Modelos ARCH univariantes y multivariantes

Alfonso Novales
Departamento de Economía Cuantitativa
Universidad Complutense

Septiembre 2013
Versión preliminar
No citar sinpermiso del autor

Contents

1 Modelos ARCH

1.1 Propiedades estadísticas

Los modelos que hemos analizado hasta ahora mantenían el supuesto de que la innovación tiene una varianza constante en el tiempo, a pesar de que la esperanza condicional es cambiante. Sin embargo, para agentes aversos al riesgo, que toman sus decisiones en un régimen de incertidumbre, la varianza condicional, es decir, la varianza de la distribución de los rendimientos en cada instante futuro de tiempo, juega un papel de la mayor importancia. Este es el aspecto que modelizamos en este capítulo.

Los modelos *ARCH* aparecen en los años 80 con el objeto de recoger los episodios de agrupamiento temporal de volatilidad que suele observarse en las series de rentabilidad de casi todo mercado financiero. Desde entonces, su variedad y su aplicación práctica ha crecido de manera espectacular. En realidad, hay precursores más antiguos[Bachelier (1900) y Mandelbrot(1963,1967)], trabajos en los que comenzó a caracerizarse las propioedades estadísticas de los precios de activos financieros.

Las características más relevantes de las series financieras recogidas con frecuencias elevadas [Ruiz (1994)], son:

1) ausencia de estructura regular dinámica en la media, lo que aparece reflejado en estadísticos Ljung-Box generalmente no significativos,

2) distribuciones leptocúrticas o exceso de curtosis,

3) suelen ser simétricas, aunque también se encuentran en algunos casos coeficientes de asimetría significativamente distintos de cero,

4) agrupamiento de la volatilidad sobre intervalos de tiempo, lo cual se refleja en funciones de autocorrelación simple significativas para los cuadrados de las variables,

5) persistencia en volatilidad: los efectos de un shock en volatilidad tardan un tiempo en desaparecer.

6) efecto apalancamiento: se observa una respuesta asimétrica de la volatilidad al nivel de los rendimientos, en el sentido de que

Muchos de estos efectos quedan recogidos en los modelos *ARCH, GARCH, EGARCH*, que vamos a analizar. Estos modelos recogen en sus formulaciones la idea de que existen agrupaciones de volatilidad, es decir, que fuertes fluctuaciones inesperadas en los mercados tienden a venir seguidas de períodos de iguales características, mientras que períodos de estabilidad tienden a venir seguidos de períodos asimismo estables. Los modelos de esta familia recogen este comportamiento inercial en volatilidad a la vez que el comportamiento dinámico, con autocorrelación que suelen presentar las series financieras.

Una de las contribuciones importantes de la literatura de procesos *ARCH* es mostrar que las variaciones que aparentemente se producen en la volatilidad de las series temporales económicas pueden expliacrse mediante una determinada forma de dependencia no lineal, que permite además predecir dichos cambios en volatilidad sin necesidad de recurrir a la modelización explícita de cambios estructurales en la varianza.

Frente a estas observaciones empíricas, fórmulas de valoración del tipo Black-Scholes suponen una volatilidad constante para el precio del activo subyacente, que permite dudicr la expresión analítica del precio teórico de una opción Europea sobre dicho activo. La fórmula BS es utilizada habitualmente para deducir de ella la volatilidad implícita, forzando el precio que de ella se deriva a coincidri con el precio observado en el mercado. La volatilidad implícita así obtenida se interpreta como el nivel de volatilidad vigente en el mercado desde el momento de inversión de la fórmula BS hasta el vencimiento de la opción. Sin embargo, no siendo dicha volatilidad constante en el tiempo, es cuestionable la interpretación del valor numérico obtenido para la volatilidad implícita. Suele interpretarse como una expectativa de mercado y en términos del valor medio de volatildad vigente para el período mencionado, pero tal interpretación no está justificada por el análisis BS.

Para poder proceder al estudio empírico de este tipo de modelos de varianza condicional cambiante en datos reales, es preciso concretar antes la estructura de las funciones de esperanza y varianza condicionales, lo que pasamos a hacer a continuación.

1.2 Primeras definiciones y propiedades. Momentos incondicionales

Los momentos de los procesos $ARCH$ han sido analizados en Engel(1982), Milhoj(1985), Bollerslev(1986) entre muchos otros. Para su cálculo, es clave la ley de iteración de expectativas: dadas dos sigma-álgebras Ω_1, Ω_2, con $\Omega_1 \subseteq \Omega_2$ y una variable aleatoria escalar y, se tiene:

$$E\left(y \mid \Omega_1\right) = E\left[E\left(y \mid \Omega_2\right) \mid \Omega_1\right]$$

En nuestro caso, las dos sigma-álgebras son las generadas por la historia pasada de las variables del modelo, en dos instantes distintos de tiempo. un caso particular de esta ley que resulta especialmente útil es cuando $\Omega_1 = \phi$, pues entonces,

$$E\left(y\right) = E\left[E\left(y \mid \Omega_2\right)\right]$$

que relaciona un moomento incondicional y un momento condicional.

Sea $\{\varepsilon_t\left(\theta\right)\}$ un proceso estocástico, definido en tiempo discreto, cuyas esperanza y varianza condicionales dependen de un vector de parámetros θ, de dimensión m. Sea θ_0 el verdadero valor de dicho vector de parámetros. Inicialmente, consideramos que $\varepsilon_t\left(\theta\right)$ es escalar, aunque la generalización al caso multivariante es relativamente simple.

Denotamos por E_{t-1} la esperanza matemática condicional en la sigma-álgebra Ω_{t-1} generada por las realizaciones pasadas de las variables observables en el instante $t-1$ o anteriores, que define *el conjunto de información disponible en* $t-1$.

Definition 1 *Decimos que* $\{\varepsilon_t\left(\theta\right)\}$ *sigue un proceso ARCH si su esperanza condicional es igual a cero:*

$$E_{t-1}\varepsilon_t\left(\theta_0\right) = 0, \quad t = 1, 2, 3, ...$$

y su varianza condicional,

$$h_t^2\left(\theta_0\right) \equiv Var_{t-1}\left[\varepsilon_t\left(\theta_0\right)\right] = E_{t-1}\left[\varepsilon_t^2\left(\theta_0\right)\right] = g\left(\varepsilon_{t-1}, \varepsilon_{t-2}, ...\right)$$

depende, en forma no trivial, del sigma-álgebra Ω_{t-1} generada por las observaciones pasadas. La notación h_t^2 hace referencia al hecho de que trabajamos con un segundo momento del proceso estocástico. Debe apreciarse que, a pesar del subíndice temporal, h_t^2 es una función de variables pertenecientes al instante $t-1$ o anteriores.

La esperanza y varianza incondicionales del proceso $\varepsilon_t\left(\theta_0\right)$ son la esperanza matemática de los momentos análogos condicionales,

$$\begin{aligned} E(\varepsilon_t) &= E\left(E_{t-1}\varepsilon_t\right) = 0 \\ Var\left(\varepsilon_t\right) &= E\varepsilon_t^2 = E\left(E_{t-1}\varepsilon_t^2\right) = Eh_t^2 \end{aligned}$$

El proceso estandarizado:

$$z_t(\theta_0) = \frac{\varepsilon_t(\theta_0)}{\sqrt{h_t^2}}$$

tendrá esperanza condicional igual a cero, y varianza condicional igual a uno,

$$E_{t-1} z_t(\theta_0) = 0, \quad t = 1, 2, 3, ...$$

$$Var_{t-1}[z_t(\theta_0)] = Var_{t-1}\left[\frac{\varepsilon_t(\theta_0)}{\sqrt{h_t^2}}\right] = \frac{1}{h_t^2} Var_{t-1}[\varepsilon_t(\theta_0)] = 1, \quad t = 1, 2, 3, ...$$

Sus momentos incondicionales serán, por tanto, iguales a los momentos condicionales, que son constantes:

$$E(\varepsilon_t) = E(E_{t-1}\varepsilon_t) = 0; \;\; Var[z_t(\theta_0)] = 1, \quad t = 1, 2, 3, ...$$

Por tanto, si bien la varianza condicional cambia en el tiempo, la varianza incondicional es constante, por lo que el proceso $ARCH$ es incondicionalmente homocedástico. Hay que notar, además, que la variable aleatoria $z_t(\theta_0)$ es independiente del pasado de $\varepsilon_t(\theta_0)$, pues la presencia de $\sqrt{h_t^2}$ en su definición no hace sino reducir su varianza[1] a 1.

Si añadimos el supuesto de Normalidad condicional para ε_t, y suponemos que la distribución condicional de $z_t(\theta_0)$ tiene momento de cuarto orden finito, se tendrá, por la desigualdad[2] de Jensen:

$$E\left[\varepsilon_t^4(\theta_0)\right] = E\left[z_t^4(\theta_0)\right] E\left[h_t^4(\theta_0)\right] \geq E\left[z_t^4(\theta_0)\right]\left[E\left(h_t^2(\theta_0)\right)\right]^2 = E\left[z_t^4(\theta_0)\right]\left[E\left(\varepsilon_t^2(\theta_0)\right)\right]^2$$

Por tanto, el coeficiente de curtosis del proceso $ARCH$ $\varepsilon_t(\theta)$ será,

$$\frac{E\left(\varepsilon_t^4\right)}{\left[E\left(\varepsilon_t^2\right)\right]^2} \geq E\left(z_t^4\right) = 3$$

y la desigualdad se cumplirá como igualdad sólo en el caso de una varianza condicional constante. En caso contrario, si la distribución de $z_t(\theta_0)$ es Normal, entonces la distribución incondicional de ε_t será leptocúrtica.

Por otra parte, si la distribución condicional de ε_t es Normal, se tiene para todo entero impar m que $E(\varepsilon_t^m(\theta_0)) = E[E_{t-1}(\varepsilon_t^m(\theta_0))] = E(0) = 0$, por lo que el coeficiente de asimetría de ε_t es nulo. Al ser ε_t una variable aleatoria continua, esto implica que su densidad es simétrica.

[1] Alternativamente, podríamos definir el proceso $ARCH$ mediante,

$$y_t = \varepsilon_t h_t$$

con $\varepsilon_t \sim N(0,1)$,independiente en el tiempo, y $h_t^2(\theta_0) \equiv g(\varepsilon_{t-1}, \varepsilon_{t-2}, ...)$. Con esta notación, la afirmación del texto equivaldría a decir que ε_t y h_t son independientes, como claramente ocurre.

[2] El lector puede comprobar la facilidad con que obtiene este resultado utilizando la notación propuesta en el pie de página previo.

1.2.1 Procesos con residuos ARCH

Aunque nos centremos en las propiedades del proceso $\{\varepsilon_t(\theta)\}$, en general, tendremos un proceso $\{y_t(\theta_0)\}$, objeto de estudio, cuya esperanza condicional será una función de θ_0,

$$E_{t-1}y_t = \mu_{t-1}(\theta_0)$$

En general, entendemos que y_t representa el rendimiento ofrecido por un activo financiero, cuyo valor actual descomponemos mediante una identidad, en dos componentes: a) el componente anticipado, $\mu_{t-1}(\theta_0)$, que pudimos haber previsto en base a información pasada, y b) la innovación en el proceso de rentabilidad. Es ésta última la que se supone que tiene una estructura de tipo *ARCH*.

Denotemos por $\{\varepsilon_t(\theta_0)\}$ el residuo de dicha relación, o error de predicción un período hacia adelante,

$$y_t = \mu_{t-1}(\theta_0) + \varepsilon_t(\theta_0) \Rightarrow \varepsilon_t(\theta_0) = y_t - \mu_{t-1}(\theta_0)$$

que satisface,

$$E_{t-1}[\varepsilon_t(\theta_0)] = 0$$

y supongamos que tiene la estructura *ARCH* definida en (**??**).

Para el proceso y_t tendremos,

$$\begin{aligned} E_{t-1}y_t &= \mu_{t-1}(\theta_0), \\ Var_{t-1}y_t &= E_{t-1}[y_t - E_{t-1}y_t]^2 = E_{t-1}[\varepsilon_t(\theta_0)]^2 = Var_{t-1}[\varepsilon_t(\theta_0)] = h_t^2 \end{aligned}$$

por lo que su varianza condicional coincide con la de $\varepsilon_t(\theta_0)$, mientras que su varianza incondicional es,

$$Var(y_t) = E\left(h_t^2\right)$$

En consecuencia, mientras que los momentos incondicionales son constantes en el tiempo aunque, como veremos, pueden no existir, los momentos incondicionales cambian a lo largo del tiempo. Un modelo *ARCH* consta de: a) una ecuación representando el modo en que la esperanza condicional del proceso varía en el tiempo, b) una ecuación mostrando el modo en que su varianza condicional cambia en el tiempo, y c) una hipótesis acerca de la distribución que sigue la innovación de la ecuación que describe el proceso seguido por su esperanza matemática.

1.3 El modelo ARCH(q)

La estructura básica de este modelo es,

$$
\begin{aligned}
y_t &= \varepsilon_t h_t \\
h_t^2 &= \alpha_0 + \sum_{i=1}^{q} \alpha_i y_{t-i}^2, \quad \alpha_0 > 0, \alpha_i \geq 0, \sum_{i=1}^{q} \alpha_i \leq 1
\end{aligned}
$$

donde, una vez más, suponemos que ε_t es un proceso ruido blanco, con $E(\varepsilon_t) = 0, Var(\varepsilon_t) = 1$. Por simplicidad, estamos suponiendo asimismo que la variable y_t carece de autocorrelación, así como de la imposibilidad de utilizar otras variables que puedan explicar su evolución temporal. En todo caso, la escasa estructura dinámica que se observa en datos frecuentes de rentabilidades de mercados financieros justifica la simplicidad en la especificación de la ecuación de la media del proceso y_t. Alternativamente, si el investigador detecta algunas variables que pueden explicar el comportamiento de y_t, posiblemente incluyendo algunos retardos de la propia variable, entonces h_t^2 sería la varianza condicional del término de error del modelo que explica el comportamiento de y_t. Esto es lo que haremos en algunos modelos analizados en las próximas secciones.

Aunque no son necesarias, las restricciones de signo de los coeficientes de la ecuación de varianza garantizan que la varianza condicional será positiva en todos los períodos. En realidad, lo que necesitamos es que, una vez que el modelo haya sido estimado, genere una serie de varianzas positiva, lo cual es compatible con que alguno de los coeficientes α_i sean negativos. Esto debe tomarse como un contraste de validez del modelo, que no sería aceptable si generase varianzas estimadas negativas. Es preferible no imponer las restricciones en la estimación del modelo, y poder contrastar la propiedad del modo que hemos descrito, que estimar bajo las restricciones de signo.

La restricción sobre la suma de los coeficientes de la ecuación de varianza garantiza que el proceso sea estacionario en varianza. Para ello, es necesario que las raíces del polinomio característico,

$$
\alpha_0 - \alpha_1 z - \alpha_2 z^2 - ... - \alpha_q z^q = 0
$$

estén fuera del círculo unidad, es decir, tengan valor absoluto mayor que uno o, si son complejas, módulo mayor que la unidad. Cuando son no-negativas, ello es equivalente a la condición sobre su suma.

De acuerdo con este modelo, una sorpresa en y_t importante en magnitud, positiva o negativa, hará que la varianza del proceso sea elevada durante un cierto número de períodos.

Si, condicional en Ω_{t-1}, ε_t sigue una distribución Normal, la distribución condicional de y_t será asimismo Normal, pues $y_t = \varepsilon_t \sqrt{\alpha_0 + \sum_{i=1}^{q} \alpha_i y_{t-i}^2}$ y el componente dentro de la raíz es conocido en $t-1$. Se tiene, además,

$$
Var\left(y_t / y_{t-1}, y_{t-2}, ...\right) = E\left(y_t^2 / y_{t-1}, y_{t-2}, ...\right) = \alpha_0 + \sum_{i=1}^{q} \alpha_i y_{t-i}^2
$$

Por el contrario, su distribución incondicional no es fácilmente caracterizable, debido a la nolinealidad de la relación entre y_t y ε_t. De hecho al no seguir incondicionalmente una distribución Normal, no se tiene la equivalencia entre ausencia de correlación e independencia, como veremos en detalle en el caso del modelo $ARCH(1)$. Es fácil probar, sin embargo, que sus momentos de orden impar son todos igual a cero, por lo que dicha distribución es simétrica.

Definiendo $v_t = y_t^2 - h_t^2$, que cumple $E(v_t) = 0$, este proceso puede escribirse,

$$y_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i y_{t-i}^2 + v_t$$

por lo que pasamos a tener un proceso *AR(q)* en el cuadrado de la variable a explicar, y_t, que podrían ser los rendimientos que ofrece un determinado activo financiero. Esta es otra interpretación del modelo $ARCH(q)$ cuando no hay estructura de variables explicativas en la ecuación de la media del proceso.

La varianza incondicional de este proceso es,

$$\sigma_y^2 = Var\ y_t = \frac{\alpha_0}{1 - \sum_{i=1}^{q} \alpha_i}$$

Aunque los sucesivos valores de y_t están incorrelacionados, no son independientes, debido a la relación que existe entre sus segundos momentos.

Para evitar trabajar con un elevado número de parámetros en ocasiones en que se percibe una alta persistencia en volatilidad, suele utilizarse una representación,

$$h_t^2 = \alpha_0 + \alpha_1 \sum_{i=1}^{q} w_i y_{t-i}^2,\ w_i = \frac{(q+1) - i}{\frac{1}{2} q (q+1)},\ \sum_{i=1}^{q} w_i = 1$$

Este es el modelo $ARCH(q)$ restringido, introducido ya por Engle(1982). En todo caso, la estructura lineal en los coeficientes de los retardos en la ecuación de la varianza pude contrastarse, frente a la alternativa formada por una estructura libre de coeficientes, siguiendo los métodos que describiremos más adelante.

Taylor (1986) prueba que la función de autocorrelación simple de y_t^2 cuando el proceso y_t tiene una estructura $ARCH(q)$ presenta la misma configuración que la función de autocorrelación simple de un proceso $AR(q)$, lo que puede servir para detectar este tipo de estructura.

1.3.1 El modelo ARCH(1)

Un caso especialmente interesante surge cuando $q = 1$, teniendo el modelo $ARCH(1)$, que puede escribirse:

$$y_t = \varepsilon_t h_t = \varepsilon_t \sqrt{\alpha_0 + \alpha_1 y_{t-1}^2}$$

siendo ε_t un proceso ruido blanco con varianza igual a 1.

Su esperanza y varianza condicionales son,

$$E_{t-1}y_t = \left(\sqrt{\alpha_0 + \alpha_1 y_{t-1}^2}\right) E_{t-1}(\varepsilon_t) = 0$$

$$Var_{t-1}(y_t) = E_{t-1}y_t^2 = \left(\alpha_0 + \alpha_1 y_{t-1}^2\right) E_{t-1}(\varepsilon_t^2) = \alpha_0 + \alpha_1 y_{t-1}^2$$

por lo que la varianza condicional varía, en función de la realización del proceso y_t.

La ley de iteración de expectativas nos dice, $E\left(\varepsilon_t\right) = E\left[E\left(\varepsilon_t \mid I_{t-1}\right)\right]$, pero como la especificación del modelo incluye el supuesto $E\left(\varepsilon_t \mid \Omega_{t-1}\right) = 0$, se tiene que $E\left(\varepsilon_t\right) = 0$, lo cual es cierto para todo modelo $ARCH(q)$.

Al ser independiente en el tiempo, ε_t también es independiente de valores pasados de y_t, por lo que la esperanza y varianza marginal o incondicional de y_t son:

$$E\left(y_t\right) = E\left(\varepsilon_t \sqrt{\alpha_0 + \alpha_1 y_{t-1}^2}\right) = E\left(\varepsilon_t\right) E\left(\sqrt{\alpha_0 + \alpha_1 y_{t-1}^2}\right) = 0$$

$$Var\left(y_t\right) = E\left(y_t^2\right) = E\left(\varepsilon_t^2\right) E\left(\alpha_0 + \alpha_1 y_{t-1}^2\right) = \alpha_0 + \alpha_1 E\left(y_{t-1}^2\right)$$

donde hemos utilizado nuevamente la independencia estadística de ε_t e y_{t-1}.

Si $\mid \alpha_1 \mid < 1$, el proceso y_t es estacionario, con $E\left(y_t^2\right) = E\left(y_{t-1}^2\right)$, lo que implica que,

$$Var\left(y_t\right) = \frac{\alpha_0}{1-\alpha_1}$$

que, a diferencia de lo que ocurre con la varianza incondicional, es constante en el tiempo.

La autocovarianza de orden τ, $\tau \geq 1$ del proceso $ARCH(1)$ es:

$$E\left(y_t y_{t-\tau}\right) = E\left[\varepsilon_t \sqrt{\alpha_0 + \alpha_1 y_{t-1}^2}\ y_{t-\tau}\right] = E\left(\varepsilon_t\right) E\left[\sqrt{\alpha_0 + \alpha_1 y_{t-1}^2}\ y_{t-\tau}\right] = 0$$

por lo que el proceso $ARCH(1)$ no está autocorrelacionado, es decir, no existen relaciones lineales entre sus valores en distintos instantes de tiempo.

