



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

INFERENCIA DE LA VOLATILIDAD DE RETORNOS FINANCIEROS
USANDO FILTRO DE PARTÍCULAS

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN CIENCIAS DE LA
INGENIERÍA MENCIÓN ELÉCTRICA

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL ELECTRICISTA

FELIPE ARTURO TOBAR HENRÍQUEZ

PROFESOR GUÍA:

MARCOS ORCHARD CONCHA.

MIEMBROS DE LA COMISIÓN:

JORGE SILVA SÁNCHEZ

JOAQUÍN FONTBONA TORRES

SANTIAGO DE CHILE

JUNIO 2010

Agradecimientos

Hay muchas personas que me acompañaron durante el desarrollo de esta Tesis, agradezco sinceramente a ellas ya que sin su apoyo este proceso no habría sido lo mismo.

A mi Profesor Guía, Dr. Marcos E. Orchard, por sus sabios consejos en cuanto a la conducción de este trabajo, por sus enseñanzas en diversas tareas de investigación, y particularmente por su constante disposición a atender mis inquietudes académico-profesionales. A los miembros de mi comisión, Dr. Jorge Silva y Dr. Joaquín Fontbona por sus valiosos comentarios sobre esta Tesis. Adicionalmente, me gustaría expresar mi gratitud al Profesor Manuel A. Duarte-Mermoud por motivar en mí el trabajo de investigación durante pregrado y ser un constante apoyo desde ese momento.

A mis compañeros en pregrado, Juan Abuhadba, Ricardo Dodds, Alvaro Jesam, Pablo Ravest, Pablo Recabal, e Ignacio Rivas, por su compañía cuando había trabajo que hacer, pero también cuando no lo había. Mención especial a mis compañeros en postgrado, Gonzalo Díaz, Javier Testart y Pablo Guerrero, por compartir el gusto por la buena mesa en el barrio Beauchef, y a Diego Muñoz por esas interminables conversaciones sobre la academia y el futuro próximo.

Finalmente, y desde lo más profundo de mi corazón, agradezco a mis padres, Clementina y Luis, por su infinito amor y generosidad, por los valores que me han inculcado, y por su dedicación sin igual para con sus hijos. A mi hermana Anita, mi compañera desde la infancia en las buenas y en las malas, por su confianza y complicidad. Y a mi polola, Pia, por sus sonrisas, por el amor que me entrega día a día, y por hacerme tan feliz. Gracias.

Felipe A. Tobar.
Santiago, Junio de 2010.

Índice General

Agradecimientos	I
Resumen	XII
Lista de Acrónimos	1
1. Introducción	2
2. Descripción del Problema	5
2.1. Máxima Verosimilitud e Identificación	
Paramétrica	6
2.2. Filtrado no-lineal y no-Gaussiano	7
2.3. Reflexión Final y Comentarios	8
3. Modelos de Volatilidad	9
3.1. Generalized AutoRegressive Conditional	
Heteroskedasticity (GARCH)	11
3.1.1. Identificación Paramétrica	13
3.2. Modelo logarítmico de Volatilidad	
Estocástica (log-VE)	15
3.2.1. Identificación Paramétrica	19
4. Inferencia Bayesiana e Integración de Monte Carlo.	21
4.1. Inferencia Bayesiana	23
4.2. Filtro de Kalman	27
4.2.1. Filtro Extendido de Kalman	29

4.2.2.	EKF en estimación de parámetros	30
4.2.3.	EKF en modelos de volatilidad	31
4.3.	Métodos Secuenciales de Monte Carlo: Filtro de Partículas	33
4.3.1.	Muestreo de Importancia (IS)	35
4.3.2.	Muestreo de Importancia Secuencial y Remuestreo	37
4.3.3.	PF en estimación de parámetros.	40
4.3.3.1.	Evolución Artificial de Parámetros	41
4.3.4.	PF en modelos de volatilidad	42
5.	Desarrollo de un Modelo de Volatilidad Estocástica Basado en el Modelo Determinístico GARCH	44
5.1.	Modelo	45
5.2.	Identificación adaptativa de parámetros	49
6.	Análisis de Resultados	52
6.1.	Indicadores de Desempeño	53
6.1.1.	Filtrado	53
6.1.2.	Predicción	54
6.2.	Resultados Usando Series de Tiempo Generadas	56
6.2.1.	Generación de Datos	56
6.2.2.	Ajuste de Modelos para Estimación de Estado	57
6.2.3.	Filtrado	61
6.2.3.1.	log-VE	61
6.2.3.2.	uGARCH	64
6.2.4.	Predicción	68
6.2.4.1.	log-VE	69
6.2.4.2.	uGARCH	71
6.2.5.	Evaluación	73
6.3.	Resultados para el Índice <i>NASDAQ Composite</i>	74
6.3.1.	Preprocesamiento	74

6.3.2. Filtrado	78
6.3.2.1. log-VE	78
6.3.2.2. uGARCH	82
6.3.3. Predicción	87
6.3.3.1. log-VE	87
6.3.3.2. uGARCH	89
6.3.4. Evaluación	91
7. Conclusiones	92
Referencias	93
Apéndices	101
A . Códigos	101

Índice de tablas

6.1. Índices de desempeño para resultados de filtrado y predicción de una señal simulada de retornos usando los modelos log-VE y uGARCH, y los filtros PF y EKF.	73
6.2. Mediana y 4 primeros momentos para los retornos del índice <i>NASDAQ Composite</i> entre Abril 18, 2006 y Abril 8, 2010	77
6.3. Índices de desempeño para resultados de filtrado y predicción del índice <i>NASDAQ Composite</i> usando los modelos log-VE y uGARCH, y los filtros PF y EKF.	91

Índice de figuras

3.1.	Ejemplos de la densidad <i>a priori</i> del modelo logarítmico de VE $p(\sigma_t \sigma_{t-1}) = \text{Log-}\mathcal{N}(\sigma_t; \gamma + \phi V_{t-1}, \sigma)$. Azul: $\gamma = -2$, $\sigma = 0,35$, $V_{t-1} = -7$, $\phi = 0,5$. Verde: $\gamma = -2$, $\sigma = 0,3$, $V_{t-1} = -9$, $\phi = 0,5$	17
3.2.	Gráfico de la función de verosimilitud del modelo logarítmico de VE $L(V_t r_t)$ para $r_t = 0,01$. Momentos aproximados: Media: $-7,94$; Desviación estándar= $2,2213$; <i>Skewness</i> = $1,53$; <i>Kurtosis</i> = $6,98$	18
4.1.	Gráfico de $p(l_t) = p(\log(\epsilon_t^2))$, $\epsilon_t \sim \mathcal{N}(0, 1)$ (rojo) y una distribución normal (azul) de igual media y varianza.	32
5.1.	Ejemplos de $p(\sigma_t^2 \sigma_{t-1}^2)$ con $\omega = 10^{-6}$, $\alpha = 0,2$, y $\beta = 0,7$ para distintos valores de σ_{t-1}^2 . Rojo: $\sigma_{t-1}^2 = 0,0004$, azul: $\sigma_{t-1}^2 = 0,00044$, y verde: $\sigma_{t-1}^2 = 0,00048$. Se grafica también el comienzo del dominio para cada distribución (línea punteada).	47
5.2.	Banco de filtros extendidos de Kalman para realizar la estimación conjunta del estado y los parámetros de un modelo uGARCH. El tilde gorro denota las estimaciones de los filtros y z^{-1} representa el operador de retraso.	51
6.1.	Resultados para la estimación de la volatilidad estocástica de la serie autogenerada $r_{1:500}$ usando el modelo log-VE. a) Negro: Volatilidad autogenerada con modelo log-VE. Azul: Estimación de PF. Magenta: Estimación de PF del Intervalo de confianza del 95 %. b) Negro: Volatilidad autogenerada con modelo log-VE. Azul: Estimación de EKF. Magenta: Estimación de EKF del Intervalo de confianza del 95 %. c) Proceso de retornos generado	62

6.2.	Versiones continuas de las estimaciones de la distribución de la volatilidad usando PF y el modelo log-VE para los 30 primeros instantes del proceso simulado. En rojo se muestran las medias.	63
6.3.	Estimaciones de los parámetros γ , ϕ y σ^n del modelo log-VE usando serie de retornos simulada. Rojo: Estimación de máxima verosimilitud para el modelo GARCH usando $r_{1:150}$ (offline). Azul: Estimación online EKF. Negro: Estimación online PF.	64
6.4.	Resultados para la estimación de la volatilidad estocástica de la serie simulada $r_{1:500}$ usando el modelo GARCH. a) Negro: Volatilidad autogenerada con modelo GARCH. Azul: Estimación de PF y uGARCH. Magenta: Estimación de PF y uGARCH del Intervalo de confianza del 95 %. b) Negro: Volatilidad autogenerada con modelo GARCH. Azul: Estimación de EKF y uGARCH. Magenta: Estimación de EKF y uGARCH del Intervalo de confianza del 95 %. c) Proceso de retornos generado mediante GARCH.	66
6.5.	Versiones continuas de las estimaciones de la distribución de la volatilidad usando PF y el modelo uGARCH para los 30 primeros instantes usando la serie de retornos simulada. En rojo se muestran las medias.	67
6.6.	Estimaciones de los parámetros α y β del modelo uGARCH para una serie de retornos simulada. Rojo: Estimación de máxima verosimilitud para el modelo GARCH usando $r_{1:80}$ (offline). Azul: Estimación online EKF. Negro: Estimación online PF.	68
6.7.	Predicción de la volatilidad estocástica autogenerada usando el modelo log-VE para $t \in [150, 200]$. Rojo: Proceso generado. Negro: Estimación de PF de media (grueso) y CI del 95 % (delgado). Azul: Estimación de EKF de media (grueso) y CI del 95 % (delgado). Para $t \leq 150$ la estimación corresponde a filtrado y para $t > 150$ a predicción considerando sólo la información hasta $t = 150$ (Abscisas = t).	70

6.8.	Predicción de la volatilidad estocástica autogenerada usando el modelo log-VE para $t = 200$ en función del horizonte de predicción. Rojo: Valor de la volatilidad estocástica en $t = 200$, σ_{200}^2 . Negro: Predicción del valor esperado (grosso) y del CI del 95 % (delgado) de σ_{200}^2 usando muestras hasta $\tau \in [150, 200]$ usando PF. Azul: Predicción del valor esperado (grosso) y del CI del 95 % (delgado) de σ_{200}^2 usando muestras hasta $\tau \in [150, 200]$ usando EKF (Abscisas = τ). . .	70
6.9.	Predicción de la volatilidad estocástica autogenerada usando el modelo uGARCH para $t \in [150, 200]$. Rojo: Proceso generado. Negro: Estimación de PF de media (grosso) y CI del 95 % (delgado). Azul: Estimación de EKF de media (grosso) y CI del 95 % (delgado). Para $t \leq 150$ la estimación corresponde a filtrado y para $t > 150$ a predicción considerando sólo la información hasta $t = 150$ (Abscisas = t).	72
6.10.	Predicción de la volatilidad estocástica autogenerada usando el modelo uGARCH para $t = 200$ en función del horizonte de predicción. Rojo: Valor de la volatilidad estocástica en $t = 200$, σ_{200}^2 . Negro: Predicción del valor esperado (grosso) y del CI del 95 % (delgado) de σ_{200}^2 usando muestras hasta $\tau \in [150, 200]$ usando PF. Azul: Predicción del valor esperado (grosso) y del CI del 95 % (delgado) de σ_{200}^2 usando muestras hasta $\tau \in [150, 200]$ usando EKF (Abscisas = τ). . .	72
6.11.	Índice <i>NASDAQ Composite</i> y sus retornos entre Abril 18, 2006 y Abril 8, 2010.	75
6.12.	Funciones de autocorrelación muestral para a) el proceso de retornos y b) el proceso de los cuadrados de los retornos del índice <i>NASDAQ Composite</i> entre Abril 18, 2006 y Abril 8, 2010.	76
6.13.	Histograma de los retornos del índice <i>NASDAQ Composite</i> entre Abril 18, 2006 y Abril 8, 2010. En rojo se presenta una distribución normal de igual media y varianza.	78

6.14. Resultados del filtrado de la volatilidad estocástica de los retornos del índice <i>NASDAQ Composite</i> $r_{1:250}$ durante el período Julio 21, 2008 - Julio 17, 2009 usando el modelo log-VE. a) Negro: Estimación suavizada de la volatilidad de índice por un modelo GARCH(1,1) y ML. Azul: Estimación de PF. Magenta: Estimación de PF del CI del 95 %. b) Negro: Estimación suavizada de la volatilidad del índice por un modelo GARCH(1,1) y ML. Azul: Estimación de EKF. Magenta: Estimación de EKF del CI del 95 %. c) Valores observados del índice. d) Retornos del índice usando composición continua.	80
6.15. Densidades estimadas para la volatilidad del índice <i>NASDAQ Composite</i> en Noviembre 5, 2008 usando el modelo log-VE. a) Magenta: Versión continua de la estimación de la fdp de σ_{76}^2 estimada mediante PF, Azul: Estimación de la fdp de σ_{76}^2 por EKF, Negro: Estimación suavizada de σ_{76}^2 por GARCH(1,1) y ML, Rojo: Valor esperado de σ_{76}^2 estimado por PF. b) Partículas correspondientes a la estimación de σ_{76}^2 de PF.	81
6.16. Estimaciones de los parámetros γ , ϕ y σ^n del modelo log-VE para los retornos del índice <i>NASDAQ Composite</i> . Rojo: Estimación de máxima verosimilitud para el modelo GARCH usando $r_{1:80}$ (offline). Azul: Estimación online EKF. Negro: Estimación online PF.	82
6.17. Resultados del filtrado de la volatilidad estocástica de los retornos del índice <i>NASDAQ Composite</i> $r_{1:250}$ durante el período Julio 21, 2008 - Julio 17, 2009 usando el modelo uGARCH. a) Negro: Estimación suavizada de la volatilidad de índice por un modelo GARCH(1,1) y ML. Azul: Estimación de PF. Magenta: Estimación de PF del Intervalo de confianza del 95 %. b) Negro: Estimación suavizada de la volatilidad de índice por un modelo GARCH(1,1) y ML. Azul: Estimación de EKF. Magenta: Estimación de EKF del Intervalo de confianza del 95 %. c) Valores observados del índice. d) Retornos del índice usando composición continua.	84

6.18. Densidades estimadas para la volatilidad del índice <i>NASDAQ Composite</i> en Noviembre 5, 2008 usando el modelo uGARCH. a) Magenta: Versión continua de la estimación de la fdp de σ_{76}^2 estimada mediante PF, Azul: Estimación de la fdp de σ_{76}^2 por EKF, Negro: Estimación suavizada de σ_{76}^2 por GARCH(1,1), Rojo: Valor esperado de σ_{76}^2 estimado por PF. b) Partículas correspondientes a la estimación de σ_{76}^2 de PF.	85
6.19. Estimaciones de los parámetros α y β del modelo uGARCH para los retornos del índice <i>NASDAQ Composite</i> . Rojo: Estimación de máxima verosimilitud de los parámetros del modelo GARCH usando $r_{1:80}$ (offline). Azul: Estimación online EKF. Negro: Estimación online PF.	86
6.20. Predicción de la volatilidad estocástica del índice <i>NASDAQ Composite</i> usan- do el modelo log-VE para $t \in [50, 100]$. Rojo: Estimación suavizada de σ_t^2 con GARCH(1,1). Negro: Estimación de PF de media (grosso) y CI del 95 % (del- gado). Azul: Estimación de EKF de media (grosso) y CI del 95 % (delgado). Para $t \leq 50$ la estimación corresponde a filtrado y para $t > 50$ la predicción considerando sólo la información hasta $t = 50$ (Abscisas = t).	88
6.21. Predicción de la volatilidad estocástica del índice <i>NASDAQ Composite</i> usando el modelo log-VE para $t = 100$. Rojo: Estimación suavizada de la volatilidad estocástica en $t = 100$, σ_{100}^2 . Negro: Predicción del valor esperado (grosso) el CI del 95 % (delgado) de σ_{100}^2 usando muestras hasta $\tau \in [50, 100]$ usando PF. Azul: Predicción del valor esperado (grosso) el CI del 95 % (delgado) de σ_{100}^2 usando muestras hasta $\tau \in [50, 100]$ usando EKF. (Abscisas = τ)	88
6.22. Predicción de la volatilidad estocástica del índice <i>NASDAQ Composite</i> usando el modelo uGARCH para $t \in [150, 200]$. Rojo: Estimación suavizada de σ_t^2 con GARCH(1,1). Negro: Estimación de PF de media (grosso) y CI del 95 % (delgado). Azul: Estimación de EKF de media (grosso) y CI del 95 % (delgado). Para $t \leq 150$ la estimación corresponde a filtrado y para $t > 150$ a predicción considerando sólo la información hasta $t = 150$ (Abscisas = t).	90

6.23. Predicción de la volatilidad estocástica del índice *NASDAQ Composite* usando el modelo uGARCH para $t = 200$. Rojo: Estimación suavizada de la volatilidad estocástica en $t = 200$, σ_{200}^2 . Negro: Predicción del valor esperado (grosso) y del CI del 95 % (delgado) de σ_{200}^2 usando muestras hasta $\tau \in [150, 200]$ usando PF. Azul: Predicción del valor esperado (grosso) y del CI del 95 % (delgado) de σ_{200}^2 usando muestras hasta $\tau \in [150, 200]$ usando EKF. (Abscisas = τ) . . 90

Resumen

En este trabajo se presenta y evalúa un modelo de volatilidad estocástica, basado en el modelo determinístico *Generalized Autoregressive Conditional Heteroskedasticity* (GARCH), para describir la relación entre los retornos de un proceso financiero y la volatilidad de éstos. Este modelo, a diferencia de la estructura GARCH, considera que las observaciones de los retornos pueden explicarse a través de un proceso de innovación, cuya distribución *a priori* se asume idéntica a la que especifica el modelo GARCH para dichos retornos. La estructura propuesta recibe el nombre de *unobserved* GARCH (uGARCH), ya que considera que un proceso de innovación no observado –en lugar de una secuencia determinística– es el que dirige la evolución de la volatilidad.

La estructura propuesta uGARCH, entre otros modelos convencionales de volatilidad, ha sido utilizada para resolver el problema de estimación de estado, en conjunto con esquemas basados tanto en filtro de partículas (PF), como en el filtro extendido de Kalman (EKF). Además, con dichos filtros se identificaron adaptativamente los parámetros del modelo usando el concepto de evolución artificial de parámetros. El PF utilizado corresponde al muestreo de importancia con remuestro SIR, mientras que la estructura basada en EKF considera un banco de filtros. Ambas estructuras estiman conjuntamente los estados y parámetros del sistema.

Las técnicas presentadas han sido evaluadas cuantitativamente –mediante la introducción de índices de desempeño– en simulación y con datos reales. En este último caso se consideró la volatilidad del índice *NASDAQ Composite*, la cual se estimó durante el período Julio 21, 2008 hasta Julio 17, 2009, y se predijo al día Diciembre 15, 2008 considerando horizontes de predicción entre 1 y 45 días. Tanto en simulación como en la aplicación mencionada, los resultados reportados por el conjunto uGARCH-PF son auspiciosos en el sentido de que mediante su implementación (i) el filtrado es altamente exacto y preciso según los índices propuestos, y (ii) las predicciones del intervalo de confianza son consistentes con el valor real (o mejor aproximación) del proceso.

Lista de Acrónimos

a.k.a	También conocido como (<i>Also known as</i>)
CI	Intervalo de Confianza (<i>Confidence Interval</i>)
c.s	Casi seguramente
EKF	Filtro Extendido de Kalman (<i>Extended Kalman Filter</i>)
fdp	Función de densidad de probabilidad
INC	Índice <i>NASDAQ Composite</i>
KF	Filtro de Kalman (<i>Kalman Filter</i>)
GARCH	<i>Generalized AutoRegressive Conditional Heteroskedasticity</i>
i.i.d.	Independiente e Identicamente Distribuido
IS	Muestreo de Importancia (<i>Importance Sampling</i>)
log-VE	Volatilidad Estocástica logarítmica
ML	Máxima Verosimilitud (<i>Maximum Likelihood</i>)
NASDAQ	<i>National Association of Securities Dealers Automated Quotations</i>
PF	Filtro de Partículas (<i>Particle Filter</i>)
SIR	Muestreo Secuencial con Resampleo (<i>Sequential Importance Resampling</i>)
SMC	Métodos secuenciales de Monte Carlo (<i>Sequential Monte Carlo methods</i>)
UKF	Filtro Unscented de Kalman (<i>Unscented Kalman Filter</i>)
VA	Variable Aleatoria
VaR	<i>Value at Risk</i>
VE	Volatilidad Estocástica

Capítulo 1

Introducción

Para comprender los conceptos de filtrado y predicción, es necesario considerar en primer lugar un sistema dinámico que permita la realización de mediciones de un conjunto de variables significativas. Si este sistema es representado por ecuaciones diferenciales –o de diferencia– e involucra variables conocidas sólo en distribución de probabilidad, se dirá que se ha realizado una modelación estocástica del sistema. En este contexto, se entiende por estado del sistema al mínimo subconjunto de variables que permite relacionar, en forma única, la densidad de probabilidad condicional de la salida del sistema en función de su entrada. Usualmente interesa conocer el estado de un sistema, pues éste contiene información fundamental para implementar rutinas de predicción, control, o detección de comportamientos anómalos. Desafortunadamente, la mayoría de las veces esta cantidad no es medible directamente, por lo que se debe inferir en base a observaciones ruidosas disponibles y el conocimiento *a priori* del sistema.

En cuanto al análisis de sistemas dinámicos, el problema de filtrado se define como la utilización de las observaciones disponibles hasta un instante t , para encontrar la mejor estimación del estado del sistema en dicho instante. Por otro lado, el problema de estimación del vector de estados al tiempo t con mediciones hasta un instante τ , recibe el nombre de predicción si $t > \tau$, o bien, suavizado si $t < \tau$. Durante años se han desarrollado distintos métodos para resolver los problemas de filtrado, predicción, y suavizado de estado en sistemas estocásticos, los cuales han permeado diversas disciplinas tales como física, eléctrica, mecánica,

biología, meteorología, y econometría. En este sentido, el filtro óptimo para sistemas lineales y Gaussianos es el presentado en (Kalman & Bucy 1961), mientras que resultados sobre el problema de filtrado no lineal pueden ser encontrados en (Stratonovich. 1959).

La estimación de estado es un tema de particular interés en la disciplina de Matemáticas Financieras, i.e. la rama de las Matemáticas aplicadas que estudia el comportamiento de las variables de un sistema financiero (mercado). Los inicios de ésta disciplina se remontan al 29 de Marzo de 1900, día en el que Louis Bachelier defendió exitosamente su tesis llamada *Théorie de la Spéculation* en *La Sorbonne*¹. Este trabajo, apoyado por Henri Poincaré, ha influenciado fuertemente el desarrollo del cálculo estocástico y las Matemáticas Financieras hasta el día de hoy. Importantes resultados en este aspecto se deben a (Black & Scholes 1973, Merton 1973); mientras que textos recomendados con aplicaciones en finanzas son (Karatzas & Shreve 1988, Karatzas & Shreve 1998, Oksendal 2007, Oksendal & Sulem 2004).

