

5. DETECCIÓN DE CAOS EN SERIES TEMPORALES.

5.1. INTRODUCCIÓN.

5.2. ANÁLISIS GRÁFICO.

5.3. COEFICIENTE DE HURST.

5.4. RECONSTRUCCIÓN DEL SISTEMA.

5.4.1. DIMENSIÓN DE CORRELACIÓN.

5.4.2. EXPONENTES DE LYAPUNOV.

5.4.3. ENTROPÍA DE KOLMOGOROV.

5.4.4. CONTRASTE BDS.

5.5. CONCLUSIONES.

5.1. INTRODUCCIÓN.

En el capítulo cuarto hemos propuesto modelos de tipo ARIMA y GARCH sobre las series temporales de rendimientos del contrato de Futuro sobre el Bono Nocional a 10 años con periodicidades de cinco minutos y de operación a operación, y de tipo ARFIMA de las series de duraciones de operaciones. Las medidas de detección de caos que se van a usar no siempre son capaces de diferenciar entre dependencias caóticas y dependencias producidas por otros modelos. Si los modelos propuestos recogen las dependencias existentes, sus residuos no tendrán ningún tipo de pauta. Si los datos siguen una pauta caótica, las series de residuos seguirán conteniendo dependencias que se tratarán de encontrar a lo largo del presente capítulo.

En el capítulo tercero planteábamos la posibilidad de que lo que podía parecer una variable aleatoria en realidad no lo fuera, sino que se tratara de una variable determinista de tipo caótico.

La duda que se nos presenta ahora es cómo saber si estamos ante una variable de tipo aleatorio o caótico. Si es del primer tipo, nuestro análisis debe terminar en este momento, pero si es caótica podemos continuar reduciendo la incertidumbre.

Lo primero que vamos a realizar es el análisis de los gráficos de recurrencias y el análisis de memoria a largo plazo (coeficientes de Hurst). Los primeros permiten identificar órbitas periódicas inestables, que son un signo inequívoco de comportamiento caótico. La memoria a largo plazo, a pesar de no ser definitoria del comportamiento caótico, sí se presenta en muchos casos, por lo que puede ser un indicio más.

En nuestro análisis partimos de una serie temporal que hemos obtenido a partir de un sistema dinámico del que desconocemos sus características. Existen toda una serie de medidas que requieren para su aplicación la reconstrucción de ese sistema original. Estas van a ser la **dimensión de correlación**, los **exponentes de Lyapunov** o la **entropía de Kolmogorov**. El análisis se cierra con el estadístico **BDS** que permite realizar un contraste sobre la presencia de no linealidad vinculada a caos.

La elección de los métodos de detección de comportamiento caótico no pretende ser exhaustiva. Hemos preferido aquéllos que se encuentran más contrastados y que han

logrado mayores éxitos, o que aportan una información interesante que no se obtiene con otros. En los últimos años se han propuesto otras medidas pero su uso práctico en estudios de búsqueda de caos ha sido escaso. Éstos son los casos del test de Kaplan¹ o el de Wang y Linton².

5.2. ANÁLISIS GRÁFICO.

El primer paso es buscar indicios que justifiquen la búsqueda de comportamientos caóticos. La aproximación gráfica más simple al problema la podemos encontrar en Edgard Peters³. Peters señala que las series de precios de los activos financieros tienen gráficamente la misma estructura, sea cual sea la escala temporal a la que la estudiemos (figura 5.2.1). Según Peters esta estructura se debe a los distintos horizontes temporales con los que operan los agentes que intervienen en un mercado financiero. Esto es a lo que Peters denomina hipótesis de mercado fractal. Dejando aparte la interpretación de Peters, sí resulta cierto que es difícil encontrar diferencias entre las gráficas de la figura 5.2.1 que permitan establecer cual es el horizonte temporal al que se corresponde esa gráfica. El hecho de que un objeto tenga la misma apariencia en distintas escalas es un indicio de que estamos ante un fractal.

¹ Kaplan, 1994.

² Wang y Linton, 1999.

³ Edgard E. Peters, 1994.

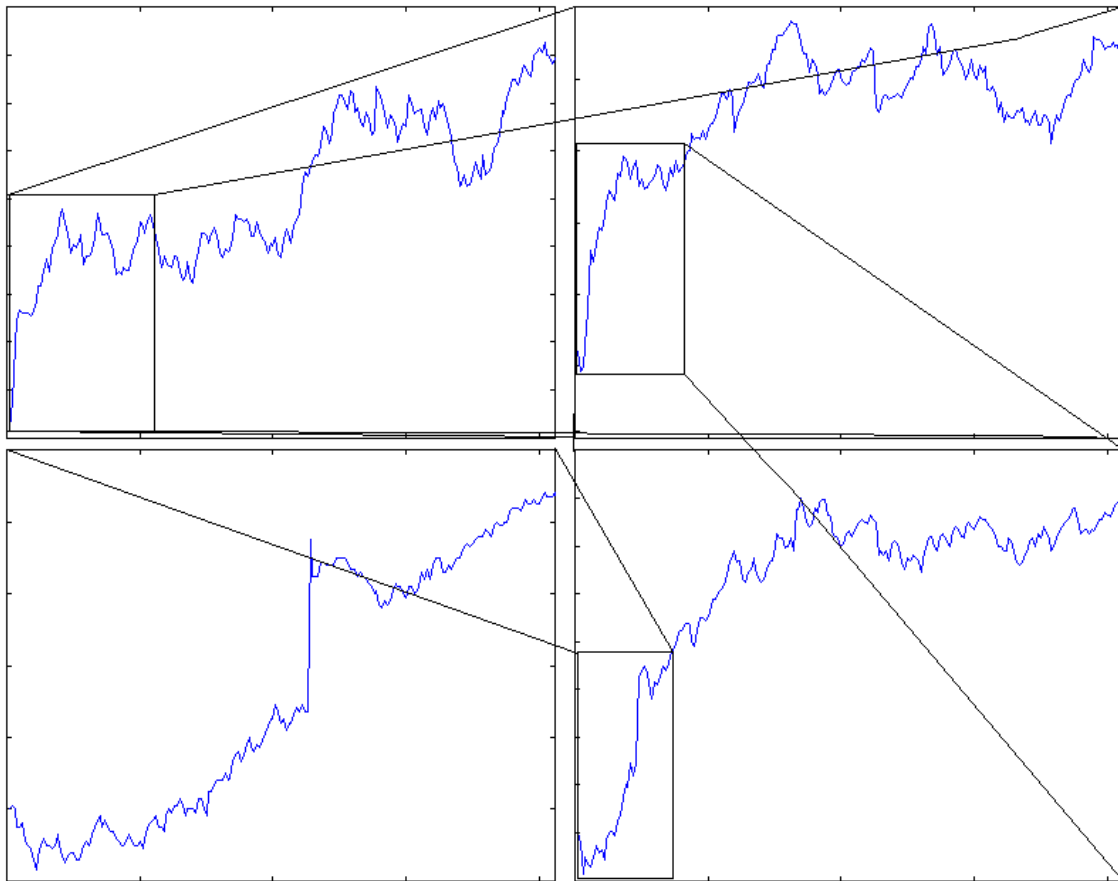


Figura 5.2.1: Cotizaciones del contrato de Futuro sobre el Bono Nocional a 10 años con vencimiento en marzo 98. Sucesivas ampliaciones en las que se ha pasado de considerar el precio de 1000 en 1000 operaciones a 250, 50 y 10 operaciones.

Sin embargo, esta característica encontrada en nuestras series no es garantía de estar ante un proceso caótico, o no lineal. El mismo Mandelbrot, en su libro *Geometría Fractal de la Naturaleza*⁴, al describir el movimiento browniano, observa esa misma característica. Puesto que el movimiento browniano es un fenómeno aleatorio y no caótico, esto nos impide utilizarlo como criterio para identificar comportamientos caóticos.

En la figura 5.2.1, las ampliaciones se realizan respecto al tiempo. En la figura 5.2.2 tenemos una nueva representación de los residuos del modelo que habíamos usado

⁴ Benoit, B. Mandelbrot, 1982.

de la serie de rendimientos operación a operación del Futuro sobre el Bono Nocional a 10 años con vencimiento en marzo del 98⁵. En este nuevo gráfico, las ampliaciones se hacen respecto al nivel de precisión del valor. En dos dimensiones, se observa como la estructura de rombo de la figura completa se reproduce en los pequeños cuerpos que la componen, y como una nueva ampliación de estos permite distinguir indicios de que la estructura se repite, aunque es difícilmente distinguible por la falta de puntos para una correcta definición.

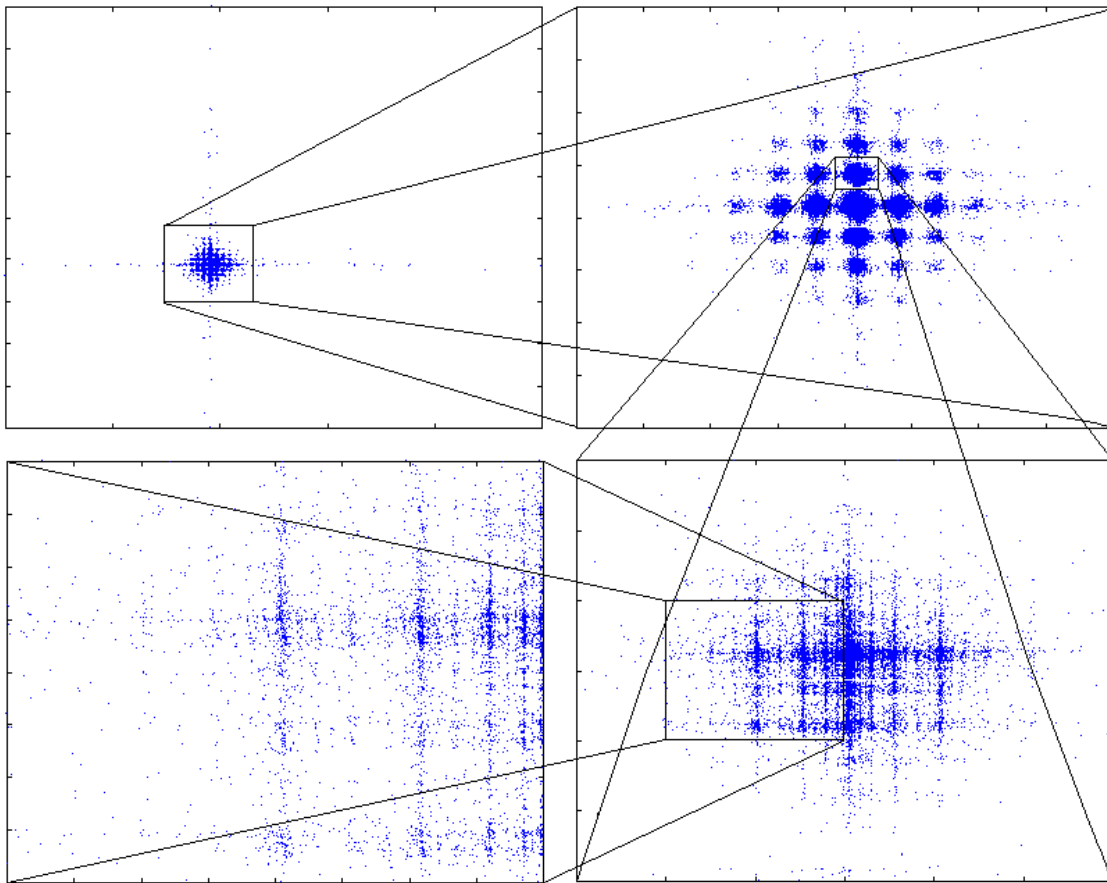


Figura 5.2.2: Gráfica de los residuos del modelo ARMA(1,1) sobre la serie de rendimientos del contrato de Futuro sobre el Bono Nocional a 10 años con vencimiento en marzo 98 y frecuencia operación a operación. En el eje x representamos a x_t y en el eje y a $x_{t+\tau}$.

⁵ Los gráficos del resto de las series obtenidas se pueden ver en el apéndice gráfico, figuras de la A.113 a la A.157.

La falta de puntos que nos impide seguir ampliando el gráfico la volveremos a encontrar cuando intentemos calcular la dimensión de correlación en el apartado quinto de este capítulo⁶.

En un principio la figura 5.2.2 parece indicar la presencia de un fractal. Las series de rendimientos sin embargo tienen una estructura que podemos denominar de rejilla, tanto las de rendimientos como las de duraciones. Esto se debe a que las cotizaciones no pueden tomar todos los valores posibles, sino sólo aquellos permitidos por el valor mínimo del cambio de precio (“tick”), que en el Futuro es el 0’01% del valor nominal del contrato. Lo mismo sucede con las series de duraciones, donde la precisión máxima que obtenemos es de un segundo. Esta estructura es la señalada por Szpiro⁷. El hecho de que exista esta estructura no es indicativo de nada, pues la serie se distribuye uniformemente dentro de los valores que se le permite tomar. Las gráficas del apéndice están elaboradas con el 90% de los datos, dejando fuera el 10% más extremo. Es interesante el hecho de que en el caso de las series de rendimientos con frecuencia de operación a operación, el 90% de los puntos se encuentre concentrado en tan sólo 9 puntos distintos mientras que en la frecuencia de cinco minutos ese 90% supone 35 valores distintos. Al agrupar la serie estamos reduciendo los casos en los que apenas ha habido cambios, que son los que dominan las series con frecuencia operación a operación.

La estructura de rejilla de las series de rendimientos se ve modificada a medida que vamos utilizando distintos modelos. Los residuos de los modelos ARIMA no sólo

⁶ Éste va a ser el límite al que llegaremos para la precisión L .

⁷ Szpiro, 1998.

difuminan los puntos de la rejilla sino que reflejan una estructura de tipo fractal. Los modelos GARCH nos generan gráficas en las que tenemos radios que se alejan del origen. Por último los modelos de tipo ARFIMA nos dan una estructura en forma de “L”. Las formas de estas gráficas muestran estructuras no lineales imposibles de detectar en las gráficas de rendimientos. Podríamos pensar que estas no linealidades, bien estaban presentes en las series pero ocultas por el “tick”, bien han sido generadas por los filtros utilizados. La segunda explicación no encaja por lo menos para los residuos de los modelos ARIMA, puesto que son modelos lineales, y no pueden provocar comportamientos no lineales. Luego, podemos suponer que existen comportamientos no lineales en las series.

La lineal de avance que representan las gráficas de las figuras 5.2.1, en una dimensión, y 5.2.2, en dos dimensiones, queda cortada al no poder proseguir con dimensiones más altas. Pero el análisis de recurrencias abre nuevas posibilidades para el análisis.

Al definir un atractor extraño decíamos que era denso en órbitas periódicas inestables. El análisis de estas órbitas inestables es una forma de estudiar las características de un atractor extraño⁸. Gilmore⁹ propone buscar estas órbitas periódicas inestables de forma gráfica sin necesidad de estadístico: elaboramos una gráfica a partir de la serie temporal en la que vamos a buscar dichas órbitas. Una vez que hayamos encontrado rastros de éstas, podremos afirmar que la serie es caótica.

⁸ Auerbach, Cvitanovic, Eckmann, Gunaratne y Procaccia, 1987; Mindlin, Hou, Solari, Gilmore y Tufillaro, 1990.

⁹ Gilmore, 1993.

Partimos de una serie temporal¹⁰ $\{x_t\}_{t=1\dots N}$, en la que queremos estudiar si ha sido generada por un proceso caótico. En caso de proceder de un atractor extraño, habrá observaciones que se encuentren próximas a una órbita periódica. En ese caso encontraremos que las siguientes observaciones también se encontrarán próximas a esa órbita, por lo menos antes de alejarse como consecuencia de la sensibilidad a las condiciones iniciales. Lo primero que haremos es calcular todas las distancias entre las observaciones de la serie

$$|x_i - x_{i+k}| - l \tag{5.2.1}$$

Si la diferencia [5.2.1] tiene un valor negativo eso indicará que ambas observaciones se encuentran a una distancia menor que l . Si para los momentos $i+1$ y siguientes, la distancia continúa siendo inferior a l , querrá decir que estas observaciones se encuentran en las proximidades de una órbita periódica inestable de periodo k .

Para detectar estos pares próximos consecutivos, elaboramos un gráfico. En el eje de abscisas se representan las observaciones i . En el eje de ordenadas se representa el retardo entre las observaciones, k . Este retardo nos informa de la periodicidad de esta órbita inestable.

¹⁰ Existen autores que proponen reconstruir el sistema antes de realizar gráficos de recurrencias, como es el caso de Grau, 1996, pero nosotros vamos a seguir a Iwanski y Bradley, 1998, que señalan que los gráficos de recurrencias no necesitan reconstruir el sistema para ser efectivos.

Cada par de valores (i,k) de la gráfica se corresponderá a un par de observaciones (x_i, x_{i+k}) . Si la distancia entre ambas observaciones es menor que l , o lo que es lo mismo, [5.2.1] tiene un valor negativo, el punto (i,k) correspondiente se representa en negro. Si la distancia es mayor, el punto aparece en blanco.

La existencia de órbitas periódicas inestables de periodo k se detectará por la presencia de líneas horizontales en la gráfica para el valor k en el eje de ordenadas.

Para la elaboración del gráfico es necesario fijar el valor del parámetro l . Resulta crucial para el éxito del análisis la elección de la precisión, l , correcta. Si el l elegido es demasiado pequeño, no habrá en el gráfico suficientes puntos negros como para identificar órbitas periódicas. Si, por el contrario, l es demasiado grande, las órbitas periódicas quedarán ocultas por un exceso de puntos. Esto podemos observarlo en la figura 5.2.3 sobre una serie de la función logística.



Figura 5.2.3: Gráficos de recurrencias de la función logística para precisiones de 0.1, 0.01 y 0.001.

Para la elección de la precisión, Mindlin y Gilmore¹¹ proponen el rango de variación de la serie, partido por cien. Dada la presencia de valores extremos en muchas de nuestras series, hemos preferido partir el rango por mil [5.2.2].

¹¹ Mindlin y Gilmore, 1992.

$$l=(\text{máx}[x_i]-\text{mín}[x_i])/1000 \quad [5.2.2]$$

El análisis de recurrencias presenta una serie de ventajas. Para empezar, es aplicable a series de datos cortas, pues lo único que buscamos son algunas líneas horizontales. Además proporciona información adicional sobre la estructura del atractor, pues el gráfico conserva el orden temporal de la serie, permitiendo conocer el lugar y la periodicidad de las órbitas periódicas.

