Capítulo 15

Modelos de Heterocedasticidad condicional

Robert F. Engle (19

Económetra estadounidense. Profesor del Departamento de Economía en la Universidad de California en San Diego. Desarrolló con Granger el concepto de cointegración e inventó los procesos ARCH. Uno de los creadores del area de estudio de la la econometría financiera, su trabajo ha influido en muchas areas del análisis de series temporales y la econometría. En 2003 recibió el Premio Nobel de Economía.

15.1 Introducción

Existe considerable experiencia empírica que las series temporales estacionarias de datos financieros y ambientales de alta frecuencia, como los datos diarios o horarios, presentan algunas características que no pueden explicarse por los modelos ARIMA. Normalmente estas series tienen poca estructura en la media y muchas de ellas siguen paseos aleatorios o procesos AR de orden bajo y coeficiente pequeño. Como ejemplo, consideremos los rendimientos diarios del indice Ibex 35 medidos por la variable $z_t = 100(\log X_t - \log X_{t-1})$, donde X_t es el indice Ibex 35 de la bolsa de Madrid. La figura 15.1 (a) incluye 1982 observaciones de esta series entre Enero de 1992 y Diciembre de 1999. Se observa que la serie tienen poca estructura y que presenta periodos de mayor varianza. Por ejemplo, la última parte de la serie tiene mayor varianza que la parte central. La figura 15.1 (b) presenta la función de autocorrelación de estos datos. Existe un débil autodependencia, ya que $r_1 = .11$ es significativamente distinto de cero, pero aparte de este primer coeficiente la serie no tiene estructura. Hay algún valor que sale fuera de los límites del 95% pero esto es esperable al considerar 50 coeficientes. La figura 15.1 (c) presenta el histograma de estos datos y se observa una distribución simétrica aunque con colas más pesadas que la normal: el coeficiente de kurtosis de estos datos es 6.88, más del doble de la distribución normal. Finalmente, la figura 15.1 (d) muestra la función de autocorrelación de los cuadrados de las observaciones. La función tiene muchos coeficientes distintos de cero con una fuerte estructura de dependencia.

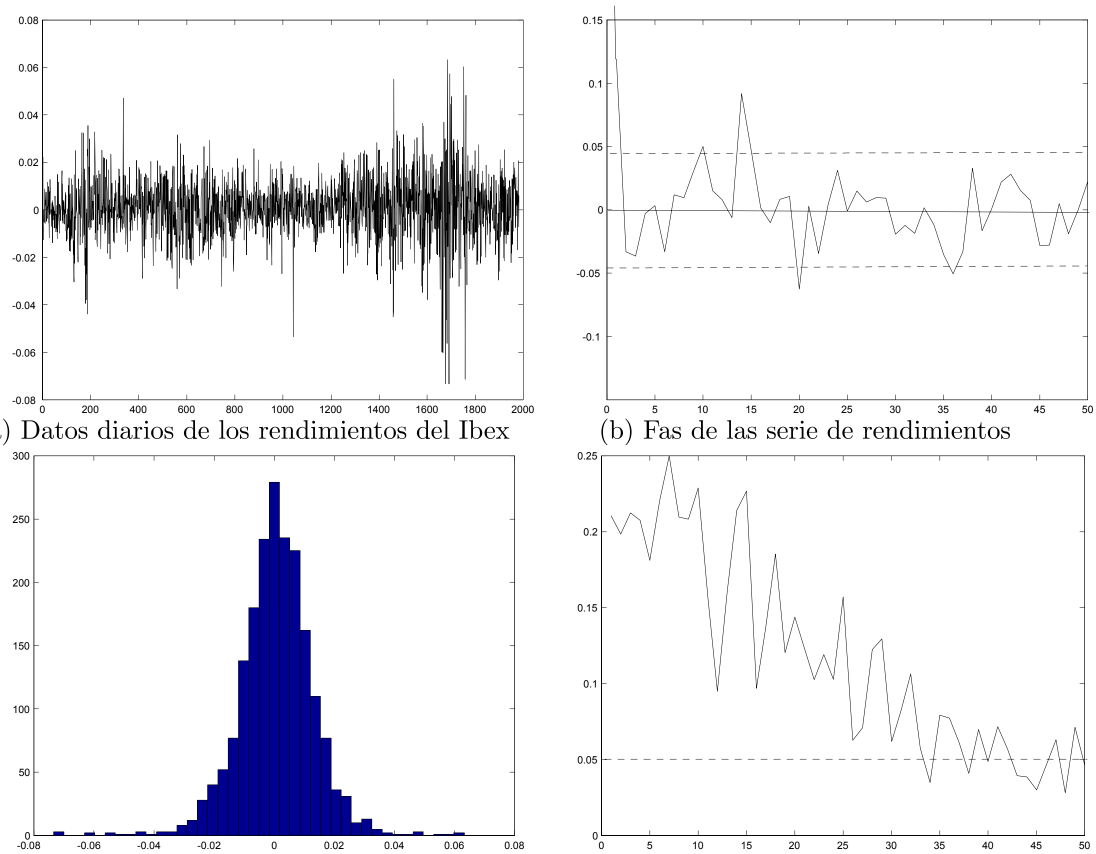
Podemos concluir que aunque la serie de rendimientos parece casi ruido blanco:

1. Su distribución no es normal, y muestra colas pesadas y alta kurtosis.

2. Los datos están casi incorrelados pero al calcular las autocorrelaciones de los cuadrados se observa un fuerte estructura de dependencia.

3. La varianza de los residuos no es constante y aparecen rachas de mayor variabilidad seguidas por otras de menor variabilidad. Esta característica, que se observa en el gráfico de los datos, se aprecia más claramente en la figura 15.2 (a), que presenta la serie de los cuadrados de los rendimientos, y en la figura 15.2 (b), que presenta las varianzas calculadas en grupos de 100 datos de la serie.