Sin embargo, su cuadrado, y_t^2, sí está autocorrelacionado. Por ejemplo, su autocovarianza de orden 1 es:

$$\begin{aligned}
\gamma_1\left(y_t^2\right) &= E\left(y_t^2 y_{t-1}^2\right) = E\left[\left(y_t^2 - \frac{\alpha_0}{1-\alpha_1}\right)\left(y_{t-1}^2 - \frac{\alpha_0}{1-\alpha_1}\right)\right] = \\
&= -\frac{\alpha_0}{1-\alpha_1}\left(Ey_t^2 + Ey_{t-1}^2\right) + \left(\frac{\alpha_0}{1-\alpha_1}\right)^2 + E\left(y_t^2 y_{t-1}^2\right)
\end{aligned}$$

pero: $y_t^2 = \varepsilon_t^2\left(\alpha_0 + \alpha_1 y_{t-1}^2\right)$, y ya hemos visto que: $Ey_t^2 = \alpha_0 + \alpha_1 E\left(y_{t-1}^2\right)$, por lo que,

$$\begin{aligned}
\gamma_1\left(y_t^2\right) &= -\frac{\alpha_0}{1-\alpha_1} 2 \frac{\alpha_0}{1-\alpha_1}+\left(\frac{\alpha_0}{1-\alpha_1}\right)^2+E\left[\varepsilon_t^2\left(\alpha_0+\alpha_1 y_{t-1}^2\right) y_{t-1}^2\right]= \\
&= -\left(\frac{\alpha_0}{1-\alpha_1}\right)^2+\left[\alpha_0 E\left(y_{t-1}^2\right)+\alpha_1 E\left(y_{t-1}^4\right)\right]
\end{aligned}$$

y de, hecho, puede probarse [Taylor (1986)] que la función de autocorrelación simple del cuadrado de un proceso $ARCH(q)$ tiene las mismas características que la función de autocorrelación simple de un proceso $AR(q)$.

Por otra parte, podemos repetir en este caso particular el análisis que hicimos antes para el caso general, acerca del momento de cuarto orden. La condición necesaria para la existencia del cuarto momento del proceso $ARCH(1)$ es $3\alpha_1^2 < 1$.Bajo este supuesto, y añadiendo la hipótesis de Normalidad de ε_t, tenemos,

$$E(y_t^4)=E\left[\varepsilon_t^4\left(\alpha_0+\alpha_1 y_{t-1}^2\right)^2\right]=3\frac{\alpha_0^2\left(1+\alpha_1\right)}{1-3\alpha_1^2}$$

por lo que la autocovarianza de orden 1 de y_t^2 es:

$$\gamma_1\left(y_t^2\right)=\frac{2\alpha_0^2\alpha_1}{\left(1-\alpha_1\right)^2\left(1-3\alpha_1^2\right)}$$

que es no nula.

Bajo estos supuestos, la curtosis del proceso $ARCH(1)$ es finita, e igual a:

$$Curtosis(y_t)=\frac{E\left(y_t^4\right)}{E\left(y_t^2\right)^2}=3\frac{1-\alpha_1^2}{1-3\alpha_1^2}$$

siendo igual a infinito en caso contrario. Si $\alpha_1 > 0$, entonces la curtosis es mayor que 3 y, por tanto, mayor que la de la distribución $N(0,1)$, por lo que el proceso $ARCH$ tiene colas más gruesas que dicha distribución. Esta es una propiedad conocida de las series financieras. Por otra parte, que el modelo $ARCH$ no imponga necesariamente una varianza finita es deseable en el sentido de que esta debe ser una propiedad del verdadero proceso generador de datos que aparezca en los resultados de la estimación.

La ausencia de autocorrelación del proceso $ARCH$ le hace deseable para la modelización de series temporales financieras. La hipótesis de mercados eficientes se describe en ocasiones como la incapacidad de predecir rentabilidades futuras a partir de rentabilidades pasadas. Si una rentabilidad r_t es un proceso $ARCH$ puro (es decir, sin variables explicativas), entonces se tiene $E\left(r_t \mid I_{t-1}\right)=E\left(r_t\right)=0$. Por tanto, la existencia de efectos $ARCH$ no contradice esta versión de la hipótesis de mercados eficientes.

La presencia de efectos $ARCH$ no afecta, teóricamente, a la predicción de valores futuros del proceso, aunque se gana eficiencia y se obtienen estimaciones puntuales distintas, una vez que se modelizan estos efectos. En cualquier caso, el potencial de un modelo $ARCH$ estriba en que proporciona una medida de riesgo cambiante en el tiempo, que puede ser un input importante en otro tipo de análisis, como por ejemplo, si se quiere cuantificar la remuneración que en un determinado mercado se ofrece al riesgo que se asume en el mismo.

Sin embargo, este modelo es susceptible de provocar algunos problemas de signo. En particular, los valores numéricos de la serie temporal de volatilidad h_t^2 que resultan del proceso de estimación, deben ser todos positivos.

Como el proceso $ARCH$ carece de autocorrelación y tiene media cero, es débilmente estacionario si existe su varianza. Una propiedad notable de este proceso es que puede no ser débilmente estacionario (porque su varianza no exista) y, sin embargo, ser estrictamente (o fuertemente) estacionario pues para este último concepto no es precisa la existencia de momentos.

1.3.2 Modelo AR(1)-ARCH(1)

Comencemos recordando el modelo $AR(1)$ sin perturbaciones $ARCH$,

$$y_t = \phi y_{t-1} + \varepsilon_t, \quad |\, \phi \,| < 1$$

siendo ε_t un proceso ruido blanco, con $E(\varepsilon_t) = 0, Var(\varepsilon_t) = \sigma_\varepsilon^2$. En este modelo se tienen momentos condicionales,

$$\begin{aligned} E_{t-1} y_t &= \phi y_{t-1}, \\ Var_{t-1} y_t &= Var_{t-1} \varepsilon_t = \sigma_\varepsilon^2 \end{aligned}$$

mientras que los momentos incondicionales son,

$$\begin{aligned} E y_t &= 0, \\ Var(y_t) &= \frac{\sigma_\varepsilon^2}{1-\phi^2} \end{aligned}$$

Como puede verse, la expresión de la esperanza condicional recoge el hecho de que es posible prever este proceso si se dispone de sus valores pasados. Como consecuencia, la varianza condicional es *inferior* a la varianza incondicional.

Más generalmente, el modelo $AR(1)$ con perturbación $ARCH(1)$ es,

$$\begin{aligned} y_t &= \phi y_{t-1} + \varepsilon_t, \quad |\, \phi \,| < 1 \\ E_{t-1} \varepsilon_t &= 0, \quad Var_{t-1} \varepsilon_t = h_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \end{aligned}$$

El supuesto $|\, \phi \,| < 1$ garantiza que el proceso es estacionario en media. Su varianza será positiva en todos los períodos si restringimos los valores de los parámetros mediante $\alpha_0 \geq 0, \alpha_1 \geq 0$.

La esperanza y varianza condicionales de y_t son,

$$
\begin{aligned}
E_{t-1}y_t &= \phi y_{t-1}, \\
Var_{t-1}y_t &= Var_{t-1}\varepsilon_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 = \alpha_0 + \alpha_1\left(y_t - \phi y_{t-1}\right)^2
\end{aligned}
$$

La varianza incondicional es finita si $\alpha_1 < 1$, y los momentos incondicionales, son entonces,

$$
\begin{aligned}
E\ y_t &= 0, \\
Var\ y_t &= \frac{\alpha_0}{1-\alpha_1^2}
\end{aligned}
$$

La varianza condicional puede escribirse,

$$
h_t^2 - \sigma^2 = \alpha_1\left(\varepsilon_{t-1}^2 - \sigma^2\right)
$$

de modo que la varianza condicional excede de la varianza incondicional siempre que la inovación (o sorpresa) al cuadra, es mayor que su esperanza incondicional, σ^2.

Aunque las innovaciones están incorrelacionadas a través del tiempo, no son independientes, puesto que están relacionadas a través de sus momentos de orden 2. Aunque y_t sigue una distribución condicional Normal, su distribución conjunta con valores en otros instantes de tiempo, no lo es. Tampoco su distribución de probabilidad incondicional o marginal es Normal, si bien será simétrica, si la distribución de probabilidad condicional de ε_t lo es. Si ε_t tiene ua distribución condicional Normal, entonces su cuarto momento incondicional excederá de $3\sigma^4$, por lo que la distribución marginal de ε_t tendrá colas más gruesas que la Normal. Su momento de orden cuatro será finito siempre que $3\alpha_1^2 < 1$.

En muchas aplicaciones empíricas, el orden del modelo ARCH que es preciso utilizar para recoger la dependencia temporal en la varianza es elevado, por lo que es útil considerar una representación más simple de este tipo de estructuras:

1.3.3 Modelos ARMA-ARCH

El análisis anterior puede generalizarse a cualquier modelo univariante de series temporales de la familia $ARIMA$, en el que puede tener perfecto sentido epecificar que la varianza del término de error es vcambiante en el tiempo. Por ejemplo, el modelo $AR(p) - ARCH(m)$ es,

$$
\begin{aligned}
y_t &= \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \varepsilon_t, \quad \mid \phi \mid < 1 \\
E_{t-1}\varepsilon_t &= 0, \quad Var_{t-1}\varepsilon_t = h_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \alpha_2\varepsilon_{t-2}^2 ... + \alpha_m\varepsilon_{t-m}^2
\end{aligned}
$$

1.3.4 El modelo ARCH(q) de regresión

Consideremos un modelo dinámico de regresión lineal,

$$y_t = x_t'\beta + \varepsilon_t, \quad t = 1, 2, ..., T$$

siendo x_t un vector $kx1$ de variables explicativas que pueden incluir retardos de la variable dependiente. El modelo $ARCH$ de regresión especifica, condicional en las observaciones pasadas de la variable dependiente y de las variables explicativas, el término de error del modelo anterior se distribuye,

$$\varepsilon_t \mid \Omega_{t-1} \sim N(0, h_t^2)$$

siendo,

$$h_t^2 = \delta_0 + \delta_1 \varepsilon_{t-1}^2 + ... + \delta_q \varepsilon_{t-q}^2$$

con $\alpha_i \geq 0, \sum_{i=1}^{q} \alpha_i \leq 1$, para asegurar que la varianza resultante sea positiva en todos los períodos. Como $\varepsilon_{t-1} = y_{t-1} - x_{t-1}'\beta$, se tiene que h_t^2 es una función de la información contenida en Ω_{t-1}. Una vez más, al ser la varianza condicional del período t una función creciente de la magnitud de las últimas innovaciones, se produce el *clustering* o agrupamiento temporal de volatilidades. El orden q de la representación ARCH es un indicador de la persistencia de los shocks en varianza.

En muchas aplicaciones empíricas, el orden del modelo ARCH que es preciso utilizar para recoger la dependencia temporal en la varianza es elevado, por lo que es útil considerar una representación más simple de este tipo de estructuras:

1.4 Modelos GARCH

Aunque en todos los modelos que se describen a continuación suponemos que las rentabilidades siguen una estructura ARCH/GARCH, es sobre su innovación sobre la que aplicaremos en la práctica el supuesto. Unicamente en el caso en que la rentabilidad carezca de autocorrelación, podemos hacer el supuesto GARCH sobre la rentabilidad pues, salvo constantes, coincide con su innovación.

1.4.1 Modelos GARCH(p,q)

En muchos casos, la especificación $ARCH$ que recoge la estructura de autocorrelación en varianza precisa de un elevado número de retardos. Para evitar que el alto número de coefcientes en términos autoregresivos, generalmente bastante relacionados, produzca una importante pérdida de precisión en su estimación, se ha propuesto una parametrización alternativa, restringida, dependiente de un número reducido de parámetros. El modelo $GARCH(p, q)$ de Bollerslev (1986) es,

$$
\begin{aligned}
y_t &= \varepsilon_t h_t \\
h_t^2 &= \alpha_0 + \sum_{i=1}^{q} \alpha_i y_{t-i}^2 + \sum_{i=1}^{p} \beta_j h_{t-j}^2, \quad \alpha_0 > 0, \alpha_i, \beta_j \geq 0, \sum_{i=1}^{q} \alpha_i + \sum_{i=1}^{p} \beta_j < 1
\end{aligned}
$$

Las condiciones anteriores garantizan (si bien no son necesarias) que la varianza condicional estimada sea positiva en todos los períodos. En realidad, Nelson y Cao (1992), mostraron condiciones más débiles que garantizan varianza positiva en todos los períodos. por ejemplo, en un $GARCH(1,2)$, es suficiente que: $\alpha_0 > 0, \alpha_1 \geq 0, \beta_1 \geq 0, \beta_1\alpha_1 + \alpha_2 \geq 0$. Este modelo puede transformarse en un modelo $ARCH$ de orden infinito [Bera y Higgins, Volatility], restringido en sus parámetros.

En la especificación anterior hemos supuesto, nuevamente por simplicidad, que y_t carece de autocorrelación, así como que no disponemos de variables epclicativas para le esperanza condicional de dicho proceso. Los mismos comentarios que hicimos acerca del modelo $ARCH(q)$ aplican a este caso.

La esperanza matemática del proceso $GARCH(p,q)$ es cero, y su varianza,

$$
Var\ y_t = \frac{\alpha_0}{1 - \sum_{i=1}^{q} \alpha_i - \sum_{i=1}^{p} \beta_j} = \frac{\alpha_0}{1 - \alpha(1) - \beta(1)}
$$

y la distribución es nuevamente leptocúrtica e incondicionalmente homocedástica.

Con la misma definición de la inovación que antes hicimos, tenemos, $h_t^2 = y_t^2 - v_t$, y el proceso $GARCH(p,q)$ puede escribirse,

$$
y_t^2 = \alpha_0 + \left(\alpha\left(L\right) + \beta\left(L\right)\right) y_{t-1}^2 - \beta\left(L\right) v_{t-1} + v_t
$$

o, lo que es lo mismo,

$$
\left(1 - \alpha\left(L\right) - \beta\left(L\right)\right) y_t^2 = \alpha_0 + \left(1 - \beta\left(L\right)\right) v_t
$$

y es necesario que todas las raíces del polinomio $1 - \alpha\left(L\right) - \beta\left(L\right)$ estén fuera del círculo unidad para que el proceso sea estacionario. En tal caso, su varianza incondicional será finita, y estará dada por la expresión anterior.

El modelo $GARCH(p,q)$ puede escribirse,

$$
y_t^2 = \alpha_0 + \sum_{i=1}^{r} \left(\alpha_i + \beta_j\right) y_{t-i}^2 - \sum_{i=1}^{r} \beta_j \left(y_{t-j}^2 - h_{t-j}^2\right) + \left(y_t^2 - h_t^2\right),
$$

$$
\alpha_0 > 0, \alpha_i, \beta_j \geq 0, \sum_{i=1}^{q} \alpha_i + \sum_{i=1}^{p} \beta_j < 1
$$

siendo $r = \max(p,q)$. Nuevamente, $E\left(y_t^2 - h_t^2\right) = 0$, por lo que puede considerarse como la innovación en la ecuación anterior. En consecuencia, un modelo $GARCH(p,q)$ para la rentabilidad y_t puede interpretarse como un modelo

$ARMA$ para y_t^2. Aunque su estimación como tal proceso $ARMA$ sería ineficiente, sin embargo las expresiones habituales para la predicción en modelos $ARMA$ son utilizables.

Examinemos el cálculo del a varianza incondicional en el caso del proceso $GARCH(1,1)$,

$$\begin{aligned} Var(\varepsilon_t) &= E\left(\varepsilon_t^2\right) = E\left[E\left(\varepsilon_t^2 \mid \Omega_{t-1}\right)\right] = E(h_t^2) = \\ &= \alpha_0 + \alpha_1 E\left(\varepsilon_{t-1}^2\right) + \beta_1 E\left(h_{t-1}^2\right) = \alpha_0 + (\alpha_1 + \beta_1) E\left(\varepsilon_{t-1}^2\right) \end{aligned}$$

que es una ecuación en diferencias en $E\left(\varepsilon_t^2\right)$ que si converge, tiene como límite,

$$Var\ y_t = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}$$

lo cual ocurre siempre y cuando $\alpha_1 + \beta_1 < 1$. En el caso general, la condición necesaria y suficiente de existencia de la varianza incondicional es $\alpha(1) + \beta(1) = \sum_{i=1}^{q} \alpha_i - \sum_{i=1}^{p} \beta_j < 1$. Bollerslev(1986) proporciona condiciones analíticas sobre los parámetros del modelo para garantizar la existencia de momentos en un proceso $GARCH(p,q)$.

Si y_t sigue un proceso $GARCH(p,q)$, su cuadrado, y_t^2 tiene una función de autocorrelación simple análoga a la de un proceso $ARMA(p^*,q)$, con $p^* = \max\{p,q\}$, parámetros autoregresivos $\phi_i = \alpha_i + \beta_i$, y parámetros de media móvil, $\theta_j = -\beta_j$, para $j = 1,2,...,q$. Precisamente esta similitud con los modelos $ARMA$ hace que se utilicen técnicas de identificación para los modelos $ARCH$ y $GARCH$ basadas en las funciones de autocorrelación simple y parcial, del mismo modo que se hace en el análisis del tipo Box-Jenkins, pero esta vez utilizando los cuadrados de los residuos. Sin embargo, la dependencia estadística de los procesos de varianza condicional hace que la estimación de dichas funciones sea poco eficiente.

El modelo más habitual dentro de esta clase es el $GARCH(1,1)$:

1.4.2 El modelo GARCH(1,1)

Este es un modelo de suavizado exponencial de la varianza, análogo a los que consideramos para la volatilidad condicional,

$$\begin{aligned} y_t &= \varepsilon_t h_t \\ h_t^2 &= \omega + \alpha y_{t-1}^2 + \beta h_{t-1}^2 \end{aligned}$$

con $\alpha > 0, \omega > 0, \beta \geq 0, \alpha + \beta < 1$.

En este modelo, la varianza condicional es,

$$Var_{t-1} y_t = h_t^2$$

mientras que la varianza incondicional o de *largo plazo*, es:

$$Var\ y_t = \frac{\omega}{1-\alpha-\beta}$$

Los retardos medio y mediano en h_t^2 son,

$$\begin{aligned} \mathrm{Re}\,tardo\ medio &= \frac{\sum_{i=1}^{\infty} i\delta_i}{\sum_{i=1}^{\infty} \delta_i} = \frac{1}{1-\beta} \\ \mathrm{Re}\,tardomediano &= -\frac{\ln 2}{\ln \beta} \end{aligned}$$

El modelo $GARCH(1,1)$ puede escribirse:

$$y_t^2 = \omega + (\alpha+\beta)\, y_{t-1}^2 - \beta\left(y_{t-1}^2 - h_{t-1}^2\right) + \left(y_t^2 - h_t^2\right)$$

donde los dos últimos términos tienen esperanza condicional igual a cero, por lo que este modelo es, en muchos aspectos, similar al modelo $ARMA(1,1)$. De hecho, el modo de identificar una estructura $GARCH(1,1)$ es porque las funciones de autocorrelación simple y parcial de los cuadrados de y_t tengan el aspecto de las funciones correspondientes a un proceso $ARMA(1,1)$. La función de autocorrelación simple del proceso $GARCH(1,1)$ es:

$$\begin{aligned} \rho(1) &= \alpha \frac{1-\alpha\beta-\beta^2}{1-2\alpha\beta-\beta^2} \\ \rho(k) &= (\alpha+\beta)^{k-1} \rho(1)\,,\ k>1 \end{aligned}$$

Como ocurría con el modelo $ARCH(q)$, aunque la distribución condicional de este proceso es Normal cuando lo es la innovación ε_t, su distribución incondicional no es conocida. Sabemos, sin embargo, que su esperanza es cero y su varianza viene dada por la expresión anterior. Es fácil probar que sus momentos impares son nulos y, por tanto, la distribución es simétrica. Además, es leptocúrtica.

Si ε_t es Normal y $\beta^2 + 2\alpha\beta + 3\alpha^2 < 1$, entonces su coeficiente de curtosis es:

$$Curtosis(y_t) = 3 + \frac{6\alpha^2}{1-\beta^2-2\alpha\beta-3\alpha^2} = 3\frac{1-(\alpha+\beta)^2}{1-(\alpha+\beta)^2-2\alpha_1^2} > 3$$

si el denominador es positivo.

1.4.3 Modelo IGARCH

En algunas aplicaciones se tiene un valor de $\alpha(1) + \beta(1)$ muy cercano a la unidad, lo que conduce al modelo $GARCH(p,q)$ *Integrado*, denotado como *IGARCH(p,q)* [Engle y Bollerslev (1986)]. En él, el polinomio autorregresivo en la ecuaciónde la varianza tiene una raíz exactamente igual a 1. En el caso particular $p=1, q=1$, el modelo *IGARCH* puede escribirse:

$$
\begin{aligned}
y_t &= \varepsilon_t h_t \\
h_t^2 &= \omega + h_{t-1}^2 + \alpha \left(y_{t-1}^2 - h_{t-1}^2 \right), \quad t = 1, 2, ...
\end{aligned}
$$

lo que hace que un shock en la varianza condicional sea persistente, no desapareciendo nunca su efecto, a diferencia de lo que ocurre en el modelo $GARCH(1,1)$. Además, la varianza no muestra reversión a la media, por lo que transcurren períodos largos antes de que la varianza vuelva a tomar su valor promedio. Esto es totalmente paralelo a la diferencia que existe entre modelos $ARMA$ y $ARIMA$ en lo relativo a las respuestas a una innovación transitoria.

El proceso puede escribirse también,

$$
y_t^2 = \omega + h_{t-1}^2 + \alpha \left(y_{t-1}^2 - h_{t-1}^2 \right) + \left(y_t^2 - h_t^2 \right), \quad t = 1, 2, ...
$$

Este proceso no es débilmente estacionario, puesto que su varianza incondicional no es finita. Sin embargo, si $\omega > 0$, el proceso es estrictamente estacionario y ergódico [Nelson (1990)].

1.5 Predicción de la varianza futura

En esta sección desarrollamos epxresiones analíticas para el cálculo de la predicción de la varianza k-períodos hacia el futuro.

