Modelos de Volatilidad Multivariada

- La volatilidad multivariada tiene muchas aplicaciones en finanzas:

  ✓ Juega un papel importante en la selección de portafolios y asignación de activos.

  ✓ Puede ser usada para calcular el valor en riesgo (VaR) de posiciones financieras que contienen varios activos.

- Considere una serie de tiempo multivariada de retornos $\{r_t\}$. Como en el caso univariado, la serie puede ser escrita como

$$r_t = \mu_t + a_t$$

donde $\mu_t = E(r_t \mid F_{t-1})$ es la media condicional de r_t dada la información pasada F_{t-1} y $a_t = (a_{1t}, \cdots, a_{kt})'$ es el shock o innovación de la serie en el período t.

  ✓ El proceso μ_t generalmente sigue un proceso ARMA multivariado con variables exógenas, es decir,

$$\mu_t = \Upsilon x_t + \sum_{i=1}^{p} \Phi_i r_{t-i} - \sum_{j=1}^{q} \Theta_j a_{t-j}$$

Donde:

$\Rightarrow$ x_t es un vector m-dimensional de variables exógenas (o variables explicativas) con $x_{1t} = 1$.

$\Rightarrow$ Υ es una matriz de k x m

$\Rightarrow$ p y q son números no negativos.

La ecuación anterior será llamada la ecuación de la media de r_t.

- Σ_t , la matriz de covarianza condicional de a_t dada F_{t-1}, es una matriz de k x k definida positiva, es decir $\Sigma_t = Cov(a_t \mid F_{t-1})$.

  ✓ La modelación de la volatilidad multivariada se relaciona con la evolución en el tiempo de Σ_t. El modelo para el proceso $\{\Sigma_t\}$ se denomina el *modelo de volatilidad* para la serie multivariada de retornos r_t.

  ✓ Hay muchas maneras para generalizar los modelos de volatilidad univariados al caso multivariado. Sin embargo, en la práctica el mayor obstáculo es *la maldición de la multidimensionalidad*, debido a que, en cada período de tiempo, hay k(k+1)/2 términos en Σ_t. Por ejemplo, para un retorno 5-dimensional, hay 15 varianzas y covarianzas condicionales. Si el retorno fuera 6-dimensional el número de parámetros a estimar en Σ_t es 21. Por tanto el objetivo es introducir algunos modelos para la volatilidad multivariada relativamente simples que sean útiles y manejables en aplicaciones reales.

Estimación Exponencialmente Ponderada de la matriz Σ_t.

- Dada la información pasada F_{t-1}=$\{a_1, \cdots a_{t-1}\}$, la matriz de covarianzas **incondicional** de la innovación se puede estimar, dado que la media de a_t es un vector nulo, como

$$\hat{\Sigma} = \frac{1}{t-1}\sum_{j=1}^{t-1} a_j a'_j$$

  ✓ Este estimador asigna la misma ponderación 1/(t-1) a todas las observaciones en la suma.

- ✓ Para permitir que la matriz sea tiempo variante y al mismo tiempo enfatizar que las observaciones recientes son más importantes, se puede usar la idea del suavizamiento exponencial y estimar la matriz de covarianzas **condicional** de a_t como

$$\hat{\Sigma}_t = (1-\lambda)a_{t-1}a'_{t-1} + \lambda\hat{\Sigma}_{t-1}$$

donde $0 < \lambda < 1$, es llamado el parámetro de persistencia. Este estimador es llamado el estimador exponencialmente ponderado de medias móviles (EWMA) de la matriz de covarianza Σ_t.

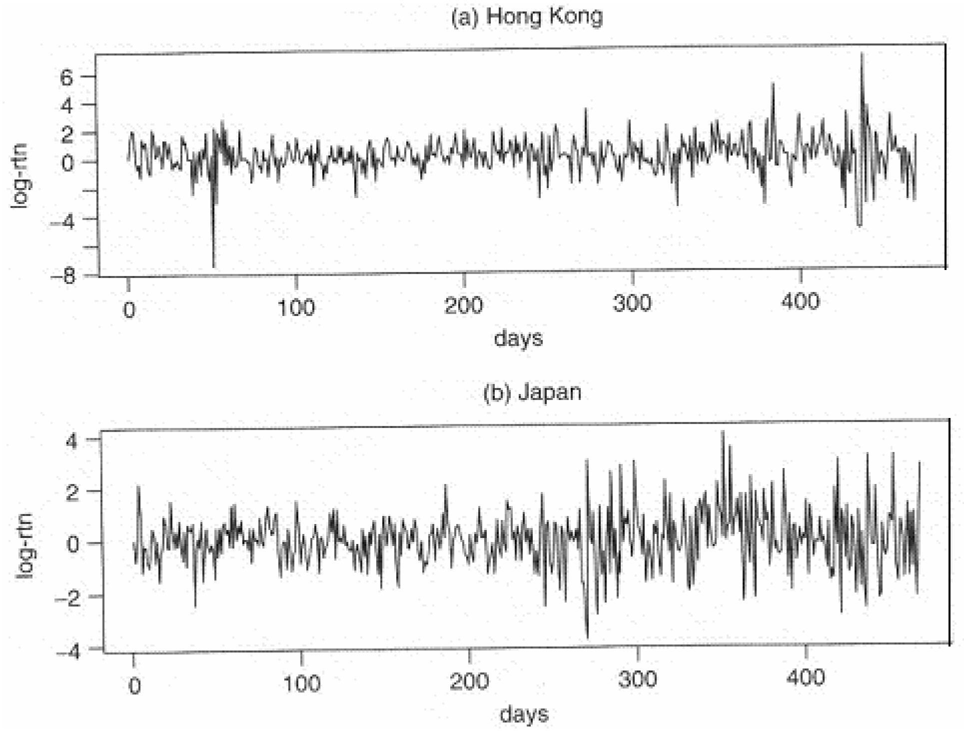
- ✓ Suponga que los datos son $\{r_1, \cdots r_T\}$. Para un valor dado de λ y una estimación inicial $\hat{\Sigma}_1$, $\hat{\Sigma}_t$ puede ser calculado recursivamente. Bajo el supuesto de que $a_t = r_t - \mu_t$ sigue una distribución normal multivariada con media cero y matriz de covarianza Σ_t, donde μ_t es una función de parámetros Θ de Υ (vea la ecuación de la media de r_t), entonces λ y Υ pueden ser estimadas conjuntamente usando máxima verosimilitud, puesto que el log de la función de verosimilitud de los datos es

$$\ln L(\Theta, \lambda) \infty - \frac{1}{2}\sum_{t=1}^{T} \ln(|\Sigma_t|) - \frac{1}{2}\sum_{t=1}^{T} (r_t - \mu_t)'\Sigma_t^{-1}(r_t - \mu_t)$$

La cual puede ser evaluada recursivamente reemplazando Σ_t por $\hat{\Sigma}_t$.

Ejemplo

Considere los retornos diarios del los índices de la bolsa de Hong Kong y Japón desde Enero1 de 1996 hasta Octubre 16 de 1997, para un total de 469 observaciones. Los retornos están en porcentajes. La siguiente figura muestra los gráficos de las dos series.



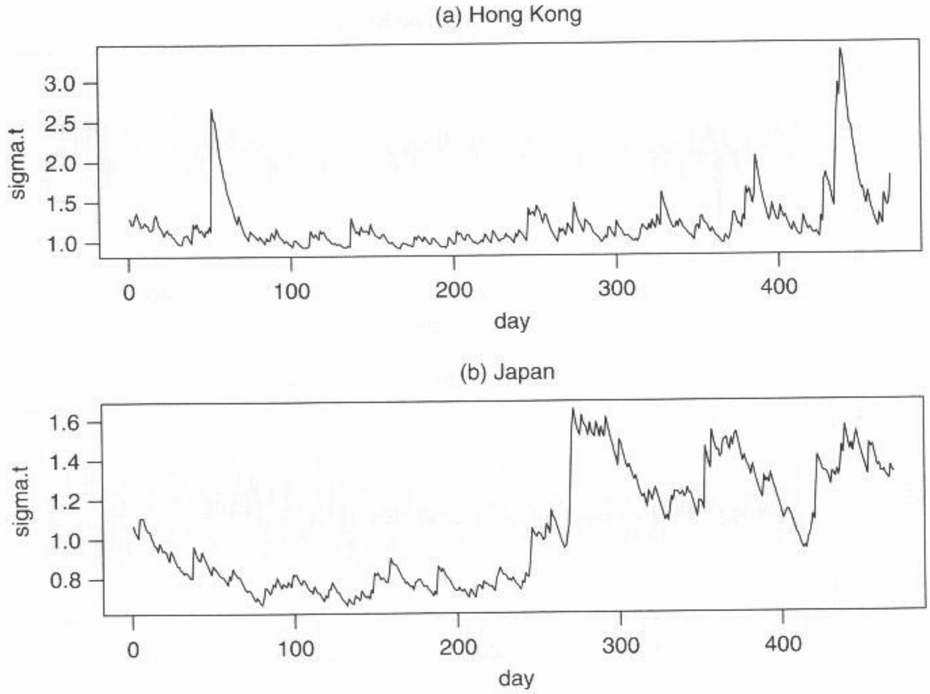
- ✓ Sean r_{1t} y r_{2t} los retornos de las bolsas de Hong Kong y Japón, respectivamente. Sus modelos GARCH univariados son

$$r_{1t} = 0.090 - 0.094 r_{1,t-6} + a_{1t}, \quad a_{1t} = \sigma_{1t}\varepsilon_{1t}$$
$$\sigma_{1t}^2 = 0.126 + 0.103 a_{1,t-1}^2 + 0.81\sigma_{1,t-1}^2$$

y

$$r_{2t} = -0.046 + a_{2t}, \quad a_{2t} = \sigma_{2t}\varepsilon_{2t}$$
$$\sigma_{2t}^2 = 0.0007 + 0.054 a_{2,t-1}^2 + 0.942\sigma_{2,t-1}^2$$

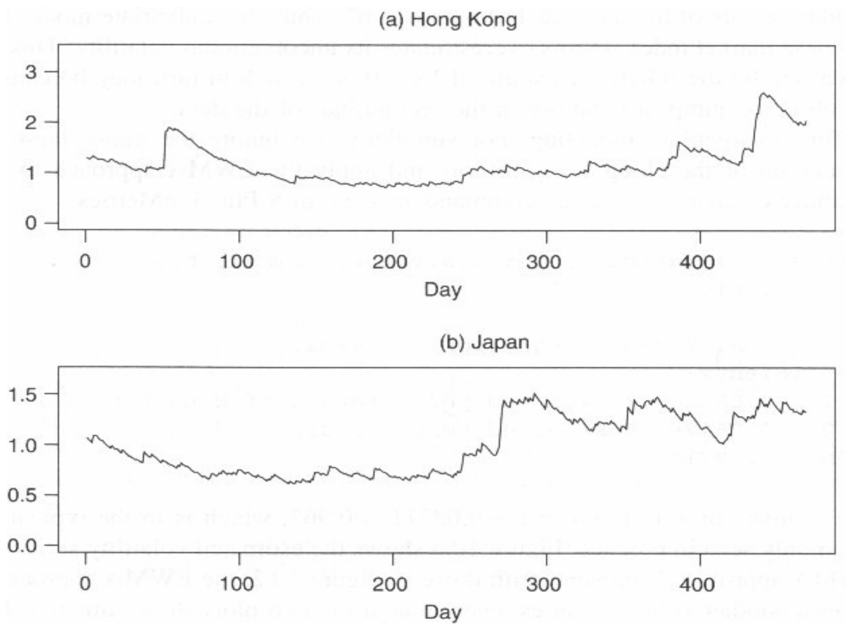
donde todos los parámetros son significativos al 5%, excepto para los términos constantes de los retornos y para el α_0 de r_{2t}. Los estadísticos de Ljung-Box para los residuales y para los residuales al cuadrado indican que el modelo es adecuado. El siguiente gráfico muestra las volatilidades de los modelos anteriores.



- ✓ La bolsa de Hong Kong parece ser más volátil que la de Japón, pero esta última presenta un incremento en la volatilidad en la segunda mitad de la muestra.

- ✓ El error estándar asintótico de los errores de los índices de retornos basados en los modelos son, respectivamente, 1.259 y 1.393. Los errores estándar de los datos son 1.296 y 1.067. Por tanto el modelo univariado para la bolsa de Japón sobrestima la volatilidad incondicional. **La causa podría ser la característica IGARCH del modelo ajustado, y esta a su vez podría ser causada por el salto en volatilidad en la segunda mitad de la muestra**.

- ✓ Para la *modelación bivariada*, ignorando la correlación serial de rezago 6 para la bolsa de Hong Kong, y aplicando la aproximación EWMA para obtener las volatilidades, se obtiene que el estimador de λ es aproximadamente 0.963, el cual cae en el rango típico encontrado en la práctica. El siguiente gráfico presenta las series de volatilidad estimada usando la aproximación EWMA. Comparada con la obtenida antes, el procedimiento EWMA suaviza las series de volatilidad, aunque los dos procedimientos producen patrones de volatilidad similares.



Algunos Modelos Multivariados GARCH

- **El modelo Diagonal, VEC**: Bollerslev, Engle y Wooldridge (1988) propusieron el siguiente modelo para generalizar la aproximación EWMA**:**

$$\Sigma_t = A_0 + \sum_{i=1}^{m} A_i \oplus (a_{t-i}a'_{t-i}) + \sum_{j=1}^{s} B_j \oplus \Sigma_{t-j}$$

donde m y s son enteros no negativos A_i y B_j son matrices simétricas, y $\oplus$ denota el producto Hadamard, es decir, la multiplicación matricial elemento por elemento. Este modelo es denominado modelo VEC(m,s) o modelo DVEC(m,s).

✓ Para entender el modelo, considere el modelo bivariante VEC(1,1),

$$\begin{bmatrix} \sigma_{11,t}^2 & \\ \sigma_{21,t}^2 & \sigma_{22,t}^2 \end{bmatrix} = \begin{bmatrix} A_{11,0} & \\ A_{21,0} & A_{22,0} \end{bmatrix} + \begin{bmatrix} A_{11,1} & \\ A_{21,1} & A_{22,1} \end{bmatrix} \oplus \begin{bmatrix} a_{1,t-1}^2 & \\ a_{1,t-1}a_{2,t-1} & a_{2,t-1}^2 \end{bmatrix}$$
$$+ \begin{bmatrix} B_{11,1} & \\ B_{21,1} & B_{22,1} \end{bmatrix} \oplus \begin{bmatrix} \sigma_{11,t-1}^2 & \\ \sigma_{21,t-1}^2 & \sigma_{22,t-1}^2 \end{bmatrix}$$

donde solamente se presenta la parte triangular del modelo.

