

Modelizar la volatilidad

Modeling volatility

Dr. Agustín ALONSO RODRÍGUEZ

Real Centro Universitario

“Escorial-María Cristina”

San Lorenzo del Escorial

Resumen: La volatilidad es un concepto clave a tener presente en la gestión de inversiones. Debido a que es inobservable, se han buscado otras magnitudes observables, y teóricamente vinculadas con ella. Entre estas hay que mencionar el *rango*: la diferencia entre el máximo y mínimo de los precios de cotización diarios de un activo. Tomando como base una muestra de 100 observaciones del Índice General de Cotización de la Bolsa de Madrid, se ajusta un modelo dinámico, de la familia de *espacio de estados*, para el logaritmo del *rango*, y una vez validado, se hace la previsión para diez valores futuros del logaritmo del *rango*, se evalúan dichas previsiones y se presentan unas conclusiones.

Abstract: Volatility is a key concept in portfolio management. Due to the fact that it is not directly observable, several other magnitudes have been devised related to it. Here it is used the *range*, the difference between the high and low daily transaction prices of a stock. Using a sample of 100 observations of the Madrid Stock Exchange Index a model of the *state space* family is established for the logarithm of the *range*. After its validation, the model is used to forecast ten future values of the log of the *range*. The forecasts are evaluated and some conclusions are presented.

Palabras clave: Volatilidad, rango, gestión de inversiones, modelos dinámicos, espacio de estados, filtro de Kalman, validación, previsión o predicción, paquetes: dlm, forecast, R.

Keywords: Volatility, portfolio management, range of quotations, dynamic linear models, state space, Kalman filter, forecasts, software packages: dlm, forecast, R.

Sumario:

- I. Introducción.**
- II. Modelos para datos financieros.**
- III. Características de la volatilidad.**
- IV. Información de la Bolsa de Madrid.**
- V. Un modelo para la evolución temporal del logaritmo del rango.**
- VI. Muestra utilizada.**
- VII. Modelización con el paquete dlm.**
- VIII. Modelización con el paquete forecast.**
- IX. Conclusiones.**
- X. Bibliografía.**

Recibido: noviembre de 2013.

Aceptado: diciembre de 2013.

I. INTRODUCCIÓN

El mes de septiembre del 2013 ha visto resurgir un ambiente de optimismo en la economía española: la Bolsa alcanza los máximos históricos de otros años, la Prima de Riesgo alcanza valores mínimos, el Tesoro logra colocar Deuda a niveles de interés no vistos desde hacía años, y desde fuentes próximas al Gobierno se afirma que está llegando a raudales dinero a España.

En este clima de euforia, se quiere considerar la posibilidad establecer una señal de fácil elaboración, que permita con rapidez y precisión, mantener este ambiente de optimismo generalizado.

Y en pos de este objetivo, la abundante información generada por la Bolsa de Valores centra el marco de referencia de este estudio. En concreto, el Índice General de Cotización de la Bolsa de Madrid proporciona la información sobre la que gira el estudio de la volatilidad que se presenta en este trabajo.

II. MODELOS PARA DATOS FINANCIEROS

El análisis de las series históricas de datos financieros se centra en los rendimientos de los activos, en lugar de sus precios. Dos razones de peso son, en primer lugar, que los rendimientos de los títulos centran la atención de los inversores, además de ser un resumen de las oportunidades de inversión, y en segundo lugar, ser más fáciles de tratar estadísticamente que los precios de los activos.

Y relacionado con los rendimientos está el tema de la volatilidad: la evolución de la varianza condicionada de los rendimientos, a lo largo del tiempo. Esta variabilidad de los rendimientos juega un papel importante en la gestión de las carteras de inversión.

Muchos de los modelos propuestos en la literatura para modelizar la volatilidad de los activos financieros se enmarcan dentro del nombre genérico de modelos

de heteroscedasticidad condicionada. La volatilidad en este ámbito se define como la desviación estándar del rendimiento del activo financiero en cuestión. Modelizar la volatilidad de una serie histórica permitiría mejorar las predicciones de la gestión.

Por último, la volatilidad del *índice* de un mercado se está convirtiendo recientemente en un instrumento financiero. Así tenemos el VIX, el *volatility index* elaborado por el *Chicago Board of Option Exchange* (CBOE), en vistas al mercado de futuros, establecido el 26 de marzo de 2004.

III. CARACTERÍSTICAS DE LA VOLATILIDAD

La característica más notable de la volatilidad de un activo es que no es directamente observable. Por ejemplo, los rendimientos diarios de IBM no permiten obtener la volatilidad diaria del activo, ya que sólo se dispone de una observación diaria. Si tuviéramos más datos diarios, por ejemplo, los rendimientos del título cada diez minutos, entonces sería posible estimar esta volatilidad diaria. Además, habría que tener presente la volatilidad diurna y nocturna. Esta falta de observabilidad de la volatilidad hace difícil evaluar las predicciones de los modelos de heteroscedasticidad condicionada.

Si en el mercado de opciones se acepta la idea de que los precios están gobernados por un modelo como la fórmula de Black-Scholes, entonces se puede utilizar el precio como un estimador de la volatilidad implícita. Sin embargo, este enfoque es criticado por utilizar un modelo que se basa en una serie de supuestos que pueden no darse en la realidad. Por ejemplo, en la fórmula de Black-Scholes, en la *European call option* la volatilidad implícita es resultado del supuesto de que el precio del título sigue un movimiento geométrico de Brown.

La experiencia, por otro lado, muestra que la volatilidad implícita obtenida con un modelo GARCH es inferior a la volatilidad implícita real del rendimiento del título. El índice VIX del CBOE es un índice de la volatilidad implícita.