1.5.1 Modelo ARCH(p)

Teniendo en cuenta la expresión de su varianza incondicional, el modelo $ARCH(q)$ puede representarse,

$$
h_t^2 - \sigma_y^2 = \sum_{i=1}^{q} \alpha_i \left(y_{t-i}^2 - \sigma_y^2 \right),
$$

por lo que,

$$
\begin{aligned}
E_t h_{t+1}^2 &= \sigma_y^2 + \sum_{i=0}^{q-1} \alpha_i \left(y_{t-i}^2 - \sigma_y^2 \right), \\
E_t h_{t+2}^2 &= \sigma_y^2 + \alpha_1 \left(y_{t-1}^2 - \sigma_y^2 \right) + \sum_{i=2}^{q-1} \alpha_i \left(y_{t-i}^2 - \sigma_y^2 \right)
\end{aligned}
$$

En general,

$$
E_t h_{t+1}^2 = \sigma_y^2 + \sum_{i=0}^{q-1} \alpha_i \left(E_t h_{t-i}^2 - \sigma_y^2 \right),
$$

donde $E_t h_s^2 = \hat{h}_s^2$ para $s \leq t$, donde $\hat{h}_s^2$ denota el valor ajustado para la varianza condicional en el período s en la estimación del modelo $ARCH$.

1.5.2 Modelo AR(1)-ARCH(1)

Escribiendo la ecuación que representa la evolución temporal de y_t en este modelo en un instante de tiempo futuro, tenemos,

$$y_{t+k} = \phi^k y_t + \sum_{i=1}^{k} \phi^{k-i} \varepsilon_{t+1}$$

de modo que la predicción óptima de la varianza condicional de y_{t+k} en el instante t es,

$$Var_t\ y_{t+k} = \alpha_0 + \alpha_1 E_t \left(h_{t+i-1}^2 - \sigma^2 \right) = \phi^2 \sum_{i=0}^{k-1} \sigma^{2i} + \alpha_1^{k-1} \left(h_{t+1}^2 - \sigma^2 \right) \sum_{i=0}^{k-1} \phi^{2i} \alpha_1^{-i}$$

que es claramente dependiente del conjunto de información disponible en el instante t. Sin embargo, al aumentar el horizonte de predicción, la dependencia respecto de $h_{t+1}^2 - \sigma^2$ va reduciéndose, y la expresión de predicción de la varianza puede aproximarse por,

$$Var_t\ y_{t+k} = \phi^2 \sum_{i=0}^{k-1} \sigma^{2i}$$

que es la expresión que utilizaríamos para prever la varianza incondicional en ausencia de estructura $ARCH$ en la innovación del proceso.

1.5.3 Modelo GARCH(1,1)

Mediante sucesivas iteraciones, es fácil probar [Engle y Bollerslev (1986)] que la predicción de la varianza que se deduce de un modelo $GARCH(1,1)$, a partir de la predicción un período hacia adelante, es:

$$\begin{aligned} E_T h_{T+1}^2 &= \omega + \alpha y_T^2 + \beta h_T \\ E_T h_{T+k}^2 &= \omega + (\alpha+\beta) h_{T+k-1}^2 = \frac{\omega}{1-\alpha-\beta} + (\alpha+\beta)^{k-1} \left(E_T h_{T+1}^2 - \frac{\omega}{1-\alpha-\beta} \right) \end{aligned}$$

que converge, según se aleja el horizonte de predicción, a la varianza incondicional, $\frac{\omega}{1-\alpha-\beta}$.

Las predicciones de volatilidad que se obtienen de un modelo como éste pueden utilizarse para valorar una opción utilizando la fórmula de Black-Scholes. Para ello, una vez obtenidas las predicciones de la volatilidad diaria desde el instante actual hasta el vencimiento de la opción, obtendríamos la volatilidad media que, anualizada, utilizaríamos en la expresión de Black-Scholes:

$$\begin{aligned}
\frac{1}{T-t}\sum_{k=1}^{T-t} E_T h_{T+k}^2 &= \frac{\omega}{1-\alpha-\beta} + \frac{1}{T-t}\sum_{k=1}^{T-t}(\alpha+\beta)^{k-1}\left(E_T h_{T+1}^2 - \frac{\omega}{1-\alpha-\beta}\right) = \\
&= \frac{\omega}{1-\alpha-\beta} + \left(E_T h_{T+1}^2 - \frac{\omega}{1-\alpha-\beta}\right)\frac{1}{T-t}\frac{1-(\alpha+\beta)^{T-t}}{1-\alpha-\beta}
\end{aligned}$$

En el modelo $IGARCH(1,1)$, la predicción de la varianza es:

$$\begin{aligned}
E_t h_{t+k}^2 &= \omega\,(k-1) + h_t^2 \\
E_t h_{t+k}^2 &= \omega k + \alpha y_t^2 + (1-\alpha)\,h_t^2
\end{aligned}$$

que no converge a la varianza condicional, pues crece linealmente con el horizonte de predicción. De hecho, puede observarse en esta expresión que una perturbación en la varianza del instante T, incluso si resulta ser de carácter transitorio, se extrapola a las predicciones de la volatilidad a todos los horizontes.Sin embargo, Kleigbergen y Van Dijk (1993) han sugerido que, tanto en el modelo $GARCH$ como en el $IGARCH$, las predicciones de la varianza se obtengan mediante simulación del modelo. [FigureII.4.3]

1.6 Modelo EGARCH(p,q)

Los modelos anteriores recogen adecuadamente las propiedades de distribuciones de colas gruesas, y de agrupamiento de volatilidades, pero son simétricos: en ellos, la varianza condicional depende de la magnitud de las innovaciones retardadas, pero no de su signo. Para recoger los efectos apalancamiento observados en series financieras fue propuesto el modelo *exponencial GARCH*, o $EGARCH(p,q)$:

$$\begin{aligned}
y_t &= \varepsilon_t h_t \\
\ln h_t^2 &= \omega + \sum_{i=1}^{q}\beta_i \ln h_{t-i}^2 + \sum_{j=1}^{p}\theta_j g\left(\varepsilon_{t-j}\right)
\end{aligned}$$

donde los ε_t tienen todos distribución N(0,1), y carecen de correlación serial, y la función de respuesta asimétrica está definida por:

$$g\left(\varepsilon\right) = \delta\varepsilon + \alpha\left(\mid \varepsilon \mid - E\left(\mid \varepsilon \mid\right)\right),$$

La estructura $ARMA(q,p)$ de $\ln(h_t^2)$ debe satisfacer las condiciones de estacionariedad habituales en estos modelos. Generalmente, se utiliza en esta formulación la innovación: $\varepsilon_t = y_t/h_t$, es decir, el *error estandarizado*, dividido por su desviación típica condicional, donde y_t será generalmente la rentabilidad del activo. En tal caso, $E\left(\mid\varepsilon\mid\right)=\sqrt{\frac{2}{\pi}}$[3] , por lo que la sucesión $g\left(\varepsilon_{t}\right)$ es independiente en el tiempo, con esperanza cero y varianza constante, si es finita. El modelo EGARCH, al estar especificado para el logaritmo de la varianzas, no precisa de restricciones de signo para ninguno de sus parámetros.

Como caso particular, cuando es Normal, la ecuación de la varianza en el modelo $EGARCH(1,1)$ es:

$$\begin{aligned} y_t &= \varepsilon_t h_t \\ \ln h_t^2 &= \omega+\beta\ln h_{t-1}^2+\delta\varepsilon_{t-1}+\alpha\left(\mid\varepsilon_{t-1}\mid-\sqrt{2/\pi}\right) \end{aligned}$$

que es un proceso estacionario si $\mid\beta\mid<1$, y tiene como varianza condicional,

$$Var_{t-1}y_t=e^{\frac{\omega}{1-\beta}}$$

como puede verse tomando esperanzas en la ecuación que define el proceso.

La persistencia en volatilidad viene indicada por el parámetro β, mientras que δ mide la magnitud del *efecto apalancamiento*. En este modelo se espera que $\delta<0$, lo que implicaría que innovaciones negativas tuviesen un mayor impacto sobre la volatilidad que innovaciones positivas de igual tamaño. El término en $\delta\varepsilon_t$ permite la existencia de correlación entre el término de error y las varianzas condicionales futuras. Si, por ejemplo, $\alpha=0$ y $\delta<0$, entonces un ε_t negativo tendería a producir una innovación positiva en el proceso de varianza. Por último, la innovación en la varianza condicional es lineal en ε_t en dos segmentos a ambos lados del origen, con pendientes $\theta\left(\delta+\alpha\right)$ cuando ε_t es positivo, y $\theta\left(\delta-\alpha\right)$ cuando ε_t es negativo, lo que genera la asimetría en la varianza condicional.

Por otra parte, la esperanza y varianza incondicionales o marginales pueden aproximarse a partir de:

$$\begin{aligned} E\left(\ln y_t^2\right) &= -1,27+\frac{\omega}{1-\beta} \\ Var\left(\ln y_t^2\right) &= \frac{\pi^2}{2}+\frac{\delta^2+\alpha^2\left(1-\frac{2}{\pi}\right)}{1-\beta^2} \end{aligned}$$

Las expresiones para la predicción de la varianza s períodos hacia adelante son bastante complejas [ver Ruiz (1994) o Nelson (1991)] para el caso del modelo $EGARCH(1,0)$]. [EII.4.4]

[3]

$$\begin{aligned} E\quad\mid\quad \varepsilon_t\mid &= \int_{-\infty}^{\infty}\mid\varepsilon_t\mid\frac{1}{\sqrt{2\pi}}e^{-\varepsilon_t^2/2}d\varepsilon_t=2\int_0^{\infty}\frac{1}{\sqrt{2\pi}}e^{-u}du= \\ &= 2\frac{1}{\sqrt{2\pi}}\left(e^{-u_t}\mid_0^{\infty}\right)=\sqrt{\frac{2}{\pi}} \end{aligned}$$

where we have made the change of variable $\varepsilon_t^2/2=u_t$

1.7 Otras especificaciones univariantes en la familia ARCH

En todas las especificaciones que siguen, mantenemos la hipótesis simplificadora de que carecemos de variables epxlicativas para y_t. En caso contrario, h_t^2 representaría la varianza condicional de ε_t, on de y_t.

El modelo $GARCH(1,1)$ no recoge a plena satisfacción las característica de asimetría y curtosis que se observan en series financieras. Para resolver el problema de la curtosis, suele utilizarse una distribución t en lugar de una distribución Normal para las innovaciones. Por otra parte, existe una versión asimétrica del modelo $GARCH$ (Engle 1990, Review of Financial Studies), el modelo $AGARCH$, que trata de recoger de modo más apropiado la asimetría de las series financieras.

El modelo $AGARCH(1,1)$ es: [EII.4.3]

$$\begin{aligned} y_t &= \varepsilon_t h_t \\ h_t^2 &= \omega + \alpha\left(y_{t-1} - \xi\right)^2 + \beta h_{t-1}^2 \end{aligned}$$

con $\omega > 0, \alpha > 0, \beta > 0$.

En este modelo, $\alpha < 0$ significa que los shocks negativos sobre los rendimientos incrementan más la volatilidad condicional que los shocks positivos, lo que constituye el efecto apalancamiento, que es habitual en los mercados financieros (Black (1976), Christie (1982)), para lo que también fueron propuestos otros modelos asimétricos que consideramos en esta sección.

La predicción de volatilidad a partir del modelo $AGARCH$ es:

$$\begin{aligned} E_T h_{T+1}^2 &= \omega + \alpha(y_T - \lambda)^2 + \beta h_T^2 \\ E_T h_{T+s+1}^2 &= \left(\omega + \lambda^2\alpha\right) + (\alpha + \beta)\, h_{T+s}^2 \end{aligned}$$

con volatilidad de largo plazo:

$$\bar{h}_t^2 = \frac{\omega + \lambda^2\alpha}{1 - (\alpha + \beta)}$$

y predicción de volatilidad:

$$\begin{aligned} E_T h_{T+1}^2 &= \hat{\omega} + \hat{\alpha}(\hat{\varepsilon}_T^2 - \hat{\lambda})^2 + \hat{\beta} h_T^2 \\ E_T h_{T+s}^2 &= \left(\hat{\omega} + \hat{\lambda}^2\hat{\alpha}\right) + \left(\hat{\alpha} + \hat{\beta}\right) E_T h_{T+s-1}^2 \end{aligned}$$

y se puede imponer un nivel de volatilidad de largo plazo mediante:

$$\begin{aligned} h_t^2 - \bar{h}^2 &= \alpha\left[\left(y_{t-1} - \xi\right)^2 - \bar{h}^2 - \zeta^2\right] + \beta\left(h_{t-1}^2 - \bar{h}^2\right) \\ h_t^2 &= \alpha\left[\left(y_{t-1} - \xi\right)^2 - \bar{h}^2 - \zeta^2\right] + \left(\beta h_{t-1}^2 + (1 - \beta)\bar{h}^2\right) \end{aligned}$$

Taylor (1986) y Schwert (1989a,b) han propuesto que sea la desviación típica quien dependa del valor absoluto de los residuos:

$$\begin{aligned} y_t &= \varepsilon_t h_t \\ h_t^2 &= \alpha_0 + \sum_{i=1}^{q} \alpha_i \mid \varepsilon_{t-i} \mid + \sum_{i=1}^{p} \beta_j h_{t-j}^2, \quad \alpha_0 > 0, \alpha_i, \beta_j \geq 0, \sum_{i=1}^{q} \alpha_i + \sum_{i=1}^{p} \beta_j < 1 \end{aligned}$$

Alternativamente, Higgins y Bera (1992) han propuesto una clase de modelos más general, denominada $NARCH$ (Non-linear $ARCH$):

$$\begin{aligned} y_t &= \varepsilon_t h_t \\ h_t^\gamma &= \alpha_0 + \sum_{i=1}^{q} \alpha_i \mid \varepsilon_{t-i} \mid^\gamma + \sum_{i=1}^{p} \beta_j h_{t-j}^\gamma, \quad \alpha_0 > 0, \alpha_i, \beta_j \geq 0, \sum_{i=1}^{q} \alpha_i + \sum_{i=1}^{p} \beta_j < 1 \end{aligned}$$

que, para γ=1, genera el modelo anterior.

Si este modelo se modifica para pasar a:

$$\begin{aligned} y_t &= \varepsilon_t h_t \\ h_t^\gamma &= \alpha_0 + \sum_{i=1}^{q} \alpha_i \mid \varepsilon_{t-i} - k \mid^\gamma + \sum_{i=1}^{p} \beta_j h_{t-j}^\gamma, \quad \alpha_0 > 0, \alpha_i, \beta_j \geq 0, \sum_{i=1}^{q} \alpha_i + \sum_{i=1}^{p} \beta_j < 1 \end{aligned}$$

para alguna constante no nula k, las innovaciones en t dependerán del tamaño, pero también del signo, de las innovaciones pasadas.

La formulación del modelo $NARCH$ con γ=2 es un caso especial del modelo $ARCH$ Cuadrático($QARCH$) con $q = 1, p = 1$, introducido por Sentana (1991), en el que la varianza condicional se modeliza a través de una forma cuadrática de las innovaciones retardadas:

$$\begin{aligned} y_t &= \varepsilon_t h_t \\ h_t^2 &= \omega + \alpha y_{t-1}^2 + \beta y_{t-1}, \end{aligned}$$

que, con el objeto de garantizar la no-negatividad de la varianza condicional, puede escribirse en función de parámetros b, c, d, tales que:

$$\alpha_1 = d > 0; \quad \omega = b^2 d + c > 0; \quad \beta = -2bd \gtreqless 0$$

por lo que eligiendo $c > 0, d > 0$ se garantiza $\alpha_1 > 0, \omega > 0$, mientras que β tendrá el mismo signo que b.

La varianza incondicional derivada de este modelo se obtiene tomando esperanzas en la ecuación de h_t^2, teniendo:

$$Eh_t^2 = \omega + \alpha E y_{t-1}^2 + \beta E y_{t-1}$$

que implica,

$$Var(y_t) = \omega + \alpha Var\left(y_{t-1}\right)$$

ya que $Ey_{t-1} = 0$. Suponiendo estacionariedad, llegamos a,

$$Var\ y_t = \frac{\omega}{1-\alpha}$$

El modelo $QARCH(1,1)$ puede generalizarse al modelo $GQARCH(1,1)$, que recoge bastante apropiadamente las características de volatilidad de los rendimientos financieros:

$$\begin{aligned} y_t &= \varepsilon_t h_t \\ h_t^2 &= \omega + \alpha y_{t-1}^2 + \theta y_{t-1} + \beta h_{t-1}^2, \end{aligned}$$

que comprende como caso particular al modelo $GARCH(1,1)$ cuando $\theta = 0$. La varianza incondicional que se deriva de este modelo es igual a la del modelo $GARCH(1,1)$,

$$Var\ y_t = \frac{\omega}{1-\alpha-\beta}$$

Ambos modelos generan una asimetría igual a cero. Genera una mayor curtosis el modelo generalizado $GQARCH(1,1)$.

Una versión sencilla de dicho modelo es el *ARCH asimétrico* ($AARCH$) [Engle (1990)], siendo el modelo $AARCH$(1,1):

$$h_t^2 = \omega + \alpha y_{t-1}^2 + \theta y_{t-1} + \beta h_{t-1}^2,$$

donde un valor negativo de significa que rendimientos positivos incrementan la volatilidad menos que rendimientos negativos (apalancamiento).

Otra especificación es el modelo *Non-linear asymmetric GARCH*, o *NA-GARCH.* El modelo $NAGARCH(1,1)$ es

$$h_t^2 = \omega + \alpha\left(\varepsilon_{t-1} + \gamma\sqrt{h_{t-1}}\right) + \theta\varepsilon_{t-1} + \beta h_{t-1}^2,$$

Un último modo de introducir efectos asimétricos es a través de la especificación:

$$h_t^\gamma = \omega + \sum_{i=1}^{q}\left[\alpha_i^+ I\left(\varepsilon_{t-i} > 0\right) \mid \varepsilon_{t-i} \mid^\gamma + \alpha_i^- I\left(\varepsilon_{t-i} \leq 0\right) \mid \varepsilon_{t-i} \mid^\gamma\right] + \sum_{j=1}^{p}\beta_j h_{t-j}^\gamma$$

donde I denota una función indicatriz que toma el valor 1 cuando se da la condición que aparece dentro del paréntesis, y toma el valor cero en caso contrario.

El modelo *Threshold ARCH* (*TARCH*) [Zakoian (1990)] corresponde al caso γ=1,

$$h_t = \omega + \sum_{i=1}^{q} \left[\alpha_i^{+} I\left(\varepsilon_{t-i} > 0\right) \mid \varepsilon_{t-i} \mid + \alpha_i^{-} I\left(\varepsilon_{t-i} \leq 0\right) \mid \varepsilon_{t-i} \mid\right] + \sum_{j=1}^{p} \beta_j h_{t-j}$$

Modelo GJR: Glosten, Jagannathan y Runkle (1993) proponen trabajar con $\gamma = 2$. Su modelo, conocido por sus iniciales, *GJR*, permite una respuesta cuadrática de la volatilidad a las sorpresas recibidas en el mercado, con distintos coeficientes para las malas noticias y para las buenas noticias, a la vez que mantiene la hipótesis de que la menor volatilidad se alcanzará cuando no haya sorpresas,

$$h_t^2 = \omega + \sum_{i=1}^{q} \left[\alpha_i^{+} I\left(\varepsilon_{t-i} > 0\right) \varepsilon_{t-i}^2 + \alpha_i^{-} I\left(\varepsilon_{t-i} \leq 0\right) \varepsilon_{t-i}^2\right] + \sum_{j=1}^{p} \beta_j h_{t-j}^2$$

En realidad, no es preciso incluir las dos variables indicadores, pudiendo utilizarse,

$$h_t^2 = \omega + \sum_{i=1}^{q} \left[\alpha_i \varepsilon_{t-i}^2 + \alpha_i^{-} I\left(\varepsilon_{t-i} \leq 0\right) \varepsilon_{t-i}^2\right] + \sum_{j=1}^{p} \beta_j h_{t-j}^2$$

Los parámetros α no son los mismos en ambos modelos, si bien existe una relación entre ambos. En el primer caso, α_i^{+} mide el efecto de una innovación pasada negativa, mientras α_i^{-} mide el efecto de una innovación pasada positiva; en el segundo caso, el efecto de una innovación pasada negativa es $\alpha_i + \alpha_i^{-}$ mientras que el de una innovación positiva es α_i. Un valor positivo de α_i^{-} en esta representación indicaría que una innovación negativa genera mayor volatilidad que una innovación positiva de igual tamaño, y la interpretación contraria se tendría para un valor negativo de α_i^{-}.

En el caso $q = p = 1$, se tendría, con esta segunda formulación,

$$h_t^2 = \omega + \left(\alpha + \gamma D_{t-1}\right) \varepsilon_{t-1}^2 + \beta h_{t-1}^2$$

donde el signo del parámetro γ es libre mientras que $\omega > 0, \alpha, \beta \geq 0$, [ver Engle y Ng (19xx)], y la variable ficticia D_t se define igual a 1 si $\varepsilon_t < 0$, e igual a cero en caso contrario. La varianza incondicional de este proceso es,

$$Var\ y_t = \frac{\omega}{1 - \alpha - \beta - \gamma/2}$$

La predicción de volatilidad a partir del modelo *GJR* es:

$$\begin{aligned} E_T h_{T+1} &= \omega + \alpha \varepsilon_T^2 + \gamma D_T \varepsilon_T^2 + \beta h_T \\ E_T h_{T+s+1} &= \omega + \left(\alpha + \beta + \frac{1}{2}\gamma\right) h_{T+s} \end{aligned}$$

Este modelo incluye como caso particular al modelo $GARCH(1,1)$ cuando γ=0. Cuando $\gamma \neq 0$, el modelo explica posibles asimetrías en la varianza de y_t : valores positivos de los parámetros α_i^- implican mayores respuestas de la volatilidad ante innovaciones negativas (malas noticias) que ante innovaciones positivas (buenas noticias), mientras que lo contrario ocurre para valores negativos de los parámetros α_i^-. Sin embargo, en mercados de renta fija, la interpretación de buenas y malas noticias es la opuestas, por lo que cabría esperar coeficientes α_i^- positivos.

