UN CONTRASTE DE EFECTOS GARCH BASADO EN LA FUNCIÓN DE COVARIACIÓN MUESTRAL ROBUSTO A INNOVACIONES CON COLAS PESADAS

Afonso Rodríguez, Julio Angel
UNIVERSIDAD DE LA LAGUNA
Departamento de Economía de las Instituciones, Estadística Económica y Econometría
Instituto Universitario de Desarrollo Regional (IUDR)

Facultad de CC. Económicas y Empresariales
Camino La Hornera, s/n. Campus de Guajara. 38202 LA LAGUNA
Tenerife. Islas Canarias

Tfno.: 922.317041, Fax: 922.317042
email: jafonsor@ull.es

Resumen

Existe un creciente interés en el análisis de series temporales de alta frecuencia empleando modelos con distribuciones de colas pesadas, con objeto de describir distintos tipos de comportamientos erráticos y recoger la influencia de observaciones anómalas u outliers frecuentemente observados en la práctica. Por aplicación de los resultados debidos a Kesten (1973) sobre la existencia de soluciones estacionarias y las características de las distribuciones de dimensión finita de ecuaciones de recurrencia estocásticas como distribuciones de variación regular con índice $\kappa > 0$ (distribuciones de colas pesadas), se han podido establecer las principales características de las distribuciones marginales de los procesos no lineales más habitualmente analizados, como los procesos GARCH (Basrak, Davis y Mikosch (2002)), Bilineales (Resnick y van den Berg (2000)) y de raíz unitaria estocástica (Yoon (2003)). En todos los casos se trata de distribuciones marginales con colas pesadas tipo Pareto, debido al mecanismo no lineal que en cada caso conecta el output con el input, aún cuando la distribución de las innovaciones de tales procesos sea de colas estrechas. Con el objetivo de detectar estos tipos de no linealidad, la literatura econométrica propone analizar la posible existencia de estructuras de dependencia no sólo en la serie observada sino también en potencias del valor absoluto mediante las funciones de autocovarianza (FACVM) y de autocorrelación muestral (FACM). En los casos citados anteriormente, las autocovarianzas y autocorrelaciones muestrales de los cuadrados convergen a límites no degenerados cuando $\kappa < 4$, de forma que resulta cuestionable la identificación de efectos no lineales tipo GARCH frente a otros tipos de no linealidad aditiva empleando estas herramientas. En este trabajo analizamos, en primer lugar, la utilización de la denominada función de covariación muestral (Cambanis y Miller (1981), Davis (1983) y Gallagher (2002)), diseñada para identificar relaciones de dependencia lineal entre procesos SαS, para identificar efectos GARCH frente a dependencia lineal o no lineal de tipo aditivo en el caso de distribuciones marginales de variación regular con índice $\kappa < 4$. En segundo lugar, proponemos una versión generalizada de segundo orden para construir un test tipo Portmanteau para detectar dependencia no lineal multiplicativa tipo GARCH.

Palabras Clave: *Función de covariación, distribuciones de variación regular, distribuciones con colas pesadas, procesos GARCH, procesos Bilineales, procesos STUR, Rendimientos de tipos de cambio*

Clasificación Código JEL: *C22, C15, C52, F31*
ÁREA TEMÁTICA: *MERCADOS MONETARIOS Y FINANCIEROS*
MÉTODOS CUANTITATIVOS EN ECONOMÍA

1. Introducción

En este trabajo estudiamos la posibilidad de utilizar una medida de dependencia alternativa a la habitual utilización de las funciones de autocovarianzas (FACVM) y autocorrelación (FACM) muestral, denominada **función de covariación muestral**, para discriminar entre procesos lineales o no lineales de tipo aditivo frente a procesos no lineales de tipo multiplicativo (es decir, procesos tipo GARCH puros o de volatilidad estocástica). En el apartado 2 se revisan los principales resultados obtenidos recientemente sobre la caracterización de las propiedades límite y de la distribución asintótica de FACVM y FACM en el caso de procesos no lineales que admitan representaciones estacionarias con colas regularmente variables de las correspondientes distribuciones marginales de dimensión finita, es decir, con colas anchas tipo Pareto con parámetro de índice de cola $0 < \kappa < 4$. Estos incluyen los procesos GARCH, Bilineales, y de raíz unitaria estocástica (STUR), utilizados habitualmente para representar muchas de las características frecuentemente observadas en series temporales de alta frecuencia. Estos resultados hacen cuestionable la utilización de los resultados estándar de inferencia para determinar la significación de los coeficientes de autocovarianza o de autocorrelación. En el apartado 3 se revisa el concepto de función de covariación, como medida de dependencia especialmente diseñada para su uso en procesos de orden $p < 2$, es decir, con varianza infinita, y en procesos de variables aleatorias con distribuciones estables simétricas con índice $0 < \alpha < 2$ ($S\alpha S$) (Cambanis y Miller, 1981). Gallagher (2002, 2006) utiliza la versión muestral de esta función de covariación para diseñar contrastes de dependencia lineal en procesos lineales ARMA con errores con distribuciones de colas anchas y posiblemente varianza infinita. En este apartado 3 estudiamos la relevancia de dicha medida para detectar este tipo de dependencia en procesos no lineales. En el apartado 4 proponemos una generalización de la función de covariación bajo el supuesto de existencia del momento de orden $m \geq 2$ para construir dos estadísticos tipo Portmanteau que explotan las diferencias en el tipo de dependencia que determinan los procesos no lineales de tipo aditivo frente a los de tipo multiplicativo y poder así discriminar entre ambos. La ventaja de la utilización de estos estadísticos frente a los habituales basados en la FACM es que su distribución parece ser robusta a distribuciones de colas anchas e incluso a la violación del supuesto de existencia de la varianza del proceso. En el apartado 5 presentamos una selección de los principales resultados de un experimento de simulación y los resultados de la aplicación de los estadísticos de contraste analizados a tres series de rendimientos de tipos de cambio. Finalmente el apartado 6 presenta las principales conclusiones.

2. Identificación de no linealidad basada en la función de autocovarianza y de autocorrelación muestral en procesos con colas anchas

En este apartado revisamos los resultados existentes sobre el comportamiento de la función de autocovarianza y de autocorrelación en procesos lineales y no lineales bajo distribución de colas anchas representada por una distribución de variación regular con exponente característico o parámetro de índice de cola $0 < \alpha < 2$, que presenta comportamiento tipo Pareto de la forma $P(X > x) \sim c_\alpha \cdot x^{-\alpha}$, $x \to \infty$.

2.1 Caso de procesos lineales con distribución de colas anchas

En el caso de procesos lineales que admiten una representación MA(∞), $X_t = \sum_{j=0}^{\infty} c_j \varepsilon_{t-j}$, con coeficientes reales c_j que verifican la condición de sumabilidad $\sum_{j=0}^{\infty} c_j^2 < 1$ e innovaciones ε_t *iid* con distribución simétrica α-estable ($S\alpha S$) con $0<\alpha<2$, Davis y Resnick (1985a,b, 1986), prueban que la función de autocorrelación muestral del proceso X_t, $\tilde{\rho}_T(h) = \sum_{t=1}^{T-h} X_t X_{t+h} / \sum_{t=1}^{T} X_t^2$, $h = 1, 2, \ldots$ es un estimador consistente de $\rho(h) = \sum_{j=0}^{\infty} c_j c_{j+h} / \sum_{j=0}^{\infty} c_j^2$ que es la función de autocorrelación del proceso X_t bajo el supuesto $\varepsilon_t \sim iidN(0,\sigma^2)$ y, además,

$$\left(T / \ln(T)\right)^{1/\alpha} (\tilde{\rho}_T(1) - \rho(1), \ldots, \tilde{\rho}_T(h) - \rho(h)) \Rightarrow (Y_1(\alpha), \ldots, Y_h(\alpha))$$

con $Y_k(\alpha) = \sum_{j=1}^{\infty} (\rho(k+j) + \rho(k-j) - 2\rho(j)\rho(k)) S_j(\alpha)/S_0(\alpha)$, $S_1(\alpha), S_2(\alpha)$, … variables aleatorias estables independientes con índice α y $S_0(\alpha)$ estable positiva con índice $\alpha/2$. A partir de estos resultados, Runde (1997) estudia el comportamiento de los contrastes tipo Portmanteau de Box y Pierce (1970), $\hat{Q}_{BP,T}(h)$, y Ljung y Box (1978), $\hat{Q}_{LB,T}(h)$,

$$\hat{Q}_{BP,T}(h) = T \sum_{k=1}^{h} \tilde{\rho}_T^2(k), \quad \hat{Q}_{LB,T}(h) = T(T+2) \sum_{k=1}^{h} \frac{\tilde{\rho}_T^2(k)}{T-k}$$

en el caso de varianza infinita del proceso. A partir de los resultados anteriores de Davis y Resnick (1986), el estadístico apropiado viene dado por

$$\left(\frac{T}{\ln(T)}\right)^{2/\alpha} \sum_{k=1}^{h} \tilde{\rho}_T^2(k) \Rightarrow Q_h(\alpha), \quad Q_h(\alpha) \overset{d}{=} S^{-2}(\alpha/2, 1) \sum_{k=1}^{h} S_k^2(\alpha, 0)$$

donde $S(\alpha,\beta)$ denota una variable aleatoria estable con parámetro de localización 0, parámetro de escala 1, exponente característico α y parámetro de asimetría $|\beta|\leq 1$. Sin embargo, el estadístico anterior presenta un pobre comportamiento en muestras finitas, especialmente para α próximo a 2, posiblemente debido a la aparente lenta convergencia a la distribución asintótica.

2.2 Caso de procesos no lineales con distribución de colas anchas

En este epígrafe revisamos los resultados existentes para el caso de los procesos no lineales de series temporales más habitualmente estudiados y utilizados en la práctica, especialmente en el contexto de series temporales financieras de alta frecuencia.

