

Proyecto de Ingeniería Financiera.

Serie Temporal Financiera: Log>Returns de Gas Natural

Máster en Técnicas Estadísticas. Curso 2010-2011

García Veiga, Mariam. Leyenda Rodríguez, María

Índice

1. Introducción	2
2. Análisis descriptivo	4
3. Estudio de la estructura principal de los modelos de volatilidad	6
3.1. Media condicional	6
3.2. Varianza condicional	7
4. Ajuste de diferentes modelos para la serie de log-returns de Gas Natural	9
4.1. Modelo ARCH	10
4.1.1. Estimación de un modelo ARCH	11
4.2. Modelo GARCH	13
4.2.1. Estimación de un modelo GARCH	13
4.3. Modelo IGARCH	15
4.3.1. Estimación de un modelo IGARCH	15
4.4. Modelo EGARCH	16
4.4.1. Estimación de un modelo EGARCH	17
4.5. Modelo TGARCH	18
4.5.1. Estimación de un modelo TGARCH	18
5. Conclusiones	19
6. Código implementado	20

1. Introducción

En este documento, se analizará la serie temporal que contiene los log-returns correspondientes a los datos de cierre del precio de las acciones de Gas Natural desde el 3 de Enero de 2000 hasta el 20 de Diciembre de 2010 (Figura 1). Observemos que las series de tiempo de los rendimientos (1) son estacionarias en media (Figura 1) y no tienen unidad de medida puesto que reflejan las tasas de cambio del precio en k-períodos.

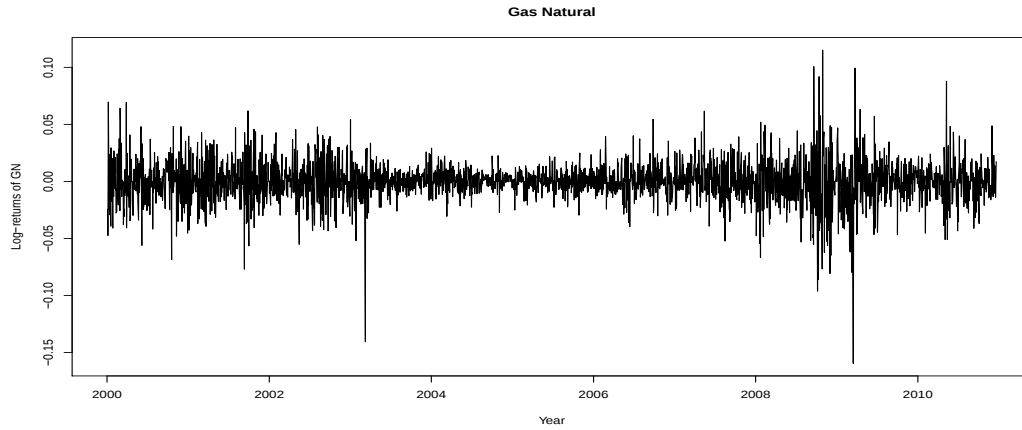


Figura 1: Serie temporal de los log-returns

$$r_{t+k} = \log \frac{P_{t+k}}{P_t} = \log \left[\frac{P_{t+k}}{P_{t+k-1}} \times \dots \times \frac{P_{t+1}}{P_t} \right] \quad (1)$$

dónde P_t , denota el precio de un activo en el instante t

Además de observar que la serie es estacionaria en media, también se puede apreciar un período de cierta estabilidad en las tasas de cambio de los precios comprendido entre el primer trimestre de 2003 hasta mediados-finales de 2007; período en el que la media es más cercana a cero. Los cambios bruscos, registrados en el período 2000-02 y fundamentalmente en el año 2008 coinciden con períodos de crisis económicas que han acarreado altas tasas de incertidumbre y volatilidad. Por consiguiente, se observa que la volatilidad no es constante en el tiempo, pudiendo diferenciar claramente períodos con varianzas distintas ("volatility clustering").

Junto a ello aparecen también algunos valores extremos positivos y negativos, siendo estos últimos son más frecuentes. Destacan sobre todo, los valores registrados en el primer trimestre de 2003 y en el primer trimestre de 2009. Como consecuencia de ello, la distribución de los rendimientos presenta una asimetría negativa (hacia la izquierda) y con colas pesadas. Además, al representar un QQ-plot (Figura 2) de dichos rendimientos se comprueba que su distribución no es gaussiana.

El objetivo final de este trabajo será testar diferentes modelos condicionalmente heterocedásticos para la volatilidad de dicha serie financiera.

2. Análisis descriptivo

Un primer paso para conocer la distribución de los datos es representar un gráfico QQ-plot (Figura 2) de dichos rendimientos, el cual indica que la asunción de normalidad es cuestionable.