Una representación bastante genérica, propuesta por Henstchel (1995), es,

$$\begin{aligned} \frac{h_t^\lambda - 1}{\lambda} &= \omega + \sum_{i=1}^{q} \alpha_i h_{t-1} \left[f\left(\varepsilon_{t-i}\right) \right]^\nu + \sum_{i=1}^{p} \beta_i \frac{h_{t-1}^\lambda - 1}{\lambda} \\ f\left(\varepsilon_{t-i}\right) &= \mid \varepsilon_t - b \mid - c\left(\varepsilon_t - b\right) \end{aligned}$$

en la que se sustituye $\frac{h_t^\lambda - 1}{\lambda}$ por $\ln h_t$ cuando $\lambda = 0$. De esta formulación pueden obtenerse muchas especificaciones como casos particulares. Así,

- $\lambda = \nu = 2$ y $b = c = 0 \Rightarrow GARCH$
- $\lambda = \nu = 2$ y $b = 0 \Rightarrow GJR - GARCH$
- $\lambda = \nu = 1$ y $b = 0, \mid c \mid \leq 1 \Rightarrow TARCH$
- $\lambda = \nu = 1, \mid c \mid \leq 1 \Rightarrow AGARCH$
- $\lambda = 0, \nu = 1$ y $b = 0 \Rightarrow EGARCH$
- $\lambda = \nu \neq 0$ y $b = c = 0 \Rightarrow NARCH$
- $\lambda = \nu = 2$ y $c = 0 \Rightarrow NAGARCH$
- $\lambda = \nu \neq 0$ y $c = 0 \Rightarrow Non-linear\ GARCH\ [Engle-Ng(1993)]$
- $\lambda = \nu \neq 0$ y $b = 0, \mid c \mid \leq 1 \Rightarrow APARCH(Asymmetric\ Power\ Arch)[Ding\ et\ al(1993)]$

Otros dos modelos propuestos recientemente son el *Structural ARCH* ($STARCH$), de Harvey, Ruiz y Sentana (1992), y el *Switching ARCH* ($SWARCH$) [Cai, Journal of Business and Economic Statistics (1994)], que postula que la variable en estudio se ajusta a una variedad de modelos $ARCH$, entre los cuales se mueve de acuerdo con la estructura de una cadena de Markov, lo cual puede ser útil para recoger episodios como el hundimiento de los mercados de valores observados en octubre de 1987 y agosto de 1998 [Campbell y Harrtschell, Journal of Financial Economics (1992)].

1.8 Modelos ARCH en media (ARCH-M)

Por último, en todos estos modelos pueden introducirse medias no nulas, lo que conduce a los modelos *ARCH* de regresión. Para ello, la primera ecuación se substituye por,

$$\begin{aligned} y_t &= \mu_t + u_t = x_t'\beta + u_t \\ u_t &= \varepsilon_t h_t \end{aligned}$$

En particular, resulta de gran interés contrastar si, cuando y_t es la rentabilidad de un activo o mercado, una de las potenciales variables explicativas x_t es precisamente la varianza condicional o la desviación típica condicional estimadas, h_t, con coeficiente positivo, lo que sugeriría que la rentabilidad del activo aumenta con el nivel de riesgo que impone al inversor.

En tal caso, tenemos los modelos denominados en media: *ARCH-M* y *GARCH-M* o sus variantes [Engle, Lilien, Roobins (1987)], en los que una variable explicativa es h_t^2, o h_t. La presencia de esta variable introducirá autocorrelación en el proceso de rentabilidades, y_t, a diferencia de los procesos sin estructura en media que hemos analizado en las secciones precedentes.

En general, estos modelos son del tipo,

$$\mu_t(\theta) = \xi_0 + \delta g\left(h_t^2(\theta), \theta\right)$$

o,

$$\mu_t(\theta) = x_t'\beta + \delta g\left(h_t^2(\theta), \theta\right)$$

donde suponemos que g es una función monótona de la varianza condicional, con $g(\alpha, \theta) = 0$, es decir, que la función es no nula unicamente si la varianza condicional es cambiante en el tiempo.

La interpretación del término $\delta g\left(h_t^2(\theta), \theta\right)$ es de una prima de riesgo, por la que un incremento en la varianza de la rentabilidad conduce a un aumento en la rentabilidad esperada. La posible existencia de tales primas en los mercados de divisas, así como en la formación de la estructura temporal de tipos de interés ha sido y es motivo de un amplio númeor de estudios.

Para analizar las propiedades de este tipo de modelos, consideremos una versión sencilla,

$$\begin{aligned} y_t &= \delta h_t^2 + \varepsilon_t, \; con\; \varepsilon_t \mid \Omega_{t-1} \sim N(0, h_t^2) \\ h_t^2 &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \end{aligned}$$

que permite escribir,

$$y_t = \delta\alpha_0 + \delta\alpha_1\varepsilon_{t-1}^2 + \varepsilon_t$$

donde ε_t es un proceso $ARCH(1)$. A partir de esta expresión, utilizando $E\left(\varepsilon_{t-1}^2\right) = \alpha_0/\left(1-\alpha_1\right)$, se tiene,

$$E\left(y_t\right)=\delta\frac{\alpha_0}{1-\alpha_1}$$

que puede interpretarse como la esperanza incondicional de la rentabilidad de mantener un activo con riesgo.

De modo análogo, tenemos,

$$Var\left(y_t\right)=\frac{\alpha_0}{1-\alpha_1}+\frac{\left(\delta\alpha_1\right)^2 2\alpha_0^2}{\left(1-\alpha_1\right)^2\left(1-3\alpha_1^2\right)} \tag{1}$$

Si no hay prima por riesgo, tendríamos: $Var\left(y_t\right)=\alpha_0/\left(1-\alpha_1\right)$. Por tanto, el segundo componente en (1) indica la presencia de una prima de riesgo, que hace que la dispersión de y_t aumente. Finalmente, el efecto $ARCH$-en media introduce autocorrelación en y_t, puesto que, én el caso del modelo,

$$\begin{aligned} y_t &= \delta h_t+\varepsilon_t,\ con\ \varepsilon_t \mid \Omega_{t-1}\sim N(0,h_t^2) \\ h_t^2 &= \alpha_0+\alpha_1\varepsilon_{t-1}^2 \end{aligned}$$

se tienen los coeficientes de autocorrelación,

$$\rho_1=Corr(y_t,y_{t-1})=\frac{2\alpha_1^3\delta^2\alpha_0}{2\alpha_1^2\delta^2\alpha_0+\left(1-\alpha_1\right)\left(1-3\alpha_1^2\right)}$$

$$\rho_k=Corr(y_t,y_{t-k})=\alpha_1^{k-1}\rho_1,\ \ k=2,3,....$$

Examinando las expresiones de ρ_1 y ρ_2 se aprecia que la región admisible para (ρ_1,ρ_2) es muy restrictiva.

En aplicaciones prácticas, las funciones más utilizadas son: $g\left(h_t^2\right)=h_t^2,\ g\left(h_t^2\right)=\sqrt{h_t^2},\ g\left(h_t^2\right)=\ln\left(h_t^2\right)$.

Bollerslev, Engle y Woolridge (1988) consideraron la versión multivariante de este modelo en el contexto del modelo de valoración de activos, *CAPM*.

1.9 Contrastes de estructura ARCH

Dada una relación del tipo: $y_t=x_t\beta+\varepsilon_t$,el contraste de los Multiplicadores de Lagrange (*ML*) propuesto por Engle (1982) considera la hipótesis nula:

$$\varepsilon_t\mid\Omega_{t-1}\sim N(0,\sigma^2)$$

donde I_t denota la información disponible en el instante t, y siendo x_t un vector de variables debilmente exógenas, o retardos de la variable dependiente. El interés de un contraste del tipo de los Multiplicadores de Lagrange (ML) reside en que, como es conocido, requiere únicamente la estimación del modelo restringido y, en este caso, la estimación del modelo bajo la hipótesis nula es muy simple. La hipótesis alternativa es que los residuos tienen una estructura de tipo $ARCH(q)$. Engle (1984) probó que T veces el R^2 de la regresión del cuadrado de los residuos obtenidos bajo la hipótesis nula, ε_t^2, sobre una constante y q retardos de los propios residuos al cuadrado, ε_{t-1}^2, $\varepsilon_{t-2}^2, ..., \varepsilon_{t-q}^2$ sigue, bajo la hipótesis nula, una distribución chi-cuadrado, con un número de grados de libertad igual al número de retardos incluidos en dicha regresión auxiliar, q. La intuición del contraste es bastante evidente: si la varianza de la perturbación es constante, entonces no podrá ser prevista a partir de los valores de los residuos pasados, cuyas fluctuaciones serán puramente aleatorias. Si, por el contrario, hay efectos *ARCH*, residuos recientes de elevado valor absoluto tenderán a sugerir un residuo corriente de elevada magnitud. Existe capacidad predictiva en la magnitud de los residuos pasados acerca de la magnitud de los residuos futuros. Dicho de otro modo, el valor absoluto del residuo mostrará autocorrelación temporal.

Sin embargo, la posible omisión de un regresor en el modelo de la media condicional, así como no tener en cuenta alguna no-linealidad o cierta autocorrelación, conduciría a un rechazo de la hipótesis nula, sugiriendo la presencia de estructura $ARCH$, incluso si ésta no existe, disminuyendo con ello el tamaño del contraste. Otra forma de llevar a cabo este contraste consiste en excluir la constante de la regresión auxiliar, restar una estimación de la varianza incondicional de la variable dependiente σ^2, y utilizar la mitad de la Suma Residual (suma de cuadrados de residuos) como estadístico de contraste. Otra posibilidad es un contraste del tipo portmanteau, como el de Ljung y Box (1978), para ε_t^2.

Como los parámetros del modelo $ARCH(q)$ deben ser positivos, el contraste debería ser de una cola, aunque para un orden q superior a 1, no es sencilla su puesta en práctica [Demos y Sentana (1991)]. Otra dificultad es que el contraste *ML* no tiene siempre mucha potencia cuando la alternativa es el modelo $GARCH(1,1)$, debido a la imposibilidad de identificar α_1 y β_1 por separado cuando el modelo *GARCH* es próximo al modelo incluido en H_0., De hecho, el contrate *ML* para $GARCH(1,1)$ es idéntico al correspondiente al modelo $ARCH(1)$, y algo similar ocurre para cualquier modelo $GARCH(p,q)$. Por otra parte, no es válido utilizar un test de Wald en el modelo $GARCH(1,1)$, que se basaría en el ratio t del coeficiente α_1 pues, en presencia de estructura *ARCH*, dicho estadístico no sigue una distribución t de Student. El contraste *RV* tiene la dificultad de que la distribución del estadístico bajo la hipótesis nula no es fácil de caracterizar, pero parece ser muy potente.

Por último, un contraste útil es el de insesgadez de las previsiones de volatilidad generadas por el modelo, para lo que se estima por mínimos cuadrados una regresión de y_t^2 sobre las varainzas h_t^2, en la que se esperaría encontrar una pendiente igual a 1 y una ordenda en el origen no significativa. Desviaciones de esta hipótesis conjunta pueden indicar preblemas de especificación en el modelo de la varianza condicional [Pagan y Schwert (1990)].

1.10 Contrastes de especificación

Los *contrastes de Normalidad* de la innovación ε_t pueden basarse en los residuos normalizados, una vez estimada la serie temporal de las varianzas condicionales. Tests habituales son: Jarque-Bera y Kolmogorov-Smirnov. También puede pensarse en un contraste mediante la χ^2 de Pearson.

La *ausencia de sesgo* en las estimaciones puede contrastarse mediante una regresión del cuadrado de la variable sobre una constante y las estimaciones de la varianza [Pagan y Schwert (1990)],

$$y_t^2 = b_0 + b_1 E_{t-1} y_t^2$$

Bajo una correcta especificación, los residuos de esta regresión no deberían presentar autocorrelación. Como las predicciones de la volatilidad un período hacia adelante deberían ser insesgadas, puede contrastarse la hipótesis nula: $H_0 : \; b_0 = 0, b_1 = 1$. Además, el R^2 de esta regresión puede utilizarse como una medida de bondad de ajuste.

Para poner en práctica las posibles desviaciones de Normalidad que puedan detectarse, se ha propuesto sustituir el supuesto de Normalidad en la estimación de Máxima Verosimilitud por las distribuciones t de Student, la distribución estándar generalizada (DEG), que incluye a la anterior como caso particular, y la distribución t generalizada, que incluye a ambas.

En general, los contrastes a llevar a cabo consisten en:

- a) contrastes de existencia de autocorrelación en media en los errores del modelo mediante técnicas Box-Jenkins,
- b) contrastes de existencia de efectos $ARCH$ no modelizados mediante técnicas Box-Jenkins aplicadas a los cuadrados de los residuos del modelo de la media,
- c) contrastes tipo Wald y de razón de verosimilitudes sobre la especificación de la ecuación de la media,
- d) contrastes de efectos asimétricos en la ecuación de la varianza, mediante el uso de variables ficticias de signo,
- e) contrastes de variables omitidas en la ecuación de la varianza,
- f) contraste de posible existencia de efectos $ARCH$ en media.

La familia de estadísticos Ljung-Box para el contraste de autocorrelación puede utilizarse tanto sobre los errores como sobre sus cuadrados (en este último caso como contraste de estructura ARCH). Su forma es,

$$Q(k) = T \sum_{i=1}^{k} \frac{T+2}{T+i} \rho_i^2$$

siento ρ_i el coeficiente de autocorrelación de orden i. Bajo la hipótesis nula, $Q(k)$ se distribuye como una chi-cuadrado con k grados de libertad.

Un contraste usualmente potente es el de los Multiplicadores de Lagrange, que utiliza una regresión de los residuos al cuadrado sobre una constante y sus k primeros retardos. El producto del tamaño muestral por el R^2 de dicha regresión se distribuye como una χ^2 con k grados de libertad. Sin embargo, este contraste no permite discriminar entre estructuras $ARCH$ y $GARCH$.

Estos contrastes se utilizan asimismo para evaluar un modelo ARCH ya estimado. En ese caso, hay que utilizar, lógicamente, los residuos de la ecuación de la media, estandarizados por la desviación típica condicional estimada, h_t.

Los contrastes tipo Wald para variables omitidas consisten en estimar el modelo más general, y contrastar la significación conjunta de los parámetros que distinguen el modelo restringido del modelo general. El contraste de razón de verosimilitudes estima ambos modelos: restringido y sin restringir, y compara la significatividad de la diferencia en los logaritmos de los máximos valores alcanzados por la función de verosimilitud en ambos casos. Para ello, se utiliza el resultado:

$$RV = -2\left(\ln L_R - \ln L_{SR}\right)$$

se distribuye asintóticamente, bajo la hipótesis nula, como una χ^2 con un número de grados de libertad igual al número de restricciones que se contrastan.

Para el contraste de asimetrías se utilizan los contrastes de signo proppuestos por Engle y Ng (1993): Definiendo unas variables ficticias S_{t-1}^{-} que toma el valor 1 si el residuo del período anterior ε_{t-1} fue negativo, y el valor cero en caso contrario, y S_{t-1}^{+},que toma el valor 1 si el residuo del período anterior ε_{t-1} fue positivo, y el valor cero en caso contrario, y definiendo el residuo estandarizado $z_t\left(\theta\right) = \frac{\varepsilon_t(\theta)}{h_t(\theta)}$ se estiman las regresiones,

Modelo I:

$$\begin{aligned} z_t^2 &= \beta_0 + \beta_1 S_{t-1}^{-} \\ H_0 &: \beta_1 = 0, \end{aligned}$$

Modelo II:

$$\begin{aligned} z_t^2 &= \beta_0 + \beta_1 S_{t-1}^{-}\varepsilon_{t-1} \\ H_0 &: \beta_1 = 0, \end{aligned}$$

Modelo III:

$$\begin{aligned} z_t^2 &= \beta_0 + \beta_1 S_{t-1}^{+}\varepsilon_{t-1} \\ H_0 &: \beta_1 = 0, \end{aligned}$$

Modelo IV:

$$\begin{aligned} z_t^2 &= \beta_0 + \beta_1 S_{t-1}^{-} + \beta_2 S_{t-1}^{+}\varepsilon_{t-1} + \beta_3 S_{t-1}^{-}\varepsilon_{t-1} \\ H_0 &: \beta_1 = \beta_2 = \beta_3 = 0, \end{aligned}$$

utilizando un estadístico tipo t en los tres primeros casos, y un estadístico tipo F en el último caso. Dada la posible existencia de autocorrelación y heterocedasticidad residual en los residuos estandarizados, dede utilizarse en los contrastes las varianzas de los parámetros estimadas del modo propuesto por Newey-West, que resultan robustas a la presencia de estos dos efectos.

Las correlaciones entre los residuos estandarizados y sus cuadrados pueden sugerir asimismo posibles asimetrías, y se utilizan a tal fin.

1.11 Estimación

[EII.4.1] [FigureII.4.2] [Ch3.5]

La estimación se lleva a cabo, generalmente, por máxima verosimilitud, para lo que suponemos una determinada densidad $f(z_t(\theta), \eta)$ para el término de error tipificado,

$$z_t(\theta) = \frac{\varepsilon_t(\theta)}{h_t(\theta)} = \frac{y_t - \mu_t(\theta)}{\left[h_t^2(\theta)\right]^{1/2}}$$

que tiene esperanza cero y varianza uno. Dado un vector de observaciones $\{y_1, y_2, ..., y_T\}$, el logaritmo de la función de verosimilitud para la observación t es:

$$l_t(y_t; \eta) = \ln\left\{f(z_t(\theta), \eta) - \frac{1}{2}\ln\left(h_t^2(\theta)\right)\right\}$$

donde el último término es el Jacobiano de la transformación que pasa de las innovaciones estandarizadas a las observaciones muestrales, que en el caso multivariante se convertirá en:

$$l_t(y_t; \eta) = \ln\left[f\left(\varepsilon_t(\theta)\left[\Gamma_t(\theta)\right]^{-1}; \eta\right)\right] - \frac{1}{2}\ln \mid \Sigma_t(\theta) \mid$$

donde Γ es una matriz no singular, de igual dimensión que Σ, tal que $\Gamma\Gamma' = \Sigma$. Es bien sabido que para toda matriz definida positiva Σ existe tal matriz Γ. Si la matriz Σ es diagonal, aunque con elementos diferentes a lo largo de la diagonal principal, entonces Γ es la matriz diagonal que tiene por elementos la raiz cuadrada de los elementos en la diagonal de Σ. Como los elementos de esta última, los $h_t^2(\theta)$ son todos positivos, no hay ninguna dificultad en este tipo de cálculo.

Por otra parte, utilizando un argumento estándar para la descomposición del error de predicción, la función de verosimilitud para la muestra completa puede escribirse como la suma de los logaritmos de la función de verosimilitud condicional:

$$L_T(y_1, y_2, ..., y_T) = \sum_{t=1}^{T} l_t(y_t; \psi) \tag{2}$$

cuya maximización generará estimadores de MV de los parámetros del modelo, $\psi = (\theta, \eta)$.

Si la función de densidad condicional y las funciones que recogen los modelos de la media y la varianza son diferenciables, el estimador de MV se obtiene resolviendo el sistema de $m + k$ ecuaciones:

$$S_T\left(y_1, y_2, ..., y_T; \psi\right) = \sum_{t=1}^{T} s_t(y_t; \psi) = 0 \tag{3}$$

donde $s_t(y_t;\psi) = \nabla_\psi l_t(y_t;\psi)$ es el vector *score* correspondiente a la observación t. Si denotamos por $f'\left(z_t\left(\theta\right);\eta\right)$ la derivada parcial de la función f respecto de su primer argumento, tendremos,

$$\nabla_\theta l_t(y_t;\psi) = \frac{f'\left(z_t\left(\theta\right);\eta\right)}{f\left(z_t\left(\theta\right);\eta\right)} \nabla_\theta z_t\left(\theta\right) - \frac{1}{2}\frac{\nabla_\theta h_t^2\left(\theta\right)}{h_t^2\left(\theta\right)}$$

expresión en la que hay que incorporar:

$$\nabla_\theta z_t\left(\theta\right) = -\frac{\nabla_\theta \varepsilon_t\left(\theta\right)}{\sqrt{h_t^2\left(\theta\right)}} - \frac{1}{2}\varepsilon_t\left(\theta\right)\frac{\nabla_\theta h_t^2\left(\theta\right)}{\left[h_t^2\left(\theta\right)\right]^{3/2}}$$

y la resolución del conjunto de $m+k$ ecuaciones (3) habrá de ser numérica.

