndose este como la incertidumbre de las predicciones en el tiempo.

La identificación de los efectos no lineales en la varianza se efectúa normalmente después de ajustar un modelo ARMA a la serie de interés para eliminar la dependencia en la media. Si existen efectos no lineales, los residuos del modelo ARMA van a estar incorrelacionados pero no serán independientes y esto podría verse en la función de autocorrelación de los residuos al cuadrado, que son los que mostrarán la correlación serial.

Los residuos, también conocidos como los errores de un modelo de la forma  $y_t = f(\mathbf{Y}_{t-1}, \beta, \mathbf{e}_t)$ , donde  $\beta$  es un vector de parámetros estimados  $\sqrt{n}$ -consistentes, son un proceso normal, iid., con media cero y varianza la unidad, y se pueden interpretar como el último proceso estocástico proveniente del modelo. Es por esta razón que al momento de querer desarrollar un modelo de predicción adecuado es importante poder identificar correctamente este último proceso estocástico para establecer la dependencia real que existe entre  $e_t$  y sus valores pasados, igual que se hace con las series de tiempo  $x_t$  observadas. Asumiendo que la función  $f$  puede invertirse, entonces  $e_t = G(y_t, \mathbf{Y}_{t-1}, \beta)$ , por lo que la técnica de identificación del retardo significativo podría aplicarse a esta serie de residuales. Sin embargo, ésta no se conoce, la que se conoce es la estimación  $\hat{e}_t = \hat{G}(y_t, \mathbf{Y}_{t-1}, \hat{\beta})$ , por lo que en su lugar se utilizará la serie de residuales estimados  $\{\hat{e}_t\}$ .

Si el modelo que genera la serie  $\{y_t\}$  está correctamente especificado, la serie  $\{\hat{e}_t\}$  es una buena estimadora de  $\{e_t\}$ . En las siguientes subsecciones se hará evidente que los residuales al cuadrado  $\{\hat{e}_t^2\}$  son buenos estimadores de los errores al cuadrado,  $\{e_t^2\}$ , por lo que al estimar la integral de correlación en los residuales al cuadrado, se obtendría una buena aproximación de la integral de correlación de los errores al cuadrado originales.

### 3.4.1. Convergencia en probabilidad de la serie de los residuales al cuadrado

Sea el proceso generador de datos de la forma (3.3), el residual (error estimado) se define como

$$\hat{e}_t = y_t - \hat{y}_t, \quad (3.26)$$

donde  $\hat{y}_t = \hat{G}(\mathbf{Y}_{t-1}, \hat{\theta})$ , entonces

$$\hat{e}_t = G + e_t - \hat{G} = (G - \hat{G}) + e_t, \quad (3.27)$$

sea

$$\begin{aligned} d^* &= G - \hat{G} \\ \hat{e}_t &= d^* + e_t \\ &= O_p(1) + e_t. \end{aligned} \quad (3.28)$$

**Proposición 3.1** Si  $\hat{\theta}$  es un estimador  $\sqrt{n}$ -consistente de  $\theta$ , y  $d^* = d^*(\mathbf{Y}_{t-1}, \theta, \hat{\theta})$  una función medible, entonces

$$\lim_{n \rightarrow \infty} (G - \hat{G}) = 0, \quad (3.29)$$

y por tanto

$$\hat{e}_t \xrightarrow{p} e_t \quad \forall t. \quad (3.30)$$

Esta proposición implica que se puede usar  $\{\hat{e}_t\}$  en lugar de  $\{e_t\}$

**Demostración:**

$\hat{\theta}$  es un estimador consistente de  $\theta$ , entonces  $\hat{\theta} \xrightarrow{p} \theta$  y  $\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{p} 0$  [De Lima, 1996], por tanto por el teorema de Slutsky

$$\sqrt{n}(\hat{G}(\hat{\theta}) - G(\theta)) \xrightarrow{p} 0 \quad \text{y} \quad \lim_{n \rightarrow \infty} (G - \hat{G}) = 0. \quad (3.31)$$

□

**Proposición 3.2** Si  $\{\hat{e}_t\} \xrightarrow{p} \{e_t\}, \forall t$ , entonces

$$\{\hat{e}_t^2\} \xrightarrow{p} \{e_t^2\}, \forall t. \quad (3.32)$$

**Demostración:**

Como  $\{\hat{e}_t\} \xrightarrow{p} \{e_t\}$  por el teorema 2.7 de ([Van Der Vaart,2000], p10), entonces  $\{\hat{e}_t\} \xrightarrow{d} \{e_t\} \forall t$ .

Sea  $g : \mathbb{R} \rightarrow \mathbb{R}^+$  tal que  $g(e_t) = e_t^2$ , por el teorema 2.3 de ([Van Der Vaart,2000], p7),  $\{\hat{e}_t^2\} \xrightarrow{p} \{e_t^2\}, \forall t$ .

Considerando ahora

$$\begin{aligned} \hat{e}_t^2 &= [G - \hat{G} + e_t]^2 = (G - \hat{G})^2 + 2(G - \hat{G})e_t + e_t^2 \\ &= (d^*)^2 + 2d^*e_t + e_t^2 \\ &= O_p(2) + e_t^2. \end{aligned} \tag{3.33}$$

□ Si  $\{e_t^2\}$  es un proceso estocástico estrictamente estacionario de variables aleatorias, se denota a  $e_t^{2,m,\tau} = (e_{t-(m-1)\tau}^2, e_{t-(m-2)\tau}^2, \dots, e_{t-\tau}^2, e_t^2)$  como la  $m$ -historia de los errores al cuadrado.

**Proposición 3.3** Si  $\{\hat{e}_t^2\} \xrightarrow{p} \{e_t^2\}, \forall t$ , entonces

$$\{\hat{e}_t^{2,m,\tau}\} \xrightarrow{p} \{e_t^{2,m,\tau}\}, \forall t. \tag{3.34}$$

**Demostración:**

Sean

$$\hat{e}_t^{2,m,\tau} = (\hat{e}_{t-(m-1)\tau}^2, \hat{e}_{t-(m-2)\tau}^2, \dots, \hat{e}_{t-\tau}^2, e_t^2); \quad e_t^{2,m,\tau} = (e_{t-(m-1)\tau}^2, e_{t-(m-2)\tau}^2, \dots, e_{t-\tau}^2, e_t^2), \tag{3.35}$$

por proposición 3.2  $\hat{e}_{t-k\tau}^2 \xrightarrow{p} e_{t-k\tau}^2 \forall t, k; 0 \leq k \leq m-1$  por tanto

$$\{\hat{e}_t^{2,m,\tau}\} \xrightarrow{p} \{e_{t-k\tau}^2\}. \tag{3.36}$$

□

**3.4.2. Convergencia en probabilidad de la integral de correlación.**

La proposición 3.3 planteada en la subsección anterior permite considerar que se tienen buenos estimadores de los errores al cuadrado originales, que captan toda la información. Este resulta-

do facilita la aplicación de la técnica propuesta para identificar el retardo óptimo de la serie de residuales al cuadrado. Sin embargo, para llevar a cabo la validación completa de la técnica, es necesario demostrar que el estadístico  $U$  de la integral de correlación aplicado a los residuales  $\hat{C}_{m,n}^\tau(\hat{e}_t^{2,m,\tau}; \varepsilon)$ , converge en probabilidad al estadístico  $U$  de la integral de correlación aplicado a los errores originales,  $C_{m,n}^\tau(e_t^{2,m,\tau}; \varepsilon)$ .

**Teorema 3.2** Sean  $\{\hat{e}_t^{2,m,\tau}\}_{t=1+(m-1)\tau}^T$  los vectores residuales retardados al cuadrado de la serie de tiempo  $\{\hat{e}_t^2\}_{t=1}^T$ , que es iid., obtenidos del ajuste a la serie de tiempo observada  $\{y_t\}$ , generada con el modelo  $y_t = G(\mathbf{Y}_{t-1}, \theta) + \mathbf{e}_t$  y sean  $\{e_t^{2,m,\tau}\}_{t=1+(m-1)\tau}^T$  los vectores error retardados al cuadrado de la serie de tiempo  $\{e_t^2\}_{t=1}^T$  que es iid., tal que  $\{\hat{e}_t^{2,m,\tau}\} \xrightarrow{p} \{e_t^{2,m,\tau}\}$ , entonces

$$\hat{C}_{m,n}^\tau(\hat{e}_t^{2,m,\tau}; \varepsilon) \xrightarrow{p} C_{m,n}^\tau(e_t^{2,m,\tau}; \varepsilon). \quad (3.37)$$

### Demostración:

Considerando las proposiciones de la sección anterior y por resultados de [De Lima,1996],  $\hat{C}_{m,n}^\tau(\hat{e}_t^{2,m,\tau}, \varepsilon)$  es un estadístico  $U$  libre de parámetros de ruido, por lo que

$$\sqrt{n}(\hat{C}_{m,n}^\tau(\hat{e}_t^{2,m,\tau}; \varepsilon) - C_{m,n}^\tau(e_t^{2,m,\tau}; \varepsilon)) \xrightarrow{p} 0. \quad (3.38)$$

Sea

$$\mu_{C_{m,n}^\tau} = \lim_{n \rightarrow \infty} E[C_{m,n}^\tau(e_t^{2,m,\tau}; \varepsilon)], \quad (3.39)$$

$$\begin{aligned} \sqrt{n}(\hat{C}_{m,n}^\tau(\hat{e}_t^{2,m,\tau}; \varepsilon) - C_{m,n}^\tau(e_t^{2,m,\tau}; \varepsilon)) &= \sqrt{n}(\hat{C}_{m,n}^\tau(\hat{e}_t^{2,m,\tau}; \varepsilon) \\ &\quad - C_{m,n}^\tau(e_t^{2,m,\tau}; \varepsilon) + \mu_{C_{m,n}^\tau} - \mu_{C_{m,n}^\tau}) \\ &= \sqrt{n}(\hat{C}_{m,n}^\tau(\hat{e}_t^{2,m,\tau}; \varepsilon) - \mu_{C_{m,n}^\tau}) \\ &\quad - \sqrt{n}(C_{m,n}^\tau(e_t^{2,m,\tau}; \varepsilon) - \mu_{C_{m,n}^\tau}), \end{aligned} \quad (3.40)$$

dividiendo ambos lados de 3.40 entre  $\sqrt{n}(\hat{e}_t^{2,m,\tau} - e_t^{2,m,\tau})$ , se tiene

$$\frac{\sqrt{n}(\hat{C}_{m,n}^\tau(\hat{e}_t^{2,m,\tau}; \varepsilon) - C_{m,n}^\tau(e_t^{2,m,\tau}; \varepsilon))}{\sqrt{n}(\hat{e}_t^{2,m,\tau} - e_t^{2,m,\tau})} = \frac{\sqrt{n}(\hat{C}_{m,n}^\tau(\hat{e}_t^{2,m,\tau}; \varepsilon) - \mu_{C_{m,n}^\tau})}{\sqrt{n}(\hat{e}_t^{2,m,\tau} - e_t^{2,m,\tau})} - \frac{\sqrt{n}(C_{m,n}^\tau(e_t^{2,m,\tau}; \varepsilon) - \mu_{C_{m,n}^\tau})}{\sqrt{n}(\hat{e}_t^{2,m,\tau} - e_t^{2,m,\tau})}, \quad (3.41)$$

$$\begin{aligned} \frac{(\hat{C}_{m,n}^\tau(\hat{e}_t^{2,m,\tau}; \varepsilon) - C_{m,n}^\tau(e_t^{2,m,\tau}; \varepsilon))}{(\hat{e}_t^{2,m,\tau} - e_t^{2,m,\tau})} + \frac{\sqrt{n}(C_{m,n}^\tau(e_t^{2,m,\tau}; \varepsilon) - \mu_{C_{m,n}^\tau})}{\sqrt{n}(\hat{e}_t^{2,m,\tau} - e_t^{2,m,\tau})} \\ = \frac{\sqrt{n}(\hat{C}_{m,n}^\tau(\hat{e}_t^{2,m,\tau}; \varepsilon) - \mu_{C_{m,n}^\tau})}{\sqrt{n}(\hat{e}_t^{2,m,\tau} - e_t^{2,m,\tau})}, \end{aligned} \quad (3.42)$$

$$\begin{aligned} \sqrt{n}(\hat{e}_t^{2,m,\tau} - e_t^{2,m,\tau}) \left\{ \frac{\hat{C}_{m,n}^\tau(\hat{e}_t^{2,m,\tau}; \varepsilon) - C_{m,n}^\tau(e_t^{2,m,\tau}; \varepsilon)}{(\hat{e}_t^{2,m,\tau} - e_t^{2,m,\tau})} \right\} + \sqrt{n}(C_{m,n}^\tau(e_t^{2,m,\tau}; \varepsilon) - \mu_{C_{m,n}^\tau}) \\ = \sqrt{n}(\hat{C}_{m,n}^\tau(\hat{e}_t^{2,m,\tau}; \varepsilon) - \mu_{C_{m,n}^\tau}). \end{aligned} \quad (3.43)$$

Usando el teorema del valor medio

$$\left\{ \frac{\hat{C}_{m,n}^\tau(\hat{e}_t^{2,m,\tau}; \varepsilon) - C_{m,n}^\tau(e_t^{2,m,\tau}; \varepsilon)}{(\hat{e}_t^{2,m,\tau} - e_t^{2,m,\tau})} \right\} \approx \frac{\partial}{\partial e_t^2} \hat{C}_{m,n}^\tau \Big|_{e_t^2 = e^*}, \quad (3.44)$$

donde  $e^* \in [e_t^2, \hat{e}_t^2]$  por lo tanto

$$\begin{aligned} \sqrt{n}(\hat{e}_t^{2,m,\tau} - e_t^{2,m,\tau}) \frac{\partial}{\partial e_t^2} \hat{C}_{m,n}^\tau \Big|_{e_t^2 = e^*} + \sqrt{n}(C_{m,n}^\tau(e_t^{2,m,\tau}; \varepsilon) - \mu_{C_{m,n}^\tau}) \\ = \sqrt{n}(\hat{C}_{m,n}^\tau(\hat{e}_t^{2,m,\tau}; \varepsilon) - \mu_{C_{m,n}^\tau}), \end{aligned} \quad (3.45)$$

el teorema de Slutsky garantiza que  $\hat{C}_{m,n}^\tau(\hat{e}_t^{2,m,\tau}; \varepsilon)$  y  $C_{m,n}^\tau(\hat{e}_t^{2,m,\tau}; \varepsilon)$  son asintóticamente equivalentes si  $\sqrt{n}(\hat{e}_t^{2,m,\tau} - e_t^{2,m,\tau}) = O_p(2)$ , esto es que  $\hat{e}_t^{2,m,\tau}$  es un estimador  $\sqrt{n}$ -consistente de  $e_t^{2,m,\tau}$ , por ([Van Der Vaart,2000], p. 72) y que

$$\lim_{n \rightarrow \infty} E \left[ \frac{\partial}{\partial e_t^2} \hat{C}_{m,n}^\tau(e_t^{2,m,\tau}, \varepsilon) \Big|_{e_t^2 = e^*} \right] = 0, \quad (3.46)$$

se tiene que (3.45) queda como:

$$\sqrt{n}(C_{m,n}^\tau(e_t^{2,m,\tau}; \varepsilon) - \mu_{C_{m,n}^\tau}) = \sqrt{n}(\hat{C}_{m,n}^\tau(\hat{e}_t^{2,m,\tau}; \varepsilon) - \mu_{C_{m,n}^\tau}), \quad (3.47)$$

por tanto  $\hat{C}_{m,n}^\tau(\hat{e}_t^{2,m,\tau}; \varepsilon) \xrightarrow{p} C_{m,n}^\tau(e_t^{2,m,\tau}; \varepsilon)$ .  $\square$

### 3.4.3. Procesos generados por modelos ARMA-GARCH

Se va a suponer que una serie de ruido blanco  $\{e_t\}$ , es la serie de residuos de un modelo ARMA. Si las variables  $e_t$  además fueran normales, se podría garantizar la independencia y no podría existir información en el pasado de la serie que fuera útil para explicar su futuro. Existe una familia de modelos denominada GARCH, que permite relajar la hipótesis de normalidad y que permite que se tengan procesos de ruido blanco formados por variables dependientes.

Los modelos GARCH son una generalización de modelos de tipo  $e_t = \sigma_t z_t$ , donde  $z_t$  y  $\sigma_t$  son dos procesos estacionarios independientes entre si. El proceso  $z_t$  es de ruido blanco normal estandarizado, esto es, formado por variables normales independientes de media cero y varianza uno. El proceso  $\sigma_t$  es estacionario, pero tiene estructura dinámica, por lo que su valor en  $t$  está en función del conjunto  $\mathbf{e}_{t-1} = (e_{t-1}, \dots, e_1)$  de los valores de la serie previos a  $t$ . La condición de independencia entre  $z_t$  y  $\sigma_t$  garantiza que la serie  $e_t$  tenga media marginal igual a cero, ya que

$$E[e_t] = E[\sigma_t]E[z_t] = 0, \quad (3.48)$$

y la media condicional es también nula:

$$E[e_t/\mathbf{e}_{t-1}] = E[\sigma_t/\mathbf{e}_{t-1}]E[z_t] = 0. \quad (3.49)$$

como el proceso es estacionario, tiene una varianza marginal constante, denotada por  $\sigma^2$ .

Esta varianza se obtiene con

$$E[e_t^2] = E[\sigma_t^2]E[z_t^2] = E[\sigma_t^2] = \sigma^2, \quad (3.50)$$

ya que coincide con la varianza del proceso  $\sigma_t$  ya que  $E[z_t^2] = 1$ . Sin embargo, el proceso  $e_t$  tiene una varianza condicionada que no es constante

$$Var[e_t^2/\mathbf{e}_{t-1}] = E[\sigma_t^2/\mathbf{e}_{t-1}]E[z_t^2] = \sigma_t^2, \quad (3.51)$$

ya que  $E[z_t^2/\mathbf{e}_{t-1}] = E[z_t^2] = 1$ . Esto significa que el proceso  $\sigma_t^2$  puede ser interpretado como la representación de la varianza condicionada de la serie en cada instante, que va variando en el tiempo con cierta estructura estacionaria.

En la figura 3.7(a) se muestra la gráfica de la evolución de un proceso generado con un modelo AR(1)+GARCH(1,1), con 500 observaciones y en la 3.7(b) se muestra la evolución de los residuos al cuadrado, en donde se puede ver que aparecen rachas de mayor variabilidad seguidas por otras de menor variabilidad.

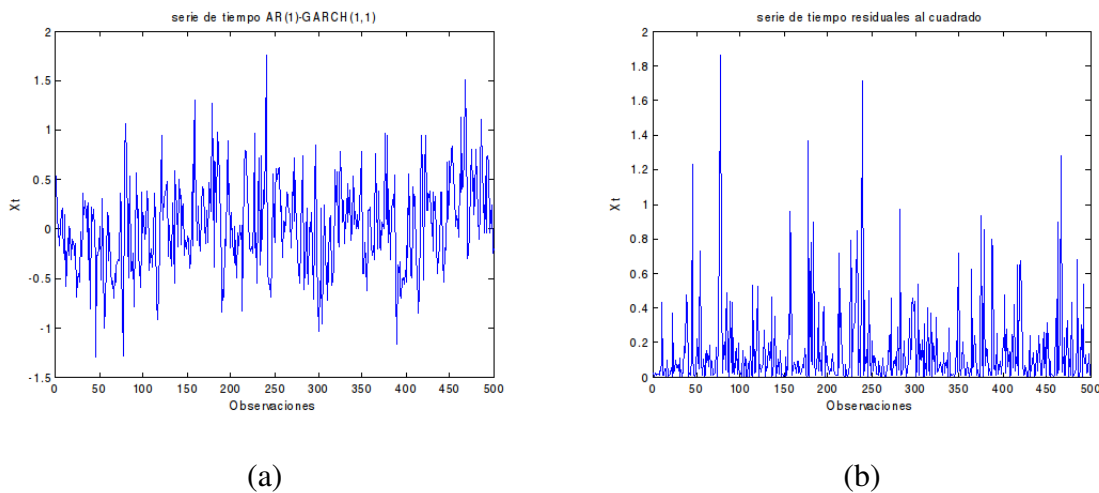


Figura 3.7: Gráficas de la evolución de (a) serie generada con un modelo AR(1)+GARCH(1,1), (b) serie de los residuales al cuadrado.

La independencia entre los procesos  $z_t$  y  $\sigma_t$  garantiza que la serie  $e_t$  carezca de autocorrelación y que forme un proceso de ruido blanco, pues  $E[e_t e_{t-k}] = E[z_t]E[\sigma_t \sigma_{t-k}] = 0$ , debido a que el proceso  $z_t$  es independiente de su pasado, y de los valores previos y actuales del proceso  $\sigma_t$ . Sin embargo a diferencia de las hipótesis de los modelos ARMA, la serie  $e_t$  no es de variables independientes, ya que  $e_t$  y  $e_{t-1}$  son función de valores comunes pasados de la serie, aún cuando estén incorrelacionados.



De acuerdo a lo señalado arriba un modelo GARCH(1,1) se expresa matemáticamente como:

$$e_t = \sigma_t z_t; \quad \sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \quad (3.52)$$

con  $\alpha_1, \beta_1 > 0$  y  $\alpha_1 + \beta_1 < 1$ , para que el proceso sea estacionario.

Por lo tanto un modelo ARMA(1,1)- GARCH(1,1) que genera una serie  $\{x_t\}$  se expresaría como

$$x_t = \alpha_0 + \alpha_1 x_{t-1} + e_t; e_t = \sigma_t z_t; \quad \sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \quad (3.53)$$

que se puede interpretar como la aplicación de dos ARMA sucesivos, el primero sobre el proceso mismo y el segundo sobre los residuos al cuadrado.

---

---

## CAPÍTULO 4

---

# METODOLOGIA PARA LA VALIDACIÓN EMPIRICA DE LA TÉCNICA PROPUESTA

La parte empírica de la validación de la técnica propuesta para identificar el retardo temporal óptimo de una serie de tiempo observada, se llevó a cabo mediante la aplicación de simulaciones con los métodos Monte Carlo y Bootstrap. El diseño de estas simulaciones se desarrolló en función del objetivo de verificar y reafirmar los resultados teóricos ya obtenidos y presentados en el Capítulo 3 de esta tesis. Por tal motivo en dicho diseño se consideraron las siguientes etapas:

- Selección de modelos de procesos generadores de series de tiempo, estacionarias y no estacionarias, lineales y no lineales, estocásticas y determinísticas (que presentan caos).
- Definición de los parámetros de inmersión requeridos para la formación de los vectores de retardos y el cálculo de la estimación de la integral de correlación.
- Definición de los tamaños de las series de tiempo a generar.
- Selección de las técnicas de simulación Monte Carlo y Bootstrap, y el número de repeticiones para ambas.

En este capítulo, en la sección 4.1 se proporciona una breve introducción de lo que son las técnicas de simulación Monte Carlo y Bootstrap, y se menciona la importancia de su uso en este trabajo. En la sección 4.2 se presentan los procedimientos de las simulaciones aplicadas y los parámetros que se definieron para la validación del buen funcionamiento de la técnica que se propone, al aplicarse a series de tiempo estacionarias y no estacionarias.

En la sección 4.3 se presenta el procedimiento de la simulación y los parámetros que se definieron para la validación del buen funcionamiento de la técnica aplicada a series de residuos estimados de un modelo ajustado a datos sin procesar. Por último en la sección 4.4 se proporciona una breve información sobre el software desarrollado para implantar los procedimientos presentados y para ejecutar las simulaciones Monte Carlo y el Bootstrap.

Los resultados de todas las simulaciones aplicadas con los procedimientos aquí descritos para validar la técnica, son presentados en el Capítulo 5, Capítulo 6 y Apéndice A (apéndice al Capítulo 5) de esta tesis.

## **4.1. Simulación Monte Carlo y Bootstrap**

En la estadística inferencial se utilizan técnicas para la estimación de parámetros de la población, cálculo de intervalos de confianza, pruebas de hipótesis y modelado, para medir la confiabilidad de los parámetros poblacionales a partir de valores obtenidos en una muestra aleatoria. La metodología tradicional que se aplica para obtener la medición y que se basa en el cálculo de la probabilidad de observar un valor del parámetro poblacional aproximado, suponiendo que la muestra fue extraída aleatoriamente de la población, requiere conocer la distribución de cualquier estadístico que se utilice en el análisis.

En situaciones en las que el estadístico en uso es bien conocido, como por ejemplo la media muestral, la aplicación de los métodos se hace analíticamente. Sin embargo, en ausencia de cualquier otra información respecto a la población que no sea la que está contenida en la propia muestra, la

distribución de los valores encontrados en una muestra aleatoria constituye la mejor orientación en cuanto a la distribución de esa población, y en este caso para llevar a cabo la inferencia es necesario utilizar métodos numéricos alternativos a los métodos analíticos clásicos.

Los métodos de remuestreo, también conocidos como métodos de cálculo intensivo [Efron and Tibshirani,1993], [Davison and Hinkley,1997], constituyen una familia de métodos que son muy utilizados para llevar a cabo simulación numérica. Esta familia incluye a los métodos de aleatorización Monte Carlo, Bootstrap y Jackknife.

En la validación estadística del buen funcionamiento del estimador de la integral de correlación, propuesto en este trabajo, en la identificación del retardo óptimo de una serie de tiempo observada, se utilizarán únicamente el método Monte Carlo y el Bootstrap, por lo que a continuación se señala brevemente en que consisten dichos métodos.

#### **4.1.1. Método Monte Carlo.**

De acuerdo a [Murdoch,2000], el término Monte Carlo originalmente se refería a cualquier tipo de simulación que involucrara caminatas aleatorias y fue primeramente usado por Jon von Neumann y S.M. Ulam en la década de los 40's del siglo XX. En la actualidad el método se refiere a cualquier tipo de simulación que involucra el uso de números aleatorios, y se emplea con éxito desde hace mucho en problemas de física, ingeniería y también en estudios de biología, ciencias naturales y economía, entre otros.