A pesar de no ser observable, los rendimientos de los activos financieros muestran otras características de la volatilidad que conviene destacar.

En primer lugar, la volatilidad se manifiesta en *clusters*, en grupos, es decir, en determinados periodos temporales es alta, y en otros periodos es baja.

En segundo lugar, presenta un comportamiento continuado, o sea, no muestra saltos, o éstos son raros.

En tercer lugar, no muestra tendencias crecientes al infinito, o en otras palabras, varía dentro de determinados rangos, lo que estadísticamente significa que es frecuentemente estacionaria.

Y por último, parece reaccionar de forma distinta ante un alza o caída notable de los precios, lo que se conoce como el efecto *leverage*.

Estas características han sugerido el establecimiento de modelos con una finalidad concreta. Así, el modelo EGARCH fue desarrollado para captar la asimetría inducida por grandes alzas o caídas de los rendimientos.

Los modelos para la volatilidad, tomando como referencia la varianza condicionada de los rendimientos bursátiles, tienen como punto de partida el modelo de Engle ARCH: Autoregressive Conditional Heteroscedasticity, de 1982, que ha generado toda una serie de modelos relacionados: GARCH, EGARCH, TGARCH, etc.

Aquí no se sigue en esta línea de razonamiento, sino en los enfoques llamados alternativos (Tsay, 2010, p. 159), utilizando la información publicada por la Bolsa de Madrid sobre precios diarios de cotización de apertura, cierre, mínimo y máximo. Y más concretamente, se toma en consideración el *rango* de variación de las cotizaciones: la diferencia entre el máximo y mínimo diarios de los precios de contratación. Esta información ha sido también considerada útil para la estimación de la volatilidad, por ejemplo, por Parkinson (1980), Garman y Klass (1980), Rogers y Satchell (1991), y Yang y Zhang (2000).

Más próximo al enfoque de este trabajo hay que citar el de Alizadeh, Brandt y Diebold (2002): *Range-Based estimation of stochastic volatility models*. Tras examinar las dificultades existentes sobre el tema, estos autores proponen un método simple y altamente eficiente para la estimación de la volatilidad basado en el *rango*.

El rango, definido como la diferencia entre el mayor y menor valor alcanzado por el precio de un activo en un día concreto y en un determinado periodo de tiempo, es considerado como una aproximación a la volatilidad. Estos autores muestran las buenas propiedades estadísticas del *rango* de los precios bursátiles, que permiten utilizar máxima-verosimilitud para estimar la volatilidad estocástica, e incluso extraer la volatilidad oculta.

Aquí no se entra en estos detalles, sino que sobre la serie histórica establecida por el estadístico *rango* utilizamos un modelo lineal dinámico, de la familia de los *espacios de estados*, para estimar y predecir el comportamiento del

logaritmo del *rango* en un horizonte de un par de semanas. La estacionariedad que muestra el *rango* permite, con la ayuda de un modelo lineal dinámico, establecer un estimador del comportamiento de la volatilidad.

Conviene señalar con Tsay, (2010, p. 164) que en la práctica diaria sólo se recoge un valor máximo observado, que probablemente sea inferior al máximo verdadero, y un mínimo, que también puede que sea mayor que el auténtico mínimo cotizado. La consecuencia que se extrae de este comportamiento es que puede que el *rango* subestime el verdadero rango, y por ende subestime también la volatilidad. No obstante, este sesgo en la estimación de la volatilidad depende de la frecuencia de cotización de los títulos, y para títulos muy cotizados el sesgo puede ser insignificante.

IV. INFORMACIÓN DE LA BOLSA DE MADRID

El Boletín de cotizaciones de la Bolsa de Madrid ofrece información sobre el precio de apertura, cierre, máximo, y mínimo de los títulos cotizados. Con estos datos, es posible establecer el *rango* para títulos individuales, contratos de futuros, índice de cotización, títulos del Tesoro, monedas, etc. Además en las publicaciones de carácter económico son populares los gráficos para resaltar esta información. En el *análisis técnico*, es uno de los indicadores importantes a tener presente, Edwards y Magee (1997).

Con todo, como resaltan Alizadeh, Brandt y Diebold (2002, p. 1048), el *rango* ha sido olvidado en los estudios recientes sobre volatilidad estocástica.

El logaritmo del *rango* es para Alizadeh, Brandt y Diebold (2002), una medida altamente eficiente de la volatilidad, al menos desde Parkinson (1980), habiendo sido formalizada por Andersen y Bollerslev (1998).

El logaritmo del *rango* presenta un comportamiento de acuerdo con la distribución de una variable normal o de Gauss, lo que permite que la función de cuasi-verosimilitud que se maximiza para estimar los parámetros de un modelo, sea realmente una verdadera función de verosimilitud.

V. UN MODELO PARA LA EVOLUCIÓN TEMPORAL DEL RANGO

El *rango*, en su evolución a lo largo del tiempo, es una serie temporal o histórica. Su logaritmo es la serie a considerar en este estudio, y lo hacemos bajo el enfoque de los modelos lineales dinámicos, dentro del enfoque más general de los modelos de *espacio de estados*.

Estos modelos vienen representados por un par de ecuaciones.

$$Y_t = F_t \theta_t + v_t; v_t \sim N_m(0, V_t) \quad (1)$$

$$\theta_t = G_t \theta_{t-1} + w_t; w_t \sim N_p(0, W_t) \quad (2)$$

siendo $\theta_0 = N_p(m_0, C_0)$ y G_t y F_t , matrices de dimensiones $(p \times p)$ y $(m \times p)$ respectivamente. Además $(v_t)_{t \geq 1}$ y $(w_t)_{t \geq 1}$ son secuencias independientes de vectores aleatorios gaussianos, con media cero y matrices de varianzas y covarianzas conocidas o a estimar: $(V_t)_{t \geq 1}$ y $(W_t)_{t \geq 1}$ respectivamente. Se supone también que θ_0 es independiente de v_t y de w_t .