Para proceder con la estimación MV hay que establecer una determinada hipótesis acerca del tipo de distribución que sigue la innovación. Si se considera que obedece a una distribución Normal, tenemos:

$$f\left(z_t\left(\theta\right)\right) = \frac{1}{\sqrt{2\pi}}\exp\left[-\frac{z_t\left(\theta\right)^2}{2}\right] = \frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(y_t - \mu_t\left(\theta\right)\right)^2\right]$$

En este caso, como la distribución está totalmente determinada por sus dos primeros momentos, sólo la media y varianza condicionales aparecen en la función de verosimilitud (2), por lo que $\psi = \theta$, y la función *score* adopta la forma:

$$s_t\left(y_t;\theta\right) = \nabla_\theta \mu_t\left(\theta\right)\frac{\varepsilon_t\left(\theta\right)}{\sqrt{h_t^2\left(\theta\right)}} + \frac{1}{2}\frac{\nabla_\theta h_t^2\left(\theta\right)}{\sqrt{h_t^2\left(\theta\right)}}\left[\frac{\varepsilon_t\left(\theta\right)^2}{h_t^2\left(\theta\right)} - 1\right]$$

En este caso, puede probarse [Hamilton (1994)] que la expresión analítica del *score* es,

$$s_t\left(y_t;\theta\right) = \frac{\varepsilon_t\left(\theta\right)^2 - h_t^2\left(\theta\right)}{2\left[h_t^2\left(\theta\right)\right]^2}\left[\begin{array}{l} -2\sum_{j=1}^{m}\alpha_j\varepsilon_{t-j}x_{t-j} \\ 1 \\ \varepsilon_{t-1}^2 \\ ... \\ \varepsilon_{t-m}^2 \end{array}\right] + \left[\begin{array}{l} x_t\varepsilon_t/h_t^2 \\ 0_{m+1} \end{array}\right]$$

donde ambos vectores columna tienen dimensión $qx1$, siendo $q = k+m+1$, con k el número de variables explicativas en el modelo de la media, y m el número de retardos del modelo $ARCH(m)$.

El gradiente de la función de verosimilitud puede entonces expresarse analíticamente como la suma de los *scores*,

$$\nabla \ln L = \sum_{t=1}^{T} s_t \left(y_t ; \theta \right)$$

o puede también evaluarse numéricamente a través de derivadas numéricas de la función de verosimilitud.

Es habitual suponer que el error del modelo tiene distribución condicional Normal, en cuyo caso,

$$\ln L = -\frac{T}{2} \ln 2\pi - \frac{1}{2} \sum_{t=1}^{T} \ln h_t^2 - \frac{1}{2} \sum_{t=1}^{T} \frac{\varepsilon_t^2}{h_t^2}$$

en el que hay que substituir las expresiones de ε_t y h_t^2 que se obtienen de la especificación del modelo para la esperanza y la varianza condicionales de y_t. En realidad, las funciones de densidad que entran en esta expresión de la función de verosimilitud son funciones de densidad condicionales, debido a la presencia de $h_t^2 = g\left(\varepsilon_{t-1}, \varepsilon_{t-2}, ...\right)$ en la densidad correspondiente a ε_t. Así, el logaritmo de la función de verosimilitud condicional en las primeras m observaciones es,

$$\sum_{t=m+1}^{T} \ln f(y_t \ / \ x_t, x_{t-1}, ..., y_{t-1}, ...) = -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=m+1}^{T} \ln(h_t^2) - \frac{1}{2} \sum_{t=m+1}^{T} \ln \left[\frac{\left(y_t - \mu_{t-1}\left(\theta\right)\right)^2}{h_t^2} \right]$$

siendo m el orden de un proceso $ARCH$, o el número de retardos de la varianza condicional en el caso de un modelo $GARCH$. En el caso de un modelo $ARCH$, para calcular el valor numérico de la función de verosimilitud, se utiliza,

$$h_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + ... + \alpha_m \varepsilon_{t-m}^2 = \omega + \alpha_1 \left(y_{t-1} - \mu_{t-2}\left(\theta\right)\right)^2 + ... + \alpha_m \left(y_{t-m} - \mu_{t-m-1}\left(\theta\right)\right)^2$$

Por ejemplo, en el caso de un simple modelo $ARCH(1)$, con

$$\begin{aligned} y_t &= \varepsilon_t, \ \ \varepsilon_t \sim N(0, h_t^2) \\ h_t^2 &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \end{aligned}$$

la función logaritmo de la función de Verosimilitud condicional es,

$$\ln L = \sum_{t=m+1}^{T} \ln f(y_t \ / \ x_t, x_{t-1}, ..., y_{t-1}) = -\frac{T}{2} \ln 2\pi - \frac{1}{2} \sum_{t=1}^{T} \ln\left(\alpha_0 + \alpha_1 y_{t-1}^2\right) - \frac{1}{2} \sum_{t=1}^{T} \frac{y_t^2}{\alpha_0 + \alpha_1 y_{t-1}^2}$$

En el caso de un modelo $GARCH(1,1)$, utilizaríamos,

$$h_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}^2 = \omega + \alpha \left(y_{t-1} - \mu_{t-2}\left(\theta\right)\right)^2 + \beta h_{t-1}^2$$

a partir de un valor inicial $h_0^2 = \frac{\omega}{1-\alpha-\beta}$. En el caso de un modelo $GARCH$ de orden superior, actuaríamos de modo aálogo, inicializando todos los retardos precisos de la varianza condicional en el valor numérico de la varianza incondicional.

1.11.1 Imponiendo un nivel de volatilidad de largo plazo en la estimación: reversión a un nivel medio de volatilidad

[EII.4.4]

Como se observa, las funciones de verosimilitud de los modelos $ARCH$ son no lineales en los parámetros del modelo, por lo que la estimación de Máxima Verosimilitud, que es el procedimiento de estimación habitualmente utilizado, requiere el uso de algoritmos numéricos de optimización. Para llevar a cabo tales procedimientos, es preciso dar valores iniciales a los parámetros del modelo. Los parámetros de la ecuación del primer momento condicional de y_t se obtienen mediante estimación de dicha ecuación, ignorando la presencia de estructura del tipo $ARCH$.

Para dar valores iniciales a los parámetros de la ecuación de la varianza condicional, existen varias posibilidades: una posibilidad consistiría en tomar para la constante la varianza incondicional obtenida para el término de error de la ecuación de la media, que se ha estimado previamente para inicializar los parámetros de dicha ecuación. En este caso, habría que dar valores iniciales a todos los restantes parámetros de la ecuación de la varianza. Otra alternativa consiste en dar valores *razonables* a los parámetros de la ecuación de la varianza, como $\alpha = .10, \beta = .80$ en el caso de un modelo $GARCH(1,1)$, pero entonces hay que dar a la constante ω un valor inicial: $\omega = (1-\alpha-\beta)\sigma_\varepsilon^2$, siendo σ_ε^2 la varianza estimada para el término de error de la ecuación de y_t.

Si el modelo $GARCH$ no tiene estructura $ARCH$ en media, entonces la estimación por separado de los parámetros en la ecuación de la esperanza condicional y de los que entran en la ecuación de la varianza condicional es eficiente. Ello se debe a que la matriz de información presenta una estructura diagonal a bloques en ambos subvectores de parámetros. Esto no ocurre en el modelo $EGARCH$.

Estimación del modelo GARCH(1,1)

Bajo el supuesto de Normalidad para las rentabilidades logarítmicas, $R_t = \sigma_t z_t$, con $z_t \sim i., i.d.N(0,1)$, tendemos la función de verosimilitud,

$$L = \prod_{t=1}^{T} \left[\frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(\frac{-R_t^2}{2\sigma_t^2} \right) \right]$$

cuya maximización equivale a maximizar su logaritmo neperiano,

$$\ln L = \frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left(\ln\sigma_t^2 + \frac{R_t^2}{\sigma_t^2} \right)$$

que se puede maximizar bien mediante algoritmos numéricos, o bien mediante procedimientos de búsqueda. En todo caso, lo primero que hemos de hacer es substitutir en la expresión anterior la volatilidad σ_t^2 por un determinado modelo dependiente de un vector de parámetros θ. En las próximas secciones veremos cómo se lleva a cabo este proceso. Como en cualquier otro problema de estimación, hemos de tener en cuenta que estamos maximizando la verosimilitud bajo el supuesto de estabilidad paramétrica, lo que puede condicionar el número de observaciones utilizado en dicho proceso de estimación.

Distribuciones no Gaussianas

Para tratar de recoger toda la leptocurtosis de la distribución empírica, se utiliza en ocasiones la distribución t estandarizada con grados de libertad $\eta>2$:

$$f\left(z_t\left(\theta\right)\right)=\frac{1}{\sqrt{\pi\left(\eta-2\right)}}\frac{\Gamma\left(\frac{\eta+1}{2}\right)}{\Gamma\left(\frac{\eta}{2}\right)}\frac{1}{\sqrt{h_t^2}}\frac{1}{\left[1+\frac{z_t(\theta)^2}{\eta-2}\right]^{(\eta+1)/2}}$$

donde Γ denota la función Gamma. La distribución t es simétrica alrededor de cero, y converge a la Normal cuando $\eta\to\infty$. Para valores $\eta>4$ tiene colas más gruesas que la Normal, con coeficiente de curtosis igual a $3(n-2)/(n-4)$, que es superior a 3. El logaritmo de la función de verosimilitud condicional en las primeras m observaciones es,

$$\sum_{t=m+1}^{T}\ln f(y_t\;/\;x_t,x_{t-1},...,y_{t-1},...)=T\;\ln\left[\frac{\Gamma\left(\frac{\eta+1}{2}\right)}{\pi^{1/2}\Gamma\left(\frac{\eta}{2}\right)}\left(\eta-2\right)^{-1/2}\right]-\frac{1}{2}\sum_{t=m+1}^{T}\ln(h_t^2)-$$

$$-\frac{\eta+1}{2}\sum_{t=m+1}^{T}\ln\left[1+\frac{\left(y_t-\mu_{t-1}\left(\theta\right)\right)^2}{h_t^2\left(\eta-2\right)}\right]$$

Se utiliza asimismo la distribución *t-Generalizada*, que depende de 2 parámetros y es simétrica, con densidad absolutamente continua, con esperanza 0 y varianza 1. Su función de densidad puede escribirse,

$$f\left(\varepsilon_t\right)=\frac{p}{2\varkappa B\left(\frac{1}{p},\frac{1}{2s}\right)}\left(1+\frac{\mid\varepsilon_t\mid^p}{\varkappa^p}\right)^{-\left(\frac{1}{p}+\frac{1}{2s}\right)}$$

donde $p>0,s>0$ son parámetros a estimar que han de satisfacer: $p-4s>0$. El parámetro $\varkappa$ es $\varkappa=\frac{\sqrt{\Gamma\left(\frac{1}{p}\right)}\sqrt{\Gamma\left(\frac{1}{2s}\right)}}{\sqrt{\Gamma\left(\frac{3}{p}\right)}\sqrt{\Gamma\left(\frac{1}{2s}-\frac{2}{p}\right)}}$ y $B\left(.\right),\Gamma\left(.\right)$ denotan las funciones Beta y Gamma, respectivamente.

Se utiliza también la *Distribución Generalizada de Error* [Nelson (1991)]:

$$f\left(z_t\left(\theta\right)\right)=\frac{\eta}{\lambda}\frac{1}{2^{1+\frac{1}{\eta}}}\frac{1}{\Gamma\left(\frac{1}{\eta}\right)}\exp\left(-\frac{1}{2}\mid\frac{z_t\left(\theta\right)}{\lambda}\mid^{\eta}\right),\;\; con\;\;\lambda=\sqrt{\frac{1}{2^{2/n}}\frac{\Gamma\left(\frac{1}{\eta}\right)}{\Gamma\left(\frac{3}{\eta}\right)}}$$

que para $\eta=2$ coincide con la densidad Normal. Para $\eta<2$, esta distribución tiene colas más gruesas que la Normal, mientras que para $\eta>2$ tiene colas más finas que la distribución Normal. Esta densidad fue propuesta en un análisis de rentabilidades diarias del mercado de valores, en exceso de las ofrecidas por el activo sin riesgo. Para ello, Nelson (1991) especificó el modelo

$$r_t = a + br_{t-1} + \gamma h_t^2 + u_t$$

con $u_t = \varepsilon_t h_t$, siendo ε_t independiente, Normal(0,1). Suponiendo una estructura $EGARCH(1,1)$ para la varianza condicional, tendríamos,

$$\ln h_t^2 = \omega + \beta \ln h_{t-1}^2 + \delta \frac{\varepsilon_{t-1}}{h_{t-1}} + \theta \left(\mid \frac{\varepsilon_{t-1}}{h_{t-1}} \mid - \sqrt{\frac{2}{\pi}} \right)$$

y, suponiendo una función de densidad generalizada, el logaritmo de la función de verosimilitud sería,

$$\ln L = T \left[\ln \frac{\eta}{\lambda} - \left(1 + \frac{1}{\eta}\right) \ln 2 - \ln \Gamma \left(\frac{1}{\eta}\right) \right] - \frac{1}{2} \sum_{t=1}^{T} \left[\mid \frac{(r_t - a - br_{t-1} - \delta h_t^2)}{\lambda h_t} \mid \right]^{\eta} - \frac{1}{2} \sum_{t=1}^{T} \ln h_t^2$$

donde, para evaluar la función $\ln L$ es preciso, una vez más, generar datos para la varianza condicional como en otros casos, utilizando iterativamente la expresión que define la varianza condicional del proceso $EGARCH$,

$$\ln h_{t+1}^2 = \omega + \beta \ln h_t^2 + \delta \frac{\varepsilon_t}{h_t} + \theta \left(\mid \frac{\varepsilon_t}{h_t} \mid - \sqrt{\frac{2}{\pi}} \right)$$

con:

$$\varepsilon_t = \frac{r_t - a - br_{t-1} - \delta h_t^2}{h_t}$$

a partir de valores paramétricos $(a, b, \gamma, \omega, \beta, \delta, \theta)$ iniciales. Los valores iniciales de la varianza condicional (uno sólo en este caso), se fijan igual a su esperanza matemática, $Eh_t^2 = \frac{\omega}{1-\beta}$.

En otros casos [Engle y González-Rivera (1991)] se ha propuesto utilizar un procedimiento de estimación semiparamétrico.

Mixtura de GARCH Normales

Markov Switching GARCH

Para simular un modelo Markov Switching GARCH hay que tener en cuenta que la probabilidad incondicional de estar en el régimen 1 es:

$$\pi = \frac{\pi_{21}}{\pi_{12} + \pi_{21}}$$

1.11.2 Estimación por Cuasi-máxima verosimilitud

En muchos casos en el ámbito de los mercados financieros, la hipótesis de Normalidad del término de error de la ecuación de la media de una rentabilidad no es aceptable.Uuna posibilidad consiste en estiomar el modelo por máxima verosimilitud bajo un supuesto distinto acerca de la distribución de dicho términon de error, ya sea mediante una distribución t de Student, una distribución

GED, una mixtura de Normales, etc.. Alternativamente, si se supone Normalidad en el cálculo de la función de verosimilitud, el estimador que resulta es de Cuasi-máxima verosimilitud, que es consistente, pero no eficiente. Todo lo que se precisa para este resultado es que las ecuaciones de la esperanza y varianza condicionales se hayan especificado correctamente, lo cual puede resumirse en las condiciones,

$$E\left(\varepsilon_t^2/x_t, y_{t-1}, y_{t-2}, ...\right) = 0; \;\; Var\left(\varepsilon_t^2/x_t, y_{t-1}, y_{t-2}, ...\right) = 1.$$

La pérdida de eficiencia en la estimación se debe precisamente a la desviación respecto de la Normal, de la verdadera distribución de probabilidad del término de error del modelo. En tal caso, debe utilizarse una estimación de la matriz de covarianzas de los parametros que sea robusta a desviaciones de Normalidad, como la propuesta por Bollerslev y Wooldridge (1992).

Esta estrategia de estimación es similar a Máxima Verosimilitud, pero requiere corregir las desviaciones típicas resultantes. La distribución asintótica del estimador es,

$$\sqrt{T}\left(\hat{\theta} - \theta\right) \rightarrow N(0, D^{-1}SD^{-1})$$

donde,

$$S = p\lim_{T\rightarrow\infty}\frac{1}{T}\sum_{t=1}^{T} s_t\left(\theta\right) s_t\left(\theta\right)'$$

siendo $s_t\left(\theta\right)$ el vector *score,* mientras que la matriz D es,

$$D = -p\lim_{T\rightarrow\infty}\frac{1}{T}\sum_{t=1}^{T} E\left[\frac{\partial s_t\left(\theta\right)}{\partial\theta'} \mid x_t, y_{t-1}, y_{t-2}, ...\right]$$

Ambas pueden estimarse consistentemente evaluando numéricamente el vector score bajo los parámetros resultantes en la estimación [ver Hamilton (1994)], obteniéndose desviaciones típicas asintóticamente robustas a errores de especificación en la densidad del término de error, tomando raíces cuadradas del producto, $\frac{1}{T}\hat{D}_T^{-1}\hat{S}_T\hat{D}_T^{-1}$. Si el modelo está correctamente especificado y la distribución del término de error es normal, entonces $S = D$, y resulta la matriz de covarianzas asintótica habitual del estimador de Máxima Verosimilitud.

Primer caso: rentabilidades incorrelacionadas con media cero Supongamos que las rentabilidades obtenidas en la unidad temporal de observación carecen de autocorrelación, lo que puede contrastarse a partir de un examen de sus funciones de autocorrelación simple y parcial, así como llevando a cabo contrastes formales del tipo Ljung-Box o Box-Pierce.

Para estimar los parámetros del modelo en una hoja de cálculo, se estima inicialmente $\sigma_{t_0}^2$ por alguno de los dos procedimientos que mencionamos antes, y comienza la recursión a partir de dicho instante temporal, después de haber fijado valores iniciales para los parámetros α, β, ω. Una vez evaluada la función de verosimilitud para los valores parámetricos inicialmente escogidos (*condiciones iniciales*), se trata de buscar en el espacio paramétrico con el objeto de obtener los valores que maximizan la función de verosimilitud,

$$\ln L(\omega,\alpha,\beta) = \frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left(\ln\sigma_t^2(\omega,\alpha,\beta) + \frac{R_t^2}{\sigma_t^2(\omega,\alpha,\beta)}\right)$$

Finalmente, la varianza de largo plazo, σ^2, se estima a partir de las expresiones anteriores y las estimaciones obtenidas para α,β,ω: $\sigma^2 = \omega/(1-\alpha-\beta)$.

La alternativa denominada *variance targetting* consiste en fijar un nivel de volatilidad de largo plazo σ^2, por ejemplo igual a la varianza muestral, y utilizando la expresión analítica de la varianza a largo plazo para fijar $\omega = \sigma^2(1-\alpha-\beta)$, estimando así sólo 2 parámetros, α y β.

Si queremos estimar un modelo de alisado exponencial como el utilizado en RiskMetrics, se fija $\omega = 0, \alpha = 1-\lambda, \beta = \lambda$, y se efectúa una búsqueda sobre el valor numérico de $\lambda, \lambda \in (0,1)$,en la función

$$\ln L(\lambda) = \frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left(\ln\sigma_t^2(\lambda) + \frac{R_t^2}{\sigma_t^2(\lambda)}\right)$$

Segundo caso: rentabilidades posiblemente correlacionadas, con media no nula Como alternativa, consideremos la posibilidad de que las rentabilidades obedezcan al modelo

$$R_t = \rho_0 + \rho_1 R_{t-1} + \varepsilon_t$$

que recoge la presencia de autocorrelación, es decir, de dependencia temporal en las rentabilidades. Tendría sentido entonces hacer el supuesto de estructura GARCH de volatilidad, pero ahora sobre la innovación ε_t del proceso estocástico de rentabilidades, por lo que σ_t^2 sería ahora: $\sigma_t^2 = Var(\varepsilon_t)$, con función de verosimilitud,

$$L = \prod_{t=1}^{T}\left[\frac{1}{\sqrt{2\pi\sigma_t^2}}\exp\left(\frac{-\varepsilon_t^2}{2\sigma_t^2}\right)\right]$$

con,

$$\ln L = -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left(\ln\sigma_t^2 + \frac{\varepsilon_t^2}{\sigma_t^2}\right) = cons\tan te - \frac{1}{2}\sum_{t=1}^{T}\left(\ln\sigma_t^2 + \frac{(R_t-\rho_0-\rho_1 R_{t-1})^2}{\sigma_t^2}\right)$$

y la estimación del modelo se lleva a cabo buscando en los parámetros $\alpha,\beta,\omega,\rho_0,\rho_1$.

En este caso habría que tener en cuenta que el procedimiento nos daría la evolución temporal de la volatilidad de la innovación ε_t, el componente no predecible de la rentabilidad, que es la volatilidad de R_t condicional en su pasado, pero no su volatilidad incondicional. En todo caso, la volatilidad incondicional (un número) es la media de la volatilidad condicional (una variable). La relación entre las volatilidades incondicionales de la Rentabilidad y su innovación es,

$$Var(R_t) = \frac{Var(\varepsilon_t)}{1-\rho_1^2}$$

1.12 Contrastación de hipótesis

Crowder probó ya en 1976 que, bajo determinadas condiciones de regularidad, el estimador MV es consistente y tiene distribución asintótica Normal en modelos con observaciones dependientes. Si la densidad condicional está correctamente especificada y el verdadero vector de parámetros ψ_0 está en el interior del espacio paramétrico considerado, un argumento del tipo utilizado en el Teorema Central del Límite conduce a:

$$T^{1/2}\left(\hat{\psi}_T - \psi_0\right) \rightarrow N(0, A_0^{-1})$$

siendo la matriz de covarianzas asintótica del estimador MV igual a la inversa de la matriz de información, evaluada en el verdadero vector de parámetros:

$$A_0 = -T^{-1}\sum_{t=1}^{T} E\left[\nabla_\psi s\left(y_t;\psi_0\right)\right]$$

que es inferior a la matriz de covarianzas de cualquier otro estimador. En la práctica, se obtiene un estimador consistente de A_0 evaluando el análogo muestral en el vector estimado de parámetros $\hat{\psi}_T$, es decir, sustituyendo $E\left[\nabla_\psi s\left(y_t;\psi_0\right)\right]$ por $\nabla_\psi s\left(y_t;\hat{\psi}_T\right)$. Además, las segundas derivadas tienen generalmente esperanza nula, y pueden omitirse. Por último, bajo el supuesto de que la densidad esté correctamente especificada, se tiene la igualdad $A_0 = B_0$, siendo:

$$B_0 = T^{-1}\sum_{t=1}^{T} E\left[s\left(y_t;\psi_0\right) s\left(y_t;\psi_0\right)'\right] \tag{4}$$

es decir, que la esperanza del producto del gradiente por sí mismo proporciona asimismo un estimador de la matriz de covarianzas asintótica. Nuevamente, esta expresión se evaluaría en el vector estimado de parámetros.