A. Procesos GARCH

Un proceso GARCH fuerte estrictamente estacionario puede representarse de forma general como

$$X_t = \varepsilon_t \upsilon_t(\theta) \tag{2.1}$$

donde ε_t es una secuencia de innovaciones *iid* con media nula, varianza unitaria e independiente de X_{t-i}, $i \geq 1$, y $\upsilon_t^2(\theta)$ es la ecuación de varianza condicional definida de forma general como

$$\upsilon_t^2(\theta) = c_0(\theta) + \sum_{i=1}^{\infty} c_i(\theta) X_{t-i}^2 \tag{2.2}$$

donde la secuencia de coeficientes reales positivos $c_i(\theta)$ verifican la condición de sumabilidad $\sum_{i=1}^{\infty} c_i^2(\theta) < 1$. Estableciendo una estructura particular para estos coeficientes en función de los parámetros estructurales del proceso θ se obtienen los casos habituales GARCH(p,q) (Bollerslev, 1986)

$$\upsilon_t^2(\theta) = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \upsilon_{t-j}^2(\theta) \tag{2.3}$$

y ARCH(p) (Engle, 1982) haciendo $\beta_j = 0$, $j = 1,\ldots,q$. Los procesos GARCH admiten una representación de ecuación de recurrencia estocástica $\boldsymbol{Y}_t = \boldsymbol{A}_t \cdot \boldsymbol{Y}_{t-1} + \boldsymbol{B}_t$, de forma que por aplicación de los resultados de Kesten (1973), Davis, Mikosch y Basrak (1999), Mikosch y Starica (2000) y Basrak, Davis y Mikosch (2002), obtienen las condiciones necesarias para la existencia de una única solución estacionaria causal de los procesos tipo GARCH y demuestran la propiedad de variación regular de las distribuciones marginales de dimensión finita de un proceso GARCH (ver también Davis y Mikosch (2007) y Lindner (2007)). En el caso GARCH(1,1),

$$\upsilon_t^2(\theta) = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \upsilon_{t-1}^2(\theta) = \alpha_0 + (\alpha_1 \varepsilon_{t-1}^2 + \beta_1)\upsilon_{t-1}^2(\theta) = \alpha_0 + A_t(\theta)\upsilon_{t-1}^2(\theta) \qquad (2.4)$$

el parámetro de índice de cola de la distribución estacionaria de colas anchas tipo Pareto (o exponente de momento máximo), $\kappa = \kappa(\theta,\nu)$ (donde ν es un parámetro adicional que depende de la distribución asumida para las innovaciones), se obtiene como solución única de la ecuación $1 = E[A_1^{\kappa}(\theta)] = E[(\alpha_1\varepsilon_0^2 + \beta_1)^{\kappa}]$. Los procesos GARCH son procesos de orden $m(\kappa)$, es decir, tienen momentos incondicionales finitos hasta el de orden m, que es el mayor entero estrictamente menor que 2κ. Un proceso IGARCH(1,1) (GARCH integrado), con $\alpha_1+\beta_1 = 1$ determina que $\kappa = 1$, de forma que tiene varianza incondicional infinita. El parámetro κ tomará un valor tanto más pequeño cuánto más pesadas sean las colas de la distribución de las innovaciones, para un valor dado de θ. La fuente de las colas anchas de la distribución de los procesos GARCH resulta del mecanismo no lineal que conecta el output con el input. Davis y Mikosch (1998, 2000) y Mikosch y Starica (2000) demuestran que bajo estacionariedad estricta la distribución asintótica de las primeros h coeficientes de la función de autocovarianzas y de autocorrelación muestral de la potencia i-ésima (i = 1, 2) de las observaciones de un proceso GARCH(p,q), $\hat{\rho}_{i,T}(k) = \hat{\gamma}_{i,T}(k)/\hat{\gamma}_{i,T}(0) = \sum_{t=1}^{T-k}\tilde{X}_t^i\tilde{X}_{t+k}^i / \sum_{t=1}^{T}\tilde{X}_t^i$, $\tilde{X}_t^i = X_t^i - \hat{\mu}_{i,T}$, $\hat{\mu}_{i,T} = (1/T)\sum_{t=1}^{T}X_t^i$, viene dada por

(a) Si $\kappa \in (0,2)$: $\quad (T^{1-2/\kappa}\hat{\gamma}_{2,T}(k))_{k=1,\ldots,h} \Rightarrow (V_k)_{k=1,\ldots,h}$, $(\hat{\rho}_{2,T}(k))_{k=1,\ldots,h} \Rightarrow (V_k / V_0)_{k=1,\ldots,h}$

(b) Si $\kappa \in (2,4)$: $\quad (T^{1-2/\kappa}(\hat{\gamma}_{2,T}(k) - \gamma_2(k)))_{k=1,\ldots,h} \Rightarrow (V_k)_{k=1,\ldots,h}$

$$(T^{1-2/\kappa}(\hat{\rho}_{1,T}(k) - \rho_1(k)))_{k=1,\ldots,h} \Rightarrow \gamma_1^{-1}(0)(V_k - \rho_1(k)V_0)_{k=1,\ldots,h}$$

donde en (a) y (b), (V_0, …, V_h) es conjuntamente $\kappa/2$ estable en $\mathbb{R}^{h+1}$, y

(c) Si $\kappa > 4$, se cumplen los resultados en (b) con factor de normalización $T^{1/2}$, donde $V_1,\ldots,V_h$ es normal multivariante con media nula y matriz de covarianzas $\boldsymbol{C}(h)$ y $V_0 = E(X_0^4)$. En el caso lineal, la tasa de convergencia, determinada por las constantes de normalización, es más rápida cuánto más pesadas sean las colas. En el caso no lineal, la tasa de convergencia de la función de autocorrelación a su contrapartida determinista es más lenta cuánto más pesadas son las colas de la distribución y si el proceso tiene varianza infinita, las autocorrelaciones muestrales tienen leyes límite no degeneradas, lo que dificulta en la práctica el uso de estos estadísticos para la identificación de efectos GARCH.

B. Procesos de volatilidad estocástica (SV)

Un proceso de volatilidad estocástica (SV) tiene, como los procesos GARCH, una representación como proceso de producto de ruidos, $X_t = \varepsilon_t \upsilon_t(\theta)$, donde ε_t es una secuencia *iid* de variables aleatorias completamente independiente de la secuencia estrictamente estacionaria de variables aleatorias no negativas, $\upsilon_t(\theta)$. La independencia de ambos procesos permite derivar de forma de forma relativamente fácil las propiedades probabilísticas básicas de X_t. La estructura de dependencia del proceso X_t se determina a través de la dependencia en el proceso de volatilidad $\upsilon_t(\theta)$. Davis y Mikosch (2001) consideran el caso general de que el logaritmo del proceso de volatilidad es un proceso lineal, $\ln \upsilon_t(\theta) = \sum_{j=-\infty}^{\infty} \varphi_j u_{t-j}$, donde u_t es una secuencia *iid* de variables aleatorias de media cero Gaussianas y $\sum_{j=-\infty}^{\infty} \varphi_j^2 < \infty$. Bajo el supuesto de que ε_t tiene distribución de variación regular con índice $0 < \alpha < 2$, demuestran que la constante de normalización de la autocorrelación muestral de orden k del proceso X_t es del mismo orden que en el caso lineal para establecer la convergencia en distribución al límite no degenerado (V_k/V_0). Davis y Mikosch (2001) discuten la posibilidad de extender estos resultados al caso de la distribución asintótica de la funciones de autocovarianza y de autocorrelación muestral del proceso $|X_t|^\delta$, $\delta > 0$, con el objeto de detectar no linealidad tipo SV.

C. Procesos Bilineales (BL)

La clase de procesos bilineales (BL) se especifica como una generalización de los modelos ARMA incorporando productos cruzados del proceso observado y de las innovaciones *iid* de la forma

$$X_t = \phi_0 + \sum_{i=1}^{p} \phi_i X_{t-i} + \varepsilon_t + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \sum_{i=1}^{r} \sum_{j=1}^{s} \phi_{ij} X_{t-i} \varepsilon_{t-j}$$

y se denota por BL(p,q,r,s). En el caso simple BL(1,0,1,1) se tiene que, condicional a la información disponible hasta el período $t-1$, el proceso bilineal presenta heterocedasticidad condicional autorregresiva tipo ARCH(1), de la forma

$$E_{t-1}(X_t^2) = \sigma_\varepsilon^2 + A_{t-1}^2(\phi) X_{t-1}^2$$

de forma que existe cierta conexión entre estos procesos y los procesos GARCH. Davis y Resnick (1996), Basrak, Davis y Mikosch (1999) y Resnick y van den Berg (2000) estudian las condiciones de estacionariedad estricta y las características de la distribución estacionaria de este tipo de procesos y encuentran que, por aplicación de los resultados de Kesten (1973), esta tiene colas regularmente variables con índice $\alpha > 0$. En el caso BL(1,0,1,1) éste se determina como solución de la ecuación $E|\phi_1+\phi_{11}\varepsilon_1|^u=1$. Estos resultados cuestionan de nuevo el uso de las herramientas estándar de selección de los órdenes del modelo y de estimación en el caso de aplicarse a procesos no lineales con distribuciones de colas anchas.

D. Procesos de raíz unitaria estocástica (STUR)

Un proceso autorregresivo de coeficientes aleatorios de primer orden, RCA(1), viene dado por

$X_t = \rho_t(\phi)X_{t-1} + \varepsilon_t,$

donde $\rho_t(\phi) = \phi + a_t$, con (ε_t, a_t) independientes, con distribución *iid* de media nula y varianzas σ_ε^2 y σ_a^2. Si σ_a^2=0, el proceso es AR(1) y si σ_a^2>0 y $\phi = 1$, se obtiene la clase de procesos de raíz unitaria estocástica (STUR). Un proceso AR(1) es débilmente estacionario si $\phi^2+\sigma_a^2 < 1$, de forma que un proceso STUR(1) no puede ser débilmente estacionario. Sin embargo, a partir de la condición de estacionariedad estricta, $E(ln|\phi+a_1|) < 0$ (Yoon (2006, 2007) y Cline (2007)), el proceso X_t es geométricamente ergódico y, de aquí, β-mixing (que implica α-mixing) con coeficientes mixing con decrecimiento geométrico, de forma que el proceso es de memoria corta. A pesar de que un proceso STUR(1) es no estacionario en covarianza y la diferenciación no conduce a la estacionariedad, se ha encontrado en diversas aplicaciones (Yoon, 2003) que es un modelos muy flexible y que permite replicar las características más relevantes de las series de rendimientos de activos financieros. Así, por ejemplo, la varianza condicional de ΔX_t viene dada por $Var_{t-h}(\Delta X_t) = (1+\sigma_a^2)^{h-1}(\sigma_a^2 X_{t-h}^2 + \sigma_\varepsilon^2)\, h \geq 1$. Cline (2007) demuestra que bajo estacionariedad estricta, un proceso STUR(1) tiene distribución marginal con colas regularmente variables tipo Pareto y de aquí de colas muy anchas, con parámetro de índice de cola $0 < \kappa < 1$, de forma que un proceso STUR no tiene momentos incondicionales finitos y es muy probable la aparición de valores extremos. Así, las funciones de autocorrelación muestral de un proceso STUR convergen a un límite no degenerado y su distribución asintótica será función de procesos $S\alpha S$ $(0<\alpha<1)$.