La ecuación en (1) se conoce como la *ecuación de observaciones* y la (2) como la *ecuación del sistema* o *ecuación de estados*.

En el marco de referencia de las series temporales univariantes, las ecuaciones anteriores se simplifican como sigue:

$$Y_t = \mu_t + v_t; v_t \sim N(0, V) \quad (3)$$

$$\mu_t = \mu_{t-1} + w_t; w_t \sim N(0, W) \quad (4)$$

es decir, $m = p = 1$, $\theta_t = \mu_t$ y $F_t = G_t = 1$.

El sistema en (3) y (4) se conoce como el *modelo paseo aleatorio más ruido* (random walk plus noise), y es el modelo apropiado cuando la serie temporal no presenta clara tendencia ni variación estacional, como es nuestro caso.

El modelo en sí manifiesta que la serie temporal Y_t evoluciona a partir de un nivel, μ_t al que se asocian variaciones aleatorias. Por esta razón, el modelo recibe también el nombre de *modelo de nivel local*.

Hay que recalcar que μ_t es una *caminata aleatoria*, de por sí no estacionaria, lo que pone de relieve que los modelos dinámicos lineales pueden utilizarse para modelar series temporales no estacionarias. En el enfoque de los modelos ARMA, es necesario transformar la serie original en estacionaria para poder identificar el modelo ARMA más apropiado.

En un modelo de *espacio de estados* el objetivo es la estimación de los no-observables estados (o factores ocultos, en la terminología de Martin, Hurn y Harris, 2013, pp. 544 y siguientes), y la predicción de observaciones futuras sobre la base de los datos observados. Tanto la estimación como la

predicción se logran computando las distribuciones de probabilidad condicionadas de las magnitudes de interés, dada la información disponible.

Para la estimación del vector de estados se computa la probabilidad condicional

$$\pi(\theta_s | y_{1:t})$$

Si $s=t$, tenemos lo que se conoce como *filtrado*, *filtering*, del vector de estados.

Si $s>t$, tenemos lo que se denomina *predicción*, del vector de estados.

Y, si $s<t$, tenemos el *suavizado*, *smoothing*, del vector de estados.

En el *filtrado* del vector de estados se supone que las observaciones se van incorporando secuencialmente, como en nuestro caso. En estas circunstancias, se trata de estimar el valor en el momento t , el actual, del vector de estados, tomado en consideración los nuevos datos en el momento $t+1$. Para ello, se computa la densidad condicional: $\pi(\theta_t | y_{1:t})$. El *filtro de Kalman* permite actualizar la inferencia en el momento t sobre el vector de estados con la llegada de los nuevos datos, lo que hace que se pueda pasar de $\pi(\theta_t | y_{1:t})$ a $\pi(\theta_{t+1} | y_{1:t+1})$.

En el paso de suavizado o análisis retrospectivo, se estima la secuencia de estados en los momentos $t:1 \dots t$, dados los valores de las observaciones $y_{1:t}$. Para ello se computa la distribución condicional de $\theta_{1:t}$ dados $y_{1:t}$ mediante el *filtro de Kalman*.

Predecir es uno de los objetivos, por no decir el principal objetivo, en el análisis de series temporales. La estimación del vector de estados es justo el paso previo para la predicción del valor futuro de la serie temporal: Y_{t+1} tomando en consideración la información en $y_{1:t}$. Y lo primero a obtener es la estimación del valor futuro θ_{t+1} del vector de estados, sobre la que se realiza la predicción. La predicción de un valor futuro (un paso adelante) de la función de densidad es: $\pi(\theta_{t+1} | y_{1:t})$ se basa sobre el filtrado de la función de densidad de θ_t y a partir de ella se hace la predicción de $\pi(y_{t+1} | y_{1:t})$, de nuevo mediante el *filtro de Kalman*.

Si se trata de hacer la predicción de k valores o k pasos hacia adelante: Y_{t+k} lo primero a realizar es la predicción k pasos adelante de la función de densidad $\pi(\theta_{t+k} | y_{1:t})$ para sobre ella establecer el valor de Y_{t+k} . Por supuesto, las predicciones se hacen más y más inseguras al alargarse el horizonte temporal, si bien es posible cuantificar esta incertidumbre a partir de la probabilidad de la función de densidad de Y_{t+1} dados los valores de $y_{1:t}$.

En concreto, la estimación de la media condicionada $E(Y_{t+1} | y_{1:t})$ constituye la predicción puntual óptima de un valor hacia delante de Y_{t+1} en cuanto que minimiza el *error cuadrático medio*. En cuánto función de los k pasos hacia adelante, $E(Y_{t+k} | y_{1:t})$ se denomina la función de predicción.

En este trabajo estimaremos el *modelo de nivel local* con ayuda del paquete *dlm*, de Giovanni Petris y colaboradores, ejecutado dentro del entorno gráfico y numérico que es el programa R.

Así mismo, se tendrá también en cuenta el paquete *forecast*, de Robert J. Hyndman y colaboradores, ejecutable dentro del programa R. La aportación de Hyndman constituye una ligera modificación del modelo de *espacio de estados* aquí considerado, y recibe el nombre de *suavizado exponencial (exponential smoothing)*.