En la estimación de modelos $ARCH$ suelen utilizarse derivadas numéricas, pues las derivadas analíticas son bastante complejas. El estimador propuesto en (4) tiene la ventaja de que sólo precisa derivadas de primer orden, pues las derivadas numéricas de segundo orden suelen ser bastante inestables. En general, el vector de parámetros de un modelo $ARCH$ puede particionarse: $\theta'=\left(\theta_1', \theta_2'\right)$, donde el primer subvector es el que aparece en el modelo de la esperanza condicional, mientras que el segundo es quien aparece en la determinación de la varianza condicional. Es, además, posible probar que, en algunos casos, la matriz de información es diagonal a bloques con esta partición. Como consecuencia, pueden calcularse estimadores asintóticamente eficientes para uno de los subvectores, a partir de una estimación consistente para el otro. Así, puede estimarse el modelo de la media por *MCO* (debe utilizarse un estimador consistente), para obtener un estimador asintóticamente eficiente de los parámetros de la varianza condicional a partir de los residuos *MCO* de la ecuación de la media. Sin embargo, la pérdida de eficiencia en los coeficientes del modelo de la media puede ser importante.

Las desviaciones típicas habituales no son apropiadas, debido a la presencia de heterocedasticidad, por lo que deben corregirse del modo sugerido por White (1980). En particular, la habitual desviación típica para los valores de la función de autocorrelación $(1/T)$ puede ser muy sesgada en presencia de estructuras *ARCH*.

La diagonalidad a bloques de la matriz de información no se cumple, sin embargo, en el modelo *EGARCH* ni en los modelos *ARCH-M*. En estos cassos, para obtener una estimación consistente es preciso que las funciones que representan la meia y varianza condicionales estén correctamente especificadas, y estimadas simultáneamente.

Si se quiere contrastar una hipótesis nula de interés, del tipo: $H_0 : r(\psi_0) = 0$,siendo el rango l de la función r inferior a $m+k$, el estadístico de Wald adopta la forma:

$$W_T = T.r\left(\hat{\psi}_T\right)' \left[\left(\nabla_\psi r\left(\hat{\psi}_T\right)\right) C_T^{-1} \left(\nabla_\psi r\left(\hat{\psi}_T\right)\right)'\right]^{-1} r\left(\hat{\psi}_T\right)$$

siendo C_T una estimación consistente de la matriz de covarianzas del vector de parámetros bajo la hipótesis alternativa. Bajo la hipótesis nula, y si se satisfacen las condiciones de regularidad, el estadístico de Wald tiene una distribución chi-cuadrado con $(m+k)-l$ grados de libertad, el número de parámetros bajo la hipótesis alternativa.

También puede utilizarse un contraste de *RV* (Razón de verosimilitudes), cuyo estadístico seguirá una distribución asimismo chi-cuadrado con $(m+k)-l$ grados de libertad, el número de restricciones (número de parámetros bajo la hipótesis alternativa).

La contrastación de hipótesis acerca de parámetros de la ecuación de varianza condicional está sujeta a dos dificultades: a) en primer lugar, dichos parámetros deben ser positivos, por lo que, como ya hemos dicho, los contrastes eficientes deberían ser de una cola, b) en segundo lugar, existen a veces problemas de identificación, a los que ya hemos hecho referencia, por lo que la matriz de información se hace singular. En el modelo $GARCH(1,1)$, bajo la hipótesis nula: $H_0 : \alpha_1 = 0$, los parámetros ω y β_1 no están identificados.

De igual modo, en el modelo $ARCH-M$, el coeficiente de la varianza (o desviación típica) condicional está identificado sólo si dicha varianza es cambiante en el tiempo, por lo que no es posible un contraste del tipo habitual para la hipótesis conjunta de presencia de efectos $ARCH$, junto con la significación del coeficiente de la ecuación de la media.

Otra cuestión de indudable relevancia se refiere a las propiedaddes en muestras finitas de los estimadores de máxima verosimilitud de modelos $ARCH$. Así, por ejemplo, con errores condicionalmente Normales, la estimación de $\alpha_1 + \beta_1$ resulta sesgada a la baja y asimétrica a la derecha en muestras finitas. El sesgo en la suma de los coeficientes proviene de un sesgo a la baja en la estimación de β_1, junto con un sesgo al alza en la estimación de α_1.

1.13 Modelos de varianza condicional como aproximaciones a difusiones.

Denotemos por Y_t el precio de un activo, y por σ_t la volatilidad instantánea de su rendimiento. Consideremos la representación conjunta de la evolución seguida por (Y_t, σ_t) a partir de valores iniciales (Y_0, σ_0) por medio del proceso en tiempo continuo,

$$\begin{aligned} dY_t &= \mu Y_t dt + Y_t \sigma_t dW_{1,t} \\ d\left[\ln \sigma_t^2\right] &= -\beta \left[\ln \sigma_t^2 - \alpha\right] dt + \psi dW_{2,t} \end{aligned} \tag{5}$$

donde $W_{1,t}$ y $W_{2,t}$ denotan movimientos brownianos independientes de las condiciones iniciales, que satisfacen:

$$\begin{pmatrix} dW_{1,t} \\ dW_{2,t} \end{pmatrix} \begin{pmatrix} dW_{1,t} & dW_{2,t} \end{pmatrix} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} dt$$

es decir, con correlación igual a ρ.

Aunque los datos se observan únicamente a intervalos de tiempo discretos, es muy útil formular representaciones continuas de los precios de un activo. Es útil, en particular, para análisis teóricos en la formación de precios de opciones. El lema de Ito permite escribir la ecuación anterior como:

$$dy_t = \left(\mu - \frac{\sigma_t^2}{2}\right) dt + \sigma_t dW_{1,t}$$

donde $y_t = \ln(Y_t)$.

Si un modelo teórico propone la representación acontinua anterior, ¿es posible formular un proceso $ARCH$ cuyas realizaciones muestrales sean indistinguibles de las generadas por el proceso de difusión cuando el intervalo de tiempo transcurrido entre observaciones sea muy reducido? Melino y Turnbull (1990) utilizan una aproximación de Euler para probar que la difusión (5) puede aproximarse por:

$$\begin{aligned} y_{t+h} &= y_t + \left(\mu - \frac{\sigma_t^2}{2}\right) h + h^{1/2} \sigma_t Z_{1,t+h}, \;\; t = h, 2h, 3h, ... \\ \ln\left(\sigma_{t+h}^2\right) &= \ln\left(\sigma_t^2\right) - h\beta \left[\ln\left(\sigma_t^2\right) - \alpha\right] + h^{1/2} \psi Z_{2,t+h}, \;\; t = h, 2h, 3h, ... \end{aligned} \tag{6}$$

siendo $(Z_{1,t}, Z_{1,t})$ una variable aleatoria Normal bivariante, con vector de esperanzas (0,0), y matriz de covarianzas:

$$Var\begin{pmatrix} Z_{1,t} \\ Z_{2,t} \end{pmatrix} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

Este proceso converge, efectivamente, a la difusión de la que hemos partido, cuando h tiende a cero. En efecto, es fácil ver que,

$$h^{-1}E_t\left[\begin{array}{l} y_{t+h}-y_t \\ \ln\left(\sigma_{t+h}^2\right)-\ln\left(\sigma_t^2\right) \end{array}\right]=\left[\begin{array}{l} \mu-\frac{\sigma_t^2}{2} \\ -\beta\left[\ln\left(\sigma_t^2\right)-\alpha\right] \end{array}\right]$$

$$h^{-1}Var_t\left[\begin{array}{l} y_{t+h}-y_t \\ \ln\left(\sigma_{t+h}^2\right)-\ln\left(\sigma_t^2\right) \end{array}\right]=\left[\begin{array}{ll} \sigma_t^2 & \sigma_t\rho\Psi \\ \sigma_t & \Psi^2 \end{array}\right]$$

que reproducen el proceso de media y la matriz de difusión en (5). Sin embargo, este no es estrictamente un proceso *ARCH*, pues σ_t^2 es la varianza condicional de $y_{t+h}-y_t$ dada toda la realización continua del proceso (no observable), pero no es la varianza condicional, dada la información reocgida a intervalos discretos de tiempo.

Para obtener un modelo $ARCH$ aproximado a la difusión anterior, reemplazamos la segunda ecuación del sistema (6) por:

$$\ln\left(\sigma_{t+h}^2\right)=\ln\left(\sigma_t^2\right)-h\beta\left[\ln\left(\sigma_t^2\right)-\alpha\right]+h^{1/2}g\left(Z_{1,t+h}\right),\quad t=h,2h,3h,...$$

para una función $g\left(.\right)$ medible, con $E\left[\mid g\left(Z_{2,t+h}\right)\mid^{2+\delta}\right]<\infty$ para algún $\delta>0$, y

$$Var\left(\begin{array}{c} Z_{1,t} \\ g\left(Z_{1,t}\right) \end{array}\right)=\left[\begin{array}{ll} 1 & \rho\psi \\ \rho\psi & 1\psi^2 \end{array}\right] \tag{7}$$

Para completar la formulación de la aproximación *ARCH*, necesitamos una especificación para la función $g(.)$. Puesto que,

$$E\left(\mid Z_{1,t}\mid\right)=\sqrt{\frac{2}{\pi}};\quad E\left(Z_{1,t}\mid Z_{1,t}\mid\right)=0;\quad Var\left(\mid Z_{1,t}\mid\right)=1-\frac{2}{\pi}$$

una posible especificación es,

$$g\left(Z_{1,t}\right)=\rho\Psi Z_{1,t}+\Psi\sqrt{\frac{1-\rho^2}{1-2/\pi}}\left(\mid Z_{1,t}\mid-\sqrt{\frac{2}{\pi}}\right)$$

que corresponde al modelo $EGARCH$.

Alternativamente, podría haberse escogido,

$$g\left(Z_{1,t}\right)=\rho\Psi Z_{1,t}+\Psi\sqrt{\frac{1-\rho^2}{2}}\left(Z_{1,t}^2-1\right)$$

que también satisface (7).

También se puede contestar a la pregunta inversa: Dado un modelo $ARCH$¿cuál es el proceso de difusión que mejor lo aproxima? Para ello, consideremos, a modo de ejemplo, una estructura de martingala con error $GARCH(1,1)$:

$$y_{t+h}=y_t+\sigma_t h z_{t+h}=y_t+\varepsilon_{t+h}$$

y:

$$\sigma_{t+h}^2 = \omega h + \left(1 - \theta h - \alpha h^{1/2}\right)\sigma_t^2 + h^{1/2}\alpha\varepsilon_{t+h}^2$$

que tiende a un modelo $IGARCH(1,1)$ cuando $h \to 0$.

Como se prueba en Nelson (1990a),

$$h^{-1}E_t\left[\begin{array}{c} y_{t+h} - y_t \\ \sigma_{t+h}^2 - \sigma_t^2 \end{array}\right] = \left[\begin{array}{l} 0 \\ \omega - \theta\sigma_t^2 \end{array}\right]$$

$$h^{-1}Var_t\left[\begin{array}{c} y_{t+h} - y_t \\ \sigma_{t+h}^2 - \sigma_t^2 \end{array}\right] = \left[\begin{array}{ll} \sigma_t^2 & 0 \\ 0 & 2\alpha^2\sigma_t^4 \end{array}\right]$$

para el que puede probarse que la difusión aproximada es,

$$\begin{aligned} dx_t &= \sigma_t dW_{1,t} \\ d\sigma_t^2 &= \left(\omega - \theta\sigma_t^2\right)dt + \sqrt{2}\alpha\sigma_t^2 dW_{2,t} \end{aligned}$$

donde $W_{1,t}$ y $W_{2,t}$ denotan movimientos brownianos independientes, lo cual puede utilizarse para estimar un proceso de difusión, y luego comparar parámetros.

1.14 Modelos multivariantes

La matriz de covarianzas entre las rentabilidades de un conjunto de activos puede escribirse, en notación matricial,

$$\Sigma_{t+1} = D_{t+1}\Gamma_{t+1}D_{t+1}$$

donde D_{t+1} es una matriz con desviaciones típicas condicionales en la diagonal y ceros fuera de la diagonal, y Γ_{t+1}es una matriz con unos en la diagonal, y con las correlaciones condicionales fuera de dicha diagonal principal.

Modelo EWMA

Un sólo parametro representa la evolución de todas las varianzas y covarianzas:

$$\begin{aligned} H_t &= (1-\lambda)H_{t-1} + \lambda\left(\varepsilon_{t-1}\varepsilon_{t-1}'\right) \Rightarrow \\ \left(\begin{array}{cc} \sigma_{1t}^2 & \\ \sigma_{12,t} & \sigma_{2t}^2 \end{array}\right) &= (1-\lambda)\left(\begin{array}{cc} \sigma_{1t-1}^2 & \\ \sigma_{12,t-1} & \sigma_{2t-1}^2 \end{array}\right) + \lambda\left(\begin{array}{cc} \varepsilon_{1,t-1}^2 & \\ \varepsilon_{1,t-1}\varepsilon_{2,t-1} & \varepsilon_{2,t-1}^2 \end{array}\right) \end{aligned}$$

es decir:

$$\begin{aligned} \sigma_{1,t}^2 &= \lambda\varepsilon_{1,t-1}^2 + (1-\lambda)\sigma_{1,t-1}^2 \\ \sigma_{12,t} &= \lambda\varepsilon_{1,t-1}\varepsilon_{2,t-1} + (1-\lambda)\sigma_{12,t-1} \\ \sigma_{1,t}^2 &= \lambda\varepsilon_{2,t-1}^2 + (1-\lambda)\sigma_{2,t-1}^2 \end{aligned}$$

que se estima maximizando el logaritmo de la función de verosimilitud:

$$\ln L(\theta,\lambda) = -\frac{1}{2}\sum_{t=1}^{T}\ln\left(\mid \Sigma_t \mid\right) - \frac{1}{2}\sum_{t=1}^{T}(r_t-\mu)'\Sigma_t^{-1}(r_t-\mu)$$

Si se aplica a rentabilidades estandarizadas, $\mid \Sigma_t \mid = 1-\rho_t^2$ en el caso bivariante, y $\Sigma_t^{-1} = \begin{pmatrix} 1 & -\rho_t \\ -\rho_t & 1 \end{pmatrix}$, por lo que:

$$\ln L(\theta,\lambda) = -\frac{1}{2}\sum_{t=1}^{T}\ln\left(1-\rho_t^2\right) - \frac{1}{2}\sum_{t=1}^{T}(r_{1t}-\mu_1)^2+(r_{2t}-\mu_2)^2-2\rho_t(r_{1t}-\mu_1)(r_{2t}-\mu_2)$$

aunque en muchos casos, se supondrá: $\mu_1 = \mu_2 = 0$.

Este es un modelo excesivamente sencillo, en el sentido de que ni permite que las innovaciones en la primera variable afecten a la varianza condicional de la segunda variable (ni viceversa), ni tampoco a la covarianza. Por tanto, las sorpresas en una variable tienen efectos exclusivamente sobre las propiedades de esa variable. Por otra parte, la covarianza sigue una evolución relativamente independiente de la magnitud de las innovaciones que experimentan las dos rentabilidades, dependiendo únicamente de su producto cruzado, pero no de su tamaño de un modo claro.

Modelo VECH Diagonal:

Es un modelo simple, nuevamente sin interdependencia dinámica entre volatilidades, aunque sin las estrictas restricciones paramétricas del modelo anterior:

$$\Sigma_t = C + \sum_{i=1}^{m} A_i \odot (\varepsilon_{t-i}\varepsilon_{t-i}') + \sum_{j=1}^{s} B_j \odot \Sigma_{t-j}$$

donde $\odot$ denota el producto elemento por elemento, y A_i y B_j son matrices simétricas. En el caso bivariante, con $m=1, s=1$:

$$\begin{aligned}\begin{pmatrix} \sigma_{11,t} & \\ \sigma_{12,t} & \sigma_{22,t} \end{pmatrix} &= \begin{pmatrix} C_{11} & \\ C_{12} & C_{22} \end{pmatrix} + \begin{pmatrix} A_{11} & \\ A_{12} & A_{22} \end{pmatrix} \odot \begin{pmatrix} \varepsilon_{1,t-1}^2 & \\ \varepsilon_{1,t-1}\varepsilon_{2,t-1} & \varepsilon_{2,t-1}^2 \end{pmatrix} + \\ &\quad \begin{pmatrix} B_{11} & \\ B_{12} & B_{22} \end{pmatrix} \odot \begin{pmatrix} \sigma_{11,t-1} & \\ \sigma_{12,t-1} & \sigma_{22,t-1} \end{pmatrix}\end{aligned}$$

es decir,

$$\begin{aligned} \sigma_{1t}^2 &= C_{11} + A_{11}\varepsilon_{1,t-1}^2 + B_{11}\sigma_{1,t-1}^2 \\ \sigma_{12,t} &= C_{12} + A_{12}\varepsilon_{1,t-1}\varepsilon_{2,t-1} + B_{12}\sigma_{12,t-1} \\ \sigma_{2t}^2 &= C_{22} + A_{22}\varepsilon_{2,t-1}^2 + B_{22}\sigma_{2,t-1}^2 \end{aligned}$$

Cada elemento sigue una estructura GARCH(1,1) y no permite dependencia dinámica cruzada entre las series de volatilidad. Es un modelo simple, pero no garantiza que la matriz de varianzas y covarianzas sea positiva definida en cada período, lo cual puede generar problemas numéricos.[4] Como vemos, con dos variables, el modelo contienen 9 parámetros.

Con 3 variables:

$$\begin{aligned}\begin{pmatrix}\sigma_{11,t} & & \\ \sigma_{12,t} & \sigma_{22,t} & \\ \sigma_{13,t} & \sigma_{23,t} & \sigma_{33,t}\end{pmatrix} &= \begin{pmatrix}C_{11} & & \\ C_{21} & C_{22} & \\ C_{31} & C_{32} & C_{33}\end{pmatrix} + \\ & \begin{pmatrix}A_{11} & & \\ A_{21} & A_{22} & \\ A_{31} & A_{32} & A_{33}\end{pmatrix} \odot \begin{pmatrix}\varepsilon_{1,t-1}^2 & & \\ \varepsilon_{2,t-1}\varepsilon_{1,t-1} & \varepsilon_{2,t-1}^2 & \\ \varepsilon_{3,t-1}\varepsilon_{1,t-1} & \varepsilon_{3,t-1}\varepsilon_{2,t-1} & \varepsilon_{3,t-1}^2\end{pmatrix} \\ & + \begin{pmatrix}B_{11} & & \\ B_{21} & B_{22} & \\ B_{31} & B_{32} & B_{33}\end{pmatrix} \odot \begin{pmatrix}\sigma_{11,t-1} & & \\ \sigma_{12,t-1} & \sigma_{22,t-1} & \\ \sigma_{13,t-1} & \sigma_{23,t-1} & \sigma_{33,t-1}\end{pmatrix}\end{aligned}$$

es decir,

$$\begin{aligned}\sigma_{11,t} &= C_{11} + A_{11}\varepsilon_{1,t-1}^2 + B_{11}\sigma_{11,t-1} \\ \sigma_{12,t} &= C_{12} + A_{12}\varepsilon_{1,t-1}\varepsilon_{2,t-1} + B_{12}\sigma_{12,t-1} \\ \sigma_{22,t} &= C_{22} + A_{22}\varepsilon_{2,t-1}^2 + B_{22}\sigma_{22,t-1} \\ \sigma_{13,t} &= C_{13} + A_{13}\varepsilon_{1,t-1}\varepsilon_{3,t-1} + B_{13}\sigma_{13,t-1} \\ \sigma_{23,t} &= C_{23} + A_{23}\varepsilon_{3,t-1}\varepsilon_{2,t-1} + B_{23}\sigma_{23,t-1} \\ \sigma_{33,t} &= C_{33} + A_{33}\varepsilon_{3,t-1}^2 + B_{33}\sigma_{33,t-1}\end{aligned}$$

El modelo de correlación condicional constante:

El modelo de correlación constante [Bollerslev (1990)] supone que las correlaciones entre cada par de rentabilidades son constantes en el tiempo: $\Gamma_{t+1} =$

---

[4] Una notación alternativa para este modelo es:

$$vech(H_t) = diag(\omega_1, \omega_2, \omega_3) + diag(\alpha_1, \alpha_2, \alpha_3) vech(\varepsilon_t \varepsilon_t') + diag(\beta_1, \beta_2, \beta_3) vech(H_{t-1})$$

En el caso de dos variables:

$$\begin{pmatrix}\sigma_{1,t}^2 \\ \sigma_{12,t} \\ \sigma_{2,t}^2\end{pmatrix} = \begin{pmatrix}\omega_1 & 0 & 0 \\ 0 & \omega_2 & 0 \\ 0 & 0 & \omega_3\end{pmatrix} + \begin{pmatrix}\alpha_1 & 0 & 0 \\ 0 & \alpha_2 & 0 \\ 0 & 0 & \alpha_3\end{pmatrix}\begin{pmatrix}\varepsilon_{1,t-1}^2 \\ \varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \varepsilon_{2,t-1}^2\end{pmatrix} + \begin{pmatrix}\beta_1 & 0 & 0 \\ 0 & \beta_2 & 0 \\ 0 & 0 & \beta_3\end{pmatrix}\begin{pmatrix}\sigma_{1,t-1}^2 \\ \sigma_{12,t-1} \\ \sigma_{2,t-1}^2\end{pmatrix}$$

es decir:

$$\begin{aligned}\sigma_{1,t}^2 &= \omega_1 + \alpha_1\varepsilon_{1,t-1}^2 + \beta_1\sigma_{1,t-1}^2 \\ \sigma_{12,t} &= \omega_2 + \alpha_2\varepsilon_{1,t-1}\varepsilon_{2,t-1} + \beta_2\sigma_{12,t-1} \\ \sigma_{1,t}^2 &= \omega_3 + \alpha_3\varepsilon_{2,t-1}^2 + \beta_3\sigma_{2,t-1}^2\end{aligned}$$

$\Gamma, \forall t$, por lo que el modelo de volatilidad consta sólo de ecuaciones para las varianzas. En el caso de dos activos:

$$\begin{pmatrix} \sigma_{11,t} \\ \sigma_{22,t} \end{pmatrix} = \begin{pmatrix} \alpha_{10} \\ \alpha_{20} \end{pmatrix} + \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \begin{pmatrix} a_{1,t-1}^2 \\ a_{2,t-1}^2 \end{pmatrix} + \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix} \begin{pmatrix} \sigma_{11,t-1} \\ \sigma_{22,t-1} \end{pmatrix}$$

y las covarianzas se estiman:

$$H_{ij}(t) = R_{ij} \sqrt{H_{ii}(t) H_{jj}(t)}$$

donde los elementos de R son parámetros estimados.

