Modelos de volatilidad estocástica

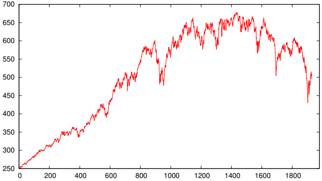
Tere García Muñoz (tgarciam@ugr.es)
Román Salmerón Gómez (romansg@ugr.es)
Alessio Gaggero (alessiogaggero@ugr.es)

Universidad de Granada
Departamento de Métodos Cuantitativos para la Economía y la Empresa

Tema 5– Econometría 3
Grado en Económicas

En los modelos econométricos convencionales, la varianza del término perturbación se supone constante. Sin embargo, muchas series temporales económicas, principalmente las financieras, exhiben periodos de alta volatilidad seguidos de periodos de relativa tranquilidad. En este caso, grandes choques (residuales) están seguidos por grandes choques en ambas direcciones, y pequeños shocks están seguidos por pequeñas perturbaciones. Por ejemplo, los mercados bursátiles se caracterizan por periodos *relajados* de baja volatilidad y alta volatilidad. Esto es particularmente cierto en series con altas frecuencias (diarias o semanales). Una forma de modelizar tales patrones es permitir que la varianza del término perturbación dependa de su propio pasado.

Volatilidad: variaciones bruscas que pueden encontrarse en una serie de tiempo. Aunque en el largo plazo la serie sea estacionaria y tenga, por tanto, varianza constante, puede presentar oscilaciones a corto plazo donde cada periodo esté influenciado por los inmediatamente anteriores, que es lo que recoge la varianza condicional.



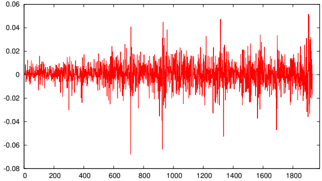


Figura : Observaciones de la bolsa de Nueva York cada 5 días (izquierda) y de las primeras diferencias regulares de su logaritmo (derecha)

Un buen modelo para abordar la volatilidad debe reflejar:

- La volatilidad tiene tendencia a aparecer agrupada por periodos, los periodos de alta volatilidad suelen mantenerse en el tiempo y lo mismo ocurre con los periodos de baja volatilidad (efecto contagio).
- A un periodo de alta/baja volatilidad, eventualmente, le sigue otro de volatilidad normal. Es decir, existe un nivel normal de volatilidad al cual ésta retorna eventualmente.

Estas características muestran la existencia de regularidades en el comportamiento que posibilitan la detección de la volatilidad y, por tanto, su modelización. Los modelos que utilizaremos reciben el nombre de **procesos con varianza condicionalmente heterocedástica** (**modelos ARCH**). En 1982, Engle introduce este tipo de modelos presuponiendo que la varianza condicional (no constante) depende de sus valores pasados siguiendo una estructura autorregresiva. Posteriormente en 1986, Bollerslev generaliza dichos modelos incorporando estructura de media móvil a la varianza condicional dando lugar a los **modelos GARCH**.

En los procesos con varianza condicionalmente heterocedástica se considera que la varianza condicionada del ruido de un modelo ARMA ($E[\varepsilon_t^2/\varepsilon_{t-1},\dots] \equiv E_{t-1}[\varepsilon_t^2]$) está condicionada por los valores pasados. Por tanto la varianza condicionada no será constante pues depende en cada instante t de los valores pasados en $t-1, t-2, \dots$

Una estrategia simple es modelizar el cuadrado de las perturbaciones en función de su pasado, por ejemplo:

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + v_t,$$

donde v_t es ruido blanco. De esta manera, la varianza condicionada de las perturbaciones depende del pasado de la perturbaciones:

$$E_{t-1}[\varepsilon_t^2] = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$

Si α_1 es cero, la varianza condicionada será constante (no habrá volatilidad). En otro caso la varianza condicional evoluciona de acuerdo a un proceso autorregresivo.

Ahora bien, la relación lineal no tiene un tratamiento algebráico cómodo, siendo más manejable especificar el ruido de forma multiplicativa.