3. La función de covariación y la función de covariación muestral: definición y propiedades

Tradicionalmente, la dependencia lineal se cuantifica mediante la correlación. En lo que sigue se considerará una medida diferente de dependencia lineal que resulta más apropiada en el caso de que los datos procedan de una distribución con colas anchas, es decir, para procesos de orden p tal que $E(|X_t|^m) < \infty 1 \leq m \leq p$, y $E(|X_t|^m) = \infty m > p$, con $p \leq 2$. Se dice que el proceso estocástico $\{X_t\}$ estrictamente estacionario y ergódico con media nula presenta dependencia lineal en el retardo k si para algún entero k se verifica que

$$E(X_t \mid X_{t-k}) = \lambda(k) X_{t-k} \tag{3.1}$$

Cambanis y Miller (1981) demuestran que la condición (3.1) se satisface para cualquier par de variables aleatorias con distribución α-estable simétrica ($S\alpha S$). Por otro lado, cualquier secuencia *iid* de media nula satisface (3.1) con $\lambda(k) = 0$ para todo $k \geq 1$. Si X_t es un proceso AR(1) con innovaciones ε_t *iid* de media nula

$$X_t = \phi_0 + \phi_1 X_{t-1} + \varepsilon_t = \phi_1^k X_{t-k} + \sum_{i=0}^{k-1} \phi_1^i (\phi_0 + \varepsilon_{t-i}) \tag{3.2}$$

se tiene que

$$E(X_t \mid X_{t-k}) = F_k(\phi) + \phi_1^k X_{t-k}, \quad F_k(\phi) = \phi_0 \sum\nolimits_{i=0}^{k-1} \phi_1^i$$

de forma que si $\phi_0 = 0$ (el proceso tiene media nula), se satisface (3.1) con $\lambda(k) = \phi_1^k$. Si se tiene que X_t sigue un proceso ARMA(1,1) con innovaciones *iid* de media nula como

$$\begin{aligned} X_t &= \phi_0 + \phi_1 X_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \\ &= \phi_1^k X_{t-k} + \phi_0 \sum_{i=0}^{k-1} \phi_1^i + \varepsilon_t + (\phi_1 - \theta_1) \sum_{i=1}^{k-1} \phi_1^{i-1} \varepsilon_{t-i} - \phi_1^{k-1} \theta_1 \varepsilon_{t-k} \end{aligned} \tag{3.3}$$

se tiene que

$$E(X_t \mid X_{t-k}) = \phi_1^{k-1}(\phi_1 X_{t-k} - \theta_1 \varepsilon_{t-k}) + \phi_0 \sum\nolimits_{i=0}^{k-1} \phi_1^i = F_k(\phi) + \phi_1^k X_{t-k}$$

con $F_k(\phi) = \phi_0 \sum\nolimits_{i=0}^{k-1} \phi_1^i - \phi_1^{k-1} \theta_1 \varepsilon_{t-k}$, de forma que si $\phi_0 = 0$ (media nula), se verifica de forma aproximada la condición (3.1).

En el caso de procesos no lineales con **nolinealidad de tipo aditivo** (bilineales, autorregresivos no lineales, ...) se espera un comportamiento similar de la función de media condicional para el retardo k. Un proceso BL(1,0,1,1), con innovaciones ε_t *iid* de media nula y varianza σ_ε^2 viene dado por,

$$X_t = \phi_0 + \phi_1 X_{t-1} + \phi_{11} X_{t-1} \varepsilon_{t-1} + \varepsilon_t$$

Mediante sustituciones recursivas, puede escribirse como

$$X_t = \phi_0\left(1+\sum_{i=1}^{k-1}\prod_{j=1}^{i} A_{t-j}(\phi)\right) + X_{t-k}\prod_{j=1}^{k} A_{t-j}(\phi) + \varepsilon_t + \sum_{i=1}^{k-1}\varepsilon_{t-i}\prod_{j=1}^{i} A_{t-j}(\phi) \qquad (3.4)$$

donde $A_{t-j}(\phi) = \phi_1 + \phi_{11}\varepsilon_{t-j}$, con $\sum_{i=1}^{0} a_i = 0$ y $\prod_{i=1}^{0} a_i = 1$. Entonces,

$$\begin{aligned} E(X_t \mid X_{t-k}) &= A_{t-1}(\phi)X_{t-1} & k = 1 \\ &= F_k(\phi) + \phi_1^k A_{t-k}(\phi)X_{t-k} & k > 1 \end{aligned}$$

con $F_k(\phi) = \phi_0\sum_{i=0}^{k-1}\phi_1^i + \phi_{11}\sigma_\varepsilon^2\sum_{i=1}^{k-1}\phi_1^{i-1}$. En el caso BL(0,0,1,1) (ϕ_1 = 0) con ϕ_0 = 0, se tiene $E(X_t \mid X_{t-k}) = \phi_{11}\varepsilon_{t-1}X_{t-1}, k = 1$ y $E(X_t \mid X_{t-k}) = \phi_{11}\sigma_\varepsilon^2, k > 1$.

Por último, la clase general de procesos no lineales con **nolinealidad de tipo multiplicativo** (denominados procesos de volatilidad estocástica), se define como

$$X_t = \varepsilon_t \upsilon_t(\theta) \qquad (3.5)$$

donde ε_t es una secuencia de innovaciones con media nula, varianza unitaria e independiente de $\{X_{t-i}, i \geq 1\}$. El proceso (3.5) incluye como casos particulares tanto los procesos tipo GARCH (Bollerslev, 1986) como los de volatilidad estocástica SV (Andersen, 1994), donde $\upsilon_t(\theta) > 0$ es la función de volatilidad (o de desviación estándar condicional). En el caso GARCH(1,1), se tiene que

$$\upsilon_t^2(\theta) = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1\upsilon_{t-1}^2(\theta)$$

mientras que en el caso SV, una de las especificaciones más general viene dada por el modelo log-AR(1)-SV(1), de la forma

$$\lambda_t(\theta) = \phi_0 + \phi_1\lambda_{t-1}(\theta) + (\delta + \alpha_1\lambda_{t-1}(\theta))u_t$$

donde $\lambda_t(\theta) = \ln \upsilon_t^2(\theta)$ y las innovaciones ε_t y u_t, ambas secuencias *iid* de media nula y varianza unitaria, son mutuamente independientes. Haciendo α_1 = 0 se obtiene la habitual especificación log-normal del modelo SV(1) asumiendo que $\varepsilon_t \sim iidN(0,1)$. En general se tiene que si las innovaciones ε_t son *iid*, para todo proceso de la forma (3.5) se verifica que $E(X_t|X_{t-k}) = 0$.

Con el objeto de diseñar un procedimiento que permita contrastar la existencia de dependencia lineal de orden k en procesos con distribuciones posiblemente de colas anchas a través del cumplimiento de la condición (3.1), Gallagher (2002) recupera el concepto de función de covariación de Cambanis y Miller (1981). Utilizando la siguiente representación para el proceso X_{t-k},

$$X_{t-k} = \delta_{t-k} \mid X_{t-k} \mid$$

donde δ_t es la función signo de X_t, $\delta_t = sign(X_t) = I_{(X_t>0)} - I_{(X_t\leq 0)}$ ($I_{(A)}$ es la función indicador de A). Entonces por expectativas iteradas, si $E|X_t| < \infty$, se tiene que

$$\begin{aligned} E(X_t) &= E[E(X_t \mid X_{t-k})] = \lambda(k)E(X_{t-k}) = \lambda(k)E(\delta_{t-k} \mid X_{t-k} \mid) \\ &= -\lambda(k)E \mid X_{t-k} \mid \quad si\, X_{t-k} < 0 \\ &= \ \lambda(k)E \mid X_{t-k} \mid \quad si\, X_{t-k} < 0 \end{aligned}$$

y $|E(X_t)| = \lambda(k)E|X_{t-k}| \leq E|X_t| < \infty$. Por otro lado, a partir del proceso $W_{t,k} = X_t\delta_{t-k}$ se tiene que $E(W_{t,k}) = E(X_t\delta_{t-k}) = |E(X_t)|$, de forma que

$$E(X_t\delta_{t-k}) = \lambda(k)E \mid X_{t-k} \mid \tag{3.6}$$

y, por tanto, $\lambda(k)$ admite la representación alternativa

$$\lambda(k) = \frac{E(X_t\delta_{t-k})}{E \mid X_{t-k} \mid} \tag{3.7}$$

donde

$$\tau(k) = E(X_t\delta_{t-k}) = \lambda(k)E \mid X_{t-k} \mid \tag{3.8}$$

es la **función de covariación** del retardo k, con $\tau(0) = E(X_t\delta_t) = E(|X_t|\delta_t^2) = E|X_t|$. Bajo estacionariedad de primer orden del proceso $\{X_t\}$, de forma que $E|X_{t-k}| = E|X_t|$, se tiene

$$\tau(k) = E(X_t\delta_{t-k}) = \lambda(k)\tau(0) \tag{3.9}$$

Si el proceso es Gaussiano de media nula y estacionario, $\lambda(k) = \tau(k)/\tau(0) = \rho(k)$. En el caso de procesos de volatilidad estocástica, como en (3.5), se tiene que $\tau(k) = \tau(-k) = 0$ con tal de que las innovaciones sean *iid* con media nula y distribución simétrica, o bien sean débilmente dependientes de orden inferior a k con media nula y distribución simétrica como, por ejemplo, en el caso de que $\varepsilon_t \sim MA(1)$ y $k>1$[1]. Esto resulta del hecho de que en procesos de la forma (3.5) se tiene que

$$\delta_{t-k} = sign(X_{t-k}) = sign(\varepsilon_{t-k}\upsilon_{t-k}(\theta)) = sign(\varepsilon_{t-k})$$

de forma que

$$\tau(k) = E(X_t\delta_{t-k}) = E(\mid X_t \mid \delta_t\delta_{t-k}) \leq \sqrt{E(X_t^2)}\sqrt{E(\delta_t^2\delta_{t-k}^2)} = \sigma_0(\theta)\sqrt{E(\delta_t^2\delta_{t-k}^2)}$$

por la desigualdad de Hölder, con $E(X_t^2) = \sigma_0^2(\theta)$ (varianza incondicional del proceso GARCH). Si la secuencia de innovaciones ε_t es *iid* o bien es m-dependiente estacionaria con $m < k$, entonces $E(\delta_t^2\delta_{t-k}^2) = E(\delta_t^2)E(\delta_{t-k}^2)$. Si $\pi = P(\delta_t = 1)$ $(1-\pi = P(\delta_t = -1))$,

[1] En el caso GARCH con innovaciones *iid*, el proceso se denomina GARCH fuerte, mientras que con innovaciones con distribución estacionaria no independiente, el proceso se denomina GARCH semifuerte.

entonces $E(\delta_t^2) = 2\pi$ de forma que $\tau(k) \leq 2\pi\sigma_0(\theta)$. Bajo simetría de la distribución (π=1/2), entonces $\tau(k) \leq \sigma_0(\theta)$. En el caso GARCH(1,1), con $\sigma_0^2(\theta)=\alpha_0/(1-(\alpha_1+\beta_1))$, la función de covariación tomará un valor tanto mayor cuánto más próximo esté a la no estacionariedad en covarianza, es decir, cuando $(\alpha_1+\beta_1)\rightarrow 1$ (IGARCH(1,1)).

Cambanis y Miller (1981) introducen el concepto de función de covariación como una medida de dependencia lineal apropiada entre procesos α-estables simétricos ($S\alpha S$) con $1 < \alpha < 2$ y para procesos con momentos finitos de orden $1 < p < 2$ (denominados de forma general procesos de orden p). La función de covariación desempeña en estos procesos un papel análogo al de la función de covarianza y autocovarianza en procesos de orden $p \geq 2$.