Las figuras 5.2.4 y 5.2.5 son representativas de los resultados que obtenemos con el resto de las series¹². Las líneas diagonales que destacan en todas las gráficas se deben a valores iguales seguidos en la serie, y que en la mayoría de los casos son rendimientos próximos a cero. Respecto a las líneas horizontales, que son las que nos indican la existencia de órbitas periódicas inestables, éstas no se distinguen en las series de cinco minutos (figura 5.2.5), donde apenas presentan rastros, mientras que en las series operación a operación (figura 5.2.4), las órbitas periódicas inestables son más frecuentes.

¹² El conjunto de las gráficas de recurrencias se pueden ver en el apéndice gráfico, figuras de la A.158 a la A.202.

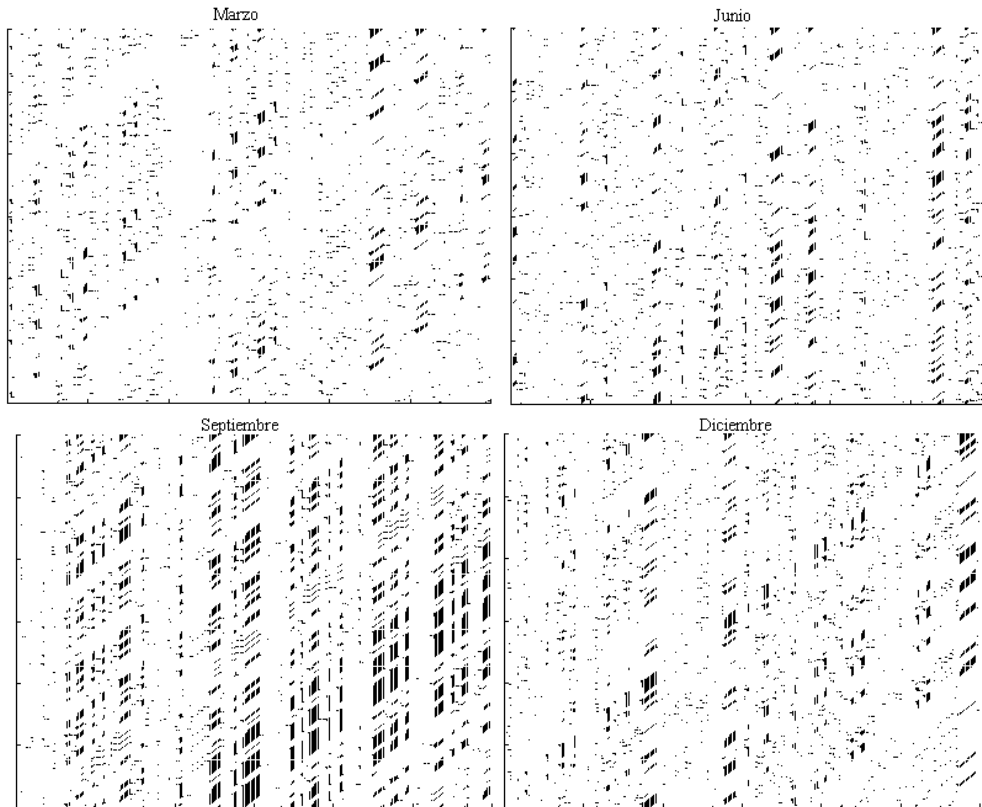


Figura 5.2.4: Gráficos de recurrencia de los primeros 300 retardos de las series de residuos de los modelos ARIMA sobre los rendimientos con frecuencia operación a operación y una precisión de $l=rango/1000$.

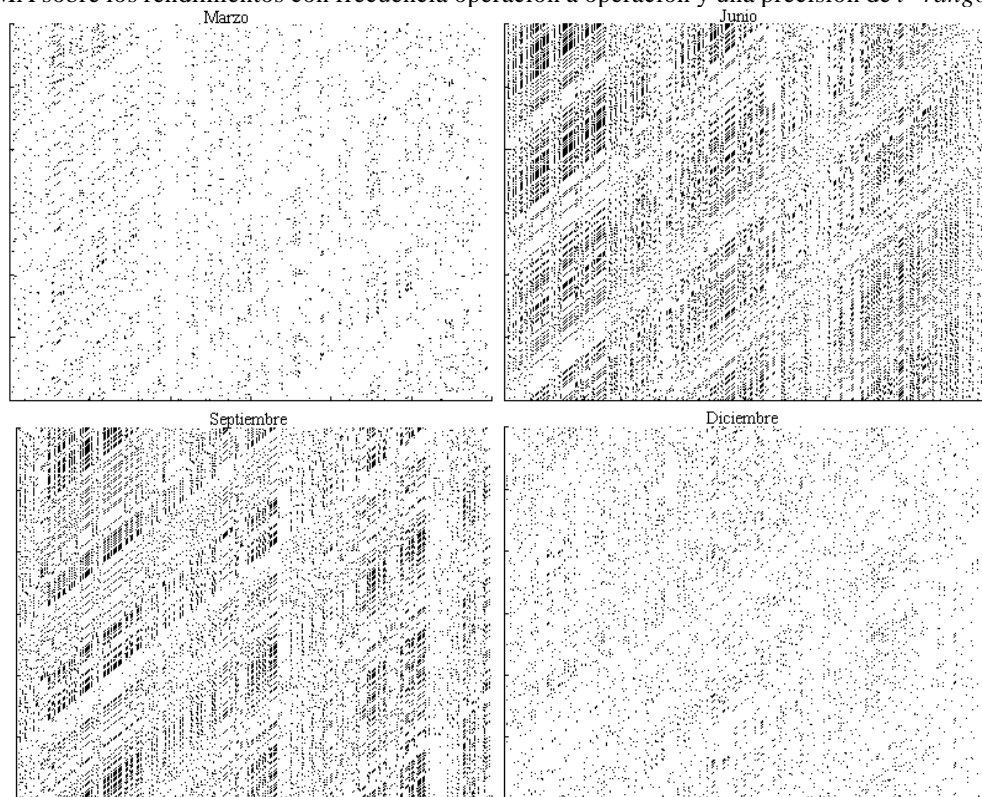


Figura 5.2.5: Gráficos de recurrencia de los primeros 300 retardos de las series de residuos de modelos ARIMA sobre rendimientos con frecuencia de cinco minutos y una precisión de $l=rango/1000$.

La diferencia entre los resultados obtenidos con unas y con otras es una primera señal de las consecuencias que tiene agrupar los datos en series de periodicidad constante, pues nos está ocultando patrones existentes en las series. Sin embargo las conclusiones que podemos sacar de los gráficos de recurrencias no son muchas ya que si bien en la figura 5.2.4, sí encontramos ciertos rastros de órbitas periódicas inestables, en la gran mayoría de las series esto no es así. La mayoría de las series nos reflejan grandes manchas negras debidas a la existencia de muchos valores repetidos. Algo que ya mencionábamos cuando tratábamos las gráficas con estructura de rejillas. Decíamos entonces que el 90% de la serie está agrupada entorno a 9 valores, por lo que es normal que cada punto se encuentre en las proximidades del otro 90% de la serie, impidiendo que las gráficas de recurrencias muestren algún tipo de órbita periódica inestable que pudiera existir.

5.3. COEFICIENTE DE HURST.

A diferencia del análisis gráfico, en el resto del capítulo vamos a utilizar distintos análisis de tipo cuantitativo que van a venir contruidos a partir de leyes de potencias. Tanto la dimensión de correlación como los exponentes de Lyapunov tienen en común ser coeficientes de unas determinadas leyes de potencia. Pero previo a éstos tenemos el coeficiente de Hurst y el estadístico R/S^{13} , a partir de otra ley de potencia, [5.3.1].

$$R/S(N)=(a*N)^H \quad [5.3.1]$$

¹³ Mandelbrot y Van Ness, 1968; Anis y Lloyd, 1976; Davies y Harte, 1987.

La expresión [5.3.1] es una ley de potencia en la que a es una constante, N es el número de observaciones, H es el exponente o coeficiente de Hurst, y $R/S(N)$ es el estadístico R/S que depende del tamaño de la serie y que se define como el rango de variación de la serie partido por su desviación típica. El coeficiente de Hurst se utiliza para la detección de memoria a largo plazo en las series temporales. En caso de encontrarlo, tendríamos una clara evidencia de comportamiento no lineal.

El análisis R/S tiene su origen en el estudio de las variaciones anuales que se producían en el cauce del Nilo. Hurst elaboró un nuevo estadístico que le permitiera contrastar si el nivel de reservas de una hipotética presa en el Nilo seguiría un paseo aleatorio.

El coeficiente de Hurst, H , indica la persistencia o la no-persistencia en una serie temporal. H es un número entre 0 y 1. En el caso de un ruido blanco, el valor obtenido sería 0'5. Una serie con un valor superior a 0'5 estará caracterizada porque los valores tienden a mantenerse en el tiempo (persistencia). Valores inferiores indicarán que, si en un periodo la serie ha aumentado, en el siguiente es más probable que baje, y al revés.

Si encontramos indicios de persistencia en la serie, será señal indudable de que la serie no es un ruido blanco y que existe algún tipo de dependencia entre los datos, sea ésta lineal o no lineal. Para modelizar series con dependencia a largo plazo se utilizan modelos estocásticos fraccionalmente integrados (ARFIMA) en los que el coeficiente de

integración fraccionaria viene dado por el coeficiente de Hurst. Este es el modelo que hemos usado con las series de duraciones.

Para nuestro análisis partimos de la serie temporal [5.3.2], sobre la que calcularemos su coeficiente de Hurst.

$$\{x_t\}_{t=1\dots T} \quad [5.3.2]$$

El primer paso es calcular la media [5.3.3] y la desviación típica [5.3.4] de la serie temporal, que dividimos en segmentos de tamaño N .

$$m(N, t_0) = \sum_{t=t_0+1}^{t_0+N} \frac{x_t}{N} \quad [5.3.3]$$

$$S(N, t_0) = \sqrt{\frac{1}{N} \sum_{t=t_0+1}^{t_0+N} (x_t - m(N, t_0))^2} \quad [5.3.4]$$

A partir de la media obtenida en [5.3.3] transformamos la serie temporal original, en una nueva serie [5.3.5] que representa la diferencia acumulada respecto a la media.

$$y(N, t_0, \mathbf{t}) = \sum_{t=t_0+1}^{t_0+\mathbf{t}} (x_t - m(N, t_0)) \quad 1 \leq \mathbf{t} \leq N \quad [5.3.5]$$

Con la serie transformada en [5.3.5], calculamos el rango de variación [5.3.6]. El estadístico R/S [5.3.7], lo obtenemos dividiendo el rango entre la desviación típica que habíamos calculado en [5.3.4].

$$R(N, t_0) = \max_t y(N, t_0, \mathbf{t}) - \min_t y(N, t_0, \mathbf{t}) \quad [5.3.6]$$

$$[R/S](N) = \frac{\sum_{t_0} R(N, t_0)}{\sum_{t_0} S(N, t_0)} \quad [5.3.7]$$

Para cada t_0 calculamos el rango de variación de la serie, y definimos la relación R/S como el cociente entre el rango y la desviación típica. Esta relación va a aumentar conforme aumente el tamaño de N . Un aumento de las observaciones lleva, bien a mantener el máximo y el mínimo y por tanto el rango o bien a lograr nuevos máximos y mínimos y con ello un rango mayor. El crecimiento de R/S sigue la ley de potencia que presentábamos en la ecuación [5.3.1]. La potencia de dicha ecuación es a la que se llama coeficiente de Hurst. El coeficiente de Hurst se obtiene de la regresión [5.3.8].

$$\log R/S = b + H \log N \quad [5.3.8]$$

El cálculo de R/S puede ser inexacto cuando el tamaño de N es demasiado grande o demasiado pequeño. Para evitar el problema hay que procurar calcular el coeficiente H usando valores intermedios de N .

Los resultados obtenidos se pueden observar en las Tablas 5.3.1 para la frecuencia de operación a operación y 5.3.2 para la frecuencia de cinco minutos. Las

series de duraciones se habían analizado en el capítulo anterior, y los residuos de los modelos ARFIMA aparecen ya sin memoria a largo plazo, por lo que no se han incluido aquí los resultados.

Tabla 5.3.1: Coeficiente de Hurst de las series de operación a operación

	marzo 98	junio 98	septiembre 98	diciembre 98
serie original	0,4907 (0,0027)	0,5053 (0,0015)	0,4727 (0,0025)	0,5715 (0,0061)
serie residuos modelo ARIMA	0,4936 (0,0028)	0,5081 (0,0015)	0,4664 (0,0036)	0,5726 (0,0061)
serie residuos modelo GARCH-M	0,4902 (0,0016)	0,4862 (0,0027)	0,5259 (0,0035)	0,5403 (0,0030)
serie residuos modelo EGARCH	0,5157 (0,0028)	0,4825 (0,0025)	0,5263 (0,0060)	0,5388 (0,0041)

Entre paréntesis desviación típica de la regresión, no desviación típica del estimador

Tabla 5.3.2: Coeficiente de Hurst de las series de cinco minutos

	marzo 98	junio 98	septiembre 98	diciembre 98
serie original	0,5096 (0,0035)	0,5239 (0,0016)	0,4694 (0,0025)	0,5491 (0,0051)
serie residuos modelo ARIMA	0,4936 (0,0028)	0,5081 (0,0015)	0,4664 (0,0036)	0,5726 (0,0061)
serie residuos modelo GARCH (*)	0,5383 (0,0035)	0,5495 (0,0031)	0,5072 (0,0040)	0,5578 (0,0031)
serie residuos modelo EGARCH	0,5433 (0,0036)	0,5465 (0,0030)	0,5064 (0,0045)	0,5445 (0,0026)

Entre paréntesis desviación típica de la regresión, no desviación típica del estimador.

(*)En el vencimiento de diciembre el modelo estimado era un GARCH-M.

Las desviación típica sólo nos sirve para reflejar la precisión de la regresión [5.3.8] que hemos tenido que utilizar para la estimación, pero en ningún caso nos da la distribución del estadístico H . Aunque es cierto que tan sólo en los casos de los modelos de volatilidades de las series de rendimientos con frecuencia de cinco minutos y vencimiento en septiembre de 1998 obtenemos estimaciones de H que podríamos considerar como 0'5, los resultados son en todos los demás casos próximos a ese 0'5. Si bien no parece que encajen los resultados con un ruido blanco, sin considerar la

distribución del estadístico, valores tan próximos a 0.5 no son justificación suficiente para el uso de modelos de tipo ARFIMA.

5.4. RECONSTRUCCIÓN DEL SISTEMA.

Para comprobar si un proceso presenta un comportamiento caótico tendremos que analizar la evolución de la órbita del sistema. La órbita viene dada por un conjunto de vectores como los definidos en [5.4.1]. En el caso de que el sistema sea caótico, el vector \bar{x} en el momento t depende de forma completamente determinista de ese mismo vector \bar{x} en el periodo anterior. Si el sistema es estocástico esta función será aleatoria.

$$\{\bar{x}_t \in \mathbb{R}^n / \bar{x}_t = \mathbf{j}(\bar{x}_{t-1})\} \quad [5.4.1]$$

Para conocer la órbita, tendríamos que conocer la evolución de todos los componentes del vector. Por desgracia, en la mayoría de las ocasiones no sólo desconocemos la función ϕ , sino cuales son las variables que intervienen en el sistema e incluso cuántas variables hay. Sólo disponemos de una serie temporal de escalares que, suponemos, proceden de ese vector, tal como en [5.4.2].

$$\hat{\mathbf{A}}^n \approx \hat{\mathbf{A}} \quad \{\bar{x}_t\} \approx \{y_t = y(\bar{x}_t)\} \quad [5.4.2]$$

Takens¹⁴ demostró que en la mayoría de los casos, la transformación [5.4.3] conserva las propiedades topológicas de la órbita original siempre que $m > 2n + 1$.

$$\hat{A} : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

$$\{y_t\} \mapsto \{y_t, y_{t+1}, \dots, y_{t+(m-1)}\} \quad [5.4.3]$$

Mediante la transformación propuesta por Takens, de los escalares observados y_t pasamos a vectores de dimensión m , cuantía que se denomina *dimensión de inmersión*. A los vectores así formados también se les llama *m-historias*. Estos vectores tienen una estructura que dibuja una trayectoria similar a la que posea la órbita original.

Podemos observar esta característica a partir de un sistema caótico concreto: el mapa de Hénon¹⁵ definido en [5.4.4].

$$\left. \begin{aligned} x_{t+1} &= 1 - 1.4x_t^2 + y_t \\ y_{t+1} &= 0.3x_t \end{aligned} \right\} \quad [5.4.4]$$

Este modelo fue introducido para explicar el movimiento de ciertos cuerpos espaciales y presenta la cualidad de ser un sistema de ecuaciones muy simple en el que la única componente no lineal es el x^2 de la primera ecuación. Sin embargo el comportamiento del sistema es tan complejo como se puede observar en la figura 5.4.1.

¹⁴ F. Takens, 1981.

¹⁵ M. Hénon, 1976.

En el caso de que, en vez del vector (x,y) que nos señala la evolución del sistema, sólo observáramos la variable x_t y no supiéramos nada del comportamiento de y_t , podríamos reconstruir todo el sistema con las m -historias de tamaño dos de la serie $\{x_t\}$.

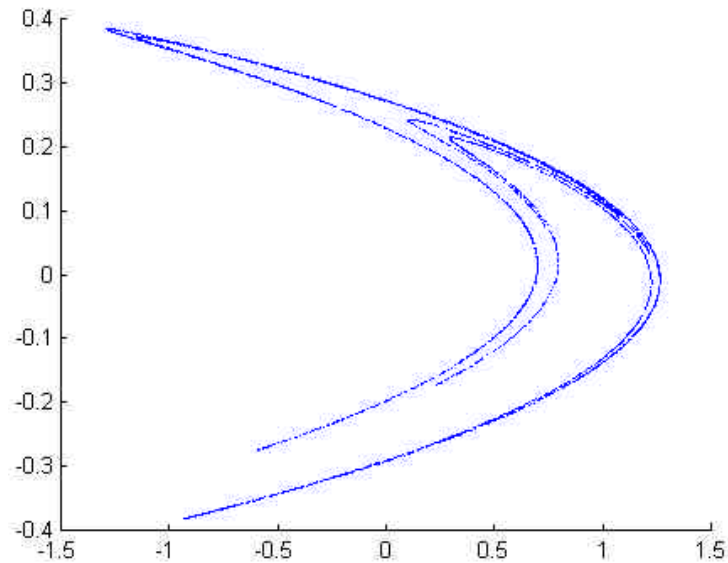


Figura 5.4.1: Mil puntos del atractor de Hénon. Semilla inicial $x_0=1$ e $y_0=1$. Los primeros cien puntos no se han incluido.