Se han propuesto distintos modelos para explicar estos efectos. Llamando $W_{t-1} = (w_{t-1}, ..., w_1)$ al conjunto de valores de la serie estacionaria w_t hasta el instante t, y μ_t y σ_t^2 a su media y varianza condicionada



(a) Datos diarios de los rendimientos del Ibex
(b) Fas de las serie de rendimientos
(c) Histograma de la serie de rendimientos
(d) Fas de los cuadrados de los datos

Figura 15.1: Analisis de la serie de ibex de rendimientos diarios

a estos valores pasados, los modelos ARMA suponen que la media condicional sigue un proceso lineal ARMA del tipo

$$\mu_t = E(w_t|W_{t-1}) = \sum \pi_i(w_{t-i} - \mu)$$

siendo μ la media marginal del proceso estacionario, y que las innovaciones:

$$e_t = w_t - E(w_t|W_{t-1})$$

tienen una varianza condicionada constante. Los modelos de *heterocedasticidad condicional* que vamos a estudiar en este capítulo, suponen que la varianza condicional no es constante, como los modelos ARIMA, sino que depende de los valores previos de la serie,W_{t-1}, y de otras variables no observadas. Es decir

$$Var(e_t|W_{t-1}) = f(W_{t-1}, \xi_t) \tag{15.1}$$

A los modelos con esta propiedad se les denomina con *varianza condicional heterocedástica* (no constante) o modelos de heterocedasticidad condicional.

Los modelos de heterocedasticidad condicional pueden clasificarse de acuerdo con la estructura que establecen para la varianza condicional (15.1). Si suponen que la dependencia es sólo de los valores pasados del proceso se obtiene la familia de procesos ARCH y GARCH. Si permitimos además la dependencia de otras variables no observadas, se obtienen los modelos de volatilidad estocástica.

Los modelos de dependencia más simple son los modelos ARCH (iniciales de AutoRegressive Conditional Heteroscedastic) introducidos por Engle (1982) que suponen una dependencia del pasado autorregresiva, y los modelos ARCH generalizados o models GARCH de Bollerslev (1986), que incorporan también términos de media móvil en la depencia (15.1). Después de estudiar estos modelos presentaremos una clase de modelos

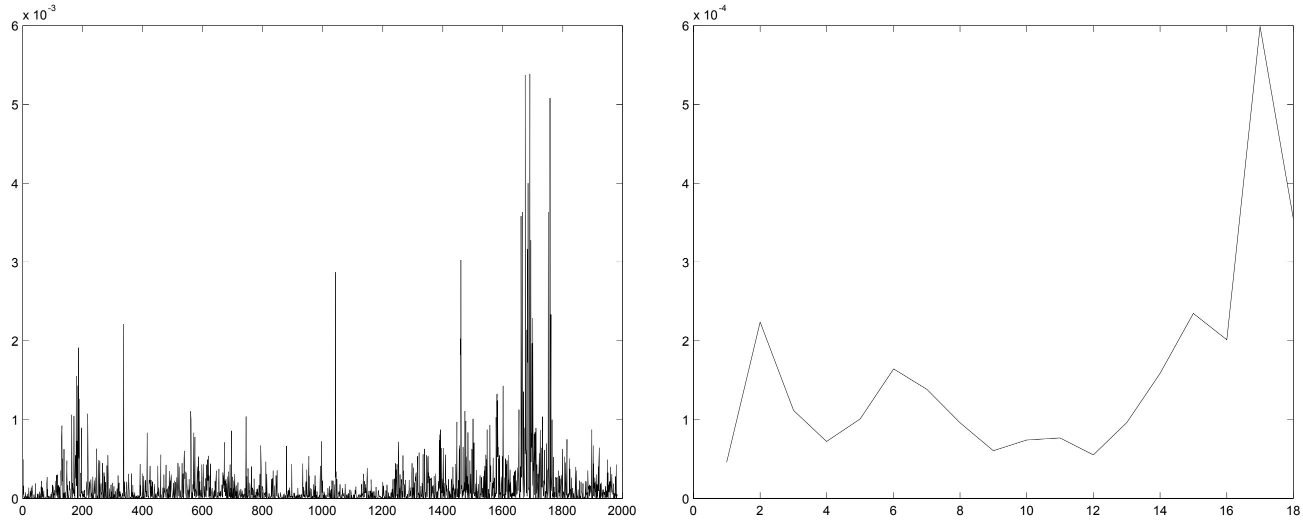


Figura 15.2: Cuadrados de los rendimientos del ibex y variación de la variabilidad

más flexibles que los anteriores, los modelos de volatilidad estocástica, introducidos por Harvey, Ruiz and Shephard (1994), y Jacquier, Polson and Rossi (1994), que incorporan un término de ruido no observado en la relación entre la varianza condicional y su pasado. Ambos tipos de modelos suponen efectos simétricos, es decir, la dependencia es de los cuadrados de las observaciones pasadas y no depende del signo de éstas. Esto puede no ser cierto y en la última sección mecionaremos brevemente otros modelos que se han propuesto para incluir estos y otros efectos adicionales.