Gallagher (2006) generaliza el contraste de dependencia de Gallagher (2002) basado en contrastar $\tau(k) = 0$ en (3.9) para $k = 1$, empleando retardos de orden superior de la función de covariación en un estadístico tipo Portmanteau para detectar dependencia lineal en procesos de series temporales con distribuciones simétricas. La base en la denominada **función de covariación muestral del retardo *k*** dada por

$$\begin{aligned}\hat{\tau}_T(k) &= (1/T)\sum_{t=1}^{T-k} X_{t+k}\delta_t \quad k > 0 \\ &= (1/T)\sum_{t=1}^{T-|k|} X_t\delta_{t+|k|} \quad k \leq 0\end{aligned} \tag{3.10}$$

es decir,

$$\hat{\tau}_T(k) = \frac{1}{T}\sum_{t=1}^{T-|k|} X_{t+(k\vee 0)}\delta_{t+|k\wedge 0|} \quad k = 0, \pm 1, \pm 2, ... \tag{3.11}$$

donde $a\vee b = max(a,b)$ y $a\wedge b = min(a,b)$), de forma que el estimador por el método de los momentos de $\lambda(k)$ en (3.9) viene dado por

$$\hat{\lambda}_T(k) = \frac{\hat{\tau}_T(k)}{\hat{\tau}_T(0)} = \frac{\sum_{t=1}^{T-|k|} X_{t+(k\vee 0)}\delta_{t+|k\wedge 0|}}{\sum_{t=1}^{T} |X_t|} \tag{3.12}$$

Bajo estacionariedad y ergodicidad de X_t, con $E|X_t|<\infty$, $\hat{\tau}_T(k)$ es un estimador fuertemente consistente de $\tau(k)$. Gallagher (2001a,b) demuestra que cuando X_t sigue un proceso ARMA causal con innovaciones $S\alpha S$, se verifica que $T^{1-1/\alpha}(\hat{\tau}_T(k) - \tau(k)) \Rightarrow S_\alpha$, donde S_α tiene distribución $S\alpha S$. Si el proceso X_t es *iid* $S\alpha S$(Gaussiano)*,* con $\alpha \in (1,2]$, entonces $(T-k)^{1-1/\alpha}\hat{\tau}_T(k)$ tiene distribución exacta $S\alpha S$ (Gaussiana). Gallagher y Okuyama (2003) estudian la tasa de convergencia de la función de covariación muestral para secuencias *iid* y encuentran que es superior a la de la usual función de autocovarianzas. Se espera así que los contrastes de hipótesis basados en la función de covariación muestral tengan buenas propiedades para tamaños muestrales pequeños.

En general, $\tau(k) \neq \tau(-k)$, pero si el proceso es *iid* con media nula, $\tau(k) = \tau(-k) = 0$. Con objeto de explotar este resultado en procesos no *iid*, se puede definir la **función de covariación simétrica** como

$$\eta(k) = \frac{1}{2}(\tau(k) + \tau(-k)) \tag{3.13}$$

que puede estimarse consistentemente por

$$\hat{\eta}_T(k) = \frac{1}{2}(\hat{\tau}_T(k) + \hat{\tau}_T(-k)) \tag{3.14}$$

A partir de estos resultados, Gallagher (2006) considera dos estadísticos de contraste tipo Portmanteau para la detección de dependencia lineal basados en la función de covariación muestral. El primero se basa en las primeras h covariaciones de retardo positivo para el caso de procesos con varianza infinita y viene dado por

$$\hat{Q}_T(\alpha, h) = \hat{d}_T^{-2} T^{2(1-1/\alpha)} \sum_{k=1}^{h} \hat{\tau}_T^2(k) \tag{3.15}$$

donde $\hat{d}_T$ es un estimador consistente de la escala del proceso d[2]. Bajo la hipótesis nula de que X_t es *iid* con distribución en el dominio normal de atracción de una ley $S\alpha S$ con $\alpha \in (0,2]$, el Teorema 1.1 de Gallagher (2006) establece que, bajo los resultados de convergencia de sumas parciales de procesos $S\alpha S$ de Davis (1983),

$$\hat{Q}_T(\alpha, h) \Rightarrow c(h)^{1-2/\alpha} \boldsymbol{Y}_h' \boldsymbol{A} \boldsymbol{Y}_h$$

donde $c(h) = 2^{h-1}$, $\boldsymbol{Y}_h$ es un vector de h variables *iid* $S\alpha S$ con escala $d = 1$ y $\boldsymbol{A}$ es una matriz idempotente de rango h. En el caso de varianza finita, con $\alpha = 2$,

$$\hat{Q}_T(2, h) = \hat{d}_T^{-2} T \sum_{k=1}^{h} \hat{\tau}_T^2(k) \Rightarrow \chi_h^2,$$

con $\hat{d}_T^2 = \hat{\sigma}_T^2$ un estimador consistente de la varianza de X_t. El segundo estadístico de contraste, indicado especialmente para el caso de varianza finita, trata de mejorar la

---

[2] En el caso de distribución estable simétrica con escala d con $\alpha \in (1,2)$, Gallagher (2006) considera dos estimadores alternativos con buen comportamiento bajo la hipótesis alternativa. En primer lugar, teniendo en cuenta que $E|X_t| = 2d\Gamma(1-1/\alpha)/\pi$, bajo estacionariedad y ergodicidad se obtiene el estimador fuertemente consistente $\hat{d}_T = \hat{\tau}_T(0)\pi / 2\Gamma(1-1/\alpha)$, con $\hat{\tau}_T(0) - E|X_t| = O(T^{1/\alpha - 1})$. En segundo lugar, considera el estimador basado en cuantiles de McCulloch (1986), que es también consistente para datos iid.

potencia en muestras finitas explotando la no simetría de la función de covariación empleando las primeras h funciones de covariación simétricas, de la forma

$$\hat{Q}_T(h) = (\hat{\mu}_T^2 + \hat{\sigma}_T^2)^{-1} T \sum_{k=1}^{h} \hat{\eta}_T^2(k) \qquad (3.16)$$

donde $\hat{\mu}_T$ es un estimador consistente de $E|X_t|$. En el caso de varianza infinita, la convergencia de $\hat{Q}_T(h)$ a un múltiplo de $\boldsymbol{Y}_h'\boldsymbol{B}\boldsymbol{Y}_h$ (donde $\boldsymbol{B}$ es una matriz idempotente de rango h e $\boldsymbol{Y}_h$ es un vector de 2^{2h-1} variables *iid* $S\alpha S$) es muy lenta, lo que hace inapropiado su uso. Ambos estadísticos divergen asintóticamente bajo la alternativa a una tasa $T^{2(1-1/\alpha)}$ y la potencia aumenta a medida que $\alpha \to 2$. El Teorema 1.4 de Gallagher (2006) establece que bajo la hipótesis de que X_t es *iid* con varianza finita, entonces, $2\hat{Q}_T(h) \Rightarrow \chi_h^2$.

Para una correcta implementación de estos contrastes es preciso realizar un par de observaciones.

Observación 1. *Corrección por media no nula.* Con el objeto de evitar un rechazo espúreo en el caso de que el proceso observado no tenga media nula, es conveniente en la práctica simplemente corregir las observaciones por la media muestral. Bajo ergodicidad del proceso estacionario, los resultados asintóticos permanecen inalterados.

Observación 2. *Corrección por no simetría del proceso*. Los procesos lineales basados en innovaciones *iid* Gaussianas son simétricos, es decir, reversibles en el tiempo. Sin embargo, esta propiedad no se verifica en general bajo no normalidad. Gallagher (2006) indica que si el proceso X_t no es simétrico, en lugar de utilizar éste se puede emplear la secuencia $Z_{2t} = X_{2t} - X_{2t-1}$. Bajo la hipótesis nula *iid*, Z_{2t} es también *iid* con distribución simétrica. Una característica similar a la de la distribución asimétrica es la distribución irreversible en el tiempo, es decir, cuando las características del proceso son distintas hacia delante que hacia atrás en el tiempo. Un ejemplo de este tipo de procesos son, en general, los procesos bilineales (ver Tong, 1990).

4. La función de covariación de orden *m*

Con el objeto de detectar dependencia lineal en momentos de orden superior al primero para procesos de orden m (m>1) definimos a continuación la función de covariación de orden m. En primer lugar, a partir de la convención $X_t^m = \delta_t \,|\, X_t \,|^m$ y asumiendo la existencia del momento absoluto de orden m del proceso,

$\lambda_m = E|X_t|^m < \infty$

de forma que el momento de orden m es finito, $\mu_m = E(X_t^m) \le \lambda_m$, se tiene que, de forma similar a (3.1), el proceso X_t presencia dependencia de orden m en el retardo k si

$$E(X_t^m \mid X_{t-k}^m) = \lambda_m(k) X_{t-k}^m \tag{4.1}$$

Consideremos a continuación tres ejemplos para el caso particular $m = 2$. En primer lugar, a partir de (3.2) para el proceso AR(1) (con $\phi_0 = 0$), se tiene que

$$E(X_t^2 \mid X_{t-k}^2) = \sigma_\varepsilon^2 \sum_{i=0}^{k-1} \phi_1^{2i} + \phi_1^{2k} X_{t-k}^2$$

En segundo lugar, a partir del proceso GARCH(1,1), $\upsilon_t^2(\theta) = \alpha_0 + A_{t-1}(\theta)\upsilon_{t-1}^2(\theta)$, mediante sustituciones se tiene la siguiente representación

$$\upsilon_t^2(\theta) = \alpha_0 \left(1 + \sum_{j=1}^{k-1} \prod_{i=1}^{j} A_{t-i}(\theta) \right) + \upsilon_{t-k}^2(\theta) \prod_{i=1}^{k} A_{t-i}(\theta)$$

de forma que, bajo distribución *iid* de las innovaciones con varianza unitaria,

$$\begin{aligned} E(X_t^2 \mid X_{t-k}^2) &= E(\varepsilon_t^2 \upsilon_t^2(\theta) \mid X_{t-k}^2) = E(\varepsilon_t^2) E(\upsilon_t^2(\theta) \mid X_{t-k}^2) = E(\upsilon_t^2(\theta) \mid \varepsilon_{t-k}^2 \upsilon_{t-k}^2(\theta)) \\ &= \alpha_0 \left(1 + \sum_{j=1}^{k-1} (\alpha_1 + \beta_1)^j \right) + (\alpha_1 + \beta_1)^{k-1} A_{t-k}(\theta) \upsilon_{t-k}^2(\theta) \end{aligned}$$

puesto que $A_{t-i}(\theta)$ es también una secuencia *iid* con $E(A_{t-i}(\theta)) = \alpha_1 + \beta_1$, y donde

$A_{t-k}(\theta)\upsilon_{t-k}^2(\theta) = \alpha_1 X_{t-k}^2 + \beta_1 \upsilon_{t-k}^2(\theta)$.

Por tanto, a partir de la definición (4.1), un proceso GARCH fuerte presenta (de forma aproximada) dependencia lineal de segundo orden para todos los retardos k. Finalmente, haciendo $\phi_0 = \phi_1 = 0$ en (3.4), se tiene que en el proceso BL(0,0,1,1),

$$E(X_t^2 \mid X_{t-k}) = \sigma_\varepsilon^2 + (\phi_{11}^{2k} \sigma_\varepsilon^{2(k-1)}) \varepsilon_{t-k}^2 X_{t-k}^2 + \phi_{11}^2 E(\varepsilon_t^4) \sum_{i=1}^{k-1} (\phi_{11}^2 \sigma_\varepsilon^2)^{i-1}$$

de forma que presenta también dependencia lineal de segundo orden para el retardo k. El nivel de dependencia medido por esta esperanza condicional será tanto menor cuánto menor sea la dispersión de la distribución de las innovaciones.