✓ Concretamente, el modelo es

$$\sigma_{11,t}^2 = A_{11,0} + A_{11,1}a_{1,t-1}^2 + B_{11,1}\sigma_{11,t-1}^2$$
$$\sigma_{21,t}^2 = A_{21,0} + A_{21,1}a_{1,t-1}a_{2,t-1} + B_{21,1}\sigma_{21,t-1}^2$$
$$\sigma_{22,t}^2 = A_{22,0} + A_{22,1}a_{2,t-1}^2 + B_{22,1}\sigma_{22,t-1}^2$$

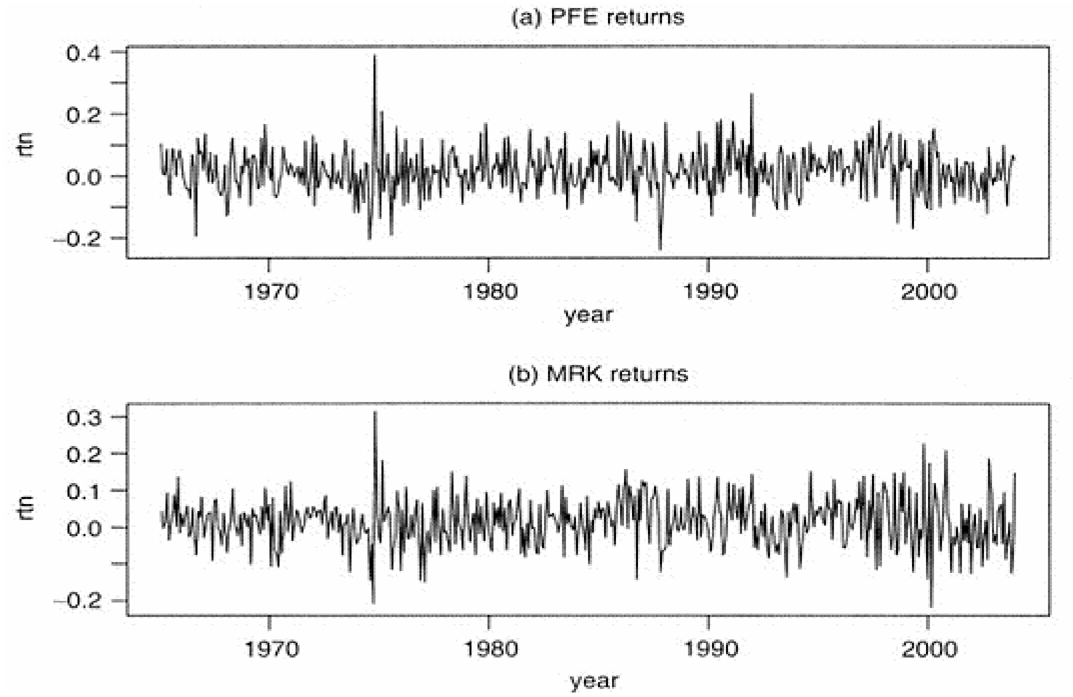
✓ De lo anterior, cada elemento del modelo DVEC sigue un modelo tipo GARCH(1,1).

✓ El modelo es simple, aunque puede no producir una matriz de covarianza definida positiva.

✓ Además, el modelo no permite la dependencia dinámica entre las series de volatilidad.

Ejemplo

Considere los retornos mensuales simples, los cuales contienen los dividendos, de dos de las mayores compañías farmacéuticas, desde Enero de 1965 a Diciembre de 2003, para un total de 468 observaciones. Sea r_{1t} la serie de los retornos de las acciones de Pfizer y r_{2t} la series de retornos de las acciones de Merck. La serie bivariante $r_t=(r_{1t}, r_{2t})'$ no tiene correlaciones seriales significativas. Por tanto la ecuación de media solamente contiene una constante. A continuación se presenta su gráfico.



La siguiente tabla presenta los resultados de la estimación de un modelo DVEC(1,1) bajo normalidad.

Coeficientes estimados:

	Valor	ErrorSTD	t	Valor p
C(1)	0.0164424	3.422e-03	4.805	1.047e-06
C(2)	0.0150987	3.139e-03	4.810	1.025e-06
A(1,1)	0.0008181	4.348e-04	1.881	3.027e-02
A(2,1)	0.0001021	4.979e-05	2.050	2.048e-02
A(2,2)	0.0001408	7.067e-05	1.992	2.348e-02
ARCH(1,1,1)	0.0727734	2.973e-02	2.448	7.363e-03
ARCH(1,2,1)	0.0259816	9.537e-03	2.724	3.343e-03
ARCH(1,2,2)	0.0518917	1.753e-02	2.961	1.614e-03
GARCH(1,1,1)	0.7777585	9.525e-02	8.165	1.554e-15
GARCH(1,2,1)	0.9407037	2.191e-02	42.928	0.000e+00
GARCH(1,2,2)	0.9203388	2.684e-02	34.296	0.000e+00

Prueba de Ljung-Box para los residuales estandarizados:

	Estadístico	Valor p	g.l de la chi-cuadrado
Pfe	10.07	0.6096	12
Mrk	14.91	0.2461	12

Prueba de Ljung-Box para los residuales estandarizados al cuadrado:

	Estadístico	Valor p	g.l de la chi-cuadrado
Pfe	18.30	0.1068	12
Mrk	5.04	0.9566	12

- ✓ De los resultados anteriores, las estimaciones para los parámetros son todas significativas al 5%.

- ✓ El modelo de volatilidad ajustado es

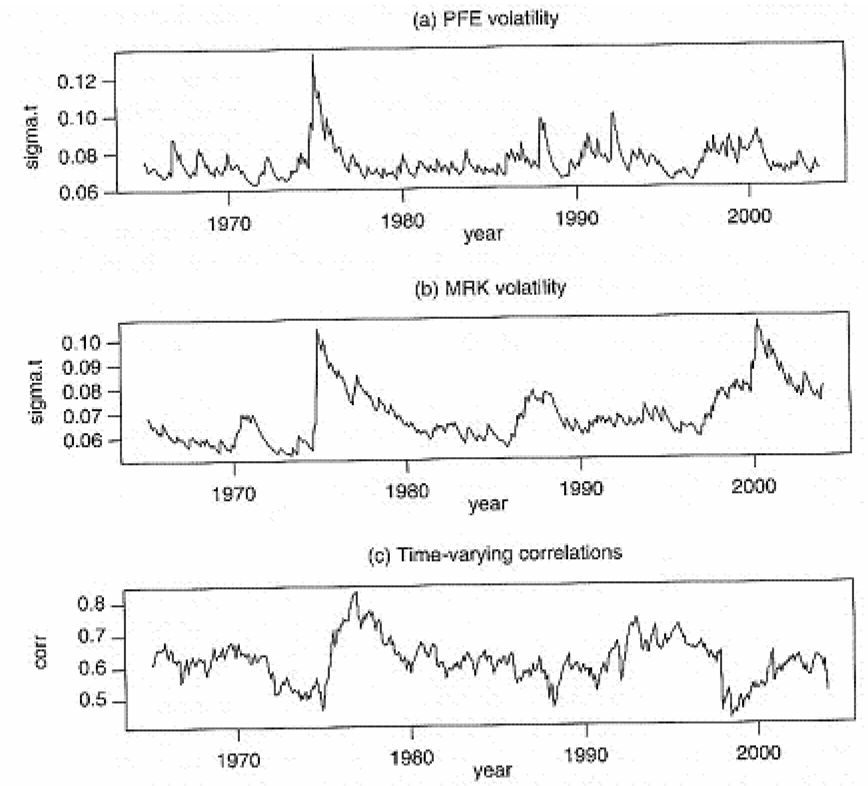
$$\sigma_{11,t}^2 = 0.00082 + 0.073a_{1,t-1}^2 + 0.778\sigma_{11,t-1}^2$$

$$\sigma_{21,t}^2 = 0.00010 + 0.026a_{1,t-1}a_{2,t-1} + 0.941\sigma_{21,t-1}^2$$

$$\sigma_{22,t}^2 = 0.00014 + 0.052a_{2,t-1}^2 + 0.920\sigma_{22,t-1}^2$$

- ✓ Los estadísticos de Ljung-Box para las series individuales de residuales y los cuadrados de los retornos muestran que el modelo DVEC(1,1) no puede ser rechazado.

- ✓ Un diagnóstico más informativo consiste en el estadístico Q multivariado a la serie bivariada de los residuales estandarizados y la serie bivariada de los residuales estandarizados al cuadrado. En este caso $Q_2(10)$=42.04(0.38) y $Q_2^*(10)$=67.33(0.004), donde Q_2^* denota el estadístico Q_2 para la series de residuales estandarizados al cuadrado. Usando estos, la ecuación de la media es adecuada a un nivel de significancia del 5%, pero el modelo DVEC(1,1) para la volatilidad es rechazado a un nivel del 1%. Los siguientes gráficos las volatilidades ajustadas y la serie de correlaciones.



El Modelo BEKK. (Baba, Engle, Kraft and Kroner). Para garantizar que la matriz Σ se definida positiva, en Engle y Kroner (1995) se propone el modelo BEKK,

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

donde A es una matriz triangular inferior y A_i y B_j son matrices de k x k.

- ✓ Basados en la parametrización simétrica del modelo, Σ_t es casi seguramente definida positiva dado que AA' es definida positiva.

- ✓ Este modelo también permite introducir dependencia dinámica entre las series de volatilidad.

- ✓ Por otro lado, el modelo tiene varias desventajas:

⇒ Los parámetros en A_i y B_j no tienen interpretaciones directas sobre los valores rezagados de las volatilidades o de los shocks.

⇒ El número de parámetros empleados es $k^2(m+s)+k(k+1)/2$, el cual incrementa rápidamente con m y s.

⇒ Alguna experiencia muestra que muchos de los parámetros estimados son estadísticamente no significativos, lo que introduce complicaciones en la modelación.

Ejemplo

Consideremos de nuevo las series de retornos mensuales de las acciones de Pfizer y Merck. Para un modelo BEKK(1,1) los resultados son:

Coeficientes estimados:

	Valor	ErrorSTD	t	Valor p
C(1)	0.0164770	0.003470	4.749e+00	1.369e-06
C(2)	0.0142816	0.003172	4.503e+00	4.255e-06
A(1,1)	0.0245803	0.008837	2.782e+00	2.815e-03
A(2,1)	0.0116134	0.005953	1.951e+00	2.584e-02
A(2,2)	0.0002018	0..267625	7.541e-04	4.997e-01
ARCH(1,1,1)	0.2994052	0.093304	3.209e+00	7.125e-04
ARCH(1,2,1)	0.1952856	0.075092	2.601e+00	4.802e-03
ARCH(1,1,2)	-0.0818745	0.097810	-8.371e-01	2.015e-01
ARCH(1,2,2)	0.0929540	0.082626	1.125e+00	1.306e-01
GARCH(1,1,1)	0.8987843	0.074407	1.208e+01	0.000e+00
GARCH(1,2,1)	-0.0674587	0.059595	-1.132e+00	1.291e-01
GARCH(1,1,2)	0.0163848	0.046402	3.531e-01	3.621e-01
GARCH(1,2,2)	0.9889547	0.040158	2.463e+01	0.000e+00

Prueba de Ljung-Box para los residuales estandarizados:

	Estadístico	Valor p	g.l de la chi-cuadrado
Pfe	10.13	0.6044	12
Mrk	15.25	0.2278	12

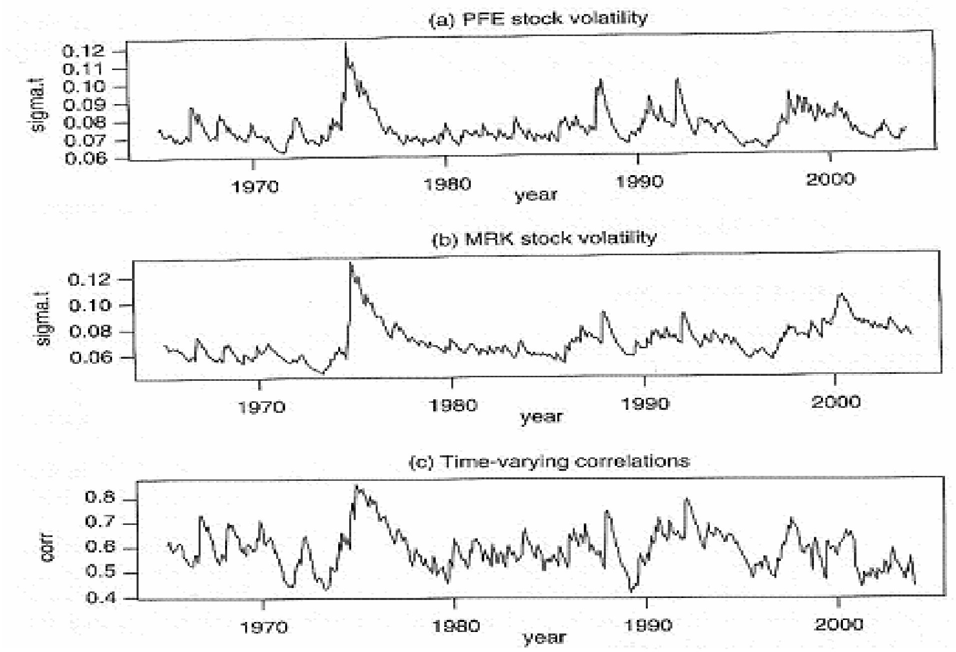
Prueba de Ljung-Box para los residuales estandarizados al cuadrado:

	Estadístico	Valor p	g.l de la chi-cuadrado
Pfe	18.314	0.1065	12
Mrk	7.174	0.8459	12

- ✓ Los estadísticos de Ljung-Box para las series individuales de residuales y los cuadrados de los retornos muestran que el modelo BEKK(1,1) no puede ser rechazado.

- ✓ Usando el estadístico Q multivariado de la serie bivariada de los residuales estandarizados y la serie bivariada de los residuales estandarizados al cuadrado obtenemos $Q_2(10)=41.57(0.40)$ y $Q_2^*(10)=65.71(0.006)$. Por tanto, en forma similar al modelo DVEC(1,1), también de rechaza el modelo para la volatilidad a un nivel de significancia del 1%.