Un significativo avance con respecto de las teorías presentadas por (Black & Scholes 1973, Merton 1973), es la introducción de los modelos de volatilidad estocástica, los que consideran a la volatilidad de un proceso de retornos financieros como un proceso estocástico. Estos modelos fueron motivados por cambios aparentemente aleatorios en la dinámica de la volatilidad de series financieras, y presentan una estructura acorde a fenómenos observados como *excess kurtosis*, *skewness*, y una notoria correlación en los cuadrados de los retornos. Si bien la habilidad de este tipo de modelos para representar fielmente el comportamiento de los procesos financieros es una ventaja con respecto a enfoques más simples, como es usual en el contexto de la modelación estocástica, la elección del modelo subyacente debe considerar el *trade-off* entre su exactitud, la complejidad asociada a la estimación de sus parámetros, y su factibilidad de implementación en línea y en altas dimensiones.

Los modelos de volatilidad estocástica son una clase particular de los modelos con estado no observable (HMM). En la mayoría de los casos, la implementación e identificación de los HMM no es posible mediante técnicas convencionales², haciendo necesaria la utilización de enfoque numéricos, los que pueden ser explotados gracias al crecimiento de la

¹Ver (Bachelier 1900)

²Como el filtro de Kalman y el método de máxima verosimilitud.

capacidad de almacenamiento y procesamiento computacional experimentada en los últimos años. En particular la inferencia Bayesiana, que para los problemas de filtrado, predicción, e identificación de parámetros involucra integrales definidas de alta dimensión, ha reportado provechosos resultados al ser utilizada en conjunto con métodos secuenciales de Monte Carlo (SMC), (Doucet, de Freitas & Gordon 2001).

Los SMC, son una clase de métodos capaces de aproximar esperanzas c.r.a una secuencia de medidas de probabilidad de dimensión creciente, mediante un conjunto de muestras ponderadas. Estos métodos han sido extensamente usados gracias a su flexibilidad y habilidad para adaptarse a distintos tipos de modelos dinámicos, sistemas con estado no observable, sistemas no lineales, y sistemas no Gaussianos. En el caso de los sistemas financieros, SMC ha captado la atención de un gran número de investigadores debido a su capacidad para estimar variables de sistemas que consideran ruido multiplicativos, e incluso difusiones de salto.

Además de una revisión de los modelos clásicos de volatilidad y de las técnicas de filtrado en base a filtro de Kalman y SMC, la contribución fundamental de esta Tesis es presentar un modelo de volatilidad estocástica –basado en la estructura de un modelo determinístico– cuyos parámetros son identificados adaptivamente, y la implementación de dicho modelo para filtrar y predecir la volatilidad instrumentos financieros. Los resultados de la estructura propuesta son comparado con modelos y técnicas convencionales de estimación de estado utilizando series de retornos simuladas y reales. Con respecto de las últimas se consideraron los retornos del índice *NASDAQ Composite*.

Esta Tesis está organizada de la siguiente forma: En el Capítulo 2 se presenta una breve introducción a los sistemas financieros y los problemas que en este contexto afloran, en el Capítulo 3 se introducen dos importantes clases de modelos de volatilidad, mientras que en el Capítulo 4 se revisan los conceptos básicos de inferencia Bayesiana, filtro de Kalman y filtro de Partículas. Un nuevo modelo de volatilidad es presentado en el Capítulo 5, cuyos resultados de simulación y experimentales –además de una comparación cuantitativa con métodos convencionales– son incluidos en el Capítulo 6. Finalmente, las conclusiones de esta tesis son discutidas en Capítulo 7.

Capítulo 2

Descripción del Problema

En los mercados financieros, las compañías utilizan distintos métodos para minimizar su riesgo. Uno de estos métodos es el uso de derivados, es decir, instrumentos financieros cuyo valor deriva de una acción, *commodity*, o índice, conocido como activo subyacente. Para entender cómo estos instrumentos ayudan a minimizar el riesgo, considérese una industria cuya materia prima es el cobre. Para esta industria, un alza en el precio del cobre afectaría directamente sus costos de producción, sin embargo, si esta industria tuviese la opción de comprar una cantidad de toneladas de cobre en el futuro, a un precio determinado, no sufriría las consecuencias de un alza inesperada del metal. Este derivado se llama opción de compra de cobre, y se tranza en la bolsa de metales de Londres.

Las opciones en el mercado financiero pueden ser de compra (opción *call*), o bien de venta (opción *put*), y el precio al cual se realiza la transacción es llamado precio de ejercicio (*strike price*). Además, el tiempo en el cual se ejecuta la opción es conocido como vencimiento (*expiration time*), y según éste existen opciones Americanas, posibles de ejecutar en cualquier instante previo al vencimiento, o Europeas, las que sólo son ejecutables en el vencimiento (Hull 2008). El derecho a tener una opción tiene un costo que depende fundamentalmente de la variabilidad del activo subyacente, ya que si fuese posible determinar el valor futuro de un activo, la transacción de derivados basados en él no tendría sentido. La práctica de determinar el precio de una opción es conocida como valorización de opciones, y los primeros

resultados se deben a los ganadores del premio Nobel¹ por la fórmula de Black-Scholes (Black & Scholes 1973, Merton 1973). Esta teoría, a pesar de asumir que el valor del activo subyacente es dirigido por un proceso estocástico, considera a su volatilidad como una función determinística.

A pesar del extensivo uso que ha tenido el modelo de Black-Scholes, observaciones del comportamiento del mercado revelan la presencia de fenómenos que este modelo no es capaz de explicar, tales como cambios de dinámica de volatilidad, *skewness* y *excess kurtosis* de los retornos, y clustering de volatilidad². Es por esta razón que una extensión natural de este modelo es considerar que la volatilidad del activo subyacente es una variable latente –i.e. no observable– que satisface una ecuación de diferencias estocástica. Esta consideración permite, mediante la estimación de la densidad de probabilidad de la volatilidad, inferir estadísticos como intervalos de confianza (CI), asimetrías, y esperanzas condicionales para hacer predicciones a corto, mediano, y largo plazo, además de cuantificar conceptos claves en finanzas como “Value at Risk” (VaR).

El seguimiento de la volatilidad de los retornos de un activo bajo los supuestos anteriores, es un problema que ha captado la atención de científicos de distintas disciplinas, los cuales en su búsqueda de nuevas teorías se han encontrado con el *trade-off* de la robustez de modelación permitida por el uso de variables latentes estocásticas, versus la complejidad de la identificación paramétrica, e implementación, de un modelo que incorpora dichas variables. Problemáticas que afloran en estos contextos se mencionan a continuación.

2.1. Máxima Verosimilitud e Identificación Paramétrica

Desde el punto de vista Bayesiano, la modelación de sistemas financieros implica la distribución de probabilidad de los retornos observados $r_{1:t}$ ³, condicional a un conjunto de parámetros θ , denotada por $p(r_{1:t}|\theta)$. Esta densidad define la función de verosimilitud del

¹F. Black falleció en 1995, no recibiendo el premio Nobel.

²Estos conceptos serán abordados en el Capítulo 3

³Adoptando la notación $r_{1:t} = \{r_j\}_{j=1..t} = \{r_j|j = 1..t\}$

parámetro θ , condicional a las observaciones $r_{1:t}$; es decir, $L(\theta|r_{1:t})$. Cuando el sistema modelado incluye variables latentes $\sigma_{1:t}$, la función de verosimilitud se puede escribir usando la versión integral del Teorema de la Partición (Grimmett & Welsh 1986, Papoulis 1984)⁴:

$$L(\theta|r_{1:t}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(r_{1:t}|\sigma_{1:t})p(\sigma_{1:t}|\theta)d\sigma_{1:t}. \quad (2.1)$$

En la mayoría de los casos, las relaciones que definen $p(r_{1:t}|\sigma_{1:t})$ y $p(\sigma_{1:t}|\theta)$ no son distribuciones usuales como $\mathcal{N}(\mu, \sigma)$ ⁵, por lo que la integral multidimensional anterior no es conocida analíticamente, y por ende, su maximización no es directa. La falta de una expresión analítica para $L(\theta|r_{1:t})$, imposibilita encontrar el estimador de máxima verosimilitud (ML) de θ , el cual es necesario para implementar una modelación confiable del proceso de retornos y su volatilidad latente.

Diversos métodos se han utilizado para resolver este problema, entre los que se encuentran el “método de los momentos” (Taylor 1986) y “quasi-máxima verosimilitud” (Harvey, Ruiz & Shephard 1994). Sin embargo, el primero de estos esquemas necesita una cantidad considerable de muestras para realizar estimaciones, mientras que el segundo corre el riesgo de usar aproximaciones de la función de verosimilitud que no sean representativas del proceso.

2.2. Filtrado no-lineal y no-Gaussiano

En el contexto de filtrado y predicción de señales, se ha demostrado en (Kalman & Bucy 1961) que para sistemas lineales y Gaussianos, el estimador óptimo de estado está dado por el filtro de Kalman-Bucy. Sin embargo, en su afán por ser consistentes con las observaciones, los intentos por modelar la volatilidad de los retornos han considerado estructuras cada vez más complejas, involucrando no-linealidades y procesos de ruido no-Gaussiano. Con la consideración de estos elementos, la implementación de rutinas de estimación de estado relacionadas con filtrado y predicción, es un problema que no tiene solución analítica para el caso general.

⁴También conocido como “Ley de alternativas” o “Teorema de Probabilidades Totales”.

⁵Donde $\mathcal{N}(\mu, \sigma)$ denota a la distribución normal de media μ y desviación estándar σ .

Para enfrentar este problema, distintos enfoques se han utilizado por más de cuarenta años. En este sentido, importantes resultados se deben a la implementación de aproximaciones de sumas de Gaussianas (Alspach & Sorenson 1972), al filtro extendido de Kalman (Anderson & Moore 1979), al modelo dinámico lineal generalizado (West, Harrison & Migon 1985), y a filtros *grid-based*. Desafortunadamente, estos métodos reportan soluciones de alta varianza, demandan considerables recursos computacionales para ser usados en altas dimensiones, e incluso algunos de ellos no permiten extraer propiedades de las estadísticas de interés de las variables (Doucet et al. 2001).

2.3. Reflexión Final y Comentarios

En base a los problemas mencionados, la modelación consistente de la volatilidad de un proceso de retornos financieros, es decir, una que explique los fenómenos observados, depende de la consideración de un modelo autoregresivo de volatilidad estocástica que permita la identificación en línea de sus parámetros mediante técnicas no convencionales, ya que las técnicas de identificación mediante ML no son aplicables en este tipo de estructuras. Dichos modelos son revisados en el Capítulo 3. Además, es igualmente necesario complementar este modelo con técnicas de estimación que reporten soluciones precisas –i.e. de baja varianza– y exactas –i.e. insesgadas–, y que también brinden una noción acabada de los estadísticos de las variables de interés en el contexto de los problemas de filtrado y predicción. Finalmente, es de esperar que estas técnicas no demanden excesivos recursos computacionales para poder ser implementadas en línea, y en horizontes de predicción extensos. Estas técnicas son tratadas en el Capítulo 4.

Capítulo 3

Modelos de Volatilidad

En finanzas, el término “volatilidad” se refiere a la desviación estándar de los retornos compuestos continuamente (*continuously compounded*) de un instrumento financiero, e.g. precio de una acción, índice bursátil, o tipo de cambio, y es representativo del riesgo que dicho instrumento involucra. En base a las hipótesis simplificatorias adoptadas para modelar un sistema financiero, la volatilidad de los retornos puede ser tratada como una secuencia determinística, o bien estocástica. Cualquiera sea la naturaleza de esta serie, en aplicaciones de análisis financiero es importante determinar la volatilidad en un instante de tiempo t dado, condicional a las mediciones hechas hasta algún instante τ .

Para introducir la primera clase de modelos de volatilidad, es necesario definir el concepto de volatilidad condicional (Rachev, Hsu, Bagasheva & Fabozzi 2008). Denotando Σ_τ a la σ -álgebra generada por las variables observadas hasta el instante τ , la volatilidad condicional $\sigma_{t|\tau}$ está definida por la relación:

$$\sigma_{t|\tau}^2 = \text{var}(r_t | \Sigma_\tau) = \mathbb{E} \left((r_t - \mu_{t|\tau})^2 | \Sigma_\tau \right), \quad (3.1)$$

donde r_t y $\mu_{t|\tau}$ son, respectivamente, el retorno y su esperanza condicional. Cuando la volatilidad condicional no es constante, se dice que los retornos son heteroscedásticos. El problema de la estimación de $\sigma_{t|\tau}^2$ es conocido como predicción para $t > \tau$, filtrado para $t = \tau$, y suavizado¹ para $t < \tau$.

¹*Smoothing.*

En relación con el comportamiento de la volatilidad de los retornos, observaciones de éstos últimos han revelado la presencia de un fenómeno llamado “*clustering* de volatilidad”. (Mandelbrot 1963) fue uno de los primeros en notarlo y mencionó que “grandes cambios [en precios del activo] tienden a ser seguidos por grandes cambios –de cualquier signo– y pequeños cambios tienden a ser seguidos por pequeños cambios”². En otras palabras, el *clustering* de volatilidad describe la propiedad del proceso de retorno de entrar en períodos de alta –o baja– volatilidad.

Otra característica observada en las series de retornos es la presencia de alta *skewness* y *kurtosis* con respecto a los presentados por la distribución normal. *Skewness*, o asimetría, es el tercer momento estandarizado de una variable aleatoria y es una medida de la asimetría de su distribución de probabilidad. Por otro lado, *kurtosis* –también conocido como el fenómeno de *fat tails*– es el cuarto momento estandarizado y denota el grado de concentración de frecuencias cercanas de la media. Usualmente el término *excess kurtosis*=*kurtosis*-3 es usado para denotar a las distribuciones que tienen mayor *kurtosis* que la distribución normal³. El riesgo asociado al considerar distribuciones Gaussianas, ignorando estos fenómenos, en la modelación de procesos financieros puede causar estimaciones erróneas de “Value at Risk” y de intervalos de confianza (Mandelbrot & Hudson 2004). Estudios asociados al cálculo de tercer y cuarto momento se encuentran en (Joanes & Gill 1998, Kim & White 2004).

Una forma simple para representar estos efectos es considerar que los retornos r_t , o bien su desviación c.r.a. un valor determinado *a priori*, pueden ser representados como el producto de dos procesos:

$$r_t = \sigma_t \epsilon_t, \quad (3.2)$$

donde ϵ_t es una secuencia i.i.d de media nula, varianza unitaria, e independiente del proceso σ_t , conocido como la volatilidad de los retornos (Carnero, Pena & Ruiz 2004). Distintas relaciones autoregresivas para la volatilidad σ_t han sido estudiadas para representar las características observadas del proceso de retornos. Dos enfoques importantes son: (i) el

²*Large changes [in asset prices] tend to be followed by large changes –of either sign– and small changes tend to be followed by small changes*

³ $\mathcal{N}(\mu, \sigma)$ tiene *kurtosis*=3 $\forall \mu, \sigma$.

modelo *Generalized AutoRegressive Conditional Heteroskedasticity* (GARCH) propuesto en (Bollerslev 1986), generalizando el modelo de (Engle 1982), y (ii) el modelo de Volatilidad Estocástica (VE), propuesto por (Taylor 1986). Si bien ambos modelos asumen que la volatilidad sigue un proceso autoregresivo con entrada exógena (ARX), éstos difieren en la naturaleza de esta entrada. El modelo GARCH asume que el proceso que dirige la volatilidad σ_t es una función determinística de ésta y de los retornos, mientras el modelo de VE sólo determina la evolución de σ_t en forma de distribución, i.e. sólo la probabilidad condicional $p(\sigma_t|\Sigma_{t-1})$ es conocida.

A continuación se presentan dichos modelos, se revisa su estructura, sus densidades de probabilidad relacionadas, y se discuten alternativas para la identificación de sus parámetros.

3.1. Generalized AutoRegressive Conditional Heteroskedasticity (GARCH)

La expresión analítica de los modelos GARCH los ha convertido en una alternativa muy usada para modelar volatilidad de retornos financieros. El modelo propuesto en (Bollerslev 1986) está dado por:

$$\sigma_{t|t-1}^2 = \omega + \alpha u_{t-1}^2 + \beta \sigma_{t-1|t-2}^2 \quad (3.3)$$

$$r_t = \mu_{t|t-1} + u_t, \quad (3.4)$$

donde el retorno r_t corresponde a la observación, la volatilidad $\sigma_{t|t-1}$ es el estado latente, $\mu_{t|t-1} = \mathbb{E}(r_t|\Sigma_{t-1})$ es la esperanza condicional del retorno, y el residuo $u_t = r_t - \mu_{t|t-1}$ tiene distribución –según (3.2)– dada por:⁴

$$u_t = \sigma_{t|t-1} \epsilon_t \sim \mathcal{N}(0, \sigma_{t|t-1}). \quad (3.5)$$

donde $\epsilon_t \sim \mathcal{N}(0, 1)$ i.i.d.

Además, $\omega > 0$, $\alpha \geq 0$, y $\beta \geq 0$ tal que $\sigma_{t|t-1}^2 > 0$ para todas las realizaciones de ϵ_t . El

⁴En el caso del modelo GARCH, la variable que es representada como producto de dos procesos (3.2) es el residuo de los retornos con respecto a su esperanza condicional, y no los retornos propiamente tal.

modelo (3.3)-(3.4) es llamado GARCH(1,1) por considerar un regresor para u_t^2 y un regresor para $\sigma_{t-1|t-2}^2$. Nótese que el subíndice $t|t-1$ en $\sigma_{t|t-1}$ y $\mu_{t|t-1}$ se debe a que estas variables están definidas sin incertidumbre dadas las observaciones de los retornos hasta el instante $t-1$, $r_{1:t-1}$.

Reescribiendo (3.3) como:

$$\sigma_{t|t-1}^2 = (1 - \alpha - \beta) \frac{\omega}{1 - \alpha - \beta} + \alpha u_{t-1}^2 + \beta \sigma_{t-1|t-2}^2, \quad (3.6)$$

es claro que el modelo GARCH(1,1) especifica la varianza condicional de los retornos como una suma ponderada de:

- La varianza a largo plazo, $\omega/(1 - \alpha - \beta)$.
- La varianza condicional en el instante anterior, $\sigma_{t-1|t-2}^2$.
- El residuo del retorno en el instante anterior, $u_t^2 = (r_t - \mu_{t|t-1})^2$.

Sustituyendo recursivamente $\sigma_{t-i-1|t-i-2}^2$, $i = 1..n$, (3.3) se puede escribir como:

$$\sigma_{t|t-1}^2 = \omega \frac{1 - \beta^{n+1}}{1 - \beta} + \alpha \sum_{j=0}^n \beta^j u_{t-j-1}^2 + \beta^{n+1} \sigma_{t-n-1|t-n-2}^2. \quad (3.7)$$

A partir de (3.7) es posible verificar que una secuencia de residuos $u_{1..t}$ de gran (pequeña) magnitud implicará valores altos (bajos) para la varianza condicional $\sigma_{t|t-1}^2$. Esta propiedad del modelo GARCH es capaz de explicar el fenómeno de *clustering* de volatilidad observado en los retornos.

Otro resultado importante sobre el modelo GARCH es que la estabilidad de $\mathbb{E}[\sigma_{t|t-1}^2]$ sólo

se garantiza si $\alpha + \beta < 1$. Para probar ésto, nótese primero que $\sigma_{t|t-1}^2$ y ϵ_t^2 son independientes:

$$\begin{aligned}
 \mathbb{E}[\sigma_{t|t-1}^2 \epsilon_t^2] &= \mathbb{E}[(\omega + \alpha \epsilon_{t-1} \sigma_{t-1|t-2}^2 + \beta \sigma_{t-1|t-2}^2) \epsilon_t^2] \\
 &= \mathbb{E}[\omega \epsilon_t^2] + \mathbb{E}[\alpha \epsilon_{t-1} \sigma_{t-1|t-2}^2 \epsilon_t^2] + \mathbb{E}[\beta \sigma_{t-1|t-2}^2 \epsilon_t^2] \\
 &= \omega \mathbb{E}[\epsilon_t^2] + \mathbb{E}[\alpha \epsilon_{t-1} \sigma_{t-1|t-2}^2] \mathbb{E}[\epsilon_t^2] + \mathbb{E}[\beta \sigma_{t-1|t-2}^2] \mathbb{E}[\epsilon_t^2] \\
 &= \mathbb{E}[\epsilon_t^2] (\omega + \mathbb{E}[\alpha \epsilon_{t-1} \sigma_{t-1|t-2}^2] + \mathbb{E}[\beta \sigma_{t-1|t-2}^2]) \\
 &= \mathbb{E}[\sigma_{t|t-1}^2] \mathbb{E}[\epsilon_t^2].
 \end{aligned} \tag{3.8}$$

Con este resultado, y considerando que $\epsilon_t \sim \mathcal{N}(0, 1)$, es posible calcular el valor esperado de $\sigma_{t|t-1}^2$ como:

$$\begin{aligned}
 \mathbb{E}[\sigma_{t|t-1}^2] &= \mathbb{E}[\omega + \alpha \epsilon_{t-1}^2 \sigma_{t-1|t-2}^2 + \beta \sigma_{t-1|t-2}^2] \\
 &= \omega + \alpha \mathbb{E}[\epsilon_{t-1}^2] \mathbb{E}[\sigma_{t-1|t-2}^2] + \beta \mathbb{E}[\sigma_{t-1|t-2}^2] \\
 &= \omega + \mathbb{E}[\sigma_{t-1|t-2}^2] (\alpha + \beta),
 \end{aligned} \tag{3.9}$$

consiguientemente, es necesario que $\alpha + \beta < 1$ para que $\mathbb{E}[\sigma_{t|t-1}^2]$ no diverja.

Con estas consideraciones sobre la estructura del modelo GARCH(1,1) es posible analizar la identificación de sus parámetros dado un conjunto de observaciones.

3.1.1. Identificación Paramétrica

La identificación de los parámetros del modelos GARCH puede ser realizada utilizando el método de máxima verosimilitud. Para ello, considérese $\mu_{t|t-1} = \mu$ durante el período estudiado, y defínase $\theta = [\omega, \alpha, \beta, \mu]^T$ los parámetros del modelo, Σ_0 el conjunto de información inicial, y $\Sigma_\tau = \Sigma_{\tau-1} \vee \{r_\tau\}$ ⁵ ⁶ el conjunto de información obtenida al instante τ . Con estas definiciones, la función de verosimilitud se puede construir usando la definición de probabilidad

⁵Donde $A \vee B$ denota la σ -álgebra generada por la unión de A y B .

