

Capítulo 13

Identificación del Retardo Temporal Significante, en la Modelación de Series de Tiempo Económicas, Lineales y no Lineales

Rina B. Ojeda Castañeda¹

Abstract: The identification of the lags that should be included in a general model description of a time series, is certainly a crucial first step to an analyst faces in the economic development of prediction models. Autocorrelation and partial autocorrelation coefficients have traditionally been used to specify the appropriate lags, even as it has well established in [Granger1983], that processes with zero autocorrelation dependence exhibit higher order or nonlinear dependence, as in the case some bilinear processes, and more so for purely deterministic chaotic models among others. In general, methods based on autocorrelations fail to be used to specify the significant lag in nonlinear models, not detect significant nonlinear relationships in the data, nor the appropriate time lags, especially in scenarios where the phenomena linear are more the rule than the exception. In this research it is proposed as a tool for modeling, a new non-parametric statistical test and a method to identify the significant time lag in the description of a linear or non-linear general model for stationary time series. These techniques are based on estimating the correlation integral.

Keywords: Correlation Integral, Lag, Time Series, Nonlinearity.

Resumen: La identificación de los retardos temporales que deben incluirse en un modelo general que describe una serie de tiempo, es

¹rbojeda@cima.uadec.mx. Centro de Investigación en Matemáticas Aplicadas, CIMA, Universidad Autónoma de Coahuila.

el primer paso crucial al que se enfrenta un analista, en el desarrollo de modelos económicos de predicción. Los coeficientes de autocorrelación y autocorrelación parcial, han sido tradicionalmente utilizados para especificar los retardos temporales apropiados, como bien se ha establecido ([Granger1983]), que procesos con una autocorrelación nula, exhiben dependencias de orden mayor o dependencias no lineales, como sería el caso de algunos procesos bilineales, y aún más para modelos puramente deterministas caóticos entre otros. En general, los procedimientos basados en las autocorrelaciones fallan al utilizarse para especificar los retardos significativos en modelos no lineales, al no detectar relaciones no lineales importantes en los datos, ni los desfases temporales apropiados, especialmente en aquellos escenarios en donde los fenómenos no lineales son más la regla que la excepción. En esta investigación se propone como herramienta para el modelado, una nueva prueba estadística no paramétrica y un procedimiento, para identificar el retardo temporal relevante en la descripción de un modelo general lineal o no lineal para series de tiempo estacionarias. Estas técnicas están basadas en la estimación de la integral de correlación.

Palabras clave: Integral de Correlación, Retardo temporal, Series de tiempo, No Linealidad.

13.1 Introducción

La selección de los retardos que deben ser incluidos en la descripción de un modelo general de una serie de tiempo es ciertamente el primer paso crucial para la selección de un modelo, y es esencial en la predicción y desarrollo de modelos económicos. Tradicionalmente, para la especificación de los retardos apropiados, se utilizan los coeficientes de autocorrelación y de autocorrelación parcial. Sin embargo, está bien establecido en [Granger1983], que procesos con autocorrelación cero pueden exhibir alto orden de dependencia o dependencia no lineal. Este es el caso de algunos procesos bilineales e incluso de modelos deterministas caóticos puros, entre otros. En general, los procedimientos basados en la autocorrelación pueden ser engañosos para los modelos no lineales, y por lo tanto pueden no detectar relaciones no lineales importantes en los datos, siendo por tanto, limitada su 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. Es bien sabido que las variables económicas más importantes son de naturaleza no lineal y que abundan las relaciones no lineales entre ellas. Desde un punto de vista econométrico y estadístico, esta situación

ha motivado el desarrollo de pruebas de dependencia serial [Tjostheim1996].

Actualmente existen muy pocas técnicas 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 de esta relativa escasez son [Granger1994], [Tjostheim1994], [Granger2003], y [Matilla-García2009]. Una característica común de la mayoría de estas técnicas es el uso de las medidas basadas en la entropía para identificar el retardo correcto. En este artículo se toma un camino diferente y con base en la muy conocida integral de correlación, se desarrolla un nuevo procedimiento estadístico no paramétrico para identificar los retardos que se deben utilizar cuando se desarrollen modelos de naturaleza desconocida (lineales o no lineales). En adición, se diseña un bootstrap basado en una prueba estadística para seleccionar entre desfases potencialmente relevantes. Esta última contribución hace una diferencia interesante con otros procedimientos disponibles, que no proporcionan un marco de prueba de hipótesis.