- ✓ El siguiente gráfico muestra las volatilidades ajustadas y las correlaciones tiempo variantes del modelo BEKK(1,1) ajustado. Comparadas con las del modelo DVEC(1,1) existen algunas diferencias: por ejemplo, las correlaciones del modelo BEKK(1,1) parecen ser más volátiles.



✓ La ecuación de volatilidad del modelo BEKK(1,1) ajustado es

$$
\begin{bmatrix} \sigma_{11,t}^2 & \sigma_{12,t}^2 \\ \sigma_{21,t}^2 & \sigma_{22,t}^2 \end{bmatrix} = \begin{bmatrix} 0.025 & 0 \\ 0.012 & 0.0002 \end{bmatrix} \begin{bmatrix} 0.025 & 0.012 \\ 0 & 0.0002 \end{bmatrix} +
$$

$$
\begin{bmatrix} 0.299 & -0.082 \\ 0.195 & 0.093 \end{bmatrix} \begin{bmatrix} a_{1,t-1}^2 & a_{1,t-1}a_{2,t-1} \\ a_{2,t-1}a_{1,t-1} & a_{2,t-1}^2 \end{bmatrix} \begin{bmatrix} 0.299 & 0.195 \\ -0.082 & 0.093 \end{bmatrix} +
$$

$$
\begin{bmatrix} 0.899 & 0.016 \\ -0.067 & 0.989 \end{bmatrix} \begin{bmatrix} \sigma_{11,t-1}^2 & \sigma_{12,t-1}^2 \\ \sigma_{21,t-1}^2 & \sigma_{22,t-1}^2 \end{bmatrix} \begin{bmatrix} 0.899 & -0.067 \\ 0.016 & 0.989 \end{bmatrix}
$$

donde seis estimaciones son insignificantes a un nivel del 5%. Específicamente, la matriz de constantes A, solamente tiene un solo elemento significativo. Para descifrar el modelo es necesario realizar los productos matriciales.

REPARAMETRIZACIÓN

- Un paso útil en la modelación de la volatilidad multivariada es reparametrizar a Σ_t haciendo uso de su propiedad de simetría.

- **Uso de las correlaciones:** La primera reparametrización de Σ_t es usar los coeficientes de correlación condicionales y varianzas condicionales de a_t. Específicamente, Σ_t puede ser escrito como

$$\Sigma_t \equiv [\sigma_{ij,t}] = D_t \rho_t D_t$$

donde ρ_t es la matriz es correlación condicional de a_t y *D*$_t$ es una matriz diagonal de k x k cuya diagonal contiene las desviaciones estándar condicionales de los elementos de a_t.

✓ Debido a que la matriz ρ_t es simétrica con los elementos de la diagonal iguales a uno, la evolución en el tiempo de Σ_t está gobernada por las varianzas condicionales $\sigma_{ii,t}$ y los elementos $\rho_{ij,t}$ de ρ_t, donde j<i y $1 \le i \le k$. Por tanto, para modelar la volatilidad de a_t es suficiente considerar las varianzas condicionales y los coeficientes de correlación condicionales de las a_{it}.

✓ Defina el vector k(k+1)/2 dimensional (ksi)

$$\Xi_t = (\sigma_{11,t}, \sigma_{22,t}, \cdots, \sigma_{kk,t}, Q_t^{'})'$$

donde $Q_t^{'}$ es un vector k(k-1)/2 dimensional que se forma colocando las columnas de ρ_t una encima de otra, pero solamente usando los elementos debajo de la diagonal principal, es decir,

$$\varrho_t = (\rho_{21,t}, \cdots, \rho_{k1,t} \mid \rho_{32,t}, \cdots, \rho_{k2,t} \mid \cdots \mid, \rho_{k,k-1,t})'$$

Para el caso de k=2, $\varrho_t = \rho_{21,t}$ y entonces

$$\Xi_t = (\sigma_{11,t}, \sigma_{22,t}, \rho_{21,t})'$$

Para k=3, se obtiene que

$$\varrho_t = (\rho_{21,t}, \rho_{31,t}, \rho_{32,t})'$$

y,

$$\Xi_t = (\sigma_{11,t}, \sigma_{22,t}, \sigma_{33,t}, \rho_{21,t}, \rho_{31,t}, \rho_{32,t})'$$

- ✓ Si a_t es un vector aleatorio normal bivariante, entonces $\Xi = (\sigma_{11,t}, \sigma_{22,t}, \rho_{21,t})'$ y la f.d.p condicional de a_t dado F_{t-1} es

$$f(a_{1t},\ a_{2t} \mid F_{t-1}) = \frac{1}{2\pi\sqrt{\sigma_{11,t}\sigma_{22,t}(1-\rho_{12.t}^2)}}\exp(-\frac{Q(a_{1t},\ a_{2t}, \Xi_t)}{2(1-\rho_{12.t}^2)}$$

donde $Q(a_{1t}, a_{2t}, \Xi_t) = \dfrac{a_{1t}^2}{\sigma_{11,t}} + \dfrac{a_{2t}^2}{\sigma_{22,t}} + \dfrac{2\rho_{21}a_{1t}a_{2t}}{\sqrt{\sigma_{11,t}\sigma_{22,t}}}$.

- ✓ Para la estimación máximo verosímil, el log de la f.d.p de a_t es

$$l(a_{1t},\ a_{2t}, \Xi_t) \infty -\frac{1}{2}\left\{\ln[\sigma_{11,t}\sigma_{22,t}(1-\rho_{12.t}^2)] + \frac{1}{(1-\rho_{12.t}^2)}\left(\frac{a_{1t}^2}{\sigma_{11,t}} + \frac{a_{2t}^2}{\sigma_{22,t}} + \frac{2\rho_{21.t}a_{1t}a_{2t})}{\sqrt{\sigma_{11,t}\sigma_{22,t}}}\right)\right\}$$

- ✓ Esta reparametrización es útil debido a que permite modelar las varianzas y las correlaciones condicionales directamente. Sin embargo, esta aproximación todavía presenta algunas debilidades:

  ⇒ La función de verosimilitud es complicada cuando $k \geq 3$.

  ⇒ La aproximación requiere una maximización restringida en la estimación de forma que asegure que la matriz Σ_t sea definida positiva. La restricción se vuelve complicada para k grande.

- **La descomposición de Cholesky:** Esta descomposición tiene algunas ventajas en la estimación debido a que no necesita restricciones sobre los parámetros para que la matriz Σ_t sea definida positiva. Además, la reparametrización es una transformación ortogonal, de manera que la función de verosimilitud resultante es extremadamente simple.

  ✓ Debido a que Σ_t es definida positiva, existe una matriz triangular inferior L_t con unos en su diagonal principal y una diagonal matriz G_t con elementos positivos en su diagonal tales que

$$\Sigma_t = L_t G_t L_t^{'}$$

  Una característica de esta descomposición es que los elementos fuera de la diagonal de L_t y los elementos de la diagonal de G_t tienen interpretaciones interesantes.

  ✓ Considere el caso bivariado. Entonces,

$$\Sigma_t = \begin{bmatrix} \sigma_{11,t} & \sigma_{21,t} \\ \sigma_{21,t} & \sigma_{22,t} \end{bmatrix}, \quad L_t = \begin{bmatrix} 1 & 0 \\ q_{21,t} & 1 \end{bmatrix}, \quad G_t = \begin{bmatrix} g_{11.t} & 0 \\ 0 & g_{22,t} \end{bmatrix}$$

donde $g_{ii,t}>0$ para i=1,2. Puesto que $\Sigma_t = L_t G_t L_t^{'}$, se obtiene

$$\Sigma_t = \begin{bmatrix} \sigma_{11,t} & \sigma_{21,t} \\ \sigma_{21,t} & \sigma_{22,t} \end{bmatrix} = \begin{bmatrix} g_{11,t} & q_{21,t} g_{11,t} \\ q_{21,t} g_{11,t} & g_{22,t} + q_{21,t}^2 g_{11,t} \end{bmatrix}$$

de donde:

- ✓ $\sigma_{11,t} = g_{11,t}$, $\sigma_{21,t} = q_{21,t} g_{11,t}$ y $\sigma_{22,t} = g_{22,t} + q_{21,t}^2 g_{11,t}$

- ✓ Resolviendo las ecuaciones anteriores, se obtiene

$$g_{11,t} = \sigma_{11,t}\,,\quad q_{21,t} = \frac{\sigma_{21,t}}{\sigma_{11,t}} \quad \text{y} \quad g_{22,t} = \sigma_{22,t} - \frac{\sigma_{21,t}^2}{\sigma_{11,t}}$$

- ✓ Considere la regresión lineal simple de a_{2t} sobre a_{1t}

$$a_{2t} = \beta a_{1t} + b_{2t}$$

donde b_{2t} denota el término de error, el cual no está correlacionado con el regresor a_{1t}. De la teoría de mínimos cuadrados se obtiene que

$$\Rightarrow \beta = \frac{\operatorname{cov}(a_{1t}, a_{2t})}{Var(a_{1t})} = \frac{\sigma_{21,t}}{\sigma_{11,t}}$$

$$\Rightarrow Var(b_{2t}) = Var(a_{2t}) - \beta^2 [Var(a_{1t})]^{-1} = \sigma_{22,t} - \frac{\sigma_{21,t}^2}{\sigma_{11,t}}$$

- ✓ En consecuencia, se obtiene que:

$g_{11,t} = \sigma_{11,t}$, $q_{21,t} = \beta$ y $g_{22,t} = Var(b_{2t})$ y $b_{2t} \perp a_{1t}$

donde $\perp$ indica no correlación.

✓ Resumiendo, la descomposición de Cholesky de una matriz Σ_t de 2x2 permite realizar una transformación ortogonal de a_t a b_t =(b_{1t}, b_{2t})', $b_t = L_t a_t$ tal que:

$\Rightarrow$ $b_{1t} = a_{1t}$, y $b_{2t} = a_{2t} - q_{21,t} a_{1t}$

$\Rightarrow$ $q_{21,t} = \beta$ se obtiene de la regresión lineal simple $a_{2t} = \beta a_{1t} + b_{2t}$

$\Rightarrow$ Cov(b_t) es una matriz diagonal con los $g_{ii,t}$ en su diagonal.

✓ Las cantidades $q_{21,t}$ y $g_{ii,t}$ tienen la siguiente interpretación:

$\Rightarrow$ El primer elemento de la diagonal de G_t es la varianza de a_{1t} .

$\Rightarrow$ El segundo elemento de la diagonal de G_t es la varianza residual de la regresión simple $a_{2t} = \beta a_{1t} + b_{2t}$.

$\Rightarrow$ El elemento $q_{21,t}$ es el coeficiente β en la regresión simple $a_{2t} = \beta a_{1t} + b_{2t}$.

El proceso anterior se mantiene para el caso de dimensión mayor. Por ejemplo considere el caso tridimensional, para el cual

$$L_t = \begin{bmatrix} 1 & 0 & 0 \\ q_{21,t} & 1 & 0 \\ q_{31,t} & q_{32,t} & 1 \end{bmatrix}, \quad G_t = \begin{bmatrix} g_{11.t} & 0 & 0 \\ 0 & g_{22,t} & 0 \\ 0 & 0 & g_{33,t} \end{bmatrix}$$

Puesto que $\Sigma_t = L_t G_t L_t^{'}$, entonces

$$\begin{bmatrix} \sigma_{11,t} & \sigma_{21,t} & \sigma_{31,t} \\ \sigma_{21,t} & \sigma_{22,t} & \sigma_{32,t} \\ \sigma_{31,t} & \sigma_{32,t} & \sigma_{33,t} \end{bmatrix} = \begin{bmatrix} g_{11,t} & q_{21,t}g_{11,t} & q_{31,t}g_{11,t} \\ q_{21,t}g_{11,t} & q_{21,t}^2 g_{11,t} + g_{22,t} & q_{31,t}q_{21,t}g_{11,t} + q_{32,t}g_{22,t} \\ q_{31,t}g_{11,t} & q_{31,t}q_{21,t}g_{11,t} + q_{32,t}g_{22,t} & q_{31,t}^2 g_{11,t} + q_{32,t}^2 g_{22,t} + g_{33,t} \end{bmatrix}$$

Igualando términos se obtiene:

$$\sigma_{11,t} = g_{11,t}, \quad \sigma_{21,t} = q_{21,t}g_{11,t}, \quad \sigma_{22,t} = q_{21,t}^2 g_{11,t} + g_{22,t}, \quad \sigma_{31,t} = q_{31,t}g_{11,t},$$

$$\sigma_{32,t} = q_{31,t}q_{21,t}g_{11,t} + q_{32,t}g_{22,t}, \quad \sigma_{33,t} = q_{31,t}^2 g_{11,t} + q_{32,t}^2 g_{22,t} + g_{33,t}$$

Resolviendo las ecuaciones anteriores, se obtiene

$$g_{11,t} = \sigma_{11,t}, \quad q_{21,t} = \frac{\sigma_{21,t}}{\sigma_{11,t}}, \quad g_{22,t} = \sigma_{22,t} - q_{21,t}^2 g_{11,t}, \quad q_{31,t} = \frac{\sigma_{31,t}}{\sigma_{11,t}},$$

$$q_{32,t} = \frac{1}{g_{22,t}}\left(\sigma_{32,t} - \frac{\sigma_{31,t}\sigma_{21,t}}{\sigma_{11,t}}\right), \quad g_{33,t} = \sigma_{33,t} - q_{31,t}^2 g_{11,t} - q_{32,t}^2 g_{22,t}$$

Estas cantidades son simplemente los coeficientes y varianzas residuales de la transformación ortogonal

$$b_{1t} = a_{1t}$$
$$b_{2t} = a_{2t} - \beta_{21}b_{1t}$$
$$b_{3t} = a_{3t} - \beta_{31}b_{1t} - \beta_{32}b_{2t}$$

donde los β_{ij} son coeficientes de las regresiones de mínimos cuadrados,

$$a_{2t} = \beta_{21}b_{1t} + b_{2t}$$
$$a_{3t} = \beta_{31}b_{1t} + \beta_{32}b_{2t} + b_{3t}$$

⇨ Resumiendo, $q_{ij,t} = \beta_{ij}$, $g_{ii,t} = Var(b_{it})$ y $b_{it} \perp b_{jt}$, para $i \neq j$.

En general, el uso de la descomposición de Cholesky permite realizar una transformación ortogonal de a_t a b_t, donde $a_{1t}=b_{1t}$ y para $1<i\leq k$, b_{it} está definida recursivamente por la regresión de mínimos cuadrados

$$a_{it} = q_{i1,t}b_{1t} + q_{i2,t}b_{2t} + \cdots + q_{i(i-1),t}b_{(i-1)t} + b_{it}$$

donde $q_{ij,t}$ es el (i,j)-ésimo elemento de la matriz triangular inferior L_t, para $1\leq j<i$.