Para llevar a cabo una simulación con este método, se requiere un modelo que represente la población o el fenómeno de interés y una forma de generar números aleatorios, de acuerdo al modelo que se proponga, usando una computadora. Los datos que se generen con el modelo pueden ser estudiados como si fueran las observaciones, y se pueden utilizar estadísticos basados en ellos, como por ejemplo: medias, medianas, modas, varianzas, sesgo, etc, para conocer la población. En el caso de este trabajo, se estudia el estimador de la integral de correlación, para conocer la relación, y el grado de ésta, del valor de la serie observada con sus valores retardados.

La simulación Monte Carlo además puede ser usada para evaluar el funcionamiento de un modelo de inferencia o bien una prueba de hipótesis en cuanto a los términos del error Tipo I y el error Tipo II. El procedimiento de la simulación para estimar el error Tipo I, dado que éste ocurre cuando se rechaza la hipótesis nula de la prueba cuando es verdadera, sería el siguiente:

1. Determinar la pseudo-población que representa cuando la hipótesis nula es verdadera.
2. Generar una muestra aleatoria de tamaño  $n$  de esta pseudo-población.
3. Realizar la prueba de hipótesis usando el valor crítico
4. Determinar si se ha cometido un error de Tipo I, esto es ¿se rechazó la hipótesis nula?. Se sabe que la hipótesis nula no debería ser rechazada porque la muestra se está generando con el modelo o con la distribución que está de acuerdo con la hipótesis nula. El resultado de este ensayo se registra en la variable contador  $I$  como sigue:

$$I_i = \begin{cases} 1 & \text{Se cometió error Tipo I,} \\ 0 & \text{Caso contrario,} \end{cases} \quad (4.1)$$

5. Se repiten los pasos 2 a 4 para  $N_r$  ensayos.
6. La probabilidad de cometer el error Tipo I es

$$\hat{\alpha} = \frac{1}{N_r} \sum_{i=1}^{N_r} I_i. \quad (4.2)$$

**Nótese** que en el paso 6, calcular la ecuación (4.2) es lo mismo que calcular la proporción de veces que la hipótesis nula fue rechazada falsamente de un total de número de repeticiones,  $N_r$ . Esto proporciona una estimación del nivel de significancia de la prueba para un valor crítico,  $\alpha$ , dado.

El procedimiento de la simulación para estimar el error Tipo II, es similar, pero este error es determinado muestreando con el modelo o la distribución cuando la hipótesis nula es falsa.

1. Determinar la pseudo-población que representa cuando la hipótesis nula es falsa.

2. Generar una muestra aleatoria de tamaño  $n$  de esta pseudo-población.
3. Realizar la prueba de hipótesis usando el nivel de significancia  $\alpha$  y los correspondientes valores críticos.
4. Determinar si se ha cometido un error de Tipo II, esto es ¿no se rechazó la hipótesis nula ?.

El resultado de este ensayo se registra en la variable contador  $I$  como sigue:

$$I_i = \begin{cases} 1 & \text{Se cometió error Tipo II,} \\ 0 & \text{Caso contrario,} \end{cases} \quad (4.3)$$

5. Se repiten los pasos 2 a 4 para  $N_r$  ensayos.
6. La probabilidad de cometer el error Tipo II es:

$$\hat{\beta} = \frac{1}{N_r} \sum_{i=1}^{N_r} I_i. \quad (4.4)$$

por lo que con (4.4) se obtiene la proporción de veces que la hipótesis nula no fue rechazada, cuando debería serlo, de un total de  $N_r$ .

Un punto importante que debe tenerse en cuenta acerca de las simulaciones de Monte Carlo, es que el experimento sólo es aplicable a la situación que ha sido simulada, por ejemplo, tamaño de la muestra y valores críticos dados; otro sería el número de repeticiones que son necesarias para las simulaciones, este frecuentemente depende de los recursos computacionales que se tengan disponibles. Si el tiempo y el equipo computacional no son un problema, entonces se recomienda que  $N_r$  tome valores tan grandes como sea posible.

En [Hope, 1968] se muestra que los resultados de una simulación Monte Carlo son insesgados para cualquier  $N_r$ , bajo la suposición que la programación es correcta. En [Mooney,1997] se señala que no hay ninguna teoría general que establezca el número de ensayos necesarios en la simulación de Monte Carlo. Sin embargo, hace algunas recomendaciones, como usar en un principio un número pequeño de ensayos para asegurarse que el programa está trabajando correctamente, y una vez validado el código, la simulación debe llevarse a cabo con una  $N_r \geq 1000$ , y que no es común

considerar valores entre 10000 y 25000. Por supuesto que, algo que debe determinar el número de ensayos debe ser el propósito de la simulación.

#### 4.1.2. Método Bootstrap

Este método fue descrito de forma sistemática por [Efron,1979] y [Efron and Tibshirani,1993]. Consiste en generar, utilizando el muestreo con reemplazo, un gran número de muestras aleatorias de tamaño  $n$ ,  $x^* = (x_1^*, \dots, x_n^*)$ , a partir de una muestra de  $n$  valores observados,  $x = (x_1, \dots, x_n)$ , de acuerdo al modelo específico que se está contrastando.

Es como escribir en papeles pequeños los valores de los datos observados, meter éstos en una urna y extraer uno de ellos con uno de los valores anotados. Después de anotar el resultado obtenido se volvería a colocar el papel en la urna, y se repetiría la operación hasta registrar  $n$  valores. Como se está muestreando con reemplazamiento de la muestra original, existe la posibilidad de que algunos puntos  $x_i$  aparezcan más de una vez en  $x^*$  o tal vez no.

En muchas situaciones los analistas están interesados en estimar algún parámetro  $\theta$  calculando un estadístico de la muestra aleatoria, o bien en determinar el error estándar en la estimación  $\hat{\theta}$ , entonces los métodos *bootstrap* son muy prácticos para calcular las propiedades estadísticas de  $\theta$ , estimándolo a partir de una muestra, especialmente cuando el parámetro no sigue una distribución estadística concreta o cuando se obtiene mediante un cálculo complicado que dificulta la estimación analítica de su error. En la *muestra bootstrap* obtenida se calcula el valor del parámetro que se está estimando. Repitiendo el proceso un gran número  $B$  de veces (por ejemplo 100, 200,  $\dots$ , o más), se obtiene una distribución de valores para el parámetro, en la que se puede calcular su media, dispersión (análogo del error estándar), u otro tipo de estimador. También puede usarse para estimar intervalos de confianza y validar la contrastación de hipótesis.

El método bootstrap es un caso especial de las simulaciones Monte Carlo en donde no se hacen suposiciones paramétricas acerca de la población subyacente que genera la muestra aleatoria, sino que en su lugar, se usa la muestra como un estimador de la población. Esta estimación es llamada la

distribución empírica  $\hat{F}$  donde cada  $x_i$  tiene una probabilidad de  $\frac{1}{n}$ , esto es, cada  $x_i$  tiene la misma probabilidad de ser seleccionado en una nueva muestra de  $\hat{F}$ . Al no utilizar más que los valores observados en la muestra, a este método se le puede denominar como autosuficiente.

El siguiente ejemplo presenta dos posibles muestras  $x^*$ , cuando se muestrea  $x$  con reemplazamiento. Sea la muestra aleatoria formada por cuatro números  $x = (7, 3, 5, 8)$ .

$$x^{*1} = (x_4, x_3, x_3, x_1) = (8, 5, 5, 7), \quad (4.5a)$$

$$x^{*2} = (x_4, x_1, x_2, x_3) = (8, 7, 3, 5), \quad (4.5b)$$

(4.5a) corresponde a una muestra bootstrap, en la cual se repite uno de los valores observados, y la (4.5b) es una muestra bootstrap, en donde no se repite ninguno de los valores. Se puede usar la notación  $x^{*b} = (x_1^{*b}, \dots, x_n^{*b})$ ;  $b = 1, \dots, B$ , para señalar la  $b$ -ésima muestra bootstrap.

Resumiendo lo anterior, el método *bootstrap* se basa en la idea de que una muestra de una población conserva las propiedades estadísticas de la población y por lo tanto, remuestrear la muestra con reemplazamiento tiene el mismo efecto que remuestrear la población [Efron and Tibshirani, 1993].

## 4.2. Procedimientos para la Validación Empírica

. El comportamiento empírico de obtener el máximo valor del promedio estimado de la correlación integral para el retardo más relevante,  $\tau_o$ , de un modelo dado es estudiado a través de la simulación Monte Carlo y la simulación Monte Carlo con Bootstrap. Para ello se seleccionaron diez modelos generadores de series de tiempo estacionarias, ver Tabla 4.1, y tres modelos generadores de series de tiempo no estacionarias lineales, ver Tabla 4.2.

Adicionalmente se llevaron a cabo dos simulaciones más que permitieron verificar el buen comportamiento de la técnica: 1) para seleccionar retardos en forma automática; 2) para seleccionar retardos independientemente del epsilon,  $\varepsilon$ , que se asigne. Para verificar el primer punto se agregó el modelo estacionario 11, ver sección 4.2.3, que incluye retardos significativos de orden 6 y 10, y



para poder comprobar la independencia de  $\hat{C}_{m,n}^\tau(\varepsilon)$  a la selección del radio epsilon,  $\varepsilon$ , se seleccionaron algunos modelos de la Tabla 4.1, dejando fija la dimensión de inmersión,  $m$ , y variando los valores de  $\varepsilon$ .

Tabla 4.1: Modelos generadores de series de tiempo estacionarias

Modelos	
Modelo 1: N(0,1)	$X_t = e_t \sim N(0, 1)$
Modelo 2: MA(2)	$X_t = e_t + 0.8e_{t-2}; e_t \sim N(0, 1)$
Modelo 3: AR(2)	$X_t = e_t + 0.8X_{t-2}; e_t \sim N(0, 1)$
Modelo 4: NMA(3)	$X_t = e_t + 0.8e_{t-3}^2; e_t \sim N(0, 1)$
Modelo 5: NAR(1)	$X_t =  X_{t-1} ^{0.8} + e_t; e_t \sim N(0, 1)$
Modelo 6: Logístico	$X_t = 4X_{t-1}(1 - X_{t-1});$
Modelo 7: BIL(2,1)	$X_t = 0.6X_{t-2}e_{t-1} + e_t; e_t \sim N(0, 1)$
Modelo 8: ARCH(2)	$X_t = \sqrt{h_t}e_t; h_t = 1 + 0.5X_{t-2}^2; e_t \sim N(0, 1)$
Modelo 9: ARCH(3)	$X_t = \sqrt{h_t}e_t; h_t = 1 + 0.1X_{t-1}^2 + 0.1X_{t-2}^2 + 0.3X_{t-3}^2; e_t \sim N(0, 1)$
Modelo 10: GARCH(1,1)	$X_t = \sqrt{h_t}e_t; h_t = 0.7 + 0.3X_{t-1}^2 + 0.4h_{t-1}; e_t \sim N(0, 1)$

El primer modelo de la Tabla 4.1 corresponde a un proceso de ruido blanco, normal estándar  $N(0,1)$ . Este modelo puede ser considerado como una referencia para la técnica propuesta, ya que todos los retardos proporcionan la misma información a la integral de correlación. Por lo tanto, no se espera encontrar un retardo relevante. Después se centró el estudio en el comportamiento del procedimiento basado en la integral de correlación para seleccionar el desfase óptimo para los procesos lineales, MA(2) y AR(2), modelos 2 y 3 de la tabla respectivamente.

Dado que el procedimiento propuesto puede ser utilizado bajo estructuras no lineales, se incluyeron en el estudio dos procesos no lineales, denominados, NMA(3) y NAR(1), modelos 4 y 5 de la tabla. El modelo autorregresivo caótico logístico, modelo 6, y el modelo bilineal, modelo 7, fueron también incluidos en el estudio, ya que han sido ampliamente examinados, dado que son ambos de distinta naturaleza (determinista y estocástico), y los métodos clásicos de identificación no son

capaces de identificarlos [Granger, 1983].

Los modelos de heterocedasticidad condicionada son comúnmente empleados en aplicaciones financieras, en consecuencia, resulta interesante conocer el comportamiento del procedimiento bajo este tipo de no linealidades en la varianza condicional. Para ello se estudiaron tres modelos diferentes de tipo GARCH con varios retardos, modelos 8,9 y 10 de la tabla respectivamente.

Para estudiar el comportamiento empírico de la técnica propuesta con el tipo de series de tiempo no estacionarias, se incorporaron al estudio los tres modelos lineales que se presentan en la Tabla 4.2.

Tabla 4.2: Modelos generadores de series de tiempo no estacionarias

Modelos	
Modelo 1: CA(1)	$X_t = X_{t-1} + e_t; e_t \sim N(0, 1)$
Modelo 2: CA(2)	$X_t = X_{t-2} + e_t; e_t \sim N(0, 1)$
Modelo 3: CA(3)	$X_t = X_{t-3} + e_t; e_t \sim N(0, 1)$

Los proceso no estacionarios en sentido débil son aquellos que no cumplen algunas o ninguna de las propiedades mencionadas en el Capítulo 2, ecuación (2.40), es decir, no tienen la media ni la varianza constante, o la covarianza entre dos variables no sólo depende de su separación, sino que también depende del tiempo.

Lo anterior significa que el proceso puede ser no estacionario en la media, en la varianza o en las autocorrelaciones o en otras características de la distribución de las variables. Cuando el nivel de la serie no es estable en el tiempo, y tiene alguna tendencia creciente o decreciente, se dice que la serie es no estacionaria en la media. Cuando la variabilidad o las autocorrelaciones se modifican con el tiempo, se dice que la serie no es estacionaria en la varianza o las autocorrelaciones, y si la distribución de la variable en cada instante varía con el tiempo, se dice que la serie no es estacionaria en distribución.

Un gran número de series observadas son no estacionarias, y su nivel varía con el tiempo, por

ejemplo, las series económicas y financieras. Los tres modelos de la Tabla 4.2, corresponden a procesos no estacionarios lineales, denominados *caminatas aleatorias*, ver ecuación (2.41) en el Capítulo 2.

#### 4.2.1. Primer procedimiento de validación: Promedio de la $\hat{C}_{m,n}^{\tau}(\varepsilon)$

La primera validación tiene el objetivo de mostrar el comportamiento de la técnica en cuanto a la identificación del retardo significativo de las series generadas por los modelos de las Tablas 4.1 y 4.2. Para ello se implantó computacionalmente una simulación Monte Carlo que obtiene el valor máximo del promedio de  $\hat{C}_{m,n}^2(\varepsilon)$ . Con este valor se verifica gráficamente que éste es obtenido con la  $\tau_o$  significativa del modelo que genera la serie. El procedimiento de la simulación es el siguiente:

##### Procedimiento 1. Identificación del valor máximo del promedio de $\hat{C}_{m,n}^2(\varepsilon)$

**Paso 1.-** Realizar  $N_r$  replicas de la serie de tiempo  $\{X_i\}_{i=1}^T$  generada con uno de los modelos seleccionados.

**Paso 2.-** Estimar las  $\hat{C}_{m,n}^{\tau}(\varepsilon)$  para las repeticiones de las series, para la dimensión de inmersión  $m = 2, 3, 4$ , y  $5$  y el tiempo de retardo  $\tau = 1, 2, 3, 4, 5$  y  $6$ .

**Paso 3.-** Obtener el promedio de las  $\hat{C}_{m,n}^{\tau}(\varepsilon)$  con las  $N_r$  repeticiones, en cada una de las dimensiones y tiempo de retardo señalados.

**Paso 4.-** Graficar, analizar y presentar los resultados obtenidos.

Para la aplicación de este procedimiento de validación, se consideró un tamaño de muestra de 500 observaciones para cada modelo presentado en la Tablas 4.1 y 4.2. Se estimó la integral de correlación con los parámetros de inmersión:  $m = 2, 3, 4$ , y  $5$ ,  $\tau = 1, 2, 3, 4, 5, 6$  y  $\varepsilon = 1.96\sigma_x$ . Se calculó el promedio de  $\hat{C}_{m,n}^{\tau}(\varepsilon)$  con  $N_r$  repeticiones de las series generadas en cada una de las dimensiones indicadas y en el retardo temporal, y, dada una dimensión de inmersión  $m$ , se graficó el promedio para  $\tau = 1, 2, 3, 4, 5, 6$ . Cada subgráfica obtenida muestra el promedio para una dimensión de inmersión dada y seis distintos retardos.

En el Capítulo 5, Sección 5.1, Tabla 5.1 y en el Apéndice A, Sección A.2, Tablas A.1 a A.3, se

presentan las gráficas que se obtuvieron al aplicar este procedimiento a las series de tiempo estacionarias, y en el Capítulo 6, Sección 6.4.1, Tablas 6.1 a 6.4, se presentan las gráficas correspondientes a las series de tiempo no estacionarias.

#### 4.2.2. Segundo procedimiento de validación: Prueba para identificar retardos en modelos lineales y no lineales.

En cada uno de los enfoques de modelado de las series de tiempo, existe una necesidad de seleccionar el retardo relevante. La relativa escasez de procedimientos estadísticos (especialmente si el proceso generador es de naturaleza no lineal) para facilitar la selección del retardo contrasta con su importancia práctica. Con el fin de mostrar el potencial uso de la integral de correlación en la identificación de qué retardo se debe usar en el desarrollo de modelos no lineales y lineales, con base en la integral de correlación evaluada en diferentes retardos, se construyó este segundo procedimiento de prueba bootstrap que ayuda al investigador en la selección del retardo adecuado para una serie de tiempo dada.

Por razones prácticas y teóricas, al estudiar una serie de tiempo, es importante fundamentar una decisión sobre la inclusión o exclusión de un retardo potencial dado mediante un procedimiento estadístico, por lo que se propone una prueba estadística, construyendo la siguiente hipótesis nula:

**Ho: El retardo temporal  $\tau_o$  no proporciona más información que  $\tau \neq \tau_o$**

Teniendo en cuenta las propiedades de información de la integral de correlación, y de acuerdo con los resultados presentados en el Capítulo 3, se mantiene que bajo la hipótesis nula

$$\delta(\tau) = \hat{C}_{m,n}^{\tau_o}(\varepsilon) - \hat{C}_{m,n}^{\tau}(\varepsilon) = 0, \quad (4.6)$$

mientras que lo contrario ocurrirá bajo la hipótesis alternativa, lo que implicaría que si la hipótesis nula se rechaza, entonces se puede identificar potencialmente un modelo. De esta manera, se puede buscar el retardo que maximiza el valor de la integral de correlación, y utilizar una estimación de  $\delta(\tau)$  para probar si el retardo temporal seleccionado  $\tau_o$  es estadísticamente diferente de cualquier

otro retardo potencial.

El procedimiento para aplicar la prueba estadística tipo bootstrap que se construye es el siguiente:

**Procedimiento 2. Realización de las pruebas estadísticas para obtener la potencia y tamaño de la prueba.**

**Paso 1.-** Considerar la realización de una serie de tiempo  $\{X_t\}_{t=1}^T$  para la cual la hipótesis nula ha sido rechazada.

**Paso 2.-** Encontrar el retardo temporal  $\tau_o$  tal que maximiza la integral de correlación de la serie  $\{X_t\}_{t=1}^T$ , para  $m$  y una  $\varepsilon$  fijas.

**Paso 3.-** Estimar  $\hat{C}_{m,n}^\tau(\varepsilon)$  de la serie  $\{X_t\}_{t=1}^T$ , para  $m$  y una  $\varepsilon$  fijas, para un conjunto de retardos  $\tau \neq \tau_o$ .

**Paso 4.-** Estimar  $\hat{\delta}(\tau) = \hat{C}_{m,n}^{\tau_o}(\varepsilon) - \hat{C}_{m,n}^\tau(\varepsilon)$

**Paso 5.-** Obtener una serie estacionaria bootstrap  $^2 \{X_t\}_{t=1}^T(b)$ , donde  $b$  indica que se refiere a la muestra bootstrap.

**Paso 6.-** Para cada serie bootstrap  $\{X_t\}_{t=1}^T(b)$  se repiten los pasos 2, 3 y 4 para obtener:

- $\hat{C}_{m,n}^{\tau_o,b}(\varepsilon)$
- $\hat{C}_{m,n}^{\tau,b}(\varepsilon)$ , para  $\tau = 1, 2, \dots$ , con  $\tau \neq \tau_o$
- $\hat{\delta}^b(\tau)$ , (4.6) de la serie bootstrap

Se repiten los pasos 5 y 6,  $B - 1$  veces para obtener  $B$  realizaciones bootstrap del estadístico  $\{\hat{\delta}(\tau)\}_{b=1}^B$ .

**Paso 7.-** Calcular el  $p$ -value bootstrap como:

$$P_{boots} = (\hat{\delta}(\tau)) = B^{-1} \sum_{b=1}^B 1(|\hat{\delta}^b(\tau)| \leq |\hat{\delta}(\tau)|); \tau = 1, 2, \dots, \quad (4.7)$$

donde  $1(\bullet)$  es la función indicadora que asigna el valor 1 si el argumento es verdadero y 0 de otra forma.

---

<sup>2</sup>Se sigue a [Politis and White,2004] para seleccionar automáticamente el movimiento de los bloques

La regla de decisión de la prueba es rechazar la hipótesis nula si  $P_{boots}$  satisface  $1 - \alpha/2 < P_{boots}$  o  $P_{boots} < \alpha/2$  para un nivel de significancia  $\alpha$  dado, y no se rechaza en caso contrario.

Es importante hacer notar que debido a que se está usando una prueba tipo bootstrap, y de acuerdo a lo remarcado en el Capítulo 3, la dimensión de inmersión  $m$  no afectaría el resultado.

Con el fin de establecer el rendimiento empírico de la prueba: tamaño y potencia, se establecieron los siguientes escenarios de simulación: para el tamaño de la prueba se consideraron tres de los modelos presentados en la Tabla 4.1; el modelo  $N(0,1)$  para el cual todos los retardos son igualmente irrelevantes y los modelos de procesos de promedios móviles  $MA(2)$ (lineal) y  $NMA(3)$ (no lineal), que tienen la propiedad de no guardar memoria. Por lo que respecta a la potencia de la prueba se consideraron los modelos 2 al 9 presentados en la misma tabla. Para la simulación y la técnica bootstrap se consideraron los parámetros de inmersión:  $m = 3$  y  $\tau = 1, 2, 3, 4, 5, 6$ , y el nivel de significancia de  $\alpha = 0.05$ . Inicialmente se consideraron tamaños de muestra pequeños  $T_1 = 50$ ,  $T_2 = 100$  y  $T_3 = 300$ , y posteriormente se tomaron dos tamaños de muestra grandes  $T_1 = 500$  y  $T_2 = 1000$ . El número de muestras bootstrap en cada caso fue  $B = 199$ , y el número de réplicas para la simulación fue de  $N_r = 1000$ .

Los resultados de la aplicación de este procedimiento a los modelos mencionados se presentan en el Capítulo 5, Sección 5.3, Tablas 5.6 a 5.10, para tamaños de muestra grandes y en el Apéndice A (Apéndice del Capítulo 5, Sección A.3, Tablas A.4 a A7, para tamaños de muestra pequeñas.

### 4.2.3. Procedimiento automático para seleccionar retardos

Considerando un caso particular de la ecuación

$$x_t = G(x_{t_{i_1}}, \dots, x_{t_{i_d}}) + e_t, \quad (4.8)$$

donde  $e_t$  es una secuencia de variables aleatorias con media cero idénticamente distribuidas con un momento de segundo orden finito y constante, es posible diseñar un esquema de selección  $i_1, \dots, i_d$ , donde  $d$  es más bien pequeño, pero donde  $i_d$  puede ser grande. Se puede observar que

las integrales de correlación, como son usadas en la prueba clásica del BDS, no pueden resolver con facilidad este tipo de situaciones. La razón principal es que la dependencia potencial tiene que estar dentro de la  $m$ -historia, de lo contrario la dependencia no sería detectada. El punto crítico es que usando una  $m$ -historia fija (y no relevante) se pueda buscar a través de la serie de tiempo por medio del retardo  $\tau_o$ . En otras palabras, para una determinada  $m$  pequeña, digamos 3, sería posible, por ejemplo, detectar una variación cíclica a largo plazo mediante el uso de un retardo apropiado ( $\tau$ ).

Esto es importante ya que el procedimiento de selección de la herramienta propuesta no se verá necesariamente afectado por el bien conocido problema de la “maldición de la dimensionalidad” que afecta regularmente a los métodos no paramétricos.

A fin de evaluar si la técnica es útil o no, se consideró el siguiente modelo:

**Modelo 11:**  $X_t = -0.4X_{t-6} + 0.4X_{t-10} + e_t$ , para el cual  $d = 2$ ,  $i_1 = 6$ ; e  $i_2 = 10$ ,

y se llevó a cabo una simulación Monte Carlo similar a la presentada en la Sección 4.2.2. En este caso se utilizaron 15 retardos posibles, y se evaluó el procedimiento con  $T = 1000$  y 500 observaciones. Los resultados obtenidos al aplicar este procedimiento a este modelo estacionario se presentan en el Capítulo 5, Subsección 5.3.3, Tablas 5.11 a 5.13, y en el Apéndice A, Sección A.4.1, Tablas A.8 y A.9 para muestras de tamaño pequeño.

#### 4.2.4. Procedimiento para seleccionar retardo con independencia del radio

Con el fin de comprobar la sensibilidad de la técnica con respecto al radio epsilon, uno de los parámetros que debe ser considerado para el cálculo de la integral de correlación, se realizaron simulaciones Monte Carlo como la mencionada en la sección 4.2.1, promedio de la  $\hat{C}_{m,n}^\tau(\varepsilon)$ , para un tamaño de muestra relativamente pequeño  $T = 50, 100, 200$  y 500, para algunos modelos de la Tabla 4.1,  $N(0,1)$ ,  $AR(2)$ ,  $NMA(3)$ , y  $ARCH(2)$ , fijando  $m = 3$ , y variando el epsilon,  $\epsilon = a\sigma$ ;  $a = 0.5, 1.0, 1.5, 2.0$ .

De esta forma se podrá comprobar que la  $\hat{C}_{m,n}^{\tau}(\varepsilon)$  es máxima para el retardo correcto independientemente del epsilon que se asigne. Los resultados obtenidos se presentan en el Capítulo 5, Sección 5.2, Tablas 5.2 a 5.5.