Operando en (4.1) por expectativas iteradas se tiene que

$$\begin{aligned} E(X_t^m) &= \lambda_m(k) E(X_{t-k}^m) = \lambda_m(k) E(\delta_{t-k} \mid X_{t-k}|^m) \\ &= -\lambda_m(k) E|X_{t-k}|^m \quad si\, \delta_{t-k} = -1 \\ &= \ \ \lambda_m(k) E|X_{t-k}|^m \quad si\, \delta_{t-k} = 1 \end{aligned}$$

con $|E(X_t^m)| = \lambda_m(k) E|X_{t-k}|^m \le E|X_t|^m < \infty$. Puesto que $|E(X_t^m \delta_{t-k})| = |E(X_t^m)|$, se define entonces la **función de covariación de orden *m*** para el retardo k como

$\tau_m(k) = E(X_t^m \delta_{t-k}) = \lambda_m(k) E \mid X_{t-k} \mid^m$

que bajo estacionariedad y la condición de existencia del momento, puede expresarse alternativamente como

$$\tau_m(k) = E(X_t^m \delta_{t-k}) = \lambda_m(k) E \mid X_t \mid^m = \lambda_m(k)\tau_m(0) \tag{4.2}$$

Con objeto de establecer la analogía con el caso m = 1 anterior y garantizar la validez de los resultados distribucionales de los contrastes propuestos es preciso definir la función de covariación del proceso X_t^m corregido como $\tilde{X}_{t,m} = X_t^m - \mu_m$, de forma que se define la **función de covariación de orden *m* corregida** como

$$\begin{aligned} \tilde{\tau}_m(k) &= E(\tilde{X}_{t,m}\delta_{t-k}) = E(X_t^m \delta_{t-k}) - \mu_m E(\delta_{t-k}) \\ &= \tau_m(k) - \mu_m(2\pi - 1) \end{aligned} \tag{4.3}$$

donde $\tilde{\tau}_m(k) = \tau_m(k)$ bajo simetría de la distribución (π = 1/2). Bajo independencia de orden m en el retardo k, se tiene que $\tilde{\tau}_m(k)$=$E(\tilde{X}_{t,m})E(\delta_{t-k})$=0, mientras que $\tau_m(k)$ = $\mu_m E(\delta_{t-k})$. En el caso GARCH, se tiene que la función de covariación de orden m corregida viene dada por

$\tilde{\tau}_m(k) = E(\mid X_t \mid^m \delta_t \delta_{t-k}) - \mu_m(\theta) E(\delta_{t-k})$

donde $\mu_m(\theta) = E(X_t^m) = E(\varepsilon_t^m)E(\upsilon_t^m(\theta))$. Entonces, por la desigualdad de Hölder

$\tilde{\tau}_m(k) \le \sqrt{E \mid X_t \mid^{2m}} \sqrt{E(\delta_t^2 \delta_{t-k}^2)} - \mu_m(\theta)E(\delta_{t-k}) = \sqrt{\lambda_{2m}(\theta)}\sqrt{E(\delta_t^2\delta_{t-k}^2)} - \mu_m(\theta)E(\delta_{t-k})$

Bajo el supuesto de distribución *iid* para las innovaciones y existencia del momento absoluto de orden 2m del proceso GARCH, entonces

$\tilde{\tau}_m(k) \le 2\pi\sqrt{\lambda_{2m}(\theta)} - \mu_m(\theta)(2\pi - 1)$

que bajo simetría (π=1/2), resulta de la forma $\tilde{\tau}_m(k) \le \sqrt{\lambda_{2m}(\theta)}$. Puesto que $\lambda_{2m}(\theta) \ge \mu_{2m}(\theta)$, la función de covariación de orden m corregida tomará un valor tanto mayor cuánto menor sea el índice de momento finito del proceso GARCH, κ(θ), es decir, cuánto mayor sea la curtosis de la distribución marginal de dimensión finita de colas anchas del proceso GARCH estrictamente estacionario. Por lo tanto, cabe esperar que si para un proceso particular se encuentra que $\tilde{\tau}(k) = \tilde{\tau}_1(k) \approx 0$ y $\tilde{\tau}_2(k) \approx 0$ (bajo existencia del momento de cuarto orden), esto podrá tomarse como indicación de que el proceso es posiblemente GARCH con innovaciones *iid* y distribución de colas no muy anchas. Sin embargo, si $\tilde{\tau}(k) = \tilde{\tau}_1(k) \approx 0$ y $\tilde{\tau}_2(k) > 0$ esto podrá tomarse como indicación de que el proceso es de tipo GARCH con distribución marginal de colas anchas.

Con el objeto de implementar un procedimiento de contraste que permita evidenciar este comportamiento de $\tilde{\tau}_m(k)$, se define la **función de covariación muestral de orden *m* corregida** como

$$\tilde{\tau}_{m,T}(k) = \frac{1}{T}\sum_{t=1}^{T-|k|} \tilde{X}_{t+(k\vee 0),m}\delta_{t+|k\wedge 0|} \qquad k = 0, \pm 1, \pm 2, ... \tag{4.3}$$

donde

$$\tilde{X}_{t,m} = X_t^m - \hat{\mu}_{m,T},\ \hat{\mu}_{m,T} = \frac{1}{T}\sum_{s=1}^{T} X_s^m \tag{4.4}$$

y, de forma equivalente a (3.14), se define la **función de covariación simétrica muestral de orden *m* corregida** como

$$\tilde{\eta}_{m,T}(k) = \frac{1}{2}(\tilde{\tau}_{m,T}(k) + \tilde{\tau}_{m,T}(-k)) \tag{4.5}$$

Así, bajo la condición de existencia del momento de orden $2m$ del proceso X_t, se definen los siguientes estadísticos tipo Portmanteau para el contraste de la existencia de dependencia lineal de orden m hasta el retardo $h > 0$, a partir de la significación de la función de covariación muestral,

$$\hat{Q}_{T,m}(\tau, h) = \frac{T}{\hat{M}_{2,T}(m)}\sum_{k=1}^{h} \tilde{\tau}_{m,T}^2(k) \tag{4.6}$$

y

$$\hat{Q}_{T,m}(\eta, h) = \frac{T}{(\hat{M}_{2,T}(m) + \hat{\lambda}_{m,T}^2)}\sum_{k=1}^{h} \tilde{\eta}_{m,T}^2(k) \tag{4.7}$$

donde

$$\hat{\lambda}_{m,T} = \frac{1}{T}\sum_{t=1}^{T} | X_t |^m \tag{4.8}$$

y

$$\hat{M}_{2,T}(m) = \frac{1}{T}\sum_{t=1}^{T} \tilde{X}_{t,m}^2 = \frac{1}{T}\sum_{t=1}^{T} (X_t^m - \hat{\mu}_{m,T})^2 \tag{4.9}$$

son, bajo ergodicidad del proceso, estimadores consistentes de los correspondientes momentos de la distribución. La distribución muestral asintótica de estos estadísticos para contrastar $H_{0,h}$: $\lambda_m(k) = 0$ para todo $k = 1, \ldots, h$ frente a la alternativa $H_{1,h}$: $\lambda_m(k) \neq 0$ para algún $k = 1, \ldots, h$, se establece a continuación en los dos siguientes teoremas.

Teorema 1. *Sea* $Y_1^m,\dots,Y_T^m$ *una secuencia de variables aleatorias iid de media nula, con* $\tilde{\mu}_{2,m} = E(|Y_1^m|^2) < \infty$. *Entonces para un entero positivo h, cuando* $T \to \infty$,

$$\hat{Q}_{T,m}(\tau, h) = \frac{T}{\hat{M}_{2,T}(m)} \sum_{k=1}^{h} \tilde{\tau}_{m,T}^2(k) \Rightarrow \chi^2(h)$$

donde $\chi^2(h)$ *denota una variable aleatoria* χ^2 *con h grados de libertad.*

Teorema 2. *Sea* $Y_1^m,\dots,Y_T^m$ *una secuencia de variables aleatorias iid de media nula, con* $\tilde{\mu}_{2,m} = E(|Y_1^m|^2) < \infty$. *Entonces para un entero positivo h, cuando* $T \to \infty$,

$$2\hat{Q}_{T,m}(\eta, h) = \frac{2T}{(\hat{M}_{2,T}(m) + \hat{\lambda}_{m,T}^2)} \sum_{k=1}^{h} \tilde{\eta}_{m,T}^2(k) \Rightarrow \chi^2(h)$$

donde $\chi^2(h)$ *denota una variable aleatoria* χ^2 *con h grados de libertad.*

La prueba de estos resultados se establece en el Apéndice bajo aplicación simple del TCL para procesos *h*-dependientes de Hoeffding y Robbins (1948) y la extensión de Berk (1973) al caso de procesos *h*-dependientes con orden de dependencia función del orden de momento máximo finito.

Considerando la aplicabilidad de esos estadísticos para discriminar entre procesos no lineales de tipo aditivo de los de no linealidad multiplicativa, cabe hacer las siguientes observaciones.

Observación 1. Como se discutió anteriormente, en el caso de un proceso GARCH, la función de covariación de segundo orden podría presentar coeficientes no nulos, especialmente en el caso de distribución con colas anchas. Sin embargo, la normalización de los coeficientes de la función de covariación muestral en los estadísticos propuestos en (4.6) y (4.7) permitirá que éstos recojan en este caso básicamente las características de las innovaciones, presentando resultados de no significación.

Observación 2. Por lo establecido anteriormente, la estrategia a considerar sería considerar en el caso GARCH como hipótesis a contrastar, H_0: $\tau_m(k) = \tau_m(-k) = 0$ (no GARCH), frente a H_1: $\tau_m(k)$, $\tau_m(-k) \neq 0$ para algún k y $m = 1$ y 2, es decir, combinando los resultados de los estadísticos de Gallagher (2006) y los propuestos en este trabajo.

Observación 3. En el caso de un proceso no lineal de tipo aditivo como el bilineal, es de esperar que el estadístico con mayor capacidad para detectar este tipo de comportamiento sea el basado en las covariaciones simétricas (4.7), debido a la característica de irreversibilidad temporal de este tipo de procesos.