13.2 Integral de Correlación

La integral de correlación, introducida por [Grassberger1983], e inicialmente diseñada para medir la dimensión fractal de datos caóticos, constituye una herramienta importante actual en las ciencias físicas y naturales para analizar algunas propiedades de las series de tiempo ([Figueiredo2011] y [Mathew2011]). Ha sido ampliamente utilizada en finanzas y macroeconomía básicamente para probar independencia serial. La contribución más relevante es la muy conocida prueba BDS [Brock1996], 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. Curiosamente, la integral de correlación también ha sido utilizada [Hiemstra1994] para detectar posibles relaciones no lineales de causalidad entre series de tiempo. Ha habido algunos comentarios seriamente críticos a la prueba BDS [Barnet1997], [Tjostheim1996], [De Lima1996], entre otros, en cuanto a una sobreestimación mayor del nivel correcto de la prueba, por lo tanto tiende a reportar una tasa alta de falsa alarma. Esto puede explicarse, en parte, debido al impacto no trivial de varias decisiones anteriores sobre diversos parámetros decisivos que deben tomarse para poder utilizar la prueba. La prueba que se propone en este artículo, a pesar de que se basa en la integral de correlación, es robusta para tomar aquellas decisiones previas, particularmente minimizando el papel que desempeñan cuando se selecciona un retardo adecuado en el modelo seleccionado. Estrechamente relacionada, otra propiedad especialmente relevante es que el procedimiento propuesto 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, está basada únicamente en la integral de correlación, que

originalmente requiere la incorporación de un parámetro de retardo de tiempo que plenamente describe 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 fue fijado en 1, 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. Otra contribución de esta investigación es que la técnica desarrollada 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, para ambos los macro-econometristas y los analizadores financieros.

13.3 Notaciones y Conceptos

Nuestro enfoque se basa en la reconstrucción de la dinámica asintótica de un sistema dinámico de las realizaciones de un proceso de series de tiempo observadas. La idea básica detrás de la reconstrucción del espacio de estados es que el pasado y el futuro de una serie de tiempo, ambos, contienen información acerca de las variables de estado inobservables y que pueden ser usadas para definir un estado en el tiempo presente. La reconstrucción se lleva a cabo a partir de una serie de tiempo escalar y todas las componentes relevantes (relacionadas con la dinámica subyacente) tienen que ser extraídas de ella. [Takens1981] mostró que este tipo de reconstrucción genera un atractor topológicamente equivalente que deja invariante la dinámica de los parámetros. La reconstrucción requerida puede empotrar observaciones univariadas en un espacio de fase multivariado. Al final, la información es encapsulada en el vector de retardo llamado la m -historia. Dado un sistema dinámico descrito por la ecuación en diferencias

$$x_{t+1} = F(x_t); x_t \in \Omega, \quad (13.3.1)$$

la reconstrucción de la dinámica supone que la evolución del sistema está limitado a un sub-espacio acotado de dimensión finita M del espacio de fase Ω , y que existe un función de medida suave h que depende únicamente del estado de la dinámica del sistema, esto es, que implícitamente depende del tiempo. Se asume por simplicidad que h es una función escalar real-valuada, esto es, h es un mapeo $h : M \rightarrow M$, tal que para cada t , la observación x_t es la imagen del estado $y_t \in M$, definida por $h(y_t)$. También $y_t \in M$ determina el estado futuro del sistema dinámico, por lo que un mapeo de dimensión m se define como:

$$\begin{aligned} \varphi_m & : M \rightarrow \mathbb{R}^m \\ \varphi_m & : y_t \mapsto (h(y_t), h(F(y_t)), \dots, h(F^{m-1}(y_t))) \end{aligned} \quad (13.3.2)$$

que mapea el estado y_t en el vector retardado

$$X_t^{m,\tau} = \{x_t, x_{t-\tau}, x_{t-2\tau}, \dots, x_{t-(m-1)\tau}\} \quad (13.3.3)$$

para un retardo de tiempo arbitrario, (13.3.3) es también conocida como la m -historia retardada.

El teorema de Takens establece que para funciones suaves como h y para una m suficientemente grande ϕ_m es generalmente un mapeo suave invertible de M a \mathbb{R}^m , por lo que los puntos representados por (13.3.3) son una fiel imagen del atractor en \mathbb{R}^m , conocido como el atractor reconstruido. Esta es una estructura geométrica que tiene una medida relacionada con la frecuencia relativa con la cual las diferentes partes del atractor son visitadas, y es invariante bajo el operador evolución. [Takens1981] mostró que se puede reconstruir la dinámica de un sistema determinista mediante una transformación paramétrica suave. Un punto central en este artículo es el concepto de la integral de correlación, denotada como $C_m^\tau(\varepsilon)$ e introducida por [Grassberger1983], la cual se construye a partir del espacio de reconstrucción retardado introducido en [Takens1981]. Dada una serie de tiempo $\{x_t\} = \{x_1, x_2, \dots, x_T\}$ de tamaño T , generada por un proceso de la forma:

$$x_t = G(X_{t-1}, \theta) + e_t, \quad (13.3.4)$$

con $t = 1, 2, \dots, T$.