### **4.3. Procedimiento para el Uso de las Técnicas Propuestas a los Residuos Estimados de un Modelo Ajustado.**

En ocasiones es muy común que al aplicar procedimientos de pruebas de diagnóstico los investigadores estén interesados en eliminar, a través de un filtro lineal, la dependencia lineal (en la media condicional) que pueda estar presente en los datos. Este proceso de eliminación consiste en ajustar un modelo lineal a los datos sin procesar, para posteriormente obtener y analizar los residuales (errores estimados).

Del mismo modo que en las secciones anteriores se aborda el tema de la utilidad de las técnicas propuestas, en esta sección se considera el caso en que el investigador necesita identificar la estructura dinámica del retardo en modelos que presentan algún tipo de dependencia, lineal o no lineal, en los momentos de condicionalidad de orden mayor. Si existe una dependencia no lineal en los residuos, la prueba propuesta será capaz de sugerir que retardo(s) usar en la modelización de los momentos condicionales, lo cual es especialmente de interés en el modelado financiero.

Sin embargo, a menudo un estadístico se comporta correctamente con los datos originales, pero no con los residuales estimados. Este hecho aparece generalmente porque el estadístico de interés toma la forma de un miembro de una familia común, a excepción de algún parámetro vital desconocido y que por lo tanto debe estimarse. Es decir, la distribución asintótica de la prueba puede verse afectada por el proceso de estimación. Esas pruebas donde el paso intermedio no afecta a la distribución de la prueba son conocidas como pruebas libres de parámetros de ruido.

Las condiciones suficientes bajo las cuales se basan las pruebas de correlación libres de parámetros de ruido, han sido estudiadas por [De Lima, 1996]. Para asegurar la propiedad de invariancia, es



necesario que el modelo sea aditivo, tal como,  $y_t = f(\mathbf{Y}_{t-1}, \beta) + e_t$ , o que los modelos se puedan transformar en este formato. Por el contrario, en nuestro caso, la prueba propuesta de bootstrap no está limitada por esas condiciones suficientes. Por lo tanto, considerando un modelo general de error de la forma  $y_t = f(\mathbf{Y}_{t-1}, \beta, e_t)$ , donde  $\beta$  es un vector de parámetros consistentemente estimados, y  $\{e_t\}$  un término de error, indirectamente se supone que la función  $f$  se puede invertir y  $e_t = G(y_t, \mathbf{Y}_{t-1}, \beta)$ , de modo que la técnica para identificar el retardo significativo puede ser aplicada a los términos del verdadero error, ver Capítulo 3, sección 3.3. Las condiciones para la estimación consistente de la  $\hat{e}_t = \hat{G}(y_t, \mathbf{Y}_{t-1}, \hat{\beta})$  y para la estimación consistente de las integrales de correlación son las mismas que en aquellas de [De Lima,1996].

Para mostrar que la prueba puede utilizarse en el diagnóstico de la serie, se considera el cuadrado de los residuales  $\{\hat{e}_t^2\}$  para varios DGP's de la forma general:

$$x_t = \mu(X_{t-1}, \beta) + \sigma(X_{t-1}, \gamma)e_t, \quad (4.9)$$

donde  $\beta$  y  $\gamma$  son parámetros  $\sqrt{n}$ -consistentes de un proceso GARCH.

Subsecuentemente, para obtener la serie de residuales estimados, se aplica un filtro lineal (un proceso de tipo ARMA(p,q) a la serie generada por (4.9) para obtener la serie residual estimada  $\hat{e}_t = x_t - \hat{x}_t$ . De la serie  $\{\hat{e}_t\}$ , se calcula la serie de residuales al cuadrado  $\{\hat{e}_t^2\}$ . La validación empírica del rendimiento de la técnica aplicada a la serie de residuos cuadrados, se llevó a cabo de manera similar a la presentada en la sección 4.2.2.

Los modelos ARMA-GARCH y los filtros de ARMA(p,q) utilizado para generar la serie de tiempo y estimar los residuales son los que aparecen en la Tabla 4.3. El primer modelo corresponde a un proceso de ruido blanco iid., normal estándar  $N(0,1)$ , mismo que ha sido considerado para evaluar el tamaño de la prueba.

Las técnicas de simulación y bootstrap se aplicaron considerando los parámetros de inmersión:  $m = 3$  y  $\tau = 1, 2, 3, 4, 5, 6$ , para un nivel de significancia de  $\alpha = 0.05$ . Los tamaños de muestra son  $T_1 = 500, T_2 = 1000, T_3 = 2000$ , y, el número de muestras bootstrap en cada caso fueron  $B = 199$ . El número de replicas de la simulación fue de  $N_r = 1000$ .

Tabla 4.3: Modelos para generar  $\{X_t\}$  y filtro para los residuales  $\{\hat{e}_t\}$

Modelos		
Modelo		Filtro lineal ARMA(p,q)
N(0,1)	$X_t = e_t \sim N(0, 1)$	ARMA(1,1)
ARMA(1,0)-GARCH(1,1)	$X_t = 0.1 + 0.5X_{t-1} + e_t;$	ARMA(1,0)
	$e_t = \sqrt{h_t}z_t; z_t \sim N(0, 1)$	
	$h_t = 0.7 + 0.3e_{t-1}^2 + 0.4h_{t-1}$	
ARMA(0,1)-GARCH(1,1)	$X_t = 0.1 - 0.3e_{t-1} + e_t$	ARMA(0,1)
	$e_t = \sqrt{h_t}z_t; z_t \sim N(0, 1)$	
	$h_t = 0.7 + 0.3e_{t-1}^2 + 0.4h_{t-1}$	
ARMA(1,1)-GARCH(1,1)	$X_t = 0.1 + 0.5X_{t-1} - 0.3e_{t-1} + e_t$	ARMA(1,1)
	$e_t = \sqrt{h_t}z_t; z_t \sim N(0, 1)$	
	$h_t = 0.7 + 0.3e_{t-1}^2 + 0.4h_{t-1}$	

Los resultados obtenidos se presentan en el Capítulo 5, Sección 5.4, Tablas 5.14 a 5.19, y en el Apéndice A (Apéndice del Capítulo 5), Sección A.5, Tablas A.10 a A.15, para tamaños de muestra pequeños.

## 4.4. Software Utilizado en las Simulaciones Monte Carlo y Bootstrap

Las simulaciones Monte Carlo y Bootstrap se llevaron a cabo ejecutando los programas computacionales que se desarrollaron y/o adaptaron para implantar los procedimientos presentados en las secciones anteriores. Los programas se elaboraron y ejecutaron dentro del entorno de trabajo del software MATLAB para los sistemas operativos Windows y Linux para PC, y cumplen las funciones de:

- Generar las series de tiempo de los procesos que representan los modelos presentados en las Tablas 4.1, 4.2, y 4.3, así como el modelo 11 presentado en la Sección 4.2.3.
- Calcular el estimador de la integral de correlación,  $\hat{C}_{m,n}^{\tau}(\varepsilon)$
- Obtener el promedio de las  $\hat{C}_{m,n}^{\tau}(\varepsilon)$  con las  $N_r$  repeticiones, en cada una de las dimensiones y tiempo de retardo fijados.
- Graficar los promedios obtenidos para las diferentes dimensiones de inmersión y retardos temporales.
- Estimar los modelos ARMA para filtrar la parte lineal de las series de tiempo generadas con modelos ARCH y GARCH.
- Obtener los residuales al cuadrado del ajuste realizado con los modelos ARMA.
- Obtener las muestras Bootstrap

- Aplicar las pruebas estadísticas para mostrar el potencial de uso de la integral de correlación en la identificación del retardo significativo del modelo generador de la serie.

Para el cálculo de la estimación de integrales de correlación se utilizó la función BDSG desarrollada por [Matilla-García et al,2004], que a diferencia de la función BDS proporcionada en [LeBaron, 1997], permite variar los valores del parámetro de retardo temporal,  $\tau$ . Por lo que respecta a la estimación de los modelos ARMA usados para filtrar la parte lineal de las series, se utilizó la función ARMAX del *System Identification Toolbox de MATLAB*.

---

---

## CAPÍTULO 5

---

# RESULTADOS EMPÍRICOS DE LA APLICACIÓN DE LA TÉCNICA

En este capítulo se presentan los resultados numéricos de las simulaciones realizadas para validar y evaluar el comportamiento de la técnica y del procedimiento estadístico no paramétrico propuestos para identificar qué retardos usar cuando se desarrollen modelos de naturaleza desconocida (lineales o no lineales). Asimismo, se presentan los resultados de la aplicación de la técnica de bootstrap basada en la prueba estadística presentada en el Capítulo 4, Sección 4.2.2 para seleccionar entre desfases potencialmente relevantes.

En la Sección 5.1 se presentan los resultados numéricos obtenidos al aplicar el Procedimiento 1, presentado en la Sección 4.2.1 del Capítulo 4, a una serie de tiempo estacionaria,  $\{X_t\}$ . Estos resultados muestran que el máximo valor promedio de la integral de correlación se obtiene al calcularse, para una  $m$  y una  $\varepsilon$  dada, con el retardo óptimo  $\tau_o$ . En la Sección 5.2 los resultados que se presentan muestran también que dada una  $\{X_t\}$  estacionaria, el máximo valor promedio de la integral de correlación se obtiene al calcularse, para una  $m$  dada y variando la distancia  $\varepsilon = a\sigma_x$ ;  $a = 0.5, 1.0, 1.5, 2.0$ , con el retardo óptimo  $\tau_o$ .

En la Sección 5.3 se presentan los resultados numéricos obtenidos al aplicar el Procedimiento 2, presentado en la Sección 4.2.2 del Capítulo 4 para obtener la potencia y tamaño de la prueba bootstrap. Esta prueba ayuda al investigador en la selección del retardo adecuado para una serie de tiempo dada. En la Sección 5.4 se muestran los resultados de la simulación Monte Carlo aplicada al Modelo 11, introducido en la Sección 4.2.3 del Capítulo 4, el cual tiene retardos significativos 6 y 10, utilizaron para su evaluación 15 retardos.

Por último en la Sección 5.5 se presentan los resultados de las simulaciones en donde se aplica el procedimiento que evalúa el comportamiento de la técnica al ser usada también para identificar el retardo óptimo de la serie de tiempo de los residuales al cuadrado, obtenidos al filtrar la serie de datos con un proceso  $ARMA(p, q)$ . Este procedimiento se presenta en la Sección 4.3 del Capítulo 4.

## **5.1. Resultados de la Aplicación del Procedimiento 1:**

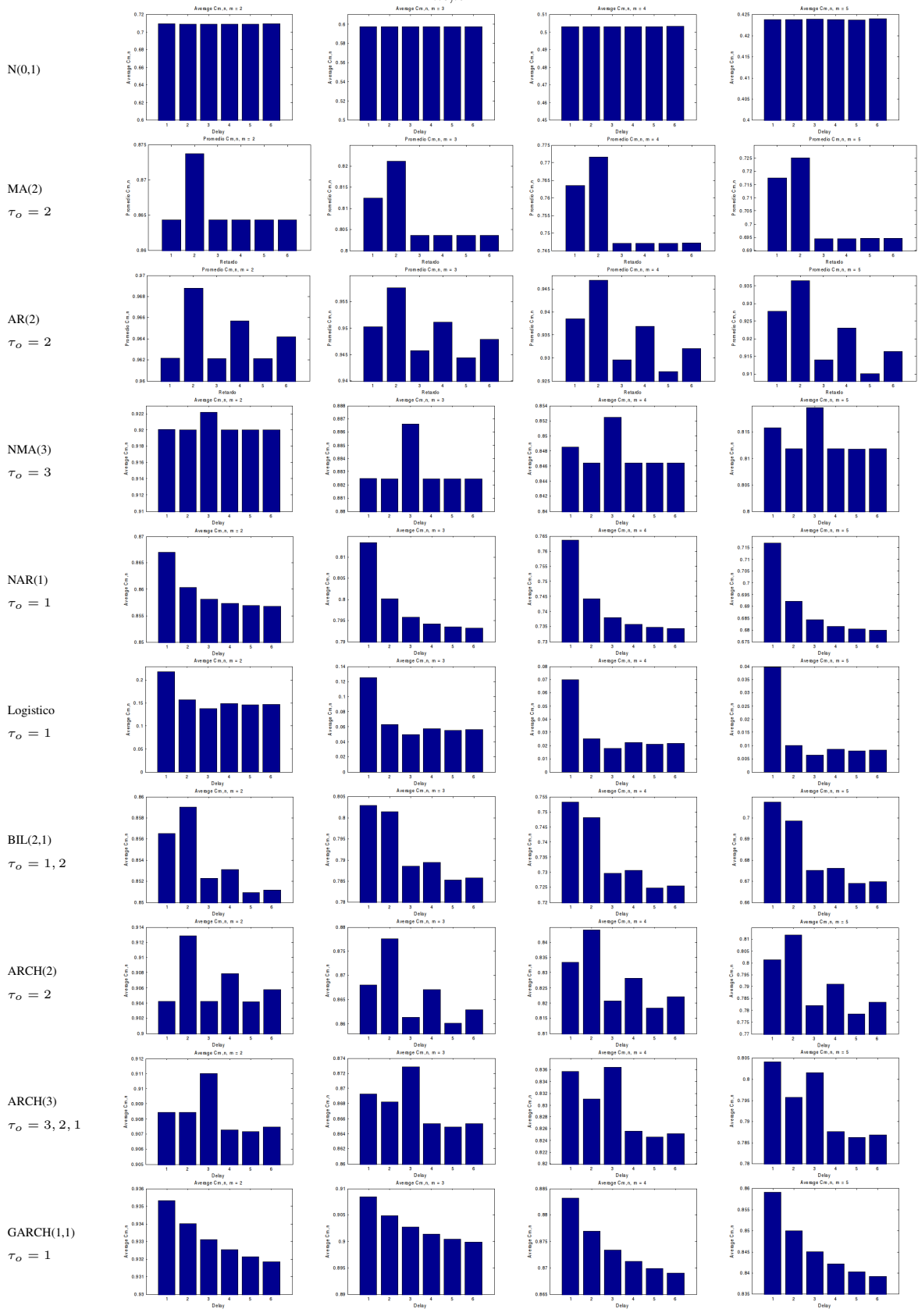
### **Identificación del Retardo a Partir del Valor Máximo del Promedio de $\hat{C}_{m,n}^{\tau}(\epsilon)$**

La aplicación del procedimiento de identificación del retardo significativo a partir del valor máximo del promedio  $\hat{C}_{m,n}^{\tau}(\epsilon)$ , como se mencionó en la Sección 4.2.1 del Capítulo 4, se llevó a cabo mediante una simulación Monte Carlo de los modelos estacionarios analizados.

Los resultados obtenidos de las simulaciones de los modelos, presentados en la Tabla 4.1, del Capítulo 4, considerando un tamaño de muestra de  $T = 500$  observaciones para cada modelo, un número de replicas para la simulación de  $N_r = 1000$ , y estimando la integral de correlación con los parámetros de inmersión:  $m = 2, 3, 4, 5$ ,  $\tau = 1, 2, 3, 4, 5, 6$  y  $\epsilon = 1.96\sigma_x$ , aparecen en la Tabla 5.1.

Antes de comentar los resultados de cada modelo, vale la señalar que una propiedad común e

Tabla 5.1: Promedios de  $\hat{C}_{m,n}^{\tau}$  para procesos estacionarios



importante es que el comportamiento que muestran las diferencias para cada retardo, no depende de la dimensión de inmersión. Esto es muy útil en la práctica, ya que no es crucial para los resultados la selección del parámetro  $m$ , lo que confirma empíricamente, que se puede evitar el problema común de la “*maldición de dimensionalidad*”.

Las gráficas del promedio de  $\hat{C}_{m,n}^\tau$  para  $N(0, 1)$  no muestran, como se esperaba, diferencias entre los valores promedios para  $\tau = 1, 2, 3, 4, 5, 6$ , y por tanto no existe un retardo relevante. Para los modelos lineales, se observa primero que el retardo correcto siempre reporta el valor más grande para la integral de correlación, y por tanto identifica el retardo apropiadamente. Particularmente, los modelos MA(2) y AR(2) tienen propiedades diferentes de memoria, y también esto es detectado claramente con el procedimiento propuesto basado en la integral de correlación: la integral de correlación para los retardos arriba de 2 del MA(2) no cambian, mientras que la propiedad de memoria descendente del AR(2) se observa claramente para el orden de retardo por encima de 2. Para los siguientes dos modelos no lineales, NMA(3) y NAR(1), la integral de correlación detecta claramente los retardos correctos y también muestra sensibilidad a las propiedades de memoria del modelo. El orden de dependencia del modelo caótico (Logístico) es detectado en el retardo correcto, y lo mismo ocurre para el proceso bilinial, BIL(2,1), para el cual dos retardos, 2 y 1, son relevantes.

En presencia de estructura de varianza condicional, el procedimiento también detecta los retardos correctos correspondientes. Nótese que el ARCH(2) depende solamente del retardo 2, mientras que el modelo ARCH(3) incorpora varios retardos relevantes con diferente fuerza en los valores de los parámetros ( 0.1 para los retardos 1 y 2, y 0.3 para el retardo 3 ), y estos hechos parecen ser capturados por la aproximación de la integral de correlación. Lo mismo ocurre con el modelo GARCH(1,1), para el cual, la dependencia es encontrada, como era de esperarse, en el retardo 1.

También se analizó el comportamiento del procedimiento para tamaños de muestra pequeños de 50, 100 y 200. Los gráficos correspondientes aparecen en el Apéndice A de este Capítulo en las Tablas A.1, A.2 y A.3 respectivamente.



## 5.2. Resultados Numéricos de la Identificación del Retardo a Partir del Valor Máximo del Promedio de $\hat{C}_{m,n}^\tau(\varepsilon)$ , Modificando el Valor del Parámetro $\varepsilon$

Como se mencionó en la Sección 4.2.4 del Capítulo 4, la sensibilidad de la técnica con respecto al radio  $\varepsilon$ , parámetro de inmersión, se verificó a través de simulaciones Monte Carlo obteniéndose el promedio de la  $\hat{C}_{m,n}^\tau(\varepsilon)$ , para tamaños de muestra relativamente pequeños  $T = 50, 100, 200$  y de un tamaño mayor de 500, y un número de replicas para la simulación  $N_r = 1000$ , para algunos modelos de la Tabla 4.1,  $N(0,1)$ ,  $AR(2)$ ,  $NMA(3)$ , y  $ARCH(2)$ , fijando el valor de la dimensión de inmersión a  $m = 3$ , y variando el valor de  $\varepsilon$ ,  $\varepsilon = a\sigma$ ,  $a = 0.5, 1.0, 1.5, 2.0$

Los valores de retardo considerados fueron  $\tau = 1, 2, 3, 4, 5, 6$ . Los resultados obtenidos se presentan en las Tablas 5.2, 5.3, 5.4 y 5.5.

Tabla 5.2: Promedios de  $\hat{C}_{m,n}^\tau$  para procesos estacionarios,  $m = 3$ ,  $T_1 = 50$ , variando el valor de  $\varepsilon = a\sigma_x$

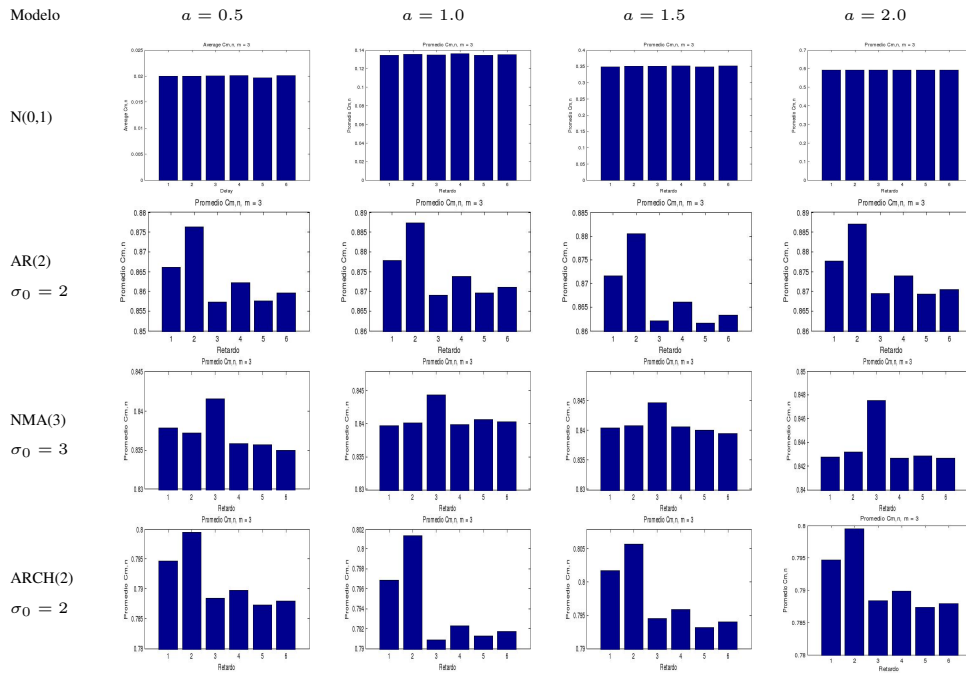


Tabla 5.3: Promedios de  $\hat{C}_{m,n}^{\tau}$  para procesos estacionarios,  $m = 3$ ,  $T_2 = 100$ , variando el valor de

$$\epsilon = a\sigma_x$$

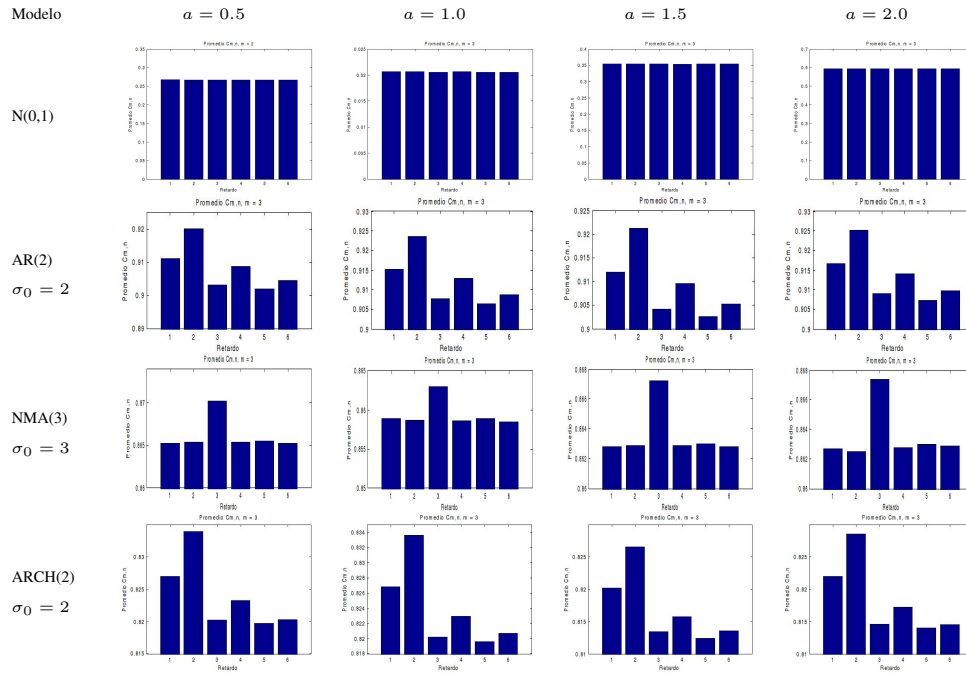


Tabla 5.4: Promedios de  $\hat{C}_{m,n}^{\tau}$  para procesos estacionarios,  $m = 3$ ,  $T_3 = 200$ , variando el valor de

$$\epsilon = a\sigma_x$$

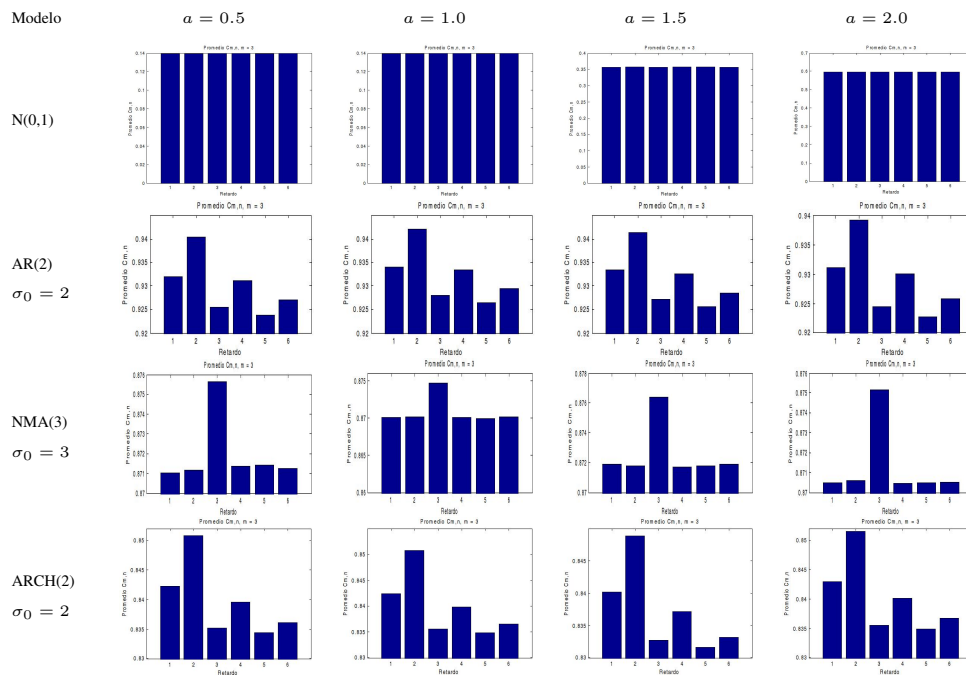
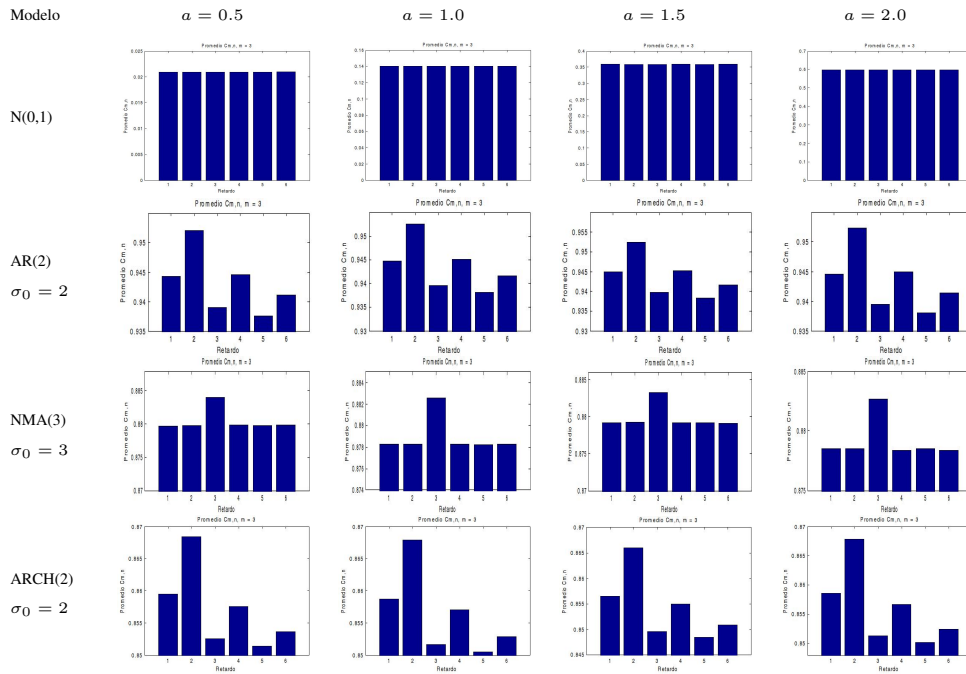


Tabla 5.5: Promedios de  $\hat{C}_{m,n}^\tau$  para procesos estacionarios,  $m = 3$ ,  $T_4 = 500$ , variando el valor de  $\epsilon = a\sigma_x$



Las gráficas del promedio de  $\hat{C}_{m,n}^\tau$  en las cuatro Tablas 5.2, 5.3, 5.4 y 5.5 para  $N(0,1)$  no muestran, como se esperaba, diferencias entre los valores promedios para  $\tau = 1, 2, 3, 4, 5, 6$ , en los cuatro distintos valores de  $a$ , y por tanto no existe un retardo relevante.