La modificación aludida constitutiva del *suavizado exponencial* consiste en reducir a una sola fuente de perturbación aleatoria las perturbaciones presentes en (3) y (4). Es decir, el sistema de ecuaciones pasa a ser:

$$Y_t = \mu_t + \varepsilon_t; \varepsilon_t \sim N(0, \sigma^2) \quad (5)$$

$$\mu_t = \mu_{t-1} + \alpha \varepsilon_t \quad (6)$$

siendo α un factor conocido o a estimar. La representación en (5) y (6) es una de las múltiples posibles dentro de la terminología desarrollada por Hyndman. En concreto es el modelo MNN (modelo de error multiplicativo, M, sin tendencia, N, ni estacionalidad, N).

Hay que señalar la muy amplia bibliografía desarrollada en torno al modelo *exponential smoothing* gracias al impulso dado por Hyndman.

VI. LA MUESTRA UTILIZADA

Los datos utilizados en nuestro estudio están tomados de la página WEB de la Bolsa de Madrid, y corresponden al Índice general de cotización diario, desde el 3 de junio de 2013 hasta el 18 de octubre de ese mismo año. Son un total de 100 observaciones diarias del Índice, referentes a cuatro series: valor de cierre, de apertura, máximo y mínimo. No ha habido una especial intencionalidad al elegir esta muestra; parecía oportuno elegir un par de meses anteriores a septiembre y continuar hasta el 18 del mes de octubre. Quedan prácticamente dos

semanas hasta finalizar el mes, del 21 al 31, las dos semanas utilizadas para la previsión o predicción de la variable de interés: el logaritmo del rango (*lrango*), fuera de la muestra utilizada para la estimación del modelo.

VII. MODELACION CON EL PAQUETE *dlnm*

Como primer paso, comprobemos si nuestra serie logaritmo del rango *lrango* se ajusta a las conclusiones de Alizadeh, Brandt y Diebold (2002) sobre el comportamiento normal de estas observaciones. El resultado de este examen exploratorio viene recogido en la figura 1.

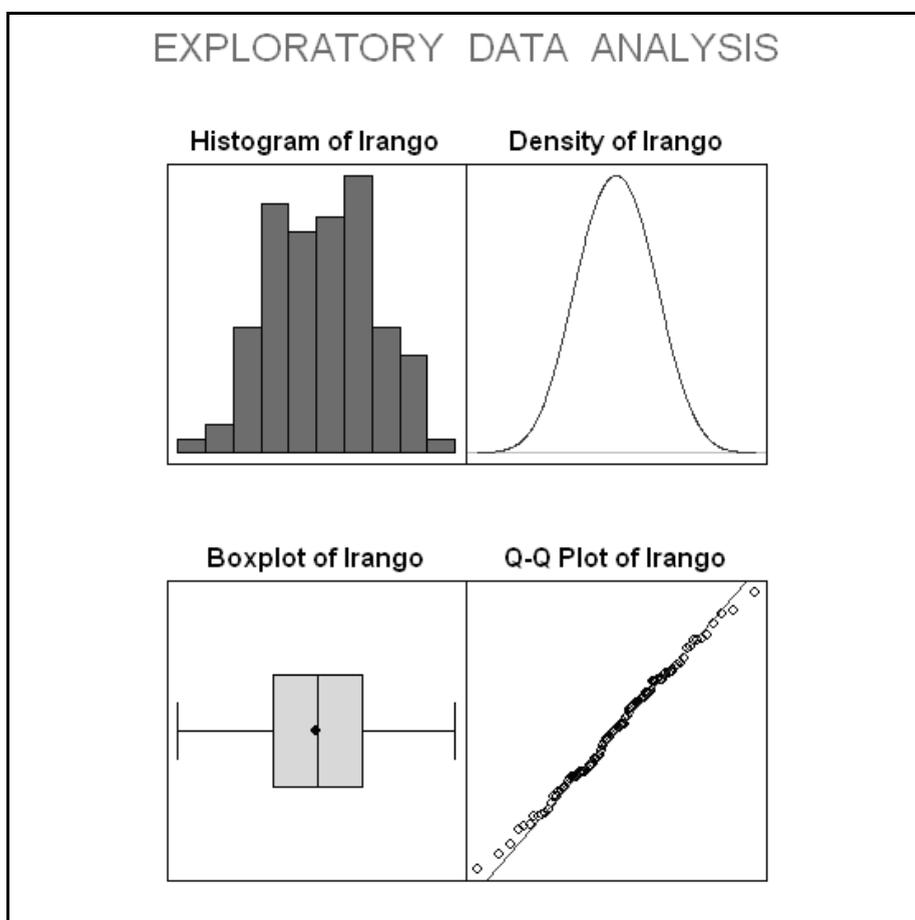


Figura 1.

Esta figura muestra un aceptable ajuste de nuestros datos al comportamiento de una variable aleatoria con distribución normal. Si añadimos el contraste de Shapiro-Wilk, obtenemos:

```
Shapiro-Wilk normality test
data:  lrango
W = 0.9927, p-value = 0.867
```

que confirma el comportamiento normal de nuestros datos.

Nuestra serie, a lo largo del tiempo, presenta el aspecto representado en la figura 2.