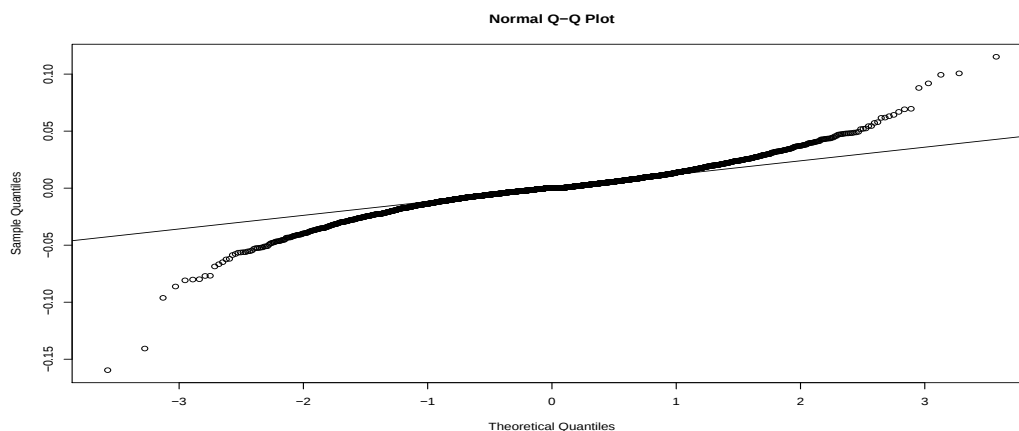


Figura 2: QQplot de los rendimientos

A la vista de los resultados que proporcionan los estadísticos descriptivos de la serie que estamos analizando (log-returns de Gas Natural)(Tabla 1) se comprueba que la media es prácticamente cero y con una desviación muy pequeña. Asimismo vemos que, tal y como se apuntó en el apartado anterior, existe una asimetría negativa en la distribución de los rendimientos y el coeficiente de curtosis es mucho mayor al de la normal.

El análisis descriptivo se complementa realizando la estimación kernel de la densidad de los datos (Figure 3) dónde observamos que esta densidad, como ya habíamos apuntado, presenta asimetría hacia la izquierda con colas pesadas. Además, se puede apreciar que la mayor masa de probabilidad recae cuando estamos muy próximos al cero; de hecho observamos una moda centrada en el cero y una moda local a la izquierda del cero. Asimismo, al compararlo con la densidad normal con media y desviación típica proporcionadas por los log-returns de los datos observamos que la estimación Kernel es más apuntada en el cero, es decir, menos suave que ésta.

	Gas Natural
n° observaciones	2854.000000
Mínimo	-0.159561
Mínimo	-0.159561
Máximo	0.115198
1° Quartile	-0.007985
3° Quartile	0.008138
Media	-0.000230
Mediana	0.000000
Suma	-0.656269
LCL Media	-0.000888
UCL Media	0.000428
Varianza	0.000321
Desviación típica	0.017922
Simetría	-0.376938
Curtosis	6.880669

Cuadro 1: Estadísticos descriptivos

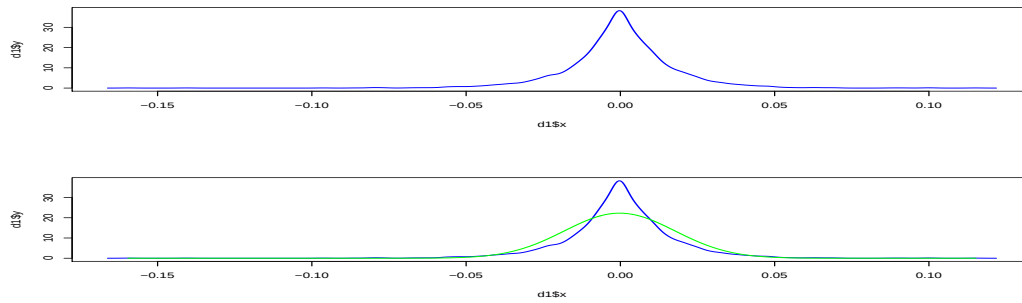


Figura 3: Estimación kernel de la densidad de los datos(azul) y la comparamos con la densidad normal de parámetros la media y la desviación típica de los datos

3. Estudio de la estructura principal de los modelos de volatilidad

A la hora de introducir la estructura principal de los modelos de volatilidad, consideramos la media y varianza condicionada de r_t dada F_{t-1} , la información obtenida hasta el instante $t-1$. Esto viene dado por $\mu_t = E(r_t|F_{t-1})$ y $\sigma_t^2 = Var(r_t|F_{t-1}) = E((r_t - \mu_t)^2|F_{t-1})$. Dónde el término σ_t^2 es la volatilidad del rendimiento en el instante t .

$$r_t = \mu_t + a_t \quad (2)$$

Notemos que a_t es usualmente denominado como shock o innovación de r_t y μ_t es conocido como la ecuación de la media para r_t .

3.1. Media condicional

En (Figura 4) se permite ver que las autocorrelaciones son muy pequeñas, prácticamente nulas, finitas y constantes. En definitiva, estamos ante una serie de rendimientos incorrelados, lo cual no quiere decir que sean independientes.

Dado que en toda la muestra los ACFs son prácticamente cero, podremos considerar que la serie es un proceso de ruido blanco. (Tsay, 2010-36).

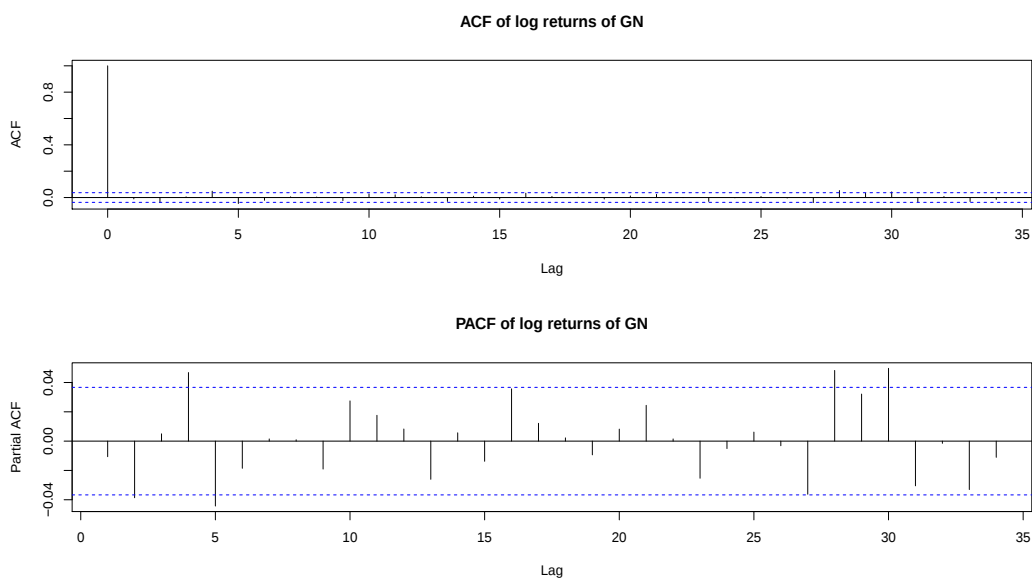


Figura 4: Función de autocorrelación y autocorrelación parcial de los rendimientos