- ✓ La transformación puede ser escrita como

$$b_t = L_t^{-1}a_t, \text{ o } a_t = L_t b_t$$

- ✓ Cov(b_t)=$L_t^{-1}\Sigma_t(L_t^{-1})'=G_t$

- ✓ El vector de parámetros relevantes al modelo de volatilidad bajo tal transformación es

$$\Xi_t = (g_{11,t}, \ \cdots, g_{kk,t}, q_{21,t}, q_{31,t}, \ q_{32,t}, \cdots, q_{k1,t}, \cdots, q_{k(k-1),t})'$$

El cual es k(k+1)/2 dimensional.

- ✓ La transformación ortogonal anterior simplifica dramáticamente la función de verosimilitud de los datos:

  ⇒ Usando el hecho de que $|L_t|=1$,

$$|\Sigma_t| = |L_t G_t L_t'| = |G_t| = \prod_{i=1}^{k} g_{ii,t} \ ,$$

⇒ Si la distribución condicional de a_t es normal multivariada, N(0, Σ_t), entonces la distribución condicional de la serie transformada b_t es N(0, G_t) y por tanto el log de la función de verosimilitud es muy simple.

⇒ El log de la fdp de a_t es (omitiendo la constante)

$$l(a_t, \Sigma_t) = l(b_t, \Xi_t) = -\frac{1}{2}\sum_{i=1}^{k}\left(\ln(g_{ii,t}) + \frac{b_{it}^2}{g_{ii,t}}\right)$$

El uso de la descomposición de Cholesky para reparametrizar a Σ_t tiene varias ventajas:

⇒ Σ_t es definida positiva si $g_{ii,t}$ >0 para todo i. por tanto la restricción de que Σ_t debe ser definida positiva, se logra modelando $\ln(g_{ii,t})$ en lugar de $g_{ii,t}$.

⇒ Los elementos del vector de parámetros Ξ_t tienen interpretaciones interesantes, pues estas cantidades son simplemente los coeficientes y varianzas residuales de regresiones lineales múltiples que ortogonalizan los shocks.

⇒ El coeficiente de correlación entre a_{1t} y a_{2t} es

$$\rho_{12,t} = \frac{\sigma_{12,t}}{\sqrt{\sigma_{11,t}\sigma_{22,t}}} = q_{21,t}\frac{\sqrt{\sigma_{11,t}}}{\sqrt{\sigma_{22,t}}}$$

el cual es tiempo variante si $q_{21,t} \neq 0$. En particular si $q_{21,t} = c \neq 0$,

entonces $\rho_{12,t} = c\frac{\sqrt{\sigma_{11,t}}}{\sqrt{\sigma_{22,t}}}$, la cual es tiempo variante si $\frac{\sigma_{11,t}}{\sigma_{22,t}}$ no es una constante.

$\Rightarrow$ Esta propiedad tiempo variante se aplica a los otros coeficientes de correlación cuando la dimensión de r_t es mayor que 2 y es la mayor diferencia entre las dos aproximaciones de reparametrización de Σ_t.

$\Rightarrow$ Usando la ecuación $a_{it} = q_{i1,t}b_{1t} + q_{i2,t}b_{2t} + \cdots + q_{i(i-1),t}b_{(i-1)t} + b_{it}$, y la ortogonalidad entre los shocks transformados b_{it}, se obtiene

$$\sigma_{ii,t} = Var(a_{it} \mid F_{t-1}) = \sum_{v=1}^{i} q_{iv,t}^2 g_{vv,t}\ , \quad \text{i=1,2,…,k.}$$

$$\sigma_{ij,t} = Cov(a_{it},\, a_{jt} \mid F_{t-1}) = \sum_{v=1}^{j} q_{iv,t} q_{jv,t} g_{vv,t}\ , \quad \text{j<i,} \quad \text{i=1,2,…,k.}$$

donde $q_{vv,t} = 1$, para v=1,2,…,k. Estas ecuaciones muestran la reparametrización de Σ_t bajo la descomposición de Cholesky.

Modelos GARCH para Retornos Bivariantes

- En una serie de retornos k-dimensional r_t, un modelo GARCH multivariado usa “ecuaciones exactas” para describir la evolución en el tiempo del vector k(k+1)/2 dimensional Ξ_t.

- El término de “ecuaciones exactas” indica que la ecuación no contiene ningún término aleatorio. Sin embargo, la ecuación puede llegar a ser complicada aún en el caso más simple de k=2, para el cual Ξ_t es tridimensional. Para producir modelos simples, generalmente se imponen restricciones sobre las ecuaciones.

Modelos de Correlación Constante

- Con el fin de mantener un bajo número de ecuaciones de volatilidad, Bollerslev (1990), considera el caso para el cual el coeficiente de correlación $\rho_{21,t} = \rho_{21}$ es invariante en el tiempo, y donde $|\rho_{21}|<1$.

- Bajo este supuesto, ρ_{21} es un parámetro constante y el modelo de volatilidad consta de dos ecuaciones para Ξ_t^*, donde $\Xi_t^* = (\sigma_{11,t}, \sigma_{22,t})'$. Un modelo GARCH(1,1) para Ξ_t^* es de la forma

$$\Xi_t^* = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \Xi_{t-1}^*$$

donde $a_{t-1}^2 = (a_{1,t-1}^2, a_{2,t-1}^2)'$, α_0 es un vector positivo bidimensional y α_1 y β_1 son matrices de 2 x 2 no definidas negativas.

- Específicamente el modelo puede ser expresado como

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

donde α_{i0} >0, para i=1,2.

- Defina $\eta_t = a_t^2 - \Xi_t^*$. Entonces el modelo anterior puede ser re-escrito como

$$a_t^2 = \alpha_0 + (\alpha_1 + \beta_1) a_{t-1}^2 + \eta_t - \beta_1 \eta_{t-1}$$

el cual es un modelo ARMA(1,1) bivariado para el proceso a_t^2. Este resultado es una generalización del modelo univariado GARCH(1,1).

- Por tanto, las propiedades del modelo

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

Se derivan de las del modelo bivariado ARMA(1,1). Específicamente,

✓ Si todos los valores propios de $\alpha_1 + \beta_1$ son positivos y menores que 1, el modelo bivariado ARMA(1,1) para a_t^2 es débilmente estacionario y por tanto E(a_t^2) existe. Esto implica que:

⇒ El proceso a_t tiene una matriz de covarianza no condicional que es definida positiva.

⇒ Las varianzas no condicionales de los elementos de a_t son $(\sigma_1^2, \sigma_2^2)' = (I - \alpha_1 - \beta_1)^{-1}\phi_0$.

⇒ La covarianza no condicional entre a_{1t} y a_{2t} es $\rho_{21}\sigma_1\sigma_2$.

✓ Si $\alpha_{12} = \beta_{12} = 0$, entonces la volatilidad de a_{1t} no depende del pasado de la volatilidad de a_{2t}. Similarmente, si $\alpha_{21} = \beta_{21} = 0$, entonces la volatilidad de a_{2t} no depende del pasado de la volatilidad de a_{1t}.

✓ Si α_1 y β_1 son diagonales, el modelo se reduce a dos modelos GARCH(1,1) univariados. En este caso, los procesos de volatilidad no están dinámicamente relacionados.

✓ Los pronósticos de la volatilidad del modelo pueden ser obtenidos usando métodos similares a los del modelo vectorial ARMA(1,1):

⇒ El pronóstico 1-paso adelante en el origen h es

$$\Xi_h^*(1) = \alpha_0 + \alpha_1 a_h^2 + \beta_1 \Xi_h^*$$

⇒ El pronóstico l-pasos adelante en el origen h es

$$\Xi_h^*(l) = \alpha_0 + (\alpha_1 + \beta_1)\Xi_h^*(l-1)\,, \quad l > 1.$$

✓ Los pronósticos anteriores son para la volatilidades marginales de las a_{it}. El pronóstico l-pasos adelante en el origen h para la covarianza entre a_{1t} y a_{2t} es $\hat{\rho}_{21}/[\sigma_{11,h}(l)\sigma_{22,h}(l)]^{0.5}$, donde $\hat{\rho}_{21}$ es la estimación de ρ_{21} y los $\sigma_{ii,h}(l)$ son los elementos de $\Xi_h^*(l)$.

Ejemplo 1.

Considere de nuevo los retornos diarios de índices de las bolsas de Hong Kong y Japón. Usando un modelo bivariado GARCH, se obtienen dos modelos que ajustan bien los datos.

1) Primer modelo:

✓ Ecuación de media:

$$r_{1t} = \underset{(0.044)}{-0.118} r_{1,t-6} + a_{1t}$$

$$r_{2t} = a_{2t}$$

✓ Ecuación de volatilidad:

$$\begin{bmatrix}\sigma_{11,t}\\ \sigma_{22,t}\end{bmatrix}=\begin{bmatrix}0.275\\(0.079)\\0.051\\(0.014)\end{bmatrix}+\begin{bmatrix}0.112 & .\\(0.032) & \\ . & 0.091\\ & (0.026)\end{bmatrix}\begin{bmatrix}a_{1,t-1}^2\\ a_{2,t-1}^2\end{bmatrix}+\begin{bmatrix}0.711 & .\\(0.068) & \\ . & 0.869\\ & (0.028)\end{bmatrix}\begin{bmatrix}\sigma_{11,t-1}\\ \sigma_{22,t-1}\end{bmatrix}$$

donde los números entre paréntesis son los errores estándar. El coeficiente de correlación estimado entre a_{1t} y a_{2t} es 0.226 con un error estándar de 0.047.

✓ Para las series de residuales estandarizados, los estadísticos de Ljung-Box $Q_2(4)$=22.29(0.10) y $Q_2(8)$=34.83(0.29). Basados en las distribuciones chi-cuadrado con 15 y 31 grados de libertad (debido a que se usa un coeficiente AR(6) en la ecuación de media), concluimos que no rechazamos el modelo de media.

✓ Para las series de residuales estandarizados al cuadrado, los estadísticos de Ljung-Box $Q_2^*(4)$=9.54(0.85) y $Q_2^*(8)$=18.58(0.96). Concluimos que no rechazamos el modelo de volatilidad.

✓ Las varianzas para las series de innovaciones son, respectivamente, 1.55 y 1.28.

✓ El modelo estimado para la volatilidad muestra dos ecuaciones de volatilidad no relacionadas, indicando que las volatilidades de los dos mercados no están relacionadas dinámicamente, pero si están correlacionadas contemporáneamente.

✓ Este modelo es denominado *modelo bivariado diagonal de correlación constante.*

2) Segundo modelo:

✓ Ecuación de media:

$$r_{1t} = -0.143 r_{1,t-6} + a_{1t}$$
$$(0.042)$$

$$r_{2t} = a_{2t}$$

✓ Ecuación de volatilidad:

$$\begin{bmatrix} \sigma_{11,t} \\ \sigma_{22,t} \end{bmatrix} = \begin{bmatrix} 0.378 \\ (0.103) \\ . \\ \end{bmatrix} + \begin{bmatrix} 0.108 & . \\ (0.030) & \\ . & 0.172 \\ & (0.035) \end{bmatrix} \begin{bmatrix} a_{1,t-1}^2 \\ a_{2,t-1}^2 \end{bmatrix} + \begin{bmatrix} . & 0.865 \\ & (0.109) \\ 0.321 & 0.869 \\ (0.135) & (0.028) \end{bmatrix} \begin{bmatrix} \sigma_{11,t-1} \\ \sigma_{22,t-1} \end{bmatrix}$$

donde los números entre paréntesis son los errores estándar. El coeficiente de correlación estimado entre a_{1t} y a_{2t} es 0.236 con un error estándar de 0.045.

✓ Para las series de residuales estandarizados, los estadísticos de Ljung-Box $Q_2(4)$=24.22(0.06) y $Q_2(8)$=35.52(0.26). Basados sobre las distribuciones chi-cuadrado con 15 y 31 grados de libertad, concluimos que no rechazamos el modelo de media a un nivel del 5%.

✓ Para las series de residuales estandarizados al cuadrado, los estadísticos de Ljung-Box $Q_2^*(4)$=17.45(0.29) y $Q_2^*(8)$=24.55(0.79). Concluimos que no rechazamos el modelo de volatilidad.

✓ Las varianzas para las series de innovaciones son, respectivamente, 1.71 y 1.32.

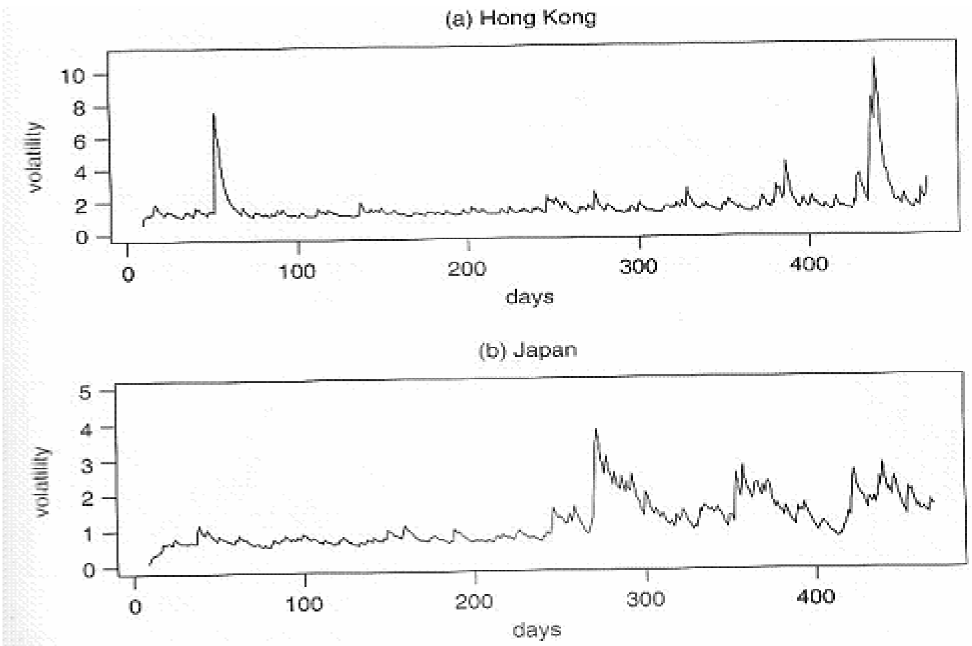
- ✓ En contraste con el modelo anterior, este modelo GARCH bivariado permite retroalimentación entre los dos mercados.