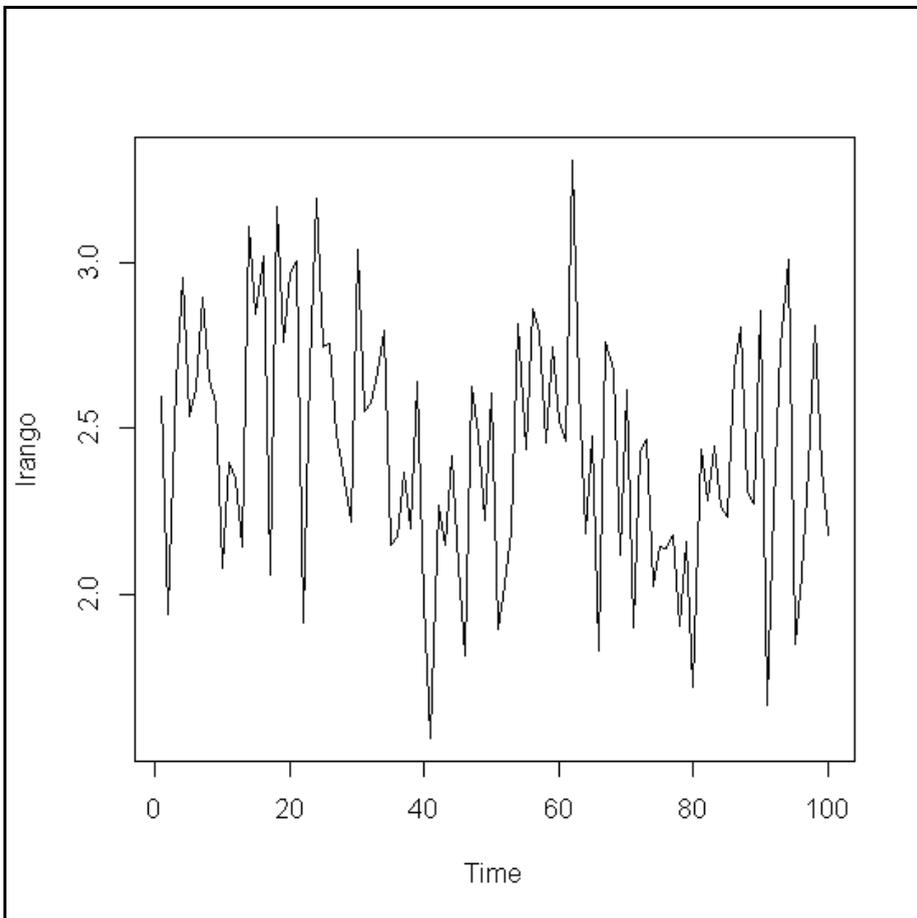


Figura 2.

Un comportamiento aparentemente estacionario y sin estacionalidad, apto para ajustar el modelo dinámico de nivel local:

$$Y_t = \mu_t + v_t; v_t \sim N(0, V) \quad (3)$$

$$\mu_t = \mu_{t-1} + w_t; w_t \sim N(0, W) \quad (4)$$

La primera tarea a cometer es la estimación de las varianzas de ambas perturbaciones. Dentro del entorno R podemos hacerlo de dos maneras. La primera con ayuda del paquete *dlm*, la segunda con la instrucción *StructS*, presente en el paquete base de la distribución de R. Nos limitamos aquí a emplear *dlm*.

El paquete *dlm* utiliza máxima verosimilitud para lograr su objetivo. Lo hace mediante las instrucciones

```
modBuild = function(par) {
  dlmModPoly(order=1, dV=exp(par[1]), dW=exp(par[2])) }
modMLE=dlmMLE(y, rep(0, 2), modBuild)
```

El resultado es:

$$\text{var}(v_t) = 0.1135707 \text{ y la } \text{var}(w_t) = 0.002561589$$

Con estas estimaciones, establecemos el modelo:

```
modelo1 = modBuild(modMLE$par)
```

y al teclear `modelo1`, obtenemos sus elementos:

```
$FF
      [, 1]
[1, ]    1

$V
      [, 1]
[1, ] 0.1135707

$GG
      [, 1]
[1, ]    1

$W
      [, 1]
[1, ] 0.002561589
```

Siendo los valores iniciales de m y C

```
$m0
[1] 0

$C0
      [,1]
[1,] 1e+07
```

Es decir, nuestro modelo queda especificado como

$$Y_t = \mu_t + v_t; v_t \sim N(0, 0.1135707) \quad (3)$$

$$\mu_t = \mu_{t-1} + w_t; w_t \sim N(0, 0.002561589) \quad (4)$$

con $m_0=0$ y $C_0=0.000000$

Todos los componentes del modelo se consiguen mediante `names(modelo1)`. Para obtener el vector de estados μ_t la instrucción `d1mFilter` lo estima:

```
modeloFilt = d1mFilter(y, modelo)
```

siendo `y`, la serie temporal *lrango*.

Su representación, superpuesta sobre la serie *lrango*, viene recogida en la figura 3.

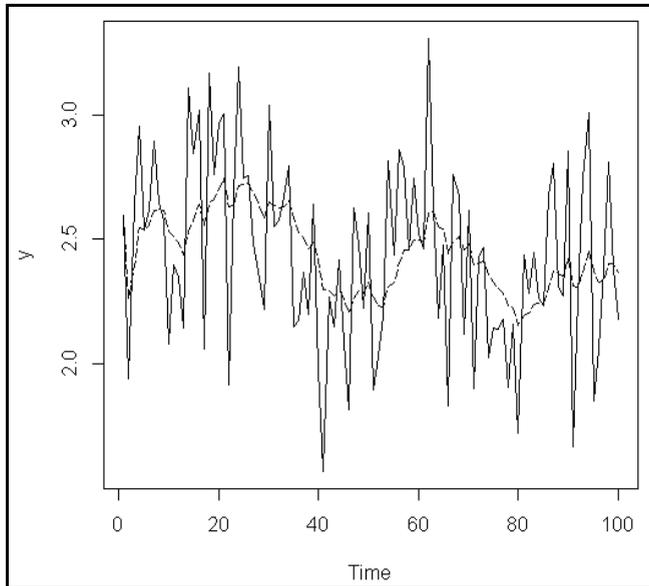


Figura 3.