Al realizar un contraste de Ljung-Box para contrastar la inexistencia de correlación en los rendimientos $H_0 : \rho = 0, H_1 : \rho \neq 0$ obtenemos que el estadístico de contraste toma un valor pequeño, $Q(3)=4.6588$, con un p-valor asociado de 0.1986; con lo que se confirma la hipótesis nula de que no hay correlación en los datos.

En consecuencia, los datos que estamos analizando no encierran una estructura ARMA para la media condicional de los log-returns.

3.2. Varianza condicional

Para estudiar la relación de dependencia, representamos la función de autocorrelación de los cuadrados de los rendimientos (Figura 5), comprobando que las correlaciones tardan mucho en desaparecer. Este hecho indica la existencia de dependencia entre ellas, pero una dependencia no lineal.

Con el fin de determinar el orden de un posible modelo ARCH, nos centramos en el estudio del gráfico de las autocorrelaciones parciales (Figura 5). En este gráfico vemos que algunos retardos resultan significativos, más concretamente hasta el parámetro 28; sin embargo, probaremos un ARCH(13) será suficiente. Pues a partir de este retardo la estructura es menor, es decir, se salen pocos retardos de las bandas de confianza.

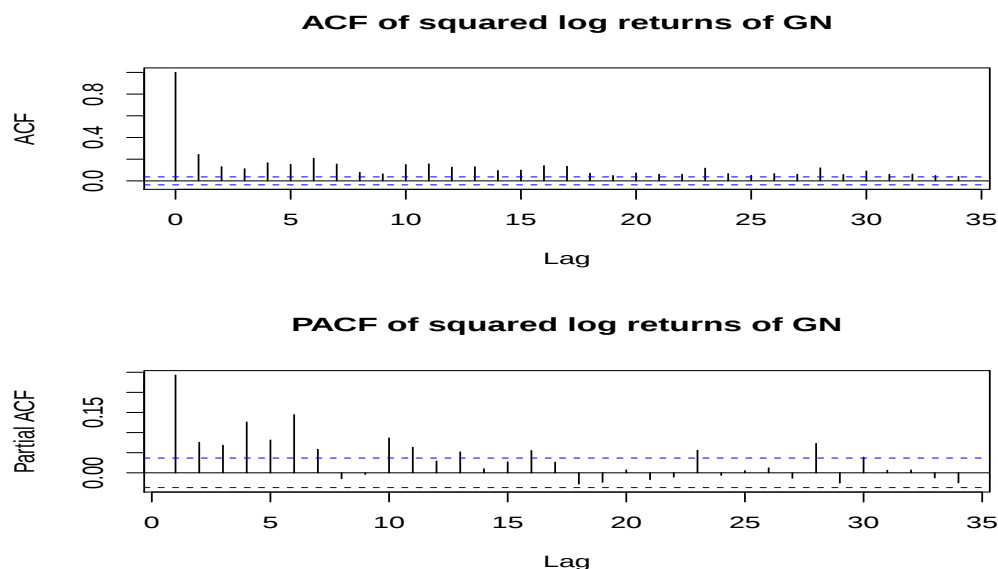


Figura 5: Función de autocorrelación y autocorrelación parcial de los rendimientos al cuadrado

A continuación, nos centraremos en el estudio de la ecuación de volatilidad,(3).

$$\sigma_t^2 = Var(r_t|F_{t-1}) = Var(a_t|F_{t-1}) \quad (3)$$

Para ello comenzaremos contrastando la presencia de heterocedasticidad condicional a través del test de Ljung-Box sobre los residuos al cuadrado. En primer lugar, enunciamos el procedimiento que vamos a utilizar

- La hipótesis nula es que los m primeros 'lags' de ACF de a_t^2 son cero. Concretamente $H_0 : \rho_1(a_t^2) = \dots = \rho_m(a_t^2) = 0$, donde $\rho_l(a_t^2)$ es el 'lag' de ACF de a_t^2 . El test de Ljung-Box sobre a_t^2 viene dado por:

$$Q(m) = T(T+2) \sum_{l=1}^m \frac{\rho_l(a_t^2)}{T-l} \quad (4)$$

Bajo H_0 , $Q(m)$ es asintóticamente una variable aleatoria chi-cuadrado con m grados de libertad. La regla decisión de rechazar H_0 es $Q(m) > \chi_\alpha^2$, donde χ_α^2 denota el percentil $(1 - \alpha)$ de una distribución chi-cuadrado con m grados de libertad.

El test de Ljung-Box arroja como resultado que el valor del estadístico de contraste de la serie a_t^2 muestra un fuerte efecto ARCH con $Q(13)=846.6681$ (4) y p-valor asociado es de $2,2 \times 10^{-16}$, por lo que el test resulta estadísticamente significativo, es decir, se comprueba que los residuos al cuadrado a_t^2 presentan heterocedasticidad condicional.

4. Ajuste de diferentes modelos para la serie de log-returns de Gas Natural

El objetivo de esta sección es presentar y estimar los modelos más importantes en la literatura para modelizar la volatilidad de los log-returns de nuestra serie Gas natural. Estos modelos son denominados **modelos condicionalmente heterocedásticos**. Recordemos que estos modelos se pueden clasificar en dos grandes categorías:

- aquéllos que utilizan una función exacta para determinar la evolución de la volatilidad
- aquéllos que utilizan una función estocástica para describir la volatilidad.