⁶Nótese que esta definición de σ -álgebras también considera al proceso $\{\sigma_{t|t-1}\}_{t=1..T}$, pues en el modelo GARCH, dado $r_{1:T}$ y σ_0 , $\{\sigma_{t|t-1}\}_{t=1..T}$ es conocido sin incertidumbre.

condicional de la forma:

$$\begin{aligned}
 L(\theta|r_{1:T}, \Sigma_0) &= p(r_{1:T}|\theta, \Sigma_0) \\
 &= p(r_1|\theta, \Sigma_0)p(r_2|\theta, \Sigma_1) \cdots p(r_T|\theta, \Sigma_{T-1}) \\
 &= \prod_{k=1}^T p(r_k|\theta, \Sigma_{k-1})
 \end{aligned} \tag{3.10}$$

Asumiendo el proceso de observación dado por (3.4), y el hecho de que $\epsilon_t \sim \mathcal{N}(0, 1)$, la función logarítmica de verosimilitud está dada por:

$$\begin{aligned}
 \log L(\theta|r_{1:T}, \Sigma_0) &= \sum_{t=1}^T \log p(r_t|\theta, \Sigma_{t-1}) \\
 &= \sum_{t=1}^T \log \left(\frac{1}{\sqrt{2\pi}\sigma_{t|t-1}} \exp \left(-\frac{(r_t - \mu)^2}{2\sigma_{t|t-1}^2} \right) \right) \\
 &= \sum_{t=1}^T \log \left(\frac{1}{\sqrt{2\pi}} \right) - \frac{1}{2} \sum_{t=1}^T \log(\sigma_{t|t-1}^2) - \frac{1}{2} \sum_{t=1}^T \frac{(r_t - \mu)^2}{2\sigma_{t|t-1}^2},
 \end{aligned} \tag{3.11}$$

utilizando $\sigma_{t|t-1}^2$ como en (3.7) se tiene:

$$\begin{aligned}
 \log L(\theta|r_{1:T}, \Sigma_0) &= \sum_{k=1}^T \log \left(\frac{1}{\sqrt{2\pi}} \right) - \frac{1}{2} \sum_{k=1}^T \log \left(\frac{\omega}{1-\beta} + \alpha \sum_{j=1}^{\infty} \beta^{j-1} (r_{k-j} - \mu)^2 \right) \\
 &\quad - \frac{1}{2} \sum_{k=1}^T \frac{(r_k - \mu)^2}{2 \left(\frac{\omega}{1-\beta} + \alpha \sum_{j=1}^{\infty} \beta^{j-1} (r_{k-j} - \mu)^2 \right)}.
 \end{aligned} \tag{3.12}$$

Notar que esta estructura se tiene siempre y cuando se asuma que la diferencia entre los retornos r_t , y su esperanza condicional μ , es normalmente distribuida dado σ_t , es decir $\epsilon_t \sim \mathcal{N}(0, 1)$.

Debido a la compleja estructura de $\log L(\theta|r_{1:T}, \Sigma_0)$, su maximización debe ser realizada mediante métodos numéricos tales como programación no lineal (Gallop 2007), o el enfoque de *Expectation Maximisation* (EM) basado en simulaciones de Monte Carlo (Schön, Wills & Ninness 2006).

Si bien el modelo GARCH, presenta una arquitectura consistente con las observaciones del proceso de retornos, sólo permite estimar el valor esperado de la volatilidad del retorno y no sus momentos de orden superior, los cuales son necesarios para determinar intervalos de

confianza, estimar la asimetría de su distribución, o la existencia de desviaciones considerables con respecto a la media. Es por esto que surge la necesidad de un modelo que considere a la volatilidad como un proceso estocástico, dicha clase de modelos es presentada a continuación.

3.2. Modelo logarítmico de Volatilidad Estocástica (log-VE)

En todo modelo discreto de volatilidad, se asume que los retornos son conocidos con una frecuencia determinada. Entre dos observaciones consecutivas del proceso de retornos, distintos agentes del sistema, e.g. personas, compañías, gobiernos, modifican el próximo valor del retorno. Una forma de reunir todos estos efectos aleatorios producidos sobre los procesos de precios del mercado es suponer que la volatilidad de éstos está expuesta a modificaciones aleatorias en cada instante (Taylor 1986).

Los modelos de volatilidad estocástica postulan que la volatilidad σ_t es una variable latente, dirigida por una fuente de incertidumbre distinta a la que dirige al proceso observado, i.e. su distribución puede ser sólo estimada en base a las observaciones disponibles. Este supuesto sobre la volatilidad permite que los modelos de VE resuelvan una deficiencia del modelo de Black-Scholes para valorización de opciones, ya que éste asume que la volatilidad del activo subyacente es constante durante la vida de los derivados, y que no se ve afectada por los cambios en el nivel de precios del activo. Esta hipótesis no explica las observaciones a largo plazo, en las cuales se ven variaciones claras en la volatilidad. De hecho, estudios empíricos de (Schmalensee & Trippi 1978, Christie 1982), muestran una correlación inversa entre los retornos de una acción y cambios en su volatilidad. Para una revisión detallada de los modelos de VE y sus aplicaciones, referirse a (Ghysels, Harvey & Renault 1996, Shephard 1996).

En cuanto a la construcción del modelo de volatilidad estocástica, es claro que valores negativos para la volatilidad deben tener probabilidad nula, por lo que un modelo que considere una distribución normal no tiene sentido. Adicionalmente, distribuciones empíricas construidas en base a observaciones de retornos de series financieras (Taylor 1986) han mostrado concentración de valores extremos a la derecha de la media. Estos hechos requieren

una distribución con asimetría derecha⁷, definida no nula sólo en \mathbb{R}^+ , como por ejemplo la distribución log-normal (Clark 1973, Tauchen & Pitts 1983):

$$\sigma_t \sim \text{Log-}\mathcal{N}(m, \sigma), \quad (3.13)$$

es decir:

$$V_t = \log \sigma_t \sim \mathcal{N}(m, \sigma), \quad (3.14)$$

donde σ_t es la volatilidad, V_t es la volatilidad logarítmica, m y σ son cantidades a definir.

Usando (3.2), dejando σ constante, y definiendo $m = \gamma + \phi V_{t-1}$ —donde γ es la volatilidad incondicional y $\phi \in (-1, 1)$ representa la autocorrelación entre valores consecutivos de $V_{1..t}$ — la relación (3.13) permite enunciar el modelo de volatilidad estocástica presentado en (Taylor 1986):

$$V_t = \gamma + \phi V_{t-1} + \eta_t \quad (3.15)$$

$$r_t = \exp\left(\frac{1}{2}V_t\right)\epsilon_t \quad (3.16)$$

donde el retorno r_t es la observación, y $\forall t \in \mathbb{N}$ las VA's $\epsilon_t \sim \mathcal{N}(0, 1)$ y $\eta_t \sim \mathcal{N}(0, \sigma)$ constituyen procesos independientes.

En el contexto de la implementación del modelo (3.15)-(3.16) como estructura para la estimación de estado mediante filtro de partículas (dichos métodos se presentan en el Capítulo 4) es importante conocer las distribuciones relacionadas con los procesos de evolución de estado y de observación.

La densidad de evolución de la volatilidad logarítmica es normal:

$$p(V_t|V_{t-1}) = \mathcal{N}(\gamma + \phi V_{t-1}, \sigma), \quad (3.17)$$

⁷Right-skewed.

por lo que la densidad condicional de la volatilidad es:

$$\begin{aligned} p(\sigma_t|\sigma_{t-1}) &= \text{Log-}\mathcal{N}(\sigma_t; \gamma + \phi V_{t-1}, \sigma) \\ &= \frac{1}{\sigma_t \sigma \sqrt{2\pi}} \exp\left(-\frac{(\log \sigma_t - \gamma - \phi V_{t-1})^2}{2\sigma^2}\right) \end{aligned} \quad (3.18)$$

Esta distribución tiene asimetría derecha y es leptokúrtica⁸ (Balakrishnan & Chen 1999), en efecto, para la distribución en (3.18) se tiene:

$$\text{Skewness} = (e^{\sigma^2} + 2)\sqrt{e^{\sigma^2} - 1}, \text{ Kurtosis} = e^{4\sigma^2} + 2e^{3\sigma^2} + 3e^{2\sigma^2} - 3,$$

lo cual es consistente con la distribución empírica de los retornos observados. Ejemplos de esta distribución para dos conjuntos de parámetros son presentados en la Figura 3.1

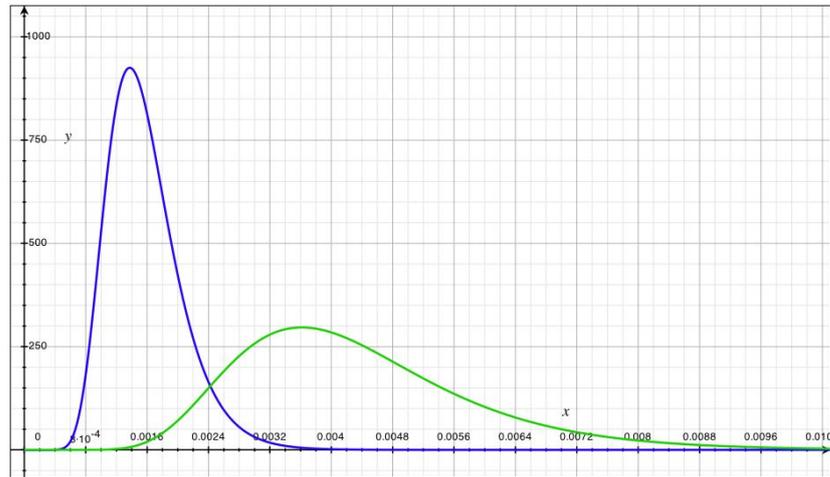


Figura 3.1: Ejemplos de la densidad *a priori* del modelo logarítmico de VE $p(\sigma_t|\sigma_{t-1}) = \text{Log-}\mathcal{N}(\sigma_t; \gamma + \phi V_{t-1}, \sigma)$. Azul: $\gamma = -2$, $\sigma = 0,35$, $V_{t-1} = -7$, $\phi = 0,5$. Verde: $\gamma = -2$, $\sigma = 0,3$, $V_{t-1} = -9$, $\phi = 0,5$.

Por otro lado, (3.2) define la densidad de las observaciones r_t condicional a la volatilidad logarítmica V_t –es decir, la función de verosimilitud del estado $L(V_r|r_t)$ – como una normal de media nula, i.e.

$$L(V_r|r_t) = p(r_t|V_t) = \mathcal{N}(0, e^{\frac{V_t}{2}}) \quad (3.19)$$

⁸Kurtosis > 3

También es importante conocer la densidad posterior del estado $p(V_t|r_t)$, pues ésta es la densidad que se va a estimar en el contexto de filtrado. Esta densidad está dada por:

$$\begin{aligned}
 p(V_t|r_t) &= p\left(\log \frac{r_t^2}{\epsilon_t^2}\right), \epsilon_t \sim \mathcal{N}(0,1) \\
 &= \frac{\partial}{\partial V_t} \mathbb{P}\left(\log \frac{r_t^2}{\epsilon_t^2} < V_t\right) \\
 &= \frac{\partial}{\partial V_t} \mathbb{P}\left(\frac{r_t^2}{\exp(V_t)} < \epsilon_t^2\right) \\
 &= \frac{\partial}{\partial V_t} \left(1 - \mathbb{P}\left(\epsilon_t^2 \leq \frac{r_t^2}{\exp(V_t)}\right)\right) \\
 &= -\chi^2\left(\frac{r_t^2}{\exp(V_t)}\right) \frac{\partial}{\partial V_t} \left(\frac{r_t^2}{\exp(V_t)}\right) \\
 &= \frac{\sqrt{\exp(V_t)}}{r_t \sqrt{2\pi}} \exp\left(-\frac{r_t^2}{2\exp(V_t)}\right) \frac{r_t^2}{\exp(V_t)} \\
 &= \frac{r_t}{\sqrt{2\pi} e^{V_t}} \exp\left(-\frac{r_t^2}{2e^{V_t}}\right)
 \end{aligned} \tag{3.20}$$

En la Figura 3.2 se presenta un ejemplo de esta distribución, y sus momentos aproximados, donde al menos para $r_t = 0,01$, la distribución tiene asimetría derecha y es leptokúrtika.

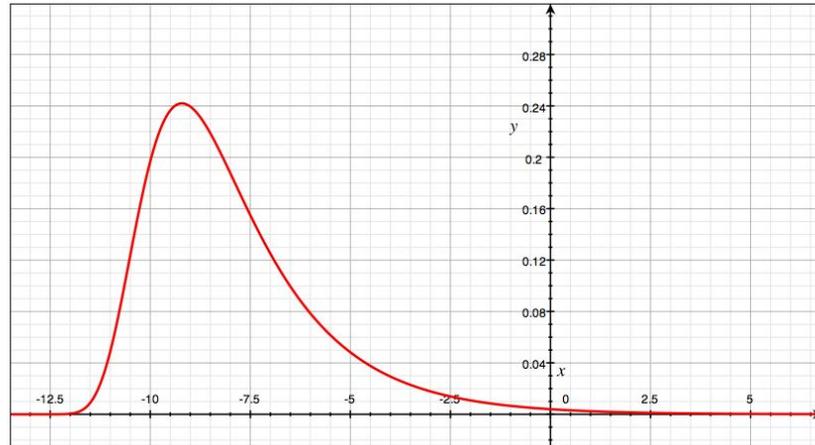


Figura 3.2: Gráfico de la función de verosimilitud del modelo logarítmico de VE $L(V_t|r_t)$ para $r_t = 0,01$. Momentos aproximados: Media: $-7,94$; Desviación estándar $=2,2213$; *Skewness* $=1,53$; *Kurtosis* $=6,98$.

Otra densidad importante en el contexto de identificación, es la función de verosimilitud

de los parámetros del modelo, la cual para el modelo de VE, no tiene forma conocida explícita como se verá a continuación.

3.2.1. Identificación Paramétrica

A diferencia del modelo GARCH, la estimación de parámetros para el modelo de VE es un problema complejo, pues la función de máxima verosimilitud de sus parámetros es una integral multidimensional que no tiene forma analítica conocida para el caso general, por lo cual, su maximización no es directa. En efecto, sea $\theta = [\gamma, \phi, \sigma]^T$ los parámetros del modelo de VE, su función de verosimilitud está dada por:

$$\begin{aligned} L(\theta|r_{1:t}) &= p(r_{1:t}|\theta) & (3.21) \\ &= \int p(r_{1:t}|V_{1:t}, \theta)p(V_{1:t}|\theta)dV_{1:t} \\ &= \int \prod_{j=1}^t p(r_j|V_j, \theta)p(V_j|V_{j-1}, \theta)dV_{1:t}, \end{aligned}$$

donde

$$p(r_{1:t}|V_{1:t}, \theta) = \prod_{j=1}^t p(r_j|V_j, \theta) \quad (3.22)$$

se tiene porque dado V_t , r_t es independiente de r_τ , y $V_\tau \forall \tau \neq t$. Además,

$$p(V_{1:t}|\theta) = \prod_{j=1}^t p(V_j|V_{j-1}, \theta) \quad (3.23)$$

se tiene porque (3.15) define $V_{1:t}$ como un proceso de Markov de primer orden.

Reemplazando (3.17) y (3.19) en (3.21) se tiene:

$$L(\theta|r_{1:t}) = \int \prod_{j=1}^t \frac{1}{2\pi\sigma e^{V_j/2}} \exp\left(\frac{-r_j^2}{2e^{V_j}} - \frac{(V_j - \mu - \phi V_{j-1})^2}{2\sigma^2}\right) dV_{1:t} \quad (3.24)$$

El cálculo de la solución de la integral en (3.24) no es trivial, sobretodo para un numero considerable de observaciones ($t \gg 1$). Para resolver este problema, se han buscado soluciones aproximadas a la maximización de $L(\theta|r_{1:T})$ mediante quasi-máxima verosimilitud en

base a distintos enfoques. En (Sandmann & Koopman 1996) se muestra que los métodos de Monte Carlo reportan soluciones más eficientes que los basados en el filtro de Kalman. Por otro lado, (Ruiz 1994) asegura que éstos últimos son “relativamente más eficientes” que el método generalizado de los momentos introducido en (Hansen 1982).

En esta subsección se ha presentado sólo el modelo de VE logarítmica, pero en general los modelos de VE son teóricamente apropiados para explicar características observadas en las series de tiempo financieras, sin embargo, éstos son conceptualmente demandantes y requieren considerables recursos computacionales para ser identificados e implementados. Es por esta razón que esquemas eficientes de identificación y estimación de estado son necesarios, en conjunto con un adecuado modelo de volatilidad, para extraer información de interés desde los procesos de retornos observados. En este sentido, las bases de la inferencia Bayesiana, además de dos importantes clases de estimadores de estado, se presentan en el siguiente capítulo.

Capítulo 4

Inferencia Bayesiana e Integración de Monte Carlo.

Distintas disciplinas científicas involucran el uso de un conjunto de observaciones para estimar cantidades desconocidas. En el caso particular de los modelos estudiados en el Capítulo 3, las observaciones corresponden al proceso de retornos del precio de un activo, y la variable que se pretende estimar es la volatilidad de dichos retornos. Desde el punto de vista Bayesiano, el conocimiento disponible del sistema en estudio se puede utilizar para formular distribuciones *a priori* de la evolución de las cantidades desconocidas, y funciones de verosimilitud para relacionar éstas con las observaciones. En este contexto, la inferencia sobre las variables latentes del sistema se basa en la distribución posterior obtenida mediante el Teorema de Bayes. Adicionalmente, para realizar estimaciones en línea es necesario actualizar esta distribución posterior en forma secuencial cada vez que nuevas observaciones son recibidas. Ejemplos en el contexto financiero en donde aflora esta problemática incluyen valorización de opciones, predicción de tipos de cambio, y como se ve en este trabajo, seguimiento de la volatilidad de instrumentos financieros.

Si las relaciones que definen los procesos de observación y evolución de estado están dadas por un modelo de espacio-estado lineal y Gaussiano, es posible derivar una expresión analítica óptima para la evolución de las distribuciones posteriores; esta expresión es el bien conocido filtro de Kalman (Kalman 1960), el cual asume los supuestos anteriores para asegurar la existencia de soluciones analíticas. Sin embargo, los procesos reales son en general bastante

complejos como para ser descritos por este tipo de relaciones, puesto que generalmente involucran no-linealidades, ruidos multiplicativos, no Gaussianos, y alta dimensionalidad; en donde consecuentemente la sucesión de densidades posteriores a estimar son integrales multi-dimensionales con argumentos no lineales que no tienen forma explícita para el caso general. Este problema es conocido por distintos nombres dependiendo del área en que se presente, entre los que se encuentran filtrado Bayesiano, filtrado óptimo no-lineal, filtrado estocástico, e inferencia en línea (Anderson & Moore 1979, Simon 2006, Oksendal 2007).

Entre las técnicas utilizadas para resolver el problema del filtrado Bayesiano, una que ha captado la atención de la comunidad científica por su facilidad de implementación, y su amplio campo de aplicación es Filtro de Partículas (PF)¹. Estos métodos son capaces de aproximar una secuencia de medidas de probabilidad de dimensión creciente mediante un conjunto de muestras ponderadas del espacio de estado, las que evolucionan en base al conocimiento previo del sistema y a las observaciones obtenidas del proceso. PF han mostrado ser una alternativa superior al filtro extendido de Kalman y al filtro *unscented* de Kalman para sistemas no lineales no Gaussianos, en el sentido de que con suficientes muestras, aproximan la estimación óptima del problema Bayesiano, de forma más exacta (Ristic, Arumlampalam & Gordon 2004). Una ventaja de PF es que sus propiedades de convergencia no dependen de la naturaleza del modelo subyacente del esquema de estimación, pudiendo incorporar no linealidades y procesos de innovación de diversas distribuciones. Las aplicaciones de PF están presentes en múltiples disciplinas, como seguimiento de trayectorias (Ristic et al. 2004), detección de fallas (Orchard 2009), pronóstico de eventos catastróficos (Orchard, Tobar & Vachtsevanos 2009), y econometría (Johannes, Polson & Stroud 2009).

En lo que sigue de este capítulo se presentará el problema de la estimación de estado desde el punto de vista Bayesiano y su solución óptima, para luego presentar dos estructuras que aproximan dicha solución.

¹A.k.a. Métodos Secuenciales de Monte Carlo (SMC).

4.1. Inferencia Bayesiana

El filtrado Bayesiano está definido como el proceso de usar observaciones ruidosas para estimar la densidad de probabilidad posterior del estado de un sistema dinámico no lineal y no necesariamente Gaussiano (Haug 2005). Además, dependiendo del conjunto de observaciones que se utilicen para la estimación, surgen también los conceptos de predicción y suavizado (Krishnan 2005). En vista de que la gran mayoría de los sistemas de interés considera observaciones secuenciales, en lo que sigue de este trabajo se pondrá énfasis en los sistemas de tiempo discreto.

El enfoque Bayesiano asume que la evolución de los estados del sistema en el tiempo puede ser descrita por un por un proceso de Markov², es decir:

$$x_0 \sim p_{x_0}(x) \quad (4.1)$$

$$x_{t+1}|x_t \sim p_{x_{t+1}|x_t}(x|x_t) \quad (4.2)$$

$$y_t|x_t \sim p_{y_t|x_t}(y|x_t) \quad (4.3)$$

donde al instante t , $x_t \in \mathcal{X}$ es el estado del sistema, e $y_t \in \mathcal{Y}$ es el proceso observado.

Un caso particular de esta estructura son los sistemas factibles de ser representados a través de un conjunto de ecuaciones de diferencia no lineales de evolución de estado y observación, respectivamente:

$$x_{t+1} = f_t(x_t, \omega_t) \quad (4.4)$$

$$y_t = h_t(x_t, v_t), \quad (4.5)$$

donde

$$f_t(\cdot) : \mathbb{R}^{n_x} \times \mathbb{R}^{n_\omega} \rightarrow \mathbb{R}^{n_x}$$

$$h_t(\cdot) : \mathbb{R}^{n_x} \times \mathbb{R}^{n_v} \rightarrow \mathbb{R}^{n_y}$$

son funciones no lineales, $\{\omega_t\}_{t \in \mathbb{N}}$ y $\{v_t\}_{t \in \mathbb{N}}$ son secuencias de variables aleatorias (VA) inde-

²A.k.a *Hidden Markov Model* (HMM) o modelo general de espacio estado.

pendientes e idénticamente distribuidas (i.i.d.), llamadas ruido de proceso y ruido de observación respectivamente, y n_x , n_y , n_ω , y n_v son respectivamente las dimensiones de x_t , y_t , ω_t , y v_t .

En los modelos presentados en el Capítulo 3, el estado $x_t = \sigma_t$ es la volatilidad, las observaciones $y_t = r_t$ son los retornos del precio del activo, los procesos de ruido son $\omega_t = \eta_t$ y $v_t = \epsilon_t$, y las dimensiones de los procesos son $n_x = n_y = n_\omega = n_v = 1$. Por simplicidad, en lo que sigue de este trabajo se tratará únicamente el caso homogéneo, i.e. las densidades de evolución de estado y observación no dependen del índice de tiempo t .

Además de la distribución posterior $p(x_{0:t}|y_{1:t})$ ³, es de interés estimar recursivamente cantidades relacionadas a ésta, tales como la densidad de filtrado $p(x_t|y_{1:t})$, la densidad de predicción $p(x_\tau|y_{1:t})$, $\tau > t$, y las esperanzas:

$$I(\phi) \doteq \mathbb{E}_{p(x_{0:t}|y_{1:t})}\{\phi(x_{0:t})\} = \int_{\mathcal{X}^{(t+1)}} \phi(x_{0:t})p(x_{0:t}|y_{1:t})dx_{0:t} \quad (4.6)$$

para alguna función de interés $\phi : \mathcal{X}^{(t+1)} \rightarrow \mathbb{R}^{n_\phi}$, integrable c.r.a $p(x_{0:t}|y_{1:t})$. Estas esperanzas permiten representar –al menos– todos los momentos del proceso $\{x_t\}_{t \in \mathbb{N}}$ mediante el uso de determinadas funciones $\phi(\cdot)$.