Entonces se generan n vectores retardados de la forma (13.3.3), donde $n = T - (m-1)\tau$ es el número de vectores retardados en una dimensión de inmersión m , y un tiempo de retardo τ dado. Para cada valor de m , todas las n m -historias forman una nube de puntos en \mathbb{R}^m . La Integral de Correlación $C_m^\tau(\varepsilon)$ puede ser estimada por $\hat{C}_{m,n}^\tau(\varepsilon)$, donde

$$\hat{C}_{m,n}^\tau(\varepsilon) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n H(\varepsilon - \|X_i^{m,\tau} - X_j^{m,\tau}\|_\infty) \quad (13.3.5)$$

$$H(Z) = \begin{cases} 1 & \text{if } Z \geq 0 \\ 0 & \text{if } Z < 0 \end{cases}$$

$$\|X_i^{m,\tau} - X_j^{m,\tau}\|_\infty = \max_{0 \leq k \leq m-1} \{|x_{i+k\tau} - x_{j+k\tau}|\}.$$

$$X_i^{m,\tau} = \{x_i, x_{i+\tau}, x_{i+2\tau}, \dots, x_{i+(m-1)\tau}\}, \text{ la } i\text{-ésima } m\text{-historia of } x_t \quad (13.3.6)$$

Esta $\hat{C}_{m,n}^\tau(\varepsilon)$ es un estadístico U generalizado que ha sido ampliamente utilizado por [Brock1996], sin embargo, no será considerado de esta forma en este artículo. Solamente mencionamos este punto, en el caso de la existencia del siguiente límite,

$$C_m^\tau(\varepsilon) = \lim_{n \rightarrow \infty} \hat{C}_{m,n}^\tau(\varepsilon). \quad (13.3.7)$$

donde

$$C_m^\tau(\varepsilon) = \iint H(\varepsilon - \|X_i^{m,\tau} - X_j^{m,\tau}\|_\infty) dF_m(X_i^{m,\tau}) dF_m(X_j^{m,\tau}). \quad (13.3.8)$$

$\hat{C}_{m,n}^\tau(\varepsilon)$ es interpretada como la fracción del total del número de parejas de vectores retardados $(X_i^{m,\tau}, X_j^{m,\tau})$, cuya distancia es menor que ε en una inmersión del espacio. Esto significa que el método de estimación está basado en una selección aleatoria del punto $X_i^{m,\tau} \in \mathbb{R}^m$ (en el espacio fase de retardo reconstruido) y en el análisis de los cambios en el número de puntos dentro de una vecindad de radio ε la cual denotaremos como $V_\varepsilon(X_i^{m,\tau})$. La integral de Correlación puede ser entendida como una medida de la frecuencia con la cual la estructura de los datos es repetida (número de parejas de vectores retardados dentro de una vecindad de radio ε). Un gran número de parejas indica que se está en contra de la hipótesis de que los datos fueron generados por un proceso estocástico independiente e idénticamente distribuido (iid).

Sea $A \in \mathbb{R}^{n \times n}$ la matriz de la función indicadora $I(X_i^{m,\tau} - X_j^{m,\tau})$, donde $X_i^{m,\tau}$; $i = 1, 2, 3, \dots, n-1$; $X_j^{m,\tau}$; $j = i+1, i+2, \dots, n$, tal que el ij -ésimo elemento de la matriz A está definido por:

$$a_{ij} = \begin{cases} 1 & \text{if } \|X_i^{m,\tau} - X_j^{m,\tau}\|_\infty \leq \varepsilon \\ 0 & \text{other case} \end{cases}; \quad (13.3.9)$$

La matriz A indica si los m -vectores retardados $X_i^{m,\tau}, X_j^{m,\tau}$ están cercanos en \mathbb{R}^m , en una vecindad de radio ε . Así que la suma $N^{m,\tau}(\varepsilon)$ de los números de unos en los renglones y columnas de la matriz A determinan el número de vectores retardados que están cercanos dentro de una vecindad dada. Por tanto, la integral de correlación (13.3.5) puede escribirse como:

$$\hat{C}_{m,n}^\tau(\varepsilon) = \frac{2N^{m,\tau}(\varepsilon)}{n(n-1)} = \frac{N^{m,\tau}(\varepsilon)}{\binom{n}{2}}. \quad (13.3.10)$$

Consecuentemente la expresión anterior es la proporción del número de vectores retardados que están dentro de la vecindad.

Como

$$\hat{C}_{m,n}^\tau(\varepsilon) = \lim_{n \rightarrow \infty} \frac{2N^{m,\tau}(\varepsilon)}{n(n-1)}, \quad (13.3.11)$$

entonces $C_m^\tau(\varepsilon)$, que representa una medida de reconstrucción μ_m asociada con los vectores de retardo de dimensión m , es definida como la probabilidad de que la distancia entre una pareja de vectores retardados arbitrariamente seleccionados, con respecto a μ_m , es menor que ε , por lo que la integral de Correlación proporcióna la probabilidad de que cualesquiera dos m -historias particulares estén

cerca:

$$\hat{C}_{m,n}^\tau(\varepsilon) = P[\|X_i^{m,\tau} - X_j^{m,\tau}\|_\infty \leq \varepsilon, \|X_i^{m,\tau} - X_{j+1}^{m,\tau}\|_\infty \leq \varepsilon, \dots, \|X_{i+1}^{m,\tau} - X_j^{m,\tau}\|_\infty \leq \varepsilon, \dots, \|X_{T-1}^{m,\tau} - X_T^{m,\tau}\|_\infty \leq \varepsilon], \quad (13.3.12)$$