Se dice que $\{\varepsilon_t\}$ sigue un **proceso autorregresivo condicionalmente heterocedástico de orden 1**, **ARCH(1)**, si:

$$\varepsilon_t = v_t \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2} \Leftrightarrow \varepsilon_t^2 = v_t^2 \left(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2\right),$$

donde v_t es ruido blanco con varianza igual a 1 e independiente de ε_{t-1}.
Denotando por: $h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$

- La media marginal (o no condicionada) vale cero:

$$E[\varepsilon_t] = E[v_t \sqrt{h_t}] = \underbrace{E[v_t]}_{0} \cdot E[\sqrt{h_t}] = 0,$$

- Teniendo en cuenta que la varianza marginal (no condicionada) es constante, se tiene que:

$$E[\varepsilon_t^2] = E[v_t^2 h_t] = \underbrace{E[v_t^2]}_{1} \cdot E[h_t] = \alpha_0 + \alpha_1 E[\varepsilon_{t-1}^2] \Rightarrow E[\varepsilon_t^2] = \frac{\alpha_0}{1-\alpha_1}.$$

  Para que la varianza sea positiva se ha de verificar que $\alpha_0 > 0$ y $0 \leq \alpha_1 < 1$.

- Covarianza:

$$E[\varepsilon_t \varepsilon_{t-j}] = E[v_t\sqrt{h_t}v_{t-j}\sqrt{h_{t-j}}] = \underbrace{E[v_t]}_{0}\underbrace{E[v_{t-j}]}_{0} \cdot E[\sqrt{h_t}\sqrt{h_{t-j}}] = 0, j \neq 0$$

Aún con la modelización ARCH, ε_t sigue siendo un ruido blanco. La influencia de la modelización ARCH cae por completo en la varianza condicionada.

- La media condicionada (a los valores pasados de la serie) vale cero:

$$\begin{aligned} E_{t-1}[\varepsilon_t] &= E_{t-1}[v_t\sqrt{h_t}] = E_{t-1}[v_t] \cdot E_{t-1}[\sqrt{h_t}] \\ &= \underbrace{E[v_t]}_{0} \cdot E_{t-1}[\sqrt{h_t}] = 0, \end{aligned}$$

- La varianza condicional tiene una estructura parecida a un AR(1):

$$\begin{aligned} E_{t-1}[\varepsilon_t^2] &= E_{t-1}[v_t^2 h_t] = E_{t-1}[v_t^2] \cdot E_{t-1}[h_t] = \underbrace{E[v_t^2]}_{1} \cdot E_{t-1}[h_t] \\ &= \alpha_0 + \alpha_1 E_{t-1}[\varepsilon_{t-1}^2] = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 = h_t, \end{aligned}$$

Para que sea positiva se impone que los dos parámetros sean positivos.

De esta expresión $E_{t-1}[\varepsilon_t^2] = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 = h_t$ se deducen las características deseables de los modelos con volatilidad:

- Un valor alto/bajo de ε_{t-1}^2 conduce a un valor alto/bajo de la varianza del siguiente valor condicionada a ε_{t-1}^2, $E_{t-1}[\varepsilon_t^2]$. De esta forma, se hace más probable que el siguiente dato, ε_t^2, sea también alto/bajo. Luego, fuertes fluctuaciones inesperadas tienden a venir seguidas de períodos de iguales características.
- Por otro lado, como tanto la media marginal como la condicionada son cero, es posible que el valor de ε_t^2 sea pequeño (retorne a su media). De esta forma disminuirá la varianza condicionada de la observación siguiente, de manera que períodos de estabilidad tienden a venir seguidos por períodos igualmente estables.

Se dice que $\{\varepsilon_t\}$ sigue un **proceso autorregresivo condicionalmente heterocedástico de orden** q, **ARCH(q)**, si:

$$\varepsilon_t = v_t \sqrt{h_t},$$

donde $h_t = \alpha_0 + \sum_{k=1}^{q} \alpha_k \varepsilon_{t-k}^2$ y v_t es ruido blanco con varianza igual a 1 e independiente de $\varepsilon_{t-1}, \ldots, \varepsilon_{t-q}$.
De manera análoga a la anterior se obtiene que:

- $E[\varepsilon_t] = 0 = E_{t-1}[\varepsilon_t]$.
- $E[\varepsilon_t^2] = \frac{\alpha_0}{1-\sum_{k=1}^{q} \alpha_k}$ de donde se deduce que $\alpha_0, \alpha_1, \ldots, \alpha_q \geq 0$ y $\sum_{k=1}^{q} \alpha_k < 1$.
- $E_{t-1}[\varepsilon_t^2] = \alpha_0 + \sum_{k=1}^{q} \alpha_k \varepsilon_{t-k}^2 = h_t$ (la varianza condicionada depende de los últimos q valores de ε_{t-1}^2, luego tiene estructura autorregresiva).

