

# Máster Interuniversitario en Estadística e Investigación Operativa UPC-UB

**Título:** Comparación de la estimación de la volatilidad de una serie financiera, a través de la metodología del filtro de Kalman y el filtro de partículas en el contexto de los modelos de espacio-estado.

**Autor:** Erick Quispe Anastacio

**Directora:** Lesly Acosta

**Departamento:** Estadística e Investigación Operativa

**Universidad:** Universidad Politécnica de Cataluña

**Fecha:** Junio 2018



*Facultad de Matemáticas y Estadística  
Universidad Politécnica de Cataluña*

*Máster en Estadística e Investigación Operativa*

*Comparación de la estimación de la volatilidad de una serie financiera, a través de la metodología del filtro de Kalman y el filtro de partículas en el contexto de los modelos de espacio-estado.*

*Autor: Erick Quispe Anastacio*

*Directora: Lesly Acosta*

*Departamento de Estadística e Investigación Operativa*

**Dedicatoria:**

A mi madre Norma Quispe, por ser ejemplo de lucha y fortaleza, por enseñarme lo importante de tener fe y no rendirse nunca.

## **Agradecimientos**

Primeramente dar gracias a Dios por tener la gran oportunidad de seguir formándome. También quiero expresar mi gratitud a mi directora de tesis Lesly. Gracias por la paciencia y por todo el apoyo y soporte que me has dado en todo este tiempo, sobre todo por ayudarme a comprender y llevar a la práctica una metodología nueva para mí.

También quiero agradecer a todos los que de una forma directa e indirecta han ayudado a la culminación de este trabajo y a todos mis amigos y familiares que han estado ahí en mis momentos duros.

A todos vosotros, muchas gracias.

---

## Resumen

Mediante un ejercicio comparativo, el objetivo del presente trabajo, es estimar la volatilidad de una serie financiera a través de dos algoritmos; el filtro de Kalman y el filtro de partículas, abordados en el contexto de los modelos de espacio de estado, y evaluar su capacidad para realizar estimaciones los más eficientes y fiables posibles.

La serie financiera que se utiliza, es la serie del índice bursátil “IBEX 35”, que es el índice de referencia de la bolsa española. La comparación de la estimación obtenida con ambos algoritmos, se realizará a través de una proxy de la volatilidad, que es el logaritmo de los rendimientos al cuadrado  $\log(r_t^2)$ ; de esta forma se analiza que método aproxima de mejor manera el comportamiento de la volatilidad de la serie de estudio.

La implementación de los algoritmos se realizará utilizando software estadístico R, versión 3.3.1., y algunos de sus paquetes ya implementados como “nimble” y “astsa”. Asimismo, de forma complementaria, se pretende explorar como la implementación en “nimble” del filtro de partículas bajo estudio difiere o no, con la implementación presentada en la tesis doctoral de Acosta L. (2013) “Particle Filtering Estimation for Linear and Nonlinear State-Space Models”.

**Palabras claves:** volatilidad, filtro de Kalman, máxima verosimilitud, filtro de partículas, modelo de espacio de estado, función de densidad posterior,

---

## Índice de contenido

<b>1.- Introducción</b> .....	1
1.1.- Objetivo .....	1
1.2.- Estimación de la volatilidad: un breve recorrido por la literatura econométrica .....	2
<b>2.- Metodología estadística</b> .....	7
2.1.- Modelos dinámicos representados en forma de espacio-estado.....	7
2.2 Formulación general de un modelo en la forma de espacio-estado.....	8
2.3.- Estimación vía filtro de Kalman .....	10
2.4.-Estimación vía Filtro de partículas .....	16
<b>3.- Modelo del modelo de volatilidad estocástica: Predicción y Filtrado</b> .....	<b>29</b>
3.1.- Análisis de los datos. ....	29
3.2.- Estimación de la volatilidad y parámetros vía filtro de Kalman y máxima verosimilitud .....	33
3.3.- Estimación de la volatilidad y parámetros vía filtro de partículas de Liu y West.....	41
<b>4. Conclusiones y discusión</b> .....	51
<b>Bibliografía</b> .....	53
<b>Anexos</b> .....	56
A. Tablas .....	56
B.- Figuras:.....	57
C.- Codigo de implementacio Filtro de Kalman y filtro de partículas .....	58

## Lista de Figuras

Figura 1. Evolución de la cotización del IBEX 35 del 01/01/2002 al 31/12/2016.....	30
Figura 2. Rendimientos de la cotización del IBEX35 .....	30
Figura 3. Gráficos descriptivos: Histograma, QQ-plot y ACF.....	32
Figura 4. Comparación de la función de densidad del $\log(\chi^2)$ y La función de densidad estimada a través de la mixtura de normales con $\pi=0.5$ .....	39
Figura 5. Volatilidad estimada $\sigma$ mediante FK + Newton Raphson.....	39
Figura 6. Gráfico de $\log(rt^2)$ y volatilidad estimada.....	40
Figura 7: Gráfico de rendimientos del IBEX35 ( $rt$ ) y la volatilidad estimada .....	40
Figura 8. Histograma y Q-Q plot de residuos: KF + Newton Raphson .....	41
Figura 9. Comparación de la estimación de la volatilidad con la implementación de Acosta L. (2013) y nimble, $N_p=5000$ , $\delta=0,95$ .....	46
Figura 10. Rendimientos del IBEX 35 y volatilidad estimada con la implementación de Acosta L. (2013) y nimble, $N_p=5000$ , $\delta=0,95$ .....	46
Figura 11. Histograma y QQ-plot de residuos del modelo SARV(1) con la implementación de Acosta L. y nimble. $N_p=5000$ , $\delta=0,95$ .....	47
Figura 12. Comparación de la volatilidad estimada con el filtro de Kalman y el filtro de partículas.....	49
Figura 13. Rendimientos del IBEX35 versus volatilidad estimada vía filtro de Kalman y el filtro de partículas de Liu y West. ....	50
Figura 1B. Evolución de los parámetros del filtro de partículas de Liu y West (2001) $N_p=5000$ , $\delta=0,95$ .....	57
Figura 2B. Densidad de probabilidad posterior de los parámetros del modelo con la implementación de Acosta L. $N_p=5000$ , $\delta=0,95$ .....	57

## Lista de Tablas

Tabla 1. Filtro de Partículas de Liu y West. ....	28
Tabla 2. Descriptivos básicos de la Serie de Rendimientos del IBEX 35 .....	31
Tabla 3. Resultados de la estimación de los parámetros .....	38
Tabla 4. Resumen de las distribuciones a priori y valores iniciales .....	43
Tabla 5. Resultados de la estimación del vector de los parámetros: Np=5000 y $\delta=0,95$ .....	45
Tabla 6. Comparación de las estimaciones obtenidas vía filtro de Kalman + Newton Raphson y el filtro de partículas de Liu y West: Np=5000, $\delta =0.95$ .....	49
Tabla 1A. Test de hipótesis sobre la media ( $H_0:\mu=0$ ) y 2 tests de normalidad de la serie de rendimientos de la cotización del índice bursátil IBEX35 .....	56
Tabla 2A. Test de hipótesis de correlación de Ljung-Box .....	56



## 1.- Introducción

### 1.1.- Objetivo

En el campo de las finanzas, hay muchas aplicaciones donde el estudio de la volatilidad es importante. Por ejemplo, es un factor clave en la valoración de activos financieros (acciones, futuros), en la gestión de carteras así como en el cálculo del VaR (Value at Risk). Además, tiene un efecto directo en las decisiones de inversión, ya que, los agentes económicos al realizar planes de inversión basan sus decisiones en el rendimiento esperado y el riesgo inherente a dicha inversión.

En series financieras, lo único que se observa es la variación en el precio del activo pero la volatilidad es un proceso latente. Esto hace que se realice una estimación de la misma, motivo que ha provocado una creciente aparición de modelos en la literatura econométrica que tratan de estimar dicha volatilidad. En términos financieros, la volatilidad, no es más que un indicador que mide el grado de variabilidad que se produce en el precio de los activos (acciones, bonos, opciones etc.), causada por la propia dinámica de los mercados financieros.

El objetivo del presente trabajo, es comparar dos métodos de estimación de la volatilidad, el filtro de Kalman y el filtro de Partículas, abordados en el contexto de los modelos de espacio de estado, y evaluar su capacidad para realizar estimaciones los más eficientes y fiables posibles.

Para poder hacer la comparativa entre filtros, se estimará la volatilidad subyacente en la serie de tiempo “IBEX 35”, el índice bursátil de referencia de la bolsa española. Una vez realizada dicha estimación, también se comparará con una proxy de la volatilidad que es el logaritmo de los rendimientos al cuadrado  $\log(r_t^2)$ ; de esta forma se podrá ver que método aproxima de mejor forma el comportamiento de la volatilidad de la serie de estudio.

La implementación de los algoritmos (Filtro de Kalman y Filtro de Partículas) se realizará utilizando el lenguaje de programación R y algunos de sus paquetes ya implementados como “Nimble<sup>1</sup>” y “astsa<sup>2</sup>”. Asimismo, de forma complementaria, se pretende explorar como la implementación en “Nimble” del filtro de partículas bajo

---

<sup>1</sup> <https://r-nimble.org/>

<sup>2</sup> <https://cran.r-project.org/web/packages/astsa/astsa.pdf>

estudio difiere o no, con la implementación presentada en la tesis doctoral “Particle Filtering Estimation for Linear and Nonlinear State-Space Models” (Acosta L. 2013).

## **1.2- Estimación de la volatilidad: un breve recorrido por la literatura econométrica**

En los últimos años se ha multiplicado el interés, e intensificado el esfuerzo por estimar la volatilidad que está presente en la serie de precios de un activo financiero. Se han publicado numerosos estudios sobre el tema, poniendo de relieve diversos enfoques para abordar el problema, como por ejemplo, los modelos heterocedásticos condicionales determinísticos tipo G(ARCH) (Engle 1982; Bollerslev et al. 1992), o los modelos de volatilidad estocástica, propuesto por Taylor (1986, 1994) y sus variantes (Jacquier et al. 1994).

Antes de la aparición de los modelo tipo G(ARCH) y los modelos de volatilidad estocástica, se utilizaba como estimador de la volatilidad “la volatilidad histórica”, que no es más que la desviación típica de los rendimientos al cuadrado. Uno de los problemas de fondo de este enfoque, es que esta desviación típica es un estimador sesgado de la volatilidad, fuertemente afectada por la presencia de valores extremos, bajo dicho enfoque, el patrón de la volatilidad depende del número de observaciones usadas para su cálculo, y además, se le da la misma ponderación a todos los rendimientos por igual.

El presente trabajo está enfocado en los modelos de volatilidad estocástica, aunque también se describe brevemente el funcionamiento de los modelos tipo G(ARCH)<sup>3</sup>, así como las diferencias y similitudes entre ambos enfoques.

Las series temporales de naturaleza financiera, al tener unas características típicas conocidas como hechos estilizados, hace imposible modelar la volatilidad utilizando enfoques clásicos, como los modelos ARIMA<sup>4</sup>, entre otros, debido a que estos modelos asumen supuestos que no tienden a cumplirse en este tipo de series, como por ejemplo, el de varianza constante etc. Por tanto es necesario utilizar métodos alternativos para poder estimar la volatilidad latente en una serie financiera.

En los siguientes párrafos, se detallan los diferentes hechos estilizados que suelen encontrarse en las series financieras, como por ejemplo, en la serie de rendimientos

---

<sup>3</sup> generalized autoregressive conditional heteroskedasticity

<sup>4</sup> Autoregressive integrated moving average

tratada en este trabajo. Estos hechos estilizados han sido documentados en diversos trabajos, como por ejemplo, en el libro de Ruey S. Tsay (2014).

Uno de los hechos estilizados que se observa frecuentemente en las series financieras, es la presencia de colas pesadas, implicando que la hipótesis de normalidad no sería del todo correcta, debido a que presentan más bien una distribución leptocúrtica. Esto explica por qué muchos investigadores han utilizado distribuciones como la *t* de student, para poder captar esa característica de colas pesadas (Mandelbrot 1963; Fama 1965).

Otro hecho estilizado que presentan las series financieras, son los agrupamiento de volatilidad (volatility clustering), es decir, altos rendimientos observados en un periodo (de cualquier signo), van seguidos de altos rendimientos en periodos siguientes, y bajos rendimientos observados en un periodo (de cualquier signo), van seguidos de bajos rendimientos en periodos siguientes (Fama 1965). Este fenómeno fue detectado por Mandelbrot en su estudio “The variation of certain speculative prices” (1963). Actualmente los modelos ARCH (Engle 1982) y sus variantes, así como los modelos de volatilidad estocástica (Ghysels, Harvey y Renault, 1996), son capaces de captar este agrupamiento de volatilidad.

El efecto leverage (acuñado por Black, 1976) es otro hecho estilizado que se halla en las series financieras, generalmente es más común detectarlo en series de precio de acciones, y es menos frecuente en series de tipos de cambio o de tasas de interés. El efecto leverage cuantifica el efecto de las perturbaciones positivas o negativas sobre la volatilidad, que en la mayor parte de los casos es asimétrico. Es decir, que la volatilidad del precio del activo reacciona de forma diferente cuando se produce una bajada en el precio, que cuando se produce una subida de la misma magnitud. Esto podría deberse al efecto asimétrico que tienen las perturbaciones positivas o negativas. (Andersen et al. 2001).

Por último, otro hecho que se observa, es la persistencia de la volatilidad a lo largo del tiempo. Se observa un comportamiento “suave” en la volatilidad; es decir, no se producen cambios abruptos de un periodo a otro. Esto es muy común en series de precios con datos diarios o intradiarios. Dicha persistencia se puede explicar a través de la presencia de raíces muy próximas a la unidad, en el proceso de la varianza condicional. Muchos autores han mostrado que las series de rendimientos presentan un patrón de volatilidad altamente persistente (Jacquier et al. 1994; Acosta L. 2013).

Como se menciona anteriormente, las series financieras presentan características típicas que no hacen posible modelar la volatilidad a través de métodos tradicionales como los modelos ARIMA, ya que no existe ninguna transformación de Box-Cox que logre estabilizar la varianza cambiante, y por tanto no se cumple el supuesto de estacionariedad que requieren este tipo de modelos.

Nótese además, que aunque la varianza marginal - subyacente en una serie de rendimientos - es constante, su varianza condicional (que depende de valores pasados previos) va cambiando con el tiempo, observándose periodos con alta volatilidad y periodos con baja volatilidad.

La literatura econométrica ha desarrollado diferentes modelos no lineales para poder estimar esta volatilidad cambiante. En la actualidad podemos distinguir dos grupos de modelos, ampliamente usados: por un lado los G(ARCH) y sus respectivas variantes, y por otro lado, los modelos de volatilidad estocástica SARV<sup>5</sup>. El modelo básico tipo GARCH y de volatilidad estocástica es el GARCH(1,1) y el SARV(1), respectivamente

Ambos modelos tienen como objetivo capturar las características típicas de las series financieras, como el exceso de curtosis, el agrupamiento de volatilidad y la persistencia de la volatilidad, así como la correlación de los residuos al cuadrado de la serie.

### **Modelo GARCH (1,1)**

Los modelos G(ARCH) se caracterizan por el hecho de que la varianza condicional en el instante t, es una función que depende de los valores anteriores de la serie de rendimientos, y de los valores de su propia volatilidad pasada. Específicamente, un modelo GARCH(1,1) asume que la varianza condicional ( $\sigma_t^2$ ), no solo depende de los pasados de la serie ( $r_{t-1}^2$ ), sino que también depende de sus propios valores pasados ( $\sigma_{t-1}^2$ ), definiendo la varianza condicional de la siguiente forma:

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

donde  $\alpha_1$  mide la reacción de la volatilidad condicional a los shocks del mercado y  $\beta_1$  es el parámetro que mide la persistencia de la volatilidad.

Se debe cumplir que  $\alpha_0 > 1$ ,  $\alpha_1 \geq 0$ ,  $\beta_1 \geq 0$  para garantizar la positividad de  $\sigma_t^2$ , además de la restricción  $\alpha_1 + \beta_1 < 1$ , para garantizar la estacionariedad.

---

<sup>5</sup> Stochastic Autorregresive Volatility

Como contrapartida a los modelos condicionales tipo G(ARCH), están los modelos de volatilidad estocástica (SV) introducidos por Taylor en 1986.

### **Modelos SARV(1)**

En esta clase de modelos, la varianza condicional ( $\sigma_t^2$ ) no depende directamente de los valores pasados de la propia serie, sino que depende de un proceso no observable, que es la propia volatilidad pasada y de una componente de ruido. Este trabajo se centra en la aplicación del modelo básico de volatilidad estocástica conocido como SARV(1). La formulación de la varianza condicional bajo el modelo SARV(1) es la siguiente:

$$x_t = \mu + \phi_1(x_{t-1} - \mu) + z_t$$

donde  $x_t = \log(\sigma_t^2)$ . Operando algebraicamente la ecuación anterior, se puede reescribir de la siguiente forma:

$$x_t = \phi_0 + \phi_1 x_{t-1} + z_t, \quad z_t \sim N(0, \sigma_z)$$

donde  $\phi_0 = \mu(1 - \phi_1)$ , es una constante que mide la volatilidad media,  $\phi_1$  es el parámetro de persistencia de la volatilidad y  $z_t$ , es el ruido del estado que se distribuye según una normal  $\sim(0, \sigma_z)$ .

En ambos modelos (GARCH(1,1), SARV(1)), la expresión básica para los valores observados es  $r_t = \sigma_* \cdot \sigma_t \cdot \epsilon_t$ , donde  $\sigma_*$  es la variabilidad marginal (opcional a modelizar),  $\sigma_t$  es una variable latente (volatilidad) y  $\epsilon_t$  se asume usualmente normal estándar,

Se observa que el modelo para los rendimientos  $r_t$  es no lineal, pero algunos autores los linearizan, por ejemplo, aplicando la transformación logarítmica de los rendimientos al cuadrado,  $\log(r_t^2)$  (Kim et al.1998).

### **Modelo SARV(1) versus GARCH(1,1)**

En ambos modelos, se considera que la varianza condicional depende de sus valores pasados, pero la diferencia esencial entre ambos, es que los tipos GARCH plantean una evolución determinista de la volatilidad al no introducir una componente de ruido, mientras que los SARV si lo hacen.

Ambos modelos son capaces de captar el agrupamiento de la volatilidad, pero en menor medida captan el efecto asimétrico existente en los rendimientos; sobre todo los modelos GARCH que son deterministas y que ante la presencia de valores extremos,

arrojan una estimación de la volatilidad más distorsionada. Los modelos SARV, en cambio suelen ser más robustos ante la presencia de valores atípicos.

A diferencia de los modelos GARCH, los modelos SV son más difíciles de estimar, debido que a veces se desconoce la función de verosimilitud y/o no puede construirse fácilmente. El proceso de estimación se complica más cuando se estiman simultáneamente los parámetros del modelo tratado junto a la variable latente (volatilidad).

En los apartados siguientes se explica de forma detallada la metodología de espacio-estado, y su uso en la estimación de la volatilidad y parámetros asociados a través del filtro de Kalman y el filtro de partículas.

## **2.- Metodología estadística**

El presente trabajo se centra en el modelo básico de volatilidad estocástica autoregresiva de primer orden SARV(1), en el contexto de los modelos de espacio-estado, utilizando la metodología del filtro de Kalman y el filtro de partículas en el proceso de estimación. Concretamente, se usa la variante del filtro de partículas de Liu y West (2001) que estima tanto el estado como los parámetros del modelo SARV(1).

### **2.1.- Modelos dinámicos representados en forma de espacio-estado**

Un modelo dinámico representado en la forma de espacio-estado, es un sistema que busca captar la dinámica de una variable observada, en función de una variable o conjuntos de variables no observadas, denominadas variables latentes o de estado, que son las que determinan dicho sistema dinámico.