Para hacer la transformación de Takens es necesario determinar el valor de dos parámetros: t y m . El tamaño del vector va a venir determinado por m , mientras que t nos indica el tiempo que debe pasar entre observación y observación dentro de la misma m -historia.

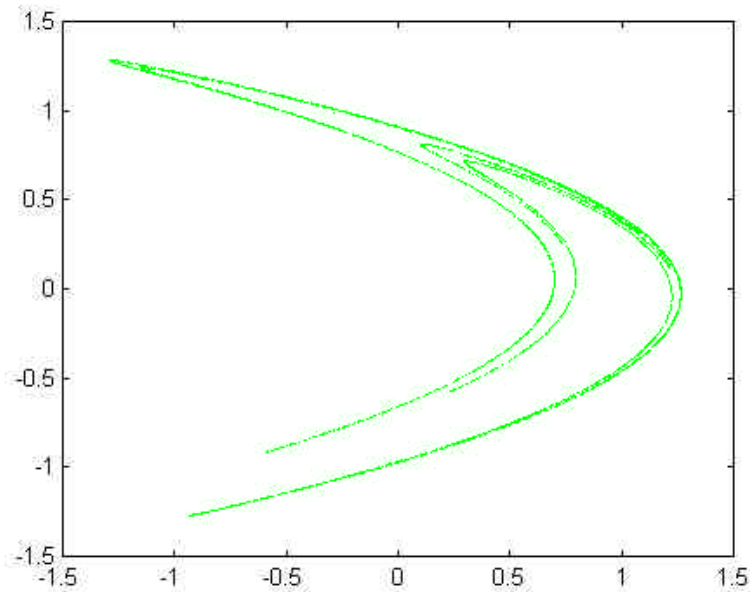


Figura 5.4.2: Reconstrucción del atractor de Hénon a partir de la serie temporal x_t que hemos transformado en una 2-historia con t igual a uno, tal que en el eje x está representando a x_{t+1} y en el eje y se representa a x_t .

Si partimos de una serie temporal escalar, desconociendo todo lo que hace referencia al sistema de partida no vamos a tener pistas sobre qué valores dar a m y t . La primera alternativa sería ir dando distintos valores a m y t , a la hora de implementar cada uno de los métodos de detección de caos. A posteriori, a la vista de los resultados, decidir cuales son los valores que debemos dar a estos parámetros. Esta forma de actuación es la que con más frecuencia se utiliza, pero tiene como inconveniente que hace crecer de forma explosiva el tiempo de cálculo, ya de por sí elevado, a la vez que complica la posterior interpretación de los resultados.

Existe una revisión completa de los métodos¹⁶ que tratan de darnos unos valores óptimos para m y t . De entre ellos hemos seleccionado aquél que proporciona unos resultados aceptables en unos plazos de tiempo moderados, la información mutua.

¹⁶ Buzug y Pfister, 1992.

El primer paso va a ser obtener el valor de t . La importancia de la elección de t radica en las consecuencias que tiene una errónea especificación de este parámetro. Para t demasiado pequeños, los elementos de la serie temporal serán muy parecidos, y cuando construyamos las m -historias los vectores tenderán a agruparse entorno a la bisectriz, resultando una mala reconstrucción del sistema pues estaremos aproximando artificialmente los puntos¹⁷. Si por el contrario tomamos valores de t demasiado elevados, corremos el riesgo de que la sensibilidad a las condiciones iniciales haga que la conexión entre dos elementos consecutivos de la serie sea demasiado débil para que seamos capaces de estimar el comportamiento futuro del sistema.

En la figura 5.4.3 podemos observar cómo una errónea especificación de t puede ocultar la dinámica de un sistema. En dicha figura se ha obtenido una serie a partir del sistema de Lorenz que habíamos definido en el capítulo tercero, y luego se ha intentado reconstruirlo con distintos valores de t . Para valores demasiado pequeños los puntos parecen concentrarse entorno a la bisectriz. Para t igual a 10 se consigue una reconstrucción óptima del atractor, pero para valores de t superiores a 10 la estructura se diluye y no resulta posible reconocer el sistema. En las series que estamos estudiando sobre los contratos de Futuro, las cotizaciones diarias exigen usar un t próximo a 2000, mientras que las series de cinco minutos pueden llegar en algunos casos a tener un t de hasta cien.

¹⁷ Kennel e Isabelle, 1992, señalan además que en caso de estar muy juntos no resulta posible distinguir un sistema caótico de ciertos tipos de ruidos.

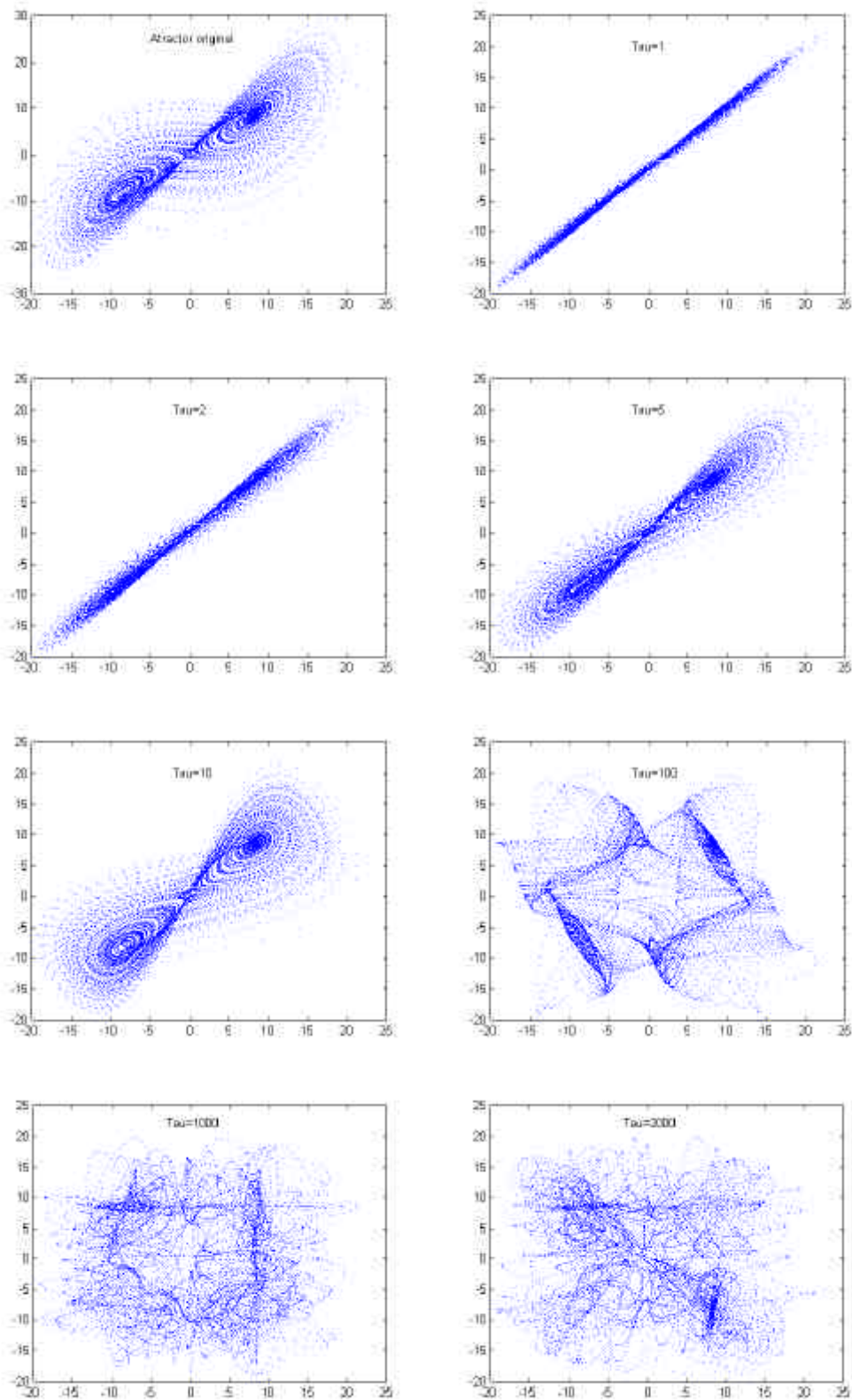


Figura 5.4.3: Serie de 10.000 datos obtenidos del sistema de Lorenz. El primer gráfico representa a las variables (x,y). Los siguientes son reconstrucciones del atractor a partir de la variable x tomando como valores de τ 1, 2, 5, 10, 100, 1000 y 2000.

En las series que estamos analizando, vamos a considerar que el t óptimo va a ser aquél que minimice la denominada información mutua. Nuestro objetivo es que las m -historias no se acumulen en la bisectriz. La forma de asegurarnos que esas m -historias se distribuyen en el espacio es reduciendo la información mutua que proporcionan x_t sobre x_{t+t} y viceversa.

La información mutua es una magnitud que se obtiene a partir de la entropía de Shannon¹⁸. La expresión de la entropía de una variable aparece en [5.4.5] y la entropía de la distribución conjunta de dos variables en [5.4.6].

$$H = -\sum_i p_i \log p_i \quad [5.4.5]$$

$$H = -\sum_i \sum_j p_{ij} \log p_{ij} \quad [5.4.6]$$

La información mutua se define como la diferencia entre la entropía conjunta y las entropías de cada una de las variables individuales [5.4.7].

$$IM = \sum_i \sum_j p_{ij} \log p_{ij} - \sum_j p_j \log p_j - \sum_i p_i \log p_i \quad [5.4.7]$$

Para el caso de una serie temporal nuestras variables van a ser x_t y x_{t+t} . Creamos marcas de clase y calculamos las frecuencias relativas, que van a servirnos como estimaciones de los valores p_i de la expresión [5.4.7].

¹⁸ Shanon, 1948.

A medida que aumenta la distancia temporal entre observaciones disminuye la información que la serie proporciona sobre el siguiente valor. El primer mínimo que se alcanza será la distancia óptima, t , para reconstruir el atractor.

Los resultados obtenidos para las series aparecen en las figuras de la A.203 a la A.247 del apéndice gráfico. En todas ellas se observa un rápido descenso de la información mutua, lo que corresponde a que efectivamente las observaciones se encontraban agrupadas entorno a la bisectriz mientras que con posterioridad parece que se estabiliza produciéndose dientes de sierra que indican que para esos retardos la estructura se está ocultando. Los primeros mínimos locales de las series se dan para los valores de t que aparecen en las Tablas 5.4.1 para los rendimientos con frecuencia operación a operación, 5.4.2 para los rendimientos con frecuencia de cinco minutos y 5.4.3 para las duraciones de operaciones.

Tabla 5.4.1: t óptimo para las series de rendimientos y de residuos de los distintos modelos (operación a operación)

	marzo	junio	septiembre	diciembre
ORIGINAL	4	6	3	2
ARIMA	7	6	3	6
GARCH-M	5	6	6	5
EGARCH	5	6	5	6

Tabla 5.4.2: t óptimo para las series de rendimientos y de residuos de los distintos modelos (cinco minutos)

	marzo	junio	septiembre	diciembre
ORIGINAL	5	5	3	4
ARIMA	5	5	3	6
GARCH(*)	2	1	4	3
EGARCH	2	4	3	4

(*)En el vencimiento de diciembre el modelo estimado era un GARCH-M.

Tabla 5.4.3: t óptimo para las series de rendimientos y de residuos de los distintos modelos (duraciones)

	marzo	junio	septiembre	diciembre
ORIGINAL	2	5	8	4
DIFERENCIA				
FRACCIONAL	5	4	6	30
ARFIMA	15	4	12	10

Se puede comprobar de estos resultados, cómo para las series de rendimientos hay pocas diferencias de un modelo a otro o de una frecuencia a otra. En todos los casos obtenemos valores inferiores a 7. En el caso de las series de duraciones de operaciones, sin embargo, parece que las diferencias fraccionales llevan a elegir valores de t muy superiores, llegando incluso a 30.

Una vez que hemos escogido el t óptimo, debemos buscar el m óptimo. Por Takens sabemos que la dimensión de inmersión esta relacionada con la dimensión del sistema original, de tal forma que para asegurarnos que estamos reproduciendo la dinámica del sistema tendremos que asegurarnos que $m=2n+1$. Nuestro problema es que desconocemos cual es ese valor de n . Una opción sería elegir un m muy alto. Pero si m toma valores demasiado altos, incrementamos artificialmente el ruido, pues puntos originalmente próximos aparecerán aislados, siendo ineficaces los métodos para distinguir un comportamiento caótico de uno aleatorio.

Al dar un valor alto a t o a m , podría parecer que se nos está reduciendo el tamaño de la serie. El hecho de que decidamos que t debe ser tres, no quiere decir que reduzcamos a una tercera parte el tamaño de la serie. Lo que vamos a hacer es solapar las m -historias como hacemos en [5.4.8].

$$\begin{aligned}
& (y_1, y_4, y_7, y_{10}, y_{13}) \\
& (y_2, y_5, y_8, y_{11}, y_{14}) \\
& \dots \\
& (y_t, y_{t+t}, y_{t+2t}, y_{t+3t}, y_{t+4t})
\end{aligned} \tag{5.4.8}$$

La mayoría de los métodos que vamos a presentar requieren unas series temporales muy grandes, por lo que no nos podemos permitir el lujo de recortar el tamaño de la serie si queremos obtener resultados significativos. El hecho de solapar las m -historias no va a producirnos problemas. El tamaño de t nos viene dado por las necesidades de separación entre los elementos de una misma m -historia, no por la separación entre m -historias distintas. Y lo mismo nos sucede con m , que no va a depender de la separación entre los vectores.

5.4.1. DIMENSIÓN DE CORRELACIÓN.

La dimensión de Hausdorff-Besicovitch, que definíamos en el capítulo tres, a pesar de ser la más corriente a la hora de identificar fractales no es la única posible. Una definición generalizada del concepto dimensión aparece en la expresión [5.4.9].

$$D_q = \lim_{l \rightarrow 0} \frac{1}{q-1} \frac{\log \sum_{i=1}^N p_i^q}{\log l} \tag{5.4.9}$$

Donde p_i es la probabilidad de que una variable aleatoria distribuida uniformemente a lo largo del fractal, caiga en el i -ésimo subconjunto de tamaño l en el que hemos dividido el fractal. El parámetro q va a variar en función del tipo de dimensión que hayamos considerado. El rango de variación de q es más y menos

infinito, o lo que es lo mismo, vamos a disponer de infinitas definiciones de dimensión¹⁹ (desde $D_{-\infty}$ hasta D_{∞}).

Así, por ejemplo, la dimensión de Hausdorff-Besicovitch se corresponde a D_0 [5.4.10].

$$D_0 = \lim_{l \rightarrow 0} \frac{1}{0-1} \frac{\log \sum_{i=1}^N p_i^0}{\log l} = \lim_{l \rightarrow 0} -\frac{\log N}{\log l} = \lim_{l \rightarrow 0} \frac{\log N}{\log(1/l)} \quad [5.4.10]$$

Para calcular la dimensión fractal se divide el espacio en subconjuntos de lado l y se comprueba cuántos de esos subespacios están ocupados por el atractor. Este cálculo resulta muy complicado en la práctica y conlleva un tiempo de computación muy elevado.

Para obtener una dimensión fácilmente computable a partir de una serie temporal vamos a utilizar D_2 también denominada dimensión de correlación²⁰. Para q igual a dos, la definición generalizada que dábamos en [5.4.9] se transforma en la expresión [5.4.11].

$$D_2 = \lim_{l \rightarrow 0} \frac{\log \sum_{i=1}^N p_i^2}{\log l} \quad [5.4.11]$$

¹⁹ Hentschel y Procaccia, 1983.

²⁰ Halsey, Jensen, Kadanoff, Procaccia y Shraiman, 1986.

En el numerador tenemos el logaritmo de $\sum_{i=1}^{M(r)} p_i^2$. Este sumatorio puede definirse²¹ como la probabilidad de que dos puntos del atractor caigan en una celda de l^d , lo que es similar a la probabilidad de que dos puntos estén separados por una distancia menor a l . Esta probabilidad vamos a calcularla mediante la integral de correlación propuesta por Grassberger y Procaccia²² que definimos en [5.4.12].

$$C^m(l) = \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \mathcal{J}\{l - |\bar{x}_i - \bar{x}_j|\} \quad [5.4.12]$$

A $\mathcal{J}\{l - |\bar{x}_i - \bar{x}_j|\}$ se la denomina función de Heaviside, y no es más que una función que vale uno si el argumento es positivo y cero si no lo es. Con $|\dots|$ representamos una norma. Por su facilidad de cálculo se utiliza la norma suprema (la magnitud del vector es el máximo de sus componentes cartesianos).

Como sucedía con la dimensión fractal, para la dimensión de correlación se verifica²³ la expresión [5.4.13].

$$C(l) \sim l^v \quad [5.4.13]$$

²¹ Schuster, 1984. pag. 105.

²² Grassberger y Procaccia, 1983 A; Grassberger y Procaccia, 1983 B.

²³ Grassberger y Procaccia, 1983 A.

Donde ν es la dimensión de correlación, y $C(l) = \lim_{m \rightarrow \infty} C^m(l)$. Tomando logaritmos en [5.4.13] obtenemos la ecuación lineal [5.4.14] en la que la pendiente es la dimensión de correlación.

$$\log C(l) = \log k + \nu \cdot \log l \quad [5.4.14]$$

Si la serie es caótica, a medida que calculamos la dimensión de correlación, para distintos valores de inmersión (n), ν tenderá asintóticamente al valor que tiene la dimensión del atractor.

Si por el contrario, nos encontráramos ante un fenómeno aleatorio, los valores de las m -historias no se concentrarían en el atractor, sino que cubrirían uniformemente todo el espacio disponible. De esta forma, si aumentamos la dimensión de inmersión, aumentamos el espacio disponible y por tanto aumentará la estimación de la dimensión de correlación. Esto nos proporcionará una regla para distinguir entre fenómenos aleatorios y fenómenos caóticos.