- ✓ Comparación de las volatilidades de los dos modelos:

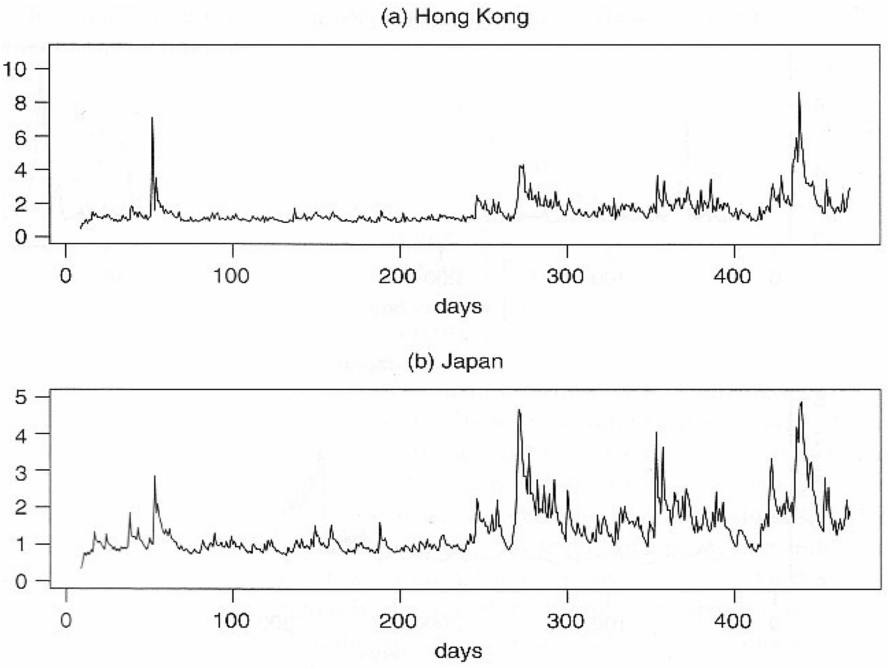
  ⇒ Primero, las varianzas incondicionales para las series de innovaciones para el segundo modelo están más cerca de las de los modelos univariados ajustados antes para dichas series.

  ⇒ Segundo, los siguientes gráficos muestran los procesos de volatilidad ajustada del primer y segundo modelo. Debido a que el primer modelo no permite dependencia dinámica de la volatilidad entre los dos mercados, sus volatilidades ajustadas son más similares a las obtenidas para los modelos GARCH univariados. En contrate, los gráficos de las volatilidades del segundo modelo presentan evidencia de impactos mutuos entre los dos mercados.

Volatilidades estimadas usando el modelo 1



Volatilidades estimadas usando el modelo 2



⇒ Tercero, la función de verosimilitud maximizada del primer modelo es -535.13 para t=8,…469, mientras que la del segundo modelo es -540.32. Por tanto, si se usa dicho criterio, se prefiere el primer modelo.

⇒ Finalmente, debido a que las implicaciones prácticas de los dos modelos difieren mucho, es necesario investigación adicional para separar los modelos. Tal investigación podría usar un período muestral más grandes o incluir más retornos (por ejemplo, usar el índice de la bolsa de E.U.)

Ejemplo 2.

Considere los retornos mensuales en porcentajes de las acciones IBM y del índice S&P 500, desde Enero de 1926 hasta Diciembre de 1999. Sean r_{1t} y r_{2t} los retornos de IBM y S&P 500, respectivamente.

Si se considera un modelo de correlación constante, se obtienen las siguientes ecuaciones:

✓ Para la media:

$$\underset{}{r_{1t}} = \underset{(0.225)}{1.351} + \underset{(0.029)}{0.0723} r_{1,t-1} + \underset{(0.034)}{0.055} r_{1,t-2} - \underset{(0.044)}{0.119} r_{2,t-1} + a_{1t}$$

$$r_{2t} = \underset{(0.155)}{0.703} + a_{2t}$$

✓ Ecuación de volatilidad:

$$\begin{bmatrix} \sigma_{11,t} \\ \sigma_{22,t} \end{bmatrix} = \begin{bmatrix} 0.298 \\ (0.59) \\ 0.209 \\ (0.47) \end{bmatrix} + \begin{bmatrix} 0.079 & . \\ (0.013) & \\ 0.042 & 0.045 \\ (0.009) & (0.010) \end{bmatrix} \begin{bmatrix} a_{1,t-1}^2 \\ a_{2,t-1}^2 \end{bmatrix} + \begin{bmatrix} 0.873 & -0.031 \\ (0.020) & (0.009) \\ -0.066 & 0.913 \\ (0.015) & (0.014) \end{bmatrix} \begin{bmatrix} \sigma_{11,t-1} \\ \sigma_{22,t-1} \end{bmatrix}$$

✓ El coeficiente de correlación contemporáneo es 0.614 con error estándar de 0.020.

✓ Para las series de residuales estandarizados, los estadísticos de Ljung-Box $Q_2(4)$=16.77(0.21) y $Q_2(8)$=32.40(0.30). Basados sobre las distribuciones chi-cuadrado con 13 y 29 grados de libertad (ajustados por los parámetros en el modelo de media), se concluye que no rechazamos el modelo de media.

✓ Para las series de residuales estandarizados al cuadrado, los estadísticos de Ljung-Box $Q_2^*(4)$=18.00(0.16) y $Q_2^*(8)$=39.09(0.10). Se concluye que no rechazamos el modelo de volatilidad a un nivel del 5%.

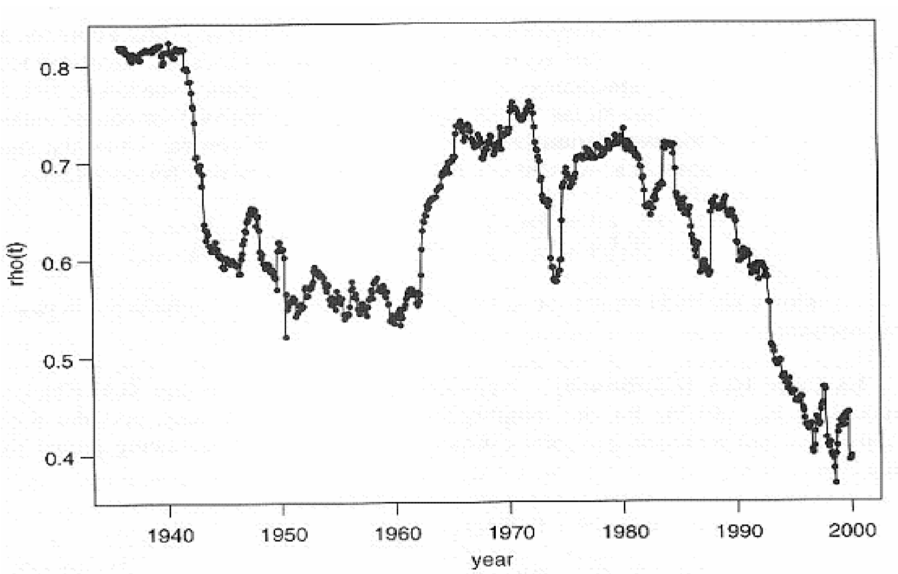
✓ Este modelo GARCH(1,1) bivariado muestra relaciones de retroalimentación entre las volatilidades de los dos retornos mensuales.

Modelos de Correlación Tiempo Variante

- Una falla importante del modelo de volatilidad de correlación constante es que en las aplicaciones reales, el coeficiente de correlación tiende a cambiar en el tiempo.

- Por ejemplo, considere los retornos mensuales de las acciones de IBM y del índice S&P 500. El siguiente gráfico muestra el coeficiente de correlación muestral entre las dos series mensuales usando una ventana móvil de 120 observaciones (10 años), para datos desde Enero de 1925 hasta Diciembre de 1999.

✓ Es difícil justificar que el índice S&P 500, el cual es un promedio ponderado, mantenga una correlación constante con los retornos de IBM durante los pasados 70 años.



✓ La correlación cambia con el tiempo y parece ser decreciente en los últimos años.

✓ La tendencia decreciente en la correlación no es sorprendente debido a que el ranking del mercado de capitalización de IBM entre las grandes compañías industriales de E.U. ha cambiado recientemente.

- Una forma simple de relajar la restricción de correlación constante dentro en los modelos GARCH, es la de especificar una ecuación exacta para el coeficiente de correlación condicional. Esto se puede hacer por dos métodos usando las reparametrizaciones vistas anteriormente.

✓ Primero, usando el **coeficiente de correlación** directamente. Puesto que el coeficiente de correlación entre los retornos de las acciones de IBM y el índice S&P 500 es positivo y debe caer en el intervalo [0,1], se emplea la siguiente ecuación

$$\rho_{21,t} = \frac{\exp(q_t)}{1+\exp(q_t)}$$

donde

$$q_t = \varpi_0 + \varpi_1 \rho_{21,t-1} + \varpi_2 \frac{a_{1,t-1}a_{2,t-1}}{\sqrt{\sigma_{11,t-1}\sigma_{22,t-1}}}$$

donde $\sigma_{ii,t-1}$ es la varianza condicional de $a_{i,t-1}$.

⇒ Esta ecuación es denominada el modelo GARCH(1,1) para el coeficiente de correlación debido a que usa el rezago 1 del coeficiente de correlación y el rezago 1 del producto cruzado de los dos shocks.

⇒ Si $\varpi_1 = \varpi_2 = 0$, el modelo para $\rho_{21,t}$ se reduce al modelo de correlación constante.

⇒ Resumiendo, un modelo bivariado GARCH(1,1) de correlación tiempo variante consta de dos conjuntos de ecuaciones:

- El primer conjunto consiste de un modelo bivariado GARCH(1,1) para las varianzas condicionales.

- El segundo es un modelo GARCH(1,1) para la correlación.

⇒ En la práctica se puede agregar un signo negativo a la ecuación vista para $\rho_{21,t}$, si la correlación es negativa.

⇒ En general, cuando el signo de la correlación es desconocido, podemos usar la transformación de Fisher para la correlación dada por

$$q_t = \ln\left(\frac{1+\rho_{21,t}}{1-\rho_{21,t}}\right) \quad \text{o} \quad \rho_{21,t} = \frac{\exp(q_t)-1}{\exp(q_t)+1}$$

y emplear un modelo GARCH para q_t, para modelar la correlación entre los dos retornos.

Ejemplo

Para los retornos mensuales de las acciones de IBM y del índice S&P 500, se obtiene el siguiente modelo:

✓ Ecuaciones para la media:

$$\begin{aligned} r_{1t} &= \underset{(0.215)}{1.318} + \underset{(0.026)}{0.076} r_{1,t-1} - \underset{(0.034)}{0.068} r_{2,t-2} + a_{1t} \\ r_{2t} &= \underset{(0.151)}{0.673} + a_{2t} \end{aligned}$$

✓ Las ecuaciones de volatilidad son

$$\begin{bmatrix}\sigma_{11,t}\\ \sigma_{22,t}\end{bmatrix}=\begin{bmatrix}2.80\\(0.58)\\1.71\\(0.49)\end{bmatrix}+\begin{bmatrix}0.084 & .\\(0.013) & \\0.037 & 0.054\\(0.009) & (0.010)\end{bmatrix}\begin{bmatrix}a^2_{1,t-1}\\ a^2_{2,t-1}\end{bmatrix}+\begin{bmatrix}0.864 & -0.020\\(0.021) & (0.009)\\-0.058 & 0.914\\(0.014) & (0.013)\end{bmatrix}\begin{bmatrix}\sigma_{11,t-1}\\ \sigma_{22,t-1}\end{bmatrix}$$

✓ La ecuación de correlación condicional es

$$\rho_{21,t}=\frac{\exp(q_t)}{1+\exp(q_t)},\ \text{donde}\quad \underset{(0.050)\ (0.090)}{q_t=-2.024+3.983\rho_{t-1}}+\underset{(0.019)}{0.088}\frac{a_{1,t-1}a_{2,t-1}}{\sqrt{\sigma_{11,t-1}\sigma_{22,t-1}}}$$

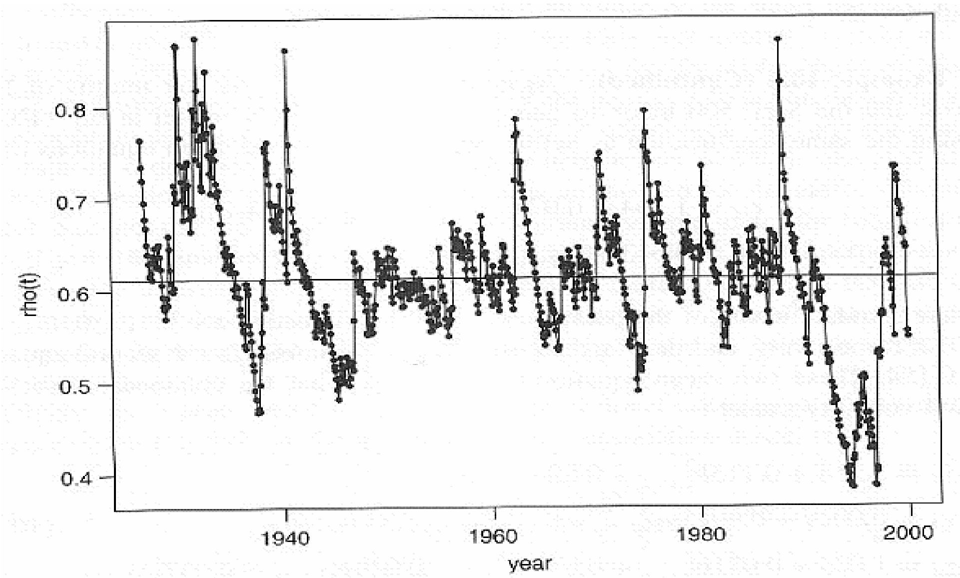
✓ Para las series de residuales estandarizados, los estadísticos de Ljung-Box $Q_2(4)=20.57(0.11)$ y $Q_2(8)=36.08(0.21)$. Se concluye que no rechazamos el modelo de media.

✓ Para las series de residuales estandarizados al cuadrado, los estadísticos de Ljung-Box $Q_2^*(4)=16.69(0.27)$ y $Q_2^*(8)=36.71(0.19)$. Se concluye que no rechazamos el modelo de volatilidad.