15.2 Modelos ARCH

Los modelos ARCH son modelos para las innovaciones e_t de una serie observada. Suponemos que se ha modelado la media condicional de la serie mediante un modelo ARMA, o no lineal, pero que las innovaciones del proceso pueden escribirse como:

$$e_t = \sigma_t \epsilon_t \tag{15.2}$$

donde ϵ_t y σ_t son dos procesos estacionarios independientes entre sí. El proceso ϵ_t es de ruido blanco estandarizado, es decir, formado por variables independientes de media cero y varianza unidad. Supondremos que la distribución de este proceso es normal, aunque todo el análisis puede generalizarse suponiendo otra distribución, como la t de Student. El proceso σ_t es estacionario y con estructura dinámica, siendo su valor en t función del conjunto $E_{t-1} = (e_{t-1}, ..., e_1)$ de las innovaciones previas a t. La independencia entre ϵ_t y σ_t implica que las innovaciones e_t forman una secuencia de variables con media marginal igual a cero, ya que:

$$E(e_t) = E(\sigma_t)E(\epsilon_t) = 0,$$

y también con media condicional nula, ya que,

$$E(e_t|E_{t-1}) = E(\sigma_t|E_{t-1})E(\epsilon_t) = 0.$$

Las innovaciones e_t tienen varianza marginal constante, que llamaremos σ^2 y viene dada por:

$$E(e_t^2) = E(\sigma_t^2)E(\epsilon_t^2) = E(\sigma_t^2) = \sigma^2 \tag{15.3}$$

ya que $E(\epsilon_t^2) = 1$. Por otro lado, la varianza condicionada es:

$$E(e_t^2|E_{t-1}) = E(\sigma_t^2|E_{t-1})E(\epsilon_t^2) = \sigma_t^2$$

ya que $E(\epsilon_t^2|E_{t-1}) = E(\epsilon_t^2) = 1$. Las innovaciones e_t van a estar, además, incorreladas, debido a al independencia entre los procesos ϵ_t y σ_t. En efecto, las covarianzas entre las innovaciones son

$$E(e_t e_{t-k}) = E(\sigma_t \sigma_{t-k})E(\epsilon_t)E(\epsilon_{t-k}) = 0.$$

y aunque las covarianzas entre las σ_t no serán en general cero, las del proceso ϵ_t sí lo serán, y, por la independencia entre ambos, el proceso de las innovaciones e_t tendrá covarianzas nulas. Sin embargo, a diferencia de los modelos ARIMA estas innovaciones no serán independientes. En efecto, si e_t y e_{t-1} son función de valores comunes pasados de la serie aunque estén incorreladas no serán independientes.

En resumen, los modelos ARCH se aplican a procesos e_t con media condicional y marginal cero, varianza marginal constante pero varianza condicionada no constante y que están formados por variables incorreladas pero no independientes. La varianza condicionada viene dada por el proceso σ_t^2, que tiene una media distinta de cero y que, según la ecuación (15.3), coincide con la varianza marginal de las innovaciones.

Vamos a estudiar el comportamiento de este modelo en los casos más simples.

15.2.1 El modelo ARCH(1)

El modelo ARCH(1) supone que la varianza condicional de las innovaciones, σ_t^2, tiene una estructura similar a un AR(1), y depende sólo de la última innovación, mediante la ecuación:

$$E(e_t^2|E_{t-1}) = \sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 \tag{15.4}$$

donde $\alpha_0 > 0$ and $\alpha_1 \geq 0$. Esta ecuación establece que si el valor de e_{t-1}^2 es alto, la varianza de la siguiente innovación será también alta, haciendo más probable que el valor de $e_t^2 = \sigma_t^2 \; \epsilon_t^2$ sea alto. Esto explica la aparición de correlaciones entre los cuadrados de las innovaciones. Por otro lado, este modelo tiene la capacidad de producir rachas de valores con alta varianza. En efecto, supongamos que los valores de e_t^2 son pequeños, entonces la varianza condicional será sólo algo mayor de α_0. Sin embargo, si aparece un valor alto de e_t^2, la varianza del siguiente valor será grande, aumentando la probabilidad de otro valor alto de e_t^2. Por otro lado, como la media marginal es cero, aunque la varianza sea alta siempre es posible que aparezca un valor pequeño de e_t^2, que disminuirá la varianza marginal de la observación siguiente y facilitará que la siguiente observación sea pequeña. De esta manera las innovaciones presentan rachas de valores altos, pero globalmente forman un proceso estacionario.

Es intuitivo que la varianza marginal de la serie, que es en definitiva el promedio de las varianzas condicionales, debe ser mayor que α_0 y será tanto mayor cuanto mayor sea el coeficiente α_1 que transmite el efecto de la última observación. En efecto, llamando $\sigma^2 = E(e_t^2)$ a la varianza marginal, podemos escribir:

$$\sigma^2 = E[E(e_t^2|E_{t-1})] = \alpha_0 + \alpha_1 E(e_{t-1}^2)$$

donde hemos utilizado que podemos obtener la esperanza de una variable calculando primero, la esperanza condicionada a un conjunto de información y segundo, tomando la esperanza respecto a ese conjunto de información. Como e_t es un proceso estacionario, tenemos que, tomando esperanzas en (15.4), como $E(e_t^2) = E(e_{t-1}^2) = \sigma^2$, se obtiene:

$$\sigma^2 = \frac{\alpha_0}{1-\alpha_1} \tag{15.5}$$

y por lo tanto $0 \leq \alpha_1 < 1$.

La autocorrelación de los cuadrados

El modelo ARCH(1) puede también escribirse poniendo de manifiesto una dependencia de tipo AR(1) de los cuadrados de las observaciónes. Para ello, definamos un nuevo proceso de variables estacionarias incorreladas:

$$v_t = e_t^2 - \sigma_t^2 = \sigma_t^2(\epsilon_t^2 - 1) \tag{15.6}$$

Estas variables tienen media cero, ya que,

$$E(v_t) = E(e_t^2) - E(\sigma_t^2) = \sigma^2 - \sigma^2 = 0,$$

y varianza constante, ya que

$$var(v_t) = E(e_t^2 - \sigma_t^2)^2 = E(e_t^4) + E(\sigma_t^4) - 2E(e_t^2\sigma_t^2)$$

y como $e_t = \sigma_t \epsilon_t$, y los procesos σ_t y ϵ_t son independientes, la esperanza de $E(\epsilon_t^2\sigma_t^4)$ será el producto de las esperanzas $E(\epsilon_t^2)$ y $E(\sigma_t^4)$. Como por la normalidad de ϵ_t se verifica que $E(\epsilon_t^4) = 3$, tenemos que:

$$var(v_t) = 2E(\sigma_t^4)$$

y como σ_t es un proceso estacionario sus momentos son constantes, y hemos demostrado que el proceso v_t es homocedástico.