Para comprobar la idoneidad de nuestro modelo, acudimos al análisis de los residuos. La instrucción `tsdiag(modeloFilt)` muestra el resultado que recoge la figura 4.

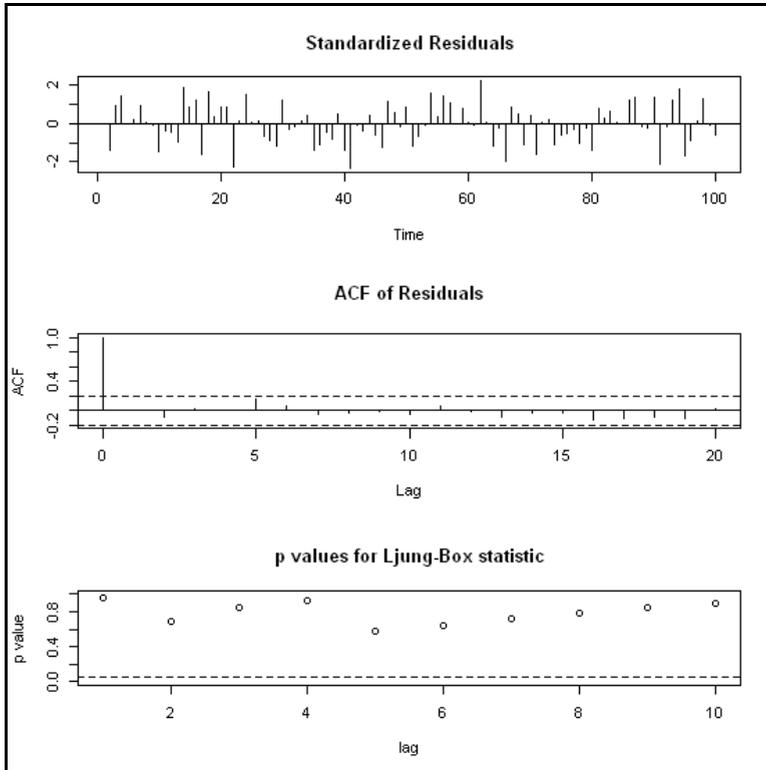


Figura 4.

es decir, los residuos son homoscedásticos, y linealmente independientes entre sí, como muestra la función de autocorrelación ACF de los residuos, y los p-valores del contraste Ljung-Box, que permiten comprobar que los coeficientes de la función de autocorrelación, aquí hasta el retardo 10, son estadísticamente hablando iguales a cero, con un 95 % de confianza.

Cabe también comprobar si los residuos de nuestro modelo estimado muestran un comportamiento normal.

Shapiro-Wilk normality test

```
data: residuals(modeloFilt, sd = F)
W = 0.9881, p-value = 0.5126
```

El modelo pues, queda validado.

Se puede, pues, pasar a la etapa de previsión o predicción de un valor hacia adelante, fuera de la muestra utilizada para la construcción del modelo. Es decir, pasamos a la predicción del valor correspondiente al 21 de octubre de 2013, lunes.

Con ayuda de la instrucción *dmlForecast*, obtenemos la predicción, aquí denominada *fore*, por abreviar:

```
fore = dlmForecast(modeloFilt, nAhead=1, sampleNew=10)
```

el valor resultante es: 2.370114.

La banda de confianza del 95% muestra el resultado:

```
Obs liminf95 limsup95 prevision puntual
101 1.772609 2.96762 2.370114
```

Lo hecho para la previsión de la observación 101, es decir, del valor correspondiente al logaritmo del *rango* del 21 de octubre, se repite para el 22 de octubre, 23 de octubre y así hasta el 31 de octubre, tras aumentar la muestra con los valores reales correspondientes a cada uno de estos días. Agrupando las predicciones con las bandas de confianza del 95, y añadiendo en la última columna los valores reales del logaritmo del rango: *lrango*, se obtiene la tabla 1

obs	liminf95	limsup95	forecast	lrango
101	1.772609	2.96762	2.370114	2.295560
102	1.765102	2.954363	2.359732	1.652497
103	1.657845	2.863674	2.26076	2.574138
104	1.705541	2.909314	2.307428	2.354228
105	1.714723	2.912702	2.313712	2.331173
106	1.719938	2.912142	2.31604	3.016025
107	1.800613	3.00801	2.404312	3.060115
108	1.893529	3.113731	2.50363	3.113515
109	2.016444	3.246481	2.631462	2.758743
110	2.049764	3.27491	2.662337	2.320425

Tabla 1 .

Se visualizan estos datos, en la figura 5.