✓ Es interesante comparar este modelo GARCH(1,1) de correlación cambiante con el modelo GARCH(1,1) de correlación constante estimado antes.

⇒ Las ecuaciones de la media y la volatilidad están cercanas.

⇒ El siguiente gráfico presenta el coeficiente de correlación condicional ajustado. Allí se muestra que el coeficiente de correlación fluctúa con el tiempo y se vuelve más pequeño en los años mas recientes, lo cual concuerda con de las correlaciones muestrales del gráfico anterior.



⇒ El promedio de la correlación ajustada es 0.612, el cual es prácticamente la misma correlación estimada de 0.614 para el modelo de correlación constante.

⇒ Usando las varianzas muestrales de r_{it} como los valores iniciales para las varianzas condicionales y las observaciones desde t=4 a t=888, la función de verosimilitud maximizada del modelo GARCH(1,1) de correlación constante es -3691, mientras que es de -3674 para el modelo GARCH(1,1) de correlación cambiante. Por tanto, hay una mejora significante para este último modelo.

⇒ Considere los pronósticos 1-paso adelante con origen h=888:

a) Para el modelo de correlación constante se tiene que $a_{1,888}$=3.075, $a_{2,888}$=4.931, $\sigma_{11,888}$=77.91, y $\sigma_{22,888}$=21.19. Por tanto, el pronóstico 1-paso adelante para la matriz de covarianza condicional es

$$\hat{\Sigma}_{888}(1) = \begin{bmatrix} 71.09 & 21.83 \\ 21.83 & 17.19 \end{bmatrix}$$

donde la covarianza se obtiene usando el coeficiente de correlación constante de 0.614.

b) Para el modelo de correlación cambiante, se tiene que $a_{1,888}$=3.287, $a_{2,888}$=4.950, $\sigma_{11,888}$=83.35, y $\sigma_{22,888}$=28.56. Por tanto, el pronóstico 1-paso adelante para la matriz de covarianza condicional es

$$\hat{\Sigma}_{888}(1) = \begin{bmatrix} 75.15 & 23.48 \\ 23.48 & 24.70 \end{bmatrix}$$

donde el pronóstico del coeficiente de correlación es 0.545.

- ✓ En el segundo método, para modelar las correlaciones tiempo variantes se usa la **descomposición de Cholesky de** Σ_t. Para el caso bivariado el vector $\Xi_t = (g_{11,t}, g_{22,t}, q_{21,t})'$. Un modelo GARCH(1,1) para a_t es

$$g_{11,t} = \alpha_{10} + \alpha_{11} b_{1,t-1}^2 + \beta_{11} g_{11,t-1}$$

$$q_{21,t} = \gamma_0 + \gamma_1 q_{21,t-1} + \gamma_2 a_{2,t-1}$$

$$g_{22,t} = \alpha_{20} + \alpha_{21} b_{1,t-1}^2 + \alpha_{22} b_{2,t-1}^2 + \beta_{21} g_{11,t-1} + \beta_{22} g_{22,t-1}$$

donde la transformación ortogonal es $b_{1t} = a_{1t}$ y $b_{2t} = a_{2t} - q_{21,t} a_{1t}$.

En este caso:

$\Rightarrow$ b_{1t} asume un modelo GARCH(1,1) univariado.

$\Rightarrow$ b_{2t} usa un modelo GARCH(1,1) bivariado.

$\Rightarrow$ $q_{21,t}$ está autocorrelacionada y usa a $a_{2,t-1}$ como variable explicativa.

$\Rightarrow$ Bajo normalidad, la fdp relevante a la estimación máximo verosímil es

$$l(a_t, \Sigma_t) = l(b_t, \Xi_t) = -\frac{1}{2}\sum_{i=1}^{k}\left(\ln(g_{ii,t}) + \frac{b_{it}^2}{g_{ii,t}}\right)$$

vista anteriormente.

Ejemplo

Consideremos de nuevo los retornos mensuales de las acciones de IBM y del índice S&P 500. Usando la misma especificación anterior, se obtiene el siguiente modelo:

- ✓ Ecuaciones para la media:

$$\underset{}{r_{1t}} = \underset{(0.219)}{1.364} + \underset{(0.027)}{0.075}r_{1,t-1} - \underset{(0.032)}{0.058}r_{2,t-2} + a_{1t}$$

$$r_{2t} = \underset{(0.154)}{0.643} + a_{2t}$$

✓ Modelo de volatilidad ajustado:

$$\underset{}{g_{11,t}} = \underset{(1.033)}{3.714} + \underset{(0.022)}{0.113 b_{1,t-1}^2} + \underset{(0.037)}{0.804 g_{11,t-1}}$$

$$q_{21,t} = \underset{(0.001)}{0.0029} + \underset{(0.002)}{0.9915 q_{21,t-1}} - \underset{(0.0004)}{0.0041 a_{2,t-1}}$$

$$g_{22,t} = \underset{(0.344)}{1.023} + \underset{(0.007)}{0.021 b_{1,t-1}^2} + \underset{(0.013)}{0.052 b_{2,t-1}^2} - \underset{(0.013)}{0.040 g_{11,t-1}} + \underset{(0.015)}{0.937 g_{22,t-1}}$$

donde $b_{1t} = a_{1t}$ y $b_{2t} = a_{2t} - q_{21,t} a_{1t}$. Todas las estimaciones son estadísticamente significativas al 1%. El modelo muestra una relación dinámica fuerte en la correlación (observe el coeficiente 0.991 en la segunda ecuación anterior).

✓ La matriz de covarianza condicional Σ_t, se puede obtener usando la descomposición de Cholesky. Para el caso bivariado, la relación vista antes, está dada por

$$\sigma_{11,t} = g_{11,t}, \quad \sigma_{21,t} = q_{21,t} g_{11,t} \quad \text{y} \quad \sigma_{22,t} = g_{22,t} + q_{21,t}^2 g_{11,t}$$

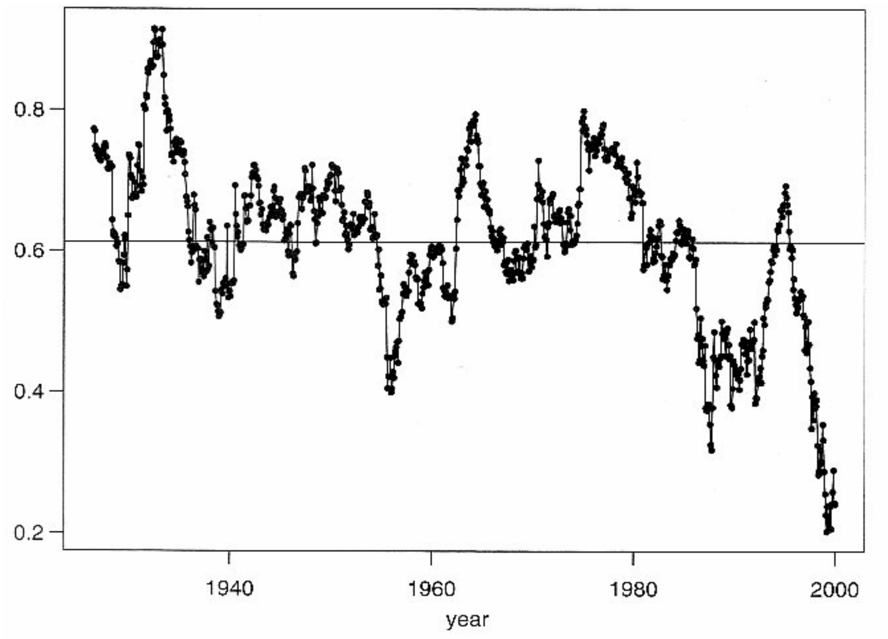
Por tanto, el coeficiente de correlación tiempo variante es

$$\rho_{12,t} = \frac{\sigma_{12,t}}{\sqrt{\sigma_{11,t}\sigma_{22,t}}} = \frac{q_{21,t}\sqrt{g_{11,t}}}{\sqrt{g_{22,t} + q_{21,t}^2 g_{11,t}}}$$

✓ Para las series de residuales estandarizados, los estadísticos de Ljung-Box $Q_2(4)$=19.77(0.14) y $Q_2(8)$=34.22(0.27). Se concluye que no rechazamos el modelo de media.

✓ Para las series de residuales estandarizados al cuadrado, los estadísticos de Ljung-Box $Q_2^*(4)$=15.34(0.36) y $Q_2^*(8)$=31.87(0.37). Se concluye que no rechazamos el modelo de volatilidad.

✓ El siguiente gráfico muestra el coeficiente de correlación tiempo variante estimado. Muestra un patrón más suave que el del gráfico del modelo anterior, y confirma su tendencia decreciente. En particular, las estimaciones del coeficiente son mas pequeñas en los últimos años que las presentadas por los otros modelos.



✓ La función de verosimilitud maximizada para este modelo es -3672, que es muy similar a la del modelo anterior. Sin embargo, la aproximación de la descomposición de Cholesky tiene varias ventajes:

⇒ Primero, no requiere restricciones paramétricas en la estimación para asegurar que Σ_t sea definida positiva. Si además se usa las transformaciones logarítmicas de $g_{ii,t}$, entonces no se necesita ninguna restricción para el modelo completo de la volatilidad.

⇒ La función de log verosimilitud es más simple.

⇒ Los parámetros tiempo cambiantes $q_{ij,t}$ y $g_{ii,t}$ tiene bonitas interpretaciones.

✓ Sin embargo, la transformación puede complicar un poco la inferencia, debido a que el modelo ajustado puede depender del ordenamiento de los elementos del vector a_t. En teoría, el orden de los elementos del vector a_t no debería tener ningún impacto sobre la volatilidad.

✓ El pronóstico 1-paso adelante con origen h=888 para la matriz de covarianza condicional es

$$\hat{\Sigma}_{888}(1) = \begin{bmatrix} 73.45 & 7.34 \\ 7.34 & 17.87 \end{bmatrix}$$

donde el pronóstico del coeficiente de correlación es 0.203, que es sustancialmente mas pequeño que los obtenidos por los modelos anteriores. Sin embargo, los pronósticos de las varianzas condicionales son parecidos a los anteriores.

Algunos Desarrollos Recientes

- Usando la reparametrización vista antes, donde $\Sigma_t \equiv [\sigma_{ij,t}] = D_t \rho_t D_t$, distintos autores han propuesto modelos parsimoniosos para ρ_t que describan las correlaciones tiempo variante.

- ✓ Para retornos k-dimensionales, Tse y Tsui (2002) suponen que la matriz de correlación condicional ρ_t sigue el modelo

$$\rho_t = (1-\theta_1-\theta_2)\rho + \theta_1\rho_{t-1} + \theta_2\psi_{t-1}$$

donde θ_1 y θ_2 son parámetros escalares, ρ es una matriz definida positiva con unos en su diagonal y ψ_{t-1} es la matriz de correlación muestral de k x k, usando los shocks desde t-m, …, t-1 para un m pre-especificado.

⇒ La estimación de θ_1 y θ_2 requiere de restricciones especiales para asegurar que la matriz de correlación sea definida positiva.

⇒ El modelo es parsimonioso, pero es difícil de implementar en aplicaciones reales, pues requiere una adecuada elección de ρ y m.

- ✓ Engle (2002) propone el modelo

$$\rho_t = J_t Q_t J_t$$

donde $Q_t = (q_{ij})_{kxk}$ es una matriz definida positiva, $J_t = diag(q_{11,t}^{-1/2},\cdots,q_{kk,t}^{-1/2})$ y Q_t satisface

$$Q_t = (1-\theta_1-\theta_2)\overline{Q} + \theta_1\varepsilon_{t-1}\varepsilon'_{t-1} + \theta_2 Q_{t-1}$$

donde ε_t es el vector de innovación estandarizado cuyos elementos son $\varepsilon_{it} = a_{it}/\sqrt{\sigma_{ii,t}}$, $\overline{Q}$ es la matriz de covarianza no condicional de ε_t y θ_1 y

θ_2 son parámetros escalares no negativos que satisfacen que $0 < \theta_1 + \theta_2 < 1$. La matriz J_t es una matriz de normalización que garantiza que R_t es una matriz de correlación.

- Una falla de los dos modelos anteriores es que θ_1 y θ_2 son parámetros escalares de modo tal que todas las correlaciones condicionales tienen la misma dinámica. Esto no es fácil de justificar, sobre todo cuando k es grande.

Modelos de Volatilidad de Dimensiones Altas

- La descomposición de Cholesky puede ser empleada para formular una estrategia para construir un modelo de volatilidad de altas dimensiones.

- Sea $r_t = \mu_t + a_t$. La media μ_t puede ser especificada usando los modelos VARMA vistos anteriormente. Con frecuencia, un modelo VAR simple es suficiente.

- La siguiente es la estrategia para la construcción del modelo multivariado para a_t:

  1. Seleccione un índice de mercado o el un retorno de alguna acción que sea el interés principal. Construya un modelo para dicha serie.

  2. Añada una segunda serie de retornos al sistema, realice la transformación ortogonal sobre el proceso del shock de esta nueva serie, y construya un modelo bivariado para el sistema. Los parámetros estimados de la serie univariada del paso 1 pueden ser usados como valores iniciales en la estimación bivariada.

3. Agregue una tercera serie al sistema, realice la transformación ortogonal sobre el proceso de shocks recientemente agregado, y construya un modelo de volatilidad tridimensional. De nuevo, los parámetros obtenidos en la estimación del modelo bivariado pueden ser usados como valores iniciales en la estimación del modelo tridimensional.