Vamos a comprobar ahora que las variables v_t estan incorreladas, escribamos

$$E(v_t v_{t-k}) = E\left[\sigma_t^2({\epsilon_t}^2 - 1)\sigma_{t-k}^2({\epsilon_{t-k}}^2 - 1)\right] = E({\epsilon_t}^2 - 1)E(\sigma_t^2\sigma_{t-k}^2)E({\epsilon_{t-k}}^2 - 1)) = 0$$

por la independencia de los procesos ${\epsilon_t}^2$ y σ_t^2 y aplicando que $E({\epsilon_{t-k}}^2 - 1)) = 0$. Utilizando este nuevo proceso de pseudo innovaciones, formado por variables incorreladas, podemos escribir la relación entre los cuadrados de las innovaciones originales utilizando (15.4) y (15.6):

$$e_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + v_t \tag{15.7}$$

que establece una dependencia markoviana de primer orden entre los cuadrados de las innovaciones. Esta ecuación es similar a un proceso AR(1) sobre los cuadrados de las innovaciones, aunque presenta la diferencia básica de que

(1) las innovaciónes v_t están incorreladas, pero no son independientes;

(2) la innovación de esta ecuación v_t está incorrelada, pero no es independiente, de las variables explicativas e_{t-1}^2.

En efecto, las innovaciones v_t no van a ser independientes ya que, por su definición (15.6), v_t depende de σ_t^2, y v_{t-1} depende de e_{t-1}^2 que están relacionadas. Por tanto v_t depende de e_{t-1}^2. Por ejemplo, como en (15.7) e_t^2 debe ser no negativo, tenemos que

$$v_t \geq -\alpha_0 - \alpha_1 e_{t-1}^2$$

Sin embargo es fácil comprobar que v_t está incorrelada con e_{t-1}^2. La covarianza es

$$E(v_t e_{t-1}^2) = E(\sigma_t^2(\epsilon_t^2 - 1)\sigma_{t-1}^2\epsilon_{t-1}^2) = E({\epsilon_t}^2 - 1)E(\sigma_t^2\sigma_{t-1}^2) = 0$$

Por tanto, la ecuación (15.7) establece una dependencia de tipo AR(1) sobre los cuadrados de las observaciones, ya que descompone una variable en dos componentes ortogonales.

Vamos a calcular la estructura de autocorrelación de los cuadrados de las innovaciones. Llamando

$$\gamma_c(k) = E(e_t^2 e_{t-k}^2) - E(e_t^2)E(e_{t-k}^2)$$

a la función de autocovarianzas de los cuadrados de las innovaciones del proceso, para $k = 1$ sustituyendo $e_t^2 = \sigma_t^2\epsilon_t^2$, y utilizando (15.4) y (15.5), se obtiene:

$$\gamma_c(1) = E((\alpha_0 + \alpha_1 e_{t-1}^2)\epsilon_t^2 e_{t-1}^2) - \alpha_0^2/(1-\alpha_1)^2$$

y como las variables ϵ_t^2 y e_{t-1}^2son independientes y $E(\epsilon_t^2) = 1$, resulta que

$$\gamma_c(1) = \alpha_0^2/(1-\alpha_1) + \alpha_1 E(e_{t-1}^4) - \alpha_0^2/(1-\alpha_1)^2$$

y dividiendo por $\gamma_c(0) = E(e_{t-1}^4) - \alpha_0^2/(1-\alpha_1)^2$ tenemos que

$$\rho_c(1) = \frac{\gamma_c(1)}{\gamma_c(0)} = \alpha_1 \tag{15.8}$$

Esta ecuación indica que el parámetro α_1 es el coeficiente de correlación entre los cuadrados de las innovaciones. Para otros retardos, tenemos que

$$\gamma_c(k) = E((\alpha_0 + \alpha_1 e_{t-1}^2)\epsilon_t^2 e_{t-k}^2) - \alpha_0^2/(1-\alpha_1)^2$$

y operando como antes, se obtiene

$$\gamma_c(k) = \alpha_1 \gamma_c(k-1)$$

con lo que

$$\rho_c(k) = \alpha_1 \rho_c(k-1)$$

La conclusión de este ejercicio es que los cuadrados de las innovaciones están ligados por una ecuación autorregresiva de primer orden.

Kurtosis

Vamos a demostrar que el modelo ARCH(1) conducirá a distribuciones con colas pesadas. En consecuencia esta distribución tendrá una probabilidad mayor que la normal de generar datos atípicos. El coeficiente de curtosis marginal será

$$K = \frac{E(e_t^4)}{Var^2(e_t)} \tag{15.9}$$

Como la variable ϵ es normal, la variable $e_t|W_{t-1}$ será también normal y tendrá kurtosis igual a tres. Entonces,

$$E(e_t^4) = E[E(e_t^4|W_{t-1})] = E(3\sigma_t^4) \tag{15.10}$$

y elevando al cuadrado y tomando esperanzas en (15.4) :

$$E(\sigma_t^4) = \alpha_0^2 + 2\alpha_0\alpha_1 E(e_{t-1}^2) + \alpha_1^2 E(e_{t-1}^4)$$

y sustituyendo en (15.10), como $E(e_t^4) = E(e_{t-1}^4)$ tenemos que

$$E(e_t^4) = \frac{3\alpha_0^2 + 6\alpha_0\alpha_1 E(e_{t-1}^2)}{(1-3\alpha_1^2)}. \tag{15.11}$$