Los modelos univariantes de volatilidad que trataremos en este capítulo, pertenecientes al primer grupo, son los siguientes.

- Modelo autorregresivo condicionalmente heterocedástico (**ARCH**) de Engle (1982).
- Modelo generalizado autorregresivo condicionalmente heterocedástico (**GARCH**) de Bollerslev (1986).
- Modelo integrado generalizado autorregresivo condicionalmente heterocedástico (**IGARCH**) de Engle y Bollerslev (1986).
- Modelo GARCH exponencial (**EGARCH**) de Nelson (1991).
- Modelo GARCH umbral (threshold) (TGARCH) de Glosten, Jagannathan y Runkle (1993) y Zakoian (1994).

Notemos que la característica más importante de la volatilidad es que ésta no es directamente observable a través de los rendimientos (Figura 1), puesto que sólo hay una observación para cada instante t . Sin embargo, podemos observar las siguientes características relativas a la volatilidad:

1. Volatilidad clúster, es decir, los rendimientos presentan períodos donde la volatilidad es muy pequeña y otros donde ésta es muy grande.
2. Se considera que la volatilidad se desarrolla en el tiempo de manera continua, es decir, no suele variar en grandes saltos.
3. Volatilidad no diverge hasta el infinito, es decir, es finita.
4. Volatilidad parece reflejar diferencias cuando un precio se incrementa o cae (efecto leverage). El mercado reacciona de forma diferente si obtengo grandes pérdidas que si obtengo grandes ganancias.

4.1. Modelo ARCH

El modelo ARCH de Engle (1982) fue el primer modelo propuesto para modelizar la volatilidad. El modelo ARCH(m) supone que:

$$a_t = \sigma_t \epsilon_t \quad (5)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \cdots + \alpha_m a_{t-m}^2 \quad (6)$$

dónde

- ϵ_t es una secuencia de variables aleatorias independiente e idénticamente distribuidas con media cero y varianza uno.
- $\alpha_0 > 0$ y $\alpha_i \geq 0$ para $i > 0$. Los coeficientes α_i deben verificar alguna condición de regularidad para garantizar que $Var(a_t)$ sea finita.

Bajo la especificación ARCH, shocks en el pasado elevados implican una gran varianza condicional σ_t^2 para la innovación a_t , así que a_t^2 tiende a asumir un elevado valor. Como consecuencia, a un shock elevado le sigue un shock elevado. Esta característica es similar a la de la volatilidad clustering observada en los log-returns.

Para comprender mejor este modelo, veamos algunas propiedades del modelo ARCH(1) dónde $\alpha_0 > 0$ y $\alpha_1 \geq 0$:

- La media no condicionada de a_t es cero.
- La varianza no concionada de a_t viene dada por

$$Var(a_t) = \frac{\alpha_0}{1 - \alpha_1} \quad (7)$$

siempre que $0 \leq \alpha_1 < 1$.

- El exceso de curtosis de a_t es positivo y la cola de la distribución de a_t es más pesada que la de una distribución normal. Esto está de acuerdo con la presencia de extremos en la serie de tiempo de rendimientos.

Sin embargo, el modelo ARCH presenta alguna debilidad, (Tsay,2010-119):

- El modelo asume que positivos y negativos shocks tienen el mismo efecto en la volatilidad, porque depende de los shocks al cuadrado anteriores. En la práctica se sabe que el precio de un activo financiero responde de forma diferente a shocks positivos y negativos.
- El modelo aplica importantes restricciones en sus parámetros, como por ejemplo tener momentos finitos.
- El modelo no proporciona ninguna idea para comprender el origen de las variaciones de la serie de tiempo financiera.
- El modelo puede que proporcione demasiada frecuencia a la volatilidad, porque responde lentamente a los grandes shocks aislados.

4.1.1. Estimación de un modelo ARCH

Utilizando el gráfico de las autocorrelaciones parciales de los rendimientos al cuadrado (Figura 5) seleccionaremos el orden del modelo ARCH correspondiente. Así pues quedaría justificado un modelo ARCH(13) (6), ya que observando el gráfico PACF de los rendimientos al cuadrado (Figura 5) tenemos que hasta el retardo 13 es donde está recogida la mayor estructura del modelo; momento a partir del cual, la mayoría de los retardos son nulos.

$$r_t = \mu + a_t$$

$$a_t = \sigma_t + \epsilon_t$$

Con la implementación del modelo ARCH(13), se ha visto que el parámetro α_8 es no significativo, por lo que se ha eliminado del modelo. Como resultado, suponiendo una distribución Gaussiana para las innovaciones estandarizadas, se obtiene que

$$r_t = 0,000528 + a_t \quad (8)$$

$$\begin{aligned} \sigma_t = & 0,000049 + 0,270402a_{t-1}^2 + 0,103599a_{t-2}^2 + \\ & + 0,103800a_{t-3}^2 + 0,055983a_{t-4}^2 + 0,036068a_{t-5}^2 + \\ & + 0,050781a_{t-6}^2 + 0,046787a_{t-7}^2 + 0,042011a_{t-9}^2 + \\ & + 0,066185a_{t-10}^2 + 0,047949a_{t-11}^2 + 0,026539a_{t-12}^2 + 0,074122a_{t-13}^2 \end{aligned} \quad (9)$$

Los errores standard de los parámetros(9) son 0.000239, 0.000006, 0.030161, 0.023927, 0.024220, 0.021577, 0.017648, 0.018280, 0.020485, 0.019719, 0.019310, 0.017062, 0.014342 y 0.022223, respectivamente. Todos los parámetros son significativos con un nivel de significación del 5 %.