En un sistema caótico, la estimación de la dimensión de correlación no crece al aumentar la dimensión de inmersión por encima de cierto valor. Si es aleatorio, la dimensión de correlación crecerá de forma indefinida ante variaciones de la dimensión de inmersión.

La dimensión de correlación es un mínimo de la dimensión fractal. Se diferencia de ésta en que la base de la dimensión de correlación está en la proximidad entre los

puntos, no en el número de cubos. Se buscan las trayectorias que tienden a volver a una región por la que ya han pasado. Es lo mismo que sucede en el solenoide de Smale de la Figura 3.3.1 cuando se han producido el número suficiente de iteraciones.

Calculada²⁴ la integral de correlación [5.4.12] de nuestras series²⁵ tendríamos las observaciones necesarias para estimar la dimensión de correlación tal como aparecía en la ecuación [5.4.14]. Sin embargo vemos como las series originales, tanto las de rendimientos como las de duraciones tienen forma escalonada. Este tipo de estructura se produce cuando el número de valores que puede tomar la serie es restringido. Esto nos sucede porque los precios se toman de “tick” en “tick” y las duraciones en segundos. En las series de residuos la curva de la integral de correlación se va suavizando, tal y como sucedía en las gráficas de x_t frente a $x_{t+\tau}$.

Según Grassberger y Procaccia²⁶, la relación entre $\log_2 C(l)$ y $\log_2 l$ es lineal. Para averiguar la dimensión de correlación de una serie basta con hallar la pendiente de esa recta.

Sin embargo, Smith²⁷ demuestra que esta relación no es lineal. La elección del segmento para el que se calcula la dimensión de correlación influye en el resultado que se obtiene.

²⁴ El cálculo de cada una de las curvas de la integral de correlación de una serie de 200.000 datos requiere, en un Pentium III-500 y con 256 Mb de RAM, 555 horas de computación. Para las series, operación a operación, hemos hecho una estimación en la que no hemos comparado 200.000*200.000 pares de puntos, sino 2.000*200.000 puntos, lo que reduce el tiempo de computación a una centésima parte.

²⁵ Las gráficas se encuentran en el apéndice gráfico, figuras de la A.248 a la A.293.

²⁶ P. Grassberger e I. Procaccia, 1983 A.

²⁷ L. A. Smith, 1988.

Smith construye la dimensión de correlación en función de l [5.4.15], obteniendo una relación que no es ni mucho menos constante como se puede observar en la figura 5.4.4.

$$\hat{\nu}(l) = \frac{d \log C(l)}{d \log l} \quad [5.4.15]$$

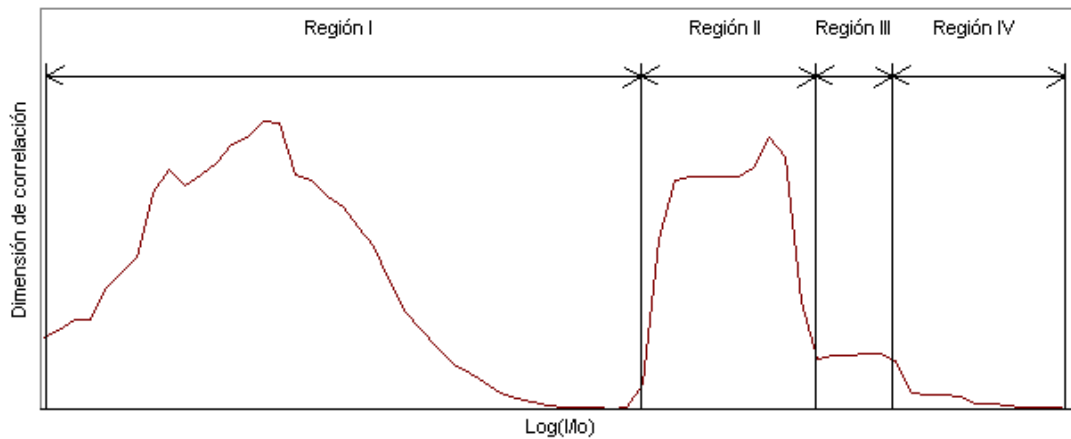


Figura 5.4.4: Dimensión de correlación en función de l presentando las cuatro regiones características.

Smith identifica en esta función cuatro regiones dominadas por fenómenos diferentes y que aparecen marcados en la figura 5.4.4.

La primera región se da para valores excesivamente pequeños de l . En esta región no encontramos puntos próximos, sino solamente puntos aislados. Esto es consecuencia de que partimos de una serie de tamaño limitado y no resulta posible hacer ningún tipo de análisis en esa región.

Aumentando l pasamos a la segunda región, que está dominada por el ruido en la observación. A pesar de que estuviéramos ante un sistema completamente determinista va a haber un ruido provocado por la observación y que es imposible de eliminar. Este ruido nos impide conocer los valores exactos del sistema. La sensibilidad a las condiciones iniciales, nos va a impedir dar predicciones a largo plazo. Al estar dominada por fenómenos aleatorios, presenta valores altos, próximos a la dimensión de inmersión.

La tercera región se alcanza cuando l tiene un tamaño superior a la desviación típica de la perturbación aleatoria y el cálculo de la dimensión de correlación ya no se ve afectada por dicha perturbación²⁸. En caso de existir el fenómeno caótico, podrá ser caracterizado en esta región mediante la dimensión de correlación. La tercera región supone un descenso en la función $v(l)$, respecto a la segunda región y una zona de cierta estabilidad en el conjunto de la función²⁹.

La última región se da para valores de l que abarcan prácticamente todos los puntos de la serie. La función $v(l)$ se caracteriza por valores que de nuevo son próximos a cero.

Cada una de estas regiones nos está dando valores distintos para $v(l)$. Tenemos que buscar la tercera región, si existe, para calcular en ella la dimensión de correlación. En las figuras de la A.294 a la A.338 del apéndice gráfico se representan las funciones

²⁸ Ben-Mizrachi, Procaccia y Grassberger, 1984.

²⁹ Ding, Grebogi, Ott, Sauer y Yorke, 1993.

$v(l)$ para las distintas series estudiadas y se marca lo que consideramos que es la tercera región observada en cada una de ellas.

Tabla 5.4.4: Dimensiones de correlación con desviaciones típicas de las estimaciones.

Series operación a operación										
	m=	2	3	4	5	6	7	8	9	10
Series originales	marzo	0,1513	0,2484	0,3685	0,4670	0,5583	0,6405	0,7143	0,8140	0,8948
		(0,0342)	(0,0560)	(0,0820)	(0,1044)	(0,1251)	(0,1437)	(0,1605)	(0,1831)	(0,2015)
	junio	0,1443	0,2031	0,2689	0,3306	0,4265	0,5066	0,5850	0,6478	0,7046
		(0,0311)	(0,0440)	(0,0581)	(0,0714)	(0,0920)	(0,1092)	(0,1260)	(0,1396)	(0,1518)
	septiembre	0,1793	0,2800	0,3847	0,4476	0,5230	0,6136	0,7112	0,7743	0,8754
		(0,0372)	(0,0585)	(0,0770)	(0,0880)	(0,1038)	(0,1226)	(0,1402)	(0,1518)	(0,1715)
	diciembre	0,3363	0,4690	0,6038	0,7409	0,8651	0,9834	1,0929	1,2577	1,4166
		(0,0768)	(0,1070)	(0,1378)	(0,1691)	(0,1974)	(0,2244)	(0,2494)	(0,2870)	(0,3232)
Series de residuos modelo ARIMA	marzo	0,1850	0,2761	0,3647	0,4571	0,5405	0,6298	0,7119	0,8037	0,8867
		(0,0259)	(0,0393)	(0,0522)	(0,0658)	(0,0777)	(0,0909)	(0,1028)	(0,1161)	(0,1280)
	junio	0,1456	0,2108	0,2758	0,3474	0,4174	0,4825	0,5476	0,6100	0,6743
		(0,0210)	(0,0304)	(0,0398)	(0,0503)	(0,0599)	(0,0694)	(0,0788)	(0,0879)	(0,0976)
	septiembre	0,1822	0,2711	0,3468	0,4174	0,4876	0,5579	0,6424	0,7052	0,7749
		(0,0267)	(0,0395)	(0,0511)	(0,0618)	(0,0727)	(0,0832)	(0,0971)	(0,1073)	(0,1181)
	diciembre	0,3402	0,5065	0,6638	0,8081	0,9662	1,1125	1,2596	1,3936	1,5410
		(0,0618)	(0,0917)	(0,1204)	(0,1476)	(0,1761)	(0,2037)	(0,2317)	(0,2568)	(0,2832)
Series de residuos modelo GARCH-M	marzo	0,2075	0,3239	0,4356	0,5392	0,6365	0,7219	0,8287	0,9234	1,0314
		(0,0231)	(0,0341)	(0,0456)	(0,0552)	(0,0650)	(0,0746)	(0,0860)	(0,0953)	(0,1012)
	junio	0,1509	0,2169	0,2873	0,3535	0,4509	0,5353	0,6116	0,6830	0,7508
		(0,0132)	(0,0194)	(0,0247)	(0,0321)	(0,0406)	(0,0471)	(0,0514)	(0,0585)	(0,0648)
	septiembre	0,1839	0,1870	0,3395	0,4061	0,4864	0,5857	0,6796	0,7650	0,8467
		(0,0157)	(0,0113)	(0,0295)	(0,0336)	(0,0410)	(0,0522)	(0,0587)	(0,0643)	(0,0704)
	diciembre	0,3404	0,5040	1,1841	0,8569	1,0161	1,1841	1,3544	1,5511	1,7284
		(0,0301)	(0,0501)	(0,1311)	(0,0904)	(0,1093)	(0,1311)	(0,1562)	(0,1787)	(0,2000)
Series de residuos modelo EGARCH	marzo	0,7548	1,2488	1,6955	2,2231	2,7843	3,2160	3,6735	4,2126	4,7421
		(0,0230)	(0,0277)	(0,0323)	(0,0417)	(0,0553)	(0,0639)	(0,0820)	(0,1042)	(0,1111)
	junio	0,1812	0,2571	0,3358	0,4086	0,5217	0,6023	0,6845	0,7644	0,8340
		(0,0211)	(0,0309)	(0,0390)	(0,0474)	(0,0600)	(0,0677)	(0,0745)	(0,0845)	(0,0929)
	septiembre	0,3999	0,6190	0,8567	1,0626	1,2405	1,3776	1,5262	1,7107	1,9003
		(0,0070)	(0,0107)	(0,0139)	(0,0196)	(0,0205)	(0,0273)	(0,0338)	(0,0415)	(0,0454)
	diciembre	0,5174	0,8150	1,0104	1,2513	1,5123	1,7108	1,9093	2,1565	2,3093
		(0,0065)	(0,0074)	(0,0077)	(0,0101)	(0,0120)	(0,0144)	(0,0208)	(0,0203)	(0,0280)

La dimensión de correlación calculada en esos intervalos aparece en las Tablas 5.4.4 para las series de frecuencia de operación a operación, 5.4.5 para las series de frecuencia de cinco minutos y 5.4.6 para las series de duraciones.

Tabla 5.4.5: Dimensiones de correlación con desviaciones típicas de las estimaciones. Series cinco minutos

	m=	2	3	4	5	6	7	8	9	10
Series originales	marzo	0,9356	1,3256	1,7328	2,0585	2,4208	2,7314	3,1010	3,3926	3,7321
		(0,0929)	(0,1339)	(0,1765)	(0,2103)	(0,2483)	(0,2796)	(0,3173)	(0,3418)	(0,3670)
	junio	0,5773	0,8196	1,0666	1,2773	1,5078	1,7097	1,9368	2,1407	2,3794
		(0,0558)	(0,0773)	(0,1006)	(0,1180)	(0,1383)	(0,1544)	(0,1747)	(0,1911)	(0,2122)
	septiembre	0,6811	0,9368	1,1495	1,3444	1,5395	1,7229	1,8971	2,0705	2,2332
		(0,0635)	(0,0909)	(0,1147)	(0,1371)	(0,1598)	(0,1824)	(0,2035)	(0,2251)	(0,2458)
	diciembre	1,0579	1,5258	1,9598	2,3810	2,7687	3,1801	3,5636	3,8773	4,2197
		(0,0912)	(0,1320)	(0,1708)	(0,2085)	(0,2433)	(0,2783)	(0,3116)	(0,3380)	(0,3667)
Series de residuos modelo ARIMA	marzo	0,9115	1,3056	1,6657	2,0043	2,3275	2,6493	2,9765	3,3112	3,6588
		(0,0569)	(0,0837)	(0,1083)	(0,1322)	(0,1551)	(0,1778)	(0,2005)	(0,2230)	(0,2459)
	junio	0,5633	0,8087	1,0380	1,2549	1,4661	1,6766	1,8819	2,0911	2,3048
		(0,0558)	(0,0793)	(0,1011)	(0,1209)	(0,1402)	(0,1592)	(0,1777)	(0,1966)	(0,2159)
	septiembre	0,6956	0,9568	1,1814	1,3865	1,5720	1,7450	1,9108	2,0703	2,2269
		(0,0659)	(0,0930)	(0,1175)	(0,1410)	(0,1631)	(0,1846)	(0,2052)	(0,2254)	(0,2453)
	diciembre(*)	0,3402	0,5065	0,6638	0,8081	0,9662	1,1125	1,2596	1,3936	1,5410
		(0,0618)	(0,0917)	(0,1204)	(0,1476)	(0,1761)	(0,2037)	(0,2317)	(0,2568)	(0,2832)
Series de residuos modelo GARCH	marzo	1,0024	1,4409	1,8889	2,3198	2,7383	3,1449	3,5249	3,9684	4,3995
		(0,0602)	(0,0892)	(0,1177)	(0,1461)	(0,1739)	(0,2007)	(0,2225)	(0,2482)	(0,2703)
	junio	1,2616	1,8650	2,4322	3,0238	3,5276	3,9577	4,3985	4,8053	5,1897
		(0,0334)	(0,0488)	(0,0634)	(0,0811)	(0,0995)	(0,1207)	(0,1451)	(0,1769)	(0,2103)
	septiembre	0,9436	1,3951	1,8663	2,3210	2,7596	3,1760	3,5502	3,9264	4,2529
		(0,0431)	(0,0624)	(0,0807)	(0,0953)	(0,1071)	(0,1227)	(0,1338)	(0,1452)	(0,1568)
	diciembre(**)	1,2319	1,8260	2,4142	3,0146	3,5505	4,1368	4,6785	5,2157	5,7905
		(0,0503)	(0,0801)	(0,1149)	(0,1537)	(0,1999)	(0,2476)	(0,2912)	(0,3299)	(0,3629)
Series de residuos modelo EGARCH	marzo	1,4094	2,0976	2,7394	3,3672	3,9340	4,4058	4,8282	5,1054	5,2492
		(0,0348)	(0,0521)	(0,0683)	(0,0841)	(0,0977)	(0,1128)	(0,1350)	(0,1622)	(0,2118)
	junio	1,4784	2,2061	2,8801	3,5456	4,1877	4,7851	5,2501	5,5797	5,5797
		(0,0294)	(0,0540)	(0,0825)	(0,1083)	(0,1488)	(0,1752)	(0,1641)	(0,1461)	(0,1461)
	septiembre	1,5678	2,3667	3,1301	3,8653	4,5793	5,1980	5,6509	5,8763	5,9291
		(0,0601)	(0,0939)	(0,1243)	(0,1493)	(0,1764)	(0,1816)	(0,1665)	(0,1634)	(0,1991)
	diciembre	1,5442	2,3108	3,1029	3,8958	4,6430	5,3184	5,8510	6,2134	6,3559
		(0,0603)	(0,0902)	(0,1299)	(0,1611)	(0,1991)	(0,2077)	(0,1904)	(0,1600)	(0,1452)

(*) En el capítulo cuarto habíamos preferido el modelo ARIMA con intervención, pero aquí analizamos ambos para comprobar si afecta al análisis la intervención.

(**) En el vencimiento de diciembre el modelo estimado era un GARCH-M.

A la vista de los resultados, se observa en todas las series que las dimensiones de correlación estimadas son muy inferiores a la dimensión de inmersión utilizada para su estimación lo que es un primer indicio que favorece la consideración de la presencia de comportamiento caótico en las series estudiadas. A pesar de que se observa que las dimensiones de correlación no llegan a converger a un valor concreto, sí se observa que la velocidad a la que crece va disminuyendo.

Tabla 5.4.6: Dimensiones de correlación con desviaciones típicas de las estimaciones. Series de duraciones

	m=	2	3	4	5	6	7	8	9	10
Series originales	marzo	0,5969	0,8656	1,1211	1,3866	1,5764	1,7725	2,0100	2,2317	2,4862
		(0,0208)	(0,0282)	(0,0411)	(0,0542)	(0,0612)	(0,0663)	(0,0782)	(0,0904)	(0,1105)
	junio	0,5881	0,8764	1,1271	1,3684	1,5923	1,8253	1,9467	2,1604	2,3237
		(0,0193)	(0,0264)	(0,0345)	(0,0396)	(0,0468)	(0,0555)	(0,0629)	(0,0731)	(0,0859)
	septiembre	0,5498	0,7973	1,0662	1,3717	1,5828	1,7364	1,9153	2,1535	2,2566
		(0,0132)	(0,0203)	(0,0271)	(0,0330)	(0,0377)	(0,0422)	(0,0441)	(0,0531)	(0,0440)
	diciembre	0,7013	0,9688	1,2912	1,6119	1,8833	2,1091	2,3451	2,5883	2,6646
		(0,0297)	(0,0425)	(0,0597)	(0,0772)	(0,0918)	(0,1023)	(0,1150)	(0,1232)	(0,1262)
Series de duraciones con diferencia fraccional	marzo	1,5828	2,1823	2,7828	3,2545	3,6608	3,9149	4,3906	5,0392	5,1136
		(0,0228)	(0,0300)	(0,0417)	(0,0441)	(0,0592)	(0,0571)	(0,0663)	(0,0980)	(0,0722)
	junio	1,1569	1,4753	1,7766	2,0740	2,2712	2,4887	2,7434	2,7869	2,8945
		(0,0238)	(0,0323)	(0,0482)	(0,0477)	(0,0535)	(0,0483)	(0,0431)	(0,0347)	(0,0455)
	septiembre	1,2685	1,6835	2,1354	2,6279	2,9283	3,2031	3,5756	3,6973	3,9614
		(0,0271)	(0,0419)	(0,0454)	(0,0421)	(0,0452)	(0,0494)	(0,0597)	(0,0709)	(0,1002)
	diciembre	1,0490	1,4334	1,7980	2,0780	2,4136	2,6179	2,7676	3,0277	3,1347
		(0,0283)	(0,0350)	(0,0446)	(0,0522)	(0,0610)	(0,0617)	(0,0596)	(0,0659)	(0,0666)
Series de residuos modelo ARFIMA	marzo	1,5872	2,2065	2,8096	3,5427	4,1354	4,6884	5,4093	5,5693	5,4641
		(0,0244)	(0,0392)	(0,0610)	(0,0828)	(0,1033)	(0,1140)	(0,0998)	(0,0737)	(0,1407)
	junio	1,6332	2,3232	2,8418	3,2655	3,6278	4,2990	4,9226	5,3902	5,5867
		(0,0235)	(0,0451)	(0,0533)	(0,0609)	(0,0681)	(0,0902)	(0,0719)	(0,0689)	(0,1142)
	septiembre	1,6887	2,4230	3,0502	3,6061	4,1169	4,3881	4,5904	4,5257	4,2320
		(0,0200)	(0,0326)	(0,0471)	(0,0622)	(0,0626)	(0,0491)	(0,0613)	(0,1070)	(0,2357)
	diciembre	1,5518	2,1750	2,8298	3,2654	3,8553	4,2511	4,7170	4,7731	4,7645
		(0,0214)	(0,0354)	(0,0450)	(0,0395)	(0,0471)	(0,0266)	(0,0393)	(0,0408)	(0,0747)

También se observa que las dimensiones son más bajas en las series operación a operación, que en las series de cinco minutos. Esto podría indicar, tal como proponíamos, que al aumentar la periodicidad entre observación y observación, aumenta también la complejidad del fenómeno que pretendemos estudiar.