Observemos que esta ecuación impone una restricción adicional sobre α_1 ya que como el cociente debe ser positivo y el numerador siempre lo es, el denominador debe ser positivo lo que implica $0 \leq \alpha_1 < 1/\sqrt{3}$. Si sustituimos esta expresión en (15.9) y utilizando la expresión de $E(e_{t-1}^2)$ en (15.5), tenemos finalmente que

$$K = \frac{3\alpha_0^2(1-\alpha_1)^2 + 6\alpha_0^2\alpha_1(1-\alpha_1)}{\alpha_0^2(1-3\alpha_1^2)} = 3\frac{1-\alpha_1^2}{1-3\alpha_1^2}$$

Como $\alpha_1 > 0$ este coeficiente de kurtosis es siempre mayor que 3, y puede ser mucho mayor si α_1 esta próximo a su valor límite $1/\sqrt{3}$. Por lo tanto, la distribución marginal resultante será de colas pesadas.

15.2.2 El modelo ARCH(r)

El análisis anterior puede generalizarse permitiendo una dependencia de la varianza condicional con r retardos. Se obtiene así el proceso ARCH(r), para $e_t = \sigma_t \epsilon_t$, donde ahora

$$\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \cdots + \alpha_r e_{t-r}^2 \tag{15.12}$$

En este proceso las posibilidades de rachas de alta volatilidad dependen no sólo del último valor observado sino de los r últimos valores. La varianza marginal será

$$Var(e_t) = E(e_t^2) = E[E(e_t^2|E_{t-1})] = \alpha_0 + \sum \alpha_i E(e_{t-i}^2)$$

que implica

$$Var(e_t) = \frac{\alpha_0}{1-\alpha_1 - ... - \alpha_r}$$

lo que exige la restricción $\sum \alpha_i < 1$.

Dependencia de los cuadrados

Si introducimos como en el caso del proceso ARCH(1) la variable $v_t = e_t^2 - \sigma_t^2$, podemos expresar la dependencia de los cuadrados de las observaciones como un proceso AR(r):

$$e_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \cdots + \alpha_r e_{t-r}^2 + v_t \tag{15.13}$$

donde la secuencia v_t tiene las mismas propiedades que en el ARCH(1): variables incorreladas de medio cero, varianza constante e incorreladas con los regresores. Por otro lado estas variables no son independientes entre sí ni de los los regresores ya que la positividad de e_t^2 exige que

$$v_t > -(\alpha_0 + \alpha_1 e_{t-1}^2 + \cdots + \alpha_r e_{t-r}^2)$$

La ecuación (15.13) indica que existirá dependencia entre los cuadrados de las observaciones, y podríamos haber utilizado esta expresión en lugar de (15.12) para definir el poceso ARCH(r). En efecto, tomando esperanzas en (15.13) condicionadas a los valores observados, $E_{t-1} = (e_{t-1}, ..., e_1)$ se obtiene (15.12).

Kurtosis

Puede comprobarse que el coeficiente de kurtosis de un ARCH(r) es siempre mayor que 3.

15.3 Modelos GARCH

La identificación de modelos ARCH conduce generalmente a modelos de orden alto. Una idea natural es aproximar un proceso AR alto mediante una media móvil en las varianzas. Esto conduce al modelo GARCH o ARCH generalizado. Vamos a estudiar las propiedades de este modelo en el caso más simple. Supondremos como en el caso de los ARCH que

$$e_t = \sigma_t \epsilon_t$$

donde ϵ_t y σ_t son dos procesos estacionarios independientes entre sí y con las mismas propiedades anteriores.

15.3.1 El modelo GARCH(1,1)

En lugar de suponer una estructura AR para los cuadrados de las innovaciones suponemos que las varianzas condicionales siguen la ecuación:

$$\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{15.14}$$

donde tanto α_1 como β_1 son coeficientes positivos. En este modelo un valor alto para e_t^2 da lugar a una mayor varianza en el periodo siguiente y el término β_1 obliga a que la varianza cambie con cierta inercia, lo que produce rachas de mayor variabilidad.

Para calcular la varianza marginal, como de nuevo se verifica:

$$Var(e_t) = E(e_t^2) = E[E(e_t^2|W_{t-1})] = E(\sigma_t^2) = \sigma^2$$

sustituyendo la varianza por su expresión (15.14), tomando esperanzas y utilizando que para que el proceso e_t sea estacionario $Var(e_t) = Var(e_{t-1})$, lo que implica $E(\sigma_t^2) = E(\sigma_{t-1}^2)$, tenemos que

$$Var(e_t) = \sigma^2 = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}$$

y, por tanto, para que el proceso sea estacionario $\alpha_1 + \beta_1 < 1$.

Dependencia de los cuadrados

Vamos a escribir este modelo en la representación de los cuadrados de las innovaciones introduciendo la variable

$$v_t = e_t^2 - \sigma_t^2$$

que, por construcción, es un proceso estacionario de variables incorreladas de media cero. Se comprueba como en los modelos ARCH que estas variables aunque incorreladas no son independientes. Escribiendo $v_t = \sigma_t^2({\epsilon_t}^2 - 1)$, las covarianzas serán

$$E(v_t v_{t-k}) = E\left[\sigma_t^2({\epsilon_t}^2 - 1)\sigma_{t-k}^2({\epsilon_{t-k}}^2 - 1)\right] = E({\epsilon_t}^2 - 1)E(\sigma_t^2\sigma_{t-k}^2({\epsilon_{t-k}}^2 - 1)) = 0$$

ya que $E({\epsilon_t}^2 - 1) = 0$ y ϵ_t es independiente del resto de componentes del producto. Sustituyendo $\sigma_t^2 = e_t^2 - v_t$ en la ecuación (15.14), podemos escribir