4. Continúe aumentado series hasta obtener el modelo de volatilidad conjunta para todas las series de interés.

Finalmente,

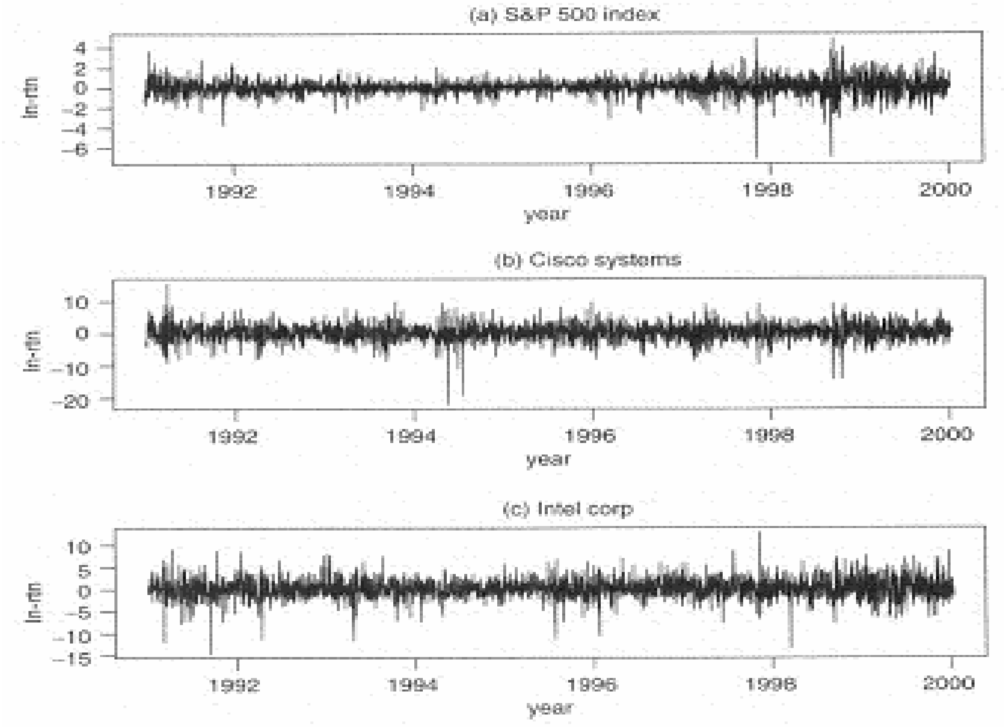
⇒ En cada paso se debe chequear la adecuación del modelo ajustado.

⇒ La experiencia ha mostrado que este procedimiento secuencial puede simplificar sustancialmente la complejidad relacionada con la construcción de un modelo de altas dimensiones para la volatilidad.

⇒ En particular, el procedimiento secuencial puede reducir sustancialmente el tiempo de computación.

Ejemplo

Se aplicará el procedimiento secuencial propuesto a la construcción de un modelo de volatilidad para los retornos diarios del índice S&P 500 y de las acciones de Cisco Systems e Intel Corporation. Se tienen 2275 observaciones, de Enero 2 de 1991 a Diciembre 31 de 1999. Los retornos están en porcentajes y se presentan en los siguientes gráficos.



Las medias muestrales, errores estándar y la matriz de correlación están dados por

$$\widehat{\mu} = \begin{bmatrix} 0.066 \\ 0.257 \\ 0.156 \end{bmatrix}, \quad \begin{bmatrix} \hat{\sigma}_1 \\ \hat{\sigma}_2 \\ \hat{\sigma}_3 \end{bmatrix} = \begin{bmatrix} 0.875 \\ 2.853 \\ 2.464 \end{bmatrix}, \quad \widehat{\rho} = \begin{bmatrix} 1.00 & 0.52 & 0.50 \\ 0.52 & 1.00 & 0.47 \\ 0.50 & 0.47 & 1.00 \end{bmatrix}.$$

- ✓ Usando el estadístico multivariado de Ljung-Box para detectar dependencia serial en el vector de series, obtenemos $Q_3(1)$=26.20, $Q_3(4)$=79.73, $Q_3(8)$=123.68. Los correspondientes p-valores están cercanos a ceros al comparar estos valores con los de una chi-cuadrado con 9, 36 y 72 grados de libertad. Por tanto hay indicios de correlación serial en los datos.

- ✓ La siguiente tabla presenta los primeros seis rezagos de la matriz de correlación cruzada muestral.

1	2	3	4	5	6
• • • • • • – • •	• • • • – • • • •	– • • • • • • • •	• • • • • • • • •	– • • – • • – • •	• • • • – • • • •

La revisión de la tabla muestra que:

⇒ Los retornos diarios de S&P 500 no dependen de los retornos pasados de Cisco o Intel.

⇒ Los retornos de las acciones de Cisco tienen alguna correlación y dependen del pasado de los retornos del índice S&P 500 (rezagos 2 y 5).

⇒ Los retornos de las acciones de Intel dependen de los retornos pasados de S&P 500 (rezagos 1 y 5).

✓ Estas dependencias son similares a las de IBM y S&P 500 vistas anteriormente y sugieren que los retornos de las acciones de grandes compañías tienden a ser afectadas por el comportamiento pasado del mercado. Sin embargo, los retornos del mercado no están afectadas significativamente por los pasados retornos de las compañías individuales.

✓ Construcción del modelo de volatilidad.

1. Comenzamos con los retornos del índice S&P 500, para el cual obtenemos el siguiente modelo

$$r_{1t} = -0.076 + 0.042 r_{1,t-1} - 0.062 r_{1,t-3} - 0.048 r_{1,t-4} - 0.052 r_{1,t-5} + a_{1t}$$

$$\sigma_{11,t} = 0.013 + 0.092 a_{1,t-1}^2 + 0.894 \sigma_{11,t-1}$$

donde:

⇒ Los errores estándar de los parámetros en la ecuación para la media son, respectivamente, 0.016, 0.023, 0.020, 0.022 y 0.02.

⇒ Los errores estándar de los parámetros en la ecuación de la volatilidad son 0.002, 0.006 y 0.007, respectivamente.

⇒ Los estadísticos de Ljung-Box de la serie de residuales estandarizados y sus cuadrados no detectan correlación serial o heterocedasticidad condicional. Se obtiene Q(10)=7.389(0.69) para los residuales estandarizados y Q(10)=3.14(0.98) para los residuales estandarizados al cuadrado.

2. Aumentado el sistema con los retornos de las acciones Cisco Systems, construimos el siguiente modelo bivariado

$$\begin{aligned} r_{1t} &= 0.065 - 0.04 r_{1,t-3} + a_{1t} \\ r_{2t} &= 0.325 + 0.195 r_{1,t-2} - 0.09 r_{2,t-2} + a_{2t} \end{aligned}$$

donde todas los estimaciones son estadísticamente significativas al 1%.
Usando la descomposición de Cholesky, obtenemos la transformación ortogonal b_t de a_t como $b_{1t} = a_{1t}$, $b_{2t} = a_{2t} - q_{21,t} b_{1t}$

Las siguientes son las ecuaciones de volatilidad:

$$\begin{aligned} g_{11,t} &= \underset{(0.001)}{0.006} + \underset{(0.005)}{0.051} b_{1,t-1}^2 + \underset{(0.006)}{0.943} g_{11,t-1} \\ q_{21,t} &= \underset{(0.156)}{0.331} + \underset{(0.099)}{0.790} q_{21,t-1} - \underset{(0.011)}{0.0411} a_{2,t-1} \\ g_{22,t} &= \underset{(0.029)}{0.177} + \underset{(0.008)}{0.082} b_{2,t-1}^2 + \underset{(0.011)}{0.890} g_{22,t-1} \end{aligned}$$

⇒ Los estadísticos bivariados de Ljung-Box de los residuales estandarizados y sus cuadrados no detectan correlación serial o heterocedasticidad condicional. Por tanto el modelo bivariado es adecuado. Se observa que la diferencia entre el modelo marginal y univariado para la volatilidad de r_{1t} es pequeña.

3. El próximo y último paso es agregar la serie de los retornos de las acciones de Intel. Las ecuaciones para la media son

$$r_{1t} = 0.065 - 0.043 r_{1,t-3} + a_{1t}$$

$$r_{2t} = 0.326 + 0.201 r_{1,t-2} - 0.089 r_{2,t-1} + a_{2t}$$

$$r_{3t} = 0.192 - 0.264 r_{1,t-1} + 0.059 r_{3,t-1} + a_{3t}$$

donde:

⇒ Los errores estándar de los parámetros en la primera ecuación son 0.016, 0.017, respectivamente.

⇒ Los errores estándar de los parámetros en la segunda ecuación son 0.052, 0.059 y 0.021, respectivamente.

⇒ Los errores estándar de los parámetros en la tercera ecuación son 0.050, 0.057 y 0.022, respectivamente.

⇒ Todas las estimaciones son estadísticamente significativas al 1%.

⇒ Las estimaciones para la media de r_{1t} y r_{12} son esencialmente las obtenidas en el modelo bivariado.

✓ Usando la descomposición de Cholesky, la trasformación ortogonal es $b_{1t} = a_{1t}$, $b_{2t} = a_{2t} - q_{21} b_{1t}$, $b_{3t} = a_{3t} - q_{31,t} b_{1t} - q_{32,t} b_{2t}$.

✓ El modelo tridimensional para la volatilidad es un poco más complicado, pero aún manejable:

$$
\begin{aligned}
g_{11,t} &= 0.006 + 0.050b_{1,t-1}^2 + 0.943g_{11,t-1},\\
q_{21,t} &= 0.277 + 0.824q_{21,t-1} - 0.035a_{2,t-1},\\
g_{22,t} &= 0.178 + 0.082b_{2,t-1}^2 + 0.889g_{22,t-1},\\
q_{31,t} &= 0.039 + 0.973q_{31,t-1} + 0.010a_{3,t-1},\\
q_{32,t} &= 0.006 + 0.981q_{32,t-1} + 0.004a_{2,t-1},\\
g_{33,t} &= 1.188 + 0.053b_{3,t-1}^2 + 0.687g_{33,t-1} - 0.019g_{22,t-1},
\end{aligned}
$$

Los errores estándar están dados en la siguiente tabla, siguiendo el mismo orden de aparición en las ecuaciones anteriores.

Ecuación	Error estándar	Ecuación	Error estándar
$g_{11,t}$	0.001 0.005 0.006	$q_{21,t}$	0.135 0.086 0.010
$g_{22,t}$	0.029 0.009 0.011	$q_{31,t}$	0.017 0.012 0.004
$g_{33,t}$	0.407 0.015 0.100 0.008	$q_{32,t}$	0.004 0.013 0.001

Con excepción del término constante de la ecuación para $q_{32,t}$, todas las estimaciones son significativas al 1%.

Sea $\hat{a}_t^s = (a_{1t}/\sigma_{1t},\ a_{2t}/\sigma_{2t},\ a_{3t}/\sigma_{3t})'$ el vector de residuales estandarizados, donde $\sigma_{it} = \sqrt{\sigma_{ii,t}}$, es el error estándar condicional ajustado para el i-ésimo retorno. Entonces:

⇒ Los estadísticos de Ljung-Box para a_t^s son $Q_3(4)$=34.48(0.31) y $Q_3(8)$=60.42(0.70), donde los grados de libertad de las chi-cuadrados son 31 y 67 respectivamente, después de ajustar por el número de parámetros en las ecuaciones de la media.

⇒ Los estadísticos de Ljung-Box para $(a_t^s)^2$ son $Q_3(4)=28.71(0.58)$ y $Q_3(8)=52.00(0.91)$.

⇒ Por tanto el modelo parece ser adecuado para modelar las medias y las volatilidades condicionales.

⇒ El modelo de volatilidad tridimensional presenta algunas características interesantes:

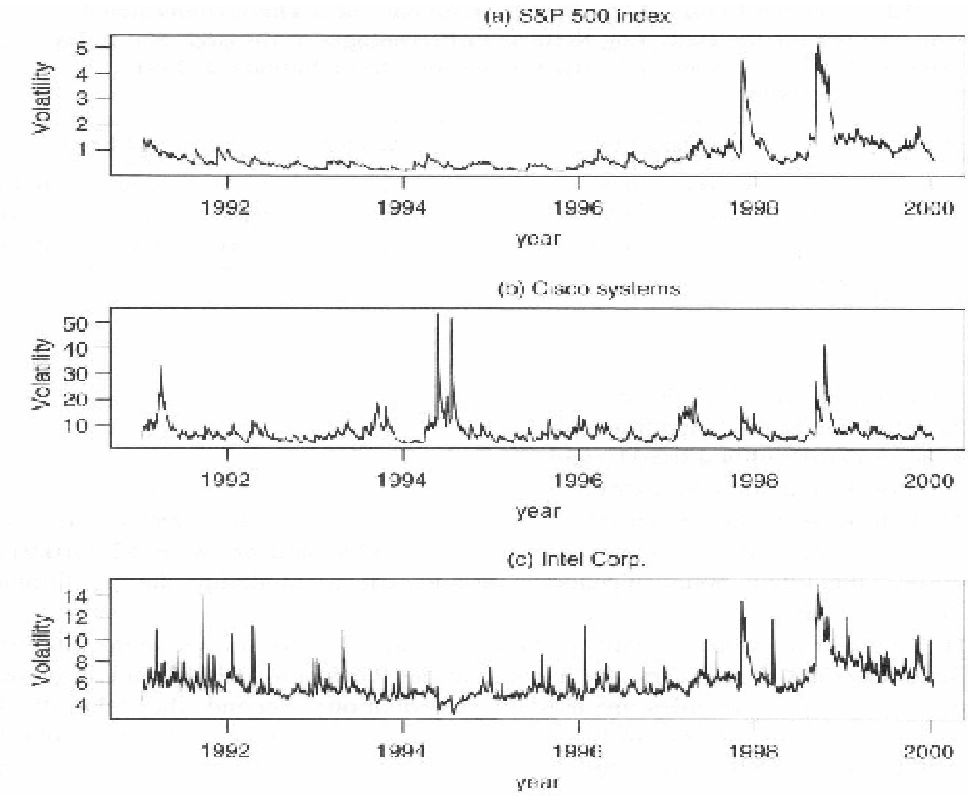
- Es esencialmente un modelo GARCH(1,1) de correlación tiempo variante, debido a que solamente rezagos de orden 1 aparecen en las ecuaciones.

- La volatilidad de los retornos diarios del índice S&P 500 no dependen de la volatilidades pasadas de los retornos de Cisco o Intel.

- Usando la transformación inversa de la descomposición de Cholesky, las volatilidades de los retornos diarios de las acciones de Cisco e Intel dependen de la volatilidad pasada del retorno del mercado (Vea la relación entre Σ_t y L_t y G_t dada anteriormente).

- Las cantidades de correlación $q_{ij,t}$ tienen alta persistencia con grandes coeficientes AR(1).