Si el valor q del *ARCH(q)* es grande, se tiene que estimar un gran número de parámetros, $\alpha_1, \ldots, \alpha_q$, con los problemas que esto puede suponer. Para evitarlos se permite que la varianza condicionada tenga estructura ARMA:
El caso más sencillo responde a la expresión:

$$h_t = E_{t-1}[\varepsilon_t^2] = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1},$$

Se dice que $\{\varepsilon_t\}$ sigue un **proceso autorregresivo condicionalmente heterocedástico generalizado de orden (1,1)**, **GARCH(1,1)**, si:

$$\varepsilon_t = v_t \sqrt{h_t},$$

donde $h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$ y v_t es ruido blanco con varianza igual a 1 e independiente de ε_{t-1}.

- La media marginal y condicionada son cero: $E[\varepsilon_t] = 0 = E_{t-1}[\varepsilon_t]$.
- Teniendo en cuenta que la varianza marginal es estacionaria y la ley de esperanzas iteradas:

$$\begin{aligned}
E[\varepsilon_t^2] &= E[v_t^2 h_t] = \underbrace{E[v_t^2]}_{1} \cdot E[h_t] = \alpha_0 + \alpha_1 E[\varepsilon_{t-1}^2] + \beta_1 E[h_{t-1}] \\
&\Downarrow \\
&\Downarrow \quad E[h_{t-1}] = E[E_{t-2}[\varepsilon_{t-1}^2]] = E[E[\varepsilon_{t-1}^2/\varepsilon_{t-2}, \ldots]] = E[\varepsilon_{t-1}^2] \\
&\Downarrow \\
&= \alpha_0 + \alpha_1 E[\varepsilon_{t-1}^2] + \beta_1 E[\varepsilon_{t-1}^2] \Rightarrow E[\varepsilon_t^2] = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}.
\end{aligned}$$

Para que la varianza sea positiva se ha de verificar que $\alpha_0, \alpha_1, \beta_1 \geq 0$ y $\alpha_1 + \beta_1 < 1$.

- La varianza condicionada corresponde a una estructura ARMA(1,1):

$$\begin{aligned} E_{t-1}[\varepsilon_t^2] &= E_{t-1}[v_t^2 h_t] = E_{t-1}[v_t^2] \cdot E_{t-1}[h_t] = \alpha_0 + \alpha_1 E_{t-1}[\varepsilon_{t-1}^2] \\ &\quad + \beta_1 E_{t-1}[h_{t-1}] = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} = h_t, \end{aligned}$$

  donde se ha usado que:

$$E_{t-1}[h_{t-1}] = E_{t-1}[E_{t-2}[\varepsilon_{t-1}^2]] = E_{t-2}[E_{t-1}[\varepsilon_{t-1}^2]] = E_{t-2}[\varepsilon_{t-1}^2] = h_{t-1}.$$

La novedad en este caso radica en la aparición de un nuevo término que hace que la varianza condicionada, h_t, dependa de un retardo suyo, h_{t-1}. Así, es de esperar que haya rachas de mayor variabilidad que con los modelos ARCH(1) debido al efecto directo de las varianzas condicionadas altas/bajas del instante anterior.

Se dice que $\{\varepsilon_t\}$ sigue un **proceso autorregresivo condicionalmente heterocedástico generalizado de orden (p,q)**, **GARCH(p,q)**, si:

$$\varepsilon_t = v_t\sqrt{h_t},$$

donde

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}$$

y v_t es ruido blanco con varianza igual a 1 e independiente de $\varepsilon_{t-1}, \ldots, \varepsilon_{t-q}$. De manera análoga a la anterior se obtiene que:

- $E[\varepsilon_t] = 0 = E_{t-1}[\varepsilon_t]$.
- $E[\varepsilon_t^2] = \frac{\alpha_0}{1-\sum_{i=1}^{q}\alpha_i - \sum_{j=1}^{p}\beta_j}$ de donde se deduce que $\alpha_0, \alpha_1, \ldots, \alpha_q \geq 0$, $\beta_1, \ldots, \beta_q > 0$ y $\sum_{i=1}^{q}\alpha_i + \sum_{j=1}^{p}\beta_j < 1$.
- $E_{t-1}[\varepsilon_t^2] = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} = h_t$ (la varianza condicionada tiene estructura ARMA(p,q)).

La modelización de modelos GARCH reside en el comportamiento del cuadrado del ruido. Por tanto, para la identificación de los mismos habrá que basarse en los residuos al cuadrado del modelo de series temporales analizado.
Los pasos a seguir son los siguientes:

- Estimar el proceso estocástico $\{Y_t\}$ con el mejor modelo ARMA posible. En tal caso, dicho proceso es estacionario (media y varianza constantes) y sus residuos deben responder a las características de un ruido blanco.
- Calcular el cuadrado de los errores y representar sus funciones de autocorrelación simple y parcial. La identificación de un patrón ARMA(p,q) para los cuadrados de los residuos sería indicativo de la adecuación de un modelo GARCH(p,q).