$$e_t^2 = \alpha_0 + (\alpha_1 + \beta_1)e_{t-1}^2 + v_t - \beta_1 v_{t-1} \tag{15.15}$$

que muestra que los cuadrados de las observaciones siguen una estructura de dependencia similar a un ARMA(1,1). Al coeficiente $\lambda = (\alpha_1 + \beta_1)$ se le denomina persistencia y es frecuente en la práctica que se estime próximo a la unidad. Se demuestra que, llamando como en la sección anterior $\rho_c(k)$ a los coeficientes de autocorrelación al cuadrado de la serie, se obtiene

$$\rho_c(1) = \frac{\alpha_1(1 - \alpha_1\beta_1 - \beta_1^2)}{(1 - 2\alpha_1\beta_1 - \beta_1^2)}$$

mientras que

$$\rho_c(k) = (\alpha_1 + \beta_1)^{k-1}\rho_c(1), \qquad k > 1$$

que indica que el decrecimiento de estos coeficientes depende de la persistencia $\lambda = (\alpha_1 + \beta_1)$, y será lenta si este coeficiente es próximo a la unidad. Observemos que en el caso particular en que $\beta = 0$ obtenemos la estructura del ARCH(1).

Kurtosis

Puede demostrarse que suponiendo que la distribución de ϵ_t es normal el coeficiente de kurtosis viene dado por

$$K = \frac{3[1 - (\alpha_1 + \beta_1)^2]}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} > 3.$$

que exige la restricción adicional $1 - 2\alpha_1^2 - (\alpha_1 + \beta_1)^2 > 0$. Por tanto este modelo implica también colas pesadas. Observemos que esta ecuación es identica a la del ARCH(1) si hacemos $\beta_1 = 0$.

15.3.2 El modelo GARCH general

El modelo anterior puede generalizarse permitiendo simultaneamente una estructura AR de orden r en la dependencia de la varianza de los cuadrados de la serie y una estructura MA de orden s en la varianza condicional. Se obtiene entonces el modelo GARCH(r,s), dado por

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{r}\alpha_i e_{t-i}^2 + \sum_{j=1}^{s}\beta_j\sigma_{t-j}^2 \tag{15.16}$$

donde la secuencia ϵ_t tiene las mismas propiedades que en los procesos ARCH. Los parámetros de esta ecuación deben verificar ciertas restricciones para que la varianza sea positiva y existan los momentos de orden superior. En particular como la varianza debe ser positiva $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, y $\sum_{i=1}^{\max(r,s)}(\alpha_i + \beta_i) < 1$.

Definiendo como antes las variables incorreladas de media cero y varianza constante $v_t = e_t^2 - \sigma_t^2$, podemos escribir la dependencia de los cuadrados como

$$e_t^2 = \alpha_0 + \sum_{i=1}^{\max(r,s)}(\alpha_i + \beta_i)e_{t-i}^2 + \eta_t - \sum_{j=1}^{s}\beta_j\eta_{t-j}. \tag{15.17}$$

ecuación que representa un proceso ARMA en los cuadrados de las observaciones. Esperamos en consecuencia que los cuadrados de las observaciones presenten estructura dinámica similar a la de un proceso ARMA. Por otro lado, tomando esperanzas en esta ecuación, se obtiene inmediatamente que

$$E(e_t^2) = \frac{\alpha_0}{1 - \sum_{i=1}^{\max(r,s)}(\alpha_i + \beta_i)}$$

que indica la propiedad que deben verificar los parametros del proceso para que sea estacionario.

Cuando el operador AR de la representación de los cuadrados tiene una raiz en el círculo unidad se obtienen procesos IGARCH. En particular el IGARCH(1,1) es, haciendo $\alpha_1 + \beta_1 = 1$ en (15.15)

$$\nabla e_t^2 = \alpha_0 + v_t - \beta_1 v_{t-1}$$

que corresponde a una varianza condicional en (15.14) dada por, poniendo $\alpha_1 = 1 - \beta_1$

$$\sigma_t^2 = \alpha_0 + (1 - \beta_1)e_{t-1}^2 + \beta_1 \sigma_{t-1}^2.$$

15.4 Construccion de modelos ARCH y GARCH

La construcción de estos modelos implica primero identificar la estructura, estimarla y comprobar su adecuación. Vamos a revisar estos aspectos.

15.4.1 Identificación

La identificación de efectos ARCH se efectúa normalmente después de ajustar un modelo ARMA a la serie para eliminar la dependencia en la media. Si existen efectos ARCH, los residuos del modelo ARIMA estarán incorrelados pero no serán independientes y este efecto será visible en la función de autocorrelación de los residuos al cuadrado, que mostrarán correlación serial. Además, si calculamos los coeficientes de autocorrelación parcial de los residuos al cuadrado y el modelo para los residuos es ARCH puro. el número de términos distintos de cero nos indicará, aproximadamente, el orden del proceso.

Para detectar estructura en los cuadrados podemos acudir a los contrastres de McLeod y Li (1983) y Peña y Rodriguez (2003) presentados en la sección anterior. Además de estos contrastes generales, que sirven para detectar una estructura general no lineal, pueden utilizarse contrastes específicos para detectarla. Rodriguez y Ruiz (2003) presentan un contraste específico y lo comparan con las alternativas existentes. Véase también Ruiz y

15.4.2 Estimación

Supongamos para simplificar un ARCH(1), Entonces los residuos del modelo ARMA deben verificar la ecuación (15.7):

$$e_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + v_t \tag{15.18}$$

que podríamos pensar en estimar por mínimos cuadrados. Como el ruido de esta ecuación v_t está incorrelado con el regresor, la estimación proporcionará estimadores consistentes (véase por ejemplo Green, 1993, pag 419), aunque no eficientes. El problema es que la varianza condicionada en esta ecuación no es constante y la regresión es heterocedástica. En efecto, utilizando (**??**), tenemos que:

$$var(v_t|e_{t-1}^2) = 2E\left[\sigma_t^4|e_{t-1}^2\right] = 2(\alpha_0 + \alpha_1 e_{t-1}^2)^2$$

y no es constante, por lo que tendríamos que utilizar mínimos cuadrados generalizados.