⇒ Los siguientes gráficos muestran los procesos de volatilidades ajustadas del modelo (es decir, los $\sigma_{ii,t}$)

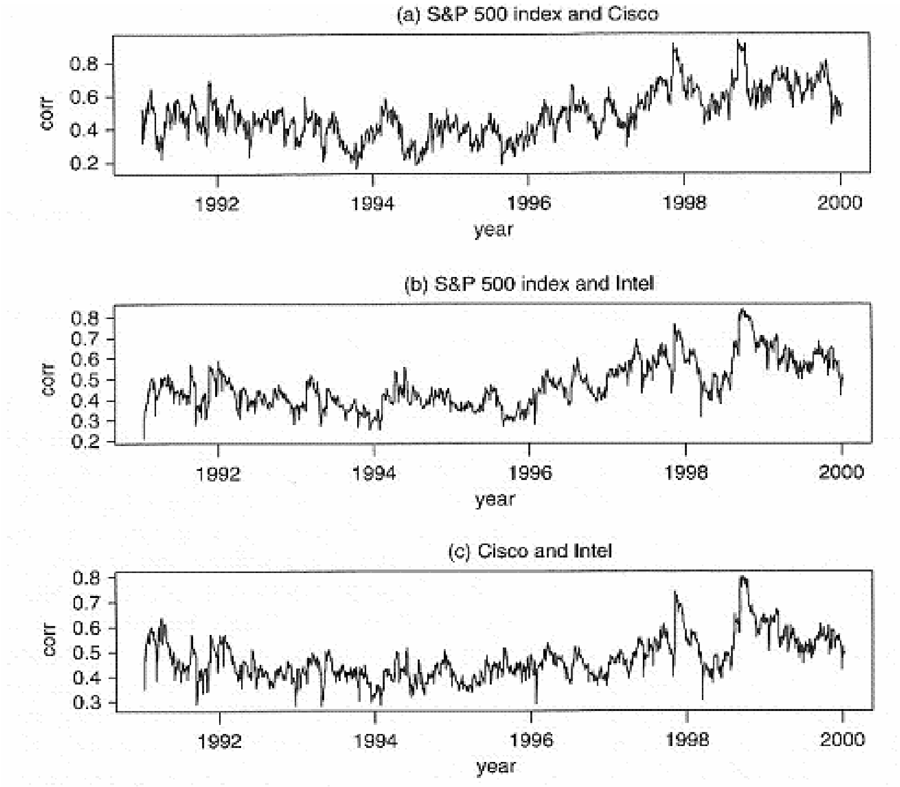


⇒ La volatilidad del retorno del índice S&P 500 es mucho menor que las de los retornos de las acciones.

⇒ Las volatilidades de los retornos del índice S&P 500 y de las acciones de Intel se han incrementado en los últimos años. No así para Cisco.

Los siguientes gráficos muestran los coeficientes de correlación condicionales entre las tres series de retornos. Es de particular interés la comparación entre las gráficas de volatilidades anteriores y las de correlación.

⇒ Ellas muestran que el coeficiente de correlación entre dos series se incrementa cuando los retornos son volátiles. Esto está de acuerdo con el estudio empírico de las relaciones entre los índices de bolsa internacionales para los cuales la correlación entre dos mercados se tiende a incrementar durante las crisis financieras.



El modelo de volatilidad estimado contiene dos conjuntos de ecuaciones:

⇒ El primer conjunto describe la evolución de las varianzas condicionales (es decir, las $g_{ii,t}$).

⇒ El segundo conjunto tiene que ver con los coeficientes de correlación condicionales (es decir, con las $q_{ij,t}$, con i>j). Para este conjunto, un modelo AR(1) podría ser suficiente.

⇒ Similarmente, un simple modelo AR también podría ser suficiente para las varianzas condicionales. Defina $v_t = (v_{11,t}, v_{22,t}, v_{33,t})'$, donde $v_{ii,t} = \ln(g_{ii,t})$, y $q_t = (q_{21,t}, q_{31,t}, q_{32,t})'$, lo anterior sugiere que podemos usar un modelo de rezagos 1, de la forma

$$v_t = c_1 + \beta_1 v_{t-1}, \quad q_t = c_2 + \beta_2 q_{t-1}$$

como las funciones exactas para modelar la volatilidad de los activos financieros, donde los c_i son vectores de constantes y β_i son matrices de 3 x 3 de valores reales.

⇒ Si se agrega un término de ruido a las ecuaciones anteriores, el modelo puede ser escrito como

$$v_t = c_1 + \beta_1 v_{t-1} + e_{1t}, \quad q_t = c_2 + \beta_2 q_{t-1} + e_{2t}$$

Donde e_{it} son shocks aleatorios con media cero y matriz de covarianza definida positiva. Esto define un modelo multivariado simple de volatilidad estocástica.

Modelos de Factores para Volatilidad

- Otra aproximación que permite simplificar la estructura dinámica de un proceso de volatilidad multivariada es usar modelos de factores.

- Debido a que los modelos de volatilidad multivariada están interesados en describir la evolución en el tiempo de la matriz de covarianza condicional de a_t, donde $a_t = r_t - \mu_t$, una manera simple de identificar los "factores comunes" en volatilidad, es realizar un análisis de componentes principales sobre a_t.

- La construcción del modelo de volatilidad multivariado consta de tres pasos:

  ✓ Seleccione unas pocas componentes principales que explique un alto porcentaje de la variabilidad en a_t.

- ✓ Construya un modelo de volatilidad para cada una de las componentes principales seleccionadas.

- ✓ Relacione la volatilidad de las a_{it} con las volatilidades de las componentes principales seleccionadas.

El objetivo de este procedimiento es reducir la dimensión pero manteniendo una aproximación precisa de la volatilidad multivariada.

Ejemplo

Considere de nuevo los retornos mensuales, en porcentajes, de las acciones de IBM y del índice S&P 500. Usando el modelo bivariante AR(3) elegido anteriormente (Capítulo de series de tiempo multivariadas), obtenemos la serie de innovaciones a_t como los residuales de dicho modelo.

- ✓ Usando PCA sobre la matriz de covarianza de a_t, se obtiene valores los propios 63.373 y 13.489. El primer valor propio explica el 82.2% de la variación total de a_t. Por tanto podemos escoger la primera componente principal, $x_t = 0.797 r_{1t} + 0.604 r_{2t}$, como el factor común.

  ⇒ Alternativamente, como se vio antes, la dependencia serial de r_t es débil, y por tanto se puede realizar el PCA directamente sobre r_t. En ese caso, los valores propios de la matriz de covarianza de r_t son 63.625 y 13.513, que son esencialmente los mismos de matriz de covarianza de a_t. El factor común es $x_t = 0.796 r_{1t} + 0.605 r_{2t}$. Por tanto, el efecto de la media condicional sobre el PCA es despreciable.

- ✓ Por simplicidad usamos el análisis basado en los retornos r_t y $x_t = 0.796 r_{1t} + 0.605 r_{2t}$ es un factor común para las dos series de retornos mesuales.

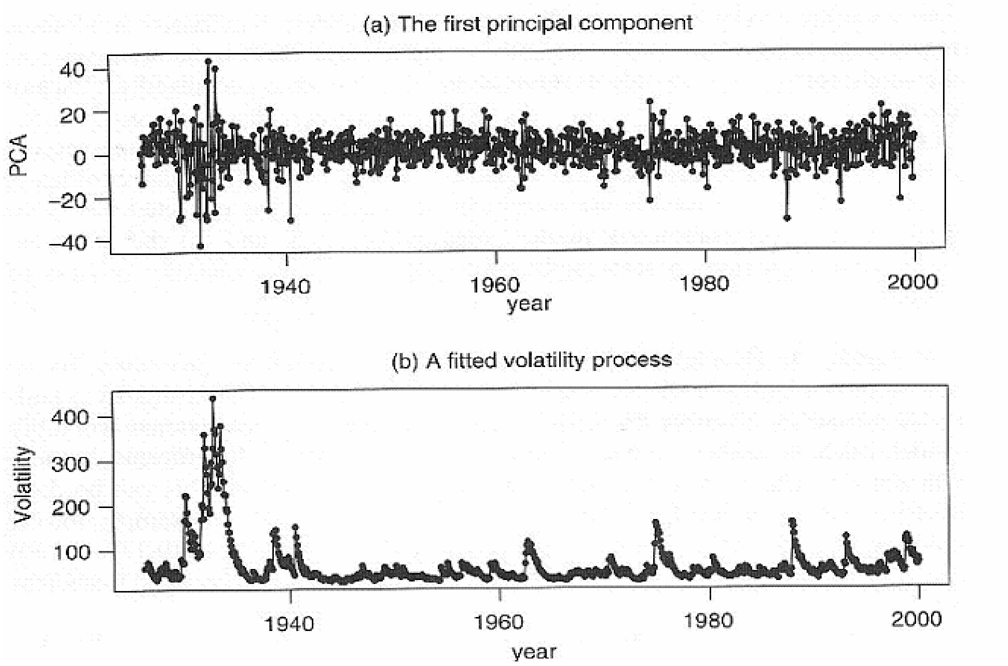
- ✓ Un modelo GARCH univariado para x_t proporciona

$$x_t = 1.317 + 0.096 x_{t-1} + a_t, \qquad a_t = \sigma_t \varepsilon_t$$
$$\sigma_t^2 = 3.834 + 0.110 a_{t-1}^2 + 0.825 \sigma_{t-1}^2 \qquad (1)$$

⇒ Todos sus parámetros son estadísticamente significativos a un nivel del 1%.

⇒ Los estadísticos de Ljung-Box para los residuales estandarizados y sus cuadrados no detectan que el modelo es inadecuado.

⇒ El siguiente gráfico muestra la evolución de la primera componente principal x_t y las volatilidades estimadas usando el modelo anterior.



✓ Usando σ_t^2 del modelo anterior, como un factor común de la volatilidad, se obtiene el siguiente modelo para los retornos mensuales originales.

Ecuación de media:

$$\underset{}{r_{1t}} = \underset{(0.211)}{1.140} + \underset{(0.030)}{0.079}r_{1,t-1} + \underset{(0.031)}{0.067}r_{1,t-2} - \underset{(0.043)}{0.122}r_{2,t-2} + a_{1t}$$

$$r_{2t} = \underset{(0.165)}{0.537} + a_{2t}$$

Ecuación de varianza condicional:

$$\begin{bmatrix} \sigma_{11,t} \\ \sigma_{22,t} \end{bmatrix} = \begin{bmatrix} 19.08 \\ (3.70) \\ -5.62 \\ (2.36) \end{bmatrix} + \begin{bmatrix} 0.098 & . \\ (0.044) & \\ . & . \end{bmatrix} \begin{bmatrix} a_{1,t-1}^2 \\ a_{2,t-1}^2 \end{bmatrix} + \begin{bmatrix} 0.833 \\ (0.076) \\ 0.596 \\ (0.050) \end{bmatrix} \sigma_t^2 \qquad (2)$$

Ecuación de correlación condicional:

$$\rho_t = \frac{\exp(q_t)}{1+\exp(q_t)}, \quad q_t = \underset{(0.025)}{-2.098} + \underset{(0.038)}{4.120}\rho_{t-1} + \underset{(0.015)}{0.078}\frac{a_{1,t-1}a_{2,t-1}}{\sqrt{\sigma_{11,t-1}\sigma_{22,t-1}}} \qquad (3)$$

⇒ Usando los residuales estandarizados, se obtiene $Q_2(4)=15.37(0.29)$ y $Q_2(8)=34.24(0.23)$. Por tanto no hay correlación serial en los residuales.

⇒ Para los cuadrados de los residuales estandarizados, $Q_2^*(4)=20.25(0.09)$ y $Q_2^*(8)=61.95(0.0004)$. Por tanto, el modelo de volatilidad anterior no es adecuado para manejar la heterocedasticidad condicional, sobre todo en rezagos grandes. Esto no es sorprendente, puesto que un solo factor solamente explica el 82.2% de la variación total de los datos.

✓ Comparando el modelo para la volatilidad y la correlación anterior con el obtenido usando el modelo de correlación cambiante, donde

Las ecuaciones para la media son:

$$r_{1t} = \underset{(0.215)}{1.318} + \underset{(0.026)}{0.076}r_{1,t-1} - \underset{(0.034)}{0.068}r_{2,t-2} + a_{1t}$$

$$r_{2t} = \underset{(0.151)}{0.673} + a_{2t}$$

Las ecuaciones de volatilidad son:

$$\begin{bmatrix} \sigma_{11,t} \\ \sigma_{22,t} \end{bmatrix} = \begin{bmatrix} 2.80 \\ (0.58) \\ 1.71 \\ (0.49) \end{bmatrix} + \begin{bmatrix} 0.084 & . \\ (0.013) & \\ 1.037 & 0.054 \\ (0.009) & (0.010) \end{bmatrix} \begin{bmatrix} a_{1,t-1}^2 \\ a_{2,t-1}^2 \end{bmatrix} + \begin{bmatrix} 0.864 & -0.020 \\ (0.021) & (0.009) \\ -0.058 & 0.914 \\ (0.014) & (0.013) \end{bmatrix} \begin{bmatrix} \sigma_{11,t-1} \\ \sigma_{22,t-1} \end{bmatrix}$$

La ecuación de correlación condicional es;

$$\rho_{21,t} = \frac{\exp(q_t)}{1+\exp(q_t)}, \qquad \underset{(0.050)}{q_t = -2.024} + \underset{(0.090)}{3.983\rho_{t-1}} + \underset{(0.019)}{0.088} \frac{a_{1,t-1}a_{2,t-1}}{\sqrt{\sigma_{11,t-1}\sigma_{22,t-1}}}$$

Se obtiene que:

⇒ Las ecuaciones de correlación de los dos modelos son prácticamente las mismas.

⇒ Como era de esperarse. El modelo de factor usa menos parámetros en la ecuación de volatilidad.

⇒ El modelo de factor común proporciona una aproximación razonable al proceso de volatilidad de los datos.

Nota: En el ejemplo anterior, se usó un procedimiento de estimación en dos etapas. En la primera etapa se construyó un modelo de volatilidad para el factor común. En la segunda etapa, dicha la volatilidad estimada se trata como dada, para estimar el modelo de volatilidad multivariada.

✓ Este procedimiento de estimación es simple, pero puede no ser eficiente.

✓ Un procedimiento de estimación más eficiente es realizar una estimación conjunta. Esta se puede realizar de una manera relativamente fácil si los factores comunes son conocidos. Por ejemplo, para los retornos mensuales del ejemplo anterior, una estimación conjunta de las ecuaciones (1), (2) y (3) se puede realizar si el factor común $x_t = 0.796 r_{1t} + 0.605 r_{2t}$ se considera como dado.

Lecturas adicionales:

Bauwens, L., Laurent, S., and Rombouts, J.V.K (2006) "Multivariate GARCH Models: A Survey", *Journal of Applied Econometrics*, 21, 79-109.

Alexander, C.O. (2000) "Orthogonal Methods for Generating Large Positive Semi-definite Covariance Matrices", ISMA Centre, University of Reading.

Alexander, C.O. and Chibumba, A.M. (2000) "Multivariate Orthogonal Factor GARCH", Mimeo, University of Sussex.