El objetivo que persiguen estos modelos, es obtener estimaciones fiables de los componentes no observables a través de observaciones contaminadas con ruido, que se han generado en el mismo sistema. Estos componentes no observables poseen información relevante referente al sistema en un punto concreto del tiempo, por lo que, su estimación permite predecir la dinámica del comportamiento futuro a través del comportamiento que el sistema ha tenido en el pasado (Acosta L. 2013)

Para poder describir y realizar inferencia de un determinado sistema dinámico formulado en forma de espacio-estado, es necesario describir dos ecuaciones importantes, la ecuación de estado, que es la que describe la evolución del estado en el tiempo, y la ecuación de observación, que es la que relaciona las observaciones o medidas contaminadas con ruido con el estado. (Ristic et al. 2004)

Los modelos dinámicos representados en forma de espacio-estado, son frecuentemente utilizados en la modelización de componentes no observables, extendiéndose su aplicación a campos como las finanzas, un ejemplo de ello, es su utilización en la estimación de la volatilidad de series financieras (Márquez, 2002), que es un elemento no observable en los mercados financieros, pero que representa un factor clave en la valoración de activos tales como opciones, futuros etc.

Además, en el campo de la economía, estos modelos han sido utilizados por muchos economistas para estimar la inflación subyacente<sup>6</sup>, y explicar porque cada vez es más difícil de preverla etc. (Stock y Watson 2007; Pellegrini 2009).

Lo que hace aún más interesante a esta clase de modelos, es que aparte de permitir la estimación de variables no observables de un sistema dinámico, ofrece la posibilidad de estimar parámetros que pueden ser constantes o no (Shumway y Stoffer 2011).

La estimación de esta clase de modelos, se realiza a través del enfoque de los filtros recursivos, consistiendo básicamente en dos pasos, el primer paso es la predicción y el segundo es el filtraje o actualización del sistema (Ristic et al. 2004), estos pasos serán explicado con más detalle en párrafos posteriores.

## 2.2 Formulación general de un modelo en la forma de espacio-estado

Dado el conjunto de los valores observados  $Y_s = \{y_1, y_2, \dots, y_s\}$  hasta el momento discreto  $s$ , se puede formular las ecuaciones de un modelo de espacio-estado en su forma general (Shumway 1988; Tanizaki 1996; Acosta L. 2013).

$$\text{Ecuación de transición:} \quad x_t = f(x_{t-1}, z_t) \quad (2.1)$$

$$\text{Ecuación de observación:} \quad y_t = g(x_t, v_t) \quad (2.2)$$

Las ecuaciones de transición (2.1) y de observación (2.2) pueden ser re-expresadas en función de su distribución condicional (Kitagawa y Sato 2001):

$$x_t | x_{t-1} \sim p_z(\cdot | x_{t-1})$$

$$y_t | x_t \sim p_v(\cdot | x_t)$$

donde  $x_t \in R^{n_x}$ , es el vector de estado que no es directamente observable,  $z_t \in R^{n_z}$  representa el error del proceso, y  $v_t \in R^{n_v}$  representa el error de la observación.

Los procesos de ruido  $z_t$  y  $v_t$  son ruido blanco, no necesariamente gaussianos; y  $p_z$  y  $p_v$ , son las funciones de densidad de probabilidad del error del proceso  $z_t$  y del error de observación  $v_t$ , respectivamente. Además, las formas funcionales  $f(x_t, z_t)$  y de observación  $g(x_t, v_t)$ , se suponen conocidas, y no necesariamente lineales.

Además, en la formulación anterior, se supone que la función de densidad de probabilidad del vector de estado inicial  $x_0$ ,  $p(x_0 | y_0) = p(x_0)$ , está disponible;  $y_0$  se

---

<sup>6</sup> Indicador en el que no se toma en cuenta la variación que han sufrido los precios energéticos ni los productos no elaborados.



denomina conjunto de no medida, es decir, en este punto aún no se ha introducido ninguna observación en el sistema.

Nótese que las formas funcionales  $f$  y  $g$ , así como las funciones de densidad de probabilidad podrían guardar dependencia con el vector de parámetros  $\theta$ .

La formulación de los modelos de espacio-estado se basa en dos supuestos. i) el estado  $x_t$  sigue un proceso markoviano de primer orden; es decir, el estado en el momento  $t$  solo depende de su estado en el momento anterior ( $t-1$ ) y ii) que dado el estado, las observaciones ( $y_t$ ) son (condicionalmente) independientes.

### **Ecuaciones generales de Predicción y Filtrado**

Una vez especificado el modelo de espacio-estado, el objetivo es estimar el estado no observado ( $x_t$ ), a través de todas las observaciones disponibles  $Y_s$  hasta el periodo  $s$ ,  $Y_s = (y_1, y_2, \dots, y_s)$ .

Nótese que si se quiere estimar  $x_t$ , dado el conjunto de observaciones  $Y_s$ , se puede tener 3 tipos de estimaciones  $x_{t|s}$  del estado: predicción, cuando  $s < t$ ; filtrado, si  $s = t$ , y el suavizado cuando  $s > t$ . Este trabajo se centra en la predicción y filtrado del estado.

Utilizando los 2 supuestos de un modelos de espacio-estado, y desde una perspectiva bayesiana (utilizando la regla de Bayes), se obtienen las siguientes expresiones generales de predicción y filtrado para el modelo especificado en las ecuaciones (2.1) y (2.2)

### **Expresión general de la función de densidad de predicción**

Suponiendo que la función de densidad de probabilidad del filtrado,  $p(x_{t-1}|Y_{t-1})$ , está disponible en el tiempo  $t-1$ , la expresión general de la función de densidad de predicción en el periodo siguiente viene dada por:

$$\begin{aligned} p(x_t|Y_{t-1}) &= \int p(x_t|x_{t-1}, Y_{t-1})p(x_{t-1}|Y_{t-1})dx_{t-1} \\ &= \int p(x_t|x_{t-1})p(x_{t-1}|Y_{t-1})dx_{t-1} \end{aligned} \quad (2.3)$$

En la ecuación (2.3), se ha utilizado la propiedad Markoviana de primer orden, por lo que  $p(x_t|x_{t-1}, Y_{t-1}) = p(x_t|x_{t-1})$ . Esta propiedad permite obtener la función de densidad del estado en el periodo  $t$  a partir del estado en el periodo en  $t-1$ .

## Expresión general de la Función de densidad de probabilidad de filtrado

Cuando la observación  $y_t$  está disponible en el periodo  $t$ , se lleva a cabo la actualización o filtrado del sistema. Es decir, se actualiza la predicción de la función de densidad de probabilidad haciendo uso de la regla de Bayes, por lo que, la predicción en el periodo posterior viene dada por la siguiente expresión:

$$\begin{aligned} p(x_t|Y_t) &= p(x_t|y_t, Y_{t-1}) \\ &= \frac{p(y_t|x_t, Y_{t-1})p(x_t|Y_{t-1})}{p(y_t|Y_{t-1})} \\ &= \frac{p(y_t|x_t)p(x_t|Y_{t-1})}{p(y_t|Y_{t-1})} \\ &\propto p(y_t|x_t)p(x_t|Y_{t-1}) \end{aligned} \quad (2.4)$$

que depende de la función de verosimilitud  $p(y_t|x_t)$  y de la densidad de predicción  $p(x_t|Y_{t-1})$  (Ristic et al. 2004; Acosta L. 2013). Nótese que  $p(y_t|Y_{t-1}) = \int p(y_t|x_t)p(x_t|Y_{t-1})dx_t$  es la constante de normalización,

## Expresión recursiva del filtrado

Si se sustituye la ecuación (2.3) en la ecuación (2.4), obtenemos la expresión recursiva del filtrado de la función de la densidad de filtrado,  $p(x_t|Y_t)$ .

$$p(x_t|Y_t) \propto p(y_t|x_t)p(x_t|Y_{t-1}) = p(y_t|x_t) \int p(x_t|x_{t-1})p(x_{t-1}|Y_{t-1})dx_{t-1} \quad (2.5)$$

La expresión (2.5), involucra a la densidad de probabilidad  $p(y_t|x_t)$  en la ecuación de observación (2.2), a la densidad de transición  $p(x_t|x_{t-1})$  en (2.1) y a la densidad del filtrado en el periodo anterior  $p(x_{t-1}|Y_{t-1})$ . Nótese que una vez conocida  $p(x_t|Y_t)$ , se puede decir que el problema de filtrado óptimo se ha solucionado (Acosta L. 2013).

## 2.3.- Estimación vía filtro de Kalman

El filtro de Kalman es un algoritmo diseñado por Rudolf E. Kalman y que plasmó en su estudio “A new approach to linear filtering and prediction problems” en 1960. Este filtro ha sido ampliamente estudiado y aplicado en diferentes campos, entre ellos la economía y las finanzas, en el que se ha utilizado para realizar estimaciones de modelos de series de tiempo estructurales y de estimación de volatilidad estocástica entre otros casos, como estimación de la inflación o del crecimiento económico (Márquez 2002; Muñoz, Márquez et al. 2007).

En la actualidad, este filtro es uno de los principales métodos que se utilizan a la hora de realizar estimaciones de sistemas dinámicos, representados en forma de espacio-estado, debido a que es un método recursivo, estable y de sencilla implementación. Es sabido que el filtro de Kalman es un método óptimo en caso de modelos de espacio-estado lineales y Gaussianos.

La estructura de este filtro se fundamenta en las propiedades de las variables aleatorias normales condicionales (Le Roux, 2003). Dada la media y la matriz de covarianzas del vector de estado inicial con distribución normal, el filtro de Kalman propaga en el tiempo los dos primeros momentos de una distribución normal. Por lo que a cada paso, la densidad posterior también será normal, y por tanto caracterizada por su media y su varianza (Ristic et al. 2004; Acosta 2013).

Si las funciones  $f(x_t, z_t)$  y  $g(x_t, v_t)$ , son lineales y los errores son Gaussianos, supuestos sobre los que descansa el filtro de Kalman, se puede reescribir las ecuaciones (2.1) y (2.2) del sistema dinámico a través de un modelo de espacio-estado Markoviano, Lineal y Gaussiano de la siguiente forma:

$$\textbf{Ecuación de transición:} \quad x_t = \Phi x_{t-1} + z_t, \quad z_t \sim WN(0, Q) \quad (2.6)$$

donde,  $x_t$ , es el vector de estado (px1) en el momento t que no es directamente observable,  $\Phi$  es una matriz (pxp), denominada matriz de transición, que relaciona el estado en t-1 con el estado en t, y  $z_t$  es el vector (px1) que contiene las perturbaciones del estado distribuidas normalmente con vector de media cero y matriz covarianza Q.

$$\textbf{Ecuación de observación:} \quad y_t = A_t x_t + v_t, \quad v_t \sim WN(0, R) \quad (2.7)$$

donde,  $y_t$ , es el vector de observación (qx1), y está relacionado con el vector de estado  $x_t$  en t.  $A_t$  es una matriz (qxp) que relaciona el estado en t con la observación en t, y  $v_t$  es el vector (qx1) de perturbaciones de las observaciones, que se distribuyen según una normal, con vector de media cero y matriz de covarianza (qxq) R.

Las perturbaciones de la ecuación de estado y de observación ( $z_t, v_t$ ) son ruido blanco Gaussianos independientes. Los parámetros ( $\Phi, Q, R, A_t$ ) podrían variar en el tiempo, pero se asume que son constantes.

Las ecuaciones (2.6) y (2.7) definen un modelo de espacio-estado Markoviano, Lineal y Gaussiano. Por tanto, dada la media,  $\mu_0 = x_{0|0}$ , y la varianza,  $\Sigma_0 = P_{0|0}$ , y que las perturbaciones no están correlacionadas entre ellas  $E(z_t \cdot v_s) = 0 \forall t \neq s$ , y ni con el

estado inicial  $E(z_t \cdot x_0) = 0 \forall t, E(v_t \cdot x_0) = 0 \forall t$ , se puede definir tres tipos de estimadores del estado y sus covarianzas

$$x_{t|s} = E(x_t|Y_s) = E(y_s, y_{s-1}, \dots, y_1)$$

$$P_{t|s} = cov(x_t|Y_s) = cov(x_t|y_s, y_{s-1}, \dots, y_1)$$

Como se comentó anteriormente, desde la perspectiva práctica nos encontramos con las siguientes situaciones:

1.- Predicción en el caso de que  $s < t$ ; en este trabajo usamos  $s=t-1$ :

$$x_{t|t-1} = E(x_t|y_{t-1}, y_{t-2}, \dots, y_1)$$

$$P_{t|t-1} = var(x_t|y_{t-1}, y_{t-2}, \dots, y_1)$$

2.- Filtrado en el caso de que  $s=t$ :

$$x_{t|t} = E(x_t|y_t, y_{t-1}, y_{t-2}, \dots, y_1)$$

$$P_{t|t} = var(x_t|y_t, y_{t-1}, y_{t-2}, \dots, y_1)$$

3.- Suavizado en el caso de que  $s > t$ , usualmente  $s=N$ , donde  $N$  es la longitud de la serie completa:

$$x_{t|N} = E(x_t|y_N, y_{N-1}, y_{N-2}, y_1)$$

$$P_{t|N} = var(x_t|y_N, y_{N-1}, y_{N-2}, \dots, y_1)$$

Nótese que el estimador de suavizado, al incorporar mayor información es mucho más preciso, pero requiere esperar a tener disponible toda la serie. Como se ha mencionado anteriormente, este trabajo se centra solo en la predicción y el filtrado.

### **Derivación del funcionamiento del algoritmo del filtro de Kalman**

El algoritmo del filtro trabaja a través de la predicción y la actualización del sistema, siendo el funcionamiento el siguiente:

1.- Se formula el modelo en forma de espacio-estado, a través de las ecuaciones (2.6) y (2.7) respectivamente:

$$x_t = \Phi x_{t-1} + z_t, \quad z_t \sim WN(0, Q)$$

$$y_t = A_t x_t + v_t, \quad v_t \sim WN(0, R)$$

2.- Se da los parámetros iniciales del estado  $x_{0|0} = \mu_0, P_{0|0} = \Sigma_0$ , equivalentes a los valores del filtrado en el momento  $t=0$ .

3.- Se calcula la estimación puntual como la varianza-covarianza de cada estimador. El filtro de Kalman consta de las siguientes ecuaciones:

**Ecuaciones de predicción para el estado:**

$$x_{t|t-1} = \Phi x_{t-1|t-1} \quad (2.8)$$

$$P_{t|t-1} = \Phi P_{t-1|t-1} \Phi' + Q \quad (2.9)$$

**Ecuaciones de predicción para la observación:**

$$y_{t|t-1} = A_t x_{t|t-1} \quad (2.10)$$

$$F_{t|t-1} = A_t P_{t|t-1} A_t' + R \quad (2.11)$$

**Ecuaciones de las innovaciones y la ganancia de Kalman:**

$$e_t = y_t - A_t x_{t|t-1} \quad (2.12)$$

$$K_t = P_{t|t-1} A_t' (A_t P_{t|t-1} A_t' + R)^{-1} \quad (2.13)$$

**Ecuaciones de filtro o actualización del estado:**

$$x_{t|t} = x_{t|t-1} + K_t e_t \quad (2.14)$$

$$P_{t|t} = (I - K_t A_t) P_{t|t-1} \quad (2.15)$$

Las expresiones (2.8), (2.9), (2.10) y (2.11), son las ecuaciones de predicción del estado y la observación dado  $Y_{t-1}$ . Para cada estimador se obtiene también la respectiva varianza-covarianza.

Con la nueva observación ( $y_t$ ) disponible, calculamos la innovación – que es el error de la predicción de la observación– y la ganancia de Kalman, necesarios para la actualización del sistema. La innovación y la ganancia de Kalman se expresan a través de las ecuaciones (2.12) y (2.13) respectivamente.

Finalmente, con las ecuaciones (2.14) y (2.15), se calcula el valor de filtrado o actualización del estado y la varianza del mismo.

En este trabajo estimamos tanto el valor de filtrado como los parámetros involucrados en el modelo formulado en espacio-estado. Un primer enfoque para realizar dicha estimación consiste en el uso combinado del filtro de Kalman y el método de máxima verosimilitud.

## Estimación Máxima Verosimilitud

Para estimar el vector de parámetros  $\theta = \{\mu_0, \Sigma_0, \Phi, Q, R\}$  de las ecuaciones (2.6) y (2.7), del modelo dinámico representado en la forma de espacio-estado, se utilizará el método de máxima verosimilitud, bajo el supuesto de normalidad del estado inicial  $x_0 \sim N(\mu_0, \Sigma_0)$ , de los errores del estado  $z_t$  y de las observaciones  $v_t$ , y el supuesto de no correlación de los errores con el estado inicial. (Shumway y Stoffer 2010).

El vector de parámetros  $\theta$ , contiene la media y las covarianzas iniciales  $(\mu_0, \Sigma_0)$ , la matriz de transición del estado ( $\Phi$ ) y las matrices de covarianzas tanto del estado como de la observación ( $Q, R$ ).

La verosimilitud se obtiene a través de las innovaciones  $e_1, e_2, \dots, e_t$ , especificadas en la ecuación (2.12), debido que las innovaciones poseen las siguientes propiedades interesantes que son:

1.- Su valor esperado es nulo.

$$\begin{aligned} E(e_t | Y_{t-1}) &= E(y_t - y_{t|t-1} | Y_{t-1}) \\ &= E(y_t - E(y_t | Y_{t-1}) | Y_{t-1}) \\ &= E(y_t | Y_{t-1}) - E(y_t | Y_{t-1}) = 0 \end{aligned}$$

2.- Su varianza, coincide con la varianza de la observación.

$$\begin{aligned} V(e_t | Y_{t-1}) &= V(y_t - y_{t|t-1} | Y_{t-1}) \\ &= V(A_t x_t + v_t - A_t x_{t|t-1} | Y_{t-1}) \\ &= V(A_t (x_t - x_{t|t-1})) + V(v_t) \\ &= A_t P_{t-1|t-1} - 1 A_t' + R = F_{t|t-1} \end{aligned}$$

3.- Las observaciones y las innovaciones pasadas son independientes.

$$\begin{aligned} Cov(e_t, y_j) &= E(e_t y_j) \\ &= E(E(e_t y_j | Y_{t-1})) \\ &= E(y_j E(e_t | Y_{t-1})) = 0 \quad j < t \end{aligned}$$

Al ser las innovaciones independientes, se puede reescribir la función de verosimilitud no en función de las observaciones  $(y_1, y_2, \dots, y_t)$ , sino en función de las innovaciones  $(e_1, e_2, \dots, e_t)$ , permitiendo una evaluación relativamente sencilla de la función de

verosimilitud en el proceso de estimación no lineal. En dicho caso, la matriz de covarianza de las innovaciones viene dada por la siguiente expresión (ecuación 2.11):

$$\begin{aligned}\Sigma_t &= \text{var}(e_t|Y_{t-1}) \\ &= A_t P_{t|t-1} A'_t + R\end{aligned}$$

Por lo tanto, el logaritmo de la función de verosimilitud tiene la siguiente forma:

$$-\ln(L_Y(\theta)) = 1/2 \sum_{t=1}^n \log|\Sigma_t(\theta)| + 1/2 \sum_{t=1}^n e(\theta)' \Sigma_t(\theta)^{-1} e(\theta) \quad (2.16)$$

En la ecuación (2.16) las innovaciones guardan dependencia con el vector de parámetros  $\theta$ , y la función de verosimilitud no es lineal. El problema de optimización de la función de verosimilitud (2.16) se suele resolver utilizando un algoritmo de optimización numérica no lineal como el de Newton-Raphson, que actualiza los valores de los parámetros de forma reiterada hasta que se maximice el logaritmo de la función de verosimilitud (Shumway y Stoffer 2010).

En el entorno del lenguaje de programación R, se puede usar la función `optim` para poder resolver el problema, pero hay que tener en cuenta que lo que hace “`optim`” es minimizar la función objetivo, por lo cual se pone el signo negativo delante del logaritmo de la función de verosimilitud (2.16). El método de Newton Raphson, implementado en “`optim`” se resume de la forma siguiente.