Ejemplo.La serie correspondiente al PIB trimestral desde 1995 a 2015 es transformada de la siguiente manera: se calcula su logaritmo neperiano y se hacen primeras diferencias regulares y estacionales. El correlograma de la serie transformada es:

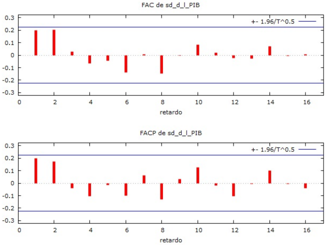


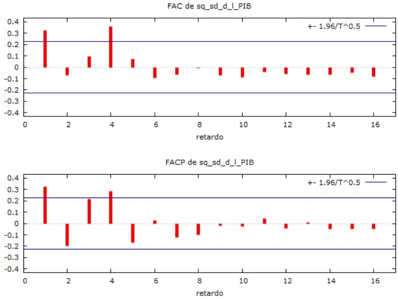
Figura : Correlogramas del logaritmo neperiano, primeras diferencias regulares y estacionales del PIB trimestral

Escribir la ecuación del modelo.

Puesto que los correlogramas se corresponden con un ruido blanco, la ecuación del modelo sería:

$$\nabla_4 \cdot \nabla \cdot \ln PIB_t = \varepsilon_t,$$

donde ε_t es ruido blanco.
El correlograma de los residuos al cuadrado de este modelo es:



En estos correlogramas se observa la existencia de estructura, por lo que en el modelo ajustado hay volatilidad y tendríamos que identificar un modelo GARCH.

Un procedimiento más formal para detectar la presencia de estructura ARCH en los errores es el siguiente:

1. Estimar el proceso estocástico $\{Y_t\}$ con el mejor modelo ARMA posible y obtener el cuadrado de los errores.
2. Realizar una regresión de dichos residuos al cuadrado sobre una constante y q retardos de los mismos:
$$e_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \cdots + \alpha_q e_{t-q}^2 + v_t$$
.
3. Si no hay un patrón ARCH, los valores estimados de α_i, con $i = 1, \ldots, q$, deberían ser estadísticamente cero y el coeficiente de determinación de esta regresión auxiliar, R_{aux}^2, muy bajo.
Se rechazaría la hipótesis nula
$$H_0 : \alpha_1 = \cdots = \alpha_q = 0$$
si $\chi_{exp}^2 = T \cdot R_{aux}^2 > \chi_q^2(1-\alpha)$

En definitiva, si se rechaza la hipótesis nula anterior se podría concluir (al nivel de significación usado) que hay estructura ARCH en los errores.

Ejemplo. En el ejemplo anterior, se ha realizado el contraste para detectar la presencia de estructura ARCH en los errores del modelo ajustado que tiene por hipótesis nula

$$H_0 : \alpha_1 = 0$$

y se obtiene un coeficiente de determinación para la regresión auxiliar igual a 0.048.
Teniendo en cuenta que $q = 1$ y que se disponen de 167 observaciones se tiene que se rechaza la hipótesis nula ya que

$$\chi^2_{exp} = 167 \cdot 0{,}048 = 8{,}016 > 3{,}81 = \chi^2_1(0{,}95)$$

.
Por tanto, en los errores hay estructura ARCH.

Ejemplo. A partir de los residuos, *e*, del modelo

$$(1 + 0{,}2 \cdot B) \cdot Y_t = (1 - 0{,}7 \cdot B) \cdot \varepsilon_t$$

se desea comprobar si hay volatilidad.
Con tal objetivo se plantea la regresión auxiliar

$$e_t^2 = \alpha_0 + \alpha_1 \cdot e_{t-1}^2 + \alpha_2 \cdot e_{t-2}^2 + \alpha_3 \cdot e_{t-3}^2 + v_t$$

obteniéndose un coeficiente de determinación $R_{aux}^2 = 0{,}018$ y utilizando 100 observaciones.
Contraste
$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0$ (no existe estructura ARCH)
$H_1 : \exists i$ tal que $\alpha_i \neq 0$ (existe estructura ARCH)
NO se rechaza la hipótesis nula ya que:

$$\chi_{exp}^2 = 100 \cdot 0{,}018 = 1{,}8 \not> 7{,}815 = \chi_3^2(0{,}95).$$

Por tanto, no existe estructura ARCH en el modelo.