Con el fin de resolver el problema de estimación de estas cantidades, considérese que la ecuación de evolución de estado (4.2) define la densidad *a priori* del proceso $\{x_t\}_{t \in \mathbb{N}}$, mientras que la ecuación de observación (4.3) define la función de verosimilitud, esto es:

$$p(x_{1:t}) = p(x_1) \prod_{j=1}^t p(x_j|x_{j-1}), \quad (4.7)$$

y

$$p(y_{1:t}|x_{1:t}) = \prod_{j=1}^t p(y_j|x_j). \quad (4.8)$$

En el contexto Bayesiano, la inferencia sobre $\{x_t\}_{t \in \mathbb{N}}$ dado un conjunto de observaciones

³Se asumirá la notación $x_{0:t} \doteq \{x_0, \dots, x_t\}$ e $y_{0:t} \doteq \{y_0, \dots, y_t\}$

$\{y_t\}_{t \in \mathbb{N}}$, se puede expresar mediante la distribución posterior⁴

$$p(x_{1:t}|y_{1:t}) = \frac{p(y_{1:t}|x_{1:t})p(x_{1:t})}{p(y_{1:t})}, \quad (4.9)$$

donde:

$$p(y_{1:t}) = \int_{\mathcal{X}^t} p(x_{1:t}, y_{1:t}) dx_{1:t}. \quad (4.10)$$

En aplicaciones que consideran la disponibilidad de observaciones en forma secuencial, es de interés contar con una relación recursiva para la densidad posterior (4.9), esta relación puede ser obtenida mediante el Teorema de Bayes de la forma⁵:

$$\begin{aligned} p(x_{1:t}|y_{1:t}) &= p(y_{1:t}|x_{1:t}) \frac{p(x_{1:t})}{p(y_{1:t})} \\ &= p(y_{1:t-1}|x_{1:t}) \frac{p(x_{1:t})p(y_t|x_t)}{p(y_{1:t-1})p(y_t|y_{1:t-1})} \\ &= p(x_{1:t}|y_{1:t-1}) \frac{p(y_t|x_t)}{p(y_t|y_{1:t-1})} \\ &= p(x_{1:t-1}|y_{1:t-1}) \frac{p(x_t|x_{t-1})p(y_t|x_t)}{p(y_t|y_{1:t-1})}, \end{aligned} \quad (4.11)$$

donde:

$$p(y_t|y_{1:t-1}) = \int_{\mathcal{X}^2} p(y_t|x_t)p(x_t|x_{t-1})p(x_{t-1}|y_{1:t-1})dx_{t-1:t}. \quad (4.12)$$

Otro resultado importante es la forma recursiva para la densidad marginal $p(x_t|y_{1:t})$, pues en general interesa estimar el estado actual y no todo el camino recorrido por éste. Esta relación es derivada mediante dos etapas, predicción y actualización. La primera de éstas consiste en estimar $p(x_{0:t}|y_{1:t-1})$ suponiendo $p(x_{0:t-1}|y_{1:t-1})$ conocida, lo cual es posible considerando la densidad de evolución de estado $p(x_t|x_{t-1})$ dada por (4.4), y la relación de Chapman-Kolmogorov. De esta forma es posible expresar $p(x_t|y_{1:t-1})$ como:

$$p(x_t|y_{1:t-1}) = \int_{\mathcal{X}} p(x_t|x_{t-1})p(x_{t-1}|y_{1:t-1})dx_{t-1}. \quad (4.13)$$

Notar que en (4.13) se asumió que $p(x_t|x_{0:t-1}, y_{1:t-1}) = p(x_t|x_{t-1})$, ya que (4.4) y las es-

⁴Para ejemplos de las distribuciones posteriores y de verosimilitud en los casos lineales, o de estado finito entre otros, referirse a (Doucet & Johansen 2008).

⁵Para una derivación alternativa de esta recursión referirse a (Doucet et al. 2001, Doucet & Johansen 2008).

tadísticas de $\{\omega_t\}_{t \in \mathbb{N}}$ definen en forma única a $\{x_t\}_{t \in \mathbb{N}}$ como un proceso de Markov de primer orden.

Por otro lado, la etapa de actualización consiste en incorporar la información recibida para mejorar la estimación de x_t realizada en la etapa de predicción. En este sentido, evaluando la densidad de observación $p(y|x_t)$ –obtenida de (4.4)– en la muestra y_t , y utilizando el Teorema de Bayes, se tiene:

$$\begin{aligned} p(x_t|y_{1:t}) &= p(y_{1:t}|x_t) \frac{p(x_t)}{p(y_{1:t})} \\ &= p(y_t|x_t) p(y_{1:t-1}|x_t) \frac{p(x_t)}{p(y_t|y_{1:t-1}) p(y_{1:t-1})} \\ &= p(y_t|x_t) \frac{p(x_t|y_{1:t-1})}{p(y_t|y_{1:t-1})}. \end{aligned} \tag{4.14}$$

En aplicaciones de tiempo real, o bien cuando los recursos computacionales son limitados, no es de interés estimar valores pasados del estado, por esta razón el cálculo de la densidad $p(x_t|y_{1:t})$ es preferido por sobre $p(x_{0:t}|y_{1:t})$.

Las relaciones (4.11) y (4.12) representan la base para la solución óptima de estimación Bayesiana⁶. Sin embargo, debido a la complejidad de la constante de normalización $p(y_t|y_{1:t-1})$, las densidades $p(y_{1:t}|x_{1:t})$ y $p(x_{1:t})$, e $I(\phi)$, la propagación recursiva de la densidad posterior es, en general, sólo una solución conceptual y muchas veces no puede ser calculada en forma explícita. Soluciones analíticas como el filtro de Kalman existirán en casos restringidos, básicamente imponiendo Gaussianidad y linealidad, mientras que en los casos en que la solución óptima no es deducible, es posible implementar métodos aproximados como el filtro de Kalman extendido, el filtro *unscented* de Kalman, métodos basados en Grilla, y los métodos secuenciales de Monte Carlo. En lo que sigue de este capítulo se revisará el filtro de Kalman y PF, y se analizará cómo éstos se pueden utilizar en el contexto de estimación de variables financieras.

⁶La solución óptima Bayesiana resuelve el problema de calcular la densidad posterior exacta. Un algoritmo óptimo es un método para deducir esta solución, mientras que un algoritmo sub-óptimo es un método para encontrar un valor aproximado a la solución óptima.

4.2. Filtro de Kalman

El filtro de Kalman (KF) es un enfoque recursivo para estimar el estado de un sistema utilizando observaciones en línea y un modelo dinámico del sistema. Para sistemas lineales y Gaussianos, el filtro de Kalman es óptimo; i.e. ningún otro filtro entrega mejores estimaciones en el sentido del mínimo error cuadrático medio⁷ (Kalman 1960).

KF asume que la densidad posterior del estado es Gaussiana en todo instante. Suponiendo que $p(x_{t-1}|y_{1:t-1})$ es Gaussiana, es posible probar que $p(x_t|y_{1:t})$ también lo es si el sistema en (4.4) cumple con (Ristic et al. 2004):

- w_{t-1} y v_t son extraídos de una densidad Gaussiana de parámetros conocidos.
- $f_{t-1}(x_{t-1}, w_{t-1})$ es una función lineal conocida de x_{t-1} y de w_{t-1} .
- $h_t(y_t, v_t)$ es una función lineal conocida de y_t y de v_t .

Esto es, suponer que (4.4) se puede escribir de la forma:

$$x_t = F_{t-1}x_{t-1} + w_{t-1} \quad (4.15)$$

$$y_t = H_t x_t + v_t \quad (4.16)$$

donde $\forall t$, $x_t \in \mathbb{R}^m$ e $y_t \in \mathbb{R}^n$ son respectivamente el estado y las mediciones del sistema, $F_t \in \mathbb{R}^{m \times m}$ y $H_t \in \mathbb{R}^{n \times m}$ son matrices variantes en el tiempo, y $w_t \sim \mathcal{N}(0, Q_t)$ y $v_t \sim \mathcal{N}(0, R_t)$ son secuencias i.i.d.

En el contexto Bayesiano, el modelo (4.15)-(4.16) puede expresarse de la forma:

$$x_t \sim \mathcal{N}(F_{t-1}x_{t-1}, Q_{t-1}) \quad (4.17)$$

$$y_t \sim \mathcal{N}(H_t x_t, R_t). \quad (4.18)$$

Utilizando las densidades en (4.17)-(4.18) y las relaciones (4.13)-(4.14), el algoritmo del

⁷Minimum Mean Squared Error (MMSE)

KF puede ser visto de la siguiente forma:

$$p(x_t|y_{1:t-1}) = \mathcal{N}(x_t; \hat{x}_{t|t-1}, P_{t|t-1}) \quad (4.19)$$

$$p(x_t|y_{1:t}) = \mathcal{N}(x_t; \hat{x}_{t|t}, P_{t|t}) \quad (4.20)$$

donde las medias y covarianzas apropiadas son calculadas en forma recursiva como en (4.21)-(4.24) (Anderson & Moore 1979, Chui & Chen 1991).

$$\hat{x}_{t|t-1} = F_{t-1}\hat{x}_{t-1|t-1} \quad (4.21)$$

$$P_{t|t-1} = Q_{t-1} + F_{t-1}P_{t-1|t-1}F_{t-1}^T \quad (4.22)$$

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t(y_t - H_t\hat{x}_{t|t-1}) \quad (4.23)$$

$$P_{t|t} = P_{t|t-1} - K_tS_tK_t^T \quad (4.24)$$

donde

$$S_t = H_tP_{t|t-1}H_t^T + R_k \quad (4.25)$$

es la covarianza de la innovación $\nu_t = y_t - H_t\hat{x}_{t|t-1}$, y

$$K_t = P_{t|t-1}H_t^T S_t^{-1} \quad (4.26)$$

es la ganancia de Kalman.

El filtro de Kalman calcula recursivamente la media y covarianza de la distribución $p(x_t|y_{1:t})$, lo cual bajo los supuestos de Gaussianidad, es suficiente para conocer completamente esta distribución. Si los restrictivos supuestos anteriores se cumplen, el filtro de Kalman entrega la solución óptima para el problema de estimación de estado, sin embargo, en general la densidad posterior no es necesariamente Gaussiana, en estos casos KF deja de ser óptimo.

En los casos en que los sistemas en estudio involucran no linealidades, una alternativa para resolver el problema de estimación de estado es el filtro extendido de Kalman (EKF).

4.2.1. Filtro Extendido de Kalman

El EKF es un filtro subóptimo que relaja la condición de que –en el sistema (4.4)-(4.5)– las funciones f_{t-1} y h_t sean lineales, pero mantiene la condición de que los procesos de ruidos sean aditivos y Gaussianos. Es decir, este filtro está diseñado para estimar el estado de sistemas de la forma:

$$x_{t+1} = f_t(x_t) + \omega_t \quad (4.27)$$

$$y_t = h_t(x_t) + v_t, \quad (4.28)$$

El filtro extendido de Kalman aproxima las funciones no lineales $f_t(\cdot)$ y $h_t(\cdot)$ en (4.27)-(4.28) mediante una expansión en serie de Taylor de orden 1, asumiendo que esta aproximación es suficiente para describir la dinámica del sistema. Con esto, la densidad posterior $p(x_t|y_{1:t})$ es aproximada por una distribución Gaussiana y las relaciones en (4.19)-(4.20) se asumen cumplidas.

Con estos supuestos, las medias y covarianzas para el estado son calculadas en forma recursiva como en (4.29) (Bar-Shalom, Li & Kirubarajan 2001).

$$\hat{x}_{t|t-1} = f_{k-1}(\hat{x}_{t-1|t-1}) \quad (4.29)$$

$$P_{t|t-1} = Q_{t-1} + \hat{F}_{t-1}P_{t-1|t-1}\hat{F}_{t-1}^T \quad (4.30)$$

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t(y_t - h_t(\hat{x}_{t|t-1})) \quad (4.31)$$

$$P_{t|t} = P_{t|t-1} - K_tS_tK_t^T \quad (4.32)$$

donde

$$S_t = \hat{H}_tP_{t|t-1}\hat{H}_t^T + R_k \quad (4.33)$$

$$K_t = P_{t|t-1}\hat{H}_t^T S_t^{-1}, \quad (4.34)$$

y \hat{F}_{t-1} y \hat{H}_t son respectivamente los Jacobianos de las funciones f_{t-1} y h_t , evaluados, respectivamente, en $\hat{x}_{t|t-1}$ y $\hat{x}_{t-1|t-1}$. Es decir, los elementos de las matrices \hat{F}_{t-1} y \hat{H}_t están

definidos según:

$$\hat{F}_{t-1}[i, j] = \left. \frac{\partial f_{t-1}[i]}{\partial x_t[j]} \right|_{x_t = \hat{x}_{t|t}} \quad (4.35)$$

$$\hat{H}_t[i, j] = \left. \frac{\partial h_t[i]}{\partial x_t[j]} \right|_{x_t = \hat{x}_{t|t-1}} \quad (4.36)$$

donde en índice entre corchetes denota la coordenada del vector (o matriz) correspondiente.

Es importante mencionar que de acuerdo a (4.35) y (4.36), si las funciones no lineales $f_{t-1}(\cdot)$ y $h_t(\cdot)$ fuesen no derivables, esta clase de filtros no podría ser aplicada. Además, EKF siempre aproxima $p(x_t|y_{1:t})$ por una distribución Gaussiana, por lo que si en (4.27) o (4.28) la no linealidad es muy severa, la distribución normal será aún menos representativa de la distribución posterior verdadera. En estos casos, la eficiencia de EKF decae significativamente, haciendo necesaria la introducción de otras técnicas de aproximación de funciones no lineales. En este contexto, el filtro *unscented* de Kalman –mediante la transformada *unscented*– realiza una aproximación mucho más eficiente de las funciones no lineales, en el sentido de conservar la media y covarianza de las variables aleatorias durante la transformación no lineal. Para una comparación de las técnicas de aproximación de no linealidades ocupadas por EKF y UKF, referirse a (Ristic et al. 2004).

A pesar de que EKF y UKF han mostrado un desempeño adecuado para sistemas no lineales, aún siguen considerando que los ruidos de evolución de estado y observación son aditivos y Gaussianos, por lo que en el contexto de estimación de estado en sistemas que involucran distribuciones asimétricas, o con alta *kurtosis*, se debe recurrir a técnicas alternativas.

4.2.2. EKF en estimación de parámetros

En el contexto de modelos con estados latentes, es necesario incorporar técnicas alternativas de estimación paramétrica, pues la función de verosimilitud del estado en dichos modelos no permite la maximización mediante técnicas convencionales.

EKF permite realizar correcciones en línea de las estimaciones de los parámetros en la medida que nuevas estimaciones para el estado son calculadas, en base a la recepción de

nuevas observaciones. Para implementar esta idea es posible considerar un banco de filtros extendidos de Kalman, en donde un filtro es usado para, dado un conjunto de parámetros, estimar el estado del sistema; mientras que el segundo filtro recibe la estimación del estado y la considera como la entrada de un sistema en donde los parámetros a estimar son el estado, en aras de estimar dichos parámetros. Un diagrama de bloques de esta estructura para estimar conjuntamente el estado y los parámetros del modelo propuesto en el Capítulo 5, es presentado en la Figura 5.2.

4.2.3. EKF en modelos de volatilidad

Como se vió en (3.2), el proceso de observación de los modelos de volatilidad está dado por:

$$r_t = \sigma_t \epsilon_t,$$

donde el retorno r_t es el proceso observado, la volatilidad σ_t es el estado latente, y el proceso $\epsilon_{t \in \mathbb{N}} \sim NID(0, 1)$ es el ruido de observación.

Esta estructura no cumple con los supuestos para utilizar el filtro de Kalman, pues el ruido ϵ_t no es aditivo. Una manera propuesta en (Harvey et al. 1994) para ocupar las técnicas del KF es aplicar la transformación no lineal $\log(\cdot^2)$ al proceso de observación, convirtiendo el modelo logarítmico de volatilidad estocástica (3.15)-(3.16) en:

$$\log(r_t^2) = V_t + \log(\epsilon_t^2) \tag{4.37}$$

$$V_t = \mu + \phi V_{t-1} + \eta_t. \tag{4.38}$$

Definiendo $p_t = \log(r_t^2)$ y $l_t = \log(\epsilon_t^2)$, el modelo (4.37) es un modelo de espacio estado lineal, invariante, de ruido aditivo, pero no Gaussiano, por lo que ni el filtro de Kalman ni sus derivados (EKF o UKF) entregarán la solución óptima para $p(\sigma_t | r_{1:t})$. Para tener una idea de la forma que tiene la distribución de l_t , su derivación se presenta en (4.39), y en la Figura 4.1 se puede ver su gráfico junto a una distribución normal de igual media $(-1, 27)$ y varianza

$(\pi^2/2 = 4,93)^8$.

$$\begin{aligned}
 p(l_t) &= p(\log(\epsilon_t^2)) & (4.39) \\
 &= \frac{\partial}{\partial l_t} \mathbb{P}[\log(\epsilon_t^2) < l_t] \\
 &= \frac{\partial}{\partial l_t} \mathbb{P}[-e^{l_t/2} < \epsilon_t < e^{l_t/2}] \\
 &= \frac{\partial}{\partial l_t} (1 - 2\mathbb{P}[\epsilon_t < -e^{l_t/2}]), \text{ pues } \epsilon_t \sim \mathcal{N}(0, 1) \\
 &= -2\mathcal{N}(-e^{l_t/2}; 0, 1) \frac{\partial}{\partial l_t} (-e^{l_t/2}) \\
 &= \frac{1}{\sqrt{2\pi}} \exp\left(\frac{l_t}{2} - \frac{e^{l_t}}{2}\right).
 \end{aligned}$$

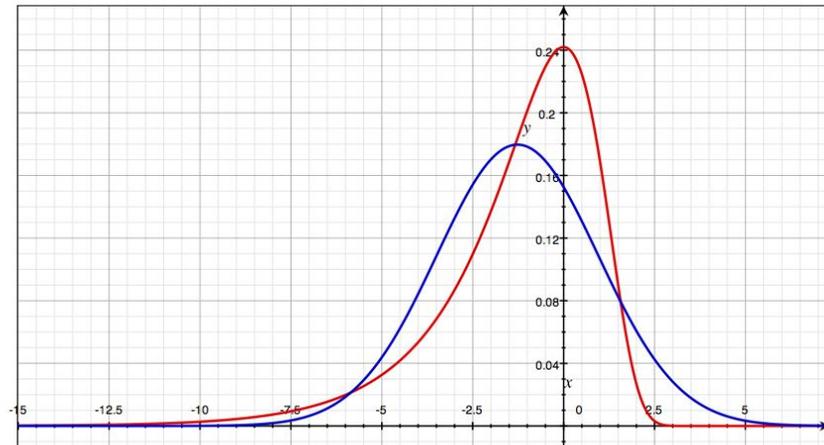


Figura 4.1: Gráfico de $p(l_t) = p(\log(\epsilon_t^2))$, $\epsilon_t \sim \mathcal{N}(0, 1)$ (rojo) y una distribución normal (azul) de igual media y varianza.

De la Figura 4.1 es posible tener una idea de los riesgos en que se incurre al asumir que las innovaciones del modelo logarítmico de VE son Gaussianas, pues $p(l_t)$ presenta una notoria asimetría izquierda.

Una ventaja del modelo logarítmico de volatilidad estocástica (3.15)-(3.16) es que, luego de la transformación no lineal mencionada, su estructura es completamente lineal, lo que

⁸Valores tomados de (Abramowitz & Stegun 1970).

—sin considerar la no Gaussianidad de la innovación— podría reportar buenos resultados en la rutina de estimación de estado mediante KF. Es importante mencionar que otro modelo de volatilidad estocástica importante para este trabajo (el que será introducido en el Capítulo 5) no cuenta con esta característica, pues luego de la transformación no lineal, si bien el ruido de observación es aditivo, se produce una no linealidad en la ecuación de observación. En este caso, la resolución del problema planteado requerirá necesariamente de una implementación basada en EKF o UKF.

El filtro de Kalman ha tenido un rol muy importante en el desarrollo de la teoría de estimación de estado, y sus aplicaciones han reportado sustantivos avances en el campo de la ingeniería. Sin embargo, después de haber revisado las condiciones necesarias para la aplicación de técnicas basadas en KF, y de considerar su posible aplicación a modelos de volatilidad, sus limitaciones quedan de manifiesto en el contexto de sistemas dinámicos complejos. Es por esto que estructuras alternativas, subóptimas, no restringidas a un determinado tipo de sistemas son necesarias. En este sentido, filtro de partículas es una herramienta que ha mostrado provechosos resultados en diversas disciplinas y es revisada a continuación.

4.3. Métodos Secuenciales de Monte Carlo: Filtro de Partículas

Métodos convencionales de Monte Carlo como *Markov Chain Monte Carlo* (MCMC) han sido extensamente usados para aproximar integrales y densidades de probabilidad. Sin embargo, su uso en el problema de inferencia Bayesiana no es directo, ya que éste involucra una secuencia de densidades de probabilidad variantes en el tiempo, mientras que MCMC asume que la densidad objetivo es invariante. Por esta razón, es necesario contar con una versión secuencial de los métodos de Monte Carlo que permita incorporar las observaciones para mejorar las estimaciones realizadas en forma recursiva.

Una herramienta que enfrenta este problema es filtro de Partículas (PF), una técnica numérica para aproximar integrales multidimensionales c.r.a una secuencia de medidas de probabilidad (Del Moral 2004). Suponiendo que la secuencia de densidades que se desea

aproximar es $\{\pi_t(x_{0:t})\}_{t \geq 1}$, donde $\forall t$ $\pi_t(x_{0:t})$ está definida en un espacio medible $(\mathcal{X}^{t+1}, \Sigma_{t+1})$ y puede ser evaluada punto a punto salvo una constante de normalización, la implementación de filtro de partículas consiste en generar una colección de $N \gg 1$ muestras aleatorias ponderadas $\{\omega_t^{(i)}, x_{0:t}^{(i)}\}_{i=1, \dots, N}$, $\omega_t^{(i)} \geq 0, \forall i, t$; que permitan aproximar $\pi_t(x_{0:t}), \forall t$ por la distribución empírica:

$$\pi_t^N(x_{0:t}) = \sum_{i=1}^N \omega_t^{(i)} \delta_{x_{0:t}^{(i)}}(x_{0:t}). \quad (4.40)$$

donde $\delta_\alpha(\cdot)$ es el delta de Dirac centrado en α .

Para resolver el problema de aproximación de esperanzas en el contexto de inferencia Bayesiana, se debe considerar que la distribución objetivo $\pi_t(x_{0:t}) = p(x_{0:t}|y_{1:t})$ es la fdp posterior de $x_{0:t}$ y reemplazar la distribución empírica (4.40) en (4.6). Con esta sustitución, las esperanzas pueden ser aproximadas mediante (Andrieu, Doucet & Punkskaya 2001):

$$\begin{aligned} \int_{\mathcal{X}^{t+1}} \phi_t(x_{0:t}) \pi_t(x_{0:t}) dx_{0:t} &\simeq \int_{\mathcal{X}^{t+1}} \phi_t(x_{0:t}) \pi_t^N(x_{0:t}) dx_{0:t} \\ &= \int_{\mathcal{X}^{t+1}} \phi_t(x_{0:t}) \sum_{i=1}^N \omega_t^{(i)} \delta_{x_{0:t}^{(i)}}(x_{0:t}) dx_{0:t} \\ &= \sum_{i=1}^N \omega_t^{(i)} \phi_t(x_{0:t}^{(i)}). \end{aligned} \quad (4.41)$$

Con estas relaciones, el problema de inferencia Bayesiana se reduce a la selección secuencial de las muestras y sus respectivos pesos tal que la sumatoria en (4.41) converja a la integral en (4.6). El algoritmo basado en PF más básico que resuelve este problema es el remuestreo secuencial de importancia (SIR)⁹.

Además, una ventaja importante de los enfoques basados en PF, es su alta flexibilidad de implementación, permitiendo la incorporación de diversas técnicas para mejorar su funcionamiento en problemas específicos, e.g. lazos de corrección externa (Orchard et al. 2009), (Tobar & Orchard 2009).