### **Pasos de la estimación utilizando el método de Newton-Raphson**

La estimación de los parámetros utilizando Newton-Raphson, se realiza a través de los siguientes pasos que se encuentran documentados en el libro de Shumway y Stoffer (2010):

- 1.- Se selecciona unos valores iniciales para el vector de parámetros  $\theta^0$ .
- 2.- Con el vector de parámetros iniciales ( $\theta^0$ ), y utilizando las ecuaciones de predicciones del estado y de las observaciones del filtro de Kalman, se calcula el conjunto de las innovaciones  $\{e_t^0; t = 1, 2, \dots, T\}$ , y las covarianzas del error  $\{\Sigma_t^0; t = 1, 2, \dots, T\}$ .
- 3.- Se realiza una iteración del algoritmo de Newton-Raphson, utilizando el logaritmo de la función de verosimilitud (2.16) como criterio a maximizar, para obtener el nuevo conjunto de estimadores ( $\theta^1$ ).

4.- Se repite el paso 2, utilizando  $(\theta^1)$  para obtener el conjunto de innovaciones  $\{e_t^1; t = 1, 2, \dots, T\}$ , y las covarianzas del error  $\{\Sigma_t^1; t = 1, 2, \dots, T\}$ , luego se repite el paso 3 una vez más.

El paso 2 y 3, se repiten hasta que las estimaciones o la likelihood se estabilicen, por ejemplo, el proceso se parará cuando el valor (pequeña cantidad predeterminada) de  $\theta^{j+1}$  difiere de  $\theta^j$  donde  $j = 1, 2, 3 \dots$

#### **2.4.-Estimación vía filtro de partículas**

En las últimas décadas, la metodología del filtrado secuencial Monte Carlo (SMC), también conocida como filtro de partículas, ha sufrido un gran desarrollo tanto teórico como práctico. Muchos autores han contribuido al desarrollo de dicha metodología, como por ejemplo Kitagawa (1996; 1998), Pitt y Shephard (1999), Doucet et. Al. (2000), Arulampalam et al. (2002), Liu y West (2001), Acosta (2007; 2013), entre otros, por ejemplo, Michaud N. et Al. (2017) que han creado una librería en R en donde implementan el filtro de Liu y West (2001).

Una de las principales contribuciones al filtrado SMC, fue la introducción del remuestreo, que juntamente con el avance de la tecnología en el campo informático y la aparición de ordenadores cada vez más potentes, provocó que esta metodología pueda ser aplicada de manera práctica (Gordon et al. 1993).

La idea principal del filtro de partículas es representar la función de densidad de probabilidad posterior a través de un conjunto de muestras aleatorias o partículas, con unos pesos asociados. A partir de esas muestras ponderadas, se realiza la estimación de la función de densidad posterior.

Conforme el número de muestras o partículas va aumentando, la caracterización de Monte Carlo se convierte en una representación equivalente a la verdadera función de densidad posterior, y la solución alcanzada se acerca al estimador Bayesiano óptimo.

Por lo que, siguiendo la metodología del filtro de partículas, se puede aproximar la función de densidad de probabilidad posterior del estado  $x_t$ , a través de muestras generadas de forma recursiva a partir de la actualización y el filtrado, a medida que se vaya introduciendo la nueva información. Una vez aproximada la función de densidad de probabilidad posterior, se puede estimar cualquier característica del estado  $x_t$  a partir de las muestras o partículas ponderadas (Acosta L. 2013)



El problema que surge, en muchos casos, es que no se puede muestrear directamente de la función de densidad posterior, ya que es multivariante, no estándar y solo se conoce hasta una constante de proporcionalidad, por lo que una solución al problema es la aplicación del método “importance sampling (IS)”<sup>7</sup> (Geweke, 1989; Ristic et al. 2004).

### Importance Sampling (IS)

Un método general de integración de Montecarlo utilizado para realizar filtrado en sistemas no lineales es el IS. La idea principal es que, al ser imposible muestrear directamente de la función de densidad posterior  $\pi(x)$ , este método permite elegir una función de densidad de probabilidad alternativa  $q(x)$ , de la que se puede generar muestras de manera mucho más sencilla, donde  $q(x)$  es similar a  $\pi(x)$ .

El método SIS utiliza una densidad de importancia, que no es más que una densidad propuesta para representar otra densidad que es muy difícil de calcular exactamente, y de la que se extraerán un conjunto de muestras para poder aproximar la función de densidad que es difícil de calcular, por lo que, la elección de la densidad de importancia es crucial en el diseño del algoritmo.

Al conjunto de muestras generadas se les da un peso adecuado, que haga posible la estimación de la función de densidad de  $q(x)$  vía estimación de Monte Carlo. Esta densidad de probabilidad alternativa es lo que se denomina “distribución de importancia”.

Es decir, si en el momento fijo  $t$  se quiere aproximar la función de densidad de probabilidad posterior del filtrado  $p(x_t|Y_t)$  de la ecuación (2.5), pero no es posible muestrear directamente de ella, una posible solución consiste en muestrear de una densidad de importancia  $q(x_t|Y_t)$ . De esa manera se obtiene un conjunto de muestras o partículas  $x_t^1, x_t^2, \dots, x_t^N$  con una masa de probabilidad discreta  $w_t^1, w_t^2, \dots, w_t^N$ , y la verdadera función de densidad de probabilidad del filtrado  $p(x_t|Y_t)$  en el tiempo fijo  $t$ , sería aproximada mediante (Doucet et al. 2000)

$$p(x_t|Y_t) \approx \sum_{i=1}^N \tilde{w}_t^i \delta(x_t - x_t^i) \quad (2.17)$$

con

$$w_t^i \approx \frac{p(x_t^i|Y_t)}{q(x_t^i|Y_t)}, \quad \tilde{w}_t^i = \frac{w_t^i}{\sum_{i=1}^N w_t^i} \quad (2.18)$$

---

<sup>7</sup> En el presente trabajo, hemos optado por preservar los nombres de los diferentes algoritmos en inglés, y no hacer una traducción de los mismos al español.

Nótese que  $w_t^i$  son los pesos de importancia, y  $\tilde{w}_t^i$  son los pesos de importancia normalizados. Conforme  $N > 1$  tiende a infinito, la aproximación (2.17) converge a la verdadera función de densidad de probabilidad de filtrado  $p(x_t|Y_t)$ , siendo  $\delta(\cdot)$  la función Dirac (Doucet 1998; Doucet et al. 2000).

### Sequential Importance Sampling (SIS)

Si el objetivo es estimar  $p(x_t|Y_t)$  recursivamente, es necesario definir una distribución de importancia que funcione bien en un marco secuencial, lo que transforma el filtro de IS “importance Sampling” en SIS “Sequential Importance Sampling” (Gordon et al. 1993).

EL IS puede modificarse para que en un contexto secuencial, en cualquier momento  $t$ , la estimación de la función de densidad de probabilidad de filtrado  $p(x_t|Y_t)$ , pueda propagarse en el tiempo sin que se vean afectadas las trayectorias que se han simulado en el pasado (Acosta L. 2013). Es decir, que en el momento  $t$ , la función de distribución de importancia, debe admitir como distribución marginal la función de densidad de probabilidad en el momento anterior  $t-1$  (Doucet 1998; Acosta L. 2013), siendo ello posible si se eligen distribuciones propuestas o de importancia que factoricen:

$$q(x_t|Y_t) = q(x_t|x_{t-1}, Y_t)q(x_{t-1}|Y_{t-1}) \quad (2.19)$$

Entonces, se puede obtener muestras  $x_t^i \sim q(x_t|Y_t)$  al aumentar cada una de las muestras existentes  $x_{t-1}^i \sim q(x_{t-1}|Y_{t-1})$ , con el nuevo estado  $x_t^i \sim q(x_t|x_{t-1}, Y_t)$  (Ristic et al. 2004; Acosta L. 2013).

Para derivar la expresión de los nuevos pesos de importancia  $w_t^i$ , bajo el SIS se sustituye la ecuación recursiva del filtrado (2.5), y la ecuación de importancia (2.19), en la ecuación (2.18), obteniéndose:

$$w_t^i \propto \frac{p(y_t|x_t^i)p(x_t^i|x_{t-1}^i)}{q(x_t^i|x_{t-1}^i, y_t)} w_{t-1}^i \quad (2.20)$$

La densidad filtrada posterior del filtrado  $p(x_t|Y_t)$ , puede ser aproximada como:

$$p(x_t|Y_t) \approx \sum_{i=1}^N \tilde{w}_t^i \delta(x_t - x_t^i) \quad (2.21)$$

donde  $\tilde{w}_t^i$  representan los pesos normalizados que se obtienen a partir de la ecuación (2.20). A las partículas que poseen una mayor probabilidad se le asignan unos mayores

pesos, por el contrario, aquellas partículas que poseen una baja probabilidad, se le asignan unos pesos menores (Acosta L. 2013)

### Selección de la densidad de importancia

En el marco de la metodología del filtrado secuencial, uno de los asuntos cruciales a la hora de utilizar el filtro de partículas, es la elección de la densidad de importancia  $q(x_t|x_{t-1}^i, y_t)$ . La densidad de importancia óptima es aquella que minimiza la varianza de los pesos de importancia  $w_t^i$  (Ristic et al. 2004).

La elección más popular de la función de densidad de importancia, y que ha sido utilizada por varios autores (Kitagawa 1996), es la denominada “transicional prior”:

$$q(x_t^i|x_{t-1}^i, y_t) = p(x_t^i|x_{t-1}^i) \quad (2.22)$$

Si ésta transicional prior es utilizada como densidad de importancia, sustituyendo la ecuación (2.22) en la ecuación (2.20), tenemos que la expresión de los pesos de importancia es la siguiente:

$$w_t^i \propto w_{t-1}^i p(y_t|x_t^i) \quad (2.23)$$

Nótese que estos pesos son proporcionales al producto de los pesos en el momento anterior y la probabilidad de la nueva observación  $y_t$ , dada la partícula. Por lo que, utilizando la transicional prior como densidad propuesta en el SIS, los pesos se propagan y se actualizan conforme llega la nueva observación a través de la ecuación (2.23) (Acosta L. 2013). Es decir, el algoritmo en vez de muestrear de  $x_t^i \sim q(x_t|x_{t-1}^i, y_t)$ , lo hace a partir de la transicional prior  $p(x_t^i|x_{t-1}^i)$ , generando  $z_t^i$ , según el ruido de la ecuación de estado (2.1), ajustando  $x_t^i = f(x_{t-1}^i, z_t^i)$ , y asignándole a cada partícula  $x_t^i$  su respectivo peso de importancia  $w_t^i$  a partir de la expresión (2.23).

El problema que surge bajo SIS es que, llegado un momento  $t$ , unas pocas muestras o partículas tendrán asignados unos pesos grandes, mientras que otras partículas tendrá un peso pequeño siendo su contribución prácticamente nula. Este problema de degeneración en las partículas provoca que el filtro SIS no funcione. En la práctica una buena elección de la función de importancia es relevante.

Este problema de la degeneración de la muestra, es uno de los grandes obstáculos que se presentó en el desarrollo de los métodos de Monte Carlo, y que de hecho, a la larga es

inevitable dentro del marco del muestreo secuencial (Doucet et al 2000; Ristic et al. 2004; Acosta L 2013).

### **Problema de degeneración de la muestra**

Este problema consiste en que a medida que  $t$  incrementa, la varianza de los pesos irá aumentando, y solo unas cuantas partículas tendrán un peso relevante, mientras que la contribución de otras será prácticamente nula. Este aumento de la varianza provoca un problema en la precisión de la predicción del filtro (Acosta L. 2013)

Una posible manera de monitorizar el grado de degeneración, consiste en calcular el llamado “effective simple size” (Acosta L. 2013) mediante:

$$\hat{N}_{eff} = \frac{1}{\sum_{i=1}^N (\tilde{w}_t^i)^2}$$

donde,  $\tilde{w}_t^i$  son los pesos normalizados, y además se puede verificar que  $1 \leq N_{eff} \leq N$ . Si el  $N_{eff}$  es pequeño, significa que existe un serio problema de degeneración, es decir, los pesos tienen una mayor varianza. Una de las formas para reducir el problema de la degeneración es la introducción del remuestreo en el filtro SIS.

### **Sequential Importance Resampling (SIR)**

Una contribución al SIS, propuesta por Gordon et al (1993) y Kitagawa (1996), introduce un nuevo paso al filtro SIS, que es el remuestreo. Estos autores utilizan la transicional prior como función de densidad propuesta y realizan el remuestreo en cada periodo de tiempo (Ristic et al. 2004).

Nótese que si el remuestreo se aplica en cada momento del tiempo, no es necesario pasar los pesos de importancia de un paso al siguiente, ya que  $w_{t-1}^i = \frac{1}{N} \forall i = 1, 2, \dots, N$ , quedando la ecuación de los pesos de importancia (2.23) simplificada a:

$$w_t^i \propto p(y_t | x_t^i)$$

El remuestreo reduce el problema de la degeneración, quitando las muestras que poseen un menor peso, pero puede introducir un problema potencial conocido como “empobrecimiento de la muestra”. Es decir, en cada remuestreo la diversidad de las partículas tiende a disminuir, ya que las partículas con mayor peso tienden a salir más, partículas con menor peso tienden a salir menos, en el caso extremo todas partículas podrían llegar a colapsar en una sola.

Este problema del empobrecimiento de la muestra, tendrá un efecto negativo en la precisión del filtro, ya que, toda la distribución posterior estaría representada por muy pocas partículas. Dicha estimación claramente no sería fiable y sería recomendable el monitorizar el posible empobrecimiento de la muestra, calculando el  $N_{eff}$  en cada remuestreo, y comparando con un umbral que el investigador crea conveniente. De esta forma se va controlando el proceso para evitar que la distribución quede representada por muy pocas partículas y también se podría determinar mediante simulación una regla empírica sobre el número mínimo de partículas a usar según el escenario concreto bajo estudio (Acosta L. 2013). Por ejemplo, para el modelo de volatilidad estocástica, la tesis de Acosta L. determina que para obtener estimaciones fiables el número recomendable a usar un mínimo de 5000 partículas.

### **Auxiliar Sequential Importance Resampling (ASIR)**

Los autores Pitt y Shephard (1999), introducen el ASIR, una variante de filtro de partículas que difiere básicamente del SIR en que usa una distribución de importancia distinta de la distribución transicional. El algoritmo realiza remuestreos hasta  $t-1$  (usando la información disponible hasta el tiempo  $t$  “look-ahead step”) antes de que las partículas se propaguen al periodo  $t$  (Ristic et al. 2004).

Este filtro introduce una variable auxiliar  $i$  en el vector de estado  $x_t$ , y por tanto genera muestras o partículas de un vector de estado de mayor dimensión  $(x_t, i)$ . Es decir, en lugar de generar muestras de la función de densidad de probabilidad del filtrado  $p(x_t|Y_t)$ , como el algoritmo SIR, genera muestras a partir de la función de densidad de probabilidad conjunta  $p(x_t, i|Y_t)$ .

Usando la regla de Bayes, se puede derivar una expresión para la densidad conjunta  $p(x_t, i|Y_t)$  como:

$$p(x_t, i|Y_t) \propto p(y_t|x_t)p(x_t|x_{t-1}^k)w_{t-1}^k, i = 1, 2, \dots, N \quad (2.24)$$

En este caso, la distribución de importancia  $q(x_t, i|Y_t)$  permite muestrear indirectamente de la función de densidad conjunta  $p(x_t, i|Y_t)$ . Por tanto, la densidad de importancia utilizada para extraer secuencialmente las partículas del estado, queda definida por la expresión

$$q(x_t, i|Y_t) \propto p(y_t|\mu_t^i)p(x_t|x_{t-1}^i)w_{t-1}^i, i = 1, 2, \dots, N \quad (2.25)$$

donde  $\mu_t^i$  denota una característica asociada a la densidad de  $(x_t|x_{t-1}^i)$ , que puede ser el valor esperado de  $E(x_t|x_{t-1}^i)$  o cualquier otra característica del estado  $x_t$  dado  $x_{t-1}^i$  (Acosta L. 2013). Se puede diseñar la distribución de importancia tal:

$$q(i|Y_t) \propto w_{t-1}^i \int p(y_t|\mu_t^i) p(x_t|x_{t-1}^i) dx_t = w_{t-1}^i p(y_t|\mu_t^k)$$

Como resume Acosta L. (2013), se remuestrea a partir de una función de densidad de importancia alternativa  $q(x_t, i|Y_t)$ , simulando el índice  $i$  con probabilidad  $\lambda_i$  (ponderaciones de primera etapa) proporcional a  $q(i|Y_t) = w_{t-1}^i p(y_t|\mu_t^k)$ , luego descartamos el índice  $i$  muestra  $x_t^i$  de la función de densidad de probabilidad marginalizada  $p(x_t|Y_t)$ . Es decir, que una vez que se muestrea  $i$ , muestreamos  $x_t^i$ , exactamente como lo hace el algoritmo SIR, y los pesos de segunda etapa  $w_t^i$  pueden derivarse de las ecuaciones (2.24) y (2.25) obteniéndose la siguiente expresión:

$$w_t^i \propto \frac{p(y_t|x_t^{k^i})}{p(y_t|\mu_t^{k^i})}$$

Tanto el SIR como el ASIR son variantes de filtro de partículas propuestas originalmente para estimar solo el estado de un sistema.

### **Filtro de partículas de Liu y West**

Los Autores Liu y West (2001) introducen un filtro de partículas que a diferencia de los filtros SIR y ASIR, que estiman solamente el estado  $x_t$ , busca estimar de forma simultanea el estado  $x_t$ , y los parámetros fijos del modelo bajo estudio.

Estos autores combinan las ideas de Pitt y Shephard (ASIR) para filtrar el estado, con las ideas de Liu y West (2001) para perturbar las partículas mediante el uso de kernels Gaussianos.

Antes de explicar el funcionamiento del algoritmo de Liu y West, se introduce unos conceptos generales y ampliaremos la formulación de los modelos de espacio-estado para contemplar la estimación simultánea del estado y los parámetros.

### **Formulación del modelo de espacio-estado aumentado**

Para poder estimar el estado  $x_t$  y los parámetros del modelo de espacio-estado, se aumenta el vector de estado  $x_t$ , introduciéndole el vector de parámetros  $\theta$  (Muñoz

1988; Kitagawa 1998), por lo que el vector de estado aumentado, queda definido de la siguiente manera:

$$l_t = \begin{bmatrix} x_t \\ \theta \end{bmatrix}$$

Con el nuevo vector del estado aumentado  $l_t = (x_t, \theta)'$ , se puede re-escribir las ecuaciones de transición y de observación (2.1) y (2.2), de la siguiente forma:

**Ecuación de transición:**  $l_t = \tilde{f}(l_{t-1}, z_t)$  (2.26)

**Ecuación de observación:**  $y_t = \tilde{g}(l_t, v_t)$  (2.27)

donde:

$$\tilde{f}(l_t, z_t) = \begin{bmatrix} f(x_t, z_t) \\ \theta \end{bmatrix}$$

$$\tilde{g}(l_t, v_t) = g(x_t, v_t)$$

$\theta$ , es el vector de parámetros del modelo, que se consideran desconocidos (pero fijos). Por último, al vector de estado aumentado inicial  $l_0$ , se le da una distribución a priori, nótese que al incorporarle el vector de parámetros  $\theta$ , al vector de estado  $x_t$ , se tiene un problema de filtrado no lineal (Acosta L. 2013).

### **Ecuaciones de predicción y filtrado**

Utilizando los supuestos básicos de los modelos de espacio-estado y la regla de Bayes e introduciendo el vector de estado aumentado  $l_t = (x_t, \theta)'$ , en la ecuación de predicción (2.3), de filtrado (2.4) y la ecuación recursiva del filtrado (2.5), se obtienen las siguientes ecuaciones:

### **Ecuación de predicción de la función de densidad de probabilidad**

Asumiendo que en el periodo  $t$ , la función de densidad de probabilidad del filtrado en el periodo  $t-1$   $p(x_{t-1}, \theta_{t-1} | Y_{t-1})$ , está disponible, se puede expresar la ecuación general de predicción de un paso hacia adelante, de la siguiente manera:

$$p(l_t | Y_{t-1}) = p(x_t, \theta_t | Y_{t-1}) = \int p(x_t | x_{t-1}, \theta_t) p(x_{t-1}, \theta_t | Y_{t-1}) dx_{t-1} \quad (2.28)$$

donde  $p(x_t | x_{t-1}, \theta_t)$ , es la evolución del estado obtenida a partir de la ecuación (2.26). Nótese que,  $\theta_{t-1}$ , no implica que los parámetros varían con el tiempo (ya que son parámetros fijos), sino que denota las estimaciones de  $\theta$  en el periodo  $t-1$ .