13.4 Identificación del retardo óptimo para una serie de tiempo dada

En esta sección se muestra teórica y experimentalmente un procedimiento para identificar el retardo relevante de una serie de tiempo generada mediante el proceso de la forma (13.3.4). Una ventaja atractiva es que, debido a su propio diseño, puede ser utilizado para una amplia clase de procesos. Supóngase que entre un conjunto de posibles retardos, existe un retardo óptimo τ_o en el sentido de que ese desfase reconstruye mejor las propiedades dinámicas de los datos de generación de propiedades. De acuerdo a las ecuaciones (13.3.10) y (13.3.12), la probabilidad de que la distancia entre una pareja de m -historias seleccionada aleatoriamente con el retardo óptimo es menor que ε , será mayor que la probabilidad que se obtiene considerando la distancia entre un par de vectores con los retardos temporales no óptimos. En consecuencia, el siguiente teorema:

Teorema 4.1 Sea $\{x_t\}_{t=1}^T$ una serie de tiempo generada por un modelo de la forma $x_t = G(X_{t-1}, \theta) + e_t$. Se denota como $\hat{C}_{m,n}^{\tau_o}(\varepsilon)$ a la estimación de la integral de correlación definida por la ecuación (13.3.5), dando una dimensión de inmersión $m \geq 2; m \in N$ y un valor $\varepsilon > 0$. Si el tiempo de retardo óptimo es τ_o para el modelo, entonces

$$\hat{C}_{m,n}^{\tau_o}(\varepsilon) = \max\{\hat{C}_{m,n}^\tau(\varepsilon) | \tau \in \mathbb{N}\}$$

Esto es, la integral de correlación dará su valor máximo en el retardo correcto, cuando se compara con los retardos de tiempo no óptimos. La probabilidad de que cualquier par de vectores en la serie de tiempo estén cerca, cuando se usa un retardo diferente al óptimo, $\tau \neq \tau_o$, implicará que alguna información importante sea perdida, y entonces será no óptima para detectar la estructura dinámica correcta de la serie de tiempo generada con el proceso subyacente. En consecuencia, este resultado puede ser útil para seleccionar el retardo de tiempo adecuado para el análisis de las propiedades del proceso. Los siguientes comentarios son importantes. El comportamiento empírico de obtener el máximo valor del promedio estimado de la correlación integral para el retardo más relevante $\tau \neq \tau_o$ de un modelo dado es estudiado a través de simulación Monte Carlo, considerando los modelos que aparecen en la Tabla 13.1:

El primer modelo de la Tabla 13.1 corresponde a un proceso de ruido blanco normal estándar, $N(0, 1)$. Este modelo puede ser considerado como una referencia para esta técnica, ya que todos los retardos proporcionan la misma información

Table 13.1: Modelos de Procesos Estocásticos generadores de Series de Tiempo.

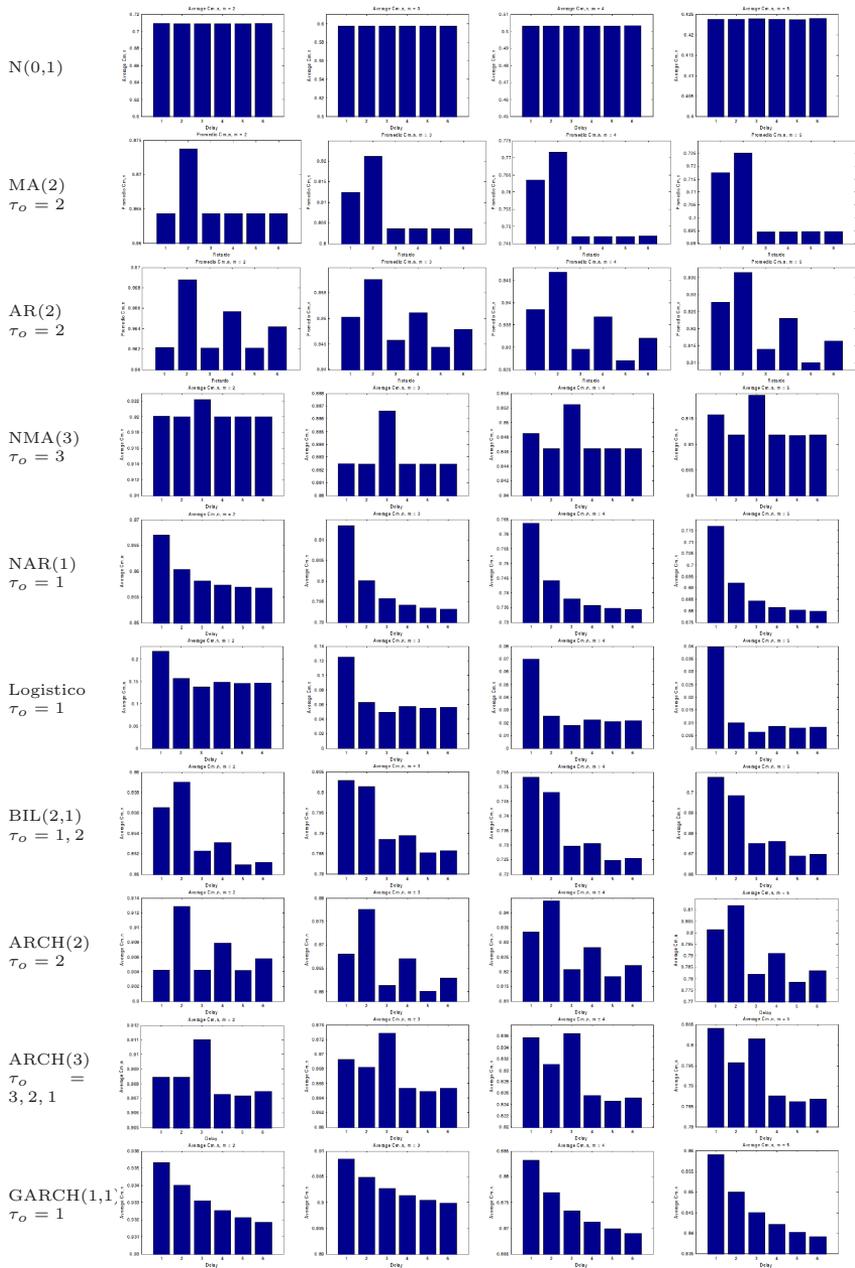
Modelos	
Modelo 1: N(0,1)	$X_t = e_t \sim N(0, 1)$
Modelo 2: MA(2)	$X_t = e_t + 0.8e_{t-2}; e_t \sim N(0, 1)$
Modelo 3: AR(2)	$X_t = e_t + 0.8X_{t-2}; e_t \sim N(0, 1)$
Modelo 4: NMA(3)	$X_t = e_t + 0.8e_{t-3}^2; e_t \sim N(0, 1)$
Modelo 5: NAR(1)	$X_t = X_{t-1} ^{0.8} + e_t; e_t \sim N(0, 1)$
Modelo 6: Logistic	$X_t = 4X_{t-1}(1 - X_{t-1});$
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)	$X_t = \sqrt{h_t}e_t; h_t = 0.7 + 0.3X_{t-1}^2 + 0.4h_{t-1};$ $e_t \sim N(0, 1)$