El estadístico de Ljung-Box de los residuos estandarizados proporciona $Q(13)=26.2413$ con un p-valor de 0.01578 y el estadístico de los residuos al cuadrado estandarizados es $Q(13)=854.6708$ con un p-valor de $2,2 \times 10^{-16}$. Por tanto, el modelo ARCH(13) sin el parámetro α_8 no es adecuado para describir la heterocedasticidad de los datos con un nivel de significación del 5 %.

4.2. Modelo GARCH

El modelo GARCH de Bollerslev (1986) es una extensión útil de el modelo ARCH que intenta resolver el problema de requerir muchos parámetros para describir adecuadamente la volatilidad. El modelo GARCH(m,s) supone que:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^m \alpha_i a_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2 \quad a_t = \sigma_t \epsilon_t$$

dónde

- ϵ_t es una secuencia de variables aleatorias independiente e idénticamente distribuidas con media cero y varianza uno.
- $\alpha_0 > 0$ y $\alpha_i \geq 0$ para $i > 0$ y $\sum_{i=1}^{max(m,s)} (\alpha_i + \beta_i) < 1$. Esta última restricción implica que $Var(a_t)$ es finita.

Para comprender mejor este modelo notemos que.

- La media no condicionada de a_t es cero.
- La varianza no condicionada de a_t viene dada por

$$Var(a_t) = \frac{\alpha_0}{1 - \sum_{i=1}^{max(m,s)} (\alpha_i + \beta_i)} \quad (10)$$

siempre que $0 \leq \sum_{i=1}^{max(m,s)} (\alpha_i + \beta_i) < 1$.

- Notemos que no es sencillo determinar el orden de los modelos GARCH(Tsay,2010-134); no obstante, los modelos más comunes son el GARCH(1,1), GARCH(2,1) y GARCH(1,2).

4.2.1. Estimación de un modelo GARCH

A continuación se presenta el modelo GARCH impletado; en este caso un modelo GARCH(1,2). Teniendo en cuenta que

$$\begin{aligned} r_t &= \mu + a_t \\ a_t &= \sigma_t \epsilon_t \end{aligned}$$

El modelo obtenido, asumiendo que ϵ_t son independientes e idénticamente distribuidos siguiendo la distribución normal standard.

$$r_t = 5,474 \times 10^{-4} + a_t \quad (11)$$

$$\sigma_t^2 = 0,000005346 + 0,1744a_{t-1}^2 + 0,4189\sigma_{t-1}^2 + 0,4023\sigma_{t-2}^2 \quad (12)$$

Los errores standard de los parámetros(12) son 0.0002357,0.000001386,0.02287,0.09271 y 0.08806,respectivamente. Todos los parámetros son significativos con un nivel de significación del 5 %.

El estadístico de Ljung-Box de los residuos estandarizados proporciona $Q(13)=26.2413$ con un p-valor de 0.01578, mientras que en el caso de los residuos al cuadrado estandarizados el estadístico es $Q(13)=854.8379$ con un p-valor de $2,2 \times 10^{-16}$. Por tanto, el modelo GARCH(1,2)no es adecuado con un nivel de significación del 5 %.

Notemos ademas que $\hat{\alpha} + \hat{\beta}_1 + \hat{\beta}_2 = 0,1744 + 0,4189 + 0,4023 = 0,9956$, muy próximo a uno. El hecho de que esta suma sea muy próxima a uno se denomina persistencia en volatilidad (Figure 6).

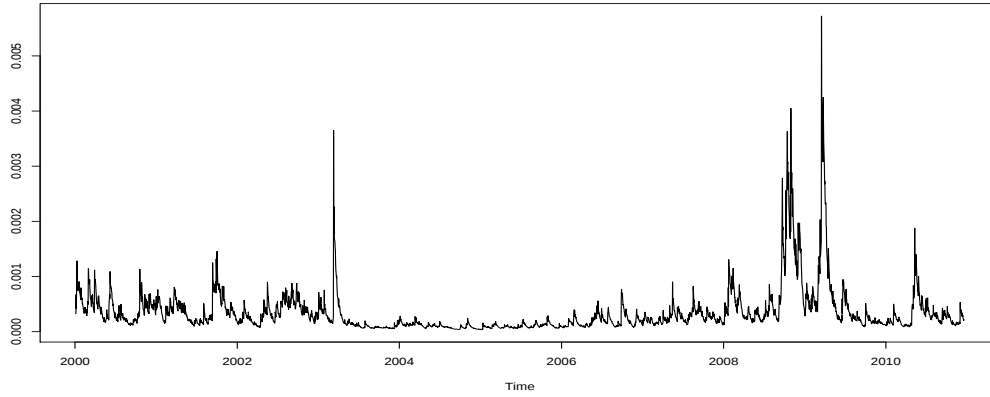


Figura 6: Volatilidad estimada

4.3. Modelo IGARCH

El modelo IGARCH propuesto por Engle y Bollerslev (1986) es un modelo Garch en el cual $\sum_{i=1}^{max(m,s)}(\alpha_i + \beta_i) = 1$. Consecuentemente, el impacto de un shock en volatilidad será persistente, mientras $Var(a_t)$ no exista. Esto parece difícil de justificar para una serie de rendimientos.

El caso real de persistencia en volatilidad está bajo investigación. Sin embargo, desde un punto de vista teórico, la persistencia puede ser causada por cambios ocasionales en el nivel de volatilidad y varios documentos han propuesto considerar modelos GARCH para diferentes conjuntos de datos (Andreou y Ghysels (2002,2004), Galeano y Tsay (2010)).