### Ecuación de filtraje de la función de densidad de probabilidad

Una vez esté disponible la observación  $y_t$  en el periodo  $t$ , haciendo uso de la regla de Bayes, y teniendo en cuenta el vector de estado aumentado, se puede expresar la ecuación de filtraje de la siguiente manera:

$$\begin{aligned}
 p(l_t|Y_t) = p(x_t, \theta_t|Y_t) &= \frac{p(Y_t|x_t, \theta_t)p(x_t, \theta_t)}{p(Y_t)} \\
 &= \frac{p(y_t, Y_{t-1}|x_t, \theta_t)p(x_t, \theta_t)}{p(y_t, Y_{t-1})} \\
 &= \frac{p(y_t|Y_{t-1}, x_t, \theta_t)p(Y_{t-1}|x_t, \theta_t)p(x_t, \theta_t)}{p(y_t|Y_{t-1})p(Y_{t-1})} \\
 &= \frac{p(y_t|Y_{t-1}, x_t, \theta_t)p(x_t, \theta_t|Y_{t-1})p(Y_{t-1})p(x_t, \theta_t)}{p(y_t|Y_{t-1})p(Y_{t-1})p(x_t, \theta_t)} \\
 &= \frac{p(y_t|x_t, \theta_t)p(x_t, \theta_t)p(x_t, \theta_t|Y_{t-1})}{p(y_t|Y_{t-1})} \\
 &\propto p(y_t|x_t, \theta_t)p(x_t, \theta_t|Y_{t-1}) \tag{2.29}
 \end{aligned}$$

donde  $p(y_t|x_t, \theta_t)$ , es la probabilidad de  $y_t$ , que se obtiene a partir de la ecuación (2.27) que mide la evolución de la densidad de las observaciones.  $p(x_t, \theta_t|Y_{t-1})$ , es la ecuación de predicción (2.28),  $p(y_t|Y_{t-1}) = \int p(y_t|x_t, \theta_t)p(x_t, \theta_t|Y_{t-1})$  es la constante de normalización, que no es fácil de calcular, de forma adicional, si se desarrolla la ecuación (2.29) un paso hacia adelante, se obtiene una ecuación de filtrado alternativa de la función de densidad de probabilidad (Acosta L. 2013).

$$p(l_t|Y_t) = p(x_t, \theta_t|Y_t) \propto p(y_t|x_t, \theta_t)p(x_t|\theta_t, Y_{t-1})p(\theta_t|Y_{t-1}) \tag{2.30}$$

Nótese que la ecuación (2.30), implícitamente muestra la contribución de la función de densidad anterior del vector de parámetros  $p(\theta_t|Y_{t-1})$ , en la actualización del vector de estado aumentado  $l_t$  (Acosta L. 2013).

Si el vector de parámetros  $\Theta$  es conocido, la ecuación (2.30), se reduce a la ecuación (2.4), ya que  $p(\theta_t|Y_{t-1})$  se degenera, por lo que los parámetros conocidos pueden ser quitados de las declaraciones de condicionamiento.

### Estimación simultánea del vector de parámetros vía filtro de partículas

Para poder realizar la estimación simultánea del estado  $x_t$  y del vector de parámetros desconocidos  $\Theta$ , se supone que en el periodo  $t$ , la función de densidad de probabilidad



del filtrado  $p(x_t, \theta_t | Y_t)$ , puede ser aproximada a través de un conjunto de muestras suficientemente grande  $\{(x_t^1, \theta_t^1), \dots, (x_t^i, \theta_t^i)\}$ , con una probabilidad discreta  $\tilde{w}_t^1, \dots, \tilde{w}_t^i$ . El conjunto de muestras se obtiene de la función de densidad propuesta alternativa  $q(x_t, \theta_t | Y_t)$ , que es similar a  $p(x_t, \theta_t | Y_t)$ .

Bajo el enfoque del filtro de partículas, las ecuaciones teóricas de predicción (2.28) y filtrado (2.29), se aproximan a través de las densidades empíricas, por lo que, asumiendo que en el periodo t-1 se dispone de las muestras  $(x_{t-1}^i, \theta_{t-1}^i)_{i=1}^N$ , obtenidas de  $p(l_{t-1} | Y_{t-1})$ , la expresión de predicción aproximada de la función de densidad de probabilidad  $p(l_t | Y_{t-1})$  (2.28) viene dada por:

### **Ecuación de predicción aproximada de la función de densidad de probabilidad**

$$p(l_t | Y_{t-1}) = p(x_t, \theta | Y_{t-1}) \approx \sum_{i=1}^N p(x_t | x_{t-1}^i, \theta^i) \tilde{w}_{t-1}^i \quad (2.31)$$

donde  $\tilde{w}_{t-1}^i$  son los pesos de importancia normalizados en el tiempo t-1. Juntando la probabilidad  $p(y_t | x_t, \theta_t)$  y la aproximación de la función de densidad de probabilidad previa, se obtiene la ecuación de filtrado aproximada de la función de densidad de probabilidad  $p(l_t | Y_t)$  (2.29), que viene dada por la siguiente expresión:

### **Ecuación de filtrado aproximada de la función de densidad de probabilidad**

$$p(l_t | Y_t) = p(x_t, \theta | Y_t) \approx p(y_t | x_t, \theta) \sum_{i=1}^N p(x_t | x_{t-1}^i, \theta^i) \tilde{w}_{t-1}^i \quad (2.32)$$

A través del filtro de partículas, la función de densidad del filtrado del vector de estado aumentado, se obtiene aproximándola haciendo uso de una distribución empírica.

$$p(l_t | Y_t) = p(x_t, \theta | Y_t) \approx \sum_{i=1}^N \tilde{w}_t^i \delta(l_t - l_t^i)$$

donde  $\sum_{i=1}^N \tilde{w}_t^i = 1$ ,  $\tilde{w}_t^i = \frac{w_t^i}{\sum_{i=1}^N w_t^i}$  y  $w_t^i$  es igual a la siguiente expresión:

$$w_t^i \propto \frac{p(y_t | x_t^i, \theta^i) p(x_t^i | x_{t-1}^i, \theta^i)}{q(x_t^i | x_{t-1}^i, \theta^i, y_t)} \tilde{w}_{t-1}^i \quad (2.33)$$

En la práctica, es común elegir la transición previa (transition prior) como densidad propuesta, en vez de la función de densidad propuesta óptima, que teóricamente tiene en cuenta la información hasta la última observación.

$$q(x_t | x_{t-1}^i, \theta^i, y_t) = p(x_t | x_{t-1}^i, \theta^i)$$

Si se opta por elegir la transición previa como densidad propuesta, la ecuación de los pesos de importancia (2.3) se ve reducida a la siguiente expresión:

$$w_t^i \propto p(y_t | x_t^i, \theta^i) \tilde{w}_{t-1}^i$$

Además, si se realiza el remuestreo en cada periodo, la formulación anterior de los pesos, tienen una forma mucho más reducida:

$$w_t^i \propto p(y_t | x_t^i, \theta^i)$$

### **Perturbación de los parámetros según Liu y West**

En este apartado se explica de forma sencilla, las características más importantes del filtro de partículas presentado por Liu y West (2001).

Asumiendo que en el periodo  $t-1$ ,  $\theta_{t-1}^i$  es el vector de parámetros de las partículas obtenido a partir de  $p(\theta_{t-1} | Y_{t-1})$ , y que los parámetros de ubicación correctos de kernel, pueden definirse y calcularse a través de la siguiente expresión.

$$m_{t-1}^i = a\theta_{t-1}^i + (1-a)\bar{\theta}_{t-1} \quad (2.34)$$

Esto implica que la correspondiente aproximación kernel para  $p(\theta_t | Y_{t-1})$ , viene dada por la siguiente expresión:

$$p(\theta_t | Y_{t-1}) \approx \sum_{i=1}^N \tilde{w}_{t-1}^i S(\theta_t; m_{t-1}^i, h^2 V_{t-1})$$

Donde  $h > 0$  es un parámetro de suavizado de control, que está definido en términos de un factor de descuento ( $\delta$  en  $(0,1]$ ), respecto al valor del factor de descuento, Liu y West aconsejan utilizar los valores 0,95-0,99.

Nótese que  $a = \frac{(3\delta-1)}{2\delta}$ ,  $h^2 = 1 - a^2$  siendo  $a = \sqrt{1 - h^2}$  y  $h^2 V_{t-1}$  es la varianza de kernel. Sustituyendo la ecuación (2.34) en la ecuación de filtrado alternativa (2.30), se obtiene la expresión para de forma simultánea el filtrado del vector de estado y los parámetros fijos desconocidos:

$$\begin{aligned} p(l_t | Y_t) &= p(x_t, \theta_t | Y_t) \\ &= p(y_t | x_t, \theta_t) p(x_t, \theta_t | Y_{t-1}) \\ &\propto p(y_t | x_t, \theta_t) p(x_t, \theta_t | Y_{t-1}) p(\theta_t | Y_{t-1}) \\ &\approx p(y_t | x_t, \theta_t) p(x_t | \theta_t, Y_{t-1}) \sum_{i=1}^N \tilde{w}_{t-1}^i S(\theta_t; m_{t-1}^i, h^2 V_{t-1}) \end{aligned}$$

Por lo que, en el momento  $t$ , se tiene una muestra combinada  $\{(x_t^i, \theta^i), i = 1, 2, \dots, N\}$ , que representa una aproximación de muestreo de importancia de la densidad posterior para los parámetros y el estado en el periodo  $t$ ,  $p(x_t, \theta_t | Y_t)$ .

Nótese que el filtro que utilizan Liu y West (2001) posee dos características principales tal como resume Acosta L. (2013):

1. Utiliza la variante del filtro ASIR para modelar la evolución del estado del sistema.
2. En cada paso de tiempo, se aplica la perturbación de los parámetros de acuerdo a lo propuesto por Liu y West para suavizar y regenerar las partículas de los parámetros fijos del modelo. De esta manera, se evita el problema de pérdida de información histórica que presenta el enfoque de la evolución artificial propuesta por Kitagawa.

Liu y West (2001) sostienen que el paso de remuestreo final puede hacerse de forma opcional, aunque ellos en su trabajo realizan el remuestreo en cada paso; ver en la Tabla 1 el pseudo código del algoritmo de Liu y West tal como se presenta en Acosta (2013).

---

**Tabla 1. Filtro de Partículas de Liu y West.**

---

**Inicialización:**  $t=0$

**for**  $i=1$  hasta  $N$

Muestrear  $x_0^i \sim p(x_0)$

Muestrear  $\theta_0^i \sim p(\theta_0)$

**end for**

Se elige un factor de descuento  $\delta (0,1]$

Calcular el parámetro  $a = \frac{(3\delta-1)}{2\delta}$

Calcular el parámetro de suavizado de control  $h^2 = 1 - a^2$

**for**  $t=1$  hasta  $T$

**for**  $i=1$  hasta  $N$

Calcular  $\mu_t^i$  asociado a la función de densidad de probabilidad condicionada  $p(x_t|x_{t-1}^i, \theta_{t-1}^i)$ , en este caso  $\mu_t^i = E(x_t|x_{t-1}^i, \theta_{t-1}^i)$ .

Calcular la media  $\bar{\theta}_{t-1}$  y la varianza  $V_{t-1}$  de la aproximación de Monte Carlo  $p(\theta_t|Y_{t-1})$ .

**end for**

**for**  $i=1$  hasta  $N$

Calcular los pesos de primera etapa  $\lambda_t^i = q(k|Y_t) \propto w_{t-1}^i p(y_t|\mu_t^i, m_{t-1}^i)$

**end for**

**for**  $i=1$  hasta  $N$

Normalizar los pesos de primera etapa  $\tilde{\lambda}_t^i = \frac{\lambda_t^i}{\sum_{i=1}^N \lambda_t^i}$

**end for**

Remuestrear con remplazo del índice  $i_{j=1}^N$  de acuerdo a los pesos de primera etapa calculados.

**for**  $i=1$  hasta  $N$

Muestrear un nuevo vector de parámetros  $\theta_t^{i(j)} \sim S(\cdot | m_{t-1}^{i(j)}, h^2 V_{t-1})$

**end for**

**for**  $i=1$  hasta  $N$

Muestrear  $x_t^{i(j)} \sim q(x_t|i^j, Y_t) = p(x_t|x_{t-1}^{i(j)}, \theta_t^{i(j)})$  (se toman muestras de las partículas de la ecuación de transición).

Calcular los pesos de segunda etapa  $w_t^i$  según:

$$w_t^i \propto \frac{p(y_t|x_t^{i(j)}, \theta_t^{i(j)})}{p(y_t|\mu_t^{i(j)}, m_{t-1}^{i(j)})}$$

**end for**

**for**  $i=1$  hasta  $N$

Normalizamos los pesos de importancia

$$\tilde{w}_t^i = \frac{w_t^i}{\sum_{i=1}^N w_t^i}$$

**end for**

En este punto, tenemos una aproximación  $(x_t^{i(j)}, \theta_t^{i(j)})$  con pesos  $\tilde{w}_t^i$

Realizamos el segundo remuestreo con remplazo de las partículas  $(x_t^{i(j)}, \theta_t^{i(j)})$ ,  $i = 1, 2, \dots, N$ , según los pesos de importancia  $\{\tilde{w}_t^1, \tilde{w}_t^2, \dots, \tilde{w}_t^i\}$  (este paso puede ser opcional)

**end for**

---

### 3.- Modelo del modelo de volatilidad estocástica: Predicción y Filtrado

#### 3.1.- Análisis de los datos.

La serie financiera que se analiza en el presente trabajo, es la cotización diaria del índice bursátil IBEX35, en el periodo comprendido entre el 01/01/2002 al 31/12/2016. Los datos han sido obtenidos de la página web de Yahoo Finance<sup>8</sup>.

El IBEX35, índice bursátil de referencia de la bolsa española, es elaborado por “Bolsas y Mercados Españoles”. Está formado por las 35 empresas con mayor liquidez, y al ser un indicador ponderado no todas las empresas que forman parte del índice tiene el mismo peso<sup>9</sup>.

En series financieras es frecuente trabajar con los rendimientos logarítmicos, debido a sus propiedades estadístico-matemáticas y porque se facilita la interpretación de cambios ocurridos en tanto por ciento.

Los rendimientos logarítmicos (de ahora en adelante rendimientos), se calculan como la diferencia del logaritmo de la cotización del índice entre el momento  $t$  y  $t-1$ :

$$\begin{aligned} r_t &= \log(P_t) - \log(P_{t-1}) \\ &= (1 - B)\log(P_t) \end{aligned} \quad (3.1)$$

donde  $B$  es el operador de retardo, y por lo tanto,  $B(\log(P_t)) = \log(P_{t-1})$ . La formulación (3.1) es aproximadamente igual a:

$$r_t = \frac{P_t - P_{t-1}}{P_t} \quad (3.2)$$

Muchas veces la transformación logarítmica es utilizada para homogeneizar la varianza en algunas series temporales, pero en series financieras dicha transformación no suele funcionar, debido a las características típicas que presentan este tipo de series.

La Figura 1, muestra la evolución de la cotización del índice durante el periodo de estudio, en la que se observa que en el periodo comprendido entre el 2008 y 2010 se produjo una fuerte caída en la cotización debido a varios factores, como por ejemplo, la crisis financiera que empezó a finales del 2007.

**Figura 1. Evolución de la cotización del IBEX 35 del 01/01/2002 al 31/12/2016**

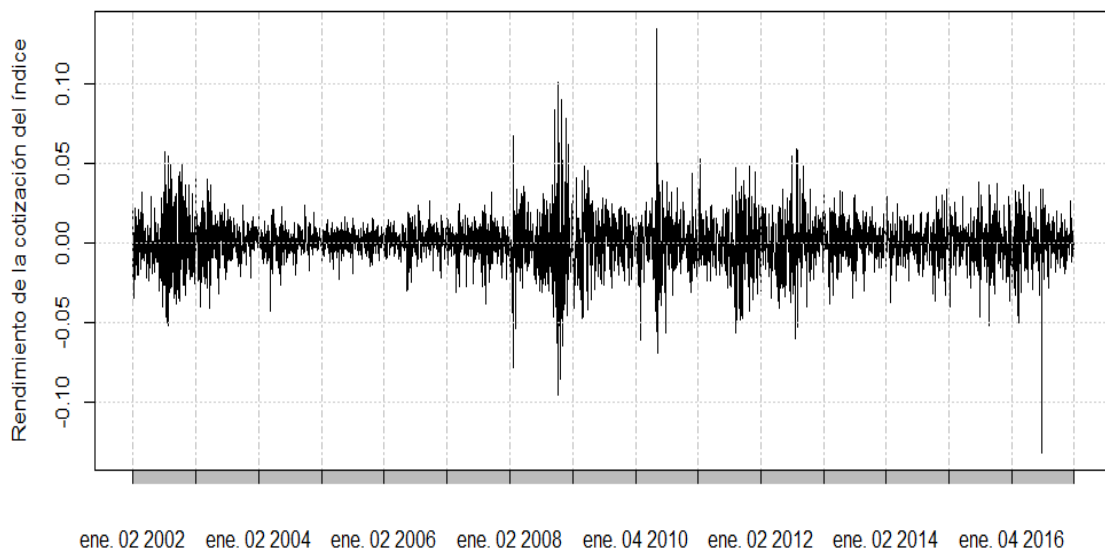
<sup>8</sup> <https://finance.yahoo.com> \_visitado el 30/01/2018

<sup>9</sup> [https://es.wikipedia.org/wiki/IBEX\\_35](https://es.wikipedia.org/wiki/IBEX_35) visitado el 05/02/2018



Por otro lado, la Figura 2 muestra los rendimientos del índice, en la que se observa los clusters de volatilidad – periodos de baja volatilidad, seguidos por periodos de alta volatilidad – que es una de las principales características de las series financieras.

**Figura 2. Rendimientos de la cotización del IBEX35**



En la Figura 2, también se observa que la serie de rendimientos no es del todo estacionaria, y aunque en media sí parece serlo no lo es en varianza. Se aprecia que la variabilidad no es constante sino que es cambiante a lo largo del tiempo, por lo que necesitaremos plantear un modelo que pueda captar esta característica de varianza cambiante.

En la tabla 2, se presenta un análisis básicamente descriptivo de la serie de rendimientos del IBEX35. De este análisis se destaca lo siguiente:

- 1.- El valor mínimo es menor que el valor máximo, lo que indica que para el periodo analizado las subidas que ha experimentado el índice han sido más importantes que las bajadas durante el periodo analizado.
- 2.- Parece que la media de la serie de rendimientos es cero, lo que se confirma con el p-valor del test sobre la media presentado en la tabla 1A (anexo). Cualquier prueba de contraste en este trabajo considera un nivel de significación del 5%.
- 3.- La serie presenta asimetría negativa (-0,082), también se observa un exceso de curtosis ( $>3$ ), lo que sugiere la existencia de colas pesadas, característica también típicamente observada en series financieras.