5. Resultados de simulación y aplicación a rendimientos de tipos de cambio

5.1 Resultados de simulación

En este apartado presentamos los principales resultados obtenidos en un experimento de simulación más amplio centrado en evaluar la distribución de los estadísticos Portmanteau propuestos por Gallagher (2006) (en el caso m = 1) y los propuestos en este trabajo, para el contraste de dependencia lineal de primer orden empleando la función de covariación muestral, en el caso de varianza finita, con procesos no lineales como los descritos en el apartado 2. Por razones de espacio, sólo se presentarán los resultados relativos a los estadísticos $\hat{Q}_{T,m}(\tau,h)$ y $\hat{Q}_{T,m}(\eta,h)$ para m = 2 propuestos en el apartado anterior. Sin embargo, es preciso comentar que los estadísticos para m = 1, tienen el comportamiento esperado: presentan coeficientes significativos de la función de covariación en el caso de no linealidad de tipo aditivo y no significativos en el caso de no linealidad multiplicativa (GARCH y SV). En primer lugar, el cuadro 1 siguiente presenta los percentiles empíricos superiores de la distribución de ambos estadísticos para un proceso *iid* (distribución nula), considerando distribuciones N(0,1), T de Student normalizada con varianza unitaria con 6 y 3 grados de libertad, ST(v=6,3) y distribución de Pareto generalizada estándar (DPG) con parámetro de forma a = 1.5. Los resultados indican la robustez de la distribución asintótica ante distribuciones de colas anchas, incluso en el caso de varianza infinita como en el caso DPG con a < 2.

Cuadro 1. *Cuantiles superiores*, $q(1-\alpha)$, *de la distribución nula de los estadísticos* $\hat{Q}_{m,T}(\tau,h)$ *y* $\hat{Q}_{m,T}(\eta,h)$, *procesos iid, m* = 2

N(0,1)		$q(0.90)$	$q(0.95)$	$q(0.975)$	$q(0.99)$
$\hat{Q}_{m,T}(\tau,h)$	h = 1	2.695	3.732	4.773	6.439
	2	4.594	5.953	7.266	9.006
	5	9.384	11.199	13.039	14.891
	10	15.932	18.324	20.987	23.903
$\hat{Q}_{m,T}(\eta,h)$	h = 1	0.909	1.295	1.666	2.243
	2	1.553	2.025	2.501	3.001
	5	3.043	3.669	4.252	5.118
	10	5.316	6.037	6.737	7.606
ST(v = 6)		$q(0.90)$	$q(0.95)$	$q(0.975)$	$q(0.99)$
$\hat{Q}_{m,T}(\tau,h)$	h = 1	2.791	4.054	5.100	6.705
	2	4.597	5.909	7.377	9.085
	3	6.329	7.892	9.378	10.937
	4	7.844	9.384	10.939	12.856
	5	9.136	10.779	12.502	15.004
	10	15.614	18.091	20.038	22.820
	20	27.948	30.902	33.987	37.049
$\hat{Q}_{m,T}(\eta,h)$	h = 1	1.168	1.617	2.099	2.959
	2	1.920	2.501	3.132	3.902
	3	2.651	3.276	3.911	4.705
	4	3.284	4.061	4.769	5.669
	5	3.821	4.706	5.402	6.365
	10	6.593	7.603	8.513	9.659
	20	11.801	13.020	14.233	15.245

Cuadro 1. (*continuación*)

ST(v = 3)		q(0.90)	q(0.95)	q(0.975)	q(0.99)
$\hat{Q}_{m,T}(\tau,h)$	h = 1	2.544	3.444	4.296	5.668
	2	4.249	5.299	6.475	7.896
	3	5.703	7.027	8.325	9.771
	4	7.185	8.654	9.863	11.440
	5	8.587	10.018	11.345	13.099
	10	15.036	16.995	18.942	21.423
	20	26.839	29.527	31.672	35.395
$\hat{Q}_{m,T}(\eta,h)$	h = 1	1.346	1.818	2.244	2.919
	2	2.193	2.738	3.311	4.018
	3	2.946	3.556	4.259	5.052
	4	3.712	4.457	5.119	5.967
	5	4.454	5.165	5.918	6.689
	10	7.641	8.549	9.455	10.530
	20	13.435	14.758	15.870	17.181
DPG (a = 1.5)		q(0.90)	q(0.95)	q(0.975)	q(0.99)
$\hat{Q}_{m,T}(\tau,h)$	h = 1	2.240	2.752	3.288	4.498
	2	2.987	3.944	4.999	5.905
	3	4.137	5.268	5.972	7.396
	4	5.264	6.058	7.289	8.705
	5	5.976	7.265	8.254	9.957
	10	10.481	11.835	13.405	15.207
	20	18.384	20.548	22.359	24.505
$\hat{Q}_{m,T}(\eta,h)$	h = 1	0.633	1.053	1.992	2.653
	2	1.260	2.213	2.808	3.334
	3	1.968	2.866	3.272	3.776
	4	2.656	3.324	3.769	4.624
	5	3.215	3.729	4.315	5.479
	10	5.138	6.124	6.853	7.857
	20	9.188	10.229	11.418	12.672

Nota: Resultados basados en 5000 replicaciones con 10000 observaciones

El cuadro 2 presenta los resultados de la distribución de los estadísticos propuestos en el caso de un proceso GARCH(1,1) con α_0 = 0.05, α_1 = 0.15 y β_1 = 0.65, para retardos h = 1, 2, 5 y 10 y distintos tamaños muestrales. Se observa la rápida convergencia de los estadísticos a sus distribuciones límite, puesto que los resultados son invariantes al tamaño muestral considerado, incluso para T = 100 o 250. Los resultados son muy similares para otros supuestos distribucionales, como en el caso $ST(v{=}6)$ o $SGED(v{=}0.5,1.0,3.0)$ (distribución generalizada del error estandarizada). Bajo innovaciones *iid*, ambos estadísticos detectarían la existencia de efectos GARCH al tomar un valor pequeño y no rechazar la hipótesis de significación nula de los coeficientes de covariación. El cuadro 3 presenta los resultados para el caso de un retardo para el proceso IGARCH(1,1) con innovaciones *iid*N(0,1) (los resultados son invariantes para distribuciones de colas anchas). A pesar de que el proceso tiene varianza infinita, en el comportamiento de los estadísticos domina el efecto multiplicativo de las innovaciones y detecta correctamente el tipo de no linealidad. El cuadro 4 siguiente presenta resultados para el caso de un proceso BL(0,0,1,1) con ϕ_{11} = 0.50 y errores *iid*N(0,1) (los resultados son similares para otras distribuciones de colas anchas). Es de destacar que el estadístico relevante para detectar este tipo de no linealidad es el basado en la covariaciones simétricas, $\hat{Q}_{T,m}(\eta,h)$, dado el carácter irreversible en el tiempo de este proceso.

Cuadro 2. *Cuantiles superiores*, $q(1-\alpha)$, *de la distribución de los estadísticos* $\hat{Q}_{m,T}(\tau,h)$ *y* $\hat{Q}_{m,T}(\eta,h)$, $m=2$, *para un proceso GARCH*(1,1) *con errores iidN*(0,1).

			$q(0.90)$	$q(0.95)$	$q(0.975)$	$q(0.99)$
$\hat{Q}_{2,T}(\tau,h)$	$h=1$	$T=100$	2.351	3.357	4.019	4.923
		250	2.636	3.889	5.009	6.554
		500	2.667	3.535	4.270	5.429
		1000	3.174	4.452	5.045	6.998
		2000	2.498	3.260	4.170	5.324
$\hat{Q}_{2,T}(\eta,h)$		$T=100$	0.882	1.256	1.528	2.007
		250	0.949	1.215	1.499	2.268
		500	1.005	1.480	1.789	2.242
		1000	0.977	1.496	1.733	2.341
		2000	0.888	1.178	1.506	1.697

			$q(0.90)$	$q(0.95)$	$q(0.975)$	$q(0.99)$
$\hat{Q}_{2,T}(\tau,h)$	$h=2$	$T=100$	4.537	5.616	7.101	9.257
		250	4.528	6.296	7.201	8.569
		500	4.206	5.603	7.978	9.476
		1000	4.859	6.630	8.282	9.193
		2000	4.752	6.203	7.234	8.957
$\hat{Q}_{2,T}(\eta,h)$		$T=100$	1.678	2.101	2.747	3.930
		250	1.657	2.148	2.261	3.126
		500	1.566	2.014	2.642	3.109
		1000	1.537	2.162	2.771	3.788
		2000	1.711	2.205	2.474	2.781

			$q(0.90)$	$q(0.95)$	$q(0.975)$	$q(0.99)$
$\hat{Q}_{2,T}(\tau,h)$	$h=5$	$T=100$	9.559	11.511	13.004	17.528
		250	8.142	9.789	11.696	13.464
		500	8.926	10.626	13.276	14.586
		1000	8.659	10.484	12.091	12.902
		2000	9.078	11.138	12.653	16.352
$\hat{Q}_{2,T}(\eta,h)$		$T=100$	3.101	3.606	4.374	5.843
		250	3.126	3.847	4.224	5.023
		500	3.037	3.566	3.966	4.783
		1000	3.361	3.839	4.418	4.793
		2000	3.198	3.894	4.658	5.867

			$q(0.90)$	$q(0.95)$	$q(0.975)$	$q(0.99)$
$\hat{Q}_{2,T}(\tau,h)$	$h=10$	$T=100$	14.086	16.643	18.132	20.259
		250	15.984	18.235	21.294	24.855
		500	15.899	18.063	20.492	23.935
		1000	15.945	19.225	21.051	22.322
		2000	16.583	19.194	21.823	25.203
$\hat{Q}_{2,T}(\eta,h)$		$T=100$	5.023	6.059	7.063	9.354
		250	5.088	5.919	6.684	8.141
		500	5.458	6.479	7.173	7.960
		1000	5.562	6.635	6.888	8.098
		2000	5.615	6.774	7.396	7.898

Nota: Resultados basados en 5000 replicaciones.

Cuadro 3. *Cuantiles superiores*, $q(1-\alpha)$, *de la distribución de los estadísticos* $\hat{Q}_{m,T}(\tau,h)$ *y* $\hat{Q}_{m,T}(\eta,h)$, $m=2$, *para un proceso IGARCH*(1,1) *con errores iidN*(0,1).

			$q(0.90)$	$q(0.95)$	$q(0.975)$	$q(0.99)$
$\hat{Q}_{2,T}(\tau,h)$	$h=1$	$T=100$	2.777	3.759	4.829	6.058
		250	2.700	3.732	4.736	6.380
		500	2.639	3.704	4.759	6.132
		1000	2.702	3.910	4.864	6.736
		2000	2.639	3.764	4.819	6.290
$\hat{Q}_{2,T}(\eta,h)$		$T=100$	1.062	1.501	1.887	2.467
		250	1.099	1.545	1.989	2.674
		500	1.209	1.706	2.234	2.900
		1000	1.373	1.886	2.559	3.379
		2000	1.433	2.067	2.779	3.566

Nota: Resultados basados en 5000 replicaciones.

Cuadro 4. *Cuantiles superiores*, $q(1-\alpha)$, *de la distribución de los estadísticos* $\hat{Q}_{m,T}(\tau,h)$ *y* $\hat{Q}_{m,T}(\eta,h)$, $m=2$, *en el caso de varianza finita para un proceso BL*(0,0,1,1) *con errores iidN*(0,1).