a la integral de correlación. Por lo tanto, no se espera encontrar un retardo relevante. Después se estudiaron varios modelos estacionarios: en primer lugar nos centramos en el comportamiento del procedimiento basado en la correlación integral para seleccionar el desfase óptimo para los procesos lineales, $MA(2)$ y $AR(2)$. Dado que nuestro procedimiento puede ser utilizado bajo estructuras no lineales, incluimos en el estudio dos procesos no lineales, denominados, $NMA(3)$ y $NAR(1)$. El modelo autorregresivo caótico logístico y el modelo bilineal han sido incluidos en el estudio, ya que han sido ampliamente examinados, dado que son ambos de distinta naturaleza (determinista y estocástico) los métodos clásicos de identificación no son capaces de identificarlos, [Granger1983]. 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 hemos estudiado tres modelos diferentes de tipo GARCH con varios retardos.

Se consideró un tamaño de muestra de 500 observaciones para cada modelo presentado en la Tabla 13.1. Se estimó la integral de correlación con los parámetros de inmersión: $m = 2, 3, 4, 5$, $\tau = 1, 2, 3, 4, 5, 6$, y $\varepsilon = 1.96\sigma_x$. Se calculó el promedio de $\hat{C}_m^\tau(\varepsilon)$ 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 $\tau = 1, 2, 3, 4, 5, 6$.

La Tabla 13.2 resume los resultados para los modelos. Cada subgráfica muestra el promedio para una dimensión de inmersión dada y seis distintos retardos.

Table 13.2: Promedios de la Integral de Correlación para procesos estacionarios, $\hat{C}_{m,n}^T$



Antes de comentar cada modelo, vale la pena darse cuenta que una propiedad común e 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, porque entonces no es crucial para los resultados la selección del parámetro m . Como se va a comentar más adelante en el documento, esta propiedad, que estaba prevista en los comentarios que figuran en la sección anterior y ahora se confirman empíricamente, permiten evitar el problema común de la maldición de dimensionalidad.

Las gráficas del promedio de la integral de correlación 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 nuestro procedimiento basado en la integral de correlación: La integral de Correlación para los retardos arriba de 2 no cambian, mientras que la propiedad de memoria descendiente del $AR(2)$ se observa claramente para el orden de retardo por encima de 2.

Para los siguientes dos modelos no lineales, 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 es detectado en el retardo correcto, y los mismo ocurre para el proceso bilineal, para el cual dos retardos (2 y 1) son relevantes. En presencia de estructura de varianza condicional, nuestro 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 $GARCH$, para el cual, la dependencia es encontrada, como era de esperarse, en el retardo 1.