**Tabla 2. Descriptivos básicos de la Serie de Rendimiento del IBEX35**

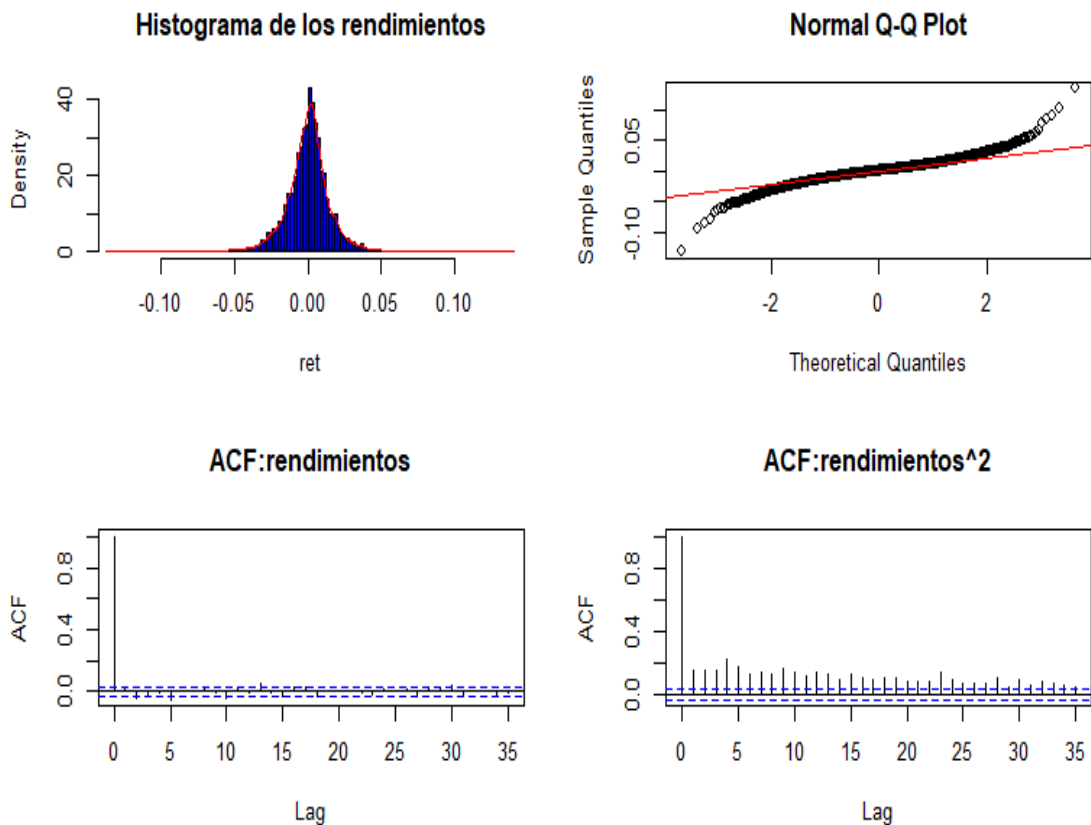
<b>Estadístico</b>	<b>Valor Estadístico</b>
Mínimo	-0,1319
Máximo	0,1348
Media	0,0000
Mediana	0,0008
Desv. Típica	0,0151
Asimetría	-0,0826
Curtosis	6,5158
Nº observ.	3814

Referente a la distribución de la serie de rendimientos, se observa que esta no se distribuye como una normal, como se confirma por 2 test de normalidad presentados en tabla 1A.

Gráficamente, también se observa la no-normalidad de los rendimientos; ver la Figura 3 donde se presentan los gráficos de normalidad, el histograma y la ACF de los rendimientos y los rendimientos al cuadrado. En el gráfico de normalidad, los valores no se ajustan del todo a la línea de los cuantiles teóricos, sino que tienen una forma de “S” lo que sugiere que hay un exceso de curtosis y también se aprecia posible presencia de valores atípicos.

En el histograma de los rendimientos, se aprecian colas pesadas, como consecuencia del exceso de curtosis que se comentaba anteriormente, y la campana un poco más puntiaguda que la campana de la distribución normal, presentando una distribución leptocúrtica<sup>10</sup>.

**Figura 3. Gráficos descriptivos: Histograma, QQ-plot y ACF**



Para finalizar este análisis estadístico preliminar de la serie de rendimientos, nos centramos en el gráfico de la ACF de la serie de rendimientos y rendimientos al cuadrado. Se observa que hay una estructura de autocorrelación entre los residuos (no son ruido blanco), ya que sobretodo en el gráfico de la ACF de la serie de rendimientos al cuadrado, se aprecia una estructura de dependencia en los valores fuera de las bandas de confianza del 95%. Todo ello evidencia presencia de volatilidad en la serie de rendimientos.

<sup>10</sup> Se dice que una serie presenta una distribución leptocúrtica (alto grado de concentración en torno a los valores centrales) cuando  $g_2 > 0$  donde  $g_2 = \mu_4 / \sigma^4 - 3$



La presencia de volatilidad, se confirma por el test conjunto de autocorrelación de Ljung-Box, donde la hipótesis nula es  $\rho=0$  (no correlación presente hasta el retardo k). Se observa que hasta los 20 primeros retardos (ver en anexos tabla 2A) se rechaza la hipótesis nula, confirmando lo observado en el ACF de la Figura 3.

Resumiendo:

- 1.- Los rendimientos logarítmicos son estacionarios en media, pero no en varianza, ya que la varianza condicional va cambiando en el tiempo, mostrando una estructura de dependencia.
- 2.- La media de los rendimientos puede considerarse igual a cero.
- 3.- Los rendimientos presentan agrupamientos de volatilidad, evidenciando una de las características de las series financieras: varianza cambiante.
- 3.- Los rendimientos no se distribuyen según una normal, sino que presentan una distribución leptocúrtica, con presencia de colas pesadas.
- 4.- Tanto las pruebas individuales como conjunta, en base a los rendimientos y los rendimientos al cuadrado respectivamente, indican la presencia de una estructura de dependencia, siendo la última evidencia de la presencia de volatilidad.

Todo lo anterior indica la necesidad de plantear un modelo para captar esta característica de varianza cambiante. Es decir, necesitamos ajustar un modelo para la volatilidad latente en la serie de rendimientos.

### **3.2.- Estimación de la volatilidad y parámetros vía filtro de Kalman y máxima verosimilitud**

Para la estimación de la volatilidad estocástica, se usa la formulación en espacio de estado del modelo SARV(1) introducido en el capítulo 1, usando la especificación propuesta por Sandmann y Koopman (1998). Para ello, es necesario llevar el modelo original de volatilidad estocástica que no es lineal, a una forma funcional que sea lineal, y formular tanto la ecuación de observación, así como la ecuación de estado.

#### **Formulación de la volatilidad estocástica en la forma de espacio-estado.**

##### **Ecuación de transición o estado:**

La ecuación de estado de un modelo SARV(1) queda definida de la siguiente forma:

$$x_t = \mu + \phi(x_{t-1} - \mu) + z_t$$

donde  $x_t = \log(\sigma_t^2)$ , operando algebraicamente la ecuación anterior, puede reescribirse como:

$$x_t = \phi_0 + \phi_1 x_{t-1} + z_t, \quad z_t \sim N(0, \sigma_z) \quad (3.3)$$

donde  $\phi_0 = \mu(1 - \phi_1)$ , es una constante que mide la volatilidad media,  $\phi_1$  es el parámetro de persistencia de la volatilidad.  $z_t$ , es el ruido del estado que se distribuye según una normal centrada en cero y varianza  $\sigma_z^2$ .

### Ecuación de observación

La ecuación de observación de los rendimientos es la siguiente:

$$r_t = \sigma_* \cdot \exp\left(\frac{x_t}{2}\right) \cdot \epsilon_t = \sigma_* \cdot \sigma_t \cdot \epsilon_t, \quad \epsilon_t \sim WN(0,1) \quad (3.4)$$

donde,  $\sigma_*$ , es la variabilidad marginal (opcional a estimar),  $\sigma_t$  es una variable latente (volatilidad) y  $\epsilon_t$ , es una variable aleatoria que se distribuye usualmente según una normal estándar.

Como se aprecia en la ecuación de observación de los rendimientos (3.4), el modelo de espacio de estado no es lineal, por lo que, para linealizarla, se transforma elevando y aplicando logaritmos, quedando la siguiente expresión lineal:

$$y_t = \log(r_t^2) = \log(\sigma_*^2) + \log(\sigma_t^2) + \log(\epsilon_t^2), \quad \log(\epsilon_t^2) \sim \log(\chi_1^2) \quad (3.5)$$

donde la nueva variable respuesta  $y_t = \log(r_t^2)$  depende de una constante  $\log(\sigma_*^2)$ , y de dos procesos estocásticos independientes entre sí,  $\log(\epsilon_t^2)$  y  $x_t = \log(\sigma_t^2)$ , el último sigue un proceso estacionario lineal especificado en la ecuación (3.3).

Un problema que se presenta al linearizar la ecuación de observación usando la transformación  $y_t = \log(r_t^2)$ , es que se obtiene un modelo no Gaussiano. Tal como demostraron Abromowitz y Stegun (1970), si  $\epsilon_t$  es normal estándar el  $\log(\epsilon_t^2)$  se distribuye como  $\log(\chi_1^2)$  con un grado de libertad, con media igual -1,27 y varianza  $(\pi/2)$ , donde la función de densidad viene dada por:

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(e^x - x)\right] \quad -\infty < x < \infty \quad (3.6)$$

La función de densidad (3.6) está fuertemente sesgada, con presencia de colas pesadas por la izquierda, por lo que la formulación de espacio de estado definida por las ecuaciones (3.3) y (3.5), es markoviana y lineal pero no es Gaussiana.

Es sabido que el filtro de Kalman es un método óptimo cuando tenemos un modelo de espacio de estado lineal y Gaussiano. Para poder usar el filtro de Kalman y mantener la simplicidad del modelo, se aproximará el  $\log(\chi_1^2)$  mediante una mixtura formada por dos componentes gaussianas; aproximación usada por varios autores (Shumway y Stoffer 2001, entre otros). Específicamente se aproximará el  $\log(\chi_1^2)$  a través de una mixtura formada por dos componentes gaussianas, fijando que una de ellas esté centrada en cero y fijando la probabilidad  $\pi=0.5$  (Shumway y Stoffer 2001). De esta forma se puede re-escribir la ecuación de observación (4.3) de la siguiente manera:

$$\begin{aligned} y_t &= \log(r_t^2) = \log(\sigma_*^2) + \log(\sigma_t^2) + \log(\epsilon_t^2) \\ &= \alpha + x_t + \eta_t \end{aligned} \quad (3.7)$$

donde  $\eta_t$ , se distribuye como una mixtura de dos componentes normales,  $\eta_t = I_t u_{t0} + (1 - I_t)u_{t1}$ .

Con  $I_t \sim \text{Bernouilli}$ ,  $\Pr[I_t = 0] = \pi_0$ ,  $\Pr[I_t = 1] = \pi_1$ . La suma de las probabilidades es igual a 1 ( $\pi_0 + \pi_1 = 1$ ), y  $u_{t0}$  y  $u_{t1}$  son normales,  $u_{t0} \sim N(0, \sigma_0^2)$ ,  $u_{t1} \sim N(\mu_1, \sigma_1^2)$ . Nótese que aparte de los parámetros a estimar de la ecuación de transición, se agregan los parámetros de la mixtura de normales.

La ventaja que ofrece representar el modelo de esta forma, es que se facilita la estimación vía el filtro de Kalman, permitiendo obtener un filtraje de Kalman robusto (Shumway y Stoffer 2001). Además, la librería SVFilter del package “asta” justo implementa esta aproximación.

Las ecuaciones del modelo de espacio de estado especificado en las ecuaciones (3.3) y (3.7), dan lugar a las ecuaciones del filtrado de Kalman (Shumway y Stoffer 2001).

$$\begin{aligned} x_{t+1|t} &= E(x_{t+1}|y_1, \dots, y_t) = \sum_{j=0}^1 E(x_{t+1}|y_1, \dots, y_t, I_t = j) \Pr(I_t = j|y_1, \dots, y_t) \\ &= \sum_{j=0}^1 (\phi_0 + \phi_1 x_{t|t-1} + K_{tj} \epsilon_{tj}) \pi_{tj} \end{aligned}$$

$$= \phi_0 + \phi_1 x_{t|t-1} + \sum_{j=0}^1 \pi_{tj} K_{tj} \epsilon_{tj}$$

donde  $I_t = j$  siendo  $j = (0,1)$ .

$$P_{t+1|t} = \phi_1^2 P_{t|t-1} + \sigma_w^2 - \sum_{j=0}^1 \pi_{tj} + K_{tj}^2 \Sigma_{tj}$$

$$e_{t0} = y_t - \alpha - x_{t|t-1}, \quad e_{t1} = y_t - \alpha - x_{t|t-1} - \mu_1$$

$$\Sigma_{t0} = P_{t|t-1} + \sigma_0^2, \quad \Sigma_{t1} = P_{t|t-1} + \sigma_1^2$$

$$K_{t0} = \frac{\phi_1 P_{t|t-1}}{\Sigma_{t0}}, \quad K_{t1} = \frac{\phi_1 P_{t|t-1}}{\Sigma_{t1}}$$

Además, para poder completar el filtrado, se ha de ser capaz de evaluar las probabilidades:

$$\pi_{t1} = \Pr(I_t = 1 | y_1, y_2, \dots, y_t), \text{ para } t = 1, 2, \dots, T$$

$$\pi_{t0} = 1 - \pi_{t1}$$

Denotando la función de densidad condicional de  $y_t$  como  $f_j(t|t-1)$ , dada las observaciones anteriores  $y_1, \dots, y_{t-1}$ , tenemos que:

$$\pi_{t1} = \frac{\pi_1 f_1(t|t-1)}{\pi_0 f_0(t|t-1) + \pi_1 f_1(t|t-1)}$$

Se considera que la distribución de  $\pi_j$ , se ha especificado con anterioridad. Como se ha comentado, en el presente trabajo se asume  $\pi_1 = 0.5$ , y se aproxima la función de densidad condicional  $f_j(t|t-1)$  por una función de densidad normal, con media  $x_{t|t-1} + \mu_j$ , y varianza  $\Sigma_{tj}$  para  $j = (0,1)$ , y  $\mu_0 = 0$

Asumiendo una mixtura equilibrada ( $\pi_1 = 0.5$ ) y fijando la media de una de las componentes de la mixtura centrada en cero ( $\mu_0 = 0$ ), el vector de parámetros a estimar es  $\theta = \{\phi_0, \phi_1, \sigma_w, \mu_1, \sigma_0, \sigma_1\}$ , donde  $\phi_0, \phi_1, \sigma_w$ , son los parámetros de la ecuación de transición, y  $\mu_1, \sigma_0, \sigma_1$ , representan los parámetros de las componentes gaussianas.

La estimación de  $\theta$ , se realiza mediante el método de máxima verosimilitud, utilizando el logaritmo de la función de verosimilitud, que viene dada por la siguiente expresión:

$$\ln L_Y(\theta) = \sum_{t=1}^n \ln \left( \sum_{j=0}^1 \pi_j f_j(t|t-1) \right) \quad (3.8)$$

La maximización de la función de verosimilitud (3.8), una función del vector de parámetros  $\theta$ , se lleva a cabo utilizando el algoritmo numérico no lineal de cuasi Newton-Raphson. Este algoritmo está implementado en la función `optim` del paquete de R. Los valores iniciales usados corresponden a los encontrados en la literatura analizada, tomando como referencia trabajos realizados por diversos investigadores; ver código en Anexo C. En este trabajo, para realizar las estimaciones se usa el software estadístico R, versión 3.3.1.

### Resultados de la estimación

Aplicando el uso combinado del FK y el cuasi Newton-Raphson, se obtienen los valores estimados del vector de parámetros  $\theta = \{\phi_0, \phi_1, \sigma_z^2, \mu_1, \sigma_0^2, \sigma_1^2\}$ , presentados en la tabla 3.

Observamos que el valor estimado del parámetro que mide la volatilidad media ( $\phi_0$ ), es -0,0104, que no es estadísticamente significativo<sup>11</sup>. El parámetro  $\phi_1$ , que mide la persistencia de la volatilidad, es un valor extremadamente alto (0,9892), por lo que la volatilidad de la serie es altamente persistente, pero es un proceso estacionario.

La desviación ( $\sigma_z$ ) de los residuos de la ecuación de estado es 0,1526 con un error estándar de 0,0228. La media de una de las mixturas es igual a -7,9343, mientras que las desviaciones típicas de las mixturas son -2.460 y 2.015 respectivamente.

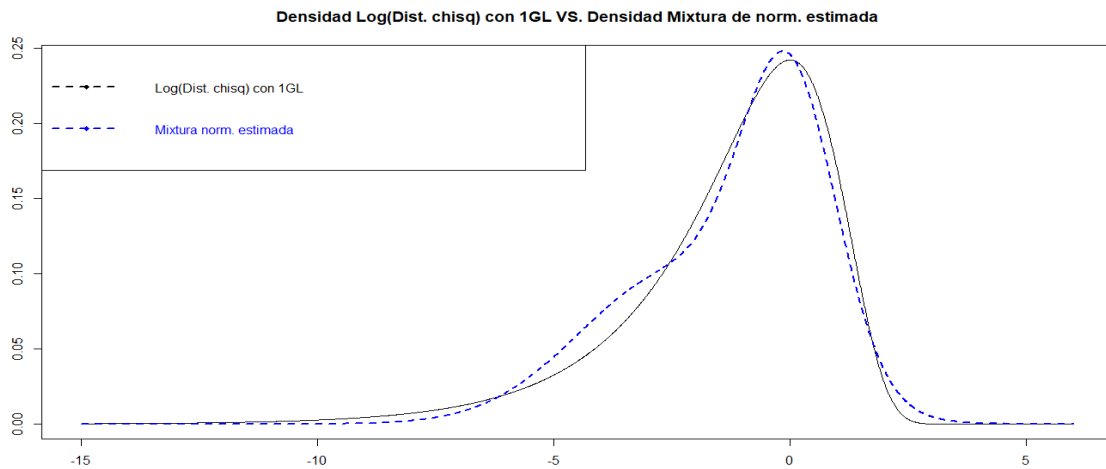
**Tabla 3. Resultados de la estimación de los parámetros**

Parámetro	Valor estimado	Error Estándar
$\phi_0$	-0,0104	0,0167
$\phi_1$	0,9892	0,0044
$\sigma_z$	0,1526	0,0228
$\alpha$	-7,9343	1,4508
$\mu_1$	0,1004	0,0391
$\sigma_0$	-2,4604	0,1154
$\sigma_1$	2,0154	0,0643

En la Figura 4, se observa que la aproximación de la función de densidad del  $\log(\chi_1^2)$ , a través de una mixtura de normales equilibrada con  $\pi=0.5$ , se ajusta relativamente bien, aunque en la cola superior no acaba de aproximar de todo bien la función de densidad del  $\log(\chi_1^2)$ , ya que los valores de las probabilidades obtenidas con la mixtura son más altos.

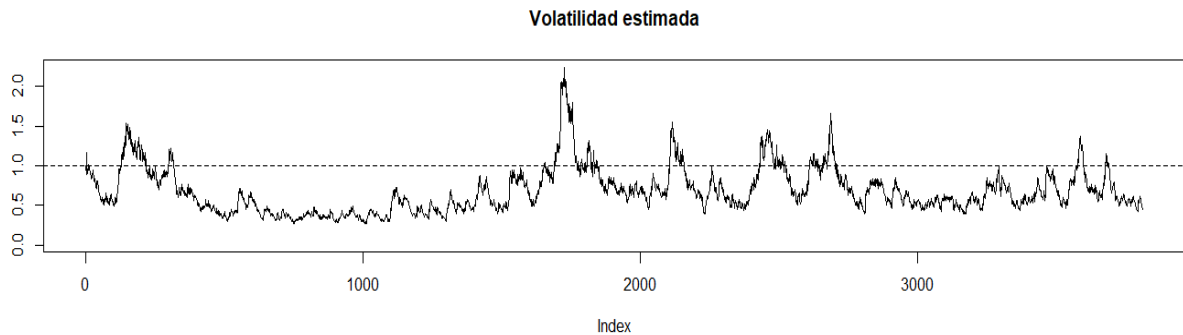
<sup>11</sup> -0,01035/0,01672 < 2

**Figura 4. Comparación de la función de densidad del  $\log(\chi_1^2)$  y La función de densidad estimada a través de la mixtura de normales con  $\pi=0.5$**



En la Figura 5, donde se grafica la volatilidad estimada ( $\sigma_t$ ), se observa que en la mayor parte del periodo estimado, esta se ha situado por debajo de la volatilidad no condicional, además, se observa que los periodos de mayor volatilidad, han sido los años de crisis financiera.