Podríamos obtener un estimador más eficiente que el de regresión si conseguimos una estimación de σ_t^2, que llamaremos $\widehat{\sigma}_t^2$, y dividimos todos los términos de esta ecuación por $\widehat{\sigma}_t^2$. llamando $y_t = {e_t}^2/\widehat{\sigma}_t^2$, $z_t = 1/\widehat{\sigma}_t^2$, $x_t = {e_{t-1}}^2/\widehat{\sigma}_t^2$ y $u_t = v_t/\widehat{\sigma}_t^2$ podemos escribir la ecuación como:

$$y_t = \alpha_0 z_t + \alpha_1 x_t + u_t \tag{15.19}$$

y al dividir todos los términos de la ecuación por la desviación típica del residuo esta ecuación será aproximadamente homocedástica.

Vamos a ver que esto es precisamente el resultado de aplicar máxima verosimilitud a un modelo ARCH. Vamos a ilustrarlo con un proceso ARCH(1). Entonces, la función de verosimilitud es

$$f(e_1, \cdots, e_T|\alpha_0, \alpha_1) = f(e_1)f(e_2|e_1)...f(e_T|e_{T-1}).$$

Como hemos visto, la función de densidad $f(e_t|e_{t-1})$ es normal. En efecto, como $e_t = \sigma_t \epsilon_t$, si condicionamos a e_{t-1}el valor σ_t es una constante y la única variable es ϵ_t que tiene distribución normal. La media condicionada es cero, y la varianza $\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2$. La función de verosimilitud condicionada a e_1 en logaritmos será

$$L(e_2, \cdots, e_T|\alpha_0, \alpha_1, e_1) = -\frac{1}{2}\sum_{t=2}^{T} \log(\alpha_0 + \alpha_1 e_{t-1}^2) - \frac{1}{2}\sum_{t=2}^{T} \frac{{e_t}^2}{\alpha_0 + \alpha_1 e_{t-1}^2}$$

y derivando respecto a los parámetros, llamando $\widehat{\sigma}_t^2 = \widehat{\alpha}_0 + \widehat{\alpha}_1 e_{t-1}^2$ e igualando a cero se obtienen las ecuaciones

$$\sum \frac{{e_t}^2}{\widehat{\sigma}_t^4} = \sum \frac{1}{\widehat{\sigma}_t^2} \tag{15.20}$$

y

$$\sum \frac{e_{t-1}^2}{\widehat{\sigma}_t^2} = \sum \frac{e_{t-1}^2 {e_t}^2}{\widehat{\sigma}_t^4} \tag{15.21}$$

para ver como resolver estas ecuaciones, multipliquemos y dividamos el primer miembro por $\widehat{\alpha}_0 + \widehat{\alpha}_1 e_{t-1}^2$ y entonces podemos escribir (15.20) como

$$\widehat{\alpha}_0 \sum \frac{1}{\widehat{\sigma}_t^4} + \widehat{\alpha}_1 \sum \frac{e_{t-1}^2}{\widehat{\sigma}_t^4} = \sum \frac{{e_t}^2}{\widehat{\sigma}_t^4} \tag{15.22}$$

y (15.21) como

$$\widehat{\alpha}_0 \sum \frac{e_{t-1}^2}{\widehat{\sigma}_t^4} + \widehat{\alpha}_1 \sum \frac{e_{t-1}^4}{\widehat{\sigma}_t^4} = \sum \frac{e_{t-1}^2 {e_t}^2}{\widehat{\sigma}_t^4}$$

que son las ecuaciones de mínimos cuadrados para obtener los parámetros en la regresión (15.19). Este resultado sugiere el siguiente procedimiento iterativo para maximizar la verosimilitud. Estimar por regresión la ecuación (15.18) y obtener unos valores iniciales de los parámetros. Calcular con estos valores $\widehat{\sigma}_t^2$ y resolver las ecuaciones (15.21) y (15.22) para obtener nuevos estimadores de $\widehat{\alpha}_0$ y $\widehat{\alpha}_1$. Con ellos calcular de nuevo $\widehat{\sigma}_t^2$ y volver a las ecuaciones (15.21) y (15.22) para obtener un nuevo estimador. Iterar hasta convergencia.

En general, la función de verosimilitud de un proceso estacionario w_t cuyas innovaciones siguen un proceso GARCH(r, s), donde suponemos $r \geq s$ es, llamando $W_t = (w_1, ..., w_t)$,

$$\begin{aligned} f(w_1, \cdots, w_n|\alpha) &= f(w_n|W_{n-1})f(w_{n-1}|W_{n-2}) \cdots f(w_{r+1}|W_r)f(w_1, \cdots, w_r|\alpha) \\ &= \prod_{t=r+1}^{n} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp[-\frac{(w_t - E(w_t|W_{t-1}))^2}{2\sigma_t^2}] \times f(w_1, \cdots, w_r|\alpha) \end{aligned}$$

ya que la varianza condicional de las innovaciones es σ_t^2. Condicionando a las primeras r observaciones, que tienen una distribución más complicada, la función soporte condicionada es

$$L(w_{r+1}, \cdots, w_n|\alpha, W_r) = -\frac{1}{2}\sum_{t=r+1}^{n} \log \sigma_t^2 - \frac{1}{2}\sum_{t=r+1}^{n} \frac{(w_t - E(w_t|W_{t-1}))^2}{\sigma_t^2} \tag{15.23}$$

que puede maximizarse con un algoritmo de optimización no lineal para obtener los estimadores de los parámetros que aparecen en la media condicional y en la varianza condicional.