Para el modelo lineal  $AR(2)$ , se observa que el retardo correcto siempre reporta el valor más grande para la integral de correlación, y por tanto identifica el retardo apropiadamente, así como la propiedad descendiente de memoria del modelo, para el orden de retardo superior a 2. Para el modelo no lineal,  $NMA(3)$  la integral de correlación detecta claramente el retardo correcto y que los retardos arriba de 3 no cambian, debido a la sensibilidad a la propiedad de memoria de este modelo.

En presencia de estructura de varianza condicional, la integral de correlación también detecta el retardo correcto para el  $ARCH(2)$  que depende solamente del retardo 2.

La conclusión que se obtiene de estos resultados es que la integral de correlación,  $\hat{C}_{m,n}^\tau$ , es máxima para el retardo correcto con independencia del radio .

### **5.3. Resultados de la Aplicación del Procedimiento 2: Prueba para Identificar Retardos en Modelos Lineales y No Lineales.**

Cómo se mencionó en la Sección 4.2.2 del Capítulo 4, con el fin de evaluar el rendimiento empírico de la prueba en cuanto a su tamaño y potencia, se utilizaron algunos de los modelos presentados en la Tabla 4.1 para generar muestras finitas. La simulación y la técnica bootstrap se aplicaron considerando los parámetros de inmersión:  $m = 3$  y  $\tau = 1, 2, 3, 4, 5, 6$ , y el nivel de significancia de  $\alpha = 0.05$ . Inicialmente se consideraron dos tamaños de muestra  $T_1 = 500$  and  $T_2 = 1000$ ; el número de muestras bootstrap en cada caso fue  $B = 199$ , y el número de réplicas para la simulación fue de  $Nr = 1000$ .

#### **5.3.1. Experimentos para verificar el tamaño de la prueba**

Para llevar a cabo los experimentos del tamaño de la prueba se consideraron tres de los modelos de la Tabla 4.1: uno de distribución normal iid.,  $N(0,1)$ , para el cual todos los retardos son igualmente poco importantes y dos modelos de procesos de medias móviles: uno lineal de orden 2, MA(2)(modelo 2), y para completez del análisis uno no lineal, NMA(3)(modelo 4), ambos sin la propiedad de guardar memoria más allá de los retardos 3 y 4 respectivamente. Las Tablas 5.6, 5.7 y 5.8 muestran los resultados obtenidos. Estas tablas deben ser leídas por renglón: por ejemplo, en la Tabla 5.6 el primer renglón con resultados presenta el tamaño estimado cuando se prueba la hipótesis nula de que el retardo 1 proporciona la misma información que los retardos 2, 3, 4, 5 y 6; la prueba de la nula para el segundo renglón es que el retardo 2 proporciona la misma información para los retardos 1, 3, 4, 5 y 6; y el último renglón presenta los resultados cuando se prueba el retardo 6 contra los retardos 1, 2, . . . , 5.

Tabla 5.6: Tamaño de la prueba para el Modelo 1 con  $T = 500$  y  $1000$

$$X_t = e_t, e_t \sim N(0, 1)$$

$\tau$	1	2	3	4	5	6
1		0.053	0.033	0.043	0.054	0.046
		0.054	0.065	0.056	0.050	0.062
2	0.048		0.039	0.043	0.055	0.051
	0.051		0.044	0.061	0.055	0.046
3	0.054	0.051		0.058	0.044	0.049
	0.051	0.034		0.061	0.052	0.050
4	0.036	0.047	0.043		0.050	0.044
	0.061	0.051	0.051		0.061	0.057
5	0.052	0.045	0.055	0.038		0.023
	0.034	0.045	0.050	0.041		0.045
6	0.047	0.048	0.051	0.055	0.047	
	0.045	0.053	0.053	0.058	0.052	

Nótese que para el Modelo 1 todos los retardos teóricamente contienen la misma información (son todos igualmente no relevantes), y por tanto se está bajo la hipótesis nula. Los resultados muestran que se está en el nivel nominal de rechazo esperado.

Las Tablas 5.7 y 5.8 deben ser leídas de la misma forma que la anterior. En este caso se están utilizando modelos con dependencia lineal y no lineal. En ambos modelos no existe una relación teórica más allá de cierto retraso, y en consecuencia se prueban diferentes configuraciones posibles de la hipótesis nula: por ejemplo en la Tabla 5.7, se prueba el retardo 3 contra los retardos 4, 5 y 6; el retardo 4 contra los retardos 3, 5 y 6 y así sucesivamente. De igual forma para la Tabla 5.8, se prueba el retardo 4 contra los retardos 5 y 6; el retardo 5 contra los retardos 4 y 6 y el retardo 6 contra los retardos 4 y 5.

Tabla 5.7: Tamaño de la prueba para el Modelo 2 con  $T = 500$  y  $1000$

$$X_t = e_t + 0.8e_{t-2}; e_t \sim N(0, 1)$$

$\tau$	3	4	5	6
3		0.046	0.036	0.050
		0.040	0.042	0.047
4	0.043		0.057	0.044
	0.047		0.046	0.047
5	0.045	0.052		0.046
	0.038	0.051		0.042
6	0.064	0.042	0.047	
	0.061	0.033	0.049	

Tabla 5.8: Tamaño de la prueba para el Modelo 4 con  $T = 500$  y  $1000$

$$X_t = e_t + 0.8e_{t-3}^2; e_t \sim N(0, 1)$$

$\tau$	4	5	6
4		0.042	0.061
		0.039	0.064
5	0.047		0.062
	0.040		0.061
6	0.053	0.055	
	0.057	0.047	

Los resultados muestran que para el nivel nominal seleccionado de la prueba, su rendimiento bajo la hipótesis nula es correcta. Esta conclusión es especialmente importante dado el hecho de que otras pruebas basadas en la integral de correlación, como el estadístico tipo BDS, tienen graves distorsiones de tamaño para muestras finitas, ver [Kanzler,1999].

En la Sección A.3 del Apéndice A de este Capítulo, en las Tablas A.4, A.5 y A.6, se presenta el

estudio de tamaño de la prueba considerando tamaños de muestra más pequeños, 50, 100 y 200, y la misma hipótesis nula.

### 5.3.2. Experimentos para verificar la potencia de la prueba

Las Tablas 5.9 y 5.10 presentan los resultados de los experimentos que se hicieron para la verificación de la potencia de la prueba. Se realizaron simulaciones Monte Carlo y Bootstrap con los modelos 2 al 9, presentados en la Tabla 4.1. Las Tablas 5.9 y 5.10 deben leerse renglón por renglón (modelo por modelo). Cada renglón (modelo) tiene una celda vacía cuya columna es indicativa del retardo que se está usando para la prueba: por ejemplo, la celda vacía para el modelo MA(2) es la columna 2, lo que significa que, para este modelo en particular, se está probando el retardo 2 contra los restantes retardos denominados, 1, 3, 4, 5 y 6. Esto es porque el retardo relevante para el modelo MA(2) es el 2, y entonces se estima la probabilidad de rechazar la hipótesis nula de que el retardo 2 no tiene mayor información que cualquier otro retardo considerado.

Tabla 5.9: Potencia de la prueba para series generadas de tamaño  $T1 = 500$

Modelo	Retardo					
	1	2	3	4	5	6
MA(2)	0.974		1.0	1.0	1.0	0.999
AR(2)	1.0		1.0	1.0	1.0	1.0
NMA(3)	0.568	0.603		0.621	0.595	0.773
NAR(1)		1.0	1.0	1.0	1.0	1.0
Logistic		1.0	1.0	1.0	1.0	1.0
BL(2,1)	0.99		0.991	0.812	0.989	0.887
ARCH(2)	0.92		0.993	0.986	0.99	0.976
ARCH(3)				0.734	0.756	0.875
GARCH(1,1)		0.569	0.703	0.843	0.874	0.898

Tabla 5.10: Potencia de la prueba para series generadas de tamaño  $T_2 = 1000$

Modelo	Retardo					
	1	2	3	4	5	6
MA(2)	1.0		1.0	1.0	1.0	1.0
AR(2)	1.0		1.0	1.0	1.0	1.0
NMA(3)	0.891	0.917		0.916	0.908	0.979
NAR(1)		1.0	1.0	1.0	1.0	1.0
Logistic		1.0	1.0	1.0	1.0	1.0
BL(2,1)	1.0		1.0	0.809	1.0	0.899
ARCH(2)	0.995		1.0	0.999	1.0	1.0
ARCH(3)				0.963	0.961	0.995
GARCH(1,1)		0.841	0.96	0.983	0.995	0.993

El rendimiento de la potencia de la prueba es extremadamente bueno, para ambos tamaños de muestra, para los modelos lineales (MA y AR), y es interesante que lo mismo aplica para el proceso autorregresivo no lineal, el mapeo logístico determinista y el bilineal.

Estos resultados son especialmente interesantes porque:

- 1) Para el modelo bilineal, el cual exhibe una dependencia a través de los dos primeros retardos, las funciones de autocorrelación y correlación parcial teóricamente son la mismas a la del ruido blanco y por tanto si se usarán los métodos clásicos basados en las correlaciones se obtendrían conclusiones erróneas.
- 2) Algo similar ocurre con el mapeo logístico, con el problema adicional de ser puramente determinista.
- 3) Como se menciona en [Granger and Lin, 1994], la prueba de Kendal tau se desempeña pobremente para procesos NAR. Para el modelo de promedios móviles no lineal generado con un retardo de orden 3, la potencia de la prueba es relativamente buena para tamaños de



muestra grande, mientras que los resultados son más discretos cuando se consideran 500 observaciones.

Los resultados anteriores permiten señalar que la nueva prueba propuesta es especialmente útil en situaciones de incertidumbre acerca de la naturaleza lineal o no lineal de la media condicional del proceso, puesto que como procesos lineal y no lineal son claramente detectados al considerar integrales de correlación en diferentes retardos.

Por lo que respecta a los modelos condicionales heteroscedásticos, el comportamiento empírico de la prueba, en términos de potencia, es ciertamente de alta calidad. Los retardos apropiados (digamos, 2 y 1, respectivamente) son detectados para los modelos ARCH(2) y GARCH(1,1). Los retardos 1, 2 y 3 son todos relevantes (siendo este último el de mayor importancia relativa, por lo tanto las celdas correspondientes están vacías) para el ARCH(3). Como el retardo 3 es el parámetro con mayor peso relativo en el modelo, la potencia reportada es probando el retardo 3 contra 4, 5 y 6. El rendimiento de la potencia es muy alto, especialmente para el tamaño de muestra grande ( $T = 1000$ ).

En el Apéndice A de este Capítulo, Sección A.3, se pueden encontrar las Tablas A.4, A.5, A.6 y A.7 para muestras de tamaño pequeño. Como se verá en este caso de muestras de tamaño pequeño, la principal conclusión obtenida de los resultados es que, dada la naturaleza no paramétrica, la potencia es buena únicamente para los modelos lineales y caóticos cuando los tamaños de muestra están por debajo de 100. Cuando hay incertidumbre acerca de la naturaleza (lineal o no lineal) del proceso, se sugiere utilizar tamaños de muestra superiores a 200.

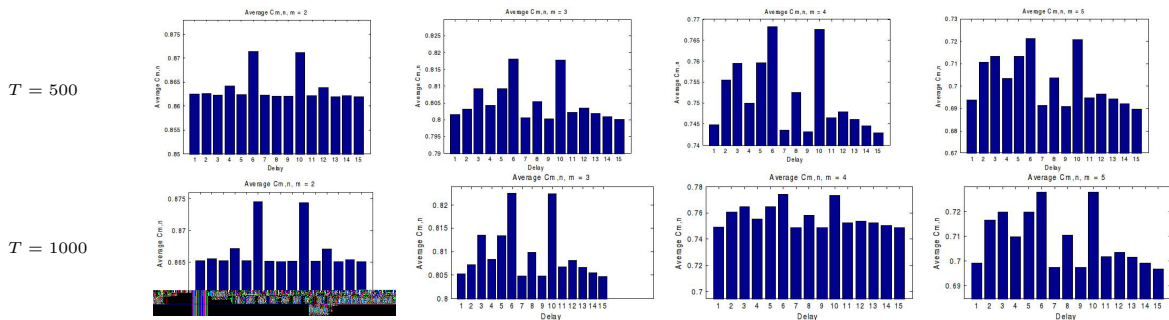
### **5.3.3. Resultados numéricos de la aplicación del procedimiento automático para seleccionar retardos**

Como se mencionó en la Sección 4.2.3 del Capítulo 4, a fin de evaluar la utilidad de las técnicas propuestas, se consideró una simulación Monte Carlo-Bootstrap similar a la presentada en la

Sección 4.2.2, utilizando el modelo AR(10),  $X_t = -0.4X_{t-6} + 0.4X_{t-10} + e_t$ , (Modelo 11) con retardos significativos 6 y 10. La evaluación del procedimiento se hizo considerando un conjunto de 15 posibles retardos, del 1 al 15. Se puede observar que esto implicaría realizar un buen número de posibles combinaciones de la estructura de configuraciones del retardo potencial, lo que significaría tener que calcular la  $\sum_{d=1}^k \binom{k}{d}$ , donde  $k$  es el número de retardos que se utilizan para generar la serie temporal y  $d$ , el número de los retardos específicos,  $i_1^* \dots, i_d^*$ . Para este caso, ya que  $k = 15$  y  $d = 1, 2, \dots, 15$ , se tendrían  $\sum_{d=1}^{15} \binom{15}{d} = 32767$  posibles combinaciones, de las cuales la única correcta es  $d = 2, i_1^* = 6, i_2^* = 10$ .

Los resultados que se obtienen al aplicar el procedimiento de identificación de los retardos significativos 6 y 10 a partir del valor promedio de la integral de correlación del modelo, se presentan en la Tabla 5.11. Las gráficas de la Tabla 5.11 corresponden a los tamaños de muestra  $T = 500$  y  $T = 1000$ ,  $m = 2, 3, 4$  y  $5$ , y  $\tau = 1, 2, \dots, 6, \dots, 10, \dots, 15$ , para un número de réplicas para las simulaciones de  $N_r = 1000$ , y para obtener el tamaño y prueba de la hipótesis, se usaron  $B = 199$  muestras bootstrap, ver resultados en las Tablas 5.12 y 5.13. En el Apéndice A de este Capítulo en la sección A.4, se presentan en las Tablas A.8 y A.9 los resultados correspondientes a los tamaños de muestra pequeños  $T = 70, 100$  y  $200$ .

Tabla 5.11: Promedios de la  $\hat{C}_{m,n}^\tau$  para el Modelo 11;  $T = 500$  y  $1000$   $\tau = 6, 10$



Se puede observar en las gráficas del promedio de  $\hat{C}_{m,n}^\tau$  para este modelo que los retardos con el valor más grande para la integral de correlación, son los correspondientes al retardo 6 y al 10, y

Tabla 5.12: Resultados Tamaño de la Prueba, para el Modelo 11,  $AR(10)$

Retardo Selecc.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
6															
T=1000	0.987	0.981	0.962	0.988	0.948		0.991	0.953	0.992	0.049	0.986	0.994	0.976	0.984	0.979
T=500	0.928	0.921	0.892	0.928	0.880		0.931	0.903	0.932	0.061	0.926	0.934	0.916	0.924	0.969
10															
T=1000	0.977	0.974	0.664	0.976	0.859	0.042	0.992	0.950	0.990		0.980	0.953	0.989	0.990	0.987
T=500	0.957	0.954	0.764	0.956	0.829	0.062	0.942	0.920	0.940		0.940	0.953	0.959	0.960	0.977

que por lo tanto se está identificando el retardo apropiadamente.

Por lo que respecta a la evaluación del rendimiento de la prueba con respecto al tamaño de la prueba, el procedimiento se evaluó para  $T = 500$  y  $1000$  observaciones. Los resultados se presentan en la Tabla 5.12. Esta tabla se debe leer renglón por renglón, los tres primeros renglones corresponden a uno de los retardos significativos del modelo, el 6, para los tamaños de muestra 1000 y 500, cuyos renglones tienen una celda vacía, cuya columna es indicativa de este retardo, lo que significa que se está probando el retardo 6 contra los restantes retardos denominados  $1, \dots, 5, 7, \dots, 15$ . Los siguientes tres renglones corresponden al segundo retardo significativo del modelo, el 10, por lo que en este caso se está probando el retardo 10 contra los restantes retardos denominados  $1, \dots, 9, 11, \dots, 15$ , para ambos tamaños de muestra. Se puede observar que cuando se prueba cada retardo verdadero (esto es, 6 y 10) contra los otros retardos potenciales, la hipótesis nula se rechaza, y no se rechaza cuando se prueba el retardo 6 contra el retardo 10, y viceversa, ya que ambos son igualmente relevantes.

En la Tabla 5.13 se muestra la proporción de veces para la cual el procedimiento seleccionó (primera columna)  $d = 1$ ,  $i^* = 6$  o  $10$ , (segunda columna)  $d = 2$ ,  $i^* = 6$  y  $10$ , y (tercera columna)  $i^* \neq 6$  y  $10$ .

El hecho más sobresaliente es la capacidad del procedimiento propuesto para seleccionar correctamente retardos evitando el riesgo de introducir retardos falsos. Además, esto se hace mediante la selección del correcto (óptimo) número de retardos.

Tabla 5.13: Proporción de veces que algunas posibles estructuras de retardos son seleccionadas.

T	Sólo retardo	Sólo retardos	Ni el retardo 6, ni
	6 o 10	6 y 10	el retardo 10
1000	1	0.891	0
500	1	0.859	0
200	0.851	0.612	0.149

Dada la naturaleza no paramétrica de la técnica, los resultados para tamaños de muestra pequeños son menos notables. En el Apéndice A de este Capítulo, Sección A.4.1, Tabla A.9, se muestran los resultados correspondientes a los tamaños de muestra pequeños,  $T = 70, 100$  y  $200$ .

## 5.4. Resultados Numéricos de la Aplicación de las Técnicas Propuestas a los Residuos Estimados de un Modelo Ajustado.

Las Tablas 5.14, 5.15 y 5.16 contienen los resultados del estadístico del tamaño de la prueba obtenidos en la simulación Monte Carlo y Bootstrap de las series de residuos cuadrados obtenidos a partir del ajuste de los datos generados con los cuatro modelos ARMA-GARCH y los filtros de ARMA(p, q), presentados en la Tabla 4.3, del Capítulo 4.

La simulación se aplicó considerando los parámetros de inmersión:  $m = 3$  y  $\tau = 1, 2, 3, 4, 5, 6$ , para un nivel de significancia de  $\alpha = 0.05$ . Los tamaños de muestra considerados fueron  $T_1 = 500$ ,  $T_2 = 1000$  y  $T_3 = 2000$ , y, el número de muestras bootstrap en cada caso fueron  $B = 199$ . El número de replicas de la simulación fue de  $Nr = 1000$ . Estas tablas indican la proporción de rechazo bajo la hipótesis nula, y deben ser leídas renglón por renglón. El primer renglón muestra la proporción de rechazo cuando la prueba considera el retardo 1 ( $\tau_o = 1$ ) contra los retardos 2, 3, ..., 6. La interpretación del otro renglón tiene que hacerse de la misma manera.

Los resultados para el modelo ARMA(1,1)-N(0,1) son los valores esperados para los que se aplicó la técnica de identificación, estos valores se encuentran dentro de un rango [0.036, 0.066], para un  $\alpha = 0.05$ . Nótese que este buen comportamiento empírico para el tamaño es independiente del tamaño de la muestra.

Tabla 5.14: Tamaño de la prueba de la serie residual con el modelo ARMA(1,1)-N(0,1),

$T_1 = 500$

		$\tau$					
$\tau_o$	1	2	3	4	5	6	
1		0.051	0.051	0.065	0.040	0.051	
2	0.049		0.050	0.040	0.042	0.058	
3	0.052	0.050		0.043	0.038	0.047	
4	0.047	0.062	0.04		0.050	0.053	
5	0.046	0.055	0.050	0.046		0.060	
6	0.053	0.058	0.047	0.050	0.051		

Tabla 5.15: Tamaño de la prueba de la serie residual con el modelo ARMA(1,1)-N(0,1).

$T_2 = 1000$

		$\tau$					
$\tau_o$	1	2	3	4	5	6	
1		0.059	0.042	0.040	0.042	0.045	
2	0.047		0.059	0.048	0.053	0.055	
3	0.036	0.063		0.050	0.054	0.044	
4	0.042	0.055	0.063		0.370	0.061	
5	0.042	0.056	0.052	0.037		0.046	
6	0.054	0.054	0.037	0.041	0.045		

Tabla 5.16: Tamaño de la prueba de la serie residual con el modelo ARMA(1,1)-N(0,1).

$T_3 = 2000$

		$\tau$					
$\tau_o$	1	2	3	4	5	6	
1		0.048	0.053	0.038	0.047	0.039	
2	0.039		0.043	0.042	0.056	0.036	
3	0.053	0.043		0.047	0.042	0.057	
4	0.038	0.058	0.053		0.045	0.055	
5	0.046	0.055	0.054	0.066		0.061	
6	0.042	0.056	0.051	0.037	0.052		

En las Tablas 5.17, 5.18 y 5.19 se presentan los resultados de los experimentos realizados para verificar la potencia empírica de la prueba para la serie de residuales al cuadrado, resultante de filtrar con un ARMA(1,0), ARMA(0,1) y ARMA(1,1) los 3 modelos ARMA-GARCH de la Tabla 4.3 presentada en la Sección 4.3 del Capítulo 4, para los tamaños de muestra  $T_1 = 500$ ,  $T_2 = 1000$ ,  $T_3 = 2000$ . El retardo significativo en los tres modelos es  $\tau_o = 1$ .

Se puede ver en estos resultados que:

- 1) El comportamiento de la prueba para los tres modelos considerados es básicamente el mismo sin importar el tamaño de la muestra.
- 2) Aumenta la potencia con el tamaño de la muestra.
- 3) Según las proporciones de rechazo, se recomienda utilizar la prueba en los residuales, para tamaños de muestra que contengan al menos 1000 observaciones.

Tabla 5.17: Potencia de la prueba en la serie de residuales del modelo:ARMA(1,0)-GARCH(1,1).

	$\tau$					
	1	2	3	4	5	6
$T_1 = 500$	0.358	0.424	0.539	0.593	0.634	
$T_2 = 1000$	0.550	0.697	0.820	0.892	0.891	
$T_3 = 2000$	0.772	0.906	0.967	0.987	0.990	

Tabla 5.18: Potencia de la prueba en la serie de residuales del modelo:ARMA(0,1)-GARCH(1,1).

	$\tau$					
	1	2	3	4	5	6
$T_1 = 500$	0.340	0.401	0.534	0.586	0.640	
$T_2 = 1000$	0.549	0.682	0.812	0.863	0.885	
$T_3 = 2000$	0.807	0.909	0.976	0.988	0.992	

Tabla 5.19: Potencia de la prueba en la serie de residuales del modelo:ARMA(1,1)-GARCH(1,1).

	$\tau$					
	1	2	3	4	5	6
$T_1 = 500$	0.358	0.433	0.536	0.606	0.642	
$T_2 = 1000$	0.563	0.700	0.812	0.858	0.902	
$T_3 = 2000$	0.772	0.908	0.964	0.987	0.989	

En el Apéndice A de este Capítulo, en la Sección A.5, Tablas A.10 a A.15, se presentan los resultados correspondientes a los tamaños de muestra pequeños,  $T = 50, 100$  y  $200$ , tanto del tamaño como de la potencia de la prueba.