⁹ *Sequential Importance Resampling Particle Filter a.k.a Bootstrap Filter*

4.3.1. Muestreo de Importancia (IS)

El problema de aproximar la integral (4.6) puede ser resuelto mediante algoritmos de Monte Carlo, es decir, generando N caminos $x_{0:t}^{(i)}$, $i = 1, \dots, N$ y ajustando sus pesos tal que su promedio tienda a la integral para $N \rightarrow \infty$. Una forma directa de resolver este problema es extraer las N muestras según $p(x_{0:t}|y_{1:t})$ y realizar una estimación de Monte Carlo de $I(\phi)$ con pesos uniformes de la forma:

$$\hat{I}(\phi) = \frac{1}{N} \sum_{i=1}^N \phi(x_{0:t}^{(i)}), x_{0:t}^{(i)} \sim p(x_{0:t}|y_{1:t}), \quad (4.42)$$

sin embargo, los casos de interés son justamente en los que $p(x_{0:t}|y_{1:t})$ no es conocida, o bien no es posible extraer muestras de ella.

Para evitar la extracción de muestras directamente de $p(x_{0:t}|y_{1:t})$, es posible suponer que existe una densidad $q(x_{0:t}|y_{1:t})$ llamada densidad o función de importancia, de la cual es posible muestrear y evaluar puntualmente salvo una constante de normalización. Considerando esta nueva densidad, $I(\phi)$ se puede escribir como (Geweke 1989):

$$\begin{aligned} I(\phi) &= \int_{\mathcal{X}^{t+1}} \phi(x_{0:t}) p(x_{0:t}|y_{1:t}) dx_{0:t} \\ &= \frac{\int_{\mathcal{X}^{t+1}} \phi(x_{0:t}) \omega(x_{0:t}) q(x_{0:t}|y_{1:t}) dx_{0:t}}{\int_{\mathcal{X}^{t+1}} \omega(x_{0:t}) q(x_{0:t}|y_{1:t}) dx_{0:t}} \end{aligned} \quad (4.43)$$

donde los pesos de importancia $\omega(x_{0:t})$ están dados por la razón (derivada de Radon-Nikodym):

$$\omega(x_{0:t}) = \frac{p(x_{0:t}|y_{1:t})}{q(x_{0:t}|y_{1:t})} \quad (4.44)$$

claramente, para que estos pesos estén bien definidos es necesario que el soporte de $q(\cdot|y_{1:t})$ contenga al soporte de $p(\cdot|y_{1:t})$. Notar además que el denominador de (4.43) permite que $q(x_{0:t}|y_{1:t})$ no sea necesariamente una fdp.

Consecuentemente, mediante la obtención de N muestras extraídas según $q(x_{0:t}|y_{1:t})$ es

posible construir una aproximación de Monte Carlo de $I(\phi)$ basado en (4.43), es decir:

$$\begin{aligned}\hat{I}(\phi) &= \frac{\sum_{i=1}^N \phi(x_{0:t}^{(i)}) \omega(x_{0:t}^{(i)})}{\sum_{i=1}^N \omega(x_{0:t}^{(i)})} \\ &= \sum_{i=1}^N \phi(x_{0:t}^{(i)}) \bar{\omega}(x_{0:t}^{(i)}),\end{aligned}\tag{4.45}$$

donde los pesos de importancia normalizados están dados por:

$$\bar{\omega}(x_{0:t}^{(i)}) = \frac{\omega(x_{0:t}^{(i)})}{\sum_{j=1}^N \omega(x_{0:t}^{(j)})}.\tag{4.46}$$

Notar que para N finito, el estimador $\hat{I}_\pi[\phi(x_{0:t})]$ es sesgado, pues es una razón de dos estimadores. Sin embargo, bajo ciertos supuestos la ley de los grandes números garantiza la consistencia asintótica de $\hat{I}_\pi[\phi(x_{0:t})]$, es decir,

$$\hat{\mathbb{E}}_\pi[\phi(x_{0:t})] \xrightarrow{c.s.} \mathbb{E}_\pi[\phi(x_{0:t})].\tag{4.47}$$

Además, el teorema central del límite garantiza convergencia independiente de la dimensión del integrando. Para una explicación detallada referirse a (Geweke 1989).

El método de muestreo de importancia es ampliamente usado en los métodos de MC para reducir la varianza de estimación. La idea detrás de esta aplicación es aprovechar el hecho de que ciertos valores tienen un mayor impacto en la estimación que otros. En este caso, IS enfatiza la utilización de éstos valores, lo cual –debido al uso de distribuciones no representativas del proceso– puede llevar a estimaciones sesgadas a no ser que se ponderen correctamente las estimaciones obtenidas según sus pesos de importancia.

Si bien el método de muestreo de importancia representa un enfoque simple para aproximar integrales definidas invariantes, no es adecuado dentro del contexto de estimación recursiva, ya que cada vez que se desee estimar $\pi(x_{0:t}|y_{1:t})$ se necesita de todo el conjunto de valores $y_{1:t}$, haciendo el procedimiento computacionalmente ineficiente debido a la alta dimensionalidad de la densidad objetivo. Además, el uso de IS en el contexto de estimación recursiva ha dejado de manifiesto que la varianza de los pesos de las estimaciones explotan

si no se incluyen rutinas de remuestreo (Doucet et al. 2001). En resumen, cada vez que una observación y_t es recibida, interesa simplemente recalculer los pesos y las muestras representativas de la densidad de filtrado de forma secuencial y sin degeneración de los pesos. En la siguiente subsección se presenta una alternativa de estimación recursiva basada en IS.

4.3.2. Muestreo de Importancia Secuencial y Remuestreo

Para evitar la necesidad de altos recursos computacionales relacionados con la utilización de toda la secuencia $y_{1:t}$ cada vez que se quiera estimar $p(x_{0:t}|y_{1:t})$, es posible derivar una relación recursiva para la estimación de la densidad posterior, en la que en el instante t se conservan las muestras anteriores del estado $\{x_{0:t-1}^{(i)}; i = 1, \dots, N\}$, y sólo se extraen las muestras $\{x_t^{(i)}; i = 1, \dots, N\}$ en base a la observación recibida y_t , y a la estimación anterior de la densidad de filtrado $p(x_{0:t-1}|y_{1:t-1})$. Adicionalmente, para solucionar el problema de la degeneración de pesos es posible introducir un paso de remuestreo de la población de partículas, el que asegura una varianza mínima de los pesos, permitiendo que la densidad empírica $\pi_t^N(x_{0:t}^{(i)})$ sea representativa de $p(x_{0:t}|y_{1:t})$.

Para asegurar que las muestras $\{x_{0:t-1}^{(i)}; i = 1, \dots, N\}$ pueden ser conservadas para la estimación de $\{x_{0:t}^{(i)}; i = 1, \dots, N\}$, se debe disponer de una función de importancia $q(x_{0:t}|y_{1:t})$ que admita a la función de importancia en el instante anterior $q(x_{0:t-1}|y_{1:t-1})$ como distribución marginal, es decir:

$$q(x_{0:t}|y_{1:t}) = q(x_{0:t-1}|y_{1:t-1})q(x_t|x_{0:t-1}, y_{1:t}), \quad (4.48)$$

i.e. el muestreo de $x_{0:t-1}$ desde $q(x_{0:t-1}|y_{1:t})$ no depende de y_t , lo cual es bastante sencillo de obtener.

Con este supuesto para la función de importancia, es posible obtener una muestra $x_{0:t}^{(i)}$ sin alterar la muestra anterior correspondiente $x_{0:t-1}^{(i)}$, esto es:

$$x_{0:t}^{(i)} = [x_{0:t-1}^{(i)}, x_t^{(i)}] \quad (4.49)$$

con

$$x_t^{(i)} \sim q(x_t|x_{0:t-1}^{(i)}). \quad (4.50)$$

También se necesita una relación recursiva para los pesos de importancia, la cual considerando (4.48), está dada por:

$$\bar{\omega}(x_{0:t}^{(i)}) \propto \bar{\omega}(x_{0:t-1}^{(i)}) \frac{p(y_t | x_{0:t}^{(i)}) p(x_t^{(i)} | x_{t-1}^{(i)})}{q(x_t^{(i)} | x_{0:t-1}^{(i)}, y_{1:t})} \quad (4.51)$$

Un caso importante es cuando se considera la densidad *a priori* del estado como la función de importancia, es decir, $q(x_{0:t} | y_{1:t}) = p(x_{0:t})$. Con esta relación, la recursión para los pesos de importancia se puede expresar como:

$$\bar{\omega}(x_{0:t}^{(i)}) \propto \bar{\omega}(x_{0:t-1}^{(i)}) p(y_t | x_{0:t}^{(i)}). \quad (4.52)$$

lo que implica que los pesos pueden ser ajustados en forma recursiva directamente de la obtención de nuevas mediciones evaluadas en la función de verosimilitud.

Este enfoque, conocido como Muestreo de Importancia Secuencial (SIS), resuelve el problema de estimación recursiva de manera mucho más eficiente que IS. Sin embargo, es bien sabido (Doucet et al. 2001) que la distribución de los pesos de importancia en SIS se vuelve cada vez más disímil hasta el punto en que sólo una partícula tiene peso no nulo, lo cual no permite una representación acabada de la distribución posterior.

Para resolver este problema, el enfoque Remuestreo de Importancia Secuencial (SIR)¹⁰ considera la estructura anterior para generar muestras recursivamente y ponderarlas de acuerdo a su verosimilitud, incorporando además un paso de remuestreo. En este paso se remueven las partículas con bajo peso –y por ende baja verosimilitud– para (i) generar nuevas partículas en una región donde la densidad objetivo es alta, y (ii) evitar el esfuerzo computacional perdido en la actualización de partículas que no son significativas para el esquema de estimación.

La pregunta de cuando es el mejor instante para ejecutar el resamplero puede ser abordada considerando el criterio del tamaño efectivo de muestras (ESS) introducido por (Liu 1996) y

¹⁰A.k.a *Bootstrap Filter*.

definido por (Kong, Liu & Wong 1994) como:

$$\begin{aligned} ESS &= \frac{N}{1 + \text{var}_{p(\cdot|y_{1:t})}[\omega(x_{0:t})]} \\ &= \frac{N}{\mathbb{E}_{p(\cdot|y_{1:t})}[\omega(x_{0:t})^2]} \leq N. \end{aligned} \quad (4.53)$$

ESS no puede ser calculado, sin embargo, una estimación \widehat{ESS} dada por (Doucet, Godsill & Andrieu 2000) es:

$$\widehat{ESS} = \frac{\left(\sum_{i=1}^N \omega(x_{0:t}^{(i)})\right)^2}{\sum_{j=1}^N \left(\omega(x_{0:t}^{(j)})\right)^2}. \quad (4.54)$$

Básicamente, cuando \widehat{ESS} cae bajo cierto umbral, la rutina de resampleo es ejecutada (Rubin 1988). Uno de los métodos más usados es el enfoque multinomial (Del Moral 2004), el cual consiste en extraer muestras (con reemplazo) del conjunto $\{x_{0:t}^{(i)}\}_{i=1..N}$, en donde cada muestra es elegida con probabilidad proporcional a su peso.

Una forma de obtener las nuevas muestras $\{\hat{x}_{0:t}^{(i)}\}_{i=1..N}$ del conjunto de partículas $\{x_{0:t}^{(i)}\}_{i=1..N}$ acorde a sus pesos es mediante:

$$\hat{x}_{0:t}^{(i)} = \text{Inv}(u_i), \quad i = 1..N \quad (4.55)$$

donde las muestras $\{u_i\}_{i=1..N}$ son uniformemente distribuidas $[0, 1]$, e $\text{Inv}(u) = x$ es la inversa de la distribución empírica acumulada de x_t ,

$$\Pi_t^N = \int_{-\infty}^x \sum_{i=1}^N \bar{\omega}^{(i)} \delta_{x_{0:t}^{(i)}}(x) dx, \quad (4.56)$$

es decir,

$$\text{Inv}(u) \doteq x_{0:t}^{(i)} \iff u \in \left(\sum_{j=1}^{i-1} \omega(x_{0:t}^{(j)}), \sum_{j=1}^i \omega(x_{0:t}^{(j)}) \right]. \quad (4.57)$$

Cabe destacar que en este esquema de remuestreo, los nuevos pesos serán uniformes, ya que las muestras fueron extraídas según la estimación de la propia estimación de la densidad posterior del estado.

Si bien el remuestreo es una herramienta fundamental para evitar la degeneración de las

partículas, la aplicación de SIR conlleva consecuencias teóricas y prácticas (Doucet et al. 2000). Teóricamente, esta rutina de remuestreo remueve la independencia estadística de las muestras, por lo que importantes resultados de los algoritmos de Monte Carlo ya no se cumplen, sin embargo, (Berzuini, Best, Gilks & Larizza 1997) establece el teorema central de límite para $\hat{I}(\phi)$ cuando SIR es aplicado en cada iteración. Por otro lado, desde el punto de vista práctico, el remuestreo –además de demandar recursos computacionales adicionales– no permite la propagación de las partículas en paralelo, pues al ocupar SIR, todas las partículas deben ser combinadas.

4.3.3. PF en estimación de parámetros.

En el contexto de la estimación de estado, es fundamental contar con un modelo representativo del sistema presentado (4.4)-(4.5), sin embargo, los parámetros del modelo generalmente son desconocidos, o bien varían durante rutinas de estimación de horizontes muy extensos. Es por esta razón que la identificación recursiva es necesaria en estas rutinas. Para abordar este tópico desde el punto de vista de PF, considérese la función de verosimilitud-logarítmica¹¹ del vector de parámetros θ , condicional a las observaciones $y_{1:T}$ según:

$$\begin{aligned}
 l(\theta) &= \log L(\theta) & (4.58) \\
 &= \log p(y_{1:T}|\theta) \\
 &= \log \prod_{t=1}^T p(y_t|y_{1:t-1}, \theta) \\
 &= \sum_{t=1}^T \log p(y_t|y_{1:t-1}, \theta)
 \end{aligned}$$

En general, la presencia de variables latentes en (4.4)-(4.5) hace que (4.58) sea una integral multidimensional cuya maximización no es directa, por lo que la determinación de θ mediante máxima verosimilitud no puede ser realizada. Incluso si la maximización de $p(y_t|y_{1:t-1}, \theta)$ fuese abordada con un enfoque numérico, e.g. realizando una estimación de Monte Carlo de

¹¹*log-likelihood*

la forma:

$$p(y_t|y_{1:t-1}, \theta) \simeq \frac{1}{N} \sum_{i=1}^N \omega_t^{(i)} y_t^i \quad (4.59)$$

el error de muestreo y el alto esfuerzo computacional relacionado hacen que el método de máxima verosimilitud pierda sentido de todas formas (Doucet et al. 2001).

En estos casos, (Kitagawa & Sato 2001) propone la estimación Bayesiana de los parámetros del modelo, aumentando el vector de estado x_t con los parámetros desconocidos θ de la forma:

$$z_t = \begin{bmatrix} x_t \\ \theta_t \end{bmatrix}, \quad (4.60)$$

con $\theta_t = \theta$, definiendo la evolución trivial de los parámetros $\theta_t = \theta_{t-1}$.

Mediante la densidad $p(z_t|y_{1:t})$ es posible obtener las distribuciones posteriores tanto del estado como de los parámetros desconocidos. Los modelos de esta forma reciben el nombre de *Self-Organising State-Space Models* y varias aplicaciones de éstos pueden ser encontradas en (Kitagawa 1998). Para poder ocupar esta formulación en conjunto con PF, es necesario que la evolución de θ_t también sea afectada por una fuente de incertidumbre, asegurando la existencia de su propia fdp no trivial¹².

4.3.3.1. Evolución Artificial de Parámetros

Al considerar el estado variante en el tiempo, (Gordon, Salmond & Smith 2002) sugiere un enfoque para reducir la degeneración de las muestras mediante la incorporación de pequeñas perturbaciones a las partículas (llamadas *roughening penalties*) además de los ruidos inherentes del modelo. Esta idea ha sido extrapolada al problema de identificación de parámetros invariantes, donde –con la inclusión de perturbaciones aleatorias en el proceso de evolución de los parámetros– éstos son vistos como si fueran variantes en el tiempo, es decir, una “evolución artificial”.

La implementación de esta idea, según (Liu & West 2001), considera la ecuación de estados

¹²Distinta a $\delta_k(\cdot)$ para algún k .

de los parámetros como :

$$\theta_t = \theta_{t-1} + \epsilon_t \quad (4.61)$$

$$\epsilon_t \sim \mathcal{N}(0, W_t) \quad (4.62)$$

para una matriz de covarianza determinada W_t , y donde θ_{t-1} y ϵ_t son condicionalmente independientes dadas las observaciones hasta el instante $t - 1$. Ejemplos en donde la matriz de varianza es manipulada en función de cierta medida de eficiencia pueden ser encontrados en (Orchard et al. 2009) y (Tobar & Orchard 2009).

La incorporación de (4.61)-(4.62) en el modelo de evolución de estado permite que enfoques de filtro de partículas como el SIR puedan ser usados para estimar conjuntamente el estado no observable y los parámetros del sistema, sin sufrir degeneración de las muestras. Sin embargo, una desventaja relacionada con la implementación de este enfoque es que, al introducir fuente de incertidumbre externas en la evolución de los parámetros, las densidades resultantes para el estado original son más difusas que las densidades teóricas resultantes para el caso de parámetros fijos. En otras palabras, la consideración de un modelo donde los parámetros están expuestos a perturbaciones aleatorias artificiales, implica que la precisión de la inferencia sea limitada.

4.3.4. PF en modelos de volatilidad

A diferencia de KF/EKF, el hecho de que el ruido de observación de los modelos de volatilidad es multiplicativo (3.2) no es un problema para implementar técnicas de estimación de estado basadas en PF. De igual manera, las características no Gaussianas de las variables en el modelo logarítmico de volatilidad estocástica –o del modelo uGARCH introducido en el capítulo siguiente–, la cuales son necesarias para explicar fenómenos observados en series de retornos, pueden sin duda ser modeladas por PF. Una aplicación de PF en sistemas econométricos se puede ver en (Johannes et al. 2009), donde se consideran procesos de Lévy para dirigir la evolución de la volatilidad estocástica e índices bursátiles. Además, en (Raggi & Bordignon n.d.) se implementa un modelo de logarítmico de VE de parámetros variantes

en el tiempo.

En el Capítulo 5 se propone un modelo de VE basado en la estructura del modelo GARCH para filtrar y predecir la volatilidad estocástica de instrumentos financieros utilizando PF. Este modelo considera distribuciones no Gaussianas, funciones no lineales, y ruido multiplicativo para explicar el *clustering* de volatilidad, y los altos 3er y 4to momento; por lo que se espera que un enfoque basado en PF entregue estimaciones confiables de los estadísticos del estado latente.

Capítulo 5

Desarrollo de un Modelo de Volatilidad Estocástica Basado en el Modelo Determinístico GARCH

En los capítulos anteriores se ha mostrado que una ventaja del modelo GARCH presentado en (3.3)-(3.4), con respecto al modelo logarítmico VE (3.15)-(3.16), es que la estimación de sus parámetros es directa dado un conjunto de observaciones del proceso de retornos. Sin embargo, como este enfoque asume que la evolución de la volatilidad es determinística, no permite la inferencia de propiedades estadísticas importantes, tales como grados de asimetría¹, *kurtosis*, e intervalos de confianza para los problemas de estimación y predicción. Diversos trabajos comparan los modelos GARCH y de VE, entre los que cabe mencionar (Lehar, Scheicher & Schittenkopf 2002), en donde modelos de volatilidad constante, GARCH, y de volatilidad estocástica son comparados en el contexto del error observado del precio de opciones, y de predicción de riesgo en situaciones hipotéticas (VaR). La conclusión general de este trabajo es que el modelo GARCH reporta satisfactorios resultados en el contexto de la valorización de las opciones, sin embargo, en cuanto al manejo de riesgo no hay una superioridad evidente de dicho modelo. Por otro lado, en (Franses, van der Leij & Paap 2008) se propone un modelo de volatilidad estocástica construido mediante la estructura del modelo GARCH, y la adición de un proceso de innovación². La estructura de este modelo es tal que, cuando

¹3er momento *skewness*

²Este modelo es llamado *stochastic* GARCH (SGARCH).

la varianza del proceso de innovación introducido es nula, el modelo propuesto colapsa al modelo estándar GARCH. Las conclusiones en este trabajo aseguran que los parámetros del modelo presentado pueden ser estimados fácilmente, y que éste presenta menores errores de estimación paramétrica al ser comparado con el modelo GARCH convencional, sin embargo, en el contexto de predicción la diferencia entre ambos modelos no es evidente.

5.1. Modelo

En este trabajo se propone un modelo de volatilidad estocástica basado en la estructura del modelo GARCH, en el cual se asume que la volatilidad en (3.3) no es dirigida por el proceso observado³ $u_t = r_t - \mu$, sino que por un proceso u'_t que tiene la misma distribución de u_t pero no es observado. El fundamento de esta hipótesis es que, dadas las observaciones del proceso de retornos, se cree que la evolución del proceso $\sigma_{1:t}$ será mejor descrita si se considera completamente la densidad de los residuos, en lugar de una realización particular u_t .

Como el modelo GARCH presentado en el Capítulo 3 considera que el proceso de innovación $u_t = r_t - \mu$ es Gaussiano $\forall t \in \mathbb{N}$, la forma de implementar esta propuesta es reemplazando el proceso $u_{t \in \mathbb{N}}$ en (3.3) por $u'_{t \in \mathbb{N}}$ definido según:

$$u'_t = \sigma_t \eta_t \quad (5.1)$$

$$\eta_t \sim \mathcal{N}(0, 1) \text{ i.i.d } \forall t. \quad (5.2)$$

Con estas consideraciones, el nuevo modelo está dado por (5.3)-(5.4) y recibe el nombre de uGARCH (*unobserved* GARCH).

$$\sigma_t^2 = \omega + \alpha \sigma_{t-1}^2 \eta_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (5.3)$$

$$r_t = \mu + \sigma_t \epsilon_t. \quad (5.4)$$

donde r_t es un proceso de retornos, σ_t es la volatilidad estocástica, $\mu \in \mathbb{R}$, $\omega \in \mathbb{R}^+$, y $\alpha, \beta \in [0, 1]^2$ son parámetros; y $\epsilon_t \sim \mathcal{N}(0, 1)$ y $\eta_t \sim \mathcal{N}(0, \sigma)$ i.i.d. $\forall t$. Nótese que a diferencia

³O bien, *realised shocks*.

del modelo GARCH convencional, el subíndice de la volatilidad en (5.3) indica que en este modelo la volatilidad al instante t , σ_t , no es conocida sin incertidumbre dado $r_{1:t}$.