4.3.1. Estimación de un modelo IGARCH

Estimaremos un modelo IGARCH(1,2). El modelo obtenido, asumiendo que e_t son independientes e idénticamente distribuidos siguiendo la distribución normal standard.

$$r_t = 0,000566 + a_t \quad (13)$$

$$\sigma_t^2 = 0,000005 + 0,178656a_{t-1}^2 + 0,422769\sigma_{t-1}^2 + 0,398574\sigma_{t-2}^2 \quad (14)$$

Los errores standard de los parámetros μ , α_0 , α_1 , β_1 (14) son 0.000237, 0.000002, 0.020101 y 0.093141, respectivamente. El parámetro β_2 es no significativo, pero es necesario para poder obtener $\alpha_1 + \beta_1 + \beta_2 = 1$. El resto de los parámetros son significativos con un nivel de significación del 5 %.

El estadístico de Ljung-Box de los residuos estandarizados proporciona $Q(13)=26.2413$ con un p-valor de 0.01578 y el estadístico de los residuos al cuadrado estandarizados es $Q(13)=854.9967$ con un p-valor de $2,2 \times 10^{-16}$. Por tanto, el modelo IGARCH(1,2) no es adecuado con un nivel de significación del 5 %.

4.4. Modelo EGARCH

El modelo GARCH exponencial (EGARCH) propuesto por Nelson (1991) intenta superar algunas de las debilidades del modelo GARCH, de modo que permita capturar la asimetría debida a la existencia de grandes rendimientos positivos y negativos. (Tsay, 2010-111).

El modelo EGARCH asume que:

$$a_t = \sigma_t \epsilon_t$$

$$\ln(\sigma_t^2) = \alpha_0 + \frac{1 + \beta_1 B + \dots + \beta_{s-1} B^{s-1}}{1 - \alpha_1 B - \dots - \alpha_m B^m} g(\epsilon_t)$$

dónde

- ϵ_t es una secuencia de variables aleatorias independiente e idénticamente distribuidas con media cero y varianza uno.
- La función $g(\epsilon_t)$ viene definida por

$$g(\epsilon_t) = \theta \epsilon_t + \gamma [\epsilon_t - E(|\epsilon_t|)] \quad (15)$$

- θ y γ son constantes reales.
- B es el operador de los retardos que satisface que $Bg(\epsilon_t) = g(\epsilon_{t-1})$, $1 + \beta_1 B + \dots + \beta_{s-1} B^{s-1}$ y $1 - \alpha_1 B - \dots - \alpha_m B^m$ son polinomios con ceros fuera del círculo unidad y no tienen factores comunes.

Para comprender mejor este modelo notemos que:

- La media no condicionada de a_t es cero.
- La media no condicionada de $\ln(\sigma_t^2)$ es α_0 .
- El modelo EGARCH utiliza la varianza condicional registrada para relajar la restricción de positividad de los coeficientes del modelo.
- El uso de $g(\epsilon_t)$ permite al modelo responder asimétricamente a los retardos positivos y negativos de los valores de a_t .

Una forma alternativa para el modelo EGARCH(m,s) es:

$$\ln(\sigma_t^2) = \alpha_0 + \sum_{i=1}^s \alpha_i \frac{|a_{t-i}| + \gamma_i a_{t-i}}{\sigma_{t-i}} + \sum_{j=1}^s \beta_j \ln(\sigma_{t-j}^2) \quad (16)$$

- Un a_{t-i} positivo contribuye $\alpha_i(1 + \gamma_i)|\epsilon_{t-i}|$ a la log-volatilidad, mientras que un a_{t-i} negativo contribuye $\alpha_i(1 - \gamma_i)|\epsilon_{t-i}|$, donde $\epsilon_{t-i} = a_{t-i}/\sigma_{t-i}$.

4.4.1. Estimación de un modelo EGARCH

El siguiente modelo implementado será EGARCH(1,2)(16).

$$\begin{aligned} r_t &= a_t \\ a_t &= \sigma_t \epsilon_t \end{aligned}$$

El modelo obtenido, asumiendo que ϵ_t son independientes e idénticamente distribuidos siguiendo la distribución normal standard:

$$\ln(\sigma_t^2) = -0,259363 - 0,077293 \frac{|a_{t-1}| + 0,295980 a_{t-1}}{\sigma_{t-1}} + 0,497836 \ln(\sigma_{t-1}^2) + 0,469098 \ln(\sigma_{t-1}^2) \quad (17)$$

Los errores standard de los parámetros (17) son 0.056406, 0.015639, 0.031885, 0.077323 y 0.077212, respectivamente. Todos los parámetros son significativos con un nivel de significación del 5 %.

El estadístico de Ljung-Box de los residuos estandarizados proporciona $Q(13)=26.2413$ con un p-valor de 0.01578 y el estadístico de los residuos al cuadrado estandarizados es $Q(13)=849.3888$ con un p-valor de $2,2 \times 10^{-16}$. Por tanto, el modelo EGARCH(1,2) no es adecuado con un nivel de significación del 5 %.

4.5. Modelo TGARCH

El modelo TGARCH propuesto por Glosten, Jagannathan y Runkle (1993) y Zokoian (1994) es una alternativa al modelo EGARCH para manejar el efecto leverage.

El modelo TGARCH(p,q) considera que;

$$\sigma_t^2 = \alpha_0 + \sum_i^m (\alpha_i + \gamma_i N_{t-i}) a_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2 \quad (18)$$

es un indicador para la negatividad de a_{t-i} dado por:

$$N_{t-i} = \begin{cases} 1 & \text{si } a_{t-i} < 0, \\ 0 & \text{si } a_{t-i} \geq 0 \end{cases}$$

- α_i, γ_i y β_j son parámetros no negativos. Además, un a_{t-i} positivo contribuye $\alpha_i a_{t-i}$ a σ_t^2 , mientras que un a_{t-i} negativo tiene un gran impacto $(\alpha_i + \gamma_i) a_{t-i}^2$ con $\gamma_i > 0$.