## 5.5. Otros Resultados de las Simulaciones

Es importante señalar que durante la ejecución de las simulaciones, se pudo comprobar que:

- Uno de los parámetros más sensibles en el cálculo de la estimación de la integral de correlación es el del radio de la vecindad,  $\varepsilon$  que se utiliza para verificar la cercanía entre las parejas de vectores de retardos en el espacio  $\mathbb{R}^m$ , principalmente cuando las series analizadas son no estacionarias.
- En el caso de las series de residuales, la estimación del modelo que se utilice para filtrar la parte lineal de la serie debe tener una muy buena aproximación para que los residuales al cuadrado estimados no difieran de manera muy significativa de los residuales al cuadrado originales, ya que con una mala aproximación la técnica no identifica el retardo óptimo de forma correcta.
- El cálculo de las estimaciones de la integral de correlación,  $\hat{C}_{m,n}^\tau$ , dentro de las  $N_r$  repeticiones de las simulaciones Monte Carlo y del Bootstrap es en extremo computacionalmente intensivo, cuando el tamaño de la serie es mayor a 200 observaciones. Por ejemplo, el tiempo de CPU necesario para ejecutar la simulación para  $N_r = 1000$  réplicas y  $B = 199$  muestras bootstrap para la validación del comportamiento de la técnica de uno de los modelos con un tamaño de serie de  $T = 1000$  es de 117210 segundos (aproximadamente 33 horas). Para ejecutar la simulación del Modelo AR(10), con los retardos significativos 6 y 10, y considerando 15 retardos temporales distintos, y un tamaño de serie de 500 se requiere un tiempo de CPU de 250178 segundos (aproximadamente 70) horas, ver Tabla 5.20.



Tabla 5.20: Tiempo de CPU(segundos) consumido en las simulaciones de los procedimientos

Procedimiento	Tamaño de la Serie						
	50	70	100	200	500	1000	2000
Promedio $\hat{C}_{m,n}^\tau$	16.72		25.41	49.61	156.47	834.51	1808.64
Modelos lineales y no lineales							
Promedio $\hat{C}_{m,n}^\tau$		56.63	70.08	131.16	373.45	2046.12	4489.12
Modelo AR(10)							
Tamaño y Potencia de la prueba	4338.8		7559.6	30378	44923	117210	435120
Modelos lineales y no lineales							
Tamaño y Potencia de la prueba		13079	17895	76386	250178.31		
Modelo AR(10)							
Tamaño y Potencia de la prueba	4300.6	6939.2	14914	98798	256960		
Serie Residuales al cuadrado							

---

---

## CAPÍTULO 6

---

# APLICACIÓN DE LA TÉCNICA PROPUESTA A SERIES DE TIEMPO NO ESTACIONARIAS

En este capítulo se presenta en forma breve la discusión y análisis que se llevó a cabo en esta investigación con respecto a la posibilidad de que la técnica propuesta para identificar el retardo significativo, pudiera ser aplicada con éxito en una serie escalar generada por un proceso que fuera no estacionario.

Como puede verse en la Sección 2.2.1 del Capítulo 2, en la definición y método de cálculo del estimador de la integral de correlación dados por [G-P, 1983], no se menciona ninguna restricción en cuanto a que los vectores de retardo deben ser obtenidos a partir de series de tiempo deterministas o bien aleatorias, que hayan sido generadas solamente mediante procesos estacionarios. De hecho la expresión matemática (3.2), del Capítulo 3, que corresponde al cálculo de la estimación de la integral de correlación,  $\hat{C}_{m,n}^{\tau}(\epsilon)$ , se refiere únicamente al conteo de las parejas de vectores retardados cuyas distancias dadas por la norma infinito, es menor que,  $\epsilon$ .

Por tal motivo, se podría plantear la hipótesis de que la técnica propuesta, basada en el cálculo del

estimador de la integral de correlación, también pudiera aplicarse a series de tiempo no estacionarias.

Las herramientas que se tienen para el análisis de series de tiempo lineales, requieren para su uso, que las series de tiempo en estudio sean estacionarias, y las herramientas aplicadas en el análisis de series de tiempo no lineales, deterministas, o no, insisten en que el operador de evolución que genera la serie, no cambie con el tiempo, esto es, si  $x_m = x_n \Rightarrow \Phi(x_m) = \Phi(x_n)$ , aún si  $m \neq n$ , lo que significaría que la dinámica del sistema que genera la serie es estacionaria. Considerando esta situación, en la Sección 6.1 se hace un recordatorio breve de lo que es un proceso lineal no estacionario y las características que tienen las series de tiempo generadas por ellos. En la Sección 6.2 se hace una pequeña introducción de lo que son los procesos no lineales con tendencia con el fin de entender por qué no es posible que se pueda obtener numéricamente la estimación de la integral de correlación para este tipo de procesos.

En la Sección 6.3 se discute la importancia que tiene la propiedad de estacionariedad de una serie de tiempo para llevar a cabo la reconstrucción del espacio fase de manera adecuada. Para ello, se presentan los resultados empíricos de la aplicación de la técnica propuesta a series de tiempo generadas por tres procesos lineales no estacionarios, tres caminatas aleatorias de diferente orden, que permiten ver el comportamiento de la técnica para estos casos.

## **6.1. Procesos Lineales No Estacionarios**

La gran mayoría de las series de tiempo observadas tienen una tendencia, ya que sus valores medios varían a lo largo del tiempo. Estas series denominadas series de tiempo no estacionarias, cuando son lineales, tienen un comportamiento parecido al que tiene un proceso de caminata aleatoria. Como ejemplo, en la siguiente figura se presentan las gráficas de algunas series de tiempo que corresponden a algunos de los Indicadores Económicos de Coyuntura de México.

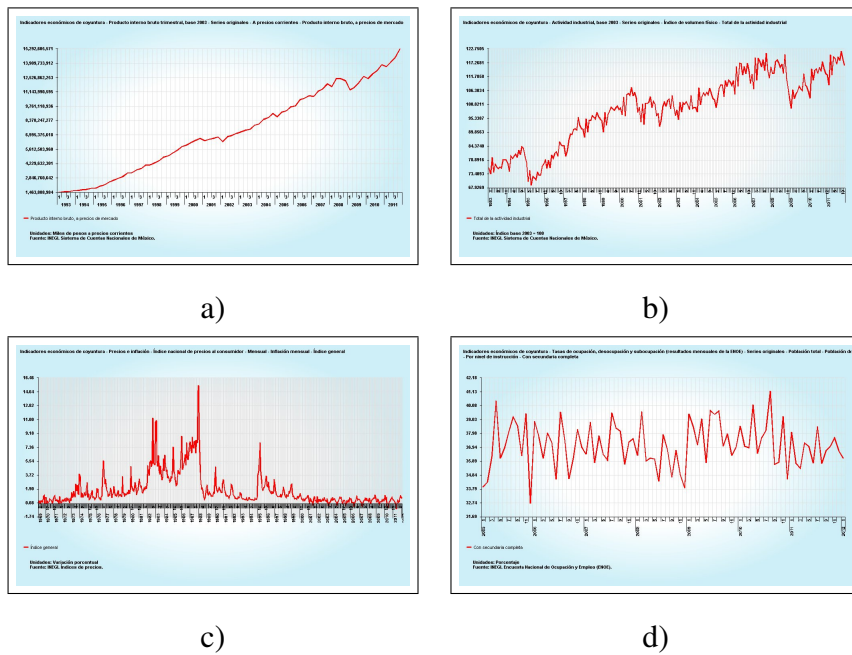


Figura 6.1: Indicadores Económicos Coyunturales de México<sup>3</sup>: a) PIB (precios corrientes), b) índice actividad industrial, base 2003, c) tasa de desempleo por nivel de instrucción, secundaria completa, d) inflación mensual.

Las dos primeras gráficas de la figura 6.1, a) y b) muestran series de tiempo con una tendencia, es decir, parecen ser series no estacionarias. Las dos siguientes c) y d) que representan la tasa de desempleo y la inflación respectivamente, no muestran una tendencia y parecen ser series estacionarias, sin embargo, un proceso generador de series puede ser no estacionario en la media, en la varianza, en las autocorrelaciones, o en otras características de la distribución de las variables.

Cuando el nivel de la serie no es estable en el tiempo, pudiendo tener tendencia creciente o decreciente, como sería el caso de las series en a) y b), se dice que la serie es no estacionaria en la media. Cuando la variabilidad o las autocorrelaciones se modifican con el tiempo, se dice que la serie no es estacionaria en la varianza o autocorrelaciones, como sería el caso de las series en c) y d), y por último si la distribución de la variable en cada instante varía con el tiempo, se dice que la serie es no estacionaria en la distribución.

<sup>3</sup><http://www.inegi.org.mx/sistemas/bie/>

Los procesos lineales MA(q) finitos son siempre estacionarios, y los procesos AR(p) sólo lo son si las raíces de la ecuación característica del proceso, están fuera del círculo unidad.

Considérese el AR(1),

$$X_t = a + \alpha X_{t-1} + e_t, \quad (6.1)$$

si,  $|\alpha| < 1$  el proceso es estacionario, pero si  $|\alpha| > 1$  se puede comprobar que se obtiene un proceso explosivo, donde los valores de la variable crecen sin límite hacia el infinito. Un caso interesante es cuando  $|\alpha| = 1$ , ya que se tiene el modelo  $X_t = X_{t-1} + e_t$  en donde  $e_t \sim N(0, \sigma^2)$ . Este proceso es no estacionario, pero no explosivo, y recibe el nombre de proceso de Caminata (Paseo) Aleatorio, ver Sección 2.3.5 del Capítulo 2. El proceso es no estacionario debido a que su varianza se incrementa a medida que aumenta el tamaño de la muestra.

Se tiene que dado  $X_0$ , valor inicial de la variable aleatoria  $X$ ,  $X_1 = X_0 + e_1$ ;  $X_2 = X_1 + e_2 = X_0 + e_1 + e_2$ ;  $X_3 = X_2 + e_3 = X_0 + e_1 + e_2 + e_3$  y así sucesivamente. De esta forma en el tiempo  $t$ ,  $X$  es la suma de todos los valores pasados y actuales del término error, que hacen que la varianza de  $X_t$  se haga más grande a medida que se incrementa el tamaño de la muestra.

Si se considera también el proceso de una caminata aleatoria pero ahora de orden 2, se tiene que si  $X_t = X_{t-2}$ ; con  $X_0, X_1$ , valores iniciales de la serie:

$$\begin{aligned} X_2 &= X_0 + e_2; & X_3 &= X_1 + e_3; \\ X_4 &= X_2 + e_4 = X_0 + e_2 + e_4; & X_5 &= X_3 + e_5 = X_1 + e_3 + e_5; \\ X_6 &= X_4 + e_6 = X_0 + e_2 + e_4 + e_6; & X_7 &= X_5 + e_7 = X_1 + e_3 + e_5 + e_7; \\ & & \vdots & \quad \quad \quad \vdots \end{aligned} \quad (6.2)$$

Lo que significa que en el tiempo  $t$  si éste es par,  $X$  es la suma de los valores pasados pares y actuales del término error, y si  $t$  es impar,  $X$  es la suma de todos los valores pasados impares y actuales del término error. Esto hace que la varianza de  $X_t$  se haga más grande pero con valor diferente si se trata de  $t$  par o impar, a medida que se incrementa el tiempo.

Si se considera también el proceso de una caminata aleatoria pero ahora de orden 3, se tiene que si

$X_t = X_{t-3} + e_t$ ; con  $X_0, X_1, X_2$  valores iniciales de la serie.

$$\begin{aligned}
 X_3 &= X_0 + e_3; & X_4 &= X_1 + e_4; & X_5 &= X_2 + e_5; \\
 X_6 &= X_0 + e_3 + e_6; & X_7 &= X_1 + e_4 + e_7; & X_8 &= X_2 + e_5 + e_8; \\
 X_9 &= X_0 + e_3 + e_6 + e_9; & X_{10} &= X_1 + e_4 + e_7 + e_{10}; & X_{11} &= X_2 + e_5 + e_8 + e_{11} \quad (6.3) \\
 &\vdots & &\vdots & &\vdots
 \end{aligned}$$

y en este caso en el tiempo  $t$  si éste es par o impar,  $X$  es la suma de valores pasados, alternando impares,  $t_r$  y pares  $t_s$  con  $s - r = 3$ , y actuales del término error. Esto hace que la varianza de  $X_t$ , se haga más grande pero con valor diferente si se trata de  $t$  par o impar, a medida que se incrementa el tiempo.

En un proceso no estacionario, a diferencia del estacionario, las constantes, si existen, son muy importantes y representan alguna propiedad permanente en el proceso. Si la constante no existe,  $a = 0$ , se dice que el proceso no tiene deriva y el nivel de la serie oscila en el tiempo. Si existe la constante,  $a \neq 0$ , se dice que el proceso tiene una deriva igual al valor de la constante, y el proceso muestra una tendencia lineal de pendiente  $a$ , lo que supone un crecimiento lineal determinista.

### 6.1.1. Vectores retardados en los procesos de caminata aleatoria

En las figuras 6.2 a 6.4 se presentan las gráficas de los vectores retardados que se obtienen a partir de series de tiempo generadas por procesos de caminata aleatoria de orden uno, dos y tres, en donde la constante  $a = 0$ ,  $\alpha = 1$ , en la expresión (6.1). La construcción de estos vectores se obtiene considerando una dimensión de inmersión  $m = 2$ , y seis diferentes valores de retardos temporales,  $\tau = 1, 2, 3, 4, 5, 6$  entre los cuales se encuentra el retardo significativo del modelo generador,  $\tau_0$ . El tamaño de la serie generada es de  $T = 100$ .

De forma similar a lo que ocurría en las series de tiempo estacionarias, en este caso la estructura de la gráfica,  $X_t$  .vs.  $X_{t-1}$ , correspondiente al retardo significativo del modelo, es diferente a la de los otros retardos, con la peculiaridad de que tanto para el orden 2 como para el 3, la de los

retardos pares (2, 4, 6 para el orden 2) e impares (3, 6 para el orden 3) que le siguen al significativo tienen un mayor parecido, lo que está relacionado a lo expuesto en las expresiones matemáticas (6.2) y (6.3).

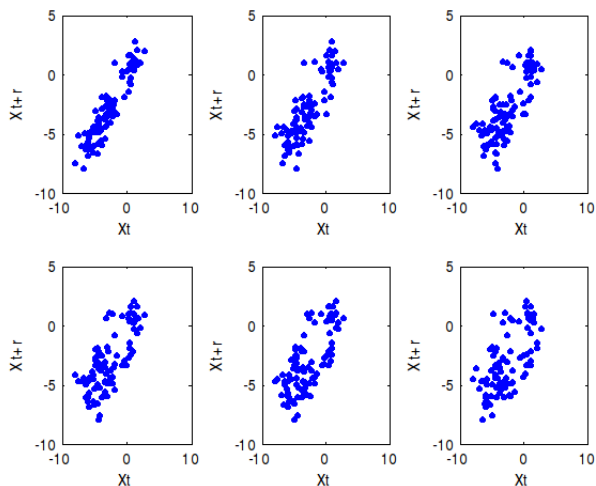


Figura 6.2: Vectores retardados obtenidos de la serie de tiempo generada con el modelo:

$$X_t = X_{t-1} + e_t; X_0 = 1.0, e_t \sim N(0, 1); m = 2, \tau = 1, 2, \dots, 6; \tau_0 = 1$$

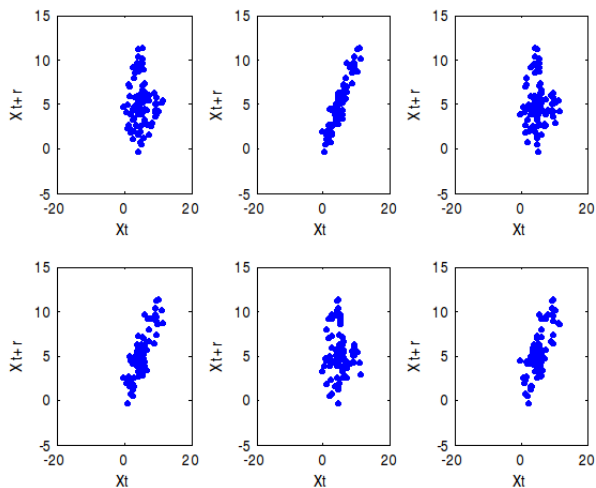


Figura 6.3: Vectores retardados obtenidos de la serie de tiempo generada con el modelo:

$$X_t = X_{t-2} + e_t; X_0 = 1.0, X_1 = 2.266, e_t \sim N(0, 1),$$

$$m = 2, \tau = 1, 2, \dots, 6; \tau_0 = 2$$

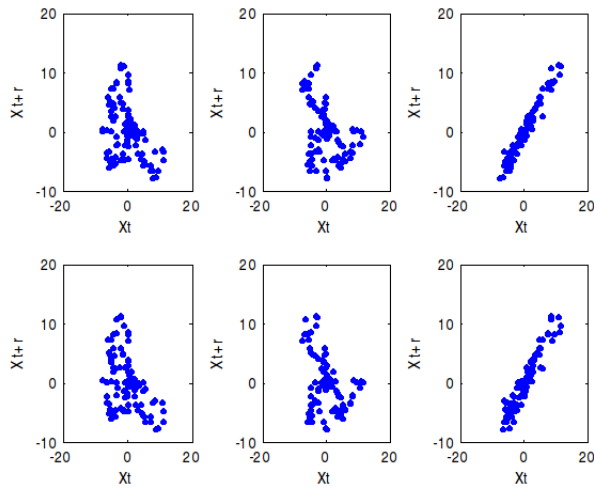


Figura 6.4: Vectores retardados obtenidos de la serie de tiempo generada con el modelo:

$$X_t = X_{t-3} + e_t; \quad X_0 = 1.0, X_1 = 2.266, X_2 = 1.249; \quad e_t \sim N(0, 1), \quad m = 2, \quad \tau = 1, 2, \dots, 6; \quad \tau_0 = 3$$

### 6.1.2. Vectores retardados en otros procesos lineales no estacionarios

Es importante señalar que en algunos procesos no estacionarios lineales, la variable tiempo,  $t$ , puede aparecer de manera explícita, y en este caso, el proceso puede volverse explosivo, esto es, los valores de la variable van creciendo sin límite hacia el infinito. El siguiente modelo es un ejemplo de esto:

$$X_t = 0.5 + t + X_{t-1} + e_t; \quad e_t \sim N(0, 1). \quad (6.4)$$

La figura 6.5 muestra las gráficas de los vectores retardados que se obtienen a partir de la serie de tiempo generada con la ecuación (6.4) considerando de igual forma, una dimensión de inmersión  $m = 2$ , y seis diferentes valores de retardos temporales,  $\tau = 1, 2, 3, 4, 5, 6$ . El retardo significativo en este modelo es  $\tau_0 = 1$ , y el tamaño de la serie generada es de  $T = 100$ . En este caso se puede ver que no se distingue ninguna diferencia entre las 6 gráficas, tomando el intervalo  $(0, 5000)$  en ambos ejes, pero lo que si se puede apreciar es como se va incrementando el valor de la variable conforme crece el tiempo.



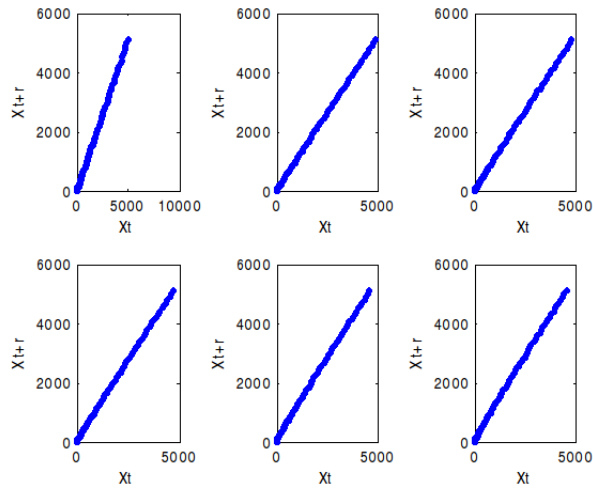


Figura 6.5: Vectores retardados obtenidos de la serie de tiempo generada con el modelo:

$$X_t = 0.5 + t + X_{t-1} + e_t; \quad X_0 = 1.0 \quad e_t \sim N(0, 1), \quad m = 2, \quad \tau = 1, 2, \dots, 6; \quad \tau_0 = 1$$

## 6.2. Procesos No Lineales con Tendencia

Se sabe que los sistemas dinámicos no lineales y no estacionarios son extremadamente difíciles de modelar a partir de una serie de tiempo, a no ser que se tenga un conocimiento a priori del sistema subyacente, ya que en estos casos el número de parámetros puede exceder en mucho al número de observaciones que se tengan.

Un ejemplo de un sistema no estacionario no lineal es aquel que está sujeto a una dependencia no lineal temporal basada en alguna influencia externa. De esta manera si se parte de la ecuación (6.4) y se toma la variable tiempo explícita de forma cuadrática, se tendría el siguiente modelo:

$$X_t = 0.5 + t^2 + X_{t-1} + e_t; \quad e_t \sim N(0, 1). \quad (6.5)$$

Al graficar  $X_t$  .vs.  $X_{t-1}$  se puede ver, figura 6.6, que la estructura que se obtiene para los seis valores de retardos dados, es similar a la presentada en la figura 6.5, a excepción del intervalo en el cual se encuentran los valores de la variable  $X_t$ ,  $(0, 4 \times 10^5)$ .

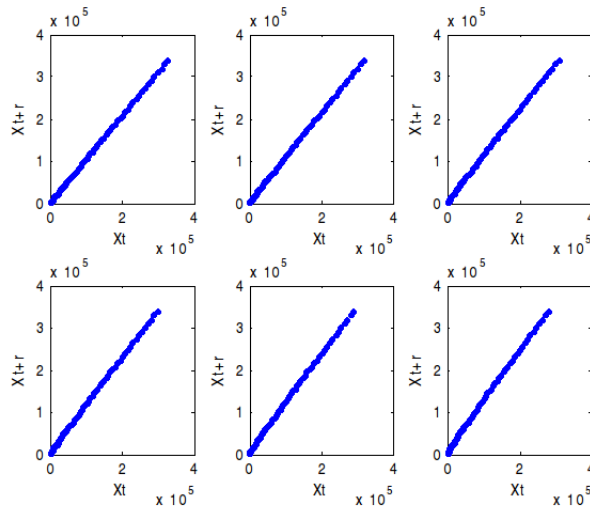


Figura 6.6: Vectores retardados obtenidos de la serie de tiempo generada con el modelo:

$$X_t = 0.5 + t^2 + X_{t-1} + e_t; X_0 = 1.0 \quad e_t \sim N(0, 1); m = 2, \tau = 1, 2, \dots, 6; \tau_0 = 1$$

### 6.3. La Estacionariedad y las Medidas de Reconstrucción.

Cuando se asume estacionariedad en los procesos no lineales, se está suponiendo que la estructura del espacio de fase reconstruido con los vectores de retardo, no presenta cambios bruscos con respecto al tiempo. Esto es lo que no puede verse en los dos ejemplos de series no estacionarias presentadas, (6.4) y (6.5), ya que la variabilidad entre las componentes de los vectores de retardo es grande, sin importar el retardo temporal seleccionado, y la distancia entre dos componentes, de los vectores  $X_i^{m,\tau}, X_j^{m,\tau}$ , también es grande.

Ante esto cabe la pregunta de si el estimador de la integral de correlación está bien definido para procesos no estacionarios, y que en el caso de poder calcularse numéricamente, pudiera aportar la información necesaria para identificar el retardo significativo del modelo que genera la serie de tiempo. Si se toma en cuenta que de forma equivalente, estacionariedad significa que todas las clases de promedios, que cuantifican el “comportamiento estadístico” de la series, están bien definidas, entonces las series no estacionarias, como lo son las series de tiempo financieras, que en

promedio se incrementan a una tasa exponencial, no poseen la propiedad de que un promedio sobre el futuro infinito esté bien definido, al menos no como un número finito. Y, en este caso, dado que el estimador de la integral de correlación, ecuación (3.2) del Capítulo 3, es el doble promedio del número de parejas de vectores retardados  $(X_i^{m,\tau}, X_j^{m,\tau})$  que se encuentran dentro de una vecindad de radio  $\varepsilon$ , tampoco estaría bien definido.

Es importante recordar que toda función no lineal de un proceso estrictamente estacionario sigue siendo estrictamente estacionario, sin embargo no sucede lo mismo con las funciones no lineales de procesos estacionarios débiles. Por ejemplo, el cuadrado de un proceso estacionario débil puede no tener una varianza finita.

Adicionalmente, se sabe que la dinámica de un proceso puede ser interpretada como un mapeo único de una variedad  $M \rightarrow M$ , por lo que una inmersión de los vectores retardados es nuevamente una asignación independiente del tiempo de  $M \rightarrow \mathbb{R}^m$ . Sin embargo, en los procesos no estacionarios existe una dependencia directa con el tiempo.