13.5 Una prueba para identificar retardos en modelos lineales y no lineales

Como se había mencionado en la sección de introducción, 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. Se ha mostrado en las secciones anteriores 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. Por tanto, su naturaleza potencial desconocida no es una limitación de la técnica. En esta sección nos basamos en la integral de correlación evaluada en diferentes retardos para construir un procedimiento de prueba bootstrap que ayuda al in-

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 con un procedimiento estadístico. Con este fin, se construye la siguiente hipótesis nula:

Ho: El retardo temporal 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 la sección anterior se mantiene, bajo la hipótesis nula que:

$$\delta(\tau) = \hat{C}_{m,n}^{\tau_o}(\varepsilon) - \hat{C}_{m,n}^{\tau}(\varepsilon) = 0, \quad (13.5.1)$$

mientras que lo contrario ocurrirá en la alternativa. Considérese ahora que la hipótesis nula es rechazada, y por tanto se puede identificar potencialmente un modelo. Con este objetivo, 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 τ_o es estadísticamente diferente de cualquier otro retardo potencial. Para ello, se puede construir una prueba de tipo bootstrap, como sigue:

Paso 1.- Considérese 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 τ_o tal que maximiza la integral de correlación de la serie $\{X_t\}_{t=1}^T$, para m y una ε fijas.

Paso 3.- Estimar $\hat{C}_{m,n}^{\tau}(\varepsilon)$ de la serie $\{X_t\}_{t=1}^T$, para m y una ε 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 ¹, $\{X_t\}_{t=1}^T(b)$ donde b indica que nos estamos refiriendo a la muestra bootstrap.

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

- $\hat{C}_{m,n}^{\tau_o,b}(\varepsilon)$
- $\hat{C}_{m,n}^{\tau,b}(\varepsilon)$, for $\tau = 1, 2, \dots$, with $\tau \neq \tau_o$
- $\hat{\delta}^b(\tau)$, (13.5.1) de la serie bootstrap

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

¹Se sigue a Politis y White (2004) para seleccionar automáticamente el movimiento de los bloques

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

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

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

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

13.5.1 Rendimiento empírico de la prueba: tamaño y potencia

Se usaron los modelos presentados en la Tabla 13.1 para evaluar el rendimiento de nuestra prueba con 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 $T_1 = 500$ 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 $N_r = 1000$.

Para los experimentos del tamaño de la prueba se consideraron particularmente tres modelos: uno de distribución normal i.i.d., para el cual todos los retardos son igualmente no importantes; un modelo de promedio móvil lineal de orden 2, el cual no tiene propiedades de memoria para retardos mayores a 2; y para completar un modelo de promedio móvil no lineal NMA(3) sin propiedades de memoria a partir del retardo 3. Las Tablas 13.3, 13.4 y 13.5 muestran los tamaños empíricos para diferentes configuraciones de la hipótesis nula. Las tablas deben ser leídas por renglón: Por ejemplo, el primer renglón con resultados en la Tabla 13.3 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, \dots , y 5. 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 estamos en el nivel nominal de rechazo esperado. Tablas 13.4 y 13.5 deben ser leídas de la misma forma que la anterior. Ahora se están utilizando modelos con dependencia lineal y no lineal. En ambos modelos no existe una relación teórica más allá de cierto retraso. En consecuencia, para ambos modelos, probamos diferentes configuraciones posibles de la hipótesis nula: En la Tabla 13.4, se probó el retardo 3 contra los retardos 4, 5 y 6; el retardo 4 contra los retardos 3, 5 y 6 y así sucesivamente. Los resultados muestran que para el nivel nominal seleccionado de la prueba, su desempeño bajo la hipótesis nula es correcta.

La conclusión anterior 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 [Kanzler1999].

Table 13.3: Tamaño de la prueba para el Modelo 1 con $T = 500$ y 1000.

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

Table 13.4: Tamaño de la prueba para el Modelo 2 con $T = 500$ y 1000.

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

Table 13.5: Tamaño de la prueba para el Modelo 4 con $T = 500$ y 1000 .
$$X_t = e_t + 0.8e_{t-3}^2; e_t \sim N(0, 1)$$

τ	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	

Las Tablas 13.6 y 13.7 presentan los resultados de los experimentos de la potencia de la prueba. Ambas 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 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. El rendimiento de la potencia es extremadamente bueno, a pesar del tamaño de la 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 i) 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, por lo que los métodos basados en correlaciones clásicas dará conclusiones erróneas. (ii) Algo similar ocurre con el modelo logístico, que es puramente determinista; y (iii) como se informó en [Granger1983], la prueba de Kendal para τ proporciona malos resultados para el proceso NAR. Para el modelo no lineal de media móvil generado con retardo de orden 3, la potencia de la prueba es relativamente buena para el mayor tamaño de la muestra, mientras que los resultados son más discretos cuando se consideran 500 observaciones. Estos resultados indican que la prueba propuesta es especialmente útil en situaciones de incertidumbre acerca de la naturaleza no lineal de la media condicional del proceso lineal o no lineal.

Table 13.6: Potencia de la prueba para series generadas de tamaño $T_1 = 500$.

Model	Lag					
	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

Table 13.7: Potencia de la prueba para series generadas de tamaño $T_2 = 1000$.