En el contexto de la utilización del modelo uGARCH en un enfoque de estimación del proceso $\sigma_{1:t}$ basado en filtro de partículas, es importante conocer las distribuciones que explican los procesos de evolución (5.3) y observación (5.4). Dado que $\eta_t \sim \mathcal{N}(0, 1)$, se tiene $\eta_t^2 \sim \chi_1^2$, con lo que la distribución condicional *a priori*, i.e., $p(\sigma_t^2 | \sigma_{t-1}^2)$, puede ser calculada de la siguiente forma:

$$\begin{aligned}
 p(\sigma_t^2 | \sigma_{t-1}^2) &= p(\omega + \alpha\sigma_{t-1}^2\eta_t^2 + \beta\sigma_{t-1}^2 | \sigma_{t-1}^2) & (5.5) \\
 &= \frac{\partial}{\partial \sigma_t^2} \mathbb{P}[\omega + \alpha\sigma_{t-1}^2\eta_t^2 + \beta\sigma_{t-1}^2 < \sigma_t^2] \\
 &= \frac{\partial}{\partial \sigma_t^2} \mathbb{P}\left[\eta_t^2 < \frac{\sigma_t^2 - \omega - \beta\sigma_{t-1}^2}{\alpha\sigma_{t-1}^2}\right] \\
 &= \chi_1^2\left(\frac{\sigma_t^2 - \omega - \beta\sigma_{t-1}^2}{\alpha\sigma_{t-1}^2}\right) \frac{\partial}{\partial \sigma_t^2} \left(\frac{\sigma_t^2 - \omega - \beta\sigma_{t-1}^2}{\alpha\sigma_{t-1}^2}\right) \\
 &= \frac{1}{\sqrt{(2\pi)(\sigma_t^2 - \omega - \beta\sigma_{t-1}^2)(\alpha\sigma_{t-1}^2)}} \exp\left(\frac{\omega + \beta\sigma_{t-1}^2 - \sigma_t^2}{2\alpha\sigma_{t-1}^2}\right), \quad \sigma_t^2 \geq \omega + \beta\sigma_{t-1}^2.
 \end{aligned}$$

En la Figura 5.1 se presentan distintos ejemplos de $p(\sigma_t^2 | \sigma_{t-1}^2)$ para diversos valores de σ_{t-1}^2 . Se puede ver que debido a la presencia del ruido η_t^2 en (5.3), esta distribución tiene una forma similar a χ_1^2 .

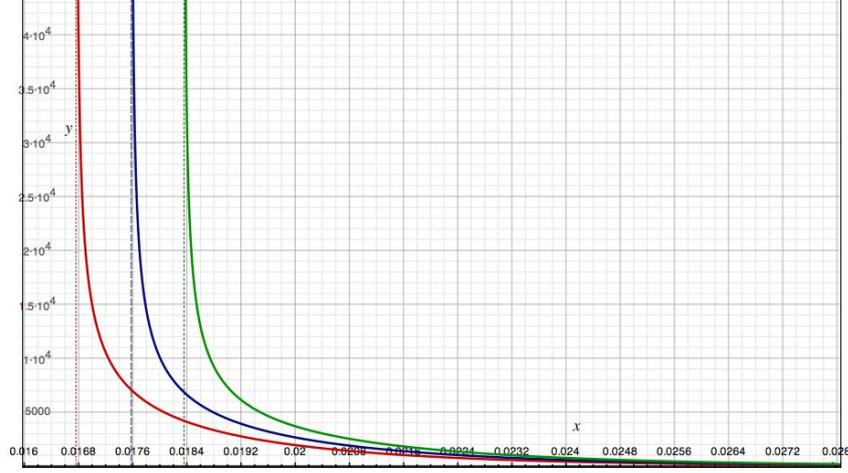


Figura 5.1: Ejemplos de $p(\sigma_t^2 | \sigma_{t-1}^2)$ con $\omega = 10^{-6}$, $\alpha = 0,2$, y $\beta = 0,7$ para distintos valores de σ_{t-1}^2 . Rojo: $\sigma_{t-1}^2 = 0,0004$, azul: $\sigma_{t-1}^2 = 0,00044$, y verde: $\sigma_{t-1}^2 = 0,00048$. Se grafica también el comienzo del dominio para cada distribución (línea punteada).

A pesar de no ser utilizada en implementación basada en PF, también se presenta la densidad *a priori* $p(\sigma_t | \sigma_{t-1})$ en (5.6) (definida igualmente en $\sigma_t^2 \geq \omega + \beta \sigma_{t-1}^2$).

$$p(\sigma_t | \sigma_{t-1}) = \frac{2\sigma_t}{\sqrt{(2\pi)(\sigma_t^2 - \omega - \beta\sigma_{t-1}^2)(\alpha\sigma_{t-1}^2)}} \exp\left(\frac{\omega + \beta\sigma_{t-1}^2 - \sigma_t^2}{2\alpha\sigma_{t-1}^2}\right). \quad (5.6)$$

Por otro lado, la densidad del proceso de observación condicional a la variable latente σ_t^2 , la cual define la función de verosimilitud del estado $L(\sigma_t^2 | r_t)$, está dada por:

$$L(\sigma_t^2 | r_t) = p(r_t | \sigma_t^2) = \mathcal{N}(\mu, \sigma_t). \quad (5.7)$$

Otra densidad importante para el problema de la estimación de estado es la densidad

posterior $p(\sigma_t^2|r_{1:t})$ dada por (5.8) según:

$$\begin{aligned}
p(\sigma_t^2|r_{1:t}) &= p\left(\frac{(r_t - \mu)^2}{\epsilon_t^2} | r_{1:t}\right) \\
&= \frac{\partial}{\partial \sigma_t^2} \mathbb{P}\left[\frac{(r_t - \mu)^2}{\epsilon_t^2} < \sigma_t^2\right] \\
&= \frac{\partial}{\partial \sigma_t^2} \mathbb{P}\left[\frac{1}{\epsilon_t^2} < \frac{\sigma_t^2}{(r_t - \mu)^2}\right] \\
&= \text{Inv-}\chi_1^2\left(\frac{\sigma_t^2}{(r_t - \mu)^2}\right) \frac{\partial}{\partial \sigma_t^2} \left(\frac{\sigma_t^2}{(r_t - \mu)^2}\right), \text{ pues } \epsilon_t \sim \mathcal{N}(0, 1) \\
&= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(r_t - \mu)^2}{2\sigma_t^2}\right) \frac{|r_t - \mu|}{\sigma_t^3}.
\end{aligned} \tag{5.8}$$

Un problema importante de esta distribución es que, tal como la distribución de Cauchy-Lorentz, no posee primer momento central, es decir, dada una secuencia de retornos $r_{1:t}$, el valor esperado de la varianza σ_t^2 del último retorno diverge, en efecto:

$$\begin{aligned}
\mathbb{E}_{p(\sigma_t^2|r_{1:t})}(\sigma_t^2|r_{1:t}) &= \int_{\mathbb{R}^+} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(r_t - \mu)^2}{2\sigma_t^2}\right) \frac{|r_t - \mu|}{\sigma_t^3} \sigma_t^2 d\sigma_t^2 \\
&> \int_1^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(r_t - \mu)^2}{2\sigma_t^2}\right) \frac{|r_t - \mu|}{\sigma_t} d\sigma_t^2 \text{ (argumento } > 0) \\
&> \frac{|r_t - \mu|}{\sqrt{2\pi}} \exp\left(-\frac{(r_t - \mu)^2}{2}\right) \int_1^\infty \frac{1}{\sigma_t} d\sigma_t^2 \text{ (} e^{\frac{-1}{x}} \text{ creciente)} \\
&= \frac{|r_t - \mu|}{\sqrt{2\pi}} \exp\left(-\frac{(r_t - \mu)^2}{2}\right) \int_1^\infty \frac{1}{\sqrt{\rho}} d\rho \text{ (cambio de variable)} \\
&= 2 \frac{|r_t - \mu|}{\sqrt{2\pi}} \exp\left(-\frac{(r_t - \mu)^2}{2}\right) \sqrt{\rho} \Big|_1^\infty \\
&= \infty.
\end{aligned} \tag{5.9}$$

Con este resultado, la estimación del estado latente parece perder sentido desde el punto de vista Bayesiano, pues las esperanzas condicionales requeridas para la solución de los problemas de filtrado y predicción no pueden ser calculadas. Sin embargo, como los esquemas de estimación de estado basados en PF consideran un número finito de muestras, y los esquemas basados en filtro de Kalman aproximan la densidad posterior por una distribución normal de media y varianza finitas, el objetivo de las estructuras de estimación propuestas será estimar los estados latentes del sistema mediante la consideración de una estimación de la densidad

posterior $\hat{p}(\sigma_t^2|r_{1:t})$, de momentos finitos, que sea representativa de la distribución $p(\sigma_t^2|r_{1:t})$ en las zonas de mayor concentración de masa de probabilidad.

Nótese que a pesar de que $p(\sigma_t^2|r_{1:t})$ no posee momentos finitos, el hecho de que sea unimodal (pues es monótona a partir de su moda) permite que sus intervalos de confianza sí sean finitos. Por esta razón, en las implementaciones de estimación y predicción de la volatilidad se pondrá énfasis en la estimación de intervalos de confianza, por ejemplo, del 95 %. Por otro lado, si bien la inexistencia de los primeros momentos de $p(\sigma_t^2|r_{1:t})$ invalida los resultados de convergencia conocidos para SMC, se espera que la utilización del modelo uGARCH para filtrar y predecir $\sigma_{1:t}$ utilizando SMC, presente resultados cuyas media sean similares a las entregadas por el modelo GARCH, pero que además entregue una noción completa de la distribución $p(\sigma_{1:t}^2|r_{1:t})$ en las zonas donde ésta concentra su masa de probabilidad.

5.2. Identificación adaptativa de parámetros

Para lograr una implementación exitosa del modelo propuesto, es necesario considerar un conjunto de parámetros que sea representativo del proceso bajo estudio. Como se explicó en los capítulos anteriores, la estimación de los parámetros del modelo uGARCH mediante ML no es directa, pues la función de verosimilitud a maximizar es una integral multidimensional. En efecto, de forma similar a (3.21) sea $\theta = [\mu, \alpha, \beta]$, la función de verosimilitud de los parámetros está dada por:

$$\begin{aligned} L(\theta|r_{1:t}) &= p(r_{1:t}|\theta) & (5.10) \\ &= \int p(r_{1:t}|\sigma_{1:t}^2, \theta)p(\sigma_{1:t}^2|\theta)d\sigma_{1:t}^2 \\ &= \int \prod_{j=1}^t p(r_j|\sigma_j^2, \theta)p(\sigma_j^2|\sigma_{j-1}^2, \theta)d\sigma_{1:t}^2 \end{aligned}$$

donde la última igualdad se tiene por las mismas razones que (3.22)-(3.23). Reemplazando (5.5) y (5.7) en (5.10), es posible expresar la función de verosimilitud de los parámetros del

modelo uGARCH según:

$$L(\theta|r_{1:j}) = \int \prod_{j=1}^j \frac{1}{2\pi\sigma_j\sigma_{j-1}\sqrt{\alpha(\sigma_j^2 - \omega - \beta\sigma_{j-1}^2)}} \exp\left(\frac{\omega + \beta\sigma_{j-1}^2 - \sigma_j^2}{2\alpha\sigma_{j-1}^2} + \frac{(r_j - \mu)^2}{2\sigma_j^2}\right) d\sigma_{1:t}^2$$

luego, es claro que la maximización de dicha función de verosimilitud no es un problema trivial de resolver.

Con el fin de obtener un conjunto de parámetros adecuados para implemetar rutinas de inferencia de estado, y basado en la similitud de las estructuras de los modelos GARCH y uGARCH, se considerará que –dada una secuencia de observaciones $r_{1:t}$ – los parámetros del modelo determinístico GARCH estimado mediante ML servirán como condiciones iniciales de las estimaciones de los parámetros del modelo uGARCH. Considerando dichos valores iniciales, es posible identificar adaptativamente los parámetros del modelo mediante el enfoque de evolución artificial de parámetros presentado en (Liu & West 2001), y explicado en la Sección 4.3.3.1, en el caso que los filtros lo permitan.

Filtro de partículas permite la consideración de sistemas que contemplen ruidos no Gaussiano y no linealidades, por lo que todos lo parámetros del modelo pueden ser estimados con este enfoque. Sin embargo, resultados experimentales preliminares revelan que la consideración de μ y ω en el contexto de estimación adaptativa mediante evolución artificial, no aporta estimaciones más certeras. Por otro lado, el filtro extendido de Kalman asume que los ruidos son Gaussianos y que los regresores son funciones del estado para realizar la estimación paramétrica, por esta razón, se considerará (5.3) de la forma:

$$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \alpha\sigma_{t-1}^2 + \alpha\sigma_{t-1}^2(\eta_{t-1}^2 - 1), \quad (5.11)$$

con esta expresión, el proceso de innovación es $(\eta_{t-1}^2 - 1)$, el cual tiene media cero y por ende es más cercano a una distribución normal que η_{t-1}^2 . Además, ambos parámetros a estimar (α y β) están acompañados de regresores que son funciones explícitas del estado, sin embargo, al estar ambos parámetros acompañados del mismo regresor, se espera que este enfoque conlleve a problemas de observabilidad, en donde si bien la estimación de $\alpha + \beta$ puede ser correcta, las

estimaciones por separado de dichos parámetros pueden no serlo necesariamente.

En la Figura 5.2 se presenta un diagrama de bloques para la estimación conjunta de los parámetros y estados del modelo uGARCH, considerando el banco de EKF como fue mencionado en la sección 4.2.2 y el enfoque de evolución artificial de parámetros presentado en la Sección 4.3.3.1.

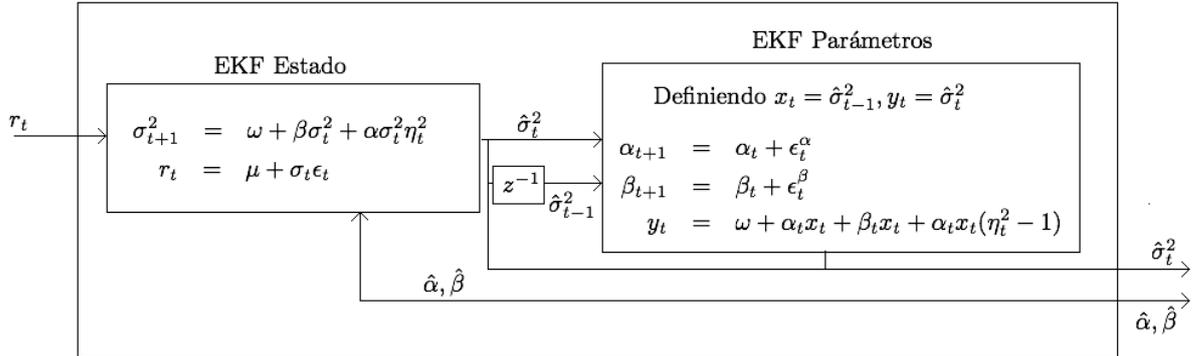


Figura 5.2: Banco de filtros extendidos de Kalman para realizar la estimación conjunta del estado y los parámetros de un modelo uGARCH. El tilde gorro denota las estimaciones de los filtros y z^{-1} representa el operador de retraso.

En el próximo Capítulo se presentan los resultados de la aplicación del modelo uGARCH y PF para el problema de filtrado y predicción de la volatilidad estocástica del índice *NASDAQ Composite*. El desempeño del modelo propuesto es comparado con el modelo logarítmico de VE y con el modelo GARCH. Además, EKF es utilizado para validar la eficiencia de PF en el filtrado y la predicción de estado en los modelos de volatilidad mencionados.

Capítulo 6

Análisis de Resultados

En este capítulo se presentan, comparan, y evalúan, los resultados de los enfoques propuestos para estimar y predecir la volatilidad estocástica de una serie de retornos financieros. Las estructuras de filtrado PF y EKF son implementadas considerando los modelos dinámicos log-VE y uGARCH, utilizando tanto datos simulados como series financieras reales. La evaluación de los resultados obtenidos es realizada por medio de un conjunto de indicadores propuestos.

La plataforma para todas las simulaciones y estimaciones fue MATLAB 7.9 (R2009b) corriendo sobre Mac OS X 10.5.8 en un Intel Core 2 Duo de 2.16GHz (32bits). El ajuste de modelos GARCH mediante ML se realizó con el *GARCH Toolbox*, mientras que todo el código utilizado para generar los resultados presentados en este capítulo fue escrito específicamente para este fin. Los algoritmos más importantes de este código pueden ser encontrados en el Apéndice.

En la implementación de los modelos log-VE y uGARCH se consideró el concepto de evolución artificial de parámetros tal como se precisó en (4.61)-(4.62), consecuentemente la identificación de los parámetros del modelo uGARCH se realizó acorde a las consideraciones del Capítulo 5. La estimación del vector de estado y parámetros variantes en el tiempo (hiperestado) fue realizada a través de esquemas Bayesianos de filtrado. En cuanto a los filtros utilizados, se consideró EKF tal como fue presentado en (4.29)-(4.34), y la implementación de FP fue el enfoque SIR presentado en la Sección 4.3.2, con 2000 partículas y un umbral de

remuestreo del 60 %.

La organización de este capítulo es de la siguiente forma, en la Sección 6.1 se presentan indicadores para cuantificar el desempeño de las estructuras propuestas, luego, en 6.2 se presentan y evalúan los resultados de los algoritmos utilizando señales de retornos simuladas. Finalmente, en 6.3 se presenta un análisis similar utilizando datos financieros reales, donde se consideró el valor del índice *NASDAQ Composite*.

6.1. Indicadores de Desempeño

Para analizar cuantitativamente el desempeño de los modelos y filtros propuestos en el contexto del filtrado y predicción de estado, se define el siguiente conjunto de indicadores. En ambientes de simulación se considerará que el estado generado, y por ende conocido, es la mejor estimación posible para el estado, mientras que en el caso de utilizar datos reales, se considerará que la mejor aproximación del estado es una estimación suavizada de éste. La motivación para usar una estimación suavizada como patrón de comparación es que, dado el Teorema Rao-Blackwell, y el hecho de que la estimación del estado al instante $t \in [1, T]$, $\bar{\sigma}_t^2$, considera todas las observaciones disponibles (es decir, $\{r_\tau | \tau \in [1, T]\}$), ésta será mejor –en el sentido de minimos cuadrados– que cualquier estimación que considere sólo las observaciones $\{r_\tau | \tau \leq t\}$, si la estructura de estimación cumple ciertas propiedades (Rao 1965).

6.1.1. Filtrado

Para un instante dado t , es posible cuantificar la exactitud de un esquema de filtrado mediante la expresión:

$$i_X^{AC}(t) = \frac{|\hat{\sigma}_t^2 - \bar{\sigma}_t^2|}{\overline{CI}_X(t)} \quad (6.1)$$

donde $\bar{\sigma}_t^2$ y $\overline{CI}_X(t)$ son respectivamente las mejores estimaciones¹ de σ_t^2 y del intervalo de confianza (CI) del X %, al tiempo t . $\hat{\sigma}_t^2$ es la estimación del estado σ_t^2 .

Con (6.1) es posible definir el índice de exactitud I_X^{AC} , con respecto al CI del X %, como

¹Es decir, una estimación suavizada o bien la propia señal generada donde corresponda.

el promedio de $i_X^{AC}(t)$ durante el horizonte de filtrado $[1, T]$, es decir:

$$I_X^{AC} = \frac{1}{T} \sum_{j=1}^T i_X^{AC}(t) = \frac{1}{T} \sum_{j=1}^T \frac{|\hat{\sigma}_t^2 - \bar{\sigma}_t^2|}{\widehat{CI}_X(t)} \quad (6.2)$$

Valores cercanos a cero del índice I_X^{AC} sugieren que las estimaciones de la estructura analizada son exactas con respecto a la propia estimación del CI del $X\%$.

En cuanto a la precisión instantánea de una estructura de filtrado, ésta puede ser cuantificada por (6.3).

$$i_X^{PR}(t) = \frac{\widehat{CI}_X(t)}{\widehat{CI}_X(t)} \quad (6.3)$$

donde $\widehat{CI}_X(t)$ es la estimación del CI del $X\%$ de σ_t^2 al tiempo t proporcionada por el filtro a evaluar.

Análogamente, con (6.3) es posible definir el índice de precisión I_X^{PR} relacionado con el intervalo de confianza del $X\%$ como el promedio de $i_X^{PR}(t)$ durante el horizonte de filtrado $[1, T]$, es decir:

$$I_X^{PR} = \frac{1}{T} \sum_{j=1}^T i_X^{PR}(t) = \frac{1}{T} \sum_{j=1}^T \frac{\widehat{CI}_X(t)}{\widehat{CI}_X(t)} \quad (6.4)$$

Valores cercanos a 1 para este indicador indican que la estimación de la estructura bajo análisis reporta niveles de incertidumbre similares al proceso en estudio. Por otro lado, valores menores que 1 reflejan que las estimaciones no son capaces de representar la dispersión real del proceso.

6.1.2. Predicción

A diferencia de la evaluación de las rutinas de filtrado, en aplicaciones de predicción de estado es de interés cuantificar el desempeño de las estructuras propuestas en el sentido de que la predicción del intervalo de confianza incluya a la mejor aproximación del estado (de acuerdo a lo expresado anteriormente). Este concepto tiene más sentido que construir un indicador basado en el error de la predicción del valor esperado del estado, pues como éste es un proceso estocástico, en general la probabilidad de que dicho proceso tome un

valor determinado es nula. En este contexto, una medida instantánea del desempeño de la predicción del estado en t , σ_t^2 , realizada en τ , i.e. $(t - \tau)$ instantes antes, se presenta en (6.5).

$$i_X^{FR}(\tau) = \sqrt{2} \frac{\sqrt{\left(\bar{\sigma}_t^2 - \widehat{CI}_X^{Sup}(t, \tau)\right)^2 + \left(\bar{\sigma}_t^2 - \widehat{CI}_X^{Inf}(t, \tau)\right)^2}}{\overline{CI}_X(t)} \quad (6.5)$$

donde $\bar{\sigma}_t^2$ es la mejor estimación de σ_t^2 , y $\widehat{CI}_X^{Inf}(t, \tau)$ y $\widehat{CI}_X^{Sup}(t, \tau)$ son respectivamente los límites inferior y superior de la estimación del CI de σ_t^2 realizada en τ .

Un índice que cuantifique el desempeño de la predicción del estado σ_t^2 durante un período de tiempo definido $[t_0, t]$, es presentado en (6.6)

$$\begin{aligned} I_X^{FR} &= \frac{1}{c} \sum_{\tau=t_0+1}^t i_X^{FR}(\tau) (\tau - t_0)^2 \\ &= \frac{\sqrt{2}}{c} \sum_{\tau=t_0+1}^t \frac{\sqrt{\left(\bar{\sigma}_t^2 - \widehat{CI}_X^{Sup}(t)\right)^2 + \left(\bar{\sigma}_t^2 - \widehat{CI}_X^{Inf}(t, \tau)\right)^2}}{\overline{CI}_X(t, \tau)} (\tau - t_0)^2 \end{aligned} \quad (6.6)$$

donde $t_0 \in \mathbb{R}^+$, y la introducción del término $(\tau - t_0)^2$ le da más importancia a las predicciones a “corto plazo”, i.e. $\tau \sim t$, pues es de esperar que las estimaciones realizadas a pocos instantes de t sean más acertadas. Además, $c = \sum_{\tau=t_0+1}^t (\tau - t_0)^2$ es una constante de normalización. Valores pequeños para este índice indican que las predicciones del intervalo de confianza del X % son precisas y que posiblemente contienen a $\hat{\sigma}_t^2(\tau)$, $\tau \in [t_0, t]$. Sin embargo, valores menores a 1 indican que las predicciones del intervalo de confianza no son representativas de la incertidumbre propia del proceso.

Con los índices presentados es posible realizar una comparación cuantitativa de los modelos y filtros introducidos en los Capítulos 3, 4 y 5, para el problema de la estimación de la volatilidad estocástica. Se debe poner énfasis en que la utilización de una estimación suavizada del estado es necesaria sólo en aplicaciones con datos financieros reales, pues cuando los algoritmos son evaluados en un ambiente de simulación, el estado a estimar $\sigma_{1:t}^2$ es conocido, y por ende, éste debe ser usado como señal de referencia absoluta. Es decir, es posible considerar $\bar{\sigma}_{1:t}^2 = \sigma_{1:t}^2$.