Considerando lo anterior se esperaría que no se podría obtener de forma numérica el valor del estimador de la integral de correlación en aquellas series de tiempo generadas por procesos no estacionarios explosivos, como por ejemplo los que tienen tendencia, ya que los valores estarían desbordando la aritmética de punto flotante de la computadora, y para las series generadas por procesos no estacionarios no explosivos, se podrían tener algunas peculiaridades al llevar a cabo la construcción de los vectores retardados y obtener con éstos la estimación de las integrales de correlación

## **6.4. Aplicación Empírica de la Técnica Propuesta a Series de Tiempo Lineales No Estacionarias**

. Con el fin de verificar empíricamente la hipótesis de que la técnica propuesta, aún tomando en cuenta las consideraciones presentadas en las secciones anteriores, basada en el cálculo del esti-

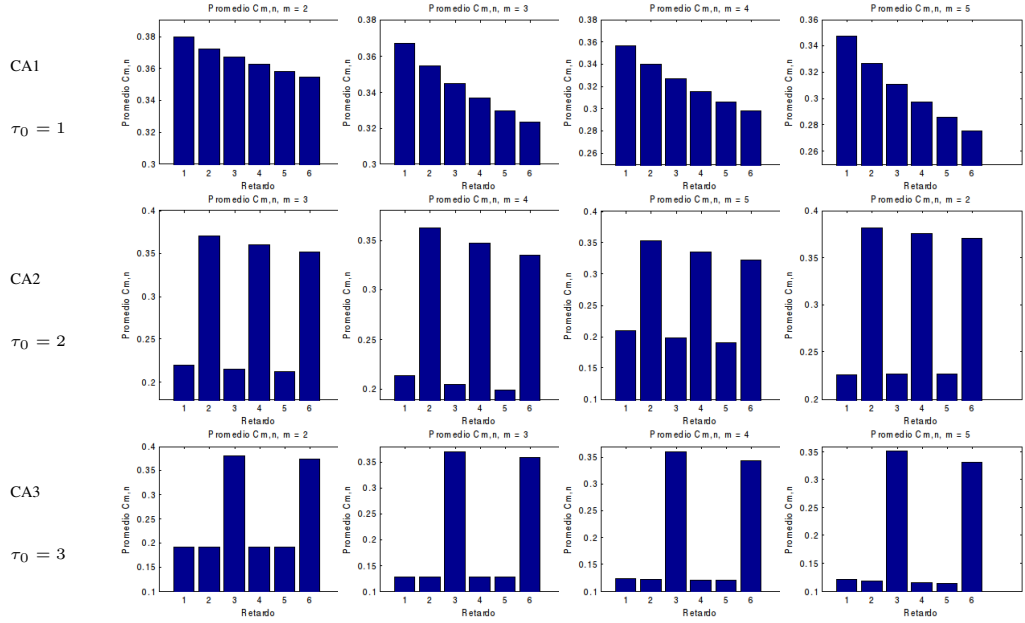
mador de la integral de correlación es adecuada para identificar los retardos significativos de una serie generada a través de procesos lineales no estacionarios y no explosivos, como las Caminatas Aleatorias, se consideró en la investigación la aplicación y análisis del comportamiento de la técnica, a través de los procedimientos computacionales ya mencionados en el Capítulo 4.

En esta sección se presentan los resultados obtenidos, mismos que permiten corroborar lo anteriormente expuesto.

#### **6.4.1. Resultados de la aplicación del Procedimiento 1: Identificación del retardo a partir del valor máximo del promedio de $\hat{C}_{m,n}^{\tau}(\varepsilon)$ .**

La aplicación del procedimiento de identificación del retardo significativo a partir del valor máximo del promedio  $\hat{C}_{m,n}^{\tau}(\varepsilon)$ , como se mencionó en la Sección 4.2.1 del capítulo 4, se llevó a cabo mediante una simulación Monte Carlo de los modelos no estacionarios analizados presentados en la Tabla 4.2 del Capítulo mencionado. Los resultados obtenidos de las simulaciones, considerando un tamaño de muestra de  $T = 500$  y  $N_r = 1000$  réplicas de la simulación para cada modelo y estimando la integral de correlación con los parámetros de inmersión:  $m = 2, 3, 4$  y  $5$ ,  $\tau = 1, 2, 3, 4, 5, 6$  y  $\varepsilon = 1.96\sigma$ , son los presentados en la Tabla 6.1 que se presenta a continuación:

Tabla 6.1: Promedios de  $\hat{C}_{m,n}^\tau$  para procesos no estacionarios,  $T = 500$



Las gráficas de la tabla muestran que el valor máximo del promedio de la integral de correlación identifica precisamente a los retardos significativos de cada uno de los modelos. En las gráficas de CA(2) y CA(3) también se puede ver que el promedio de la integral de correlación presenta el comportamiento peculiar para los respectivos múltiplos del retardo significativo de los modelos,  $\tau = 4, 6$  para CA(2) y  $\tau = 6$  para CA(3), que se mencionó en la Sección 6.1, ecuaciones (6.2) y (6.3), con lo que se puede constatar que para estas caminatas aleatorias los vectores retardados de reconstrucción del espacio de fase correspondiente presentan una estructura de periodo  $t = (m - 1)\tau$ ;  $X_t^{3,2} = \{x_t, x_{t-2}, x_{t-4}\}$ ;  $X_t^{3,3} = \{x_t, x_{t-3}, x_{t-6}\}$ .

También se analizó el comportamiento del procedimiento para tamaños de muestra pequeños de 50, 100 y 200. Los gráficos correspondientes aparecen en las tablas 6.2, 6.3 y 6.4 respectivamente.

Tabla 6.2: Promedios de  $\hat{C}_{m,n}^\tau$  para procesos no estacionarios,  $T_1 = 50$

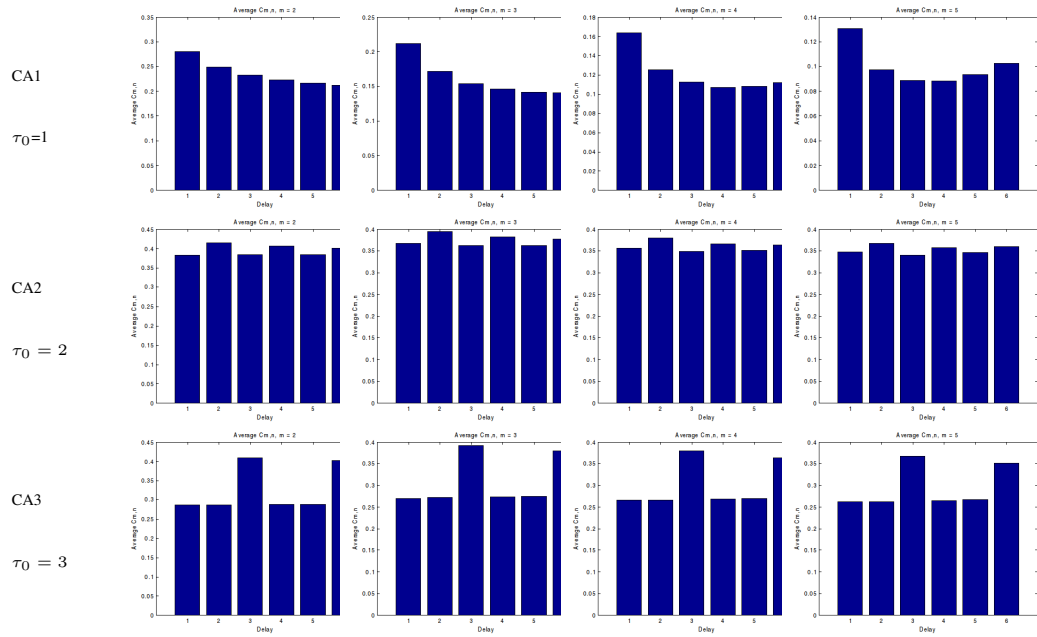


Tabla 6.3: Promedios de  $\hat{C}_{m,n}^\tau$  para procesos no estacionarios,  $T_2 = 100$

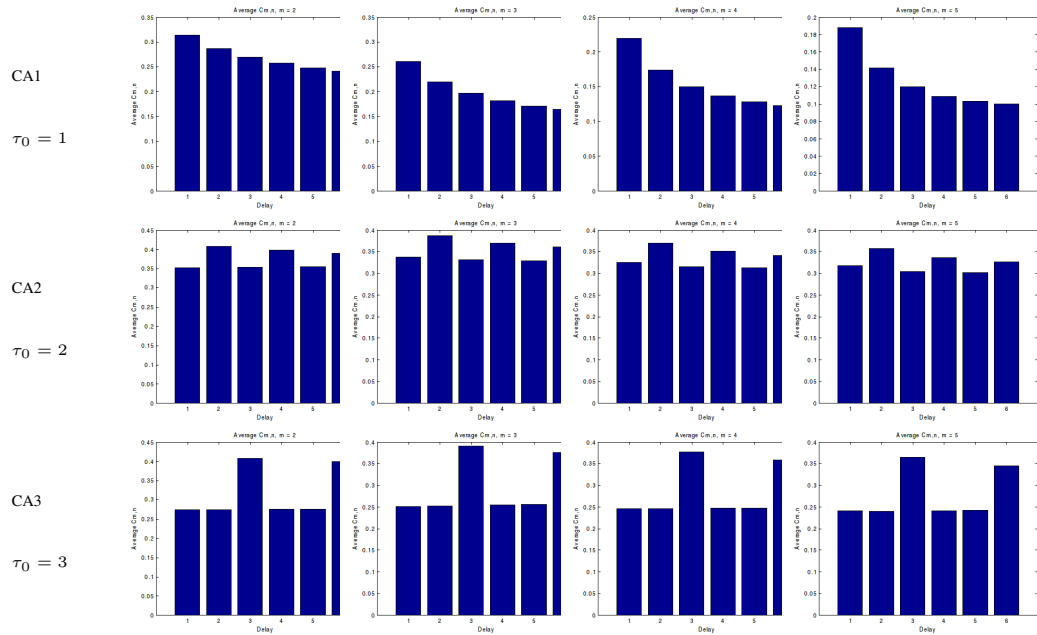
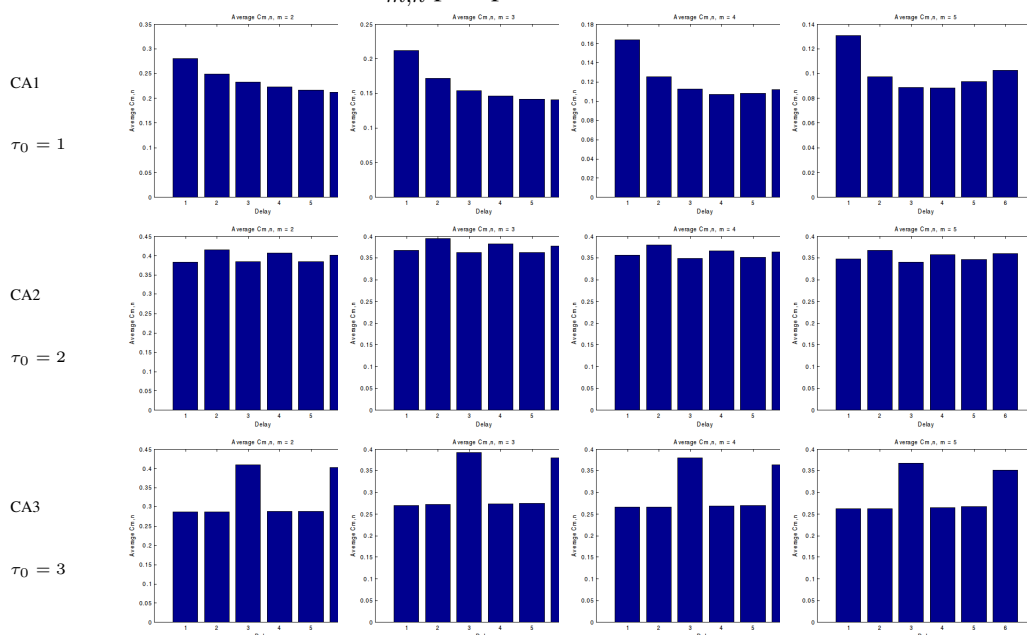


Tabla 6.4: Promedios de  $\hat{C}_{m,n}^\tau$  para procesos no estacionarios,  $T_3 = 200$



Al igual que con las simulaciones de series de tamaño grande, el comportamiento de la técnica es muy similar a lo anteriormente comentado.

### 6.4.2. Resultados de la aplicación del procedimiento 2: Prueba para identificar retardos en modelos no estacionarios lineales y no lineales.

Para evaluar el rendimiento de la prueba con muestras finitas en relación a la potencia de la prueba, se realizaron simulaciones Monte Carlo y bootstrap considerando los parámetros de inmersión:  $m = 3$  y  $\tau = 1, 2, 3, 4, 5, 6$ , y el nivel de significancia de  $\alpha = 0.05$ . Inicialmente se consideraron dos tamaños de muestra grandes  $T_1 = 500$  y  $T_2 = 1000$ , y posteriormente se llevaron a cabo simulaciones para tamaños de muestra pequeños  $T_1 = 50$ ,  $T_2 = 100$ , y  $T_3 = 200$ . El número de muestras bootstrap en cada caso fue  $B = 199$ , y el número de réplicas para la simulación fue de  $Nr = 1000$ . En las Tablas 6.5 a 6.7 se presentan los resultados de los experimentos de la potencia de la prueba con los tres modelos no estacionarios. Estas tablas deben leerse renglón por renglón (modelo por modelo). Cada renglón (modelo) tiene una celda vacía cuya columna es indicativa del

retardo que se está usando para la prueba: por ejemplo, la celda vacía para el modelo CA(2) es la columna 2, lo que significa que, para este modelo en particular, se está probando el retardo 2 contra los restantes retardos denominados, 1, 3, 4, 5 y 6. Esto es porque el retardo relevante para el modelo CA(2) es el 2, y entonces se estima la probabilidad de rechazar la hipótesis nula de que el retardo 2 no tiene mayor información que cualquier otro retardo considerado.

Tabla 6.5: Potencia de la prueba para series no estacionarias generadas de tamaño,  $T_1 = 500$

Modelo	Retardo					
	1	2	3	4	5	6
CA(1)		0.747	0.845	0.907	0.932	0.947
CA(2)	0.991		0.991	0.785	0.989	0.868
CA(3)	0.985	0.987		0.985	0.981	0.766

Tabla 6.6: Potencia de la prueba para series no estacionarias generadas de tamaño  $T_2 = 1000$

Modelo	Retardo					
	1	2	3	4	5	6
CA(1)		0.791	0.892	0.943	0.958	0.965
CA(2)	0.999	0	0.999	0.815	0.998	0.886
CA(3)	0.999	0.999		0.999	0.999	0.765

La tabla 6.7 con sus incisos a) a c), presenta los resultados de los experimentos de la potencia de la prueba para tamaños de muestra pequeños, 50, 100 y 200. Al igual que con los de tamaño grande, la tabla debe leerse renglón por renglón (modelo por modelo).

Comparando estos resultados con los que se obtuvieron para los tamaños de muestra grandes, ver tablas 6.5 y 6.6, se puede concluir que para que el rendimiento de la prueba, en cuanto a la potencia, sea bueno los tamaños de muestra que se recomiendan utilizar deben ser mayores de 200.



Tabla 6.7: Potencia de la prueba para muestras de tamaño pequeño modelos no estacionarios.

a)  $T_1 = 50$

Modelo	Retardo					
	1	2	3	4	5	6
CA(1)		0.491	0.489	0.482	0.463	0.454
CA(2)	0.533		0.568	0.464	0.519	0.472
CA(3)	0.557	0.613		0.528	0.464	0.316

b)  $T_2 = 100$

Modelo	Retardo					
	1	2	3	4	5	6
CA(1)		0.568	0.589	0.615	0.622	0.624
CA(2)	0.76		0.771	0.594	0.724	0.622
CA(3)	0.749	0.770		0.743	0.720	0.553

c)  $T_3 = 200$

Modelo	Retardo					
	1	2	3	4	5	6
CA(1)		0.645	0.712	0.771	0.799	0.81
CA(2)	0.911		0.911	0.694	0.902	0.767
CA(3)	0.913	0.918		0.903	0.894	0.663

## 6.5. Futuras aplicaciones de la técnica propuesta a series de tiempo no estacionarias

En la metodología clásica de análisis de series de tiempo se enfatiza el hecho de que para llevar a cabo el estudio de una serie de tiempo no estacionaria, previamente debe transformarse esta en una estacionaria para así poder aplicarle los mismos métodos de análisis. En algunos casos se tiene una idea bastante clara sobre el tipo de no estacionariedad que presenta la serie de tiempo, y este conocimiento puede ser utilizado para definir la transformación adecuada. Por ejemplo una tendencia lineal en la serie temporal no estacionaria como la que se presenta en la ecuación 6.4 puede ser removida fácilmente utilizando un procedimiento de ajuste de mínimos cuadrados. Para el caso de la ecuación 6.5 que presenta una tendencia cuadrática, sería mejor ajustar la tendencia con un polinomio o una función no lineal adecuada, para lo cual se tendría que hacer un ajuste de regresión no lineal. En el estudio de series de tiempo económicas tales como los precios diarios de acciones es común aplicar métodos de transformación que consisten en diferenciar la serie del logaritmo de los precios, u obtener la serie de rendimientos diarios.

Un trabajo de investigación que podría llevarse a cabo en el futuro, sería aplicar la técnica propuesta a series temporales resultantes de la transformación de series generadas con procesos estocástico no estacionarios como los pesentados en este capítulo mediante simulaciones Monte Carlo, utilizando algunos de los métodos de transformación aquí mencionados. Para este caso particular es necesario considerar lo siguiente:

- Que el modelo generador de la serie no estacionaria asegure la evaluación numérica para toda  $t$ , con  $t = 1, 2, \dots, T$ ,  $T$  tamaño de la muestra.
- Que la aproximación del ajuste para la transformación de la serie sea lo suficientemente buena para evitar contaminar los resultados por errores de aproximación o redondeo de la aritmética de punto flotante.
- Que el tamaño de la muestra sea mayor a 500 observaciones

---

---

## CAPÍTULO 7

---

### CONCLUSIONES Y CONSIDERACIONES FINALES

En este trabajo de investigación se obtuvieron resultados que contribuyen, de manera relevante, a la especificación adecuada de un modelo general lineal o no lineal de series de tiempo estacionarias. Estos resultados corresponden al desarrollo de un procedimiento y una nueva prueba estadística no paramétrica que permite seleccionar los retardos temporales significativos de un modelo generador de series.

Dado que de manera tradicional, el primer paso crucial para seleccionar los retardos significativos de un modelo, y así definirlo correctamente, se realiza a través de los coeficientes de autocorrelación y autocorrelación parcial, que como bien señala [Granger, 1983] es un procedimiento incorrecto; este nuevo procedimiento y prueba propuestos, pueden ser utilizados como una buena herramienta de diagnóstico.

En general los procedimientos basados en las autocorrelaciones, fallan cuando se tienen modelos no lineales, al no detectar apropiadamente los retardos temporales, especialmente en aquellos escenarios donde los fenómenos no lineales son más bien la regla que la excepción. Es ampliamente conocido que la mayoría de las variables económicas son de naturaleza no lineal y sus relaciones

no lineales son hechos comúnmente aceptados.

Desde un punto de vista econométrico y estadístico, esta situación ha impulsado el desarrollo de pruebas para verificar la independencia serial, sin embargo, las pruebas que se tienen actualmente, si bien son de naturaleza no paramétrica, por construcción no están diseñadas para seleccionar los retardos temporales relevantes. En este sentido, la propuesta de esta nueva técnica es de importancia, ya que basada en el cálculo de la bien conocida Integral de Correlación, es posible identificar que retardo temporal debe ser utilizado al construir modelos de naturaleza desconocida, lineales o no lineales. Adicionalmente el procedimiento propuesto, es de fácil aplicación, ya que para decirlo de forma resumida, consiste en seguir los pasos del siguiente algoritmo:

- 1) Proporcionar como datos iniciales, la serie de tiempo observada y los parámetros de inmersión  $m$  (dimensión de inmersión) y  $\varepsilon$  (distancia entre parejas de vectores en  $\mathbb{R}^m$ );
- 2) Generar el conjunto de vectores retardados (las  $m$ -historias de la serie), necesarios para el cálculo de la estimación de la integral de correlación considerando varios valores para el tercer parámetro de inmersión,  $\tau \in \mathbb{Z}^+$ ;
- 3) Calcular las integrales de correlación utilizando la ecuación 3.2 del Capítulo 3;
- 3) Comparar los valores del estimador de la integral de correlación  $\hat{C}_{m,n}^\tau(\varepsilon)$ , para identificar el retardo temporal en el cual se alcanzó el valor máximo, ya que éste, como se mostró en el desarrollo teórico y empírico de este trabajo, es precisamente el retardo temporal óptimo que se está buscando.

En las siguientes secciones se señalan las conclusiones a las que se ha llegado con este trabajo de investigación en sus diferentes etapas, específicamente relacionadas con la aplicación de los procedimientos propuestos y de los resultados relevantes que se obtuvieron.

## 7.1. Identificación Correcta del Retardo Temporal Significativo

Después de estudiar el comportamiento de la técnica propuesta, aplicándola a diez modelos de tipo lineal y no lineal, estacionarios (ver Tabla 4.1 del Capítulo 4), un modelo lineal autorregresivo AR(10) con dos retardos significativos, 6 y 10 (ver Modelo 11 en el Capítulo 4, Sección 4.2.3, y tres modelos lineales no estacionarios y no explosivos, (ver Tabla 4.2 del Capítulo 4), se puede concluir lo siguiente:

- La técnica identifica correctamente el retardo temporal en todas las series generadas por los 14 modelos mencionados y el comportamiento que mostraron las diferencias para cada retardo, no dependió de la dimensión de inmersión,  $m$ .
- Las gráficas del promedio de  $\hat{C}_{m,n}^{\tau}(\varepsilon)$  para el modelo  $N(0, 1)$  no muestran, como se esperaba, diferencias entre los valores promedios para  $\tau = 1, 2, 3, 4, 5, 6$ , y por tanto no existe un retardo relevante. Ver gráficas en las Tablas 5.1 a 5.5 y 5.11 del Capítulo 5, Tablas A.1 a A.3 y A.8 del Apéndice A (Apéndice al Capítulo 5 para tamaños de muestra pequeños), y Tablas 6.1 a 6.4 del Capítulo 6.
- En presencia de estructura de varianza condicional, el procedimiento también detecta los retardos correctos correspondientes. El ARCH(2) depende solamente del retardo 2, y el ARCH(3) incorpora varios retardos relevantes con diferente fuerza en los valores de los parámetros (0.1 para los retardos 1 y 2, y 0.3 para el retardo 3). Para el modelo GARCH la dependencia identificada es como era de esperarse, con el retardo 1. Ver Tablas 5.1 a 5.5 del Capítulo 5 y Tablas A.1 a A.3 del Apéndice A (Apéndice al Capítulo 5 para tamaños de muestra pequeños).

## 7.2. Identificación Correcta del Retardo Temporal Significativo con Distintos Valores del Parámetro $\varepsilon$ .

Los resultados obtenidos en cuanto a la prueba de sensibilidad de la técnica con respecto al radio  $\varepsilon$ , parámetro de inmersión, permiten concluir lo siguiente:

- La técnica identifica correctamente el retardo temporal en las cuatro series de tiempo que se utilizaron para llevar a cabo la prueba de sensibilidad:  $N(0,1)$ ,  $AR(2)$ ,  $NMA(3)$  y  $ARCH(2)$ , y el comportamiento que mostraron las diferencias para cada retardo, no dependió del valor de  $\varepsilon$  dado. Ver Tablas 5.2 a 5.5 del Capítulo 5.
- Las gráficas en las Tablas arriba mencionadas muestran que el promedio de  $\hat{C}_{m,n}^{\tau}(\varepsilon)$  para el modelo  $N(0,1)$  no muestran, como se esperaba, diferencias entre los valores promedios para  $\tau = 1, 2, 3, 4, 5, 6$ , y los valores de  $\varepsilon = a\sigma$ ;  $a = 0.5, 1.0, 1.5, 2.0$ , por tanto no existe un retardo relevante.

## 7.3. Tamaño y Potencia de la Prueba Estadística Propuesta

A partir de los resultados obtenidos con la realización de simulaciones Monte Carlo y Bootstrap a los modelo  $N(0,1)$ ,  $MA(2)$  y  $NMA(3)$  para obtener el tamaño de la prueba estadística y a los modelos 2 a 10 de la Tabla 4.1 para obtener la potencia, ver Secciones 4.2.2 del Capítulo 4 y 5.3.1 y 5.3.2 del Capítulo 5, se puede concluir que:

- Para el modelo  $N(0, 1)$  todos los retardos contienen la misma información (son todos igualmente no relevantes), y por tanto se está bajo la hipótesis nula. Las Tablas 5.6 y A.4, del Capítulo 5 y del Apéndice A respectivamente, muestran que se está en el nivel nominal de rechazo esperado;

- Al probar las diferentes configuraciones posibles de la hipótesis nula, a los modelos MA(2) y NMA(3), ver Tablas 5.7 y 5.8, Capítulo 5 para tamaños de muestra grandes y las Tablas A.5 y A.6 del Apéndice A para tamaños de muestra pequeña, los resultados muestran que para el nivel nominal seleccionado de la prueba, su desempeño bajo la hipótesis nula es correcta, considerando la sensibilidad a las propiedades de memoria que tienen los dos modelos usados. Esta conclusión es especialmente importante ya que otras pruebas basadas en la integral de correlación, como el estadístico BDS, tienen graves distorsiones de tamaño para muestras finitas [Kanzler,1999].
- El rendimiento de la potencia de la prueba es extremadamente bueno, a pesar del tamaño de la muestra, para los modelos lineales (MA y AR), al igual que para el proceso autorregresivo no lineal.
- Para el modelo bilinial, que exhibe una dependencia a través de los dos primeros retardos, las funciones de autocorrelación y correlación parcial teóricamente son la mismas a la del ruido blanco y por tanto si se usarán los métodos clásicos basados en las correlaciones se obtendrían conclusiones erróneas. Algo similar ocurre con el mapeo logístico, con el problema adicional de ser puramente determinista.
- Como se menciona en [Granger and Lin, 1994], la prueba de Kendal tau se desempeña pobremente para procesos NAR. Para el modelo de promedios móviles no lineal generado con un retardo de orden 3, la potencia de la prueba propuesta es relativamente buena para tamaños de muestra grande, mientras que los resultados son más discretos cuando se consideran 500 observaciones.
- Por lo que respecta al comportamiento empírico de la prueba, en términos de potencia, para los modelos condicionales heteroscedásticos es ciertamente de alta calidad.
- Se puede señalar que la nueva prueba propuesta es especialmente útil en situaciones de incertidumbre acerca de la naturaleza lineal o no lineal de la media condicional del proceso,

puesto que como procesos lineal y no lineal son claramente detectados al considerar integrales de correlación en diferentes retardos.

- Para el caso de muestras de tamaño pequeño, la principal conclusión es que, dada la naturaleza no paramétrica, la potencia es buena únicamente para los modelos lineales y caóticos cuando los tamaños de muestra están por debajo de 100 observaciones. Cuando hay incertidumbre acerca de la naturaleza (lineal o no lineal) del proceso, es necesario utilizar tamaños de muestra superiores a 200 observaciones.