La estimación puede realizarse en dos etapas ya que la correlación entre los parámetros ARMA y la de los GARCH suele ser pequeña. En la primera etapa se estiman los parámetros de la media condicional, es decir, el modelo ARMA, y se construyen las innovaciones $e_t = w_t - E(w_t|W_{t-1})$. En la segunda se estiman los parámetros de la varianza condicional maximizando la verosimilitud de los residuos. Alternativamente, es posible estimar las ecuaciones de la media condicional y la varianza condicional conjuntamente, con lo que se obtiene una estimación más precisa.

15.4.3 Diagnosis

Si llamamos e_t a los residuos del modelo ARIMA y $\widehat{\sigma}_t$ a las varianzas condicionadas estimadas los residuos estandarizados $e_t/\widehat{\sigma}_t$ deben seguir un proceso de ruido blanco y les podemos aplicar los contraste habituales estudiados para procesos ARIMA. Además sus cuadrados no deben mostrar dependencia y podemos aplicar los contraste sobre las autocorrelaciones de los cuadrados descritos anteriormente.

Los valores atípicos pueden interpretarse como un aumento de la varianza en ese instante y , especialmente si aparecen en rachas, pueden confundirse con efectos heterocedascicidad condicional. Por otro lado, una serie que sigue un modelo ARCH puede mostrar muchos atípicos si se analiza como si tuviese varianza constante y siguiese un modelo ARMA. Es por lo tanto importante diferenciar ambos fenómenos.

Un procedimiento que funciona bien en la práctica es limpiar la serie inicialmente de las observaciones que presentan innovaciones tan grandes que no pueden ser debidas a heterocedasticidad condicional y que son muy probablemente valores atípicos. Puede comprobarse (véase Carnero, Peña y Ruiz, 2003) que observaciones mayores de siete desviaciones típicas pueden considerarse razonablemente como atípicas. A continuación estimaremos los efectos ARCH o GARCH con la serie corregida y construiremos la serie de innovaciones estandarizadas, e_t/σ_t. Sobre esta serie de innovaciones podemos de nuevo buscar valores atípicos y, si se detectan, corregirlos y volver a modelar las varianzas condicionales. El proceso se itera hasta convergencia.

15.5 Modelos de Volatilidad Estocástica

Un procedimiento alternativo para modelar la heterocedasticidad es suponer que la varianza condicional depende de factores no observados. Se obtienen así los modelos de volatilidad estocástica, o modelos SV (Stochastic volatility) que han sido estudiados por Harvey, Ruiz and Shephard (1994) y Jacquier, Polson and Rossi (1994). La ecuación de estos modelos es

$$e_t = \sigma_t \epsilon_t$$

donde, como en los modelos GARCH, los errores ϵ_t son iid $N(0, 1)$. Sin embargo, a diferencia de los GARCH donde la varianza depende de factores observables, en este modelo se supone que los logaritmos de las varianzas condicionadas siguen un proceso AR(p),

$$\ln(\sigma_t^2) = \alpha_0 + \sum_{i=1}^{p} \alpha_i \ln(\sigma_{t-i}^2) + \xi_t \tag{15.24}$$

donde el término de error en esta ecuación, ξ_t, es iid $N(0, \sigma_\xi^2)$ e independiente de ϵ_t. El coeficiente α_0 es una constante y los parámetros del proceso verifican las condiciones estandar de estacionaridad. Es decir, los ceros del polinomio $1 - \sum_{i=1}^{r} \alpha_i B^i$ deben estar fuera del círculo unidad. Vamos a estudiar este modelo en el caso más simple.

15.5.1 El modelo SV(1)

Suponiendo que el logaritmo de la volatilidad sigue un proceso AR(1), tenemos que

$$\ln \sigma_t^2 = \alpha_0 + \alpha_1 \ln \sigma_{t-1}^2 + \xi_t$$

donde suponemos $|\alpha_1| < 1$. Llamando $h_t = \ln \sigma_t^2$ este proceso es AR(1) con ruidos normales por lo que la distribución de la variable h_t será también normal. Los parámetros de la distribución marginal son los de un AR(1), es decir, la media es $\mu_y = \alpha_0/(1 - \alpha_1)$ y la varianza $\sigma_y^2 = \sigma_\xi^2/(1 - \alpha_1)$. Entonces, la variable σ_t^2 será lognormal y puede demostrarse que:

$$E(e_t^2) = e^{\mu_y} e^{(\sigma_y^2/2)}$$

y el coeficiente de kurtosis es

$$K = 3e^{\sigma_y^2}$$

que será siempre mayor que 3. Este modelo tiene pues también la capacidad de generar distribuciones con colas pesadas. Finalmente puede comprobarse que la estructura de correlacion de los cuadrados de este modelo es

$$\rho_c(k) = \frac{\exp((\sigma_y^2 \alpha_1^k) - 1]}{[3\exp(\sigma_y^2) - 1]}, \qquad k \geq 1$$

si la varianza σ_y^2 es pequeña, tomando $e^{\sigma_y^2} \approx 1$ y aproximando $\exp(\alpha_1^k) \approx 1 + \alpha_1^k$, tenemos que

15.6 Otros Enfoques

Además de estos modelos exiten otros modelos propuestos para explicar la volatilidad condicional. Entre ellos están los modelos GARCH exponenciales de Nelson (1991) que conducen a efectos asimétricos de las innovaciones, los modelos ARMA de heterocedasticidad condicional introducidos por Tsay (1987), y los modelos de coeficientes aleatorios de Nicholls and Quinn (1982). Vease Peña, Tiao y Tsay (2001) y Tsay (2002) para una descripción de estos modelos y sus limitaciones.