En los casos de series de retornos reales, es posible ocupar el modelo GARCH (3.3) ajustado fuera de línea mediante ML como estructura de estimación suavizada.

6.2. Resultados Usando Series de Tiempo Generadas

En esta sección se presentan resultados de la utilización de los modelos y filtros propuestos para resolver los problemas de filtrado y predicción de una señal de retornos generada mediante simulación. La consideración de resultados en un ambiente simulado permite una comprensión acabada del comportamiento de los métodos implementados, pues en este caso el proceso latente a estimar es conocido sin incertidumbre.

Ocupando el modelo logarítmico de volatilidad estocástica (3.15)-(3.16) y el modelo uGARCH (5.3)-(5.4) es posible generar secuencias de T muestras para la volatilidad, i.e. $\sigma_{1:T}^2$ y $V_{1:T}$, y construir un proceso de retornos $r_{1:T}$ según (3.2). Considerando dicho proceso como las observaciones de los retornos de cierto instrumento financiero, se han implementado estructuras, basadas en EKF y PF, para estimar y predecir el proceso de volatilidad generado.

6.2.1. Generación de Datos

Para evaluar la adaptatividad de las estructuras propuestas en simulación, las secuencias generadas $r_{1:T}$ y $\sigma_{1:T}^2$ (o $V_{1:T}$ en el caso del modelo de log-VE), consideran un cambio de parámetros.

Para validar el modelo de log-VE, se generó un proceso de $T = 500$ muestras $V_{1:500}$, con condición inicial $V_0 = -9,5$ según la estructura del modelo logarítmico de volatilidad estocástica, cambiando dos de sus parámetros para la mitad de las muestras. Las primeras 250 muestras ($t \in [1, 250]$) fueron generadas según el modelo:

$$V_t = -1 + 0,85V_{t-1} + \eta_t, \eta_t \sim \mathcal{N}\left(0, \frac{1}{10}\right) \quad (6.7)$$

$$r_t = \exp\left(\frac{1}{2}V_t\right)\epsilon_t, \epsilon_t \sim \mathcal{N}(0, 1) \quad (6.8)$$

mientras que las 250 muestras restantes ($t \in [251, 500]$) fueron generadas según:

$$V_t = -1 + 0,8V_{t-1} + \eta_t, \eta_t \sim \mathcal{N}\left(0, \frac{1}{5}\right) \quad (6.9)$$

$$r_t = \exp\left(\frac{1}{2}V_t\right)\epsilon_t, \epsilon_t \sim \mathcal{N}(0, 1). \quad (6.10)$$

De forma similar al modelo de log-VE, para la validación del modelo uGARCH, se generó un proceso de $T = 500$ muestras $\sigma_{1:500}^2$, con condición inicial $\sigma_0^2 = 0,002$ según la estructura del modelo GARCH(1,1), cambiando dos de sus parámetros para la mitad de las muestras. Las primeras 250 muestras ($t \in [1, 250]$) fueron generadas según el modelo:

$$\sigma_t^2 = 10^{-6} + 0,6\sigma_{t-1}^2 + 0,2u_{t-1}^2 \quad (6.11)$$

$$r_t = 9 \cdot 10^{-4} + \sigma_t\epsilon_t, \epsilon_t \sim \mathcal{N}(0, 1) \quad (6.12)$$

mientras que las 250 muestras restantes ($t \in [251, 500]$) fueron generadas según:

$$\sigma_t^2 = 10^{-6} + 0,85\sigma_{t-1}^2 + 0,1u_{t-1}^2 \quad (6.13)$$

$$r_t = 9 \cdot 10^{-4} + \sigma_t\epsilon_t, \epsilon_t \sim \mathcal{N}(0, 1). \quad (6.14)$$

Considerando subconjuntos de las series generadas según los modelos presentados, es posible ajustar modelos para implementar rutinas de filtrado y predicción. En la siguiente sección se presenta la identificación de dichos modelos.

6.2.2. Ajuste de Modelos para Estimación de Estado

Para implementar rutinas de filtrado utilizando los modelos log-VE y uGARCH, en algoritmos PF y EKF, es necesario definir valores para los parámetros del modelo y condiciones iniciales para el hiperestado. En el caso del modelo log-VE, es posible utilizar el subconjunto del proceso generado según (6.7)-(6.10) $r_{1:150}$ para ajustar un modelo GARCH(1,1) mediante ML, el cual reporta una estimación $\bar{\sigma}_{1:150}^2$ de la volatilidad. Luego, considerando $\bar{\sigma}_{1:150}^2$, es posible estimar el proceso latente de volatilidad logarítmica en el período $t \in (1, 150)$ $V_{1:150}$

asumiendo:

$$V_{1:150} \simeq \bar{V}_{1:150} = 2 \log(\bar{\sigma}_{1:150}), \quad (6.15)$$

y ajustar el modelo log-VE mediante ML (pues de esta forma se dispone de una estimación del proceso latente). Además, la condición inicial para el estado será considerada según:

$$V_0 = \log(r_1^2). \quad (6.16)$$

Con estas consideraciones, el modelo a usar para estimar la volatilidad logarítmica, y los parámetros del modelo, usando filtro de partículas es:

$$V_t = \gamma_t + \phi_t V_{t-1} + \eta_t, \eta_t \sim \mathcal{N}(0, \sigma_t^\eta) \quad (6.17)$$

$$\gamma_{t+1} = \gamma_t + \eta_t^{(\gamma)} \quad (6.18)$$

$$\phi_{t+1} = \phi_t + \eta_t^{(\phi)} \quad (6.19)$$

$$\sigma_{t+1}^\eta = \sigma_t^\eta + \eta_t^{(\sigma^\eta)} \quad (6.20)$$

$$r_t = \exp\left(\frac{1}{2}V_t\right) \epsilon_t \quad (6.21)$$

donde los procesos de ruido $\eta_t^{(\gamma)}$, $\eta_t^{(\phi)}$, $\eta_t^{(\sigma^\eta)}$ son normales de media nula y baja varianza, con respecto a la magnitud de sus respectivas condiciones iniciales², siendo éstas últimas:

$$V_0 = -11,2 \quad (6.22)$$

$$\gamma_0 = -1,08 \quad (6.23)$$

$$\phi_0 = 0,85 \quad (6.24)$$

$$\sigma_0^\eta = 0,37 \quad (6.25)$$

Por otro lado, los resultados de la consideración del modelo (6.17)-(6.25) en una rutina de estimación basada en EKF no son óptimos, pues éste esquema asume que el sistema dinámico sólo considera ruido aditivo, por lo que el parámetro σ_t^η no podría ser estimado. Además, resultados experimentales preliminares evidencian que el asumir γ_t variante en el tiempo introduce ruido excesivo en la estimación al usar EKF.

²10% de la condición inicial de cada parámetro.

Con estas consideraciones, el modelo a usar para filtrar la volatilidad logarítmica usando el filtro extendido de Kalman es:

$$V_t = -1,08 + \phi_t V_{t-1} + \eta_t, \eta_t \sim \mathcal{N}(0, 0,37) \quad (6.26)$$

$$\phi_{t+1} = \phi_t + \eta_t^{(\phi)} \quad (6.27)$$

$$r_t = \exp\left(\frac{1}{2}V_t\right) \epsilon_t \quad (6.28)$$

donde los procesos de ruido y las condiciones iniciales V_0 y ϕ_0 cumplen con las mismas propiedades que el modelo anterior.

Nótese que el resultado de aplicar ML con el fin de obtener los modelos dinámicos que se utilizarán en las rutinas de estimación, implican ligeras diferencias entre las estimaciones de los valores óptimos de las condiciones iniciales (6.23)-(6.25), y los valores reales de los parámetros que se definen en (6.7), a pesar del supuesto introducido en (6.15). Sin embargo, en el caso particular de la condición inicial (6.25), pueden apreciarse diferencias más significativas debido a que el parámetro σ^n está asociado a un regresor desconocido, lo que conlleva un problema de observabilidad.

Con respecto de la implementación del modelo uGARCH en las rutinas de estimación de estado, del proceso de retornos generado según (6.11)-(6.14) $r_{1:500}$, el subconjunto $r_{1:150}$ se utilizó para identificar un modelo determinístico GARCH(1,1) mediante ML. El sistema identificado fue:

$$\sigma_t^2 = 8,4 \cdot 10^{-7} + 0,6756\sigma_{t-1}^2 + 0,184u_{t-1}^2 \quad (6.29)$$

$$r_t = 7,7 \cdot 10^{-5} + \sigma_t \epsilon_t. \quad (6.30)$$

De forma similar a la identificación del modelo log-VE, las estimaciones de los parámetros son bastante similares a los valores reales, de hecho, las estimaciones de $\alpha = 0,2$ y $\beta = 0,6$ tienen un error menor a 15 %. Se pone énfasis en estas cantidades ya que son las responsables de la estabilidad del modelo. Además, éstas fueron consideradas como parámetros variantes en el tiempo tal como se verá a continuación.

Considerando que las estructuras de los modelos GARCH y uGARCH son iguales salvo

las propiedades de la distribución del proceso de innovación³, los parámetros en el modelo identificado (6.29)-(6.30) fueron utilizados como condiciones iniciales en un modelo uGARCH de parámetros variantes en el tiempo. Este modelo sirvió de estructura subyacente para la implementación de rutinas de filtrado utilizando PF y EKF, y está dado por:

$$\alpha_t = \alpha_{t-1} + \eta_{t-1}^{(\alpha)} \quad (6.31)$$

$$\beta_t = \beta_{t-1} + \eta_{t-1}^{(\beta)} \quad (6.32)$$

$$\sigma_t^2 = 8,4 \cdot 10^{-7} + \beta_t \sigma_{t-1}^2 + \alpha_t \sigma_{t-1}^2 \eta_{t-1}^2 \quad (6.33)$$

$$r_t = 7,7 \cdot 10^{-5} + \sigma_t \epsilon_t. \quad (6.34)$$

donde los procesos de ruido $\eta_t^{(\alpha)}$ y $\eta_t^{(\beta)}$ son Gaussianos de baja varianza con respecto a la magnitud de sus respectivas condiciones iniciales, y $\eta_t^2 \sim \mathcal{N}(0, 1)$. Además, las condiciones iniciales están dadas por:

$$\alpha_0 = 0,184 \quad (6.35)$$

$$\beta_0 = 0,6756 \quad (6.36)$$

$$\sigma_0^2 = r_1^2 \quad (6.37)$$

Con los modelos log-VE y uGARCH identificados, es posible implementar rutinas de estimación y predicción de la volatilidad estocástica de las series generadas según (6.7)-(6.10) y (6.11)-(6.14). En las siguientes secciones, se presentan dichos resultados utilizando las estructuras de estimación PF y EKF.

³Como se explicó en el Capítulo 5

6.2.3. Filtrado

6.2.3.1. log-VE

En la Figura 6.1 se presentan los procesos generados $\sigma_{1:500}^2 = 2 \log(V_{1:500})$ y $r_{1:500}$, además de las estimaciones del valor esperado del estado y su intervalo de confianza del 95% para todo el período simulado $[1,500]$ utilizando PF y EKF. En esta figura se puede ver claramente el cambio de parámetros introducido en el proceso simulado a partir de $t = 250$, y cómo responde cada filtro a este cambio de comportamiento. La media de la estimación de PF es similar al proceso generado por el modelo log-VE, asimilando rápidamente el cambio de parámetros. Por otro lado, EKF reporta estimaciones altamente sesgadas y su estimación del intervalo de confianza no comprende al proceso generado en la mayoría del período considerado. Adicionalmente, como las estimaciones de PF se encuentran en la forma de distribuciones discretas, es posible construir versiones continuas de éstas mediante la utilización del *kernel* de Epanechnikov (Li & Racine 2007). En la Figura 6.2 se muestran estas distribuciones para los 30 primeros instantes del período considerado, donde se ve cómo las estimaciones de PF abandonan el valor de la condición inicial a medida que se reciben nuevas observaciones.

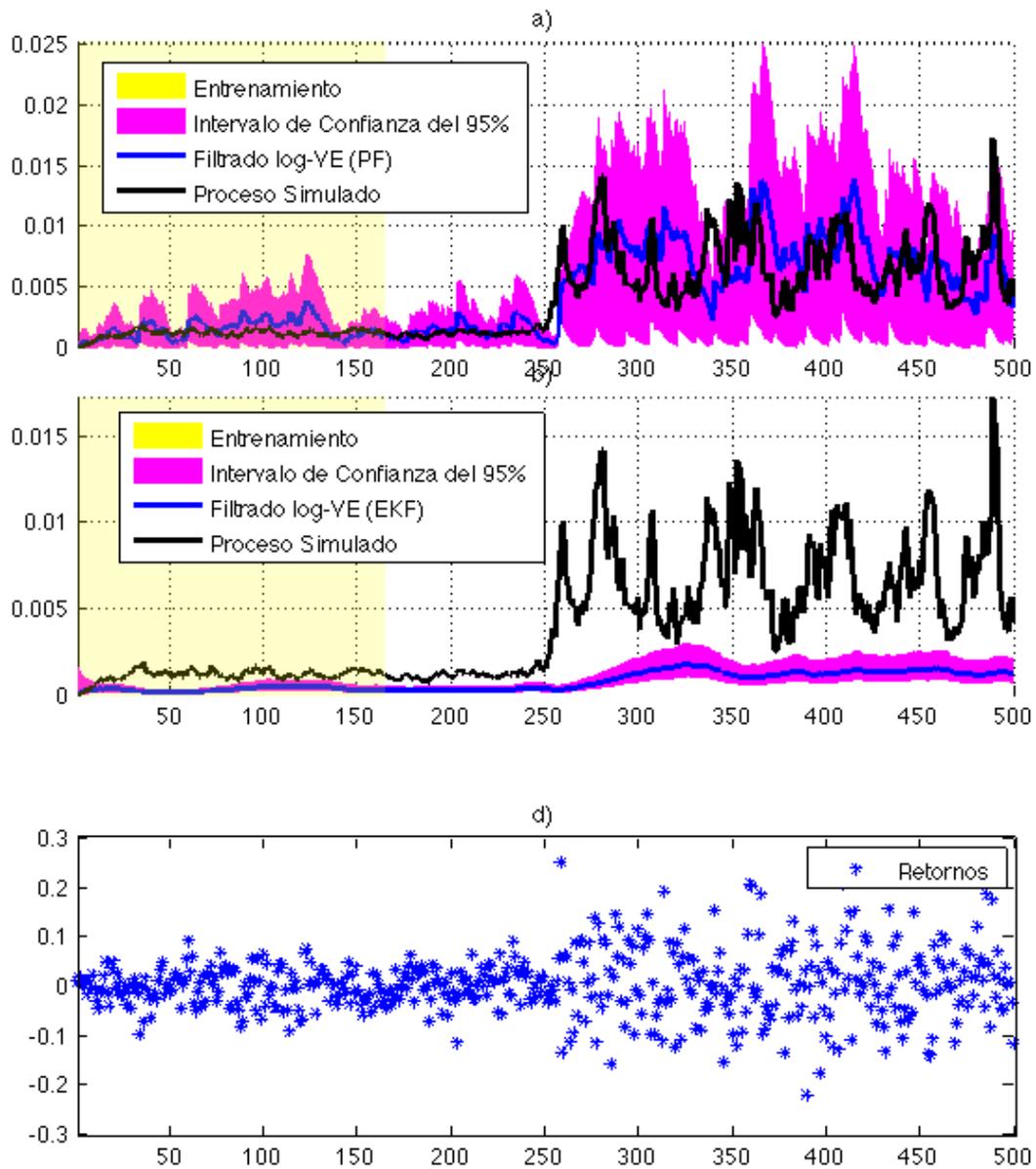


Figura 6.1: Resultados para la estimación de la volatilidad estocástica de la serie autogenerada $r_{1:500}$ usando el modelo log-VE. **a)** Negro: Volatilidad autogenerada con modelo log-VE. Azul: Estimación de PF. Magenta: Estimación de PF del Intervalo de confianza del 95%. **b)** Negro: Volatilidad autogenerada con modelo log-VE. Azul: Estimación de EKF. Magenta: Estimación de EKF del Intervalo de confianza del 95%. **c)** Proceso de retornos generado

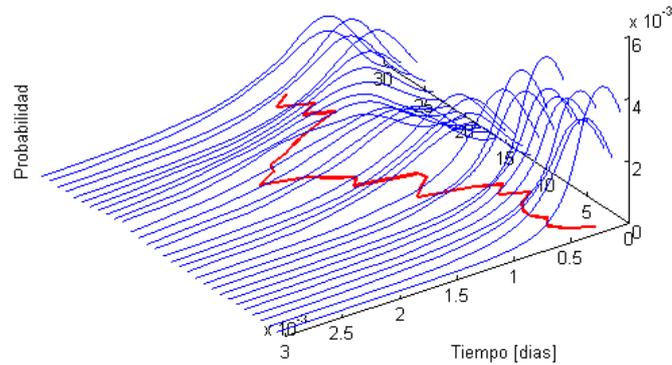


Figura 6.2: Versiones continuas de las estimaciones de la distribución de la volatilidad usando PF y el modelo log-VE para los 30 primeros instantes del proceso simulado. En rojo se muestran las medias.

En cuanto a la identificación paramétrica, ésta fue realizada adaptativamente en base a evolución artificial con los respectivos filtros EKF y PF. Estas estimaciones se presentan en la Figura 6.3. Se debe notar que la implementación basada en EKF sólo estima el parámetro ϕ_t por las razones que se enunciaron anteriormente, no así PF, el cual es capaz de estimar todos los parámetros. La Figura 6.3 muestra cómo el cambio de parámetros fue detectado por ambos filtros, pues a pesar del ruido presente en las estimaciones es posible notar una disminución en la estimación de ϕ_t y una aumento para σ_t^η en $t = 250$.

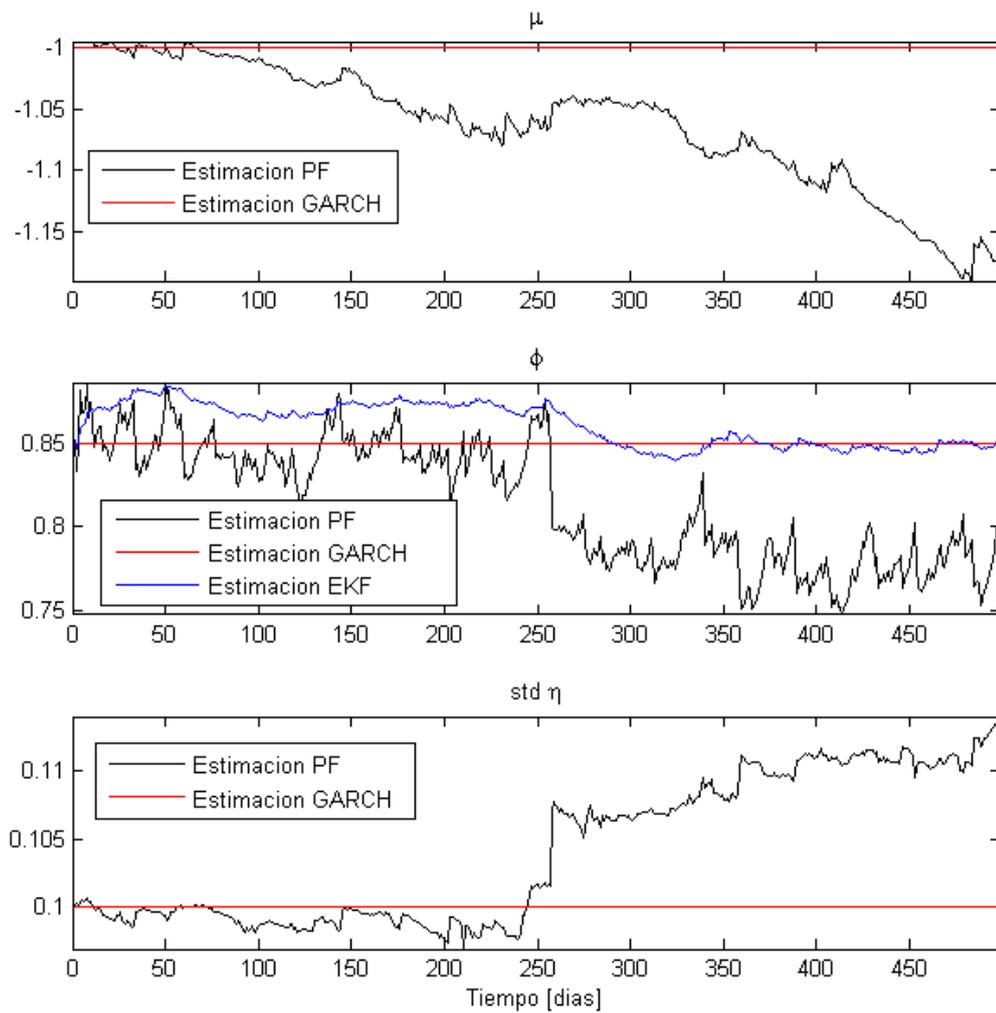


Figura 6.3: Estimaciones de los parámetros γ , ϕ y σ^η del modelo log-VE usando serie de retornos simulada. Rojo: Estimación de máxima verosimilitud para el modelo GARCH usando $r_{1:150}$ (offline). Azul: Estimación online EKF. Negro: Estimación online PF.

6.2.3.2. uGARCH

En la Figura 6.4 se presentan los procesos generados por el modelo GARCH $\sigma_{1:500}^2$ y $r_{1:500}$, además de las estimaciones del valor esperado del estado y su intervalo de confianza del 95%, para todo el período simulado $[1,500]$ realizadas por PF y EKF, usando el modelo uGARCH. En esta figura se puede ver que, a pesar del cambio de parámetros en la generación

del proceso $\sigma_{1:500}^2$, el filtrado usando PF sigue a dicho proceso de forma muy exacta, y su estimación del CI siempre abarca al proceso generado. Por otro lado, la estimación reportada por EKF es evidentemente sesgada, y su estimación del intervalo de confianza no comprende a $\sigma_{1:500}^2$ durante la mayoría del período considerado.