## **7.4. Aplicación Automática del Procedimiento de Identificación del Retardo Significativo**

Los resultados obtenidos en la aplicación que se hizo de la técnica propuesta al modelo lineal autorregresivo AR(10), con dos retardos significativos, 6 y 10 (ver Modelo 11 en el Capítulo 4, Sección 4.2.3), considerando 15 retardos posibles, del 1 al 15, permiten corroborar como ya se dijo en la Sección 7.1 que, la técnica identifica correctamente los dos retardos significativos, así como los resultados del tamaño y potencia de la prueba permiten concluir que, la técnica propuesta tiene una gran capacidad para seleccionar correctamente los retardos, evitando el riesgo de introducir retardos falsos, considerando series de tamaño mayor a 200 observaciones (ver Tablas 5.12 y 5.13, Capítulo 5, y Tabla A.9 del Apéndice al Capítulo 5).

## **7.5. Uso de la Prueba Propuesta como Herramienta de Diagnóstico.**

Las simulaciones Monte Carlo y Bootstrap hechas a las series de residuos al cuadrado obtenidas después de aplicarles filtros ARMA(p,q) a cuatro modelos ARMA-GARCH (ver Tabla 4.3 del



Capítulo 4, permitió verificar la capacidad de la técnica propuesta en la identificación de la estructura de desfase dinámico en modelos que presentan algún tipo de dependencia (lineal o no lineal) en su variabilidad condicionada. De esta manera se puede concluir que si existe una dependencia lineal o no lineal en los residuos, la prueba es capaz de identificar el retardo que debe ser utilizado en el modelo seleccionado, por lo que se considera que la técnica sería de gran utilidad a los investigadores interesados en el modelado financiero.

Los resultados obtenidos durante la validación (ver Tablas 5.14 a 5.19 del Capítulo 5 y Tablas A.10 a A.15 del Apéndice A) permiten además establecer las siguientes dos observaciones:

- 1) El tamaño y potencia de la prueba aumenta con el tamaño de la muestra, por lo que se sugiere utilizar series de residuos de al menos 1000 observaciones.
- 2) La estimación del ajuste que se le haga al modelo para filtrar la parte lineal de la serie debe tener una muy buena aproximación numérica, para que los residuos al cuadrado estimados no difieran de manera significativa con los residuos al cuadrado originales.

## **7.6. Aplicación del Procedimiento a Series de Tiempo Lineales No Estacionarias y No Explosivas.**

La aplicación de la técnica propuesta a tres series de tiempo generadas por modelos no estacionarios que no son explosivos (caminatas aleatorias de orden 1,2,3) permite, a través de los resultados obtenidos, concluir lo siguiente:

- Si se puede aplicar la técnica con resultados satisfactorios, en cuanto a la identificación del retardo significativo de series de tiempo lineales no estacionarias, que no sean de tipo explosivo, esto es, series de tiempo generadas por procesos estocásticos de caminata aleatoria, ver Sección 2.3.5 del Capítulo 2, y Sección 6.1 del Capítulo 6.

- Las gráficas de la Tabla 6.1 del Capítulo 6, muestran su buen desempeño en la identificación, aún cuando presentan un comportamiento peculiar para los respectivos múltiplos del retardo significativo, que se debe a que para estas caminatas aleatorias los vectores retardados presentan una estructura de período  $t - (m - 1)\tau$ . Lo mismo ocurre al analizar el comportamiento del procedimiento para tamaños de muestra pequeños, 50, 100 y 200 observaciones, ver Tablas de la 6.2 a la 6.4, Capítulo 6.
- A partir de los resultados obtenidos para la potencia de la prueba, se puede concluir que para el nivel nominal seleccionado de la prueba, su desempeño bajo la hipótesis nula es relativamente correcta. En este caso también se presenta la peculiaridad de dependencias entre los múltiplos de periodicidad par e impar del retardo significativo. Ver Tablas 6.5, 6.6 y 6.7 del Capítulo 6.
- Ya que las series de tiempo no lineales y no estacionarias son en general de tipo explosivo, y el valor observado de la variable  $x_t$ , sufre cambios bruscos tomando valores infinitos para la aritmética de punto flotante de la computadora, cuando se incrementa el tiempo, no es posible obtener numéricamente los vectores retardados que se requieren para calcular el estimador de la integral de correlación,  $\hat{C}_{m,n}^\tau(\varepsilon)$ , razón por la cual no se realizó ningún estudio con este tipo de series. Esta situación permite concluir que para obtener resultados correctos con la técnica propuesta es necesario que la serie de tiempo que se quiera analizar cumpla con el requerimiento de estacionariedad.

---

---

## APÉNDICE A

---

### APENDICE DEL CAPITULO 5: RESULTADOS NUMERICOS DE LAS SIMULACIONES, TAMAÑOS DE MUESTRA PEQUEÑOS.

#### A.1. Introducción

En este apéndice del Capítulo 5 de la tesis, se presentan los resultados numéricos que se obtuvieron al ejecutarse las simulaciones Monte Carlo y Bootstrap a las series estudiadas utilizando tamaños de muestra pequeños, esto es,  $T = 50, 100$  y  $200$ .

En la sección A.1 se presentan los resultados de la aplicación del Procedimiento 1, presentado en la sección 4.2.1 del Capítulo 4, con el que se identifica el retardo significativo de la serie a partir del valor máximo del promedio de la integral de correlación  $\hat{C}_{m,n}(\epsilon)$ . En la sección A.2 se presentan los resultados de la aplicación del Procedimiento 2, presentado en la sección 4.2.2 del Capítulo 4, con el que se evalúa el rendimiento empírico de la prueba en cuanto a su tamaño y potencia, en los modelos de tipo lineal o no lineal. En la sección A.3 se muestran los resultados obtenidos en

la simulación Monte Carlo y Bootstrap realizada con el Modelo 11, ver sección 5.3.3 del capítulo 5, mismo que corresponde a un proceso autorregresivo lineal de orden 10, AR(10), que presenta dos retardos significativos, el 6 y el 10. En la sección A.4 se muestran los resultados de la aplicación de las técnicas propuestas a los residuos al cuadrado estimados de un modelo ajustado, ver sección 5.4 del Capítulo 5.

## **A.2. Resultados de la Aplicación del Procedimiento 1: Identificación del Retardo a Partir del Valor Máximo del Promedio de $\hat{C}_{m,n}^{\tau}(\epsilon)$ .**

En esta sección, en las Tablas A.1, A.2 y A.3, se presentan los resultados numéricos del comportamiento del Procedimiento 1, al ser aplicado a series de tiempo generadas por los modelos estacionarios que aparecen en la Tabla 4.1 del Capítulo 4. Los tamaños de muestra utilizados fueron  $T = 50, 100$  y  $200$ .

De la misma forma que para los tamaños de muestra grande, las gráficas del promedio de la integral de correlación para el modelo  $N(0,1)$ , muestran que no existe un retardo relevante. Para el caso de los modelos lineales AR(2) y MA(2), en ambos se identifica que el retardo correcto siempre reporta el valor más grande para la estimación de la integral de correlación para los tres tamaños de muestra, y para un tamaño de muestra  $T \geq 100$ , también se detecta la propiedad de memoria del modelo MA(2), esto es, para los retardos arriba del 2 el valor de la integral no cambia.

Para los siguientes dos modelos no lineales, NMA(3) y NAR(1), el estimador de la integral de correlación detecta claramente los retardos correctos y muestra sensibilidad a las propiedades de memoria de ambos modelos a partir de tamaños de muestra  $T \geq 100$ . El orden de dependencia del modelo caótico (logístico) es detectado en el retador correcto para los tres tamaños de muestra, ocurriendo lo mismo para el proceso bilinial BL(2,1), para el cual dos retardos, 2 y 1 son relevantes.

Por lo que respecta a los modelos con estructura de varianza condicional, el procedimiento detecta los retardos correctos correspondientes para el ARCH(2) y el GARCH(1,1), retardos 2 y 1 respectivamente, para cada una de las  $m$  dimensiones y para todos los tamaños pequeños de muestra analizados.

Para el ARCH(3), el procedimiento detecta el retardo correcto para las dimensiones  $m = 2$  y  $m = 3$  para los tres tamaños de muestra,  $T = 50, 100, 200$ .

Tabla A.1: Promedios de la Integral de Correlación para procesos estacionarios,  $T_1 = 50$

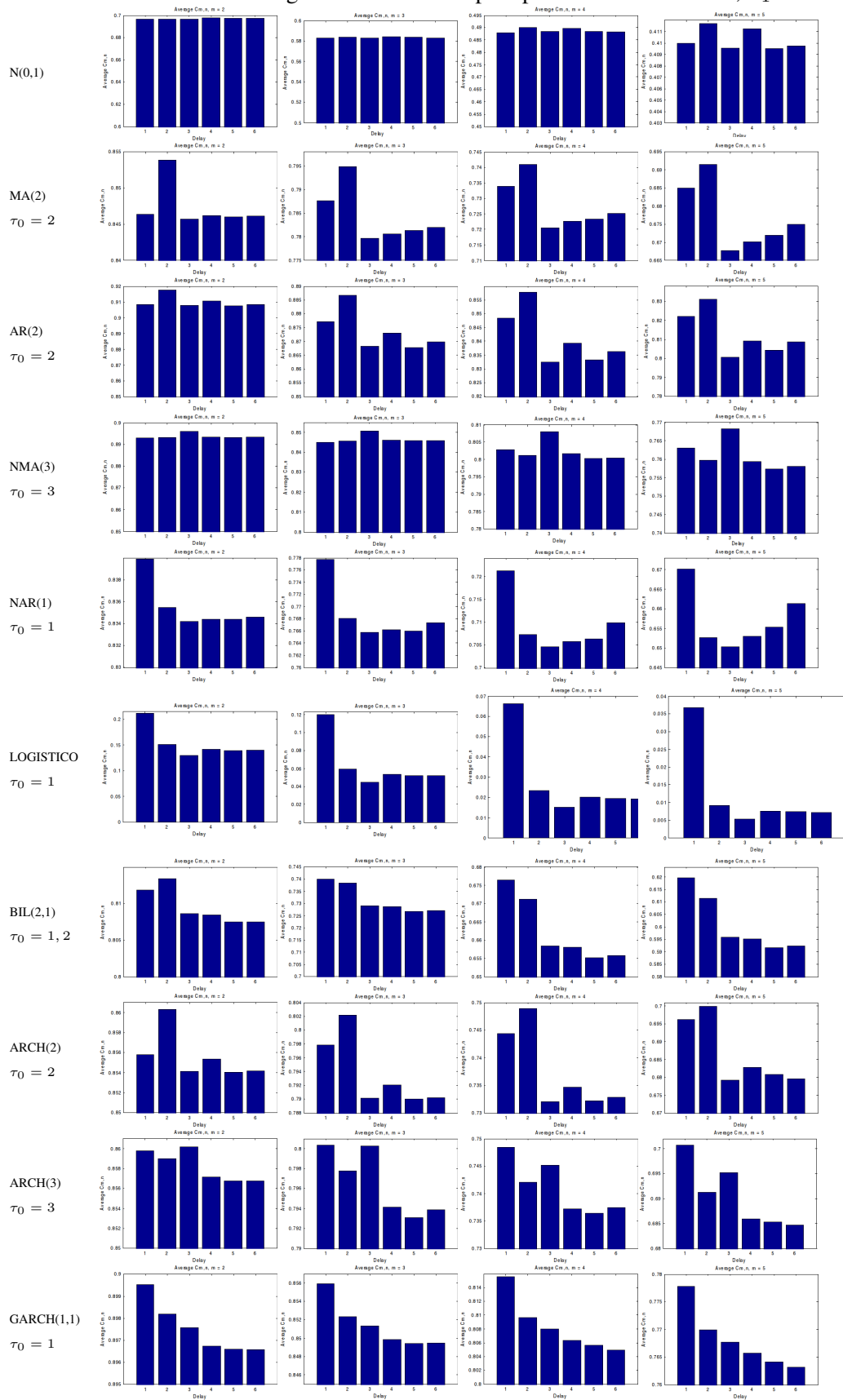


Tabla A.2: Promedios de la Integral de Correlación para procesos estacionarios,  $T_2 = 100$

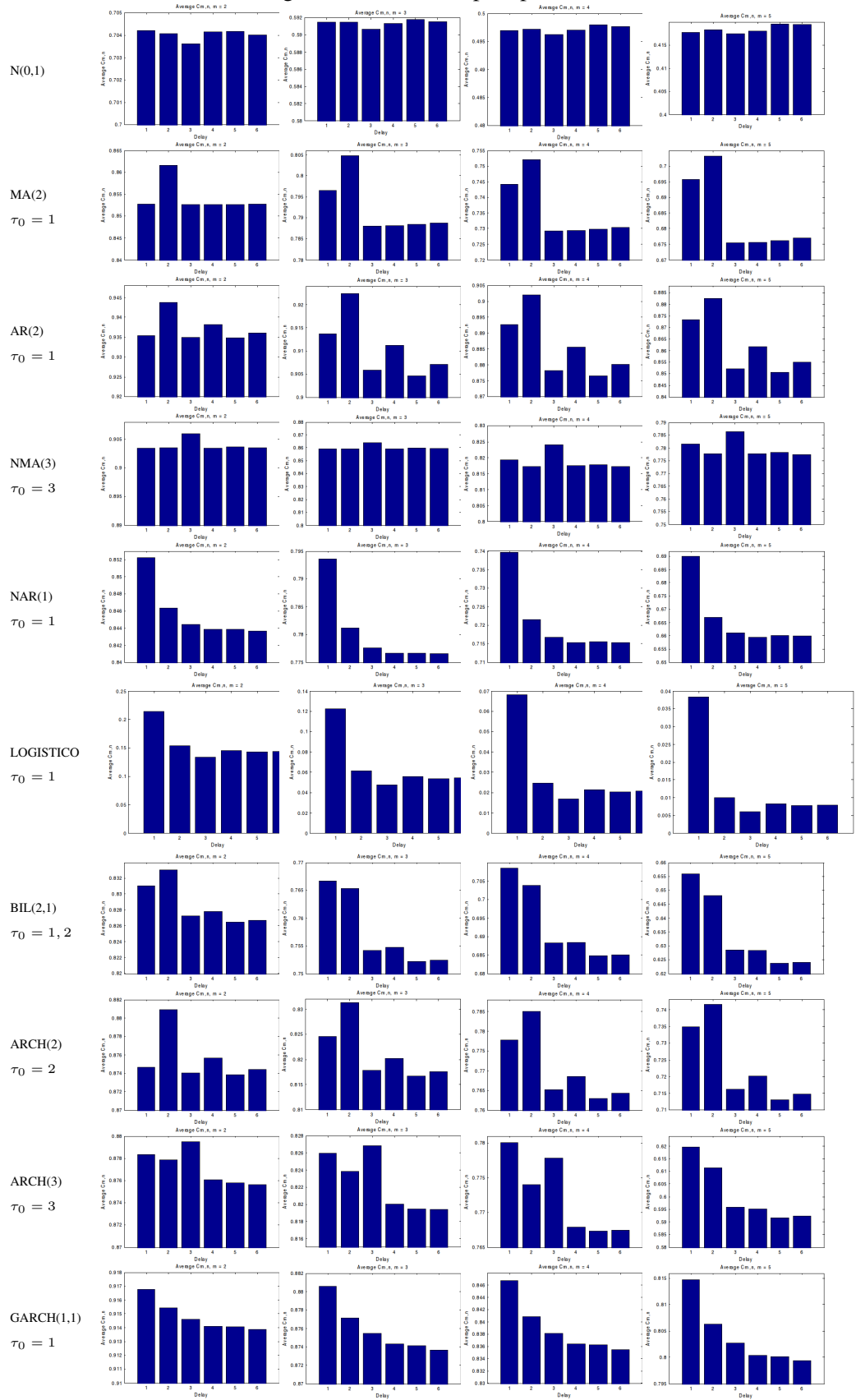
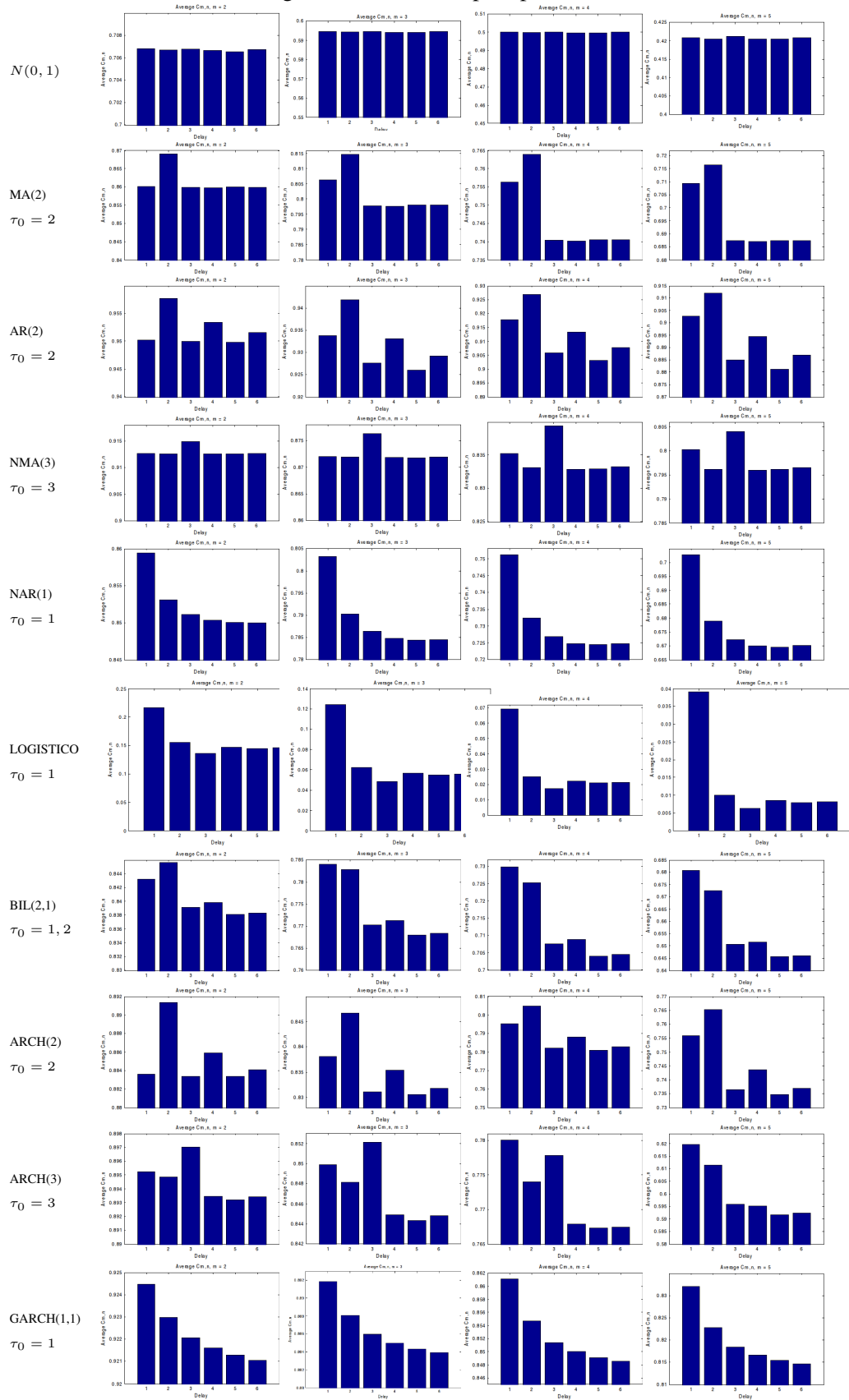


Tabla A.3: Promedios de la Integral de Correlación para procesos estacionarios,  $T_3 = 200$





### **A.3. Resultados de la Aplicación del Procedimiento 2: Prueba para Identificar Retardos en Modelos Lineales y No Lineales.**

En esta sección, se presentan los resultados numéricos de la simulación Monte Carlo y el Bootstrap, con la que se hizo el estudio del tamaño y de la potencia de la prueba, ver sección 4.2.2 del Capítulo 4, y secciones: 5.3.1, 5.3.2 del Capítulo 5, considerando tamaños de muestra pequeños de: 50, 100 y 200, y la misma hipótesis nula utilizada para tamaños de muestra grandes.

Las Tablas A.4, A.5 y A.6 presentan los resultados correspondientes al estudio del tamaño de la prueba, utilizando algunos de los modelos presentados en la Tabla 4.1, del Capítulo 4, el  $N(0,1)$ , el MA(2) y el NMA(3). Estas tablas deben ser leídas por renglón: por ejemplo, el primer renglón con resultados en la tabla A.4 presenta el tamaño estimado cuando se prueba la hipótesis nula de que el retardo 1 proporciona la misma información que los retardos 2, 3, 4, 5 y 6; la prueba de la nula para el segundo renglón es que el retardo 2 proporciona la misma información para los retardos 1, 3, 4, 5 y 6; y el último renglón presenta los resultados cuando se prueba el retardo 6 contra los retardos 1, 2, . . . , 5.

Nótese que, al igual que ocurrió con el tamaño de muestras grandes, para el Modelo 1 todos los retardos teóricamente contienen la misma información (son todos igualmente no relevantes), y por lo tanto se está bajo la hipótesis nula. Los resultados muestran que se está en el nivel nominal de rechazo esperado.

Las Tablas A.5 y A.6 que deben ser leídas de la misma forma que la anterior, corresponden a modelos con dependencia lineal y no lineal, en los que no existe una relación teórica más allá de cierto retraso, y por tal motivo se prueban diferentes configuraciones posibles de la hipótesis nula; en la Tabla A.5 se probó el retardo 3 contra los retardos 4, 5 y 6, y en la tabla A.6 el retardo 4 contra los retardos 5 y 6.

Tabla A.4: Tamaño de la prueba para muestras de tamaño pequeño Modelo 1.

$$X_t = e_t e_t \sim N(0, 1)$$

<b>Lag, T</b>	1	2	3	4	5	6
<b>1, 50</b>		0.052	0.033	0.044	0.045	0.032
100		0.040	0.039	0.053	0.040	0.047
200		0.056	0.048	0.041	0.049	0.057
<b>2</b>	0.040		0.043	0.036	0.060	0.035
	0.051		0.046	0.052	0.047	0.049
	0.039		0.046	0.066	0.050	0.057
<b>3</b>	0.038	0.043		0.052	0.046	0.041
	0.041	0.040		0.057	0.038	0.043
	0.047	0.051		0.047	0.056	0.051
<b>4</b>	0.045	0.045	0.034		0.041	0.029
	0.048	0.057	0.058		0.044	0.050
	0.052	0.056	0.060		0.045	0.049
<b>5</b>	0.044	0.049	0.038	0.043		0.040
	0.051	0.056	0.056	0.042		0.050
	0.049	0.042	0.037	0.055		0.042
<b>6</b>	0.051	0.045	0.039	0.044	0.043	
	0.042	0.054	0.060	0.043	0.056	
	0.041	0.045	0.051	0.041	0.047	

Tabla A.5: Tamaño de la prueba para muestras de tamaño pequeño Modelo 2.

$$X_t = e_t + 0.8e_{t-2}; e_t \sim N(0, 1)$$

<b>Lag, T</b>	3	4	5	6
<b>3, 50</b>		0.044	0.029	0.054
100		0.036	0.038	0.067
200		0.044	0.041	0.055
<b>4</b>	0.042		0.043	0.034
	0.047		0.062	0.044
	0.054		0.052	0.039
<b>5</b>	0.044	0.045		0.057
	0.046	0.050		0.051
	0.042	0.051		0.045
<b>6</b>	0.061	0.043	0.039	
	0.048	0.032	0.042	
	0.066	0.040	0.053	

Los resultados que se muestran en estas tablas, considerando la sensibilidad a las propiedades de memoria que tienen los modelos, permiten concluir que también para tamaños pequeños de muestra, el desempeño de la prueba bajo la hipótesis nula es correcto.

Tabla A.6: Tamaño de la prueba para muestras de tamaño pequeño Modelo 4.

$$X_t = e_t + 0.8e_{t-3}^2; e_t \sim N(0, 1)$$

<b>Lag, T</b>	4	5	6
<b>4, 50</b>		0.030	0.048
100		0.050	0.039
200		0.028	0.037
<b>5</b>	0.048		0.040
	0.045		0.046
	0.046		0.037
<b>6</b>	0.042	0.042	
	0.039	0.037	
	0.043	0.048	

Con lo que respecta a los experimentos de la potencia de la prueba, en la Tabla A.7 en sus incisos a) al c), se presentan los resultados obtenidos. La tabla debe leerse renglón por renglón (modelo por modelo). Cada renglón (modelo) tiene una celda vacía cuya columna es indicativa del retardo que se está usando para la prueba: por ejemplo, la celda vacía para el modelo MA(2) es la columna 2, lo que significa es que, para este modelo en particular, se está probando el retardo 2 contra los restantes retardos denominados, 1, 3, 4, 5 y 6. La razón de esto es porque el retardo relevante para el modelo MA(2) es el 2, y por lo tanto se estima la probabilidad de rechazar la hipótesis nula de que el retardo 2 no tiene mayor información que cualquier otro retardo considerado.

Comparando estos resultados con los que se obtuvieron para los tamaños de muestra grandes, ver tablas 5.9 y 5.10 del Capítulo 5, se puede concluir que para que el rendimiento sea bueno, a excepción del mapeo logístico determinista caótico (el tamaño de muestra puede ser a partir de 50), y el modelos lineal AR(2) (el tamaño de muestra puede ser mayor a 200), los tamaños de muestra que se recomiendan utilizar son mayores de 200.