Smith³⁰ considera que para la adecuada estimación de la dimensión de correlación es necesario disponer de una serie de al menos 10^v datos³¹, donde v es la dimensión de correlación. Esta cifra no se alcanza en las series de cinco minutos pero sí se logra con las series operación a operación. Dada la cantidad de puntos con las que estamos trabajando, no pueden considerarse como muy fiables los resultados obtenidos en las series de cinco minutos.

³⁰ Smith, 1988.

³¹ Tanto Nerenberg y Essex, 1990, como Ramsey y Yuan, 1989, proponen valores ligeramente inferiores. Sin embargo Galka, Maaß y Pfister, 1998, señalan que esa cifra va a depender de cada caso y que para series del orden de los 100.000 datos se puede llegar hasta la dimensión 8 sin problema.

Tal como señala Sayers³², la dimensión de correlación tiende a sesgar los resultados a la baja para las series aleatorias y al alza para las series caóticas. Ese sesgo es mayor cuanto más pequeña es la serie, por lo que es necesario tomar con precaución las series de cinco minutos.

En las series de rendimientos con frecuencia operación a operación la dimensión sólo pasa de dos en los residuos de los modelos EGARCH, que son en todos los casos la cota superior. Las dimensiones son muy similares entre los distintos vencimientos siendo los valores más altos los del vencimiento de diciembre y los más bajos los de junio del 98.

En las series de rendimientos con frecuencia de cinco minutos, los valores de la dimensión de correlación son en todo caso superiores a los de frecuencia operación a operación. Los valores más bajos se dan para las series de rendimientos y para las de residuos de los modelos ARIMA. Los valores de los residuos de los modelos GARCH, y EGARCH son muy similares, aunque ligeramente superiores para el caso EGARCH.

Comprobamos que en el caso de la serie de residuos del modelo ARIMA sobre los rendimientos con frecuencia de cinco minutos con vencimiento en diciembre, en el que teníamos la serie intervenida y sin intervenir, los resultados sin intervención son más bajos que con la serie intervenida.

³² Sayers, 1989.

Las series de duraciones, presentan dimensiones de correlación bajas (entre 2'2 y 2'6) para las series originales lo que es coherente con la presencia de memoria a largo plazo. Para las series diferenciadas fraccionalmente la dimensión de correlación aumenta y en la de los residuos de los modelos ARFIMA la dimensión alcanza el máximo que si bien es alta (el máximo es 5'5) resulta un valor bajo para haber sido obtenido con una dimensión de inmersión 10 con 160.000 datos.

Los resultados que hemos obtenido se deben a la utilización de una región determinada de forma subjetiva. Una alternativa es escoger la región en función de un criterio objetivo. En las Tablas 5.4.7, 5.4.8 y 5.4.9 se incluyen los resultados si utilizamos como región crítica los valores comprendidos entre un medio y dos veces la desviación típica. Los resultados que obtenemos en estos casos son mucho más irregulares pues corresponden a distintas regiones que las descritas por Smith. Con lo que preferimos quedarnos con los resultados que obteníamos en las Tablas anteriores.

Tabla 5.4.7: Dimensiones de correlación con desviaciones típicas de las estimaciones. Región tomada en función de la desviación típica de la serie operación a operación

	m=	2	3	4	5	6	7	8	9	10
Series originales	marzo	0,8729	1,3315	1,7715	2,1634	2,5897	2,9516	3,4066	3,8378	4,2502
		(0,2375)	(0,3623)	(0,4821)	(0,5887)	(0,7047)	(0,8032)	(0,9271)	(1,0444)	(1,1566)
	junio	0,3027	0,4505	0,5716	0,6990	0,8327	0,9528	1,0654	1,1981	1,3339
		(0,1154)	(0,1718)	(0,2180)	(0,2665)	(0,3175)	(0,3633)	(0,4063)	(0,4569)	(0,5086)
	septiembre	0,3522	0,5349	0,7008	0,8328	1,0129	1,1530	1,2762	1,4224	1,5459
		(0,0869)	(0,1320)	(0,1729)	(0,2055)	(0,2499)	(0,2845)	(0,3149)	(0,3510)	(0,3814)
	diciembre	0,1771	0,2505	0,3150	0,3777	0,4398	0,4933	0,5425	0,6134	0,6787
		(0,0817)	(0,1138)	(0,1468)	(0,1808)	(0,2111)	(0,2405)	(0,2678)	(0,3092)	(0,3496)
Series de residuos modelo ARIMA	marzo	0,8945	1,3341	1,7847	2,2300	2,6347	3,0978	3,5131	3,9678	4,4014
		(0,1333)	(0,2015)	(0,2698)	(0,3389)	(0,4025)	(0,4733)	(0,5372)	(0,6067)	(0,6702)
	junio	0,3888	0,5847	0,7614	0,9406	1,0998	1,2740	1,4482	1,6004	1,7465
		(0,0851)	(0,1283)	(0,1695)	(0,2111)	(0,2503)	(0,2904)	(0,3310)	(0,3718)	(0,4091)
	septiembre	0,4113	0,6003	0,7817	0,9604	1,1273	1,2997	1,4816	1,6535	1,8212
		(0,0641)	(0,0950)	(0,1241)	(0,1549)	(0,1842)	(0,2145)	(0,2477)	(0,2779)	(0,3079)
	diciembre	0,8126	1,2094	1,6004	1,9847	2,3640	2,7516	3,1222	3,4564	3,8243
		(0,2534)	(0,3803)	(0,5083)	(0,6390)	(0,7637)	(0,8986)	(1,0219)	(1,1344)	(1,2564)
Series de residuos modelo GARCH-M	marzo	1,6657	2,4393	3,1542	3,8480	4,6268	5,2441	5,9835	6,8444	7,4313
		(0,1911)	(0,2633)	(0,3361)	(0,4100)	(0,4970)	(0,5556)	(0,6536)	(0,7630)	(0,8444)
	junio	0,4629	0,6766	0,8416	1,0120	1,1668	1,3239	1,4838	1,6718	1,8360
		(0,0743)	(0,1110)	(0,1426)	(0,1790)	(0,2157)	(0,2442)	(0,2741)	(0,3093)	(0,3463)
	septiembre	0,4199	0,6948	0,7785	0,9740	1,1925	1,3394	1,5211	1,6799	1,8576
		(0,0397)	(0,0746)	(0,0739)	(0,0926)	(0,1122)	(0,1331)	(0,1508)	(0,1714)	(0,2025)
	diciembre	0,6195	0,9204	2,0971	1,5247	1,8079	2,0971	2,4360	2,6903	2,9553
		(0,0676)	(0,1008)	(0,2483)	(0,1787)	(0,2119)	(0,2483)	(0,2890)	(0,3260)	(0,3687)
Series de residuos modelo EGARCH	marzo	0,8067	1,3125	1,8116	2,4198	3,0637	3,5351	4,1059	4,7585	5,3230
		(0,0377)	(0,0278)	(0,0344)	(0,0448)	(0,0535)	(0,0528)	(0,0620)	(0,0646)	(0,0711)
	junio	0,4176	0,6193	0,7814	0,9275	1,0699	1,2319	1,3784	1,5303	1,6714
		(0,0645)	(0,0967)	(0,1237)	(0,1548)	(0,1878)	(0,2150)	(0,2404)	(0,2726)	(0,3088)
	septiembre	0,4820	0,7102	0,9463	1,1434	1,3732	1,5157	1,6767	1,8622	2,0516
		(0,0246)	(0,0314)	(0,0353)	(0,0402)	(0,0524)	(0,0623)	(0,0717)	(0,0804)	(0,0856)
	diciembre	0,1899	0,2846	0,3884	0,4641	0,5584	0,6347	0,7088	0,7890	0,8938
		(0,0106)	(0,0194)	(0,0388)	(0,0468)	(0,0551)	(0,0653)	(0,0742)	(0,0834)	(0,0925)

Tabla 5.4.8: Dimensiones de correlación con desviaciones típicas de las estimaciones. Región tomada en función de la desviación típica de las series de cinco minutos

	m=	2	3	4	5	6	7	8	9	10
Series originales	marzo	1,0178	1,4391	1,8812	2,2300	2,6202	2,9515	3,3508	3,6634	4,0357
		(0,1228)	(0,1772)	(0,2335)	(0,2790)	(0,3296)	(0,3720)	(0,4224)	(0,4557)	(0,4897)
	junio	1,5495	2,1889	2,8477	3,3242	3,7494	3,9981	4,1460	4,2369	4,3176
		(0,1919)	(0,2738)	(0,3560)	(0,4087)	(0,4389)	(0,4356)	(0,3980)	(0,3584)	(0,3193)
	septiembre	0,8329	1,1615	1,4622	1,7267	1,9744	2,2222	2,4714	2,7106	2,9501
		(0,0613)	(0,0876)	(0,1126)	(0,1346)	(0,1547)	(0,1756)	(0,1968)	(0,2172)	(0,2381)
	diciembre	0,5250	0,7503	0,9528	1,1436	1,3220	1,5290	1,7175	1,8946	2,0643
		(0,0680)	(0,0985)	(0,1251)	(0,1514)	(0,1744)	(0,2023)	(0,2283)	(0,2497)	(0,2726)
Series de residuos modelo ARIMA	marzo	1,0394	1,5010	1,9264	2,3289	2,7145	3,0961	3,4807	3,8644	4,2521
		(0,0812)	(0,1218)	(0,1604)	(0,1980)	(0,2340)	(0,2685)	(0,3019)	(0,3312)	(0,3570)
	junio	1,5305	2,2007	2,8131	3,3408	3,7341	3,9898	4,1469	4,2537	4,3412
		(0,1898)	(0,2752)	(0,3510)	(0,4106)	(0,4362)	(0,4271)	(0,3954)	(0,3565)	(0,3215)
	septiembre	0,8582	1,1906	1,4837	1,7544	2,0091	2,2517	2,4859	2,7086	2,9261
		(0,0638)	(0,0899)	(0,1136)	(0,1358)	(0,1572)	(0,1779)	(0,1979)	(0,2169)	(0,2353)
	diciembre(*)	0,0215	0,0303	0,0416	0,0527	0,0629	0,0708	0,0800	0,0893	0,0978
		(0,0034)	(0,0049)	(0,0062)	(0,0074)	(0,0087)	(0,0099)	(0,0113)	(0,0123)	(0,0134)
Series de residuos modelo GARCH	marzo	0,3936	0,5704	0,7627	0,9271	1,0964	1,2525	1,4252	1,6104	1,7900
		(0,0409)	(0,0586)	(0,0784)	(0,0988)	(0,1159)	(0,1346)	(0,1517)	(0,1738)	(0,1960)
	junio	1,3198	1,9695	2,5890	3,2344	3,8085	4,2873	4,7334	5,0470	5,3064
		(0,0504)	(0,0798)	(0,1103)	(0,1422)	(0,1741)	(0,1971)	(0,2055)	(0,1816)	(0,1613)
	septiembre	1,4427	2,1507	2,8630	3,5429	4,1666	4,6470	5,0205	5,1721	5,2332
		(0,0665)	(0,0997)	(0,1307)	(0,1579)	(0,1810)	(0,1902)	(0,2011)	(0,2078)	(0,2326)
	diciembre(**)	1,1218	1,6608	2,1942	2,7389	3,2270	3,7568	4,2477	4,7371	5,2671
		(0,0524)	(0,0807)	(0,1110)	(0,1435)	(0,1774)	(0,2145)	(0,2485)	(0,2792)	(0,3065)
Series de residuos modelo EGARCH	marzo	1,1719	1,7514	2,3062	2,8648	3,4282	3,9305	4,3995	4,8170	5,1727
		(0,0633)	(0,0961)	(0,1263)	(0,1564)	(0,1814)	(0,1995)	(0,2196)	(0,2253)	(0,2307)
	junio	1,2176	1,8276	2,3613	2,9189	3,4182	3,8964	4,3966	4,8399	4,8399
		(0,0455)	(0,0667)	(0,0856)	(0,1055)	(0,1204)	(0,1352)	(0,1517)	(0,1645)	(0,1645)
	septiembre	1,4447	2,1799	2,8713	3,5320	4,1599	4,6975	5,0831	5,2943	5,3930
		(0,0642)	(0,0976)	(0,1269)	(0,1496)	(0,1726)	(0,1788)	(0,1711)	(0,1692)	(0,1848)
	diciembre	1,3749	2,0558	2,7508	3,4441	4,0868	4,6640	5,1144	5,4263	5,6164
		(0,0628)	(0,0946)	(0,1313)	(0,1630)	(0,1924)	(0,2037)	(0,1966)	(0,1818)	(0,1690)

(*) En el capítulo cuarto habíamos preferido el modelo ARIMA con intervención, pero aquí analizamos ambos para comprobar si afecta al análisis la intervención.

(**) En el vencimiento de diciembre el modelo estimado era un GARCH-M.

Tabla 5.4.9: Dimensiones de correlación con desviaciones típicas de las estimaciones. Región tomada en función de la desviación típica de las series de duraciones

	m=	2	3	4	5	6	7	8	9	10
Series originales	marzo	0,3559	0,5103	0,6297	0,7417	0,8301	0,9602	1,0613	1,1350	1,2238
		(0,0145)	(0,0217)	(0,0269)	(0,0324)	(0,0367)	(0,0427)	(0,0469)	(0,0518)	(0,0547)
	junio	0,1782	0,2624	0,3327	0,3782	0,4431	0,5484	0,6272	0,6928	0,7271
		(0,0042)	(0,0118)	(0,0130)	(0,0128)	(0,0134)	(0,0127)	(0,0173)	(0,0195)	(0,0196)
	septiembre	0,3201	0,4401	0,5813	0,7007	0,8094	0,8996	1,0147	1,0767	1,2287
		(0,0056)	(0,0073)	(0,0170)	(0,0248)	(0,0288)	(0,0316)	(0,0366)	(0,0364)	(0,0512)
	diciembre	0,3504	0,4993	0,6310	0,7517	0,8562	0,9334	1,0210	1,1165	1,2622
		(0,0125)	(0,0139)	(0,0180)	(0,0220)	(0,0295)	(0,0319)	(0,0380)	(0,0420)	(0,0511)
Series de duraciones con diferencia fraccional	marzo	0,4339	0,5932	0,7286	0,8886	1,0214	1,1276	1,2385	1,3065	1,4193
		(0,0236)	(0,0327)	(0,0385)	(0,0442)	(0,0447)	(0,0521)	(0,0586)	(0,0653)	(0,0705)
	junio	0,2385	0,3118	0,3664	0,4328	0,5014	0,5571	0,6164	0,6732	0,7315
		(0,0163)	(0,0188)	(0,0202)	(0,0284)	(0,0283)	(0,0336)	(0,0354)	(0,0370)	(0,0359)
	septiembre	0,3880	0,5166	0,6723	0,7929	0,9234	0,9952	1,1223	1,2035	1,3165
		(0,0180)	(0,0241)	(0,0277)	(0,0326)	(0,0404)	(0,0397)	(0,0456)	(0,0502)	(0,0585)
	diciembre	0,3476	0,4808	0,6145	0,6973	0,8130	0,8712	0,9849	1,0625	1,1742
		(0,0184)	(0,0273)	(0,0323)	(0,0361)	(0,0423)	(0,0449)	(0,0515)	(0,0569)	(0,0664)
Series de residuos modelo ARFIMA	marzo	0,4516	0,6250	0,8018	0,9366	1,0973	1,2218	1,3809	1,5437	1,6735
		(0,0261)	(0,0395)	(0,0499)	(0,0548)	(0,0636)	(0,0698)	(0,0805)	(0,0930)	(0,1029)
	junio	0,1984	0,2761	0,3679	0,4355	0,5012	0,6062	0,6505	0,6685	0,6950
		(0,0098)	(0,0171)	(0,0175)	(0,0202)	(0,0220)	(0,0225)	(0,0203)	(0,0195)	(0,0156)
	septiembre	0,3044	0,4641	0,5573	0,6600	0,7186	0,8317	0,9039	0,9699	1,0539
		(0,0147)	(0,0185)	(0,0194)	(0,0297)	(0,0316)	(0,0328)	(0,0362)	(0,0379)	(0,0405)
	diciembre	0,3472	0,4730	0,6432	0,7920	0,8913	0,9614	1,0738	1,1679	1,2640
		(0,0140)	(0,0204)	(0,0243)	(0,0346)	(0,0411)	(0,0396)	(0,0413)	(0,0451)	(0,0470)

5.4.2. EXPONENTES DE LYAPUNOV.

Hasta ahora hemos basado nuestro análisis en la existencia de puntos periódicos densos en torno al atractor. Los exponentes de Lyapunov tratan de medir otra de las características de un sistema caótico: la sensibilidad a las condiciones iniciales.