El modelo puede utilizarse con varianzas estimadas mediante esquemas EWMA o mediante modelos univariantes GARCH [ver EII.4.6]. Las diferencias entre las covarianzas obtenidas se deben a las diferencias que surgen al estimar la volatilidad de cada rentabilidad individual mediante un modelo GARCH univariante o mediante EWMA.

Bajo ambas opciones, cada covarianza se obtiene multiplicando el coeficiente de correlación entre las rentabilidades *estandarizadas/sin estandarizar*,[5] supuesto constante y calculado con toda la muestra, por el producto de las desviaciones típicas obtenidas a partir de los modelos de volatilidad condicional previamente estimados mediante EWMA o GARCH. Si se trabaja con rentabilidades diarias, hay que recordar multiplicar por 250 (número de sesiones diarias en un año) y sacar la raiz cuadrada para obtener la volatilidad anual, y multiplicar asimismo por 250 el producto de ambas volatilidades anuales y el coeficiente de correlacion muestral, al calcular la covarianza condicional. Podemos estimar el coeficiente de correlacion con toda la muestra, como se ha dicho, o hacer el ejercicio "en tiempo real", estimando la correlacion con datos hasta cada día. Excepto posiblemente al inicio de la muestra, las diferencias son pequeñas. Ver [EII.4.6] para el cálculo de la covarianza condicional con correlación constante, utilizando modelos asimétricos A-GARCH para las rentabilidades, en este caso de los tipos de cambio libra/USdólar y euro/USdólar.

Lógicamente, una vez estimado este modelo, no puede pretenderse generar correlaciones cambiantes en el tiempo a partir de las covarianzas. La estimación de las covarianzas es el punto final del ejercicio de estimación bajo esta especificación.

El modelo de correlación condicional dinámica:

Nuevamente en este modelo puede utilizarse un esquema EWMA para estimar las varianzas de las rentabilidades individuales, o estimar modelos GARCH univariantes. En el ejercicio [EII.4.6], [Figures II.4.6] se utiliza un procedimiento EWMA, con $\lambda = 0,94$, para generar volatilidades de las rentabilidades, sin media, de dos tipos de cambio: la libra y el euro, ambos respecto del dólar, $r_{it} - \mu_i$,

[5] Para estimar la correlación debe utiolizarse rentabilidades estandarizadas, pues de este modo, la varianza es constante en el tiempo, y el denominador del coeficiente de correlación está bien definido. Si se utilizan rentabilidades estandarizadas: $z_{it} = (r_{it} - \mu_i)/\sigma_{it}$, entonces $\rho_{ij,t} = E(z_{it} z_{jt})$, estimándose mediante la media muestral del producto de rentabilidades estandarizadas (ver [Ch3.6]).

con $\mu_i = T^{-1}\sum_{t=1}^{T} r_{it}$, y se generan: $\sigma_{it}^2 = \lambda\sigma_{it-1}^2 + (1-\lambda)\left(r_{it-1} - \mu_i\right)^2$. Las volatilidades anuales vienen dadas por $250\sigma_{it}^2$, y la covarianza podria estimarse mediante: $250\rho\sqrt{250\sigma_{1t}^2 250\sigma_{2t}^2}$, si bien este no es el cálculo más interesante en este modelo, con el que pretendemos estimar correlaciones cambiantes en el tiempo.

El modelo de correlación condicional dinámica utiliza las rentabilidades estandarizadas: $z_{1t} = (r_{1t} - \mu_1)/\sigma_{1t}, z_{2t} = (r_{2t} - \mu_2)/\sigma_{2t}$ para generar variables auxiliares $q_{ij,t}$:

$$q_{ij,t+1} = (1-\lambda)\, z_{i,t} z_{j,t} + \lambda q_{ij,t} \forall i, j$$

Es decir, en el caso de dos variables:

$$\begin{aligned}
q_{11,t+1} &= (1-\lambda)\, z_{1t}^2 + \lambda q_{11,t} \\
q_{12,t+1} &= (1-\lambda)\, z_{1t} z_{2t} + \lambda q_{12,t} \\
q_{11,t+1} &= (1-\lambda)\, z_{1t}^2 + \lambda q_{22,t}
\end{aligned}$$

Al ser una covarianza entre rentabilidades estandarizadas, la serie temporal q_{it} ya nos proporciona una estimación de la correlación condicional entre dos rentabilidades. Pero para garantizar que dicha correlación esté siempre en el intervalo $(-1, 1)$, utilizamos la normalización:

$$\rho_{ij,t+1} = \frac{q_{ij,t+1}}{\sqrt{q_{ii,t+1}}\sqrt{q_{jj,t+1}}}.$$

El algoritmo recursivo anterior puede inicializarse tomando como valor inicial para $q_{11,1}$ y $q_{22,1}$ su esperanza matemática, que es 1.Como condición inicial para $q_{ij,1}$ podemos tomar el promedio de los productos $z_{i,t} z_{j,t}$ a lo largo de toda la muestra. Esto es útil en el caso en que queremos estimar a posterior cómo ha variado dicha correlación condicional. En alguna otra situación podemos no querer imponer como condición inicial la media de toda la muestra, y preferimos utilizar el promedio de un número inicial de observaciones, 50 por ejemplo, y actualizar $q_{ij,t}$ a partir de la observación siguiente, desechando los primeros 50 datos.

Si tomamos la opción de estimar las varianzas condicionales mediante modelos GARCH univariantes, una vez estimados tales modelos, estandarizamos las rentabilidades, y definimos las variables $q_{ij,t}$ y el coeficiente de correlación de modo análogo al descrito arriba. en [EII.4.6] Se comparan las correlaciones entre los dos tipos de cambio obtenidas a partir de un esquema EWMA y de modelos univariantes asimétricos A-GARCH para cada tipo de cambio. Las diferencias entre las correlaciones resultantes son mínimas. En [Ch3.7] se analiza la correlacion condicional entre S&P500 y el tipo de cambio USdólar/Yen, así como entre S&P500 y el tipo de interés de las Treasury bills a 3 meses, mediante un esquema EWMA, estimando en ambos ejercicios el parámetro λ por máxima verosimilitud.

Para permitir reversión a la media, es generalmente preferible utilizar representaciones:

$$q_{ij,t+1} = \rho_{ij} + \alpha z_{i,t} z_{j,t} + \beta q_{ij,t} \forall i, j$$

Es decir, en el caso de dos variables:

$$\begin{aligned} q_{11,t+1} &= \rho_{11} + \alpha z_{1,t}^2 + \beta q_{11,t} \\ q_{12,t+1} &= \rho_{12} + \alpha z_{1,t} z_{2,t} + \beta q_{12,t} \\ q_{11,t+1} &= \rho_{22} + \alpha z_{2,t}^2 + \beta q_{22,t} \end{aligned}$$

con la diferencia de que ahora tenemos 5 parámetros, frente al parámetro único, λ, de la representación anterior. Para evitar problemas numéricos, podemos imponer objetivos de correlación de largo plazo, haciendo: $\bar{\rho}_{ij} = E(z_{it} z_{jt})$, y sustituyendo la esperanza matemática por la media muestral de cada uno de estos productos. En tal caso, $\bar{\rho}_{11} = \bar{\rho}_{22} = 1$, y solo tenemos dos parámetros por estimar (ver [Ch3.8], con estimación por máxima verosimilitud).

En general, para n variables, comenzamos generando unas matrices auxiliares Q_t:

$$Q_t = (1 - a - b) Q_0 + a \varepsilon_{t-1} \varepsilon_{t-1}' + b Q_{t-1}$$

siendo Q_0 la matriz de covarianzas incondicional. Sin embargo, la secuencia de matrices Q_t no son matrices de covarianzas, sino matrices auxiliares que se utilizan en la estimación de la matriz de correlaciones, mediante:

$$\rho_{ij}(t) = \frac{Q_{ij}(t)}{\sqrt{Q_{ii}(t) Q_{jj}(t)}}$$

La matriz de covarianzas H_t se genera combinando las series temporales de las varianzas condicionales, obtenidas a partir de modelos GARCH univariantes, $H_{ii}(t) H_{jj}(t)$, con las correlaciones generadas a partir de las matrices Q_t :

$$H_{ij}(t) = \sqrt{H_{ii}(t) H_{jj}(t)} \frac{Q_{ij}(t)}{\sqrt{Q_{ii}(t) Q_{jj}(t)}}$$

La opción de modelizar la correlación como constante, frente a representarla como una variable cambiante en eltiempo afecta por supuesto a muchos cálculos financieros de interés. Entre ellos, la estimación del VaR de una cartera configurada sobre dos activos, por ejemplo. En [Ch3.7] y [Ch3.8] se presenta el cálculo del VaR bajo estas dos modelizaciones alternativas. En ambos casos el cálculo del VaR se realiza mediante:

$$VaR_p = \Phi^{-1}(p) \sqrt{T} \sqrt{\omega_1^2 \sigma_{1t}^2 + \omega_2^2 \sigma_{2t}^2 + 2 \omega_1 \omega_2 \rho_{12,t} \sigma_{1t} \sigma_{2t}}$$

donde $\rho_{12,t}$ puede ser constante si hemos utilizado el modelo de correlación constante, o cambiante en el tiempo, si hemos utilizado el modelo de correlación condicional dinámica.

Modelo VECH: El modelo VECH completo permite una interacción entre las volatilidades de todas las variables, a costa de un elevado número de parámetros. Número de parámetros:[6] $k(k+1)[k(k+1)+1]/2$.[Para $k=2$, resultan 21 parámetros; para $k=3$, resultan 78 parámetros]. El modelo no es muy aconsejable para más de dos variables, debido a los problemas numéricos que pueden surgir en la estimación. El modelo se expresa:

$$vech\,(H_t) = C + A \cdot vech\,(\varepsilon_t \varepsilon_{t-1}{}') + B \cdot vech\,(H_{t-1})$$

que considera ecuaciones con una estructura similar a las del modelo BEKK:

$$\begin{pmatrix} \sigma_{1,t}^2 \\ \sigma_{12,t} \\ \sigma_{2,t}^2 \end{pmatrix} = \begin{pmatrix} \beta_{10} \\ \beta_{20} \\ \beta_{20} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \\ \beta_{31} & \beta_{32} & \beta_{33} \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t-1}^2 \\ \varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \varepsilon_{2,t-1}^2 \end{pmatrix} + \begin{pmatrix} \delta_{11} & \delta_{12} & \delta_{13} \\ \delta_{21} & \delta_{22} & \delta_{23} \\ \delta_{31} & \delta_{32} & \delta_{33} \end{pmatrix} \begin{pmatrix} \sigma_{1,t-1}^2 \\ \sigma_{12,t-1} \\ \sigma_{2,t-1}^2 \end{pmatrix}$$

$$\begin{aligned} \sigma_{1,t}^2 &= \beta_{10} + \beta_{11}\varepsilon_{1,t-1}^2 + \beta_{12}\varepsilon_{1,t-1}\varepsilon_{2,t-1} + \beta_{13}\varepsilon_{2,t-1}^2 + \delta_{11}\sigma_{1,t-1}^2 + \delta_{12}\sigma_{12,t-1} + \delta_{13}\sigma_{2,t-1}^2 \\ \sigma_{12,t} &= \beta_{20} + \beta_{21}\varepsilon_{1,t-1}^2 + \beta_{22}\varepsilon_{1,t-1}\varepsilon_{2,t-1} + \beta_{23}\varepsilon_{2,t-1}^2 + \delta_{21}\sigma_{1,t-1}^2 + \delta_{22}\sigma_{12,t-1} + \delta_{23}\sigma_{2,t-1}^2 \\ \sigma_{2,t}^2 &= \beta_{30} + \beta_{31}\varepsilon_{1,t-1}^2 + \beta_{32}\varepsilon_{1,t-1}\varepsilon_{2,t-1} + \beta_{33}\varepsilon_{2,t-1}^2 + \delta_{31}\sigma_{1,t-1}^2 + \delta_{32}\sigma_{12,t-1} + \delta_{33}\sigma_{2,t-1}^2 \end{aligned}$$

Modelo BEKK: Este modelo es un caso particular del modelo VECH. Número de parámetros:[7] $k(5k+1)/2$. [Para $k=2$, resultan 11 parámetros; para $k=3$, resultan 24 parámetros].

Garantiza que la matriz de varianzas y covarianzas sea positiva definida en cada período, mediante una estructura de evolución temporal:

$$\Sigma_t = CC' + \sum_{i=1}^{m} A_i(\varepsilon_{t-i}\varepsilon_{t-i}')A_i' + \sum_{j=1}^{s} B_j \Sigma_{t-j} B_j'$$

donde C es una matriz triangular inferior y A_i y B_j son matrices kxk. Habitualmente, con $m=1, s=1$. El modelo permite dependencia dinámica entre volatilidades, que se influyen mutuamente, pero tiene el inconveniente de que el número de parámetros crece muy rápidamente con el número de variables. El tamaño de una innovacion en una cualquiera de las variables afecta a la varianza condicional de esa variable, pero tambien a la varianza condicional de las demás variables, así como a su covarianza condicional.

Cada término en la expresión anterior es semidefinido positivo por construcción, lo que evitará entrar regiones inapropiadas de la superficie de verosimilitud. La contrapartida es que la superficie de verosimilitud es excesivamente plana,

[6] Con k variables, hay $k(k+1)/2$ elementos (varianzas y covarianzas) y, por tanto,e cuacioens en el modelo VECH. Como vemos en la representación, el número de parámetros es igual a $n+2n^2$, siendo n el número de ecuaciones. por tanto, el número de parámetros es: $k(k+1)/2 + 2\,[k(k+1)/2]^2$.

[7] Como se ve enl ar epresentación posterior, el número de parámetros es igual a $2k^2 + k(k+1)/2$, siendo k el número de variables.

dificultando la identificación de los parámetros. De hecho, la expresión muestra claramente que pueden cambiarse los signos de todos los parámetros de las matrices C, A_i, B_j, sin generar ningun efecto sobre la función de verosimilitud.

En el caso bivariante, el modelo es:

$$\begin{aligned}
\begin{pmatrix} \sigma_{1t}^2 & 0 \\ \sigma_{12,t} & \sigma_{2t}^2 \end{pmatrix} &= \begin{pmatrix} C_{11} & 0 \\ C_{12} & C_{22} \end{pmatrix} \begin{pmatrix} C_{11} & C_{12} \\ 0 & C_{22} \end{pmatrix} + \\
& \begin{pmatrix} A_{11} & A_{21} \\ A_{12} & A_{22} \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t-1}^2 & \varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \varepsilon_{1,t-1}\varepsilon_{2,t-1} & \varepsilon_{2,t-1}^2 \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} + \\
& \begin{pmatrix} B_{11} & B_{21} \\ B_{12} & B_{22} \end{pmatrix} \begin{pmatrix} \sigma_{1t}^2 & 0 \\ \sigma_{12,t-1} & \sigma_{2t}^2 \end{pmatrix} \begin{pmatrix} B_{11} & B_{11} \\ B_{21} & B_{22} \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
\sigma_{1t}^2 &= \beta_1 + \beta_{11}\varepsilon_{1,t-1}^2 + \beta_{12}\varepsilon_{1,t-1}\varepsilon_{2,t-1} + \beta_{13}\varepsilon_{2,t-1}^2 + \delta_{11}\sigma_{1t}^2 + \delta_{12}\sigma_{12,t-1} + \delta_{13}\sigma_{2t}^2 \\
\sigma_{12,t} &= \beta_2 + \beta_{21}\varepsilon_{1,t-1}^2 + \beta_{22}\varepsilon_{1,t-1}\varepsilon_{2,t-1} + \beta_{23}\varepsilon_{2,t-1}^2 + \delta_{21}\sigma_{1t}^2 + \delta_{22}\sigma_{12,t-1} + \delta_{23}\sigma_{2t}^2 \\
\sigma_{2t}^2 &= \beta_3 + \beta_{31}\varepsilon_{1,t-1}^2 + \beta_{32}\varepsilon_{1,t-1}\varepsilon_{2,t-1} + \beta_{33}\varepsilon_{2,t-1}^2 + \delta_{31}\sigma_{1t}^2 + \delta_{32}\sigma_{12,t-1} + \delta_{33}\sigma_{2t}^2
\end{aligned}$$

Modelo BEKK diagonal: Es una versión simplificada del anterior, cuando las matrices A_i, B_j tienen estructura diagonal, que excluye las interacciones entre volatilidades, que siempre pueden tener interés. Es un caso particular del modelo VECH diagonal. En el caso de dos variables, con $m = 1, s = 1$, el modelo BEKK diagonal tendría 7 parámetros, frente a los 9 parámetros del modelo VECH diagonal

$$\begin{aligned}
\begin{pmatrix} \sigma_{1t}^2 & 0 \\ \sigma_{12,t} & \sigma_{2t}^2 \end{pmatrix} &= \begin{pmatrix} C_{11} & 0 \\ C_{12} & C_{22} \end{pmatrix} \begin{pmatrix} C_{11} & C_{12} \\ 0 & C_{22} \end{pmatrix} + \\
& \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t-1}^2 & \varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \varepsilon_{1,t-1}\varepsilon_{2,t-1} & \varepsilon_{2,t-1}^2 \end{pmatrix} \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix} + \\
& \begin{pmatrix} B_{11} & 0 \\ 0 & B_{22} \end{pmatrix} \begin{pmatrix} \sigma_{1t-1}^2 & 0 \\ \sigma_{12,t-1} & \sigma_{2t-1}^2 \end{pmatrix} \begin{pmatrix} B_{11} & 0 \\ 0 & B_{22} \end{pmatrix}
\end{aligned}$$

es decir:

$$\begin{aligned}
\sigma_{1t}^2 &= C_{11}^2 + A_{11}^2\varepsilon_{1,t-1}^2 + B_{11}^2\sigma_{1t-1}^2 \\
\sigma_{12,t} &= C_{11}C_{12} + A_{11}A_{22}\varepsilon_{1,t-1}\varepsilon_{2,t-1} + B_{11}B_{22}\sigma_{12,t-1} \\
\sigma_{2t}^2 &= \left(C_{12}^2 + C_{22}^2\right) + A_{22}^2\varepsilon_{2,t-1}^2 + B_{22}^2\sigma_{2t-1}^2
\end{aligned}$$

La diferencia respecto al modelo VECH diagonal:

$$\begin{aligned}
\sigma_{1t}^2 &= C_{11} + A_{11}\varepsilon_{1,t-1}^2 + B_{11}\sigma_{1,t-1}^2 \\
\sigma_{12,t} &= C_{12} + A_{12}\varepsilon_{1,t-1}\varepsilon_{2,t-1} + B_{12}\sigma_{12,t-1} \\
\sigma_{2t}^2 &= C_{22} + A_{22}\varepsilon_{2,t-1}^2 + B_{22}\sigma_{2,t-1}^2
\end{aligned}$$

puede verse en los signos de los coeficientes de las dos ecuaciones de las varianzas, así como en el menor número de parámetros (7 frente a 9).

1.14.1 Factor GARCH models

Si consideramos un modelo factorial:

$$Y = A + XB + E$$

donde Y es una matriz Txm de T datos sobre m activos, X es Txk, con $k << m$, A es Txm siendo cada columna $j = 1, 2, ..., m$, la rentabilidad en exceso del activo j. La matriz B es kxm, siendo su elemento (i, j) la sensibilidad del activo j al factor de riesgo i-ésimo, y E es una matriz Txm de residuos. Tomando varianzas:

$$V \approx B'\Omega B$$

siendo Ω la matriz de covarianzas de los factores, V la matriz de covarianzas de los activos, y donde hemos ignorado el componente de varianza específica de cada activo. Esta expresión nos permite estimar, de manera aproximada, la matriz de covarianzas de un amplio número de activos a partir de la matriz de covarianzas de un número reducido de factores y la estimación de las sensibilidades de cada activo con respecto a cada factor. Si estimamos un modelo GARCH multivariante para los factores, supuesto que su número sea reducido, podemos aproximar:

$$V_t \approx B'\Omega_t B$$

Si los factores fuesen componentes principales, la tarea se simplifica enormemente, debido a la ausencia de correlación entre ellos.