Tabla A.7: Potencia de la prueba para muestras de tamaño pequeño

a)  $T_1 = 50$

Modelo	Retardo					
	1	2	3	4	5	6
MA(2)	0.112		0.199	0.241	0.169	0.142
AR(2)	0.353		0.595	0.361	0.420	0.309
NMA(3)	0.076	0.072		0.062	0.063	0.077
NAR(1)		0.250	0.203	0.210	0.204	0.162
Logistico		0.999	1.0	0.999	0.998	0.997
BIL(2,1)	0.091		0.084	0.116	0.113	0.099
ARCH(2)	0.092		0.169	0.134	0.140	0.129
ARCH(3)	0.079	0.060		0.052	0.088	0.108
GARCH(1,1)		0.062	0.066	0.084	0.109	0.115

b)  $T_2 = 100$

Modelo	Retardo					
	1	2	3	4	5	6
MA(2)	0.265		0.539	0.653	0.429	0.387
AR(2)	0.868		0.975	0.807	0.906	0.760
NMA(3)	0.126	0.141		0.121	0.126	0.130
NAR(1)		0.589	0.543	0.537	0.479	0.462
Logistico		1.0	1.0	1.0	1.0	1.0
BIL(2,1)	0.104		0.177	0.225	0.226	0.214
ARCH(2)	0.218		0.416	0.297	0.359	0.283
ARCH(3)	0.105	0.105		0.142	0.168	0.206
GARCH(1,1)		0.164	0.182	0.209	0.221	0.219

c)  $T_3 = 200$

Modelo	Retardo					
	1	2	3	4	5	6
MA(2)	0.599		0.910	0.965	0.859	0.832
AR(2)	0.996		1.0	0.989	1.0	0.985
NMA(3)	0.266	0.278		0.273	0.268	0.345
NAR(1)		0.942	0.927	0.916	0.896	0.875
Logistico		1.0	1.0	1.0	1.0	1.0
BIL(2,1)	0.108		0.367	0.493	0.490	0.443
ARCH(2)	0.511		0.766	0.665	0.730	0.614
ARCH(3)	0.154	0.186		0.306	0.333	0.390
GARCH(1,1)		0.256	0.329	0.410	0.444	0.448

## A.4. Resultados Numéricos de la Aplicación del Procedimiento Automático para Seleccionar Retardos.

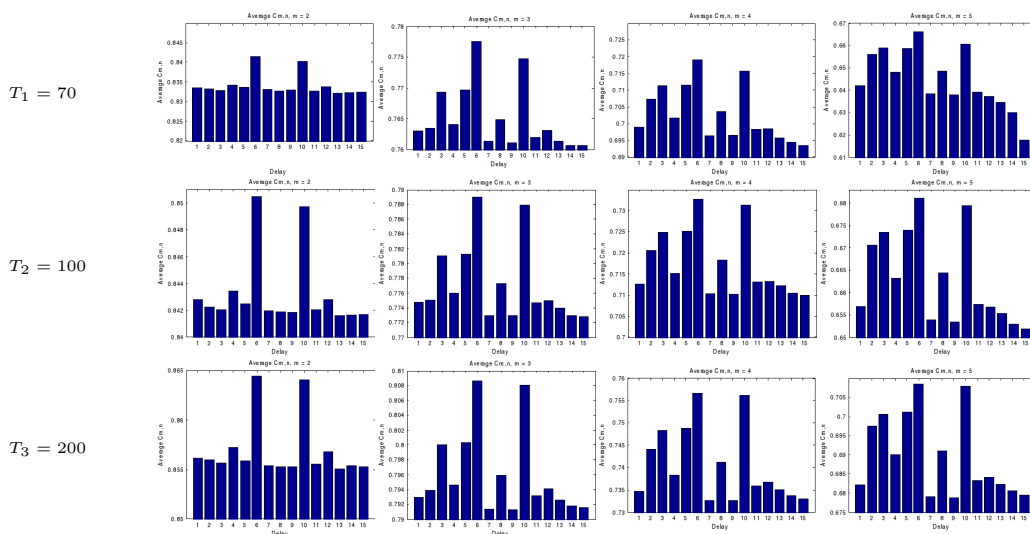
En esta sección se presentan los resultados de simulaciones Monte Carlo y Bootstrap similares a las aplicadas a los modelos de la Tabla 4.1 ya antes mencionados, tanto para la identificación del retardo significativo a partir del valor máximo del promedio  $\hat{C}_{m,n}^\tau(\varepsilon)$ , como para obtener el tamaño y potencia de la prueba, pero ahora aplicadas al Modelo 11 que aparece en la sección 4.2.3 del capítulo 4.

### A.4.1. Resultados de la identificación del retardo significativo a partir del valor máximo del promedio $\hat{C}_{m,n}^\tau(\varepsilon)$

En la tabla A.8 se muestran los resultados de la identificación de los retardos significativos 6 y 10 a partir del valor promedio de la integral de correlación del Modelo 11, modelo autorregresivo lineal de orden 10, AR(10). Las gráficas corresponden a los tamaños de muestra  $T = 70, 100$  y  $200$ ,  $m = 2, 3, 4$  y  $5$ , y  $\tau = 1, 2, \dots, 6, \dots, 10, \dots, 15$

Tabla A.8: Promedios de la Integral de Correlación para el Modelo 11.

$$X_t = -0.4X_{t-6} + 0.4X_{t-10} + e_t; T = 70, 100, 200, \tau = 6, 10$$



Al igual que ocurrió con los tamaños de muestra grandes, se puede observar en estas gráficas del promedio de  $\hat{C}_{m,n}^{\tau}$ , que los retardo con el valor más grande para la integral de correlación, son los correspondientes al retardo 6 y al 10, y que por lo tanto se está identificando el retardo apropiadamente, también para muestras de tamaño pequeño.

#### A.4.2. Resultados del rendimiento de la prueba con muestras pequeñas

La Tabla A.9, muestra los resultados del procedimiento evaluado para  $T = 70, 100$  y  $200$  observaciones. Esta tabla se debe leer renglón por renglón, los cuatro primeros renglones corresponden a uno de los retardos significativos del Modelo, el 6, para los tamaños de muestra mencionados, cuyos renglones tienen una celda vacía, cuya columna es indicativa de este retardo, lo que significa que se está probando el retardo 6 contra los restantes retardo denominados  $1, \dots, 5, 7, \dots, 15$ . Los siguientes cuatro renglones corresponden al segundo retardo significativo del modelo, el 10, por lo que en este caso se está probando el retardo 10 contra los restantes retardos denominados,  $1, \dots, 9, 11, \dots, 15$ , para los tres tamaños de muestra mencionados.

Tabla A.9: Resultados del Tamaño de la Prueba, para el Modelo 11, AR(10)

Retardo	Retardo														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
<b>6</b>															
<b>T = 70</b>	0.148	0.159	0.097	0.187	0.150		0.316	0.193	0.260	0.097	0.164	0.191	0.181	0.183	0.178
<b>T = 100</b>	0.239	0.235	0.162	0.299	0.211		0.476	0.281	0.432	0.105	0.295	0.351	0.309	0.313	0.284
<b>T = 200</b>	0.701	0.673	0.470	0.681	0.378		0.814	0.594	0.783	0.115	0.656	0.714	0.621	0.634	0.596
<b>10</b>															
<b>T = 70</b>	0.083	0.090	0.059	0.107	0.070	0.084	0.205	0.146	0.280		0.229	0.194	0.186	0.200	0.154
<b>T = 100</b>	0.174	0.181	0.096	0.190	0.128	0.097	0.388	0.242	0.448		0.378	0.318	0.344	0.324	0.323
<b>T = 200</b>	0.559	0.494	0.247	0.519	0.360	0.102	0.764	0.552	0.783		0.694	0.631	0.716	0.716	0.694

Se puede observar que para un tamaño de muestra pequeño los resultados no son tan buenos como

los que se presentan para los tamaños de muestra grande (ver Tabla 5.12 en el Capítulo 5); sin embargo, si se puede ver que cuando se prueba el retardo 6 contra el 10 y viceversa, los resultados que se obtienen son muy similares, lo que indicaría que ambos son igualmente relevantes.

## **A.5. Resultados Numéricos de la Aplicación de las Técnicas Propuestas a los Residuos Estimados de un Modelo Ajustado.**

En esta sección se muestran los resultados del estadístico del tamaño y potencia de la prueba obtenidos con la simulación Monte Carlo y Bootstrap a las series de residuos al cuadrado obtenidos a partir del ajuste de los datos generados con los cuatro modelos ARMA-GARCH y los filtros de ARMA(p,q), presentados en la Tabla 4.3, del Capítulo 4, para tamaños de muestra pequeños, esto es,  $T = 50, 100$  y  $200$ . Las técnicas de simulación y bootstrap se aplicaron considerando los parámetros de inmersión:  $m = 3$  y  $\tau = 1, 2, 3, 4, 5, 6$ , para un nivel de significancia de  $\alpha = 0.05$ . El número de muestras bootstrap en cada caso fueron  $B = 199$ , y el número de replicas de la simulación fue de  $N_r = 1000$ .

En las Tablas A.10, A.11 y A.12 se muestran los resultados del estadístico del tamaño de la prueba de la serie residual con el modelo ARMA(1,1)-N(0,1), para los tres tamaños de muestra mencionados arriba. Estas tablas indican la proporción de rechazo bajo la hipótesis nula, y deben ser leídas renglón por renglón. El primer renglón muestra la proporción de rechazo cuando la prueba considera el retardo 1 ( $\tau_0 = 1$ ) contra los retardos 2, 3 . . . , 6. La interpretación del otro renglón tiene que hacerse de la misma manera.



Tabla A.10: Tamaño de la prueba de la serie residual con el modelo ARMA(1,1)-N(0,1)

$$T_1 = 50$$

$\tau_0$	$\tau$					
	1	2	3	4	5	6
1		0.052	0.050	0.055	0.043	0.043
2	0.038		0.052	0.050	0.044	0.055
3	0.055	0.058		0.045	0.044	0.050
4	0.045	0.043	0.052		0.046	0.060
5	0.059	0.060	0.054	0.056		0.050
6	0.045	0.042	0.054	0.055	0.052	

Tabla A.11: Tamaño de la prueba de la serie residual con el modelo ARMA(1,1)-N(0,1)

$$T_2 = 100$$

$\tau_0$	$\tau$					
	1	2	3	4	5	6
1		0.049	0.053	0.059	0.045	0.053
2	0.045		0.052	0.050	0.059	0.048
3	0.060	0.057		0.053	0.054	0.041
4	0.042	0.040	0.042		0.052	0.057
5	0.056	0.053	0.055	0.055		0.050
6	0.045	0.042	0.054	0.055	0.052	

Tabla A.12: Tamaño de la prueba de la serie residual con el modelo ARMA(1,1)-N(0,1)

$$T_3 = 200$$


---

$\tau_0$	$\tau$					
	1	2	3	4	5	6
1		0.050	0.041	0.041	0.045	0.045
2	0.050		0.070	0.056	0.040	0.057
3	0.050	0.047		0.055	0.053	0.056
4	0.039	0.042	0.054		0.051	0.046
5	0.041	0.053	0.048	0.056		0.045
6	0.062	0.048	0.052	0.050	0.044	

---

Los resultados para el modelo ARMA(1,1)-N(0,1) son los valores esperados para los que se aplicó la técnica de identificación, estos valores se encuentran dentro de un rango  $[0.038, 0.060]$ , para un  $\alpha = 0.05$ . Nótese que este buen comportamiento empírico para el tamaño es independiente del tamaño de la muestra.

En las Tablas A.13, A.14 y A.15 se presentan la potencia empírica de la prueba para la serie de residuales al cuadrado resultante de filtrar con un ARMA(1,0), ARMA (0,1) y ARMA(1,1), los 3 modelos ARMA-GARCH presentados en la Tabla 4.3, en el Capítulo 4 para los tamaños de muestra pequeños ya mencionados. El retardo significativo en los tres modelos es  $\tau_0 = 1$

Tabla A.13: Potencia de la prueba en la serie de residuales del modelo ARMA(1,0)-GARCH(1,1)

ARMA(1,0)-GARCH(1,1)						
	$\tau$					
	1	2	3	4	5	6
$T_1 = 50$		0.110	0.075	0.090	0.080	0.079
$T_2 = 100$		0.122	0.113	0.123	0.139	0.137
$T_3 = 200$		0.193	0.199	0.244	0.280	0.270

Tabla A.14: Potencia de la prueba en la serie de residuales del modelo ARMA(0,1)-GARCH(1,1)

ARMA(0,1)-GARCH(1,1)						
	$\tau$					
	1	2	3	4	5	6
$T_1 = 50$		0.092	0.058	0.083	0.072	0.080
$T_2 = 100$		0.135	0.120	0.125	0.128	0.130
$T_3 = 200$		0.197	0.210	0.248	0.271	0.282

Tabla A.15: Potencia de la prueba en la serie de residuales del modelo ARMA(1,1)-GARCH(1,1)

ARMA(1,1)-GARCH(1,1)						
	$\tau$					
	1	2	3	4	5	6
$T_1 = 50$		0.102	0.081	0.089	0.071	0.070
$T_2 = 100$		0.126	0.121	0.135	0.146	0.142
$T_3 = 200$		0.199	0.222	0.251	0.266	0.282

Las conclusiones a las que se llegan con estos resultados son las siguientes:

- 1) El comportamiento de la prueba para los tres modelos es básicamente el mismo y no es tan bueno como el reportado para tamaños de muestra grandes (ver sección 5.4 del Capítulo 5).
- 2) Al igual que con las muestras de tamaño grande, la potencia de la prueba va aumentando.
- 3) Resumiendo los resultados de las proporciones de rechazo para tamaños de muestra grandes y pequeños, se recomienda utilizar la prueba en los residuales para tamaños de muestra que contengan al menos 1000 observaciones.

---

## BIBLIOGRAFÍA

- [Albano et al, 1988] Albano A.M., Muench J., Schwartz C., Mees A.I., Rap P.E. (1988). *Singular-value decomposition and the Grassberger-Procaccia algorithm*. Physical Review A, Vol.38, No. 6.
- [Ashley and Patterson, 1989] Ashley R.A. and Patterson D.M. (1989). *Linear Versus Nonlinear Macroeconomics: A Statistical Test*. International Economic Review, Vol.30, No.3. pp.685-704.
- [Ashley, Altug and Patterson, 1997] Ashley R., Altug S. and Patterson D. (1997). *Are thecnology shocks nonlinear?*. Unpublished Manuscript, VPI&SU.
- [Barnett and Chen,1988] Barnett W. and Chen P. (1988). *The Aggregation-Theoretic Monetary Aggregates Are Chaotic and Have Strange Attractors*, in Barnett W., Berndt E. and White H.,eds., *Dynamic Econometric Modelling*. Proceedings of the Third Austin Symposium, Cambridge: Cambridge University Press.
- [Barnett et al, 1997] Barnett William A. & Gallant A. Ronald & Hinich Melvin J. & Jungeilges Jochen A. & Kaplan Daniel T. & Jensen Mark J., (1997). *A single-blind controlled competition among tests for nonlinearity and chaos*. Journal of Econometrics, Vol.82(1), pp.157-192.
- [Bollerslev,1986] Bollerslev Tim, (1986). *Generalized Autoregressive Conditional Heteroskedasticity*. Journal of Econometrics, Vol.31, issue 3, pp.307-327.

- [Brock and Dechert,1988a] Brock W. and Dechert W., (1988a). *Theorems on Distinguishing Deterministic and Random Systems*, in Barnett W., Berndt E. and White H., eds., *Dynamic Econometric Modelling, Proceedings of the Third Austin Symposium*. Cambridge: Cambridge University Press.
- [Brock et al,1993] Brock William A., Hsieh David A., LeBaron Blake, (1993). *Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence*. The Mit Press, Cambridge, Massachusetts, London, England, ISBN 0-262-02329-6.
- [Brock et al,1990] Brock W.A., Dechert W.D., Scheinkman J.A., LeBaron B., (1990). *A Test for Independence Based on the Correlation Dimension*. Department of Economics, University of Wisconsin, Madison, University of Houston, University of Chicago.
- [Broer,Takens,2009] Broer Henk, Takens Floris, (2009) *Dynamical Systems and Chaos*. Springer, New York, ISBN 978-1-4419-6869-2.
- [Buzug et al,1999] Buzug Th., Reimers T., and Pfister G., (1999). *Optimal Reconstruction of Strange Attractors from Purely Geometrical Arguments*. Europhysics, Letters, Vol.13, pp.605.
- [Cees,1999] Cees Diks, (1999). *Nonlinear Time Series Analysis, Nonlinear Time Series and Chaos*. Vol.4, World Scientific Publishing Co. Pte. Ltd. USA, ISBN 9810235054.
- [Clinton,2003] Clinton Sprott Julien, (2003). *Chaos and Time-Series Analysis*. Oxford University Press, Great Britain, ISBN 978-0-19-850839-7.
- [Davison and Hinkley,1997] Davison A.C., Hinkley D.V., (1997). *Bootstrap Methods and their Application*. Cambridge University Press, Cambridge, N.Y.
- [De Lima,1996] De Lima Pedro J.F., (1996). *Nuisance Parameter Free Properties of Correlation Integral Based Statistics*. Econometric Reviews, Vol.15(3), pp.237-259.
- [Denker and Keller,1983] Denker, M. and Keller, G. (1983). *On U-Statistics and V. Mises Statistics for Weakly Dependent Processes*. Z. Wahrscheinlichkeitstheorie und verwandte Gebiete Vol.64, pp. 505-522. Springer-Verlag.

- [Efron,1979] Efron B., (1979). *Computers and the theory of statistics: thinking the unthinkable*. SIAM Review, Vol.21, pp.460-479.
- [Efron,1982] Efron B., (1982). *The Jackknife, The Bootstrap and Other Resampling Plans*. Society of Industrial and Applied Mathematics CBMS-NSF Monographs, 38.
- [Efron and Tibshirani,1986] Efron B., and Tibshirani R., (1986). *Bootstrap Methods for Standard Error, Confidence Intervals, and Other Measures of Statistical Accuracy*. Statistical Science Vol.1, pp.54-77.
- [Efron and Tibshirani,1993] Efron B., and Tibshirani R., (1993). *An Introduction to the Bootstrap*. London:Chapman and Hall.
- [Engle,1982] Engle Robert, (1982). *Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation*. Econometrica, Vol.50, pp.987-1008.
- [Eubank and Farmer,1990] Eubank S., and Farmer J.D., (1990). *An introduction to chaos and randomness, in 1989 Lectures in Complex Systems*. SFI Studies in the Sciences of Complexity, Lect. Vol 2, Erica Jen, ed., pp.75, Addison-Wesley.
- [Figueiredo et al,2011] Figueiredo C., Diambra L. and Pereira C., (2011). *Convergence Criterium of Numerical Chaotic Solutions Based on Statistical Measures*. Applied Mathematics, Vol.2, pp. 436-443.
- [Franses and van Dijk, 2004] Franses Philip Hans, van Dijk Dick, (2004). *Non-linear time series models in empirical finance*. Cambridge University Press, Cambridge N.Y., ISBN 0-521-77041-6.
- [Fraser and Swinney,1986] Fraser A.M., and Swinney H.L., (1986). *Independent coordinates for strange attractors from mutual information*. Phys. Rev. A, Vol.33, pp.1134.
- [Fuller,1976] Fuller Wayne A., (1976). *Introduction to Statistical Time Series*. John Wiley & Sons, Inc. ISBN 0-471-28715-6.

- [Granger and Anderson,1978] Granger C.W.J. and Anderson A.A., (1978). *An Introduction to Bilinear Time Series Models*. Vandenhoeur and Ruprecht: Gottingen.
- [Granger, 1983] Granger C.W.J. and Weiss A., (1983) *Time Series Analysis of Error-Correcting Models*. Studies in Econometrics, Time Series, and Multivariate Statistics. Academic Press, New York, pp.255-278.
- [Granger and Lin, 1994] Granger C.W.J. and Lin J-L., (1994). *Using the mutual information coefficient to identify lags in nonlinear models*. J. Time Series Anal., Vol.15, pp.371-384.
- [Granger et al, 2003] Granger C.W., Maasoumi E., Racine J., (2003). *A Dependence Metric for Possibly Nonlinear Processes*. Department of Economics & Center for Policy Research, Syracuse University, Syracuse, N.Y., 13244-1020.
- [G-P, 1983] Grassberger Peter and Procaccia Itamar, (1983). *Measuring the Strangeness of Strange Attractors*. Physica 9D, pp.189-208, North-Holland Publishing Company.
- [Grassberger and Procaccia, 1983] Grassberger Peter and Procaccia Itamar, (1983). *Characterization of Strange Attractors*. Physical Review Letters, Vol. 50, No.5, pp.346-349.
- [Hiemstra and Jones,1994] Hiemstra C. and Jones J.D., (1994). *Testing for linear and nonlinear Granger causality in the stock price-volumen relation*. Journal of Finance, Vol.49, pp.1639-1665.
- [Hinich,1982] Hinich, M.J., (1982). *Testing for Gaussianity and Linearity of a Stationary Time Series*. Journal of Time Series Analysis Vol.3, pp.169-176.
- [Hope, 1968] Hope A.C.A, (1968). *A simplified Monte Carlo Significance test procedure*. Journal of the Royal Statistical Society, Series B, Vol.30, pp.582-598.
- [Johnson and McClelland,1998] Johnson David and McClelland Robert, (1998). *General Dependence Test and Applications*. Journal of Applied Econometrics, Vol.13, Issue 6, pp.627-644.



- [Kantz and Schreiber,2005] Kantz Holger and Schreiber Thomas, (2004). *Nonlinear Time Series Analysis*. Cambridge University Press, Cambridge N.Y. ISBN-13 978-0-521-82150-6.
- [Kanzler,1999] Kanzler L., (1999). *Very Fast and correctly sized estimation of the BDS statistic*. Department of Economics, Oxford University.
- [Kocenda and Briatka,2004] Kocenda E., Briatka L., (2004). *Advancing the iid Test Based on Integration across the Correlation Integral: Ranges, Competition, and Power*. Econometrics 0409001, EconWPA.
- [LeBaron, 1997] LeBaron B., (1997) *A Fast Algorithm for the BDS Statistic*. Studies in Nonlinear Dynamics and Econometrics, Vol.2(2), pp.53-59.
- [McLeod and Li,1983] McLeod A.L. and Li W.K., (1983). *Diagnostic checking ARMA time series models using squared-residual autocorrelations*. Journal of Time Series Analysis, Vol.4, pp.269-273.
- [Mathew and Picu,2011] Mathew N. and Picu R.C., (2011). *Molecular conformational stability in cyclotrimethylene trinitramine crystals*. Journal of Chemical Physics, Vol.135(2).
- [Matilla-García et al,2004] Matilla-García M., Queralt R., Sanz P. and Vazquez F.J., (2004). *A Generalized BDS Statistic*. Computational Economics Vol.24, pp.277-300, Kluwer Academic Publishers.
- [Matilla-García,2007] Matilla-García, M, (2007). *Nonlinear Dynamics in Energy Futures* The Energy Journal, Vol.28, No. 3, pp.7-30.
- [Matilla-García and Ruiz,2009] Matilla-García Mariano, Ruiz Marín Manuel, (2009). *Detection of non-linear structure in time series*. Economic Letters, Vol.105, pp.1-6.
- [Mooney,1997] Mooney C.Z., (1997). *Monte Carlo Simulation*. London:Sage Publications
- [Murdoch,2000] Murdoch Duncan J., (2000). *Markov chain Monte Carlo*. Chance, Vol.13, pp.48-

- [Peña, 2005] Peña Daniel, (2005). *Análisis de series temporales*. Alianza Editorial, S.A. Madrid, España.
- [Packard et al, 1980] Packard, N., Crutchfield, J., Farmer, D., Shaw, R., (1980). *Geometry from a time series*. Physical Review Letters Vol.45, pp.712-715.
- [Politis and White,2004] Politis D.N. and White H., (2004). *Automatic block-length selection for the dependent bootstrap*. Econometric Reviews, Vol. 23, No.1, pp.53-70.
- [Prichard and Theiler,1995] Prichard D., and Theiler J., (1995). *Generalized redundancies for time series analysis*. Physica D, Vol.84, pp.476-493.
- [Ramsey et al,1990] Ramsey J., Sayers C. and Rothman P., (1990). *The Statistical Properties of Dimension Calculations Using Small Data Sets: Some Economic Applications..* International Economic Review, Vol.31, No.4, November, pp.991-1020.
- [Ross,1996] Ross Sheldon M., (1996). *Stochastic Processes*. Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons Inc. USA, ISBN 0-471-09942-2.
- [Small,2005] Small Michael, (2005). *Applied Nonlinear Time Series Analysis: Applications in Physics, Psychology and Finance*. World Scientific Series on Nonlinear Science, Series A, Vol.52.
- [Serfling, 1980] Serfling Robert J., (1980). *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons Inc. USA, ISBN 0-471-02403-1.
- [Solé and Manrubia,2001] Solé Ricard V., Manrubia Susanna, (2001). *Orden y caos en sistemas complejos. Fundamentos*. Ediciones Universidad Politécnica de Cataluña, España. ISBN: 84-8301-430.0.
- [Subba Rao, 1980] Subba Rao T. and Gabr M.M., (1980). *An Introduction to Bispectral Analysis and Bilinear Time Series Models*. Vol. 24 of Lecture Notes in Statistics, Springer-Verlag, New York.

- [Takens,1981] Takens Floris, (1981). *Detecting Strange Attractors in Turbulence*. Lecture Notes in Mathematics Vol.898, pp.355-381.
- [Takens,2003] Takens Floris, (2003). *Linear versus nonlinear time series analysis-smoothed correlation integrals*. Chemical Engineering Journal Vol.96, pp.99-104.
- [Theiler,1986] Theiler J., (1986). *Spurious dimension from correlation algorithms applied to limited time series data*. Phys. Rev. A., Vol.34, pp.2427.
- [Theiler,1990] Theiler J., (1990). *Estimating Fractal dimension*. J.Opt. Soc. Am. A, Vol.7, pp.1055.
- [Tjostheim and Auestad,1994] Tjostheim, D. and Auestad, B., (1994). *Nonparametric identification of nonlinear time series: selecting significant lags*. Journal of the American Statistical Association, Vol.89, pp.1410-1419.
- [Tjostheim,1996] Tjostheim, D., (1996). *Measures and tests of independence: a survey*. Statistics, Vol.28, pp.249-284.
- [Van Der Vaart,1998] Van Der Vaart A.W., (1998). *Asymptotic Statistics*. Cambridge University Press, UK, ISBN 0-521-49603-9.
- [Van Der Vaart,2000] Van Der Vaart A. and Wellner J.A., (2000). *Weak Convergence of Empirical Processes*. Springer Series in Statistics, New York: Springer.