Lyapunov³³ define el exponente característico para cuantificar el grado de estabilidad de un sistema dinámico midiendo la tasa de divergencia de trayectorias próximas.

Lyapunov establece que dado un sistema de ecuaciones diferenciales, una solución $x(t)$ será estable si soluciones que empiezan próximas a ella, continúan

³³ A.M. Lyapunov, 1992.

estándolo con el paso del tiempo. O bien, expresado matemáticamente como en [5.4.16].

$$\forall \epsilon > 0 \exists \delta > 0 / \text{ si } |x'(0) - x(0)| < \delta \text{ entonces } |x'(t) - x(t)| < \epsilon \quad [5.4.16]$$

Si estudiamos un sistema lineal de ecuaciones diferenciales encontraremos una solución general de la forma de [5.4.17]. Para estudiar la estabilidad de esta solución bastará con estudiar el signo de sus exponentes característicos, λ_i . La presencia de exponentes negativos supone que la solución es estable. Si aparecen exponentes positivos estaremos ante una solución inestable.

$$x(t) = \sum_i C_i e^{\lambda_i t} \quad [5.4.17]$$

Para el caso de sistemas de ecuaciones diferenciales no lineales, su significado no es idéntico, pues la presencia de exponentes positivos, no indica inestabilidad, sino comportamiento caótico. El cálculo de los exponentes característicos también resulta más complicado. Para ello tenemos que linealizar el sistema para obtener unos exponentes que van a presentar las mismas cualidades descriptivas de la estabilidad del sistema que presentaban los exponentes característicos en el caso de los sistemas lineales.

Consideremos³⁴ una órbita $x(t)$, correspondiente a un sistema dinámico no lineal en un espacio de fase m -dimensional que representamos por [5.4.18].

$$\dot{x} = F(x) \quad [5.4.18]$$

Definimos por \mathbf{x} las pequeñas desviaciones respecto a esta trayectoria. Si linealizamos la ecuación [5.4.18], para \mathbf{x} tendremos unas ecuaciones de la forma de [5.4.19].

$$\dot{\mathbf{x}} = T(x_t) \mathbf{x} \quad [5.4.19]$$

Donde T es el jacobiano de F . Este jacobiano [5.4.20] es el sistema lineal en el que hemos transformado el sistema no lineal. Nuestro objetivo va a ser calcular los exponentes del sistema lineal.

$$T = DF = \frac{\partial F}{\partial x} \quad [5.4.20]$$

El exponente de Lyapunov se define, entonces, a partir de la ecuación [5.4.21].

$$I(x(0), \mathbf{x}(0)) = \lim_{t \rightarrow \infty} \frac{1}{t} \ln \frac{\|\mathbf{x}(t)\|}{\|\mathbf{x}(0)\|} \quad [5.4.21]$$

³⁴ Seguimos aquí a M. Sano e Y. Sawada, 1985.

El exponente de Lyapunov definido en [5.4.21] nos estaría indicando la tasa de variación de la distancia entre dos trayectorias próximas, o como la definen Eckmann y Ruelle³⁵, la razón exponencial a la que una perturbación en el estado inicial de una serie, crece o disminuye. En [5.4.22] podemos ver la expresión del exponente. Si definimos D como la distancia relativa entre dos trayectorias próximas, esto es, la razón que aparece en el logaritmo de la expresión [5.4.21], podremos expresar esta distancia en función de λ , el exponente de Lyapunov.

$$D = e^{\lambda t} \quad [5.4.22]$$

Si λ es positivo, la distancia entre las trayectorias aumenta con el tiempo, y si λ es negativo, la distancia disminuye. Si el sistema es estable, las trayectorias deben tender hacia un atractor, y por tanto la distancia entre ellas debe disminuir. Si el sistema es inestable, las trayectorias no tienden a ningún atractor, y la distancia entre trayectorias próximas tenderá a aumentar.

Va a haber tantos exponentes de Lyapunov como dimensiones tenga el espacio de fases en que venga definido el sistema que estemos estudiando. El estudio original de Lyapunov establecía que todo sistema en que exista algún exponente positivo era inestable. Sin embargo, también señala que si existen exponentes tanto positivos como negativos, el sistema va a mantener cierta estabilidad³⁶. Adaptando esto al estudio de sistemas caóticos, si existe un atractor en el sistema, sea este puntual, cíclico, o extraño,

³⁵ J.P. Eckmann y D. Ruelle, 1985.

³⁶ A.M. Lyapunov, 1992. p. 91.

la suma de los exponentes debe ser negativa, puesto que las trayectorias tienden a quedar confinadas en el atractor.

En los atractores extraños, sin embargo, a pesar de que el sistema en su conjunto se contraiga, en determinadas direcciones se expande, siendo esto la causa de que aparezca la sensibilidad a las condiciones iniciales. En estas direcciones en las que el sistema se expande, los exponentes serán positivos. Los exponentes de Lyapunov nos sirven, por tanto, para indicar la presencia de sensibilidad a las condiciones iniciales, y con ello, de comportamiento caótico.

El hecho de que un sistema tenga exponentes positivos y negativos y que el sistema se este contrayendo se puede observar gráficamente en la figura 5.4.5. En esta figura se considera inicialmente una circunferencia, que con el paso del tiempo evoluciona y se deforma.

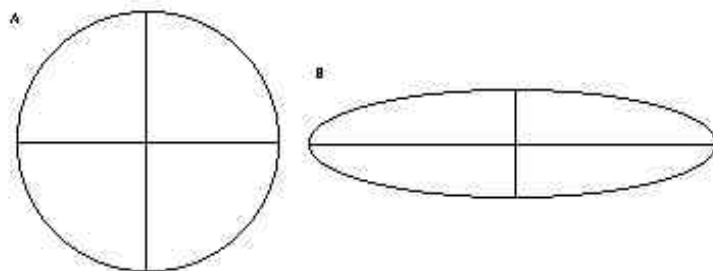


Figura 5.4.5: El sistema evoluciona de A a B. En su conjunto, el sistema se ha contraído, pero en la componente horizontal se ha expandido.

La figura 5.4.5 es una circunferencia que está definida en un espacio de fases de dos dimensiones. A estas dos dimensiones les corresponden dos exponentes característicos, uno negativo, que se identifica con la componente vertical, indicando que se ha contraído, y otro positivo, correspondiente a la componente horizontal, como

consecuencia de su expansión. La suma de los dos debe ser negativa si el sistema tiende a un equilibrio.

La definición de exponente de Lyapunov dada en [5.4.21] nos va a proporcionar un único exponente. En un sistema dinámico es posible definir tantos exponentes como dimensiones tiene el espacio de fases, de tal forma que cada exponente nos va indicar si en esa dimensión el sistema se expande o se contrae. Va a ser necesario, por tanto, establecer una nueva definición de exponentes de Lyapunov.

Definimos una base m -dimensional $\{e_i\}$ (donde m es la dimensión del espacio de fases). En este caso λ_i será la tasa de variación de la distancia entre dos trayectorias próximas en la dirección del vector e_i de la base que hemos definido. Al conjunto de estos exponentes λ_i ordenadas de mayor a menor $\lambda_1, \lambda_2, \dots, \lambda_m$ se les denomina espectro de los exponentes de Lyapunov.

Cuando trabajamos con una serie temporal, el primer problema es que no tenemos el sistema de ecuaciones, con lo que éste es nuestro primer objetivo a la hora de calcular los exponentes. No disponemos de varias soluciones para calcular la velocidad a la que estas se separan. Una única solución debe servir para el cálculo. Para obtener las series con las que comparar se utiliza esa misma serie en momentos de tiempo posteriores, en los que volvemos a pasar por las proximidades de la serie.

Existen multitud de algoritmos que calculan los exponentes de Lyapunov de una serie temporal. Algunos buscan el espectro completo de exponentes de Lyapunov. Para

ello estiman un modelo sobre el que calculan los exponentes. Estos modelos³⁷ pueden ser desde redes neuronales³⁸, funciones de base radial³⁹ a funciones localmente lineales como Sano y Sawada⁴⁰, Eckmann y Ruelle⁴¹. Estos dos algoritmos, los más habituales, son muy similares, y van a proporcionar el mismo valor para el máximo exponente, difiriendo sólo en los siguientes⁴². Otros se concentran en el máximo exponente, que es el relevante para decidir si estamos ante una serie aleatoria o caótica. Estos son los casos de Wolf *et al.*⁴³, Ellner *et al.*⁴⁴ o Rosenstein *et al.*⁴⁵. De entre todos los más corrientes son los de Eckmann y Ruelle entre los primeros y el de Wolf entre los segundos. Nosotros hemos escogido el primero, porque proporciona más información y porque la forma de calcularlo muestra elementos comunes con la predicción por métodos locales que veremos en el capítulo siguiente y de la que este algoritmo es una de sus referencias.

De nuevo partimos de una serie temporal $\{x_i\}_{i \in \mathbb{N}}$. Reconstruimos el sistema transformando la serie en vectores de m -historias. Para obtener las soluciones próximas a la órbita producida por las m -historias buscamos en la misma serie. Escogemos un punto \bar{x}_i . A partir de este punto, buscamos vecinos de ese punto, que denominamos \bar{x}_j , y que están contenidos en una bola de radio l centrada en \bar{x}_i . Este criterio queda expresado en [5.4.23]⁴⁶.

³⁷ McCaffrey, Ellner, Gallant y Nychka, 1992.

³⁸ Casdagli, 1989.

³⁹ Smith, 1992.

⁴⁰ Sano y Sawada, 1985.

⁴¹ Eckmann y Ruelle, 1985.

⁴² Karantonis y Pagitsas, 1996.

⁴³ Wolf, Swift, Swinney y Vastano, 1985.

⁴⁴ Ellner, Gallant, McCafrey y Nychka, 1991.

⁴⁵ Rosenstein, Collins, y De Luca, 1993.

⁴⁶ donde $\|\dots\|$ es la norma de Takens ($\|\bar{x}_j - \bar{x}_i\| = \max \{x_{j+a} - x_{i+a}\}$)

$$\| \bar{x}_j - \bar{x}_i \| \ll 1 \quad [5.4.23]$$

Dejamos evolucionar el sistema tanto con \bar{x}_i como con \bar{x}_j . Para calcular los exponentes necesitamos linealizar el sistema, cosa que hacemos en [5.4.24]. A través de la matriz T_i aproximamos la forma en que las órbitas se acercan o se separan.

$$T_i (\bar{x}_j - \bar{x}_i) \approx \bar{x}_{j+t} - \bar{x}_{i+t} \quad [5.4.24]$$

Dando el valor t al paso de tiempo la matriz T_i tendrá la forma de la expresión [5.4.25]. Sólo tendremos que estimar los coeficientes de la última fila, pues el paso de una m -historia a la siguiente supone correr cada componente una posición, lo que hacemos en las $m-1$ primeras filas de la matriz, siendo la última fila una regresión del próximo componente en función de los m -anteriores.

$$T_i = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ a_1 & a_2 & a_3 & \dots & a_m \end{pmatrix} \quad [5.4.25]$$

Para estimar los valores de los a_i que forman la matriz basta con utilizar mínimos cuadrados ordinarios para estimar los valores de la última fila de la matriz.

El valor l de [5.4.23] cambia con cada nuevo \bar{x}_i . A la hora de elegir l hay que tener en cuenta que este ha de ser suficientemente pequeño para asegurarnos que estamos considerando órbitas próximas, pero suficientemente grande para que haya suficientes observaciones como para estimar los parámetros de la regresión.

Tenemos así una matriz T_i para cada \bar{x}_i . Hay un exponente por cada dirección. Es necesario determinar las direcciones de variación de los exponentes. Según la base en la que estemos, tendremos unos exponentes diferentes. Vamos a llamar Q_{i-1} a la base adecuada para el cálculo de los exponentes. Esta base busca la dirección en la que se produce la máxima expansión. Tendremos que ir variando la base a medida que avanzamos a lo largo de la serie. Para lograr esto multiplicamos ambas matrices, $T_i Q_{i-1}$, y descomponemos la nueva matriz en dos matrices Q_i y R_i , siendo la primera de ellas ortogonal y la segunda triangular con elementos positivos diagonales. La matriz Q_i nos va a servir de base para la siguiente matriz de transformación T_{i+1} . El proceso es de la forma [5.4.26].

$$\begin{array}{l} T_1 \ Q_0 = Q_1 \ R_1 \\ T_2 \ Q_1 = Q_2 \ R_2 \\ \\ T_{j+1} \ Q_j = Q_{j+1} \ R_{j+1} \end{array} \qquad [5.4.26]$$

Los exponentes de Lyapunov I_k vienen dados entonces por los elementos de la diagonal principal de las matrices R_i [5.4.27]. En esta expresión, k es el número de matrices disponibles.

$$\lambda^m_k = \frac{1}{\tau_k} \sum_{j=0}^{k-1} \ln R_{(j)kk} \quad [5.4.27]$$

Los máximos exponentes así calculados aparecen en las Tablas 5.4.10 para los rendimientos del contrato de Futuro sobre el Bono Nocional a 10 años con frecuencia operación a operación, 5.4.11 para la frecuencia de cinco minutos y 5.4.12 para las series de duraciones. Los espectros completos aparecen en las figuras de la A.339 a la A.383 del apéndice gráfico.

Tabla 5.4.10: Máximo exponente de Lyapunov de las series de rendimientos, operación a operación.

	m=	2	3	4	5	6	7	8	9	10
Series originales	marzo	-0,1507	-0,1769	-0,1445	-0,1576	-0,2036	-0,0896	-0,1333	-0,1149	-0,1421
	junio	-0,1624	-0,3207	-0,1377	-0,2545	-0,0273	-0,0376	-0,0728	-0,1316	-0,1194
	septiembre	-0,2242	-0,1761	-0,1476	-0,2337	-0,1246	-0,0897	-0,1235	-0,2169	-0,1484
	diciembre	-0,1130	-0,2153	-0,0643	-0,0581	0,0004	-0,0967	-0,0398	-0,1425	-0,1094
Series de residuos modelo ARIMA	marzo	3,3427	1,2138	0,5240	0,2958	0,2594	0,1282	0,1101	-0,0016	0,0022
	junio	5,4132	2,9531	1,2442	0,9217	0,2172	0,1647	0,0686	0,0750	0,0204
	septiembre	7,3140	8,2785	6,1789	3,9251	1,4233	0,2359	0,1055	0,0382	0,0352
	diciembre	2,2361	1,2788	0,7942	0,5512	0,3357	0,2771	0,1176	0,1069	0,0610
Series de residuos modelo GARCH-M	marzo	3,8891	2,2910	1,6108	0,5756	0,5121	0,3784	0,1786	0,0402	0,1740
	junio	3,5046	1,3817	0,9727	0,7944	0,4417	0,2454	0,0277	0,0566	0,0748
	septiembre	3,0023	1,7976	1,1191	0,9916	0,4907	0,2209	0,3838	0,2938	0,0332
	diciembre	3,0695	2,1129	0,9190	0,6371	0,4592	0,1742	0,1080	0,0615	-0,0438
Series de residuos modelo EGARCH	marzo	2,7048	1,8946	0,9444	0,5927	0,4568	0,3418	0,1687	0,0861	0,1459
	junio	5,9624	2,7665	1,0168	0,5984	0,4141	0,4096	0,1206	0,1483	0,1375
	septiembre	1,7688	0,8327	0,5972	0,4285	0,4400	0,2673	0,1053	0,0908	0,0790
	diciembre	1,5721	0,8441	0,5790	0,6705	0,4180	0,4192	0,3861	0,2397	0,1194

Para la regresión consideramos 200 vecinos.

En las series de rendimientos con frecuencia operación a operación se puede observar que existen exponentes positivos, salvo para las series originales de rendimientos que son negativos. Esto puede deberse a la estructura de rejilla que habíamos observado en el comienzo del capítulo en las series originales y que ocultaban

cualquier estructura que intentáramos detectar en la serie. Una vez filtrada las series estas estructuras se hacen patentes, y nos surgen los exponentes positivos. Sin embargo, las series de residuos son positivas con exponentes tan bajos que prácticamente valen cero.

Tabla 5.4.11: Máximo exponente de Lyapunov de las series de rendimientos con frecuencia de cinco minutos.

	m=	2	3	4	5	6	7	8	9	10
Series originales	marzo	4,8237	2,5426	0,9272	0,8519	0,6640	0,6915	0,6302	0,8240	1,1609
	junio	0,0667	0,1352	0,1209	0,1419	0,1312	0,2502	0,2675	0,3804	0,7515
	septiembre	0,0338	0,1186	0,1963	0,1226	0,1616	0,2352	0,2505	0,3511	0,7891
	diciembre	0,1881	0,1484	0,1076	0,0709	0,1574	0,1878	0,2797	0,3893	1,0786
Series de residuos modelo ARIMA	marzo	2,3976	1,4774	0,8709	0,5473	0,4982	0,4695	0,5551	0,6365	1,1562
	junio	2,9648	-0,1638	0,2618	0,0985	0,0156	0,0154	0,0031	0,0143	0,0011
	septiembre	4,7566	2,6393	0,9377	0,6272	0,8326	0,7418	0,7453	0,8948	1,2766
	diciembre	1,8799	1,0532	0,6703	0,5440	0,4887	0,5896	0,5420	0,6809	1,0155
	diciembre intervenido(*)	2,4410	1,4509	0,7375	0,4996	0,4940	0,4769	0,4874	0,5897	1,0786
Series de residuos modelo GARCH	marzo	1,6491	0,7528	0,5432	0,5592	0,5058	0,4447	0,5404	0,6441	1,1139
	junio	2,5606	1,6346	0,7934	0,6260	0,6191	0,5034	0,6183	0,8196	1,2044
	septiembre	2,5606	1,6346	0,7934	0,6260	0,6191	0,5034	0,6183	0,8196	1,2044
	diciembre(**)	1,7437	0,9455	0,6029	0,5114	0,4851	0,4892	0,4680	0,5711	1,1118
Series de residuos modelo EGARCH	marzo	1,6043	0,9175	0,4921	0,5544	0,4495	0,4779	0,5139	0,6492	1,0437
	junio	1,3924	0,7030	0,4083	0,4894	0,4501	0,5181	0,6181	0,6855	1,1186
	septiembre	1,4272	0,9309	0,4447	0,4825	0,4974	0,5534	0,5350	0,6901	1,1808
	diciembre	1,7437	0,9455	0,6029	0,5114	0,4851	0,4892	0,4680	0,5711	1,1118

Para la regresión consideramos 50 vecinos.