En [EII.4.7] se realiza un ejercicio con rentabilidades de acciones y con un único factor (el indice de mercado). Este modelo es excesivamente sencillo, con:

$$\begin{aligned}
\sigma_{it}^2 &= \beta_i^2 \sigma_t^2 \\
\sigma_{ijt} &= \beta_i \beta_j \sigma_t^2 \\
\sigma_t^2 &= \omega + \alpha(\varepsilon_{t-1} - \zeta)^2 + \beta \sigma_{t-1}^2
\end{aligned}$$

Aunque las covarianzas varian en el tiempo, todas evolucionan con correlación igual a 1, y no tiene sentido calcular las correlaciones entre los activos que implica el modelo. Podriamos incorporar betas estimadas mediante EWMA, con lo que la correlación sistemática cambiaría en el tiempo.

1.14.2 Orthogonal GARCH models

El modelo GARCH ortogonal (OGARCH) es una caso particular del anterior cuando los factores son componentes principales. En tal caso:

$$P = XW \Rightarrow X \approx P^* W^{*\prime}$$

donde X es la matriz de datos originales y P es la matriz de datos de los componentes principales, mientras que P^* tiene por columnas las correspondientes a los componentes principales que hayamos escogido. La expresión de la derecha es la *representación* de los datos en función de los componentes principales, que ya conocemos.

El modelo OGARCH es especialmente apropiado para representar una estructura temporal de tipos, por la elevada correlación que muestran entre ellos. Tomando varianzas:

$$V_t \approx W^* \Omega_t W^{*\prime}$$

donde Ω_t es ahora una matriz diagonal. El modelo OGARCH requiere estimar un GARCH univariante para cada variable. Como la matriz Ω_t es siempre definida positiva, la matriz V_t definida por le modelo OGARCH también lo será:

$$x' V_t x = x' W^* \Omega_t W^{*\prime} x = y' \Omega_t y$$

donde $y = W^{*\prime} x$.

Computacionalmente, este modelo es muy eficiente. Al tender el horizonte de predicción a infinito, las predicciones de la matriz de coviarianzas que resulta del modelo OGARCH revertirán a los niveles de volatilidad medios, es decir, a una matriz de covarianzas de largo plazo.

[Case Study 4.6.2]

1.15 Algunas aplicaciones de los modelos GARCH

- *Pricing path-dependent European options:* [EII.4.8] muestra como valorar una opción Asiática utilizando un modelo GARCH. [EII.4.9] muestra cómo valorar una opción barrera a partir de un modelo GARCH para la rentabilidad del activo subyacente.
- *Calibración de parámetros del modelo GARCH a precios de opciones*
- *VaR histórico, Monte Carlo VaR, VaR lineal analítico*
- *Estimación de sensibilidades cambiantes en el tiempo,* por contraste al cálculo utilizando EWMA
- *Optimización de carteras:* [EII.4.10]

1.16 Modelos de varianza condicional y medidas de volatilidad

Las estructuras $ARCH$ tienen el atractivo de recoger, de modo bastante adecuado, la agrupación de episodios de alta volatilidad que se observa en series temporales financieras de alta frecuencia.

Foster y Nelson (1992) probaron que, incluso si la varianza no cambia a lo largo de un mes, el procedimiento de utilizar promedios de rentabilidades diarias al cuadrado como estimador de volatilidad es ineficiente y sesgado. Parkinson (1980) sugirió utilizar un estimador basado en los precios alto y bajo para aproximar la varianza de un proceso de camino aleatorio continuo, lo que se demuestra más eficiente que el uso de observaciones de final de período. Otra alternativa es el cálculo de volatilidades implícitas a través de la fórmula de valoración de opciones de BS, pero si la varianza condicional del precio de la opción es cambiante en el tiempo, no es evidente qué se obtienen de dicha fórmula. Day y Lewis (1992) muestran que para opciones sobre el índice bursátil, un modelo GARCH(1,1) simple para la varianza condicional del rendimiento implícito en el índice proporciona información estadísticamente significativa, que es adicional a las estimaciones de volatilidad implícita de BS. En esta misma línea, Engle y Mustafá (1992) probaron que ...

Buena parte de la investigación reciente [Amin y Ng (1992), Heston (1993), Hull y White (187), Melino y Turnbull (1990), Scott (1987) y Wiggins (1987)] se ha destinado a desarrollar fórmulas teóricas de valoración de opciones en presencia de volatilidad estocástica. Aunque una expresión analítica de los precios resultantes puede obtenerse sólo en algunos casos relativamente sencillos, es generalmente cierto en todos ellos que cuanto más volátil es el activo subyacente, más elevado es el precio resultante para la opción.

1.16.1 Canina, L. y S. Figlewski: "The informational content of implied volatility"

Review of Financial Studies,(1993)

En este trabajo se analiza la capacidad que tiene la volatilidad implícita obtenida a partir de opciones sobre $S\&P100$ para predecir la volatilidad futura de dicho índice. El interés del análisis se basa en el hecho de que la volatilidad implícita se interpreta, generalmente, como la *predicción del mercado* acerca del nivel de volatilidad futuro. Se encuentra que la capacidad de la volatilidad implícita para predecir la volatilidad futura es mínima, tanto cuando se trabaja con todas las opciones existentes (que garantizan una liquidez mínima y que no incumplen las relaciones básicas que deben satisfacer los precios de las opciones), como cuando se trabaja con clases de opciones, según su vencimeinto y su precio de ejercicio.

Una posible explicación de este resultado negativop sería la posible dificultad para prever la volatilidad durante el período muestral analizado en el mercado considerado. Sin embargo, cuando se utiliza la desviación típica anualizada del logartimo de las rentabilidades del $S\&P100$ sobre una ventana móvil de 60 días previos al momento de cálculo de la volatilidad implícita, se encuentra que esta medida de volatilidad histórica tiene cierta capacidad de prever la volatilidad futura. La amplitud de la ventana considerada no es crítica en estos resultados. Sin embargo, esta medida de volatilidad histórica incumple el test de *racionalidad* de la predicción.

En el trabajo se considera asimismo la posibilidad de que la volatilidad implícita incorpore la información contenida en la volatilidad recientemente observada en el mercado, rechazando asimismo dicha hipótesis.

1.16.2 Day, T.E. y C.M. Lewis, *"Forecasting futures market volatility",*

The Journal of Derivatives, winter 1993.

Se compara la capacidad predictiva de diversos métodos para anticipar la volatilidad del precio en el mercado de futuros sobre petróleo. Para ello, se calcula la volatilidad condicional resultante de un modelo $GARCH$ para el precio del futuro sobre el barril de petróleo, así como la volatilidad implícita a partir de opciones *call* sobre dicho futuro.

Se encuentra que ambas medidas contienen cierta capacidad explicativa sobre la volatilidad futura del precio del futuro. Se considera asimismo la posibilidad de utilizar un modelo $EGARCH$, pero no se detecta evidencia de efectos asimétricos en volatilidad. Como las volatilidades $GARCH$ se calculan para cada uno de los días que quedan entre el instante de valoración y el vencimiento, dichas predicciones deben consolidarse en un único nivel de volatilidad asociado al día de vencimiento de la opción. En el trabajo se utiliza un promedio simple de las volatilidades prevista para cada uno de dichos dias, pero es claro que podrían utilizarse otras alternativas. Cada día se estima el modelo $GARCH$ con una ventana móvil, se obtienen las previsiones para cada uno delos días desde el último contenido en la muestra hasta el vencimiento de la opción, y se calcula su promedio. Este procedimiento de ventana móvil permite generar una serie temporal de predicción $GARCH$ de la volatilidad para el instante (día) de vencimiento de la opción.

La capacidad predictiva de una serie temporal de volatilidades se estima mediante el ajuste de una regresión,

$$\sigma^2_{H,t+1} = b_0 + b_1\sigma^2_{F,t+1} + \eta_{t+1}$$

donde $\sigma^2_{H,t+1}$ denota la volatilidad realizada (observada) durante los períodos desde que se calcula la predicción hasta el vencimiento de la opción, y $\sigma^2_{F,t+1}$ es la predicción de dicha volatilidad, calculada con información hasta el instante t.

Las volatilidades $GARCH$ y $EGARCH$ incumplen la propiedad de reacionalidad, mientras que las volatilidades implícitas satisfacen dicha propiedad. La cuarta medida utilizada es una medida *ingenua*, pero sus resultados son peores que los obtenidos con la volatilidad implícita. Una peculiaridad no discutida en el trabajo es que el modelo $GARCH$ que se utiliza es un modelo con componente en media, siendo un $GARCH(1,1) - M$, mientras que el modelo $EGARCH$ no tiene tal componente, incorporando en cambio una estructura $AR(1)$ en rentabilidad.

Se muestra asimismo que las predicciones extra-muestrales a partir de modelos tipo $GARCH$ no contienen información que no estuviese ya recogida en la serie temporal de volatilidad implicita. Las predicciones que se obtienen llevando a cabo ajustes de sesgo predictivo o combinando predictores no tienen un comportamiento significativamente mejor que la volatilidad implícita sin ajustar.

Para calcular la volatlidad implícita se utiliza la técnica de árbol binomial, dado que las opciones sobre futuro de barril de petróleo permiten el ejercicio anticipado, al ser opciones Americanas.

1.16.3 Day, T.E. y C.M. Lewis, "Stock market volatility and the information content of stock index options"

Journal of Econometrics (1992), 52:267-287.

Este trabajo compara la capacidad predictiva de modelos *GARCH* y *EGARCH* estimados para el índice S&P100 acerca de la volatilidad futura del exceso de rentabilidad ofrecido por el índice. A pesar de disponer de observaciones diarias, el trabajo se lleva a cabo con rentabilidades semanales. Para evitar (a la vez que estimar) posibles efectos dia de la semana, el estudio se realiza tanto con los datos correspondientes a los miércoles, como con los correspondientes a los viernes. Otra razón para ello es evitar la autocorrelación existente en datos diarios que, aparentemente surge por problemas de negociación no simultánea en el índice (*nonsynchronous trading*). Se utilizan dos series de rentabilidad semanal: por un lado, las rentabilidades (sin ajustar), que se obtienen de los datos de cierre; por otro, las estimaciones del nivel del índice implícito en el precio de opciones *call* sobre dicho índice. Sólo se reportan los resultados obtenidos con la primera de las medidas.

Se eliminan los precios diarios de opciones con menos de 100 contratos de negociación, o aquellas cuyo precio a cierre de mercado difiere sustancialmente del precio de ejercicio (en más de $15). Se eliminan asimismo las opciones conprecio de mercado muy reducido (inferior a 0,25$), porque la horquilla *bid-ask* es entonces un porcentaje muy elevado del precio de la opción. Los dividendos efectivamente pagados a posteriori se toman como *proxy* de las expectativas de dividendos a recibir durante la vida de la opción. Para estimar la volatilidad implícita se utiliza la fórmula Black-Scholes ajustada de una tasa constante de dividendos.

Las predicciones obtenidas a partir de volatilidades condicionales deducidas de modelos *GARCH* y *EGARCH* se comparan con la volatilidad implícita, interpretada como estimador de la volatilidad del índice a vencimiento de la opción.

Resumir los precios observados para todas las opciones negociadas sobre el índice en determinado momento requiere cierto trabajo estadístico. Si denotamos por $C_k(\sigma_0(\tau))$ el precio teórico de una opción con tiempo a vencimiento τ y precio de ejercicio indicado por k, dada una estimación $\sigma_0(\tau)$ de la volatilidad de la rentabilidad del índice hasta el instante de vencimiento de la opción, construimos la función objetivo,

$$F_\tau = \sum_{k=1}^{N_\tau} \left[\delta_{k\tau}\left(C_{k\tau} - C_k(\sigma_0(\tau))\right)\right]^2$$

donde $\delta_{k\tau}$ denota la proporción del volumen de negociación que se lleva a cabo en opciones con vencimiento τ que corresponde al contrato con precio de ejercicio k, y N_τ es el número de precios de ejercicio diferentes de opciones con vencimiento τ. Por tanto, la función de pérdida anterior se asocia al vencimiento τ.

En cada iteración, la nueva estimación de la volatilidad viene dada por,

$$\sigma(\tau) = \sigma_0(\tau) + \left[(\Omega X)'(\Omega X)\right]^{-1}(\Omega X)'(\Omega Y)$$

siendo Ω la matriz diagonal $N_\tau x N_\tau$ que tiene por elementos el porcentaje de volumen de negociación en las opciones *call* con vencimiento τ en cada uno de los precios de ejercicio negociados, X es un vector $N_\tau x1$ que contiene las derivadas parciales de los precios de las opciones *call* respecto de la volatilidad del subyacente, $\sigma_0(\tau)$. Por último, Y es un vector $N_\tau x1$ cuyos elementos son las diferencias entre precios teóricos, calculados con la estimación $\sigma_0(\tau)$ dela volatilidad, y los precios de mercado.

La volatilidad histótica con cuyas realizaciones se comparan las predicciones de volatilidad se calcula de dos maneras diferentes: a) mediante el cuadrado de la rentabilidad semanal, y b) mediante la varianza de las rentabilidades diarias, multiplicada por el número de dias de negociación en dicha semana.

Los resultados apuntan a que la volatilidad implícita contiene información no contenida en la volatilidad condicional que se deriva de los modelos *GARCH* y *EGARCH*. Se obtiene asimismo el resultado dual: la volatilidad condicional que surge de los modelos *GARCH* y *EGARCH* contiene información adicional a la incorporada en la volatilidad implícita. Por tanto, ambas deben combinarse al predecir la volatilidad futura del índice. La volatilidad condicional obtenida a partir del modelo *EGARCH* no contiene información significativa que no esté ya incorporada en la volatilidad condicional del modelo *GARCH*.

Se utiliza un procedimientoo de ventanas móviles, al igual que en el trabajo anterior, para obtener predicciones de volatilidad a partir de modelos *GARCH* y *EGARCH*. Las estimaciones de estos modelos parecen ser insesgadas, al contrario de lo que ocurre con las volatilidades implícitas. Sin embargo, la volatilidad semanal parece difícil de predecir, y los ajustes entre predicción de volatilidad y volatilidad observada futura no presentan valores muy altos del estadístico R^2. Los modelos *GARCH* parecen ofrecer mejores resultados que los modelos *EGARCH*.

1.16.4 Engle, R.F., y C. Mustafa: *"Implied ARCH models from option prices"*:

Estimación de los procesos estocásticos que para la volatilidad de un activo se deducen de los precios de las opciones que tienen a dicho activo como subyacente. Se supone que dicha volatilidad responde a una representación del tipo *GARCH*. Se propone estimar el modelo *GARCH* mediante un procedimiento de minimización de los cuadrados de los errores en precio; para ello, partiendo de unas pre-estimaciones para los parámetros del modelo *GARCH*, se obtiene por *simulación* el precio de la opción, y se compara con su precio de mercado. Inicialmente, se toma como función objetivo la suma de los cuadrados de las diferencias en precio, que se minimiza, iterando en el espacio de parámetros $GARCH$. La propuesta se generaliza en la forma de un método de mínimos cuadrados generalizados, ponderando los errores cometidos en el precio de cada opción de manera inversa a la precisión con que se calcula su precio teórico.

Conocer la persistencia de la volatilidad es importante para los agentes que operan en un mercado de derivados, en el que estarán dispuestos a pagar un precio más alto poropciones de largo plazo si perciben que los shocks actuales en volatilidad son altos y permanentes, en relación con la vida residual de la opción.

Se obtiene que la persistencia de los shocks de volatilidad que se obtiene en el modelo que se infiere a partir de precios de opciones del $S\&P500$es similar a la que se estima a partir de datos históricos sobre el índice. Sin embargo, la persistencia después del *crash* bursátil de 19 Octubre 2987 se estima como significativamente más débil.

1.16.5 Noh, J., R.F. Engle, y A. Kane, "Forecasting volatility and option prices of the S&P500 index"

Journal of Derivatives, (1994), 17-30.

Este trabajo compara la capacidad de la volatilidad implícita, con la de la volatilidad condicional, paraanticipar la volatildad futura de la rentabilidad ofrecida por el índice $S\&P500$. La comparación se efectúa mediante la gestión de una cartera de *straddles*, llevada a cabo utilizando las predicciones de volatilidad proporcionadas por ambos métodos.

El modelo $GARCH$ se supone que rep resenta la volatilidad condicional del error de un modelo $AR(1)$ para la rentabilidad del índice $S\&P500$. En el modelo de volatilidad condicional se incorpora una corrección por el número de dias naturales transcurridos entre dos días sucesivos de negociación.

Los procedimientos para resumir en una única medida de volatilidad las predicciones dle modelo $GARCH$, así como para trabajar con todas las opciones negociadas, son los mismos que se han expuesto para otros trabajos. Una aportación de este trabajo es obtener predicciones de volatilidad implícita mediante relacioens lineales de la misma sobre sus valores previos, obtenidos tanto a partir de opciones *put* como de opciones *call*, de la rentabilidad pasada del mercado, y de 2 variables ficticias que tratan de incorporar el *efecto día de la semana* que se ha observado en volatilidad los lunes y viernes.

Se construyen *straddles* con al menos 15 dias hasta vencimiento, y una negociación mínima de 100 contratos diarios. Se considera cada día el *straddle* con precio de ejercicio más cercano al índice $S\&P500$. Se usa el tipo de interés en Letras del Tesoro a un mes de vencimiento como tipos de interés sin riesgo. Si la predicción del precio del *straddle* es superior al precio de mercado, se compra dicho *straddle*. En caso contrario, se vende. Invertimos $100 en el contrato más at-the money. Cuando se vende un *straddle*, se invierte el dinero en el activo sin riesgo.

En sucesivas repeticiones del ejercico de simulación, se aplican filtros, comprando o vendiendo únicamente si la diferencia entre precio teórico y precio de mercado es superior a \$0,05 o \$0,25 sin que esto afecte al resultado fundamental, que es que las predicciones de volatilidad del modelo $GARCH$ generan beneficios significativos, superiores a los costes de transacción.

1.16.6 French, K.R., G.W. Schwert, y R.F. Stambaugh, "Expected stock returns and volatility"

Journal of Financial Economics (1987), 19, 3-29.

En este trabajo se examina la relación entre rentabilidades de activos de renta variable, y la volatilidad del mercado. Se encuentra evidencia favorable a lque la prima de riesgo esperada (definida como diferencia entre la rentabilidad esperada de una determinada cartera de renta variable y la rentabilidad ofrecida por una cartera de letras del Tesoro), depende positivamente del componente predecible de la volatilidad de la rentabilidad del activo de renta variable.

Se encuentra asimismo una relación negativa entre el componente no anticipado de la rentabilidad del mercado y el cambio no anticipado en la volatilidad de la rentabilidad. Este resultado proporciona evidnecia indirecta a favor de una relación positiva entre primas de riesgo y volatilidad.

1.17 Referencias

1.17.1 Libros:

Mills, T.C., The econometric modelling of financial time series, Cambridge U. Press, 1993

Taylor, S., Modelling financial time series, Wiley, Nueva York, 1986.

Novales, A., Econometría, McGraw-Hill, 1993, 1996.

Campbell, J.Y., A.W. Lo, y A.C.MacKinlay, The Econometrics of financial markets, Princeton U. Press, 1997.

1.17.2 Artículos:

Bollerslev, T., R.F. Engle y J.M. Wooldridge, A capital asset pricing model with time-varying covariances, Journal of Political Economy, 96, 1, 116-132, 1988.

Engle, R.F. y M. Rothschild, ARCH models in Finance, Journal of Econometrics, 52, 5-59, 1992.

Bollerslev, T., R.F. Engle y D.B. Nelson, ARCH models, The Handbook of Econometrics, vol.4, capítulo 11, 1994.

Ruiz, E., Modelos para series temporales heterocedásticas,Cuadernos económicos de ICE, 1994.

Engle, R.F., T. Ito, y W.L. Lin, Meteor showers or heat waves? Heteroskedastic intra daily volatility in the foreign exchange market, Econometrica, 58, 525-542, 1990.

Engle,R. y T.Bollerslev, 1986, "Modelling the persistence of conditional variances", Econometric Reviews, 5, 1-50.

Bollerslev,T., R.Y.Chou y K.F.Kroner, 1992, "ARCH modeling in finance: A review of the theory and empirical evidence", Journal of Econometrics, 52, 5-59.

Engle, R.F., 1982, "Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation", Econometrica, 50, 987-1008.

Engle, R.F., D.Lilien y R.Robins, 1987, "Estimating time varying risk premia in the term structure: the ARCH-M model", Econometrica, 55, 391-408.

Engle, R.F., 1982, Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation, Econometrica, 50, 4.

Bollerslev, T., 1986, Generalized autoregressive conditional heteroskedasticity, Journal of Econometrics, 31.

1.17.3 1^a Parte: Estructura temporal de volatilidades. Evidencia empírica desde los mercados.

Bessembinder, Coughenour, Seguin, Smoller: "Is there a term structure of volatilities? Reevaluating the Samuelson hypothesis", The Journal of Derivatives, winter 1996, 45-58.

Heynen, Kemna, Vorst, "Analysis of the term structure of implied volatilities", The Journal of Business, v.29, 1994,

Xu y Taylor, "The term structure of volatility implied by foreign exchange options", Journal of Financial and Quantitative Analysis, 1994.

1.17.4 2^a Parte: Transmisión de volatilidades entre mercados

Koutmos y Tucker, "Temporal relationships and dynamic interactions between spot and futures stock markets", Journal of Futures Markets, 1996

Iihara, Hato y Tokunaga, "Intraday return dynamics between the cash and futures markets in Japan", Journal of Futures Markets, 1996

1.17.5 3^a Parte: Implicaciones para la cobertura de carteras.

Myers, "Estimating time-varying optimal hedge ratioson futures markets", Journal of Futures Markets, 1991.

Engle y Chowdhury, "Implied ARCH models from option prices", Journal of Econometrics, 1992.

Noh, Engle y Kane, "Forecasting volatility and option prices of the S&P 500 index", Journal of Derivatives, 1994.

Lien y Luo, "Multiperiod hedging in the presence of conditional heteroskedasticity", Journal of Futures Markets, 1994