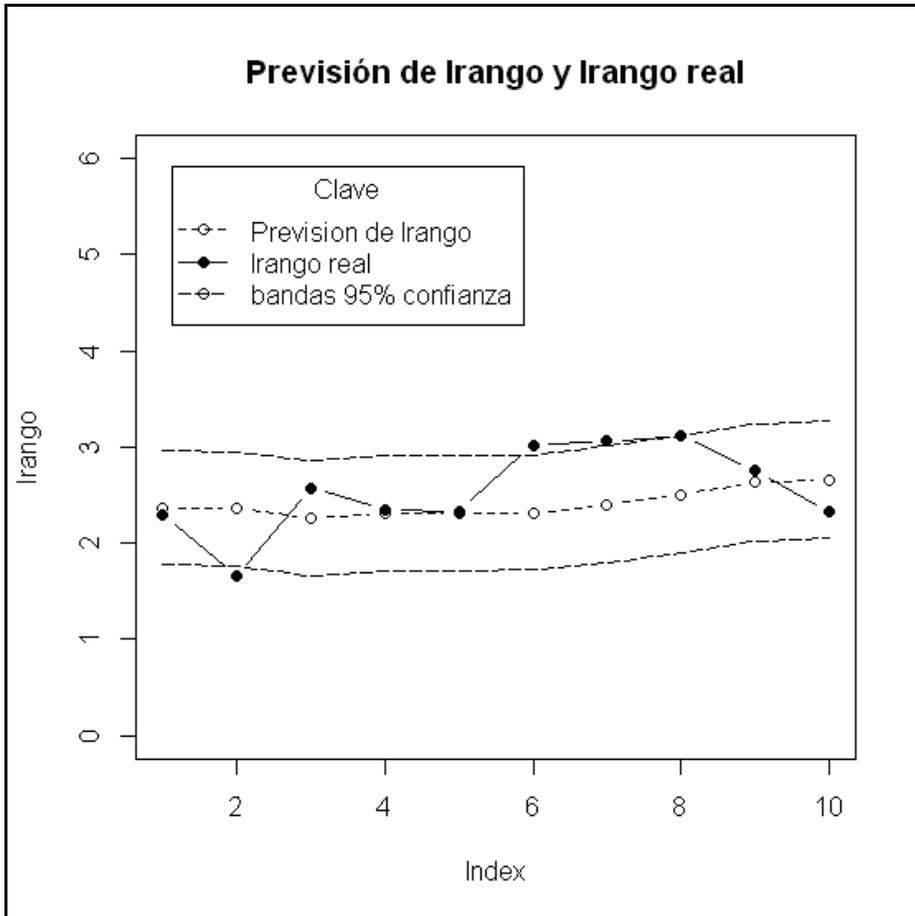


Figura 5.

La figura muestra que la banda del 95% de confianza, en torno al vector de predicciones del logaritmo del *rango*, la línea de círculos vacíos, abarca a la mayoría de los valores del logaritmo del *rango* real, los círculos rellenos en negro.

VIII. MODELACION CON EL PAQUETE *forecast*

Esta biblioteca o paquete de programas, que se ejecuta en el entorno R, fue diseñada con la idea de permitir la actualización de las predicciones de miles de series de manera automática, en los servicios de estudio, empresas de marketing, etc.

Lo primero a realizar es la incorporación del paquete a la memoria activa del ordenador, una vez descargado desde uno de los muchos lugares donde están almacenados los más de 5000 paquetes que configuran el entorno R, a la hora de redactar este trabajo. Hecho lo cual, basta con teclear:

```
library(forecast)
```

para tenerlo operativo.

Si la serie de los datos del logaritmo del rango, la denominamos y , la siguiente instrucción:

```
modelo2 = ets(y)
```

Nos devuelve las estimaciones siguientes:

```
modelo2
ETS(M,N,N)
Call:
ets(y = y)

Smoothing parameters:
  alpha = 0.1454

Initial states:
  l = 2.5433

sigma: 0.1472

      AIC      AICc      BIC
260.2677 260.3914 265.4781
```

Es decir, el programa ha identificado y estimado un modelo MNN (error multiplicativo, no tendencia y no estacionalidad), para la serie $y = lrango$. Cf. Fórmulas (5) y (6). El análisis de los residuos, no mostrado aquí, valida el modelo.

Para obtener la predicción del logaritmo del rango en el día 21 de octubre, basta con teclear:

```
forecast(modelo2,1,level=95)
```

que devuelve la predicción puntual, junto con la banda del 95% de confianza.

```
Point Forecast      Lo 95      Hi 95
101           2.370481 1.686471 3.054491
```

Si al igual que se hizo anteriormente, se repite el proceso para obtener las predicciones puntuales de los días 22, 23,..., hasta el 31 de octubre, tras aumentar la muestra, sucesivamente, con los valores reales del logaritmo del *rango* correspondientes a estos días, y resumimos la previsión puntual alcanzada, las bandas del 95% de confianza y, en la última columna, el logaritmo real del *rango* obtenemos la tabla 2:

obs	forecast	liminf	limsup	lrango
101	2.370481	1.686471	3.054491	2.295560
102	2.359606	1.681997	3.037215	1.652497
103	2.255866	1.597908	2.913825	2.574138
104	2.306279	1.634493	2.978064	2.354228
105	2.312793	1.642324	2.983263	2.331173
106	2.315275	1.647304	2.983247	3.016025
107	2.405618	1.701363	3.109874	3.060115
108	2.512935	1.768642	3.257227	3.113515
109	2.643582	1.854863	3.432301	2.758743
110	2.666876	1.874544	3.459208	2.320425

Tabla 2.

Más informativa que la tabla misma, es su representación en la figura 6.