(*) En el capítulo cuarto habíamos preferido el modelo ARIMA con intervención, pero aquí analizamos ambos para comprobar si afecta al análisis la intervención.

(**) En el vencimiento de diciembre el modelo estimado era un GARCH-M.

Las series de rendimientos con frecuencia de cinco minutos presentan en todos los casos exponentes positivos y muy superiores a sus equivalentes con frecuencia operación a operación. En las series de residuos ARIMA no se observan grandes diferencias entre la serie intervenida y sin intervenir con vencimiento en diciembre. Tampoco hay grandes diferencias entre distintos modelos o entre distintos vencimientos.

Tabla 5.4.12: Máximo exponente de Lyapunov de las series de duraciones.

	m=	2	3	4	5	6	7	8	9	10
Series originales	marzo	1,5199	1,3582	1,0575	0,9544	0,8178	0,6106	0,5620	0,3974	0,3698
	junio	1,4893	1,3419	1,2177	0,9233	0,7961	0,6647	0,5067	0,4677	0,3473
	septiembre	2,2119	1,8448	1,2630	1,0105	0,8596	0,7977	0,6166	0,5460	0,4843
	diciembre	1,5975	1,1883	0,9890	0,8364	0,6616	0,4781	0,5606	0,4251	0,3548
Series con diferencias fraccionales	marzo	2,1043	0,9792	0,6669	0,4549	0,3494	0,2577	0,2304	0,1718	0,1569
	junio	2,0632	1,1825	0,7555	0,5120	0,4095	0,3454	0,2915	0,2305	0,2161
	septiembre	2,1602	1,0506	0,7821	0,5855	0,4327	0,3048	0,2237	0,2368	0,2348
	diciembre	1,9930	1,0174	0,6132	0,4565	0,4156	0,2376	0,2312	0,1779	0,1605
Series de residuos modelo ARFIMA	marzo	2,2176	1,1310	0,7430	0,4778	0,3981	0,3687	0,2519	0,2327	0,1893
	junio	2,1297	1,1925	0,8199	0,6234	0,4649	0,3424	0,3106	0,2843	0,1930
	septiembre	2,1532	1,2573	0,8272	0,5688	0,5483	0,3831	0,2866	0,3391	0,2419
	diciembre	2,0657	0,9939	0,6712	0,4016	0,3880	0,2504	0,2550	0,1954	0,2019

Para la regresión consideramos 200 vecinos.

En las series de duraciones de operaciones, obtenemos máximos exponentes siempre positivos. Los resultados son más altos en las series originales que en las series diferenciadas fraccionalmente o en las de residuos de los modelos ARFIMA. Los resultados no son muy distintos de un vencimiento a otro.

Los resultados de todas las series, aún que nos hayan salido positivos no podemos considerarlos como definitivos para establecer la presencia de comportamiento caótico, pues incluso hay procesos estocásticos que también pueden presentar exponentes positivos cercanos a cero⁴⁷.

5.4.3 ENTROPÍA DE KOLMOGOROV.

La dimensión de correlación y los exponentes de Lyapunov son las dos formas principales de estudiar si un sistema es caótico. Ambas magnitudes se relacionan entre

⁴⁷ Tanaka, Aihara y Taki, 1996.

sí a través de la entropía de Kolmogorov. El cálculo de la entropía no aporta mucha más información de la que ya tenemos, aunque sí sobre el sistema como tal.

La entropía es un concepto que surge de la segunda ley de la Termodinámica. Esta establece que la entropía del universo es creciente. En termodinámica la entropía mide el grado de desorden de un sistema. Los sistemas no se ordenan espontáneamente, sino que tienden a desordenarse. El concepto de entropía introduce la dirección del *tiempo* en la ciencia. El tiempo transcurrirá en la dirección en la que la entropía aumenta, no siendo posible la dirección contraria. Un momento es anterior a otro si su entropía es menor.

La segunda ley de la termodinámica impide que los jarrones rotos se recompongan o que el agua fluya en los ríos desde los mares hasta las montañas. Podemos recomponer lo que habíamos señalado sobre sistemas disipativos, conservativos y expansivos en términos físicos a través de la entropía. Los sistemas disipativos tienden al equilibrio y tienen una entropía creciente. Los sistemas conservativos permanecen en el estado en que se encuentran y tienen una entropía constante. Los sistemas expansivos que huyen del equilibrio equivalen a sistemas con entropía creciente.

La entropía puede considerarse como una medida del grado en que un sistema es predecible. Un sistema perfectamente ordenado (determinista) tendrá una entropía baja, mientras que un sistema completamente desordenado, tendrá una entropía muy alta. La

entropía nos indica, entonces, la cantidad de información extra sobre los próximos valores de la serie, que da el continuar observando la serie.

En una serie determinista de la forma:

$$0, 1, 2, 3, 4, 5, 6, 7, 8, 9, \dots$$

no necesitamos observar su comportamiento futuro para saber cual va a ser éste:

$$\dots 10, 11, 12, 13, 14, \dots$$

Las series deterministas tienen una entropía cero. No supone una mejora de la información el continuar observándola. Si la serie fuese aleatoria, por ejemplo:

$$4, 3, 8, 3, 1, 3, 9, 2, 6, 8, 7, 0, 1, 3, 5, \dots$$

no sabemos a priori como continua la serie. La única forma de conocer su comportamiento futuro es observarlo:

$$\dots 3, 0, 8, 7, 0, 0, 1, 3, 8, 2, 4, 0, 2, 1, 3, \dots$$

La información que proporciona el futuro de la serie es fundamental, por lo que la entropía es máxima, tiende a infinito. En una serie caótica como la siguiente que proviene de la función logística⁴⁸:

$$3400, 8976, 3676, 9299, 2606, 7707, 7067, 8289, 5670, 9820, 706, 2627, 7748, \dots$$

podemos llegar a conocer el comportamiento futuro a corto plazo pero la sensibilidad a las condiciones iniciales hace que nuestro conocimiento no sea perfecto, sino que mejore con la observación, con lo que la entropía sería mayor que cero pero menor que la de una serie aleatoria. La entropía en este caso será un número positivo finito.

⁴⁸ Concretamente es consecuencia de redondear al cuarto decimal y multiplicar por diez mil una serie generada con la función $x_{t+1} = 4x_t(1-x_t)$ con valor inicial: 0.34

La serie predicha será:

6979, 8432, 5286, 9967, 130, 516, 1958, 6299, 9324, 2520, 7541, 7416,...

mientras que la real habría sido:

6979, 8432, 5285, 9967, 130, 514, 1951, 6283, 9340, 2462, 7424, 7648,...

Existen múltiples definiciones de entropías, pero por la afinidad con la integral de correlación vamos a estudiar la *Entropía de Kolmogorov*.

En la definición generalizada de dimensión, habíamos pasado de $q=0$ que se correspondía a la denominada Dimensión de Hausdorff-Besicovitch, a la dimensión de correlación, de $q=2$. Para q igual a uno tendríamos un denominador con valor cero, por lo que no podemos usarla directamente. Pero si hacemos que q tienda a uno, obtenemos la expresión [5.4.28].

$$D_1 = \lim_{r \rightarrow 0} \frac{\sum_{i=1}^N p_i \log p_i}{\log r} \quad [5.4.28]$$

A D_1 se le denomina dimensión de información y permite describir la pérdida de información a medida que un sistema caótico evoluciona. El numerador se corresponde con la definición de entropía de una distribución de probabilidad discreta⁴⁹. En este contexto se le denomina *Entropía de Kolmogorov*⁵⁰.

⁴⁹ Es la Entropía de Shannon que definíamos al hablar de información mutua. Shannon, 1948.

⁵⁰ Grassberger y Procaccia, 1983 C.

Sea un sistema dinámico de m variables. Supóngase que el espacio de fase m -dimensional se divide en cubos de volumen l^m . Supóngase que hay un atractor en el espacio de fase y que la trayectoria x_t pertenece a su base de atracción.

Sea $P(i_1, i_2, \dots, i_d)$ la probabilidad conjunta de que x_t esté en el cubo i_1 , x_{t+1} esté en el cubo i_2 , ... y x_{t+d-1} esté en el cubo i_d . La entropía de Kolmogorov será entonces la que aparece en [5.4.29].

$$K = - \lim_{t \rightarrow 0} \lim_{l \rightarrow 0} \lim_{d \rightarrow \infty} \frac{1}{d \cdot t} \sum_{i_1, \dots, i_d} P(i_1, \dots, i_d) \cdot \ln P(i_1, \dots, i_d) \quad [5.4.29]$$

Donde t es el tiempo que transcurre entre observación y observación, y l es el lado del cubo. Dependiendo del valor que tome K podremos caracterizar el sistema de tal manera que si es completamente determinista vale cero, y si es aleatorio tenderá a infinito. Si el sistema es caótico estará en valores intermedios.

Debido a las dificultades que presenta el cálculo de la Entropía de Kolmogorov, Grassberger y Procaccia⁵¹ definen K_2 [5.4.30], con el mismo significado que K . K_2 es una magnitud relacionada con la integral de correlación que nosotros ya hemos calculado. K_2 va a ser una cota superior de la entropía de Kolmogorov⁵².

$$K_2 = - \lim_{l \rightarrow 0} \lim_{d \rightarrow \infty} \frac{1}{d} \cdot \ln \sum_{i_1, \dots, i_d} P^2(i_1, \dots, i_d) \quad [5.4.30]$$

⁵¹ Grassberger y Procaccia, 1983 C.

⁵² Exactamente igual que la Dimensión de correlación es una aproximación a la dimensión de Hausdorff Besicovich dentro de una familia de infinitas dimensiones, la entropía de kolmogorov y K_2 forman parte de una familia de infinitas entropías. Renyi, 1971.

Para $d=1$, nos encontramos con una serie escalar como con las que nos hemos enfrentado hasta ahora. A partir de aquí introducimos de nuevo la integral de correlación, puesto que $\sum_i P_i^2 = C(l)$ donde P_i es la probabilidad de visitar el cubo i , y con i se incluye a todos los cubos necesarios para cubrir el atractor.

Partimos de la serie temporal $\{x_i\}_{i=1}^N$. A partir de la definición de integral de correlación [5.4.12], se puede reescribir de forma que se cumpla la expresión [5.4.31].

$$C'_d(l) \equiv \sum_{i_1, \dots, i_d} P^2(i_1, \dots, i_d) \quad [5.4.31]$$

A partir de [5.4.13] y [5.4.30] podemos expresar, cuando d tiende a infinito y l tiende a cero en [5.4.32].

$$C'_d(l) \sim l^v \exp(-d \mathbf{t} K_2) \quad [5.4.32]$$

De la que es posible despejar K_2 y obtener la expresión [5.4.33].

$$K_{2,d}(l) = \frac{1}{d} \ln \frac{C_d(l)}{C_{d+1}(l)} \quad [5.4.33]$$

En la que $K_{2,d}(\mathbf{e}) \sim K_2$ cuando d tiende a infinito y l tiende a cero. Los resultados que obtenemos pueden verse en las figuras del apéndice gráfico (de la A.384 a la A.430). En todos ellos se observa cómo la entropía llega a una cota. Si bien también se

ve cómo esa cota máxima se alcanza al llegar a una precisión que se corresponde con la primera región que definíamos al tratar la dimensión de correlación y que se corresponde a la zona en la que carecemos de suficientes observaciones como para hacer una buena estimación. Por lo tanto, los resultados obtenidos no se pueden tener demasiado en cuenta.

5.4.4. CONTRASTE BDS.

Hemos definido la integral de correlación $C^m(l)$ en [5.4.12], donde m es la dimensión de inmersión y l es la precisión, como la probabilidad de que dos m -historias se encuentren a una distancia menor a l [5.4.34].

$$C^m(l) = P[|\bar{x}_r - \bar{x}_s| < l] \quad [5.4.34]$$

De esta manera, por ejemplo, tendremos que para distintos valores de m , se obtienen las expresiones [5.4.35] y [5.4.36].

$$C^1(l) = P[|x_r - x_s| < l] \quad [5.4.35]$$

$$C^2(l) = P[|\bar{x}_r - \bar{x}_s| < l] = P[(|x_r - x_s| < l) \cap (|x_{r+1} - x_{s+1}| < l)] \quad [5.4.36]$$

Si los x_t son variables aleatorias independientes, podemos dar un paso más. La probabilidad de que se den conjuntamente los dos sucesos será igual a la probabilidad de que se dé uno de ellos por la probabilidad de que se dé el otro [5.4.37].

$$C^2(l) = P\left[\left(|x_r - x_s| < l\right) \cap \left(|x_{r+1} - x_{s+1}| < l\right)\right] = P\left(|x_r - x_s| < l\right) \cdot P\left(|x_{r+1} - x_{s+1}| < l\right) \quad [5.4.37]$$

Si todos los valores de la serie están idénticamente distribuidos, ambas probabilidades serán idénticas [5.4.38].

$$P\left(|x_r - x_s| < l\right) = P\left(|x_{r+1} - x_{s+1}| < l\right) \quad [5.4.38]$$

Con lo que la integral de correlación con dimensión de inmersión dos será igual al cuadrado de la integral de correlación con dimensión de inmersión uno. O bien, generalizando para cualquier dimensión de inmersión, obtenemos la expresión [5.4.39].

$$C^m(l) = \left(C^1(l)\right)^m \quad [5.4.39]$$

La igualdad [5.4.39] se cumplirá siempre que la serie sea IID. Brock, Dechert y Scheinkman demuestran que bajo la hipótesis de IID, la distribución de probabilidad del estadístico [5.4.40] converge a una normal de media cero y desviación típica $\mathbf{s}_m(l)$, cuando el tamaño de la serie es suficientemente grande. A esta distribución se llega sin necesidad de añadir ningún tipo de hipótesis sobre el tipo de ruido que estemos considerando. Es un test no paramétrico, que va a servirnos frente a cualquier tipo de dependencia o diferencias en la distribución.

$$\sqrt{N} \left\{ C^m(l) - C^1(l)^m \right\} \xrightarrow{d} N \left[0, \mathbf{s}_m(l) \right] \quad [5.4.40]$$

La varianza del estadístico, $\mathbf{s}_m^2(l)$, puede ser estimada a partir de [5.4.41].

$$\sigma_m^2(l) = 4 \left[K(l)^m + 2 \sum_{j=1}^{m-1} K(l)^{m-j} C^1(l)^{2j} + (m-1)^2 C^1(l)^{2m} - m^2 K(l) C^1(l)^{2m-2} \right] \quad [5.4.41]$$

La expresión de la varianza [5.4.41] va a depender de $C^1(l)$ que definíamos en [5.4.35] y que se calcula como una integral de correlación más, y de $K(l)$ que es la probabilidad de que tres puntos se encuentren a menos de una distancia l [5.4.42]. Esta probabilidad puede estimarse a partir de la expresión [5.4.43].

$$K(l) = P[|x_r - x_s| < l) \cap (|x_r - x_t| < l) \cap (|x_t - x_s| < l)] \quad [5.4.42]$$

$$\hat{K}(l) = \frac{6}{N(N-1)(N-2)} \sum_{1 \leq i < j < k \leq N} \frac{1}{3} \left[\begin{aligned} &\theta(l - |x_i - x_j|) \theta(l - |x_j - x_k|) + \\ &+ \theta(l - |x_i - x_k|) \theta(l - |x_j - x_k|) + \\ &+ \theta(l - |x_i - x_k|) \theta(l - |x_i - x_j|) \end{aligned} \right] \quad [5.4.43]$$

Este estadístico nos permite elaborar el contraste BDS⁵³, en el que adoptamos como hipótesis nula que x_t es una variable aleatoria independiente e idénticamente distribuida. Aceptaremos la hipótesis nula de IID si el valor que obtengamos en el estadístico está dentro de los límites que se aceptan en una Normal $N[0, \mathbf{s}_m(l)]$ para el nivel de significación que fijemos.

Dada la naturaleza no paramétrica del contraste, la información que nos proporciona el rechazo de la hipótesis nula no es muy amplia. Dado que estamos

⁵³ Brock, Dechert y Scheinkman, 1987. Kanzler, 1999, hace un estudio en profundidad del estadístico.

aplicando el contraste a los residuos de un filtro lineal, el rechazo de la hipótesis de IID va a venir de la presencia de un comportamiento no lineal en la serie.

Para el cálculo de $C^1(l)$ y $C^m(l)$, aprovechamos los resultados obtenidos en el apartado anterior, cuando estudiábamos la dimensión de correlación de las series. La estimación de $K(l)$ siguiendo la fórmula de [5.4.43]⁵⁴ resulta inviable dado que requiere la consideración de T^3 pares de puntos. Como alternativa se propone un estadístico⁵⁵ que sí resulta posible utilizar pues aprovecha los cálculos hechos para la estimación de $C^1(l)$, que aparece en [5.4.44].

$$\hat{K}(l) = \frac{1}{N^3} \sum_{i=1}^N \left(\sum_{j=1}^N \left\{ -|\bar{x}_i - \bar{x}_j| \right\} \right)^2 \quad [5.4.44]$$

Los valores de l para los que se ha calculado el estadístico BDS han sido 2; 1 y 0,5 veces la desviación típica de la serie que estamos estudiando, tal y como proponen Brock, Dechert y Scheinkman en el artículo original. Los resultados obtenidos aparecen en las Tablas 5.4.13 y siguientes, en las que se ha resaltado en negrita los casos en los que se aceptaría la hipótesis nula de IID. Esta se produce en muy pocos casos, concentrados en valores pequeños de m , lo que podría venir explicado, por ser una dimensión de inmersión insuficiente para reproducir las características del atractor que estemos considerando.

⁵⁴ Brock, Dechert y Scheinkman, 1987.

⁵⁵ Dechert, 1994.