**Figura 5. Volatilidad estimada  $\sigma_t$  mediante FK + Newton Raphson**



En la Figura 6, se representa el  $\log(r_t^2)$  y la volatilidad estimada, se observa que el modelo capta bastante bien la característica de varianza cambiante que se encuentra en las series financieras. La volatilidad estimada se ajusta bien a los movimientos que se han producido en la variación de los rendimientos.

Si se observa conjuntamente la volatilidad estimada y los rendimientos del índice bursátil ( $r_t$ ) (Figura 7), se puede apreciar mejor el ajuste, ya que los rendimientos están en su forma original. Por lo que viendo este gráfico y los anteriores, se puede comentar que el modelo es útil a la hora de estimar el comportamiento de la volatilidad de esta serie financiera.

Figura 6. Gráfico de  $\log(r_t^2)$  y volatilidad estimada

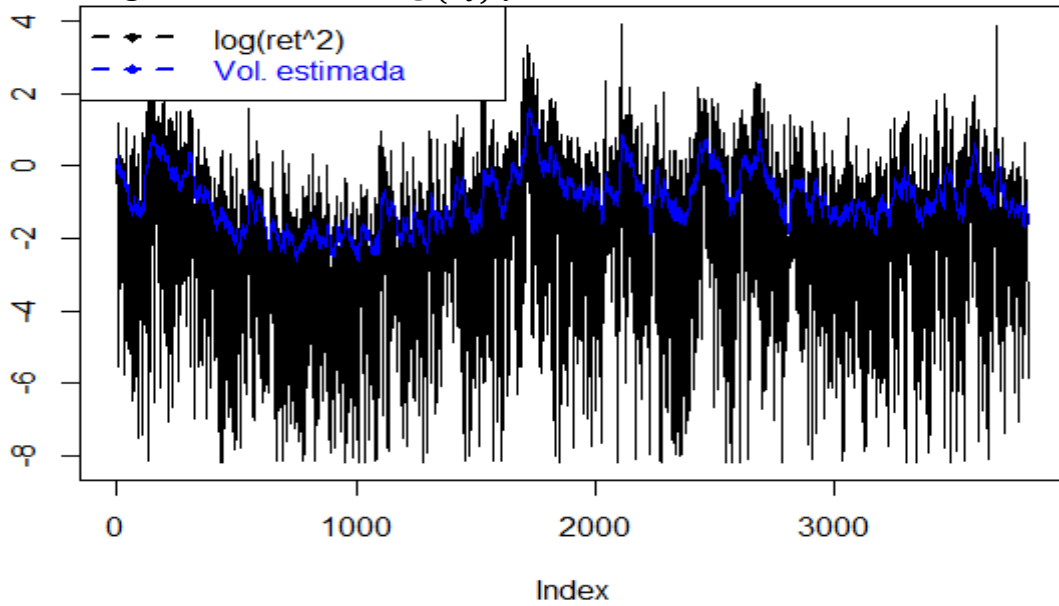
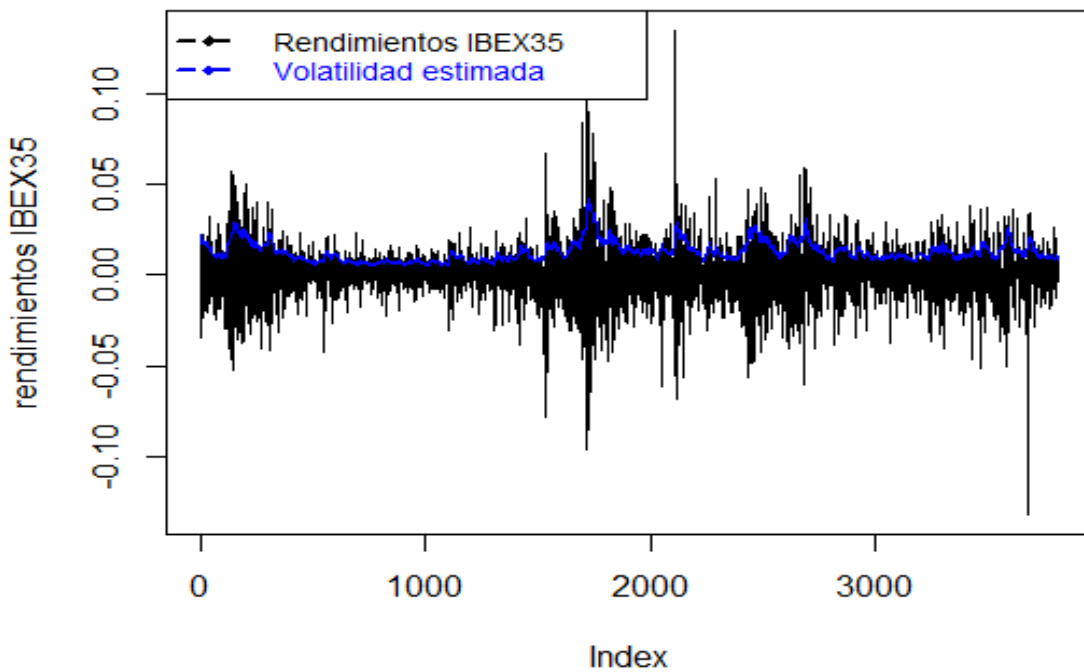


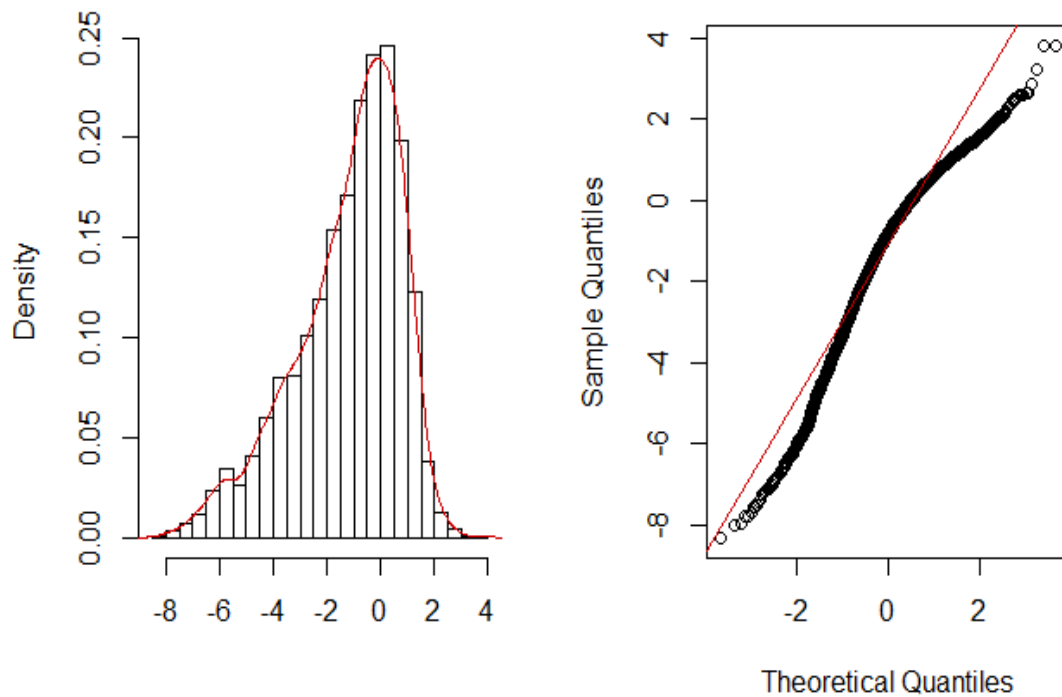
Figura 7: Gráfico de rendimientos del IBEX35 ( $r_t$ ) y la volatilidad estimada



Por último, realizando un análisis gráfico de los residuos, en la Figura 8 se observan que estos no se comportan según una normal, tal como se esperaba. De hecho el  $\log(\epsilon_t^2) \sim \log(\chi_1^2)$  y se ha aproximado mediante una mezcla de normales equilibrada  $\pi=0.5$ .



**Figura 8. Histograma y Q-Q plot de residuos: KF + Newton Raphson**



### 3.3.- Estimación de la volatilidad y parámetros vía filtro de partículas de Liu y West

En este apartado realizaremos la estimación de la volatilidad, mediante el filtro de partículas de Liu y West (2001), descrito en el capítulo 2.

En este caso, no es necesario linearizar la ecuación de observación original ya que el filtro de partículas no tiene restricciones de linealidad y/o normalidad. Por tanto, se usa el modelo básico de volatilidad estocástica SARV(1) en su forma original, dado por las siguientes expresiones:

**Ecuación de observación:** 
$$r_t = \exp\left(\frac{x_t}{2}\right) \cdot \epsilon_t = \sigma_t \cdot \epsilon_t, \quad \epsilon_t \sim WN(0,1)$$

donde,  $\sigma_t$  es una variable latente y  $\epsilon_t$ , es una variable aleatoria que se distribuye usualmente según una normal estándar. En este apartado no estimaremos  $\sigma_*$ , como se hizo en el filtro de Kalman, ya que la implementación usada del filtro de partículas, no lo considera. Como se mencionó anteriormente, la estimación de  $\sigma_*$  es opcional.

A continuación, se presenta la ecuación de estado aumentado  $l_t = (x_t, \theta)' = (\mu, \phi_1, \sigma_z)'$ , que incorpora además de la volatilidad latente los parámetros del modelo a estimar.

## Ecuación de estado

Adaptando la ecuación (2.26) al modelo SARV(1), la ecuación de estado queda expresada de la siguiente forma:

$$l_t = \begin{bmatrix} x_t \\ \theta \end{bmatrix} = \tilde{f}(l_{t-1}, z_t) = \begin{bmatrix} f(x_{t-1}, z_t) \\ \theta \end{bmatrix} = \begin{bmatrix} \mu(1 - \phi_1) + \phi_1 x_{t-1} + z_t \\ \mu \\ \phi_1 \\ \sigma_z \end{bmatrix}$$

donde, como antes,  $\phi_0 = \mu(1 - \phi_1)$  y la ecuación del estado original  $x_t$  puede reescribirse como  $x_t = \phi_0 + \phi_1 x_{t-1} + z_t$ .

Una vez formulado el modelo de espacio de estado, se procede a especificar las distribuciones a priori para la volatilidad latente y los parámetros del modelo asociados.

Las distribuciones a priori que se utilizan en este apartado, se escogen en base a la literatura analizada, tomando como referencia trabajos realizados por diversos investigadores.

## Distribuciones a priori

Para cada parámetro del vector  $\theta$ , elegimos una componente independiente, es decir  $p(\mu), p(\phi_1), p(\sigma_z)$  (Kim et al. 1998; Acosta L. 2013)

Para el vector de estado inicial,  $x_0 \sim N(\mu_{x_0}, \Sigma_{x_0})$ , al igual que para el parámetro  $\mu_0 \sim N(b_\mu, B_\mu)$ , con hiperparámetros  $\mu_{x_0}=0, \Sigma_{x_0}=\sigma_{z_0}^2, b_\mu=-8$  y  $B_\mu=25$ .

Para el parámetro que mide la persistencia,  $\phi_1 \in (-1,1)$ , se escoge una distribución beta  $\frac{\phi_1+1}{2} \sim B(a_0, b_0)$ , (Kastner G 2016); se usan los valores de los hiperparámetros  $a_0=18$  y  $b_0=1$  (Pitt y Shephard 1998b), con valor inicial es de 0,9702. Kastner comenta que si el número de observaciones de los datos financieros es demasiado pequeño (<1.000), la elección de los hiperparámetros  $(a_0, b_0)$  pueden tener una alta influencia en la forma de la distribución posterior del parámetro  $\phi_1$ .

Para la varianza se utilizará una distribución inversa gamma  $\sigma_{z_0}^2 \sim IG(c_0, d_0)$ , con hiperparámetros  $c_0=5$  y  $d_0=1/20$  (Pitt y Shephard 1999b), y valor inicial  $\sigma_{z_0}^2=0,03168$

La tabla 4, contiene un resumen de las distribuciones a priori de los parámetros del modelo SARV(1) usadas en el presente trabajo, juntamente con los valores iniciales o de arranque.

**Tabla 4. Resumen de las distribuciones a priori y valores iniciales**

Parámetros	Hiperparámetros		Valores iniciales
$\mu_0 \sim N(\mathbf{b}_\mu, \mathbf{B}_\mu)$	$b_\mu=-8$	$B_\mu=5$	$-8, 5^2$
$\phi_1 \sim 2B(\mathbf{a}_0, \mathbf{b}_0) - \mathbf{1}$	$a_0=18$	$b_0=1$	0,9702
$\sigma_z^2 \sim IG(c_0, d_0)$	$c_0=5$	$d_0=1/20$	0,03168

Para la estimación conjunta del vector de parámetros y del estado (la volatilidad), utilizaremos el filtro de partículas de Liu y West (2001) implementado en la tesis doctoral de Acosta L. (2013) y también en el paquete de la librería nimble (para hacer una comparación entre los resultados obtenidos), se usan  $N_p=5000$  partículas y para el valor de descuento  $\delta$  utilizaremos 0,95.

### **Funcionamiento del paquete Nimble**

Nimble es un paquete implementado en R por Michaud et al (2017), que permite definir un modelo en espacio de estado y de forma separada permite compilarlo mediante C++, sin necesidad de programar en dicho lenguaje; los autores comentan que ello hará que la ejecución del proceso sea un poco más rápida.

El paquete pretende ser flexible, ya que proporciona un sistema en el que se puede programar algoritmos genéricos en lenguaje BUGS<sup>12</sup> y aplicarlos a cualquier modelo especificado en dicho lenguaje. Particularmente nimble permite la estimación e inferencia de modelos en forma de espacio-estado utilizando la metodología de los filtros de partículas como el SIS, el SIR, el ASIR o el filtro de partículas de Liu y West (2001) aplicado en este trabajo.

En el caso del filtro de Liu y West (2001), los autores del paquete han implementado el algoritmo inicial que propusieron dichos autores (Liu y West). El pseudo código implementado en el paquete es el siguiente:

<sup>12</sup> Bayesian inference Using Gibbs Sampling, proyecto que nació en 1989 y que puede ejecutarse en Microsoft Windows, Linux.

---

**Algorithm 3** Liu and West filter

---

```
1: for  $k$  in  $1 : K$  do
2:   Generate  $x_0^{(k)} \sim p_0(x_0)$ 
3:   Generate  $\theta_0^{(k)} \sim p_0(\theta)$ 
4:   Set  $\pi_0^{(k)} = \frac{1}{K}$ 
5: end for
6: for  $t$  in  $1 : T$  do
7:   for  $k$  in  $1 : K$  do
8:     Compute  $x_{t|t-1}^{(k)} = E(x_t | x_{t-1}^{(k)}, \theta_{t-1}^{(k)})$ 
9:     Compute  $\theta_{t|t-1}^{(k)} = a\theta_{t-1}^{(k)} + (1-a)\bar{\theta}_{t-1}$ 
10:    Compute  $w_{t|t-1}^{(k)} = \pi_{t-1}^{(k)} g(y_t | x_{t|t-1}^{(k)}, \theta_{t|t-1}^{(k)})$ 
11:    Normalize  $w_{t|t-1}^{(k)}$  as  $\pi_{t|t-1}^{(k)} = \frac{w_{t|t-1}^{(k)}}{\sum_{i=1}^K w_{t|t-1}^{(i)}}$ 
12:    Sample an index  $j_k$  from the set  $\{1, \dots, K\}$  with probabilities
     $\{\pi_{t|t-1}^{(i)}\}_{i=1}^K$ 
13:    Sample  $\theta_t^{(k)} \sim N(\theta_{t|t-1}^{(j_k)}, h^2 V_{t-1})$ 
14:    Sample  $x_t^{(k)} \sim f(x_t | x_{t-1}^{(j_k)}, \theta_t^{(k)})$ 
15:    Calculate  $w_t^{(k)} = \frac{g(y_t | x_t^{(k)}, \theta_t^{(k)})}{g(y_t | x_{t|t-1}^{(j_k)}, \theta_{t|t-1}^{(j_k)})}$ 
16:    Normalize  $w_t^{(k)}$  as  $\pi_t^{(k)} = \frac{w_t^{(k)}}{\sum_{i=1}^K w_t^{(i)}}$ 
17:   end for
18: end for
```

---

Imagen tomada del artículo “Sequential Monte Carlo Method in the nimble R package” N Michaud N et Al. 2017

### Modelo a estimar con la librería Nimble

Para hacer comparables los resultados los resultados del algoritmo desarrollado por Acosta L. (2013) con los de la librería nimble, estimaremos el mismo modelo descrito en el apartado 3.3, con las mismas distribuciones a priori y los parámetros fijos que se pueden encontrar en la tabla 4. Utilizaremos el mismo número de partículas (5000) y el mismo factor de descuento (0,95). Resumiendo, el modelo de espacio de estado aumentado es el siguiente:

**Ecuación de observación:** 
$$r_t = \exp\left(\frac{x_t}{2}\right) \cdot \epsilon_t = \sigma_t \cdot \epsilon_t, \quad \epsilon_t \sim WN(0,1)$$

**Ecuación de estado:**

$$l_t = \begin{bmatrix} x_t \\ \theta \end{bmatrix} = \tilde{f}(l_{t-1}, z_t) = \begin{bmatrix} f(x_{t-1}, z_t) \\ \theta \end{bmatrix} = \begin{bmatrix} \mu(1 - \phi_1) + \phi_1 x_{t-1} + z_t \\ \mu \\ \phi_1 \\ \sigma_z^2 \end{bmatrix}$$

donde  $\phi_0 = \mu(1 - \phi_1)$ , por lo cual  $x_t = \phi_0 + \phi_1 x_{t-1} + z_t$

## Resultados de la estimación

La Tabla 5, contiene los resultados de la estimación de los parámetros (en tiempo final T) vía filtro de partículas, utilizando la implementación del filtro de partículas de Liu y West realizada por Acosta L. (2013) y mediante la librería nimble de Michaud et. al. (2017).

**Tabla 5. Resultados de la estimación del vector de los parámetros.  $N_p=5000$  y  $\delta=0,95$**

Algoritmo	Parámetro	Valor estimado	I.C*	$\Delta_{I.C}^{**}$
Acosta L. (2013)	$\mu$	-8,8515	(-9,2210, -8,4882)	0,7328
Nimble	$\mu$	-8,7972	----	----
Acosta L. (2013)	$\phi_0$	-0,0725	(-0,1080, -0,0482)	0,0598
Nimble	$\phi_0$	-0.0768	----	----
Acosta L. (2013)	$\phi_1$	0,9918	(0,9883, 0,9943)	0,0060
Nimble	$\phi_1$	0,9912	----	----
Acosta L. (2013)	$\sigma_z$	0,1465	(0,1264, 0,1679)	0,0415
Nimble	$\sigma_z$	0,1374	----	----

\* Intervalo de credibilidad (percentiles 2,5 y 97,5) \*\* Longitud del intervalo de credibilidad

La primera diferencia práctica que se detecta en la estimación realizada con la librería nimble, es que, aunque proporciona la evolución del estado estimado a lo largo del tiempo, no lo hace en el caso de los parámetros, solo proporciona la estimación de los parámetros en el momento final. En cambio con la implementación de Acosta L. (2013) podemos obtener la evolución tanto del estado estimado como de los parámetros en cualquier punto de tiempo<sup>13</sup>.