4.5.1. Estimación de un modelo TGARCH

Para ilustrar este tipo de modelos, vamos a estimar un modelo TGARCH(1,2) (18);

$$r_t = \mu + a_t = a_t \quad (19)$$

$$\sigma_t^2 = 0,0000007 + (0,115505 + 0,129181 N_{t-1}) a_{t-1}^2 + 0,410941 \sigma_{t-1}^2 + 0,399310 \sigma_{t-2}^2 \quad (20)$$

Los errores standard de los parámetros (20) son 0.000002, 0.020126, 0.028206, 0.089137 y 0.084771, respectivamente. Todos los parámetros son significativos con un nivel de significación del 5 %.

El estadístico de Ljung-Box de los residuos estandarizados proporciona $Q(13)=26.2413$ con un p-valor de 0.01578 y el estadístico de los residuos al cuadrado estandarizados es $Q(13)=849.3888$ con un p-valor de $2,2 \times 10^{-16}$. Por tanto, el modelo TGARCH(1,2) no es adecuado con un nivel de significación del 5 %.

5. Conclusiones

De los modelos construidos deducimos que ninguno de los modelos condicionalmente heterocedásticos estimados en los apartados anteriores no son adecuados con un nivel de significación de 5 % para modelizar la volatilidad. Por ello propondríamos la estimación de un modelo que utilice una función estocástica para describir la volatilidad. Como por ejemplo;

- Modelo bilinear de Granger y Andersen (1978).
- "Threshold autorregresive model"(TAR) de Tong y Lin (1980).
- Modelo de transición suave AR (STAR) de Chan y Tong.
- Modelo de Markov "switching"de Tong (1983).

Otra posibilidad sería probar con una distribución t-student en lugar de la normal para resolver el problema de las colas pesadas. También sería útil probar con una mixtura de dos distribuciones normales (Bai,Rusell y Tiao, 3003) o con una mixtura infinita de distribuciones gaussianas para solucionar el problema de la asimetría (Ausín, Galeano y Ghosh, 2010).

A pesar de que se han implementado varios modelos y que ninguno de ellos ha resultado apropiado por el efecto de "jumps", se considera oportuno estimar un modelo que tenga en cuenta estos efectos.

6. Código implementado

```
gas= read.table("GN.txt",header=TRUE,sep=" ",dec=".",fill=TRUE)

names(gas)
attach(gas)
length(gas[,2]) #2854
gas[,2]

library(moments)
library(tseries)
library(fBasics)
library(fGarch)
library(numDeriv)
library(chron)
library(Rsolnp)
library(sandwich)
library(Rcpp)
library(rgarch)

# Gráfico de los rendimientos

gas[,1]<-as.Date(gas[,1])
x11()
plot(gas,type="l",xlab="Year",ylab="Log-returns of GN",main="Gas Natural")

# QQ-plot

windows()
qqnorm(gas[,2])
qqline(gas[,2])

# Momentos

mean(gas[,2]) #-0.0002299469
sd(gas[,2]) #0.01792231
skewness(gas[,2]) #-0.3769381
kurtosis(gas[,2]) #6.880669
basicStats(gas[,2])
```

```

# Estimacion kernel de la densidad de los datos

d1 = density(gas[,2])
x11()
par(mfrow=c(2,1))
plot(d1$x,d1$y,type='l',lwd = 2, col="blue")
plot(d1$x,d1$y,type='l',lwd = 2, col="blue")
lines(gas[,2][order(gas[,2])],dnorm(gas[,2][order(gas[,2])],mean(gas[,2]),
sd(gas[,2])),type='l',lwd=2,col="green")

# ACF de los rendimientos

x11()
windows()
par(mfrow=c(2,1))
acf(gas[,2],main="ACF of log returns of GN")
pacf(gas[,2],main="PACF of log returns of GN")

# Contraste Ljung-Box sobre los rendimientos

Box.test(gas[,2],lag=3,type="Ljung-Box")
#X-squared = 4.6588, df = 3, p-value = 0.1986

# ACF y PACF de los rendimientos al cuadrado
x11()
par(mfrow=c(2,1))
acf(gas[,2]^2,main="ACF of squared log returns of GN")
pacf(gas[,2]^2,main="PACF of squared log returns of GN")

# Ljung-Box sobre los residuos
at=gas[,2]-mean(gas[,2])
Box.test(at^2,lag=13,type="Ljung-Box")
# X-squared = 846.6681, df = 13, p-value < 2.2e-16

#PACF de los log-rendimientos al cuadrado permite conocer el orden del modelo.
#Probaremos un modelo ARCH(13).