Tabla 5.4.13: Estadístico BDS de las series de rendimientos con frecuencia de operación a operación

	marzo			junio			septiembre			diciembre		
	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$
m=2	-33,02	-6,19	-6,19	-50,35	-13,68	-13,68	21,95	-11,00	-11,00	-830,13	0,10	-5,36
m=3	-7,69	-2,15	-2,15	-4,47	-2,36	-2,36	2,36	-2,74	-2,74	-8,31	2,25	1,03
m=4	-87,66	-6,45	-6,45	-30,92	-1,94	-1,94	0,33	-6,07	-6,07	-367,28	12,09	10,60
m=5	-78,40	-4,02	-4,02	-27,82	-0,04	-0,04	8,93	-0,75	-0,75	-627,66	7,88	9,18
m=6	-64,73	-3,18	-3,18	-36,47	-0,01	-0,01	11,14	-1,19	-1,19	-603,26	8,34	9,85
m=7	-54,86	-1,60	-1,60	-34,94	1,04	1,04	10,08	0,30	0,30	-566,53	8,98	11,37
m=8	-43,95	-1,65	-1,65	-33,17	2,09	2,09	6,53	1,68	1,68	-601,91	8,82	11,86
m=9	-41,83	-1,57	-1,57	-27,67	2,07	2,07	8,43	2,22	2,22	-574,99	6,73	10,58
m=10	-36,44	-1,25	-1,25	-22,48	1,97	1,97	6,67	2,99	2,99	-528,64	5,74	10,33

Tabla 5.4.14: Estadístico BDS de las series de rendimientos con frecuencia de cinco minutos

	marzo			junio			septiembre			diciembre		
	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$
m=2	21,45	14,88	10,74	1,06	1,72	1,73	42,22	33,96	33,27	34,18	26,34	20,86
m=3	4,13	4,29	2,81	1,65	3,73	3,74	12,06	9,83	8,38	11,33	7,88	5,83
m=4	16,11	17,34	17,31	1,97	4,47	4,47	9,72	33,94	38,47	30,38	26,15	24,27
m=5	14,58	19,03	23,35	2,50	6,47	6,48	16,45	39,47	47,62	30,57	27,19	26,05
m=6	14,52	19,22	25,30	2,80	7,78	7,78	20,37	44,01	58,66	25,59	25,23	26,12
m=7	13,03	19,80	30,73	3,22	10,05	10,06	21,05	48,49	70,68	26,07	25,08	25,99
m=8	12,71	19,33	32,05	3,46	11,83	11,84	22,22	53,52	85,00	25,35	24,94	26,41
m=9	11,53	19,42	40,32	3,80	14,33	14,34	23,19	59,61	105,1	24,89	24,86	27,42
m=10	11,15	18,76	47,02	3,93	16,43	16,45	23,28	66,20	129,2	24,83	25,33	28,74

Tabla 5.4.15: Estadístico BDS de las series de residuos ARIMA de los rendimientos con frecuencia de operación a operación

	marzo			junio			septiembre			diciembre		
	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$
m=2	-37,28	-6,89	-6,59	-20,69	3,29	14,63	11,03	22,93	28,02	68,42	58,25	31,15
m=3	-1,28	0,20	1,33	2,10	1,24	3,69	3,87	8,74	10,15	15,64	13,34	7,71
m=4	-2,99	-0,93	1,72	23,00	10,09	17,47	38,81	32,16	37,00	57,77	54,47	24,17
m=5	0,01	-0,70	3,07	18,62	10,43	18,00	51,60	32,06	36,80	57,51	58,53	23,89
m=6	7,18	5,16	8,47	14,87	13,88	22,29	57,13	34,34	39,32	62,79	59,75	24,65
m=7	8,92	2,58	6,08	23,65	14,92	22,93	62,83	34,06	39,16	69,71	62,92	21,95
m=8	11,45	4,73	7,92	25,52	14,75	22,18	57,81	29,84	35,16	74,40	63,60	23,47
m=9	9,54	3,14	6,08	28,01	15,75	24,61	61,60	29,73	35,02	75,49	68,17	27,51
m=10	13,05	3,58	6,48	32,10	18,06	27,06	62,09	29,86	35,06	78,13	69,31	26,90

Tabla 5.4.16: Estadístico BDS de las series de residuos ARIMA de los rendimientos con frecuencia de cinco minutos

	marzo			junio			septiembre			diciembre		
	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$
m=2	6,59	10,87	12,87	145,4	242,0	533,8	16,31	25,16	26,12	15,69	16,43	17,51
m=3	1,90	2,74	2,45	25,55	36,79	56,41	5,29	7,28	6,54	2,99	1,92	1,49
m=4	9,06	16,52	20,14	83,72	218,1	1056	20,99	35,77	40,96	21,35	28,11	56,65
m=5	9,55	18,39	23,925	74,07	247,7	1944	22,08	39,73	48,67	23,43	42,50	178,7
m=6	9,49	19,94	28,02	67,72	295,5	3924	22,89	44,10	58,00	25,70	75,91	765,1
m=7	9,19	20,96	32,12	63,22	363,1	8362	23,39	48,85	69,78	27,86	165,7	3836
m=8	8,78	21,56	36,15	59,90	455,3	18382	23,81	54,10	84,87	30,47	437,7	20719
m=9	8,24	21,82	41,71	57,30	577,4	42427	24,19	60,13	105,2	33,79	1304	116498
m=10	7,69	21,73	49,32	55,25	737,5	102246	24,48	66,94	132,0	38,17	4262	673701

Tabla 5.4.17: Estadístico BDS de las series de residuos GARCH-M de los rendimientos con frecuencia de operación a operación

	marzo			junio			septiembre			diciembre		
	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$
m=2	-45,09	-56,95	-61,07	-50,01	-13,87	-9,01	8,79	-11,70	-19,10	-1343	-10,20	1,28
m=3	-44,51	-7,76	-6,96	-3,87	-2,40	-1,32	21,13	0,99	-6,20	-8,47	-0,13	-0,36
m=4	-37,81	-54,35	-45,28	-28,87	-2,21	0,85	13,12	2,45	-4,31	-6800	-104,6	-58,53
m=5	-32,19	-37,62	-34,68	-21,89	-0,38	2,85	10,47	1,75	-3,84	-1574	-3,20	-7,63
m=6	-29,54	-30,42	-28,32	-32,70	-0,47	2,92	9,94	3,21	-4,87	-1550	-2,16	-5,97
m=7	-26,47	-20,00	-20,57	-31,83	0,66	3,54	6,10	3,21	-3,69	-1549	-0,90	-5,43
m=8	-24,51	-17,97	-19,57	-27,40	1,75	4,70	4,12	3,14	-2,97	-1427	-0,46	-5,37
m=9	-22,84	-15,59	-16,92	-23,92	1,71	4,54	3,85	3,88	-2,48	-1439	-2,33	-5,70
m=10	-22,85	-20,75	-15,72	-20,48	1,67	4,44	3,05	3,38	-2,82	-1315	-1,90	-5,44

Tabla 5.4.18: Estadístico BDS de las series de residuos GARCH de los rendimientos con frecuencia de cinco minutos

	marzo			junio			septiembre			diciembre(*)		
	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$
m=2	15,18	7,29	4,70	-8,03	5,72	6,90	17,67	-0,10	-2,70	33,42	6,10	3,82
m=3	3,53	1,19	0,39	0,57	1,43	0,74	3,11	0,58	0,04	6,56	1,26	0,62
m=4	11,41	4,82	2,89	3,05	6,88	8,37	15,14	5,26	0,79	-20,47	-2,50	3,16
m=5	9,77	2,56	1,36	0,13	6,56	11,58	11,45	5,29	0,80	-12,87	-1,19	5,80
m=6	7,47	2,57	2,94	-5,55	6,56	19,34	8,51	6,71	2,87	-8,68	-1,59	7,67
m=7	8,19	4,07	7,41	-5,92	8,06	39,75	9,37	7,14	4,79	-8,22	-1,52	12,18
m=8	8,47	4,84	13,54	-4,74	10,38	83,82	8,69	7,88	12,76	-5,54	0,01	18,21
m=9	8,01	4,82	30,88	-3,95	12,54	230,1	9,24	8,82	32,58	-4,22	0,24	28,27
m=10	8,01	5,10	64,37	-9,70	14,25	636,9	7,94	8,00	84,23	-2,24	1,66	46,92

(*)En el vencimiento de diciembre el modelo estimado era un GARCH-M.

Tabla 5.4.19: Estadístico BDS de las series de residuos EGARCH de los rendimientos con frecuencia de operación a operación

	marzo			junio			septiembre			diciembre		
	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$
m=2	-33,53	-18,98	-9,72	-13,22	-13,06	-1,17	-26,42	8,74	31,27	-1889	-1716	-1594
m=3	-10,72	-8,35	-6,04	-0,66	-2,20	-0,12	-11,85	-0,85	4,57	-552,0	-363,6	-187,9
m=4	-47,01	-31,57	-19,82	0,14	-1,52	3,88	-37,43	-5,83	18,05	-843,1	-692,4	-569,8
m=5	-47,47	-32,59	-21,62	0,69	0,36	6,28	-32,66	-5,60	18,86	-670,4	-526,9	-413,4
m=6	-45,50	-32,15	-21,51	-7,34	0,29	6,37	-25,23	-2,45	18,04	-557,7	-420,7	-316,8
m=7	-36,94	-26,95	-17,62	-6,17	1,31	6,61	-19,98	0,17	22,96	-477,5	-347,0	-252,3
m=8	-33,72	-24,71	-16,15	-5,17	2,38	7,21	-20,59	1,25	25,75	-417,7	-292,9	-206,5
m=9	-32,75	-23,42	-15,56	-3,63	2,34	7,55	-20,55	0,84	26,04	-370,9	-251,7	-172,8
m=10	-30,84	-21,52	-14,27	-1,45	2,21	7,93	-18,71	0,72	28,09	-333,3	-219,2	-147,0

Tabla 5.4.20: Estadístico BDS de las series de residuos EGARCH de los rendimientos con frecuencia de cinco minutos

	marzo			junio			septiembre			diciembre		
	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$
m=2	20,65	14,45	10,94	-2,86	-2,89	-1,90	3,33	-1,55	-0,86	8,13	2,09	0,26
m=3	4,22	1,91	1,40	0,73	-0,18	-0,21	2,53	0,31	0,12	0,97	0,44	0,37
m=4	15,82	11,91	12,79	2,63	1,21	2,74	8,88	1,16	0,10	5,65	3,87	19,08
m=5	14,24	11,23	13,13	1,54	1,27	7,88	10,25	1,37	0,84	9,16	5,28	66,41
m=6	11,39	10,87	15,40	1,16	2,32	13,85	7,93	2,33	3,30	7,79	5,85	240,2
m=7	12,07	12,23	18,15	0,25	2,38	27,19	6,53	2,39	8,79	7,87	5,66	1032
m=8	11,86	12,28	19,73	1,17	2,70	66,72	7,20	3,19	22,49	7,32	6,60	4646
m=9	11,58	13,24	22,83	1,67	4,05	179,1	6,96	3,37	57,06	7,64	8,47	20278
m=10	-70,79	-17,76	-8,86	1,41	4,22	511,7	7,18	3,18	145,7	8,56	10,39	95691

Tabla 5.4.21: Estadístico BDS de las series de duraciones

	marzo			junio			septiembre			diciembre		
	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$
m=2	21,53	8,33	-2,50	-141,4	-68,54	-4,92	36,46	-4,49	4,65	-2,35	14,07	17,93
m=3	3,21	3,61	0,64	-18,01	-12,09	0,61	3,23	-0,99	2,49	5,13	5,47	6,90
m=4	16,27	15,49	8,65	-57,56	-23,38	6,87	10,13	6,69	7,68	16,57	15,54	21,00
m=5	14,71	16,81	11,24	-33,57	-10,19	13,16	10,46	10,36	8,90	17,98	16,71	21,72
m=6	9,37	17,07	13,56	-30,86	-9,35	12,21	8,89	10,93	9,79	17,45	19,14	22,87
m=7	8,84	16,43	13,50	-22,57	-9,39	9,32	10,31	12,50	12,52	15,73	20,38	26,05
m=8	8,29	16,58	14,78	-16,72	-4,22	8,61	11,50	12,65	12,58	12,24	20,75	26,87
m=9	7,44	17,52	17,06	-10,59	-1,19	9,90	14,05	13,92	15,42	12,78	21,41	28,82
m=10	6,28	16,88	18,33	-12,79	-0,23	10,78	15,91	15,47	13,58	14,00	21,98	28,89

Tabla 5.4.22: Estadístico BDS de las series de duraciones diferenciadas fraccionalmente

	marzo			junio			septiembre			diciembre		
	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$
m=2	18,02	15,49	20,35	-86,55	-63,58	-15,12	31,18	34,52	33,47	9,52	18,94	29,77
m=3	3,35	5,59	6,30	-9,19	-4,70	1,50	3,63	10,16	10,98	5,70	7,85	9,24
m=4	17,59	21,72	32,53	-30,55	-4,15	17,57	9,21	22,77	30,82	17,43	22,54	29,22
m=5	19,81	19,71	33,12	-33,13	0,03	16,34	13,91	25,47	34,70	9,47	20,45	31,32
m=6	19,82	18,04	37,49	-13,39	7,99	20,33	18,63	28,01	35,78	12,27	21,44	32,22
m=7	16,16	18,84	41,68	-11,28	12,24	21,45	20,25	29,20	42,47	13,15	23,14	36,62
m=8	16,12	20,21	47,67	-7,05	14,77	22,56	19,50	29,28	44,33	13,12	23,05	36,94
m=9	14,70	23,09	59,32	-1,80	17,59	23,63	19,29	30,77	51,03	13,32	24,16	39,92
m=10	15,15	23,18	66,41	0,15	17,31	24,13	20,28	32,53	55,37	13,39	25,00	41,26

Tabla 5.4.23: Estadístico BDS de las series de residuos ARFIMA de las duraciones

	marzo			junio			septiembre			diciembre		
	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$	$l=2*s$	$l=s$	$l=s/2$
m=2	7,33	13,13	14,29	-27,91	14,02	27,53	28,07	28,33	32,95	25,06	19,92	12,51
m=3	3,17	5,53	4,77	-3,27	5,46	7,81	6,33	5,45	7,11	2,04	6,05	4,74
m=4	14,02	19,65	19,95	-12,35	11,64	19,04	12,27	13,48	26,46	11,80	14,10	11,76
m=5	4,94	17,50	21,91	-7,60	13,49	19,84	9,19	18,91	27,65	9,10	13,62	10,41
m=6	6,36	18,63	22,85	-2,21	14,96	20,69	9,66	20,31	32,22	11,54	16,77	13,57
m=7	7,20	20,07	26,22	-1,15	12,05	18,17	10,20	18,78	32,10	11,21	17,09	16,14
m=8	8,02	20,01	27,24	-1,05	12,31	19,06	10,09	19,34	34,62	10,24	16,12	15,63
m=9	9,90	20,66	27,74	-1,12	13,84	21,28	9,22	19,29	37,45	10,17	16,53	16,57
m=10	10,87	22,06	30,22	-0,77	14,00	22,54	11,01	20,13	40,44	10,79	16,74	17,43

5.5. CONCLUSIONES.

Las conclusiones que sacamos de este capítulo son las siguientes:

1. Las gráficas de x_t frente a $x_{t-\tau}$ de las series originales de rendimientos presentan estructura de rejilla, debido al tamaño mínimo del cambio del precio. De igual forma, las series de duraciones presentan esa misma estructura, debido a que el

máximo de la precisión es el segundo. El 90% de los datos de las series de rendimientos con frecuencia operación a operación se acumulan entorno al valor cero y a un “tick” de distancia de este.

2. Las gráficas de x_t frente a $x_{t-\tau}$ de los residuos de los distintos modelos ya no presentan estructura de rejilla que podamos imputar a la precisión de la observación, sino estructuras más complejas como las de los modelos ARMA(1,1) de las series de rendimientos con frecuencia operación a operación que tienen la misma forma a distintas escalas, esto es, son fractales. Esto indica que la modelización previa de las series de rendimientos no sólo es necesario para eliminar las dependencias lineales que existan en las series sino que permite encontrar estructuras no lineales ocultas por la existencia de un tamaño mínimo de los cambios de precios.
3. Las gráficas de recurrencias no muestran de forma generalizada órbitas periódicas inestables. Esto se debe a la gran cantidad de valores iguales o alrededor de cero. Tan solo en las series de residuos ARIMA sobre las series de operaciones esto se hace presente, por lo que sólo en este caso habría un indicio claro de la presencia de caos.
4. Los coeficientes de Hurst de las series de rendimientos y sus residuos no muestran la presencia de persistencia (memoria a largo plazo) que justifique el uso de modelos ARFIMA.

5. Dados los indicios de no linealidades mostrados por las gráficas tanto de x_t frente a $x_{t-\tau}$ se ha procedido a la reconstrucción del sistema para la posterior aplicación de análisis de tipo cuantitativo. Para la reconstrucción del sistema se ha utilizado el método de la información mutua que permite obtener el retardo óptimo, t , para el análisis posterior.
6. Se ha comprobado como el uso de un t excesivo lleva a la negación de la existencia de caos aun siendo este el generador de la serie. Este resultado es coherente con los resultados obtenidos por análisis de caos anteriores que usaban la frecuencia diaria. Se hace, por tanto, necesario partir de una serie con la más alta frecuencia posible. Esto justifica el uso de las series de operación a operación y de cinco minutos.
7. El estudio de la dimensión de correlación muestra la presencia de una región lineal en la que es posible la estimación de la dimensión de correlación en todas las series.
8. El cálculo de la dimensión de correlación, tanto en las series de rendimientos con frecuencia operación a operación como en las de duración de operaciones, nos llevan a estimaciones muy bajas del valor de la dimensión de correlación. Estos valores son lo suficientemente bajos como para ser considerados como provenientes de sistemas no lineales. Estos valores son superiores en las series de cinco minutos, aunque en cualquier caso inferiores a la dimensión de inmersión. Siguiendo la literatura existente, los resultados de las series de cinco minutos no son concluyentes, pero los obtenidos en las series de operación a operación sí se pueden considerar como tales.

9. Se han encontrado exponentes de Lyapunov positivos en la mayoría de las series analizadas, lo que demuestra la presencia de inestabilidades propias de sistemas caóticos. Tan solo en el caso de las series originales de los rendimientos de operación a operación no llegamos a encontrar exponentes positivos. De todas formas, la no posibilidad de calcular intervalos de confianza en ninguno de los casos nos impide asegurar si efectivamente son o no positivos.
10. La entropía de Kolmogorov se mantiene acotada y no tiende infinito, situación propia de sistemas caóticos y no de sistemas aleatorios. Si bien estos resultados no deben ser tomados en cuenta pues la cota se alcanza para unos valores de l que no dejan suficientes puntos disponibles para su adecuada estimación.
11. El contraste BDS es el único, de los utilizados, que nos permite hacer una contrastación de hipótesis. Los resultados, en todas las series analizadas nos llevan a aceptar la existencia de no linealidades en la serie.
12. Los resultados conjuntos muestran presencias de no linealidades en todas las series. Estas son, en general, más claras en las series de operación a operación que en las de cinco minutos.