			$q(0.90)$	$q(0.95)$	$q(0.975)$	$q(0.99)$
$\hat{Q}_{2,T}(\tau,h)$	$h=1$	$T=100$	2.375	3.159	3.882	4.896
		250	2.837	3.620	4.648	5.645
		500	3.808	4.912	5.986	7.254
		1000	5.402	6.727	7.805	9.471
		2000	8.349	9.909	11.467	13.243
$\hat{Q}_{2,T}(\eta,h)$		$T=100$	2.219	2.724	3.137	3.739
		250	3.869	4.555	5.171	5.844
		500	6.316	7.246	8.087	9.242
		1000	10.580	11.782	12.880	14.302
		2000	18.509	20.351	21.810	23.590

Nota: Resultados basados en 5000 replicaciones.

5.2 Aplicación

Existe una amplia evidencia en la literatura econométrica sobre la validez de los modelos tipo GARCH para describir el comportamiento y las principales características de las series de rendimientos de tipos de cambio en altas frecuencias, bajo distintos supuestos sobre la distribución condicional el proceso. Algunas referencias son: Bollerslev (1987) y Hsieh (1989) que consideran el caso de la distribución *T* de Student, Nelson (1991) y Hsieh (1989) la distribución de error generalizada, Alexander y Lazar (2004) un proceso GARCH(1,1) con distribución mezcla de normales, y Jensen y Lunde (2001), Forsberg y Bollerslev (2002) y Kiliç (2007) la distribución normal inversa en modelos GARCH y FIGARCH, respectivamente.

A continuación se presentan los resultados de los contrastes Portmanteau basados en la función de covariación muestral y la función de covariación muestral simétrica para m = 1 y 2 para tres series de rendimientos diarios de los tipos de cambio dólar/libra esterlina, euro/dólar y yen/dólar para el período 03.01.2000-27.02.2009 con un total de T = 2303 observaciones. Los resultados se presentan en los siguientes cuadros 5-7. Para m = 1, el estadístico Portmanteau $\hat{Q}_T(\alpha, h)$ se ha calculado para α = 1.1,1.2,…,2.0.

Cuadro 5. *Resultados de los estadísticos Portmanteau basados en la función de covariación muestral para m* = 1, 2, *para los rendimientos del tipo de cambio dólar/libra esterlina.*

			h = 1	2	3	4	5	10	20
(m = 1)	$\hat{Q}_T(\alpha, h)$	α = 1.1	0.0802	0.4781	0.7261	1.7419	1.8508	2.2772	3.8735
		1.2	0.0727	0.4338	0.6587	1.5804	1.6790	2.0659	3.5141
		1.3	0.0986	0.5881	0.8932	2.1427	2.2765	2.8011	4.7645
		1.4	0.1471	0.8771	1.3320	3.1955	3.3951	4.1774	7.1056
		1.5	0.2225	1.3268	2.0149	4.8340	5.1359	6.3193	10.7490
		1.6	0.3321	1.9802	3.0074	7.2149	7.6654	9.4318	16.0432
		1.7	0.4842	2.8875	4.3854	10.5206	11.1776	13.7533	23.3939
		1.8	0.6879	4.1022	6.2300	14.9460	15.8794	19.5385	33.2344
		1.9	0.9523	5.6785	8.6239	20.6892	21.9812	27.0464	46.0051
		2.0	1.2862	7.6698	11.6482	27.9445	29.6895	36.5309	62.1381
	$\hat{Q}_T(h)$		0.1046	0.3344	1.0156	2.1537	2.2652	4.0395	6.0984
(m = 2)	$\hat{Q}_{T,m}(\tau, h)$		0.0014	3.2361	3.2951	3.7191	3.7216	12.5969	28.7092
	$\hat{Q}_{T,m}(\eta, h)$		0.7073	2.0742	2.0884	2.0884	2.5759	8.1996	12.0722

Cuadro 6. *Resultados de los estadísticos Portmanteau basados en la función de covariación muestral para m* = 1, 2, *para los rendimientos del tipo de cambio euro/dólar.*

			h = 1	2	3	4	5	10	20
(m = 1)	$\hat{Q}_T(\alpha, h)$	α = 1.1	0.0087	0.0094	0.0468	0.4393	1.2866	2.3738	3.5201
		1.2	0.0079	0.0086	0.0425	0.3986	1.1672	2.1536	3.1935
		1.3	0.0107	0.0116	0.0576	0.5404	1.5825	2.9199	4.3299
		1.4	0.0159	0.0173	0.0859	0.8059	2.3601	4.3546	6.4574
		1.5	0.0242	0.0262	0.1299	1.2192	3.5702	6.5874	9.7684
		1.6	0.0360	0.0391	0.1938	1.8197	5.3286	9.8319	14.5796
		1.7	0.0525	0.0569	0.2827	2.6534	7.7701	14.3367	21.2597
		1.8	0.0747	0.0809	0.4016	3.7695	11.0386	20.3673	30.2024
		1.9	0.1033	0.1120	0.5559	5.2179	15.2803	28.1937	41.8081
		2.0	0.1396	0.1513	0.7508	7.0478	20.6387	38.0806	56.4692
	$\hat{Q}_T(h)$		0.2199	0.3549	0.5340	2.3088	3.9968	8.7693	12.3823
(m = 2)	$\hat{Q}_{T,m}(\tau, h)$		2.5814	2.9207	3.4174	4.8506	5.2074	15.1591	25.8951
	$\hat{Q}_{T,m}(\eta, h)$		0.5444	0.5525	1.0055	2.0032	2.0276	6.2686	14.1456

Para el tipo de cambio euro/dólar existe cierta evidencia de dependencia no lineal de tipo aditivo (no GARCH) al 10% de significación, empleando ambas versiones del estadístico Portmanteau, aunque puede considerarse débil y está basada en un comportamiento poco regular de ambos estadísticos.

Cuadro 7. *Resultados de los estadísticos Portmanteau basados en la función de covariación muestral para* $m = 1, 2$, *para los rendimientos del tipo de cambio yen/dólar.*

			$h = 1$	2	3	4	5	10	20
($m = 1$)	$\hat{Q}_T(\alpha, h)$	$\alpha = 1.1$	0.0252	0.0254	0.2009	0.2715	0.2925	0.5446	1.3482
		1.2	0.0228	0.0231	0.1823	0.2463	0.2653	0.4941	1.2231
		1.3	0.0309	0.0313	0.2472	0.3339	0.3597	0.6699	1.6583
		1.4	0.0461	0.0467	0.3687	0.4980	0.5365	0.9991	2.4731
		1.5	0.0697	0.0706	0.5577	0.7534	0.8116	1.5113	3.7411
		1.6	0.1040	0.1054	0.8323	1.1244	1.2113	2.2557	5.5837
		1.7	0.1517	0.1536	1.2137	1.6397	1.7663	3.2892	8.1421
		1.8	0.2155	0.2183	1.7242	2.3294	2.5093	4.6727	11.5669
		1.9	0.2983	0.3021	2.3868	3.2244	3.4735	6.4683	16.0117
		2.0	0.4029	0.4081	3.2238	4.3552	4.6916	8.7366	21.6267
	$\hat{Q}_T(h)$		0.5465	0.7255	0.9398	0.9398	1.0491	1.6142	5.2121
($m = 2$)	$\hat{Q}_{T,m}(\tau, h)$		0.6802	10.9014	10.9529	11.1421	11.1435	15.2036	25.6498
	$\hat{Q}_{T,m}(\eta, h)$		0.1966	3.0187	3.1443	3.1964	3.3915	7.0577	9.0516

Para el tipo de cambio yen/dólar y retardos superiores al segundo se encuentra también cierta evidencia de dependencia no lineal no multiplicativa aunque de nuevo al 10% de significación. La conclusión general en el caso de las tres series podría considerarse la existencia de evidencia a favor de dependencia no lineal de tipo multiplicativo, es decir, evidencia de efectos GARCH significativos.

6. Conclusiones

En este trabajo revisamos los resultados más recientes relativos a las características y distribución asintótica de los procedimientos habituales de inferencia para detectar procesos no lineales, tanto de tipo aditivo como multiplicativo, basados en las funciones de autocovarianzas y de autocorrelación muestral. Las dificultades que pueden existir con la utilización de estos estadísticos, especialmente cuando los procesos subyacentes se caracterizan por distribuciones estacionarias de colas anchas tipo Pareto, pueden superarse haciendo uso de nuevas herramientas, como la función de covariación muestral. En este trabajo se extienden los resultados de Gallagher (2002, 2006) al caso de detectar dependencia tipo lineal pero en momentos superiores al de primer orden del proceso. La evidencia obtenida indica que incluso ante el incumplimiento de la condición de existencia de momento requerida para establecer la distribución asintótica de los estadísticos propuestos, estos presentan un comportamiento estable en muestras finitas y permiten discriminar entre ambos tipos de no linealidad.

Apéndice

Prueba del Teorema 1. Sea (a_1, ..., a_h) $\in \mathbb{R}^h$ y para toda secuencia ergódica e *iid*, X_t, con momento de orden *2m* finito, $E|X_t^m|^{2+\delta} \leq C$ ($\delta > 0$), se define

$$W_{t,m}(h) = X_t^m \sum_{k=1}^{h} a_k \delta_{t-k}$$

Considerando que $E(\delta_t) = 2\pi - 1$, con $\pi = P(\delta_t = 1)$ y $1-\pi = P(\delta_t = -1)$, se tiene que

$$E(W_{t,m}(h)) = \mu_m (2\pi - 1)\sum_{k=1}^{h} a_k \ ,$$

$$\gamma_m(0) = Var(W_{t,m}(h)) = \tilde{\mu}_{2,m}\left\{2\pi\sum_{k=1}^{h} a_k^2 + (2\pi-1)^2\left[\left(\sum_{k=1}^{h} a_k\right)^2 - \sum_{k=1}^{h} a_k^2\right]\right\}$$

y

$$\gamma_m(k) = Cov(W_{t,m}(h), W_{t+k,m}(h)) = 0 \quad k \geq 1$$

Se tiene entonces que por el TCL para procesos estacionarios *h*-dependientes (Hoeffding y Robbins (1948) y Berk (1973)), bajo el supuesto adicional de distribución simétrica ($\pi = 0.5$),

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} W_{t,m}(h) \Rightarrow N(0, \gamma_m(0))$$

con $\gamma_m(0) = \tilde{\mu}_{2,m}\sum_{k=1}^{h} a_k^2$. Por otro lado, el proceso

$$\tilde{W}_{t,m}(h) = \tilde{X}_{t,m}\sum_{k=1}^{h} a_k \delta_{t-k}, \quad \tilde{X}_{t,m} = X_t^m - \hat{\mu}_{m,T}$$

tiene el mismo comportamiento que $W_{t,m}(h)$, donde $E(\tilde{W}_{t,m}(h)) = o(1)$ incluso bajo distribución no simétrica bajo ergodicidad del proceso X_t. Si se define ahora el vector de dimensión *h*, $\tilde{\boldsymbol{\tau}}_{m,T}(h) = (\tilde{\tau}_{m,T}(1), ..., \tilde{\tau}_{m,T}(h))'$, con

$$\tilde{\tau}_{m,T}(k) = (1/T)\sum_{t=1}^{T-k} \tilde{W}_{t,m}(k),$$

y haciendo $a_i = 1$ para todo $i = 1, \ldots, k$, entonces

$$\sqrt{T}\,\tilde{\boldsymbol{\tau}}_{m,T}(h) \Rightarrow \mathbf{Z}_h \sqrt{\gamma_m}$$

donde $\mathbf{Z}_h$ es un vector de *h* variables *iid* normales estándar, y $\gamma_m = \tilde{\mu}_{2,m}$. El resultado final del Teorema 1 se sigue directamente de este resultado.