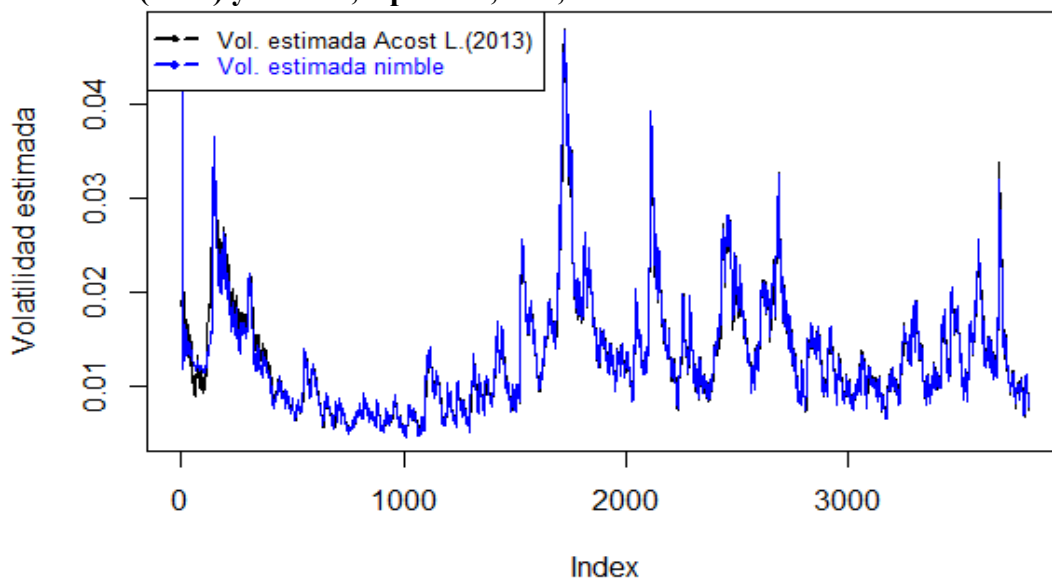
Otra diferencia relevante entre ambas implementaciones, es que nimble no provee los intervalos de credibilidad y su costo computacional es mucho mayor, tal como se reporta más adelante.

<sup>13</sup> El gráfico de la evolución de los parámetros en el tiempo, lo encontramos en el apartado de anexos

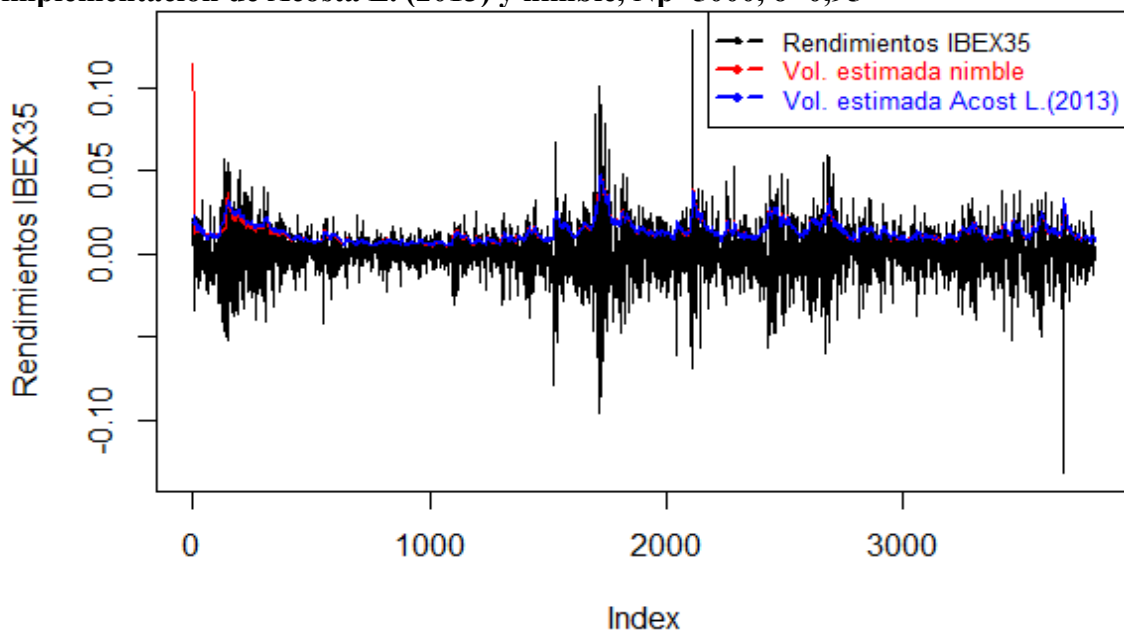
En base a los resultados de la estimación en Tabla 5, con  $N_p=5000$  y  $\delta=0,95$ , se obtiene que ambas implementaciones arrojan valores estimados de los parámetros similares. Se observa que el valor estimado de  $\sigma_z$  con nimble es algo menor que el de la implementación de Acosta L. (2013) que a su vez es más parecido al valor estimado mediante el filtro de Kalman+Newton Raphson.

Respecto a la estimación de la variable latente (volatilidad), ver Figuras 9 y 10, se observa que en los primeros periodos hay diferencias entre ambas estimaciones, pero que ésta va desapareciendo a medida que el proceso de estimación avanza.

**Figura 9. Comparación de la estimación de la volatilidad con la implementación de Acosta L. (2013) y nimble,  $N_p=5000$ ,  $\delta=0,95$**

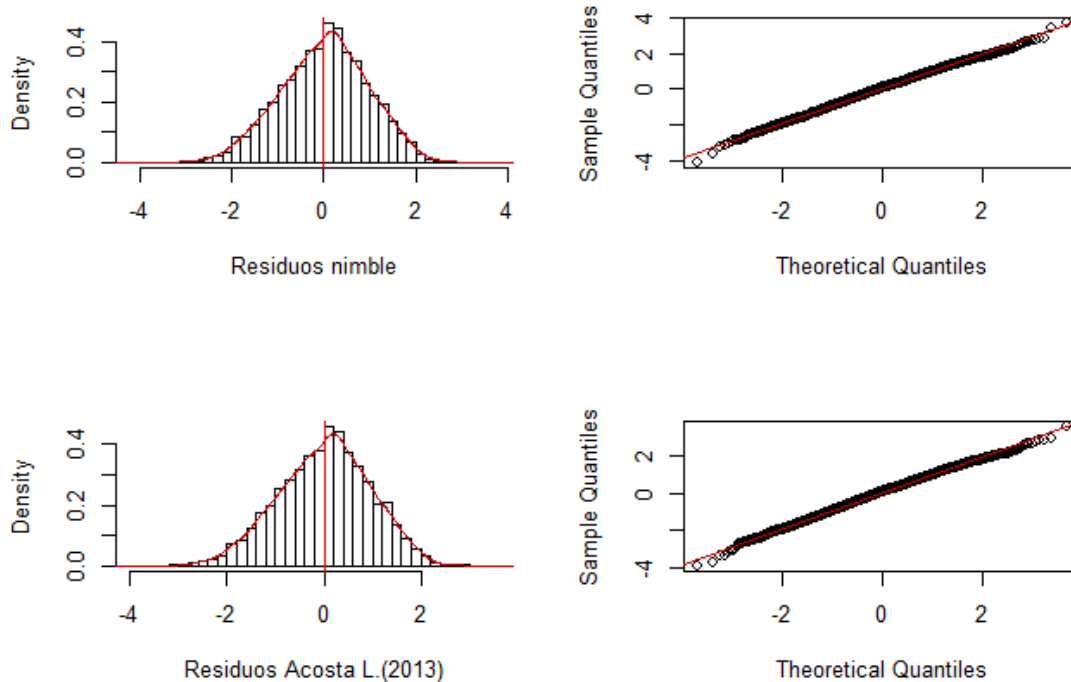


**Figura 10. Rendimientos del IBEX 35 y volatilidad estimada con la implementación de Acosta L. (2013) y nimble,  $N_p=5000$ ,  $\delta=0,95$**



Por último, en la Figura 11, se analizan los residuos del modelo SARV(1) con  $N_p=5000$  y  $\delta=0,95$ , en la que se puede observar que en ambas implementaciones los residuos tienden a distribuirse más o menos según una normal estándar, tal como está especificado en la ecuación de observación de los rendimientos; obsérvese aún la presencia de algunos atípicos

**Figura 11. Histograma y QQ-plot de residuos del modelo SARV(1) con la implementación de Acosta L. y nimble.  $N_p=5000$ ,  $\delta=0,95$**



### **Comparación de la estimación del modelo SARV(1): filtro de Kalman + Newton Raphson versus el filtro de partículas de Liu y West (2001)**

Una de las principales diferencias que encontramos en las metodologías utilizadas en el presente trabajo, es el coste computacional existente en la estimación de la volatilidad, y parámetros asociados al modelo SARV(1). Para el filtro de partículas se usan  $N_p=5000$  y  $\delta=0,95$ .

Utilizando la implementación realizada por Acosta L. y la librería nimble, apreciamos que el coste computacional del paquete nimble es mucho más alto; 2 minutos versus 40 minutos, respectivamente. Si comparamos el coste computacional de la implementación del filtro de Kalman + Newton Raphson, con el filtro de partículas de Liu y West

implementado por Acosta L., la diferencia existente es de segundos, por lo que podemos decir que el coste computacional es muy similar. Es decir, la ganancia computacional es mínima entre ambas implementaciones (filtro de Kalman y filtro de partículas implementado por Acosta L.).

<b>Tiempo (CPU)</b>	<b>Filtro de Kalman</b>	<b>PF (Acosta L.)</b>	<b>FP (nimble)</b>
<b>Minutos</b>	1,430	2,045	40,319

Respecto a la estimación de los parámetros, el valor del parámetro que mide la volatilidad media,  $\phi_0$ , con el filtro de Kalman es de -0,01035 (no es estadísticamente significativo), mientras que con el filtro de Liu y West, el valor es más alto, situándose alrededor de -0,07252 y -0,07675 respectivamente según la implementación realizada (Tabla 6).

Tanto con la metodología del filtro de Kalman como con el filtro de partículas, se obtiene que la volatilidad de la serie es altamente persistente con valores estimados de  $\phi_1$ , cerca de la unidad, siendo el valor estimado vía FK algo menor (0,989) que el estimado con el filtro de Liu y West (aprox. 0,992).

Por último, en el caso de la estimación de  $\sigma_z$ , apreciamos el valor estimado de  $\sigma_z$  con el filtro de Liu and west utilizando la implementación de Acosta se aproxima más al valor estimado con el filtro de Kalman, mientras que con la implementación de nimble vemos que el valor es menor. Se destaca que las estimaciones obtenidas mediante la metodología de filtro de partículas son más precisas, ver Tabla 6.

En la Figura 12, se realiza una comparación de la volatilidad estimada con ambas metodologías, se puede apreciar que la volatilidad estimada con el filtro de Kalman y el filtro de partículas es muy similar. En el caso de la volatilidad estimada con el filtro de partículas mediante nimble, se observa que en los primeros periodos la volatilidad tiende a ser más alta que la estimada con las otras 2 implementaciones (Figura 13). En términos generales la estimación de la volatilidad es muy similar

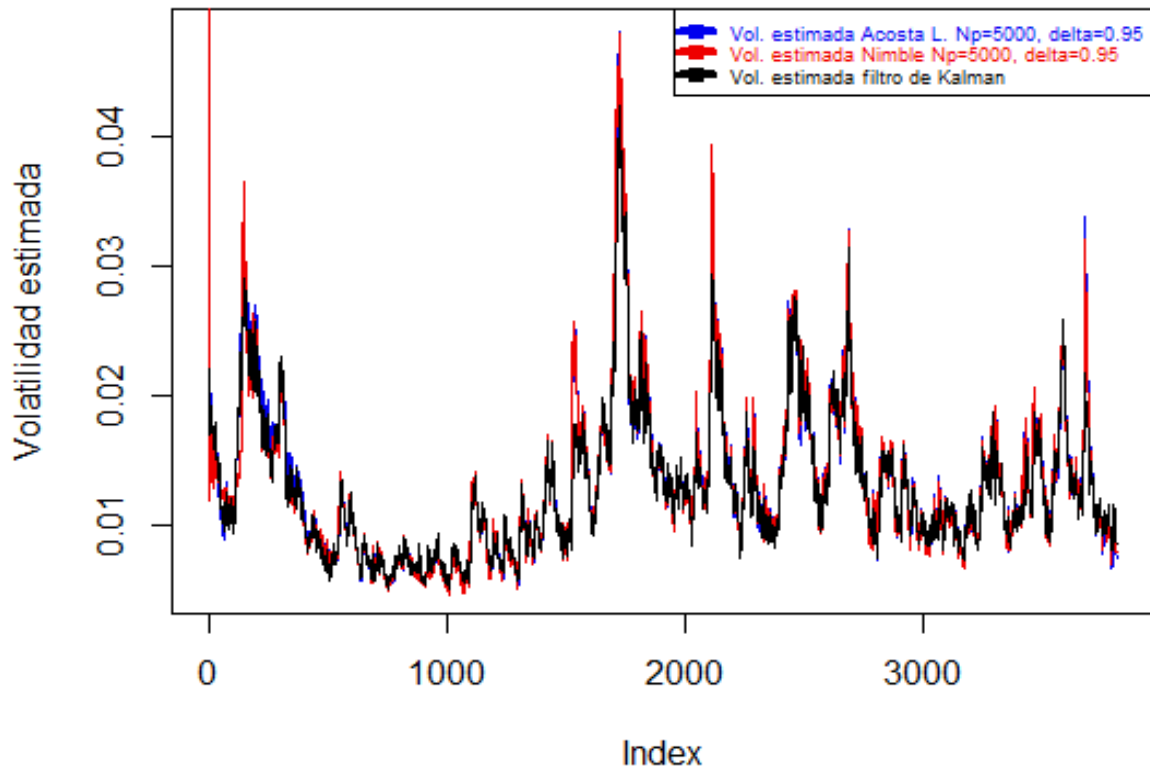


**Tabla 6. Comparación de las estimaciones obtenidas vía filtro de Kalman + Newton Raphson y el filtro de partículas de Liu y West:  $N_p=5000$ ,  $\delta =0.95$**

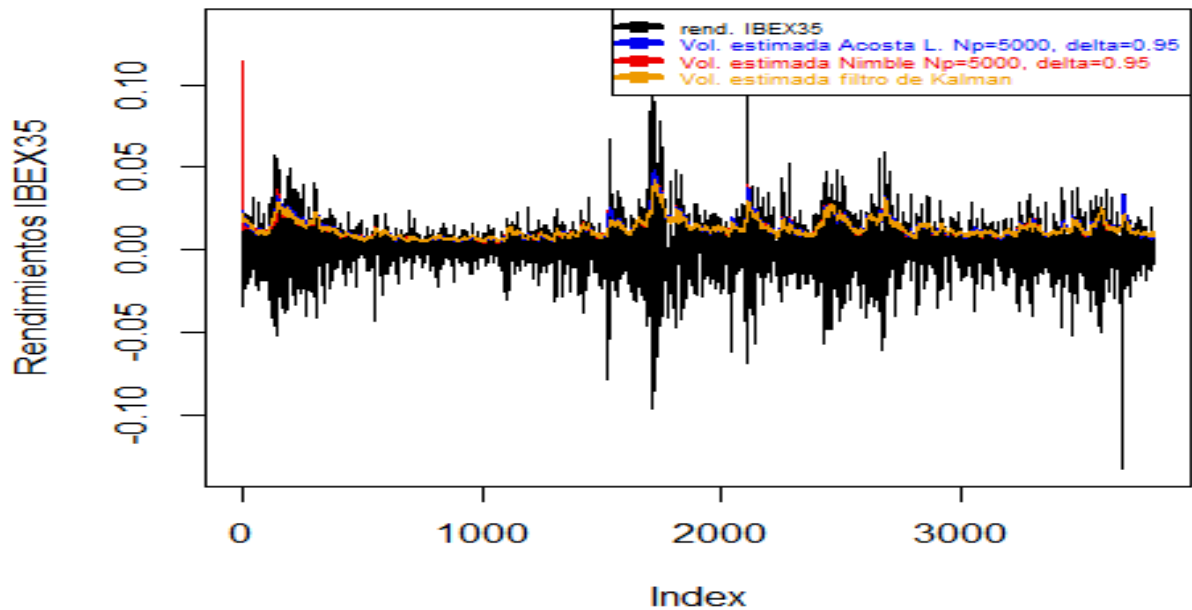
Parámetros	Filtro de Kalman	Filtro de partículas de Liu y West (2001)	
		Implement. Acosta L. (2013)	Nimble
$\phi_0$	-0,0104 (-0,0438, 0,0231)* [0,0669]***	-0,0725 (-0,1080, -0,0482)** [0,0598]	-0,0768 ---
$\phi_1$	0,9892 (0,9804, 0,9980) [0,0176]	0,9918 (0,9883, 0,9943) [0,0060]	0,9913 ---
$\sigma_z$	0,1526 (0,1071, 0,1982) [0,0911]	0,1465 (0,1264, 0,1679) [0,0415]	0,1375 ---

\* Intervalos de confianza \*\* intervalos de credibilidad \*\*\* Longitud de los intervalos

**Figura 12. Comparación de la volatilidad estimada con el filtro de Kalman y el filtro de partículas.**



**Figura 13. Rendimientos del IBEX35 versus volatilidad estimada vía filtro de Kalman y el filtro de partículas de Liu y West.**



Comentado lo anterior, se puede decir que el patrón de la varianza cambiante es captado de manera similar por ambas metodologías, siendo la diferencia más relevante entre metodologías el coste computacional a la hora de estimar la volatilidad. Nótese que en caso de implementación con librería nimble y filtro de Kalman + Newton raphson solo se obtienen los valores estimados de los parámetros al final de la serie y por tanto no podemos comparar su evolución. En el caso de implementación de Acosta L. se obtiene toda la evolución de los parámetros estimados y también su distribución a posteriori (ver Figuras 1B y 2B en Anexo B ).

#### 4. Conclusiones y discusión

En el presente trabajo, el objetivo principal planteado es estimar la volatilidad latente (y parámetros asociados) en la serie financiera IBEX35. Se propone realizar dicha estimación en el contexto de los modelos de espacio de estado, mediante el modelo de volatilidad estocástica SARV(1) y dos metodologías de estimación: el uso combinado del filtro de Kalman + máxima verosimilitud y el método Monte Carlo secuencial conocido como el filtro de partículas (bayesiano). Específicamente, se usa la variante de filtro de partículas de Liu y West (2001) que permite estimar simultáneamente la volatilidad y los parámetros asociados.

Se plantea asimismo, realizar luego una comparativa utilizando una proxy de la volatilidad (basada en el logaritmo de los rendimientos al cuadrado) para evaluar la capacidad de cada uno de los métodos para realizar estimaciones los más eficientes y fiables posibles. En caso de los parámetros, se reportan los valores estimados puntuales más el intervalo de confianza/credibilidad y la longitud del intervalo correspondiente.

En el caso del filtro de partículas de Liu y West (2001), en un primer momento se plantea solo utilizar la librería nimble, pero al percatarnos de algunas limitaciones<sup>14</sup>, se decidió también utilizar la implementación realizada por Acosta L. (2013) en su tesis doctoral; comparando en la medida de lo posible los resultados obtenidos mediante ambas implementaciones.

Tomando la serie del índice bursátil IBEX35, se realiza la estimación con las dos metodologías mencionadas y se destacan las siguientes diferencias o similitudes:

Entre ambos métodos, la diferencia en el costo computacional es negligible. El tiempo que se tarda en estimar la volatilidad mediante el filtro de Kalman + máxima verosimilitud es de 1,43 minutos, mientras que con el filtro de partículas de Liu y West utilizando la implementación de Acosta L. el tiempo es de 2,05 minutos. En cambio, el coste computacional con la librería nimble, es muy alto, alrededor de 40 minutos.

Otra diferencia a comentar es que con la implementación de la librería nimble y filtro de Kalman + Newton raphson solo se obtienen los valores estimados de los parámetros al final de la serie y por tanto no es posible comparar su evolución. Por otro lado, con la

---

<sup>14</sup> Se escribió a los autores del paquete para preguntar, por ejemplo, si era posible extraer la evolución de los parámetros y cómo, pero no se obtuvo respuesta.

implementación realizada por Acosta L., se obtiene la evolución de los parámetros estimados, así como también su distribución a posteriori.

Tanto con el filtro de Kalman como con el filtro de partículas se observa que la estimación de la volatilidad es bastante similar. Las estimaciones obtenidas con el filtro de partículas son algo más precisas como se puede apreciar en la Tabla 6.