```

```
#####
#           Modelos           #
#####
```

```
# Modelo ARCH
```

```
ajusteARCH <- garchFit(gas[,2] ~garch(13,0), data = gas[,2],trace=F, include.mean=FALSE)
ajusteARCH
s0 <- ugarchspec(variance.model = list(model = "fGARCH", garchOrder = c(13, 0),
submodel = "GARCH", external.regressors = NULL, variance.targeting = FALSE),
mean.model = list(armaOrder = c(0, 0), include.mean = TRUE, garchInMean = FALSE,
inMeanType = 1, arfima = FALSE, external.regressors = NULL), distribution.model = "norm",
start.pars = list(), fixed.pars = list(alpha8=0))
ajusteARCH <- ugarchfit(s0,gas[,2])
ajusteARCH
```

```
Box.test(residuals(ajusteARCH),lag=13,type="Ljung-Box")
# X-squared = 26.2413, df = 13, p-value = 0.01578
Box.test(residuals(ajusteARCH)^2,lag=13,type="Ljung-Box")
# X-squared = 854.6708, df = 13, p-value < 2.2e-16
```

```
res=residuals(ajusteARCH)
Box.test((res-mean(res))/sd(res),lag=13,type="Ljung-Box")
# X-squared = 26.2413, df = 13, p-value = 0.01578
Box.test((res^2-mean(res^2))/sd(res^2),lag=13,type="Ljung-Box")
# X-squared = 854.6708, df = 13, p-value < 2.2e-16
```

```
# Modelo GARCH
```

```
ajusteGARCH <- garchFit(gas[,2] ~ garch(1,2), data = gas[,2], trace = F)
ajusteGARCH
```

```
Box.test(residuals(ajusteGARCH),lag=13,type="Ljung-Box")
# X-squared = 26.2413, df = 13, p-value = 0.01578
Box.test(residuals(ajusteGARCH)^2,lag=13,type="Ljung-Box")
# X-squared = 854.8379, df = 13, p-value < 2.2e-16
```

```
res=residuals(ajusteGARCH)
Box.test((res-mean(res))/sd(res),lag=13,type="Ljung-Box")
```

```
# X-squared = 26.2413, df = 13, p-value = 0.01578
Box.test((res^2-mean(res^2))/sd(res^2),lag=13,type="Ljung-Box")
# X-squared = 854.8379, df = 13, p-value < 2.2e-16
```

```
# Gráfico de la volatilidad estimada
```

```
volatilityGARCH = matrix(NA,nrow=2854,ncol=2)
volatilityGARCH[,1] = gas[1:2854,1]
volatilityGARCH[,2] = ajusteGARCH$h.t
volatilityGARCH = as.data.frame(volatilityGARCH)
volatilityGARCH[,1] = as.Date(volatilityGARCH[,1])
windows()
plot(volatilityGARCH,type="l",ylab="",xlab="Time")
```

```
# Modelo IGARCH
```

```
s1 <- ugarchspec(variance.model = list(model = "iGARCH", garchOrder = c(1, 2),
submodel = NULL, external.regressors = NULL, variance.targeting = FALSE),
mean.model = list(armaOrder = c(0, 0), include.mean = TRUE, garchInMean = FALSE,
inMeanType = 1, arfima = FALSE, external.regressors = NULL), distribution.model = "norm",
start.pars = list(), fixed.pars = list())
```

```
ajusteIGARCH <- ugarchfit(s1,gas[,2])
```

```
ajusteIGARCH
```

```
Box.test(residuals(ajusteIGARCH),lag=13,type="Ljung-Box")
# X-squared = 26.2413, df = 13, p-value = 0.01578
Box.test(residuals(ajusteIGARCH)^2,lag=13,type="Ljung-Box")
# X-squared = 854.9967, df = 13, p-value < 2.2e-16
```

```
res=residuals(ajusteIGARCH)
Box.test((res-mean(res))/sd(res),lag=13,type="Ljung-Box")
# X-squared = 26.2413, df = 13, p-value = 0.01578
Box.test((res^2-mean(res^2))/sd(res^2),lag=13,type="Ljung-Box")
# X-squared = 854.9967, df = 13, p-value < 2.2e-16
```

```
# Modelo EGARCH
```

```
spec1 <- ugarchspec(variance.model = list(model = "eGARCH", garchOrder = c(1, 2),
submodel = NULL, external.regressors = NULL, variance.targeting = FALSE),
mean.model = list(armaOrder = c(0, 0), include.mean = FALSE, garchInMean = FALSE,
inMeanType = 1, arfima = FALSE, external.regressors = NULL), distribution.model = "norm",
start.pars = list(), fixed.pars = list())
```

```
ajusteEGARCH <- ugarchfit(spec1,gas[,2])
```

```
ajusteEGARCH
```

```
Box.test(residuals(ajusteEGARCH),lag=13,type="Ljung-Box")
```

```
# X-squared = 26.2413, df = 13, p-value = 0.01578
```

```
Box.test(residuals(ajusteEGARCH)^2,lag=13,type="Ljung-Box")
```

```
# X-squared = 849.3888, df = 13, p-value < 2.2e-16
```

```
res=residuals(ajusteTGARCH)
```

```
Box.test((res-mean(res))/sd(res),lag=13,type="Ljung-Box")
```

```
# X-squared = 26.2413, df = 13, p-value = 0.01578
```

```
Box.test((res^2-mean(res^2))/sd(res^2),lag=13,type="Ljung-Box")
```

```
# X-squared = 849.3888, df = 13, p-value < 2.2e-16
```

```
# Modelo TGARCH
```

```
spec1 <- ugarchspec(variance.model = list(model = "gjrgARCH", garchOrder = c(1, 2),
submodel = NULL, external.regressors = NULL, variance.targeting = FALSE),
mean.model = list(armaOrder = c(0, 0), include.mean = FALSE, garchInMean = FALSE,
inMeanType = 1, arfima = FALSE, external.regressors = NULL), distribution.model = "norm",
start.pars = list(), fixed.pars = list())
```

```
ajusteTGARCH <- ugarchfit(spec1,gas[,2])
```

```
ajusteTGARCH
```

```
Box.test(residuals(ajusteTGARCH),lag=13,type="Ljung-Box")
```

```
# X-squared = 26.2413, df = 13, p-value = 0.01578
```

```
Box.test(residuals(ajusteTGARCH)^2,lag=13,type="Ljung-Box")
```

```
# X-squared = 849.3888, df = 13, p-value < 2.2e-16
```

```
res=residuals(ajusteTGARCH)
Box.test((res-mean(res))/sd(res),lag=13,type="Ljung-Box")
# X-squared = 26.2413, df = 13, p-value = 0.01578
Box.test((res^2-mean(res^2))/sd(res^2),lag=13,type="Ljung-Box")
# X-squared = 849.3888, df = 13, p-value < 2.2e-16
```