Model	Lag					
	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

13.5.2 Uso de la prueba como una herramienta de diagnóstico

Es una práctica común en los procedimientos de pruebas de diagnóstico, que en ocasiones los investigadores están interesados en eliminar, a través de un filtro lineal, la dependencia lineal (en la media condicional) que pueda estar presente en los datos. Del mismo modo que en las secciones anteriores, se considera el caso en que el investigador necesita identificar la estructura dinámica del retardo en modelos que presentan algún tipo (lineal o no lineal) de dependencia en los momentos de condicionalidad de orden mayor. Si existe una dependencia no lineal en los residuos, nuestra prueba 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 en bruto, 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 por 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 son libres de parámetros de ruido, han sido estudiadas por [De Lima1996]. Para asegurar esta propiedad de invariancia, se encuentra que el modelo tiene que ser aditivo, tal como, $y_t = f(Y_{t-1}, \beta) + e_t$, o para los modelos que se pueden transformar en este formato. Por el contrario, nuestra prueba de bootstrap no está limitada por esas condiciones suficientes. Por lo tanto, consideramos un modelo general de error de la forma $y_t = f(Y_{t-1}, \beta, e_t)$, donde β es un vector de parámetros consistentemente estimados, y un término de error. Indirectamente, se supone que la función f se puede invertir, $e_t = G(y_t, Y_{t-1}, \beta)$, de modo que la técnica de identificar el retardo significativo puede ser aplicada a los términos del verdadero error. Las condiciones para la estimación consistente de la $\hat{e}_t = \hat{G}(y_t, Y_{t-1}, \hat{\beta})$ y para la estimación consistente de las integrales de correlación son los mismos que en aquellas de [De Lima1996].

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

$$x_t = \mu(X_{t-1}, \beta) + \sigma(X_{t-1}, \gamma)e_t. \quad (13.5.2)$$

Donde β y γ 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 (13.5.2) para obtener la serie residual estimada $\hat{e}_t = x_t - \hat{x}_t$. De la serie $\{e_t\}$, se calcula la serie de residuales al cuadrado $\{\hat{e}_t^2\}$. Se llevó a cabo la validación empírica del rendimiento de la técnica aplicada a la serie de residuos cuadrados, de manera similar a la presentada en la sección 13.4. Los modelos ARMA-GARCH y los filtros de ARMA (p, q) utilizado para generar la serie de tiempo y estimar los residuales son:

Table 13.8: Modelos para generar $\{X_t\}$ y filtro para los residuales $\{\hat{e}_t\}$.

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)	$X_t = 0.1 + 0.5X_{t-1} + e_t;$ $e_t = \sqrt{h_t}z_t; z_t \sim N(0, 1)$ $h_t = 0.7 + 0.3e_{t-1}^2 + 0.4h_{t-1}$	ARMA(1,0)
ARMA(0,1)-GARCH(1,1)	$X_t = 0.1 + 0.0X_{t-1} - 0.3e_{t-1} + e_t$ $e_t = \sqrt{h_t}z_t; z_t \sim N(0, 1)$ $h_t = 0.7 + 0.3e_{t-1}^2 + 0.4h_{t-1}$	ARMA(0,1)
ARMA(1,1)-GARCH(1,1)	$X_t = 0.1 + 0.5X_{t-1} - 0.3e_{t-1} + e_t$ $e_t = \sqrt{h_t}z_t; z_t \sim N(0, 1)$ $h_t = 0.7 + 0.3e_{t-1}^2 + 0.4h_{t-1}$	ARMA(1,1)

Como se muestra en la Tabla 13.8, el primer modelo corresponde al proceso de ruido blanco i.i.d. normal estándar $N(0,1)$ que ha sido considerado para evaluar el tamaño de la prueba. Los resultados del estadístico del tamaño de la prueba son presentados en la Tablas 13.9, 13.10 y 13.11. 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 $T_1 = 500$, $T_2 = 1000$ y $T_3 = 2000$, y el número de muestras bootstrap en cada caso fueron $B = 199$. El número de réplicas de la simulación fue de $N_r = 1000$. Estas tablas indican la proporción de rechazo bajo la hipótesis nula. Las tablas 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 en el renglón son los valores esperados para los que se aplicó la técnica de identificación². Nótese que este buen comportamiento empírico para el tamaño es independiente del tamaño de la muestra.

²Los valores están dentro del rango [0.036, 0.066] para $\alpha = 0.05$.

Table 13.9: Tamaño de la prueba de la serie residual con el modelo ARMA(1,1)-N(0,1) Tamaño $T_1 = 500$.

τ_o	τ					
	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	

Table 13.10: Tamaño de la prueba de la serie residual con el modelo ARMA(1,1)-N(0,1) Tamaño $T_2 = 1000$.

τ_o	τ					
	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	

Table 13.11: Tamaño de la prueba de la serie residual con el modelo ARMA(1,1)-N(0,1), Tamaño $T_3 = 2000$.

τ_o	τ					
	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	

Las Tablas 13.12-13.14 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 13.6, para los tamaños de muestra $T_1 = 500$, $T_2 = 1000$ y $T_3 = 2000$. El retardo significativo en los tres modelos es $\tau_o = 1$.