Prueba del Teorema 2. Sea (a_1, ..., a_h) $\in \mathbb{R}^h$ y para toda secuencia ergódica e *iid*, X_t, con momento de orden *2m* finito, $E|X_t^m|^{2+\delta} \leq C$ ($\delta > 0$), se define

$$Y_{t,m}(h) = \frac{X_t^m}{2}\sum_{k=1}^{h} a_k(\delta_{t-k} + \delta_{t+k})$$

Teniendo en cuenta que, en general, se tiene que $E(\delta_t) = 2\pi-1$, con $\pi = P(\delta_t = 1)$ y $1-\pi = P(\delta_t = -1)$, resulta que

$$E(Y_{t,m}(h)) = \mu_m(2\pi-1)\sum_{k=1}^{h} a_k\ ,$$

$$\gamma_m(0) = Var(Y_{t,m}(h)) = \frac{\tilde{\mu}_{2,m}}{2}\left\{2\pi\sum_{k=1}^{h} a_k^2 + (2\pi-1)^2\left(\sum_{k=1}^{h} a_k\right)^2\right\}$$

con $\tilde{\mu}_{2,m} = E((X_t^m - \mu_m)^2), \mu_m = E(X_t^m)$, y

$$\begin{aligned}\gamma_m(k) = Cov(Y_{t,m}(h), Y_{t+k,m}(h)) &= \frac{a_k^2}{4}(\lambda_m - \mu_m(2\pi-1))^2 \quad && k = 1,\ldots,h \\ &= 0 && k > h\end{aligned}$$

con $\lambda_m = E|X_t|^m$, de forma que el proceso $Y_{t,m}(h)$ es h-dependiente. Entonces, bajo simetría de la distribución del proceso X_t, de forma que π = 1/2, se tiene que $E(Y_{t,m}(h)) = 0$ y $\upsilon_m(h) = \gamma_m(0) + 2\sum_{k=1}^{h}\gamma_m(k) = \frac{\tilde{\mu}_{2,m} + \lambda_m^2}{2}\left(\sum_{k=1}^{h} a_k^2\right)$

con $0 < \upsilon_m(h) < \infty$. Entonces, siempre que $h^{2+2\delta}/T \to 0$ cuando $T \to \infty$ para algún $\delta > 0$, se tiene que por el TCL de Berk (1973) para procesos h-dependientes con h posiblemente no acotado,

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} Y_{t,m}(h) \Rightarrow N(0, \upsilon_m(h))$$

Si se define el proceso $\tilde{X}_{t,m} = X_t^m - \hat{\mu}_{m,T}$ y en lugar de $Y_{t,m}(h)$ se considera

$$\tilde{Y}_{t,m}(h) = \frac{\tilde{X}_{t,m}}{2}\sum_{k=1}^{h} a_k(\delta_{t-k} + \delta_{t+k})$$

resulta que, bajo ergodicidad y consistencia del estimador del momento de orden *m*,

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} \tilde{Y}_{t,m}(h) \Rightarrow N(0, \upsilon_m(h))$$

con $E(\tilde{Y}_{t,m}(h)) = o(1)$, incluso bajo no simetría de la distribución de X_t. Si se define ahora el vector de dimensión *h*, $\tilde{\mathbf{T}}_{m,T}(h) = (\tilde{T}_{m,T}(1),\ldots,\tilde{T}_{m,T}(h))'$, con

$\tilde{T}_{m,T}(k) = (1/T)\sum_{t=1}^{T-k}\tilde{Y}_{t,m}(k)$, y $a_i = 1$ para todo $i = 1, \ldots, k$, entonces

$$\sqrt{T}\,\tilde{\mathbf{T}}_{m,T}(h) \Rightarrow \mathbf{Z}_h \sqrt{\upsilon_m}$$

donde $\mathbf{Z}_h$ es un vector de h variables *iid* normales estándar, y $\upsilon_m = (\tilde{\mu}_{2,m} + \lambda_m^2)/2$. El resultado final se sigue directamente de este resultado.

Bibliografía

Alexander, C., E. Lazar (2004). Normal Mixture GARCH(1,1): Applications to Exchange Rate Modelling. ISMA Centre Finance Discussion Paper no.2004-06, University of Reading.

Andersen, T. (1994). Stochastic Autoregressive Volatility: A Framework for Volatility Clustering. *Journal of Mathematical Finance*, 4 (75–102).

Basrak, B., R.A. Davis, T. Mikosch (1999). The Sample ACF of a Simple Bilinear Process. *Stochastic Processes and Their Applications*, 83(1) (1-14).

Basrak, B., R.A. Davis, T. Mikosch (2002). Regular Variation of GARCH Processes. *Stochastic Processes and Their Applications*, 99(1) (95-115)

Berk, K.N. (1973). A Central Limit Theorem for *m*-Dependent Random Variables with Unbounded *m*. *The Annals of Probability*, 1(2) (352-354).

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity. *Journal of Econometrics*, 31 (307-327).

Bollerslev, T. (1987). A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. *Review of Economics and Statistics*, 69 (542-547).

Cambanis, S., G. Miller (1981). Linear Problemas in pth Order and Stable Processes. *SIAM Journal of Applied Mathematics*, 41(1) (43-69).

Cline, D.B.H. (2007). Regular variation of order 1 nonlinear AR-ARCH models. *Stochastic Processes and their Applications*, 117(7) (840-861).

Davis, R.A. (1983). Stable Limits for Partial Sums of Dependent Random Variables. *The Annals of Probability*, 11(2) (262-269).

Davis, R.A., S.I. Resnick (1985a). Limit theory for moving averages of random variables with regularly varying tail probabilities. *Annals of Probability*, 13 (175-195).

Davis, R.A., S.I. Resnick (1985b). More limit theory for the sample correlation function of moving averages. *Stochastic Processes and their Applications*, 20 (257-279).

Davis, R.A., S.I. Resnick (1986). Limit theory for the sample covariance and correlation function of moving averages. *Annals of Statistics*, 14 (533-558).

Davis, R.A., S.I. Resnick (1996). Limit Theory for Bilinear Processes with Heavy Tailed Noise. *Annals of Applied Probability*, 6(4) (1191-1210)

Davis, R.A., T. Mikosch (2000). The sample autocorrelations of financial time series models. En: Nonlinear and nonstationary signal processing (W.J. Fitzgerald, R.L.Smith, A.T. Wolden, P.Young, eds.). Cambridge University Press, Cambridge (247-274).

Davis, R.A., T. Mikosch (2001). Point process convergence of stochastic volatility processes with application to sample autocorrelation. *Journal of Applied Probability*, 38(A) (103-114).

Davis, R.A., T. Mikosch (2007). Extreme Value Theory for GARCH Processes. En: Andersen, T.G., R.A. Davis, J.P. Kreib, T. Mikosch (eds.), *Handbook of Financial Time Series*, Springer

Davis, R.A., T. Mikosch, B. Basrak (1999). Limit theory for the sample autocorrrelations of solutions to stochastic recurrence equations with applications to GARCH processes. Manuscrito.

Engle, R.F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50 (987-1007).

Forsberg, L., T. Bollerslev (2002). Bridging the Gap between the Distribution of Realized (ECU) Volatility and ARCH Modelling (of the EURO): The GARCH-NIG Model. *Journal of Applied Econometrics*, 17 (535-548).

Gallagher, C.M. (2001a). Estimating the Autocovariation Function with Applications to Heavy-Tailed ARMA Modeling. Department of Mathematical Sciences Technical Report.

Gallagher, C.M. (2001b). Estimating the Autocovariation from Stationary Heavy-Tailed Data, with Applications to Time Series Modeling. Department of Mathematical Sciences Technical Report GAL0101.

Gallagher, C.M. (2002). Testing for Linear Dependence in Heavy-Tailed Data. *Communications in Statistics: Theory and Methods*, 31(4) (611-623).

Gallagher, C.M. (2006). Detecting Dependence in Heavy-Tailed Time Series using Portmanteau-type Dependence Tests. International Mathematical Forum, 10(1) (455-469).

Gallagher, C.M., T. Okuyama (2003). Some Differences in the Rates of Convergence of the Sample Covariance and Covariation Functions. *Interstat*.

Hoeffding, W., H. Robbins (1948). The Central Limit Theorem for Dependent Random Variables. *Duke Mathematical Journal*, 15 (773-780).

Jensen, M. B., A. Lunde (2001). The NIG-S and ARCH model: a Fat Tailed, Stochastic, and Autoregressive Conditional Heteroscedastic Volatility Model. *Econometrics Journal*, 4 (319-342).

Kesten, H. (1973). Random difference equations and renewal theory for products of random matrices. *Acta Mathematica*, 131 (207-248).

Kiliç, R. (2007). Conditional Volatility and Distribution of Exchange Rates: GARCH and FIGARCH Models with NIG Distribution. *Studies in Nonlinear Dynamics and Econometrics*, 11(3).

Lindner, A.M. (2007). Stationarity, Mixing, Distributional Properties and Moments of GARCH(p,q). En: Andersen, T.G., R.A. Davis, J.P. Kreib, T. Mikosch (eds.), *Handbook of Financial Time Series*, Springer

McCulloch, J.H. (1986). Simple Consistent Estimators of Stable Distribution Parameters. *Communication in Statistics – Computation and Simulation*, 15(4) (1109-1136).

Mikosch, T., C. Starica (2000). Limit Theory for the Sample Autocorrelations and Extremes of a GARCH(1,1) Process. *The Annals of Statistics*, 28(5) (1427-1451)

Nelson, D.B. (1991). Conditional Heteroscedasticity in Asset Returns: A New Approach. *Econometrica*, 59 (347-370).

R.A. Davis, T. Mikosch (1998). The Sample Autocorrelations of Heavy-Tailed Processes with Applications to ARCH. *The Annals of Statistics*, 26(5) (2049-2080).

Resnick, S.I., E.van den Berg (2000). Sample Correlation Behavior for the Heavy Tailed General Bilineal Processes. *Stochastic Models*, 16(2) (233-258).

Runde, R. (1997). The Asymptotic Null Distribution of the Box-Pierce Q-Statistic for Random Variables with Infinite Variance. An Application to German Stock Returns. *Journal of Econometrics*, 78 (205-216).

Tong, H. (1990). Non-linear time series: A dynamical system approach. Oxford University Press, Oxford.

Yoon, G. (2003). A Simple Model that Generates Stylized Facts of Returns. UCSD Economics Working Paper no. 2003-04.

Yoon, G. (2006). A note on some properties of STUR processes. *Oxford Bulletin of Economics and Statistics*, 68(2) (253-260).

Yoon, G. (2007). Geometric ergodicity and regular variation of stochastic unit root processes. Brunel QASS Conference. Brunel University West London.