En la práctica, implementar el filtro de Kalman puede resultar más sencillo que implementar el filtro de partículas, ya que el último requiere de un mayor conocimiento experto, como por ejemplo, en la elección de la densidad de importancia o del método de remuestreo. Ambos métodos requieren ciertas condiciones iniciales, pero en filtro de partículas es algo más complicado seleccionar las distribuciones a priori para el estado y cada uno de los parámetros que intervienen en el modelo.

Otro inconveniente potencial que presenta todo filtro de partículas, incluyendo el de Liu y West implementado en este trabajo, es el problema de la degeneración y empobrecimiento de las partículas. Ello conlleva a la necesidad de monitorizar el llamado ESS para garantizar que las estimaciones obtenidas sean fiables. Todo esto requiere conocimiento experto por parte del investigador.

Cabe destacar que la metodología del filtro de partículas es mucho más flexible, ya que a diferencia del filtro de Kalman, no hay requerimientos de linealidad y/o normalidad. Además, al proveer la distribución a posteriori de la volatilidad y los parámetros, hace posible calcular cualquier característica estadística de interés.

En el caso del filtro de Kalman, al requerirse linealidad y normalidad, hizo necesario el aproximar el término de error de la ecuación de observación mediante una mixtura de dos normales. Muchas veces esta aproximación no acaba de ajustarse del todo bien como se puede apreciar Figura 4.

Como conclusión final, podemos subrayar que ambos métodos son buenas herramientas para la estimación de la volatilidad, por lo que ambos filtros pueden ser utilizados de forma complementaria para estimar tanto la volatilidad latente en el índice bursátil IBEX35 como los parámetros incluidos en el modelo estocástico SARV(1) ajustado.

En el caso delo modelo SARV(1), algunos investigadores pueden preferir el uso del filtro de partículas frente al filtro de Kalman, al considerarlo más flexible. Otros sin embargo pueden decantarse por el filtro de Kalman al considerarlo de más fácil aplicación.

## **Bibliografia**

Acosta Argueta, L. (2013). Particle Filtering Estimation for Linear and Nonlinear State-Space Models. Universitat Politècnica de Catalunya.

Andersen, T., T. Bollerslev, F. Diebold, and H. Ebens (2001). The distribution of realised stock return volatility. *Journal of Financial Economics* 61(1), 43–76.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), pp.307-327.

Arulampalam, S., S. Maskell, N. Gordon, and T. Clapp (2002). A tutorial on particle filters for online non-linear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing* 50(2), 174–188.

Bollerslev, T., R. Chou, and K. Kroner (1992). Arch modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics* 52(1-2), 5–59.

Black, F. (1976). Studies in stock price volatility changes. In *American Statistical Association, Proceedings of the Business and Economic Statistics Section*, pp. 177–181.

Carpenter, J., P. Clifford, and P. Fearnhead (1999). An improved particle filter for non-linear problems. *IEEE Proceedings on Radar and Sonar Navigation* 146(1), 2–7.

Doucet, A. (1998). On sequential simulation-based methods for Bayesian filtering. Technical Report CUED/F-INFENG/TR.310, Department of Engineering, University of Cambridge.

Doucet, A., S. Godsill, and C. Andrieu (2000). On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing* 10(3), 197–208.

Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of the United Kingdom inflation. *Econometrica* 50(4), 987–1007.

Fama, E. (1965). The behavior of stock market prices. *Journal of Business* 38(1), 34–105.

Geweke, J. (1989). Bayesian inference in econometrics models using Monte Carlo integration. *Econometrica* 57, 1317–1339.

Ghysels, E., A. Harvey, and E. Renault (1996). Stochastic volatility. Technical Report 95s-49, CIRANO: Centre Interuniversitaire de Recherche en Analyse des Organisations, Montréal.

- Gordon, N. J., D. Salmond, and A. Smith (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE. Proceedings-F* 140(2), 107–110.
- Jacquier, E., N.G. Polson, and P.E Rossi (1994). Bayesian analysis of stochastic volatility models. *Journal of Business & Economic Statistics* 12(4), 69–87
- Kastner, G. (2016). Dealing with Stochastic Volatility in Time Series Using the R Package stochvol. *Journal of statistical software*, Volume 69.
- Kim, S., N. Shephard, and S. Chib (1998). Stochastic volatility: likelihood inference and comparison with ARCH models. *Review of Economic Studies* 65, 361–393.
- Kitagawa, G. (1998). Self-organizing state-space model. *Journal of the American Statistical Association* 93, 1203–1215.
- Kitagawa, G. (1996). Monte Carlo filter and smoother for non-Gaussian nonlinear state-space model. *Journal of Computational and Graphical Statistics* 5, 1–25.
- Kitagawa, G. and S. Sato (2001). Monte Carlo smoothing and self-organizing state-space model. In A. Doucet, d. N., and D. Gordon (Eds.), *Sequential Monte Carlo Methods in Practice*, pp. 177–195. Springer Verlag.
- Le Roux, J. (2003) “An Introduction to Kalman filter”, University of Nice.
- Liu, J. and M. West (2001). Combined parameter and state estimation in simulation-based filtering. In A. Doucet, N. de Freitas, and D. Gordon (Eds.), *Sequential Monte Carlo Methods in Practice*, pp. 197–223. Springer Verlag.
- Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business* 36(4), 394–419.
- Muñoz, M. P. (1988). Estimació dels paràmetres de models ARMA(p,q) mitjançant algorismes de filtratge òptim, tesi doctoral. Technical report, Universitat Politècnica de Catalunya, Barcelona, España.
- Muñoz, M. P., M. D. Márquez, and L. M. Acosta (2007). Forecasting volatility by means of threshold models. *Journal of Forecasting* 26, 343–363.
- Márquez, M. D. (2002). Modelo SETAR aplicado a la volatilidad de la rentabilidad de las acciones: algoritmos para su identificación. Ph. D. thesis, Universitat Politècnica de Catalunya, Barcelona España.

- Márquez, M. D., M. P. Muñoz, C. Villazón, M. Martí-Recober, and L. M. Acosta (2005). Rendimiento y volatilidad del IBEX 35: Capturando las asimetrías y el exceso de curtosis. Technical Report DR 2005/1, Universitat Politècnica de Catalunya.
- Michaud M., de Valpine, P., Turek, D., and Paciorek, C. (2017). Sequential Monte Carlo Methods in the nimble R Package. Retrieved from <https://arxiv.org/pdf/1703.06206.pdf>
- Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica: Journal of the Econometric Society*, 59(2), pp.347–370.
- Pellegrini, S. (2009). Predicción en modelos de componentes inobservables condicionalmente heteroscedásticos. Ph. D. thesis, Universidad Carlos III de Madrid, Departamento de Estadística.
- Pitt, M. K. and N. Shephard (1999). Filtering via simulation: auxiliary particle filters. *Journal of the American Statistical Association* 94, 590–599.
- Ristic, B., Arulampalam, S. and Gordon, N. (2004). Beyond the Kalman filter. Boston, Mass.: Artech House.
- Stock, J. and M. Watson (2007). Why has U.S. inflation become harder to forecast? *Journal of Money, Credit and Banking* 39, 13–33.
- Sandmann, G. and S.J. Koopman (1998). Estimation of stochastic volatility models via Monte Carlo maximum likelihood. *J. Econometrics*, 87, 271-301.
- Shumway, R. (1988). Applied Statistical Time Series Analysis. New Jersey: Prentice Hall.
- Shumway, R. and Stoffer, D. (2011). Time series analysis and its applications. New York: Springer.
- Tanizaki, H. (1991). Nonlinear Filters: Estimation and Applications. Ph. D. thesis, University of Pennsylvania.
- Taylor, S. (1986). Modeling Financial Time Series. New York: John Wiley.
- Taylor, S. (1994). “Modeling stochastic volatility: a review and comparative study”. *Mathematical Finance*, 4, pp.183-204.
- Tsay, R. (2014). An Introduction to Analysis of Financial Data with R. Somerset: Wiley

## Anexos

### A. Tablas

**Tabla 1A. Test de hipótesis sobre la media ( $H_0:\mu=0$ ) y 2 tests de normalidad de la serie de rendimientos de la cotización del índice bursátil IBEX35**

Hipótesis nula	Valor estadístico	P-valor
$H_0: \mu=0$ (Test de Wilcoxon para la media)	37179500	0,1873
$H_0$ : Normalidad (Test de Shapiro Wilk)	0,9427	0,0000
$H_0$ : Normalidad (Test de Jarque Bera)	6772,4	0,0000

**Tabla 2A. Test de hipótesis de correlación de Ljung-Box**  
**Test Ljung-Box  $H_0: \rho=0$  (no correlación)**

Retardos	Valor estadístico	P-valor
Lag=1	0,3549	0,5513
Lag=5	24,989	0,0001
Lag=10	34,882	0,0001
Lag=15	47,385	0,0000
Lag=20	58,347	0,0000

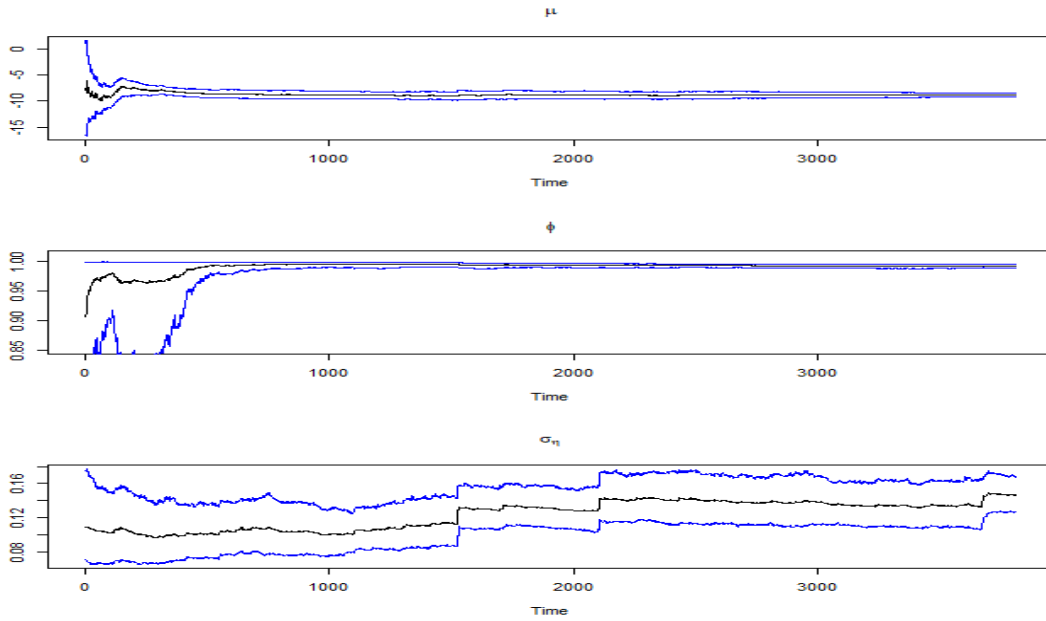
El estadístico del test de Jung-Box viene definido por la siguiente expresión:

$$LB = n(n + 2) \sum_{i=1}^m \frac{\hat{\rho}_k^2}{n - k} \sim \chi_m^2$$

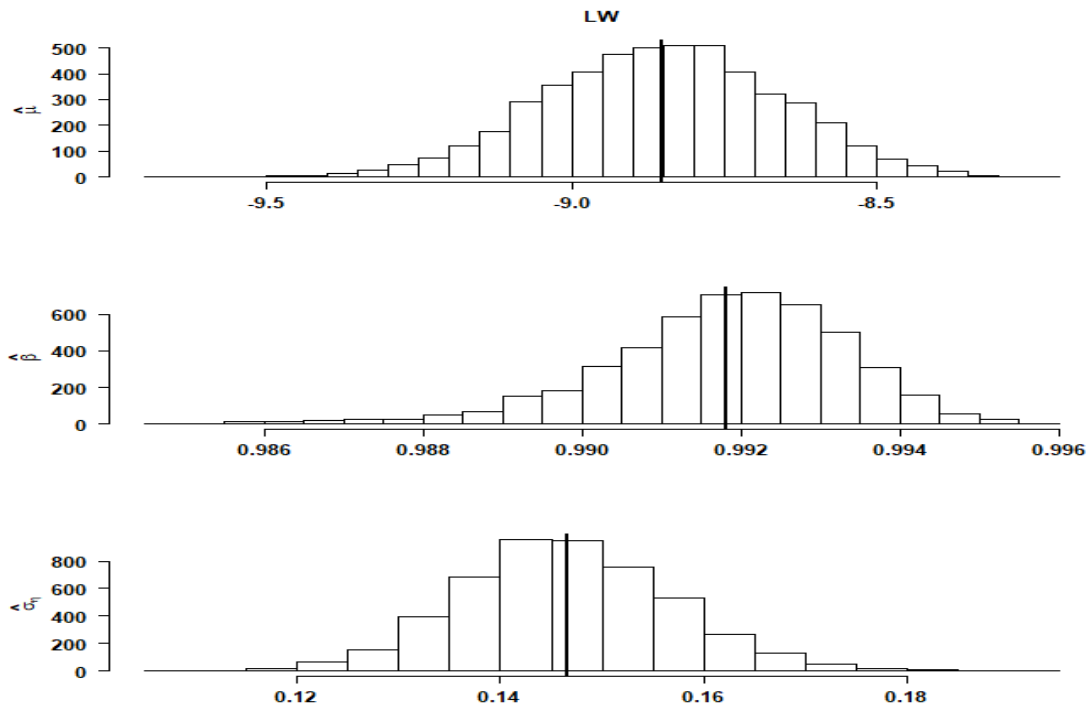


**B.- Figuras:**

**Figura 1B. Evolución de los parámetros del filtro de partículas de Liu y West (2001)  $N_p=5000$ ,  $\delta=0,95$**



**Figura 2B. Densidad de probabilidad posterior de los parámetros del modelo con la implementación de Acosta L.  $N_p=5000$ ,  $\delta=0,95$**



## C.- Código de implementación Filtro de Kalman y filtro de partículas

### C1.- Código: estimación vía filtro de Kalman y máxima verosimilitud (Newton raphson)

Cargamos los datos

```
setwd("G:/")
library(openxlsx)
library(astsa)
IBEX35<-read.xlsx("IBEX35.xlsx")
```

Calculamos los rendimientos

```
set.seed(123)
ret<-diff(log(IBEX35[,2]))
num=length(ret)
```

Definimos la ecuación de las observaciones de la volatilidad

```
eps=.Machine$double.eps
y=log(ret^2+1e-7)
```

Definimos los Parámetros iniciales que utilizaremos para maximizar la función de verosimilitud

```
phi0=0
phi1=.9702
sQ=.178
alpha=mean(y,na.rm=T)
sR0=1
mu1=-3
sR1=2
initpar=c(phi0,phi1,sQ,alpha,sR0,mu1,sR1)
```

Evaluación de la verosimilitud de la volatilidad Estocástica. Utilizamos la función SVFilter del package “astsa”, que lo que hace es aproximarla distribución del logaritmo de la Chi-cuadrado con un grado de libertad por una mixtura equilibrada de dos componentes gaussianas, una de ellas centrada en el cero.

```
Linn=function(para){
phi0=para[1]
phi1=para[2]
sQ=para[3]
alpha=para[4]
sR0=para[5]
mu1=para[6]
sR1=para[7]
sv = SVfilter(num,y,phi0,phi1,sQ,alpha,sR0,mu1,sR1)
return(sv$like)
}
```

Estimación de los parámetros y su error estándar usando un algoritmo de optimización no-lineal

```
est=optim(initpar, Linn, NULL, method = "BFGS", hessian = TRUE,
control=list(trace=1,REPORT=1))
stderr=sqrt(diag(solve(est$hessian)))
cbind(est$par,stderr)
```

Estimamos la volatilidad con los parámetros estimados

```

phi0=est$par[1]
phi1=est$par[2]
sQ=est$par[3]
alpha=est$par[4]
sR0=est$par[5]
mu1=est$par[6]
sR1=est$par[7]
sv = SVfilter(num,y,phi0,phi1,sQ,alpha,sR0,mu1,sR1)
Volatilidad.estimada<-exp(unlist(sv$xp)/2)
# para poder comparar la volatilidad de KF con PF hacemos el cambio de
escala, incluyendo alpha en la ecuación de observación.
Volatilidad.KF<-exp(unlist(sv$xp+alpha)/2)

```

Cálculo de los residuos

```

resid=y-alpha-sv$xp
hist(resid,freq = F, breaks = 20)
lines(density(resid), col="red3")

qqnorm(resid)
qqline(resid)

```

C2: Código R: estimación de la volatilidad vía filtro de partículas (paquete nimble)

```

setwd("C:/ ")
library(openxlsx)
IBEX35<-read.xlsx("IBEX35.xlsx")
ret<-diff(log(IBEX35[,2]))
L=length(ret)

```

Creamos un objeto que contiene el código con las especificaciones de las ecuaciones del modelo, juntamente con cada una de las distribuciones a priori que hemos decidido utilizar.

```

library("nimble")
stochVCode <- nimbleCode({
  x[1] ~ dnorm(mu+phi1*(x0-mu), var= sigmaSquaredInv)
  y[1] ~ dnorm(0, var = exp(x[1]))
  for(t in 2:T){
    x[t] ~ dnorm(mu+phi1*(x[t-1]-mu), var= sigmaSquaredInv)
    y[t] ~ dnorm(0, var = exp(x[t]))
  }
  x0 ~ dnorm(0, 1)
  mu ~ dnorm(mean=-8, var=25)
  phi1 <- 2 * phiStar - 1
  phiStar ~ dbeta(18, 1)
  sigmaSquaredInv ~ dinvgamma(5, 1/20)
})

```

Creamos un segundo objeto, que es el modelo a estimar en el cual introducimos el objeto creado anteriormente, los datos de la serie y los valores iniciales.

```

stochVolModel <- nimbleModel(code = stochVCode, name = 'stochVol',
                             constants = list(T = L), data = list(y = ret),
                             inits = list(mu=-8, phi1 =0.9702,
                                           sigmaSquaredInv = 0.031684))

```

**Compilamos el objeto que contiene el modelo con los datos y los valores iniciales en C++ (el mismo paquete proporciona una función para realizar la compilación)**

```
cstochVolModel <- compileNimble(stochVolModel, showCompilerOutput = F)
```

Creamos otro objeto con la función que estima el filtro de Liu y West, introduciendo los nodos “x” y los parámetros “params” que es lo que estimará conjuntamente. Además introducimos unos controles, como es el factor de descuento que queremos utilizar (0,95), y si queremos que nos dé la evolución de todo el estado (saveAll=T) o si queremos que solo nos de la estimación de estado en t (saveAll=F)

```
stochVolLiuWestFilter <- buildLiuWestFilter(model = stochVolModel,  
      nodes = 'x', control = list(saveAll = TRUE, d=0.95),  
      params = c('mu', 'phiStar', 'sigmaSquaredInv'))
```

Volvemos a compilar todo el modelo, y luego introducimos el número de partículas que queremos utilizar.

```
CstochVolLiuWestFilter <- compileNimble(stochVolLiuWestFilter, project =  
stochVolModel)  
CstochVolLiuWestFilter$run(5000)
```

Extraemos los resultados en una matriz y mostramos las primeras filas

```
resultados <- as.matrix(CstochVolLiuWestFilter$mvEWSamples)  
head(resultados[, c(1:6)])
```

Calculamos Phi0, Phi1, sigmaSquared y Sigma

```
(phi1 <- mean(2*resultados[, 2]-1))  
(phi0 <- mean(resultados[, 1]*(1-phi1)))  
(sigmaz <- mean(resultados[, 3]^0.5))  
(sigmasquared <- mean(resultados[, 3]))
```

Calculamos los residuos

```
res <- ret/apply(resultados[, c(4:3817)], 2, mean)  
hist(res, freq = F, breaks = 50, main="Histograma residuos", xlab="Residuos")  
lines(density(res), col="red3")  
abline(v=0, col="Red2")
```