Se puede ver 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.

Table 13.12: Potencia de la prueba en la serie de residuales del modelo ARMA(1,0)-GARCH(1,1).

	τ					
	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

Table 13.13: Potencia de la prueba en la serie de residuales del modelo ARMA(1,0)-GARCH(1,1).

	τ					
	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

Table 13.14: Potencia de la prueba en la serie de residuales del modelo ARMA(1,0)-GARCH(1,1).

	τ					
	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

13.6 Conclusiones y recomendaciones finales

De los resultados teóricos y empíricos presentados, se puede concluir que la técnica de identificación del tiempo de retardo significativo de series de tiempo no lineales puede ser aplicada a series de tiempo estacionarias originales (series generadas por los modelos presentados en la Tabla 13.1), y series de residuales (series

obtenidas después de haberse aplicado a una serie original un filtro lineal del tipo ARMA(p,q), ya que en ambos casos se obtuvieron resultados teóricos asintóticos. En el caso de las series de residuos esta técnica debe ser usada con series con un número de observaciones mayor a 1000, ya que la convergencia demanda una cantidad mayor de datos.

El procedimiento de la técnica propuesta es fácil de aplicar, ya que se trata de aplicar el siguiente algoritmo:

- 1) Proporcionar como dato inicial la serie de tiempo observada y los parámetros de inmersión m (dimensión de inmersión) y ε (distancia entre las parejas de vectores en \mathbb{R}^m).
- 2) Generar el conjunto de vectores retardados (las m -historias de la serie) requeridas para el cálculo del estimador de la integral de correlación, considerando varios valores para el tercer parámetro de inmersión $\tau \in \mathbb{Z}^+$.
- 3) Comparar los valores de la integral de correlación, para identificar el retardo en el cual se alcanza su valor máximo.

La implantación computacional de este algoritmo es simple. Para llevar a cabo las simulaciones que sirvieron para validar la técnica propuesta, se desarrollaron programas que se implantaron en MATLAB, utilizando las funciones de simulaciones Monte Carlo y técnicas de Bootstrap que tiene.

Bibliografía

- [Barnet1997] Barnett William A. & Gallant, A. Ronald & Hinich, Melvin J. & Jungeilges, Jochen A. & Kaplan, Daniel T. & Jensen, Mark J., **A single-blind controlled competition among tests for nonlinearity and chaos**, Journal of Econometrics, 82(1),157-192
- [Brock1996] Brock, W., Dechert, D., LeBaron, J., Scheinkman, B., & Dechert, W. **A Test for Independence Based on the Correlation Dimension**. Wisconsin, USA: Department of Economics, Univerty of Wisconsin.
- [De Lima1996] De Lima, P. **Nuisance Parameter Free Propereties of Correlation Integral Based Statistics**, Econometrics Reviews, 15(3), 237-259.
- [Figueiredo2011] Figueiredo C., Diambra, L., & Pereira, C. (2011). **Criterion of Numerical Chaotic Solutions Based on Statistical Measures**, Applied Mathematics, 2, 436-443.
- [Granger1994] Granger, C., & Lin, J. **Using the Mutual Information Coefficient to Identify Lags in Nonlinear Model**, Journal of Time Series Analysis, Vol 15, Issue 4, 371-384.
- [Granger1983] Granger, C., & Weiss, A. **Time Series Analysis of Error-Correcting Models**, Studies in Econometrics, Time Series, and Multivariate Statistics, New York, Academic Press, 255-278.
- [Granger2003] Granger, C., Masoumi, E., & Racine, J. **A Dependence Metric for Possibly Nonlinear Processes**, San Diego, California: Department of Economics, University of California.
- [Grassberger1983] Grassberger, P., & Procaccia, I. (1983). **Measuring the Strangeness of Strange Attractors**, Physica 9D, 189-208.
- [Hiemstra1994] Hiemstra, C., & Jones, J. **Testing for Linear and Nonlinear Granger Causality in the Stock Price-Volume Relation**, Journal of Finance, vol. 49, issue 5, 1639-1664.

- [Kanzler1999] Kanzler, L. (1999). **Very Fast and Correctly Sized Estimation of the BDS Statistic**, Oxford, England: Christ Church and Department of Economics University of Oxford.
- [Mathew2011] Mathew, N., & Picu, R. (2011). **Molecular conformational stability in cyclotrimethylene trinitramine crystals**, Journal of Chemical Physics 135 (2).
- [Matilla-García2009] Matilla-García, M., & Ruiz Marin, M. Detection of non-linear structure in time series. Economic Letters, 105, 1-6.
- [Takens1981] Takens, F. **Detecting Strange Attractors in Turbulence**, Lecture Notes in Mathematics, 898, 355-381.
- [Tjostheim1996] Tjostheim, D. (1996). **Measures and tests of independence: a survey**, Statistics, 28, 249-284.
- [Tjostheim1994] Tjostheim, D., & Auestad, B. (1994). **Nonparametric identification of nonlinear time series: selecting significant lags**, Journal of the American Statistical Association, 89., 1410-1419.