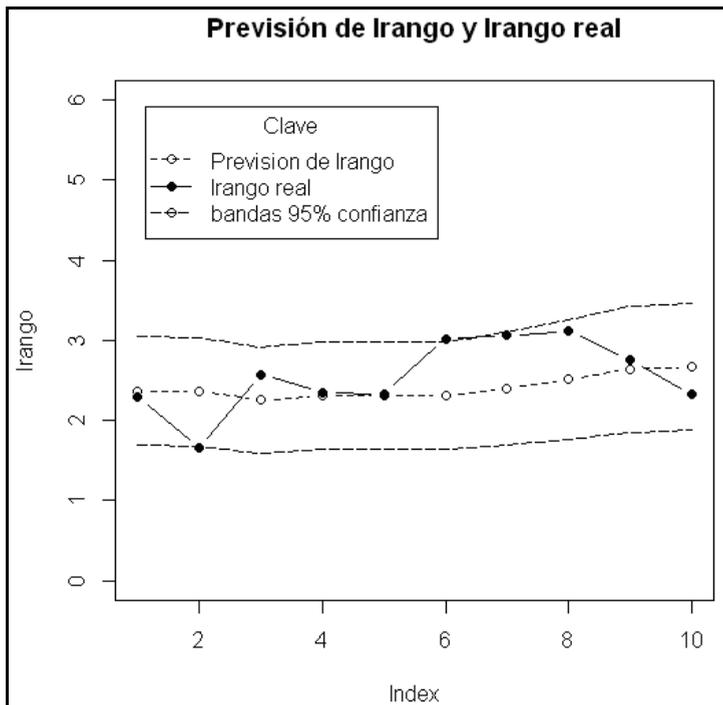


Figura 6.

Al igual que antes, la línea con los círculos en blanco representa las predicciones del logaritmo del *rango*, y los círculos en negro, los valores reales del logaritmo del *rango*. En esta ocasión, parece, visualmente, que la banda del 95% de confianza en torno a las predicciones del logaritmo del *rango* abarcan casi todos los valores del logaritmo del *rango* real.

Como una última aportación, indicamos los estadísticos más usuales para juzgar de la bondad de las predicciones: el error cuadrático medio, ECM, el error absoluto medio, EAM, el porcentaje del error absoluto medio PEAM, de estas predicciones:

Modelo 1	Modelo 2
ECM: 0.2031561	ECM: 0.2022954
EAM: 0.3594294	EAM: 0.358407
PEAM: 0.1493001	PEAM: 0.1488303

Las diferencias son irrelevantes, no permitiendo tomar una decisión sobre cuál de los dos sea el modelo preferible.

IX. CONCLUSIONES

Llegados aquí parecen claras las siguientes conclusiones.

Primera. Los datos del logaritmo del *rango* parecen gozar de las buenas propiedades estadísticas resaltadas por Alizadeh, Brandt y Newbold.

Segunda. El logaritmo del *rango* es susceptible de un tratamiento desde la óptica de los modelos dinámicos, en el marco de referencia de los *espacios de estados*. Desde este punto de vista, se ha podido establecer y validar un modelo para su comportamiento.

Tercera. El *quid* de la cuestión radica en si es o no posible considerar al *rango* de las cotizaciones como una aproximación a la volatilidad de los activos cotizados en Bolsa. Si se acepta esta posibilidad, la volatilidad permite un cierto tipo de modelización, y por ende, de control de su comportamiento.

Cuarta. En relación con la idea motivadora de este artículo, enunciada al comienzo, el logaritmo del *rango* de las cotizaciones puede ser un indicador útil sobre la situación general de la economía, en cuanto reflejada en la Bolsa.

X. BIBLIOGRAFÍA

- ALIZADEH, S., BRANDT, M.W., y DIEBOLD, F. X., “Range-Based Estimation of Stochastic Volatility Models”, en *Journal of Finance*, 57, 3 (2002) 1047-1091.
- ANDERSEN, T.G., y BOLLERSLEV, T., “Answering the skeptics: Yes, standard volatility models do provide accurate forecasts”, en *International Economic Review*, 39 (1998) 885-905.
- BECKERS, S., “Variance of security price returns based on high, low, and closing prices”, en *Journal of Business*, 56 (1983) 97-112.
- BOLSA DE MADRID, www.bolsamadrid.es
- EDWARDS, R. D., y MAGEE, J. *Technical Analysis of Stocks Trends*, Amacon, New York, 1997.
- GARMAN, M. K., y KLASS, M. J., “On the estimation of price volatility from historical data”, en *Journal of Business*, 53 (1980) 67-78.
- HYNDMAN, R.J., y KHANDAKAR, Y., “Automatic Time Series Forecasting: The forecast Package for R”, en *Journal of Statistical Software*, 27, 3 (2008) 1-22; <http://www.jstatsoft.org/>
- HYNDMAN, R.J.; KOEHLER, A.B.; ORD, J. K., y SNYDER, R. D., *Forecasting with Exponential Smoothing. The State Space Approach*, Springer Verlag, Berlin, 2008.
- MARTIN, V., HURN, S., y HARRIS, D., *Econometric Modelling with Time Series, Specification, Estimation and Testing*, Cambridge University Press, New York, 2013.
- PARKINSON, M., “The extreme value method for estimating the variance of the rate of return”, en *Journal of Business*, 53 (1980) 61-65.
- PETRIS, G., SONIA, P., y CAMPAGNOLI, P., *Dynamical Linear Models with R*, Springer, New York, 2009.
- PETRIS, G., “An R Package for Dynamical Linear Models”, en *Journal of Statistical Software*, 36, 12 (2010) 1-16.

- R CORE TEAM (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.

Paquetes estadísticos utilizados con R, instalables desde: cran.r-project.org/

- dlm, Giovanni Petris et al.
 - forecast, Robert Hyndman et al.
 - timeSeries (Rmetrics), <https://www.rmetrics.org>
 - PASWR, Probability and Statistics with R, Maria Dolores Ugarte et. Al.
- ROGERS, L.C.G., y SATCHEL, S. E., “Estimating variances from high, low, and closing prices”, en *Annals of Applied Probability*, 1 (1991) 504-512.
 - TSAY, R. S., *Analysis of Financial Time Series*, tercera edición, J. Wiley and Sons, Hoboken, New Jersey, 2010.
 - YANG, D., y ZHANG, Q., “Drift-independent volatility estimation based on high, low, open and close prices”, en *Journal of Business*, 73 (2000) 477-491.