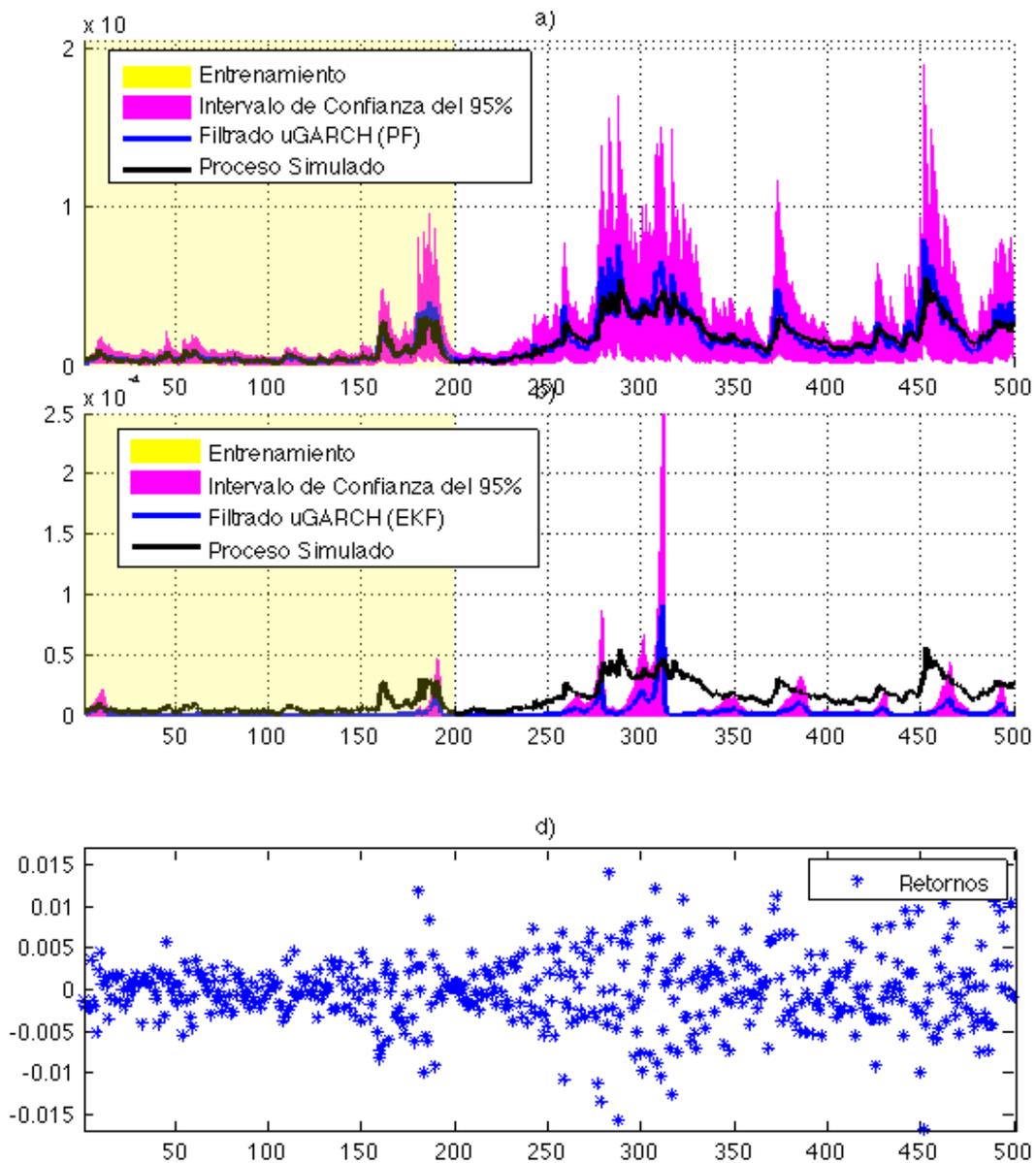


Figura 6.4: Resultados para la estimación de la volatilidad estocástica de la serie simulada $r_{1:500}$ usando el modelo GARCH. **a)** Negro: Volatilidad autogenerada con modelo GARCH. Azul: Estimación de PF y uGARCH. Magenta: Estimación de PF y uGARCH del Intervalo de confianza del 95 %. **b)** Negro: Volatilidad autogenerada con modelo GARCH. Azul: Estimación de EKF y uGARCH. Magenta: Estimación de EKF y uGARCH del Intervalo de confianza del 95 %. **c)** Proceso de retornos generado mediante GARCH.

Versiones continuas de la densidad estimada mediante PF (*kernel* de Epanechnikov) se muestran en la Figura 6.5 para las 30 primeras estimaciones, en esta figura se puede ver

cómo las estimaciones se alejan de la media y varianza iniciales a medida se reciben las observaciones de los retornos.

Además, en este caso los parámetros también han sido estimados adaptativamente mediante estimación artificial con los filtros EKF y PF. Estas estimaciones se presentan en la Figura 6.6, donde se puede ver: (i) el cambio de comportamiento a partir de $t = 250$, sugiriendo un cambio de parámetros, y (ii) el problema de observabilidad asociado a la estimación conjunta de los parámetros α y β , tal como fue mencionado en el Capítulo 5 .

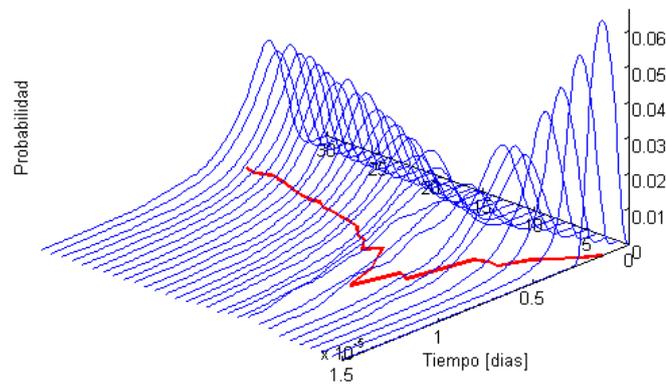


Figura 6.5: Versiones continuas de las estimaciones de la distribución de la volatilidad usando PF y el modelo uGARCH para los 30 primeros instantes usando la serie de retornos simulada. En rojo se muestran las medias.

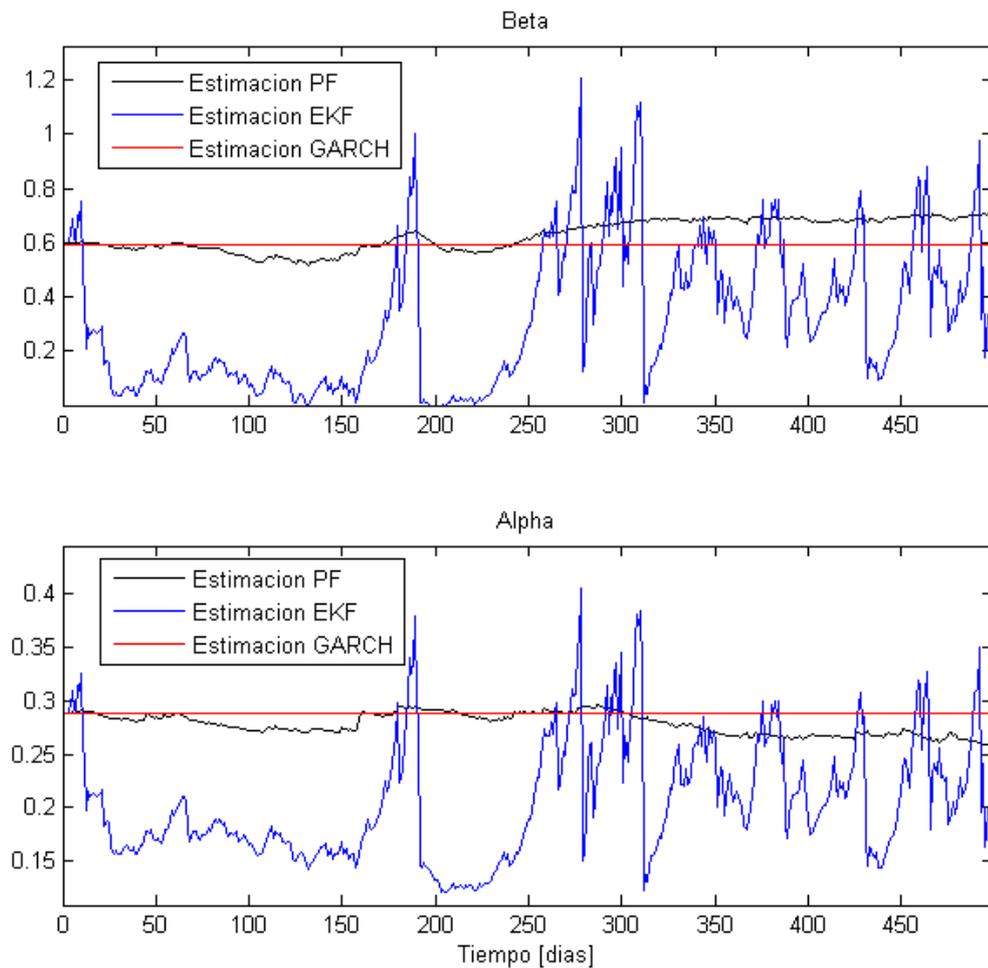


Figura 6.6: Estimaciones de los parámetros α y β del modelo uGARCH para una serie de retornos simulada. Rojo: Estimación de máxima verosimilitud para el modelo GARCH usando $r_{1:80}$ (offline). Azul: Estimación online EKF. Negro: Estimación online PF.

6.2.4. Predicción

Basado en resultados preliminares, el criterio para realizar las predicciones usando filtro de partículas se basó en realizar únicamente la etapa de predicción (4.13), ignorando la etapa de actualización (4.14). Esto es equivalente a realizar ambas etapas y asumir que la densidad de verosimilitud de las observaciones $p(r_t|\sigma_t^2)$ es uniforme; es decir, se asumió que cualquier valor de σ_t^2 tiene la misma verosimilitud de generar los retornos (no observados en el problema de

predicción). Con este supuesto, ni siquiera es necesario calcular el valor esperado de dichos retornos. Igualmente, el criterio para predecir la volatilidad usando el filtro extendido de Kalman fue considerar sólo la etapa de predicción; esto es asumir que el valor esperado de los retornos es igual a la estimación predicha, lo que implica que el término de error en (4.31) sea nulo.

Para implementar dichas rutinas de predicción de la volatilidad estocástica en un ambiente simulado, se han considerado los procesos $r_{1:200}$, $\sigma_{1:200}^2$ generados por los modelos (6.7)-(6.8) (log-VE), y (6.11)-(6.12) (uGARCH). Además, el subconjunto $r_{1:150}$ se ha utilizado para identificar un conjunto apropiado de parámetros (o de las condiciones iniciales de los parámetros cuando corresponda) para predecir la volatilidad al instante $t = 200$, σ_{200}^2 considerando muestras según $\tau \in [150, 200]$. Los resultados de esta rutina de predicción para cada uno de los modelos introducidos en los Capítulos 3 y 5 se presentan a continuación.

6.2.4.1. log-VE

El modelo utilizado para generar los procesos $r_{1:200}$ y $\sigma_{1:200}^2 = \exp(V_{1:200})$ fue (6.7)-(6.8). Además, la identificación de cada modelo log-VE (para utilizar en conjunto con PF y con EKF) fue realizada considerando el subconjunto $r_{1:150}$ de la misma forma que en la Subsección 6.2.2.

En la Figura 6.7 se presentan los resultados de la predicción de la señal generada $\sigma_{150:200}^2$, utilizando las observaciones $r_{1:150}$, y se puede ver cómo las predicciones tienden al punto de equilibrio a medida que el horizonte de predicción aumenta. Estos puntos de equilibrio son distintos debido a que los enfoques que utilizan EKF y PF usan distintos modelos dinámicos (EKF no considera todos los parámetros como variantes en el tiempo), y distintos criterios de predicción.

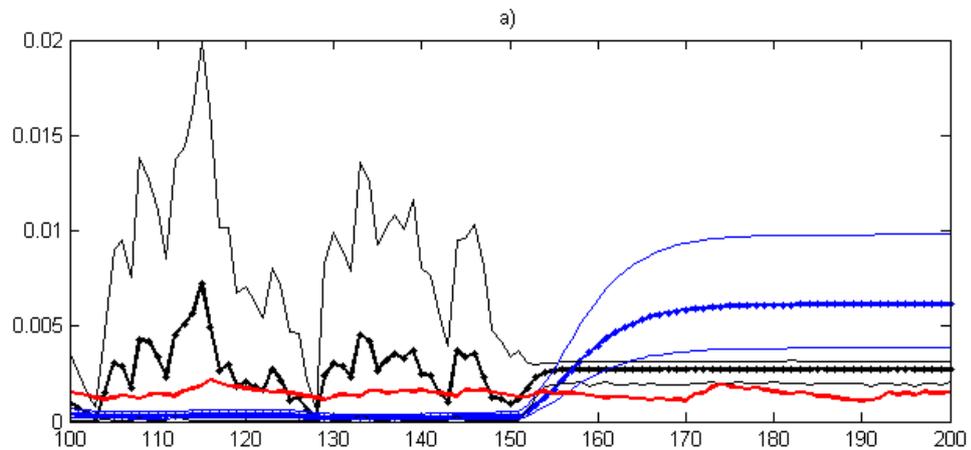


Figura 6.7: Predicción de la volatilidad estocástica autogenerada usando el modelo log-VE para $t \in [150, 200]$. Rojo: Proceso generado. Negro: Estimación de PF de media (grosso) y CI del 95 % (delgado). Azul: Estimación de EKF de media (grosso) y CI del 95 % (delgado). Para $t \leq 150$ la estimación corresponde a filtrado y para $t > 150$ a predicción considerando sólo la información hasta $t = 150$ (Abscisas = t).

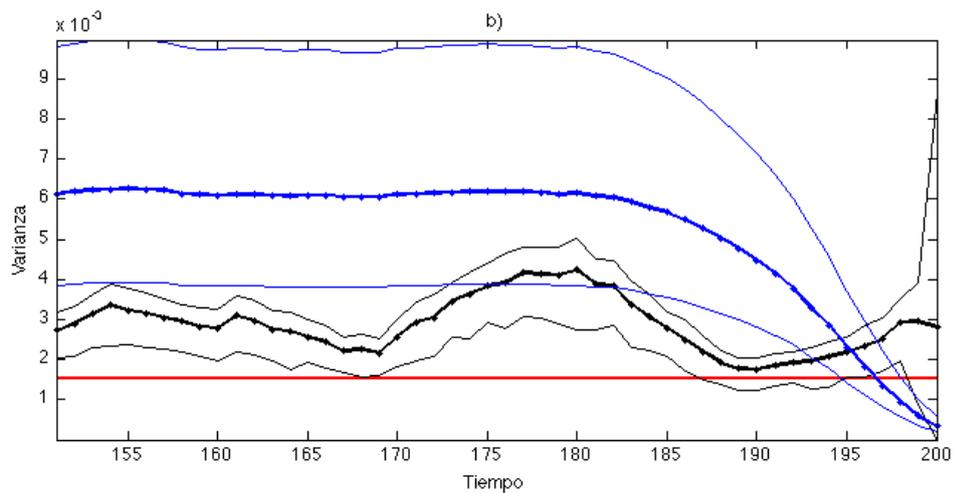


Figura 6.8: Predicción de la volatilidad estocástica autogenerada usando el modelo log-VE para $t = 200$ en función del horizonte de predicción. Rojo: Valor de la volatilidad estocástica en $t = 200$, σ_{200}^2 . Negro: Predicción del valor esperado (grosso) y del CI del 95 % (delgado) de σ_{200}^2 usando muestras hasta $\tau \in [150, 200]$ usando PF. Azul: Predicción del valor esperado (grosso) y del CI del 95 % (delgado) de σ_{200}^2 usando muestras hasta $\tau \in [150, 200]$ usando EKF (Abscisas = τ).

Con respecto de la mejora de las predicciones de σ_{200}^2 a medida que el horizonte de predicción disminuye, en la Figura 6.8 se ve que para ambas estructuras, EKF y PF, el

valor esperado de las estimaciones se aproxima a σ_{200}^2 . Además, se ve que el tamaño del intervalo de confianza estimado por PF –a diferencia del estimado por EKF– presenta una baja dependencia con respecto al horizonte de predicción. Es importante notar también que la media de la estimación proporcionada por PF no mejora monótonamente, esto se tiene porque PF incorpora rápidamente los cambios en el proceso de retornos observado, por lo que el valle en la señal generada σ_t^2 poco antes de $t = 170$ ciclos provoca que las predicciones tengan un valor menor.

6.2.4.2. uGARCH

Siguiendo la lógica de la implementación de la rutina de predicción basada en el modelo log-VE, el modelo para generar los procesos $r_{1:200}$ y $\sigma_{1:200}^2$ en esta etapa fue (6.11)-(6.12). Además, la identificación de cada modelo uGARCH (para PF y EKF) también fue realizada considerando el subconjunto $r_{1:150}$.

En la Figura 6.9 se puede ver, como es de esperar, que las predicciones igualmente tienden a su punto de equilibrio. A diferencia de la implementación basada en el modelo log-VE, en este caso los puntos de equilibrios son muy similares, esto es debido a que ambas estructuras de estimación consideran a ambos parámetros como variantes en el tiempo, a pesar de que las estimaciones de los parámetros por ambas estructuras no son iguales (ver Figura 6.6), pues la estimación de los parámetros del modelo uGARCH con el EKF tiene problemas de observabilidad tal como se comentó en el Capítulo 5.

Además, de la Figura 6.10 se puede ver que la predicción de PF mejora de forma monótona a medida que el horizonte de predicción disminuye, a diferencia de la predicción de EKF. Es importante notar que el intervalo de confianza del 95 % predicho por PF incluye a σ_{200}^2 para horizontes de predicción menores a 20 ciclos, en cambio el predicho por EKF no se aproxima a σ_{200}^2 , lo que era de esperar basado en las estimaciones presentadas en la Figura 6.4 b).

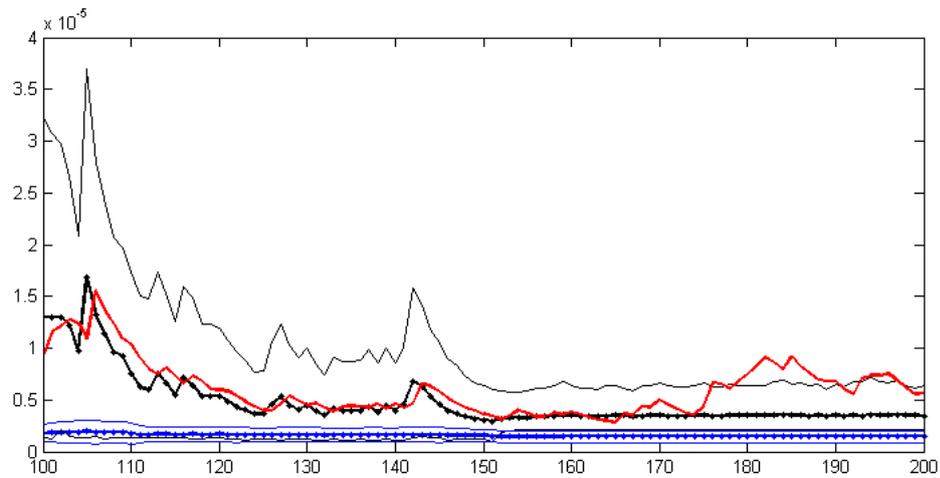


Figura 6.9: Predicción de la volatilidad estocástica autogenerada usando el modelo uGARCH para $t \in [150, 200]$. Rojo: Proceso generado. Negro: Estimación de PF de media (grosso) y CI del 95 % (delgado). Azul: Estimación de EKF de media (grosso) y CI del 95 % (delgado). Para $t \leq 150$ la estimación corresponde a filtrado y para $t > 150$ a predicción considerando sólo la información hasta $t = 150$ (Abcisas = t).

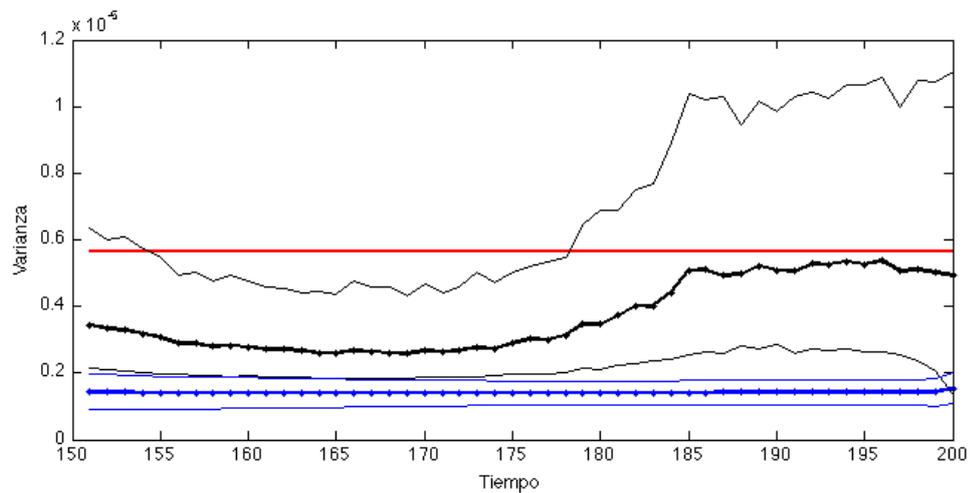


Figura 6.10: Predicción de la volatilidad estocástica autogenerada usando el modelo uGARCH para $t = 200$ en función del horizonte de predicción. Rojo: Valor de la volatilidad estocástica en $t = 200$, σ_{200}^2 . Negro: Predicción del valor esperado (grosso) y del CI del 95 % (delgado) de σ_{200}^2 usando muestras hasta $\tau \in [150, 200]$ usando PF. Azul: Predicción del valor esperado (grosso) y del CI del 95 % (delgado) de σ_{200}^2 usando muestras hasta $\tau \in [150, 200]$ usando EKF (Abcisas = τ).

6.2.5. Evaluación

Utilizando los indicadores presentados en la Sección 6.1 es posible cuantificar el desempeño de los modelos y filtros presentados en un ambiente de simulación. En la Tabla 6.1 se presentan dichos indicadores para cada uno de los casos analizados en las secciones anteriores. De la información en esta Tabla, se puede inferir que –para el problema de filtrado– las estimaciones de PF son más exactas que las de EKF para ambos modelos. Sin embargo, las estimaciones de EKF son más precisas que las de PF, lo cual no es tan bueno como parece, pues las estimaciones de EKF son altamente sesgadas. Finalmente, las estimaciones usando el modelo uGARCH son en general más precisas y exactas que las hechas usando el modelo log-VE para cualquier filtro.

Tabla 6.1: Índices de desempeño para resultados de filtrado y predicción de una señal simulada de retornos usando los modelos log-VE y uGARCH, y los filtros PF y EKF.

Modelo	log-VE		uGARCH	
	PF	EKF	PF	EKF
I_{95}^{AC}	0.57	1.98	0.23	1.62
I_{95}^{PR}	2.9	0.84	1.91	1.31
I_{95}^{FR}	5.9	14.7	1.896	2.195

En cuanto al desempeño de las estructuras propuestas para el problema de predicción, se puede ver que el modelo uGARCH reporta mejores predicciones que el modelo log-VE en el sentido del tamaño del CI predicho \widehat{CI}_X , y su posición con respecto al proceso de volatilidad generado. De hecho, el modelo uGARCH muestra valores cercanos a 1 para el índice I_{95}^{FR} , lo que quiere decir que la mayor parte del tiempo, \widehat{CI}_X comprende a la estimación suavizada tal como se vió en la Figura 6.10.

En esta sección, se han presentado, comparado, y evaluado los resultados del filtrado y predicción de la volatilidad estocásticas de un proceso simulado de retornos utilizando dos estructuras de modelación, y dos estructuras de filtrado. A continuación se presenta un análisis similar, pero utilizando una serie de retornos reales.

6.3. Resultados para el Índice *NASDAQ Composite*

En funcionamiento desde el 5 de Febrero de 1971, el índice *NASDAQ Composite* (INC) representa a los más de 3000 componentes que cotizan en el mercado *NASDAQ*⁴, y es mundialmente reconocido como indicador del funcionamiento del mercado de la tecnología. El estudio de la volatilidad de INC –motivado por inversionistas interesados en transar derivados sobre éste y por agentes que pretenden entrar al mercado de la tecnología– es un desafío interesante dadas las fuertes fluctuaciones y los altos niveles de incertidumbre que este índice ha mostrado desde sus inicios. Estas características convierten a INC en un buen candidato para validar las estructuras propuestas.

6.3.1. Preprocesamiento

En vista de que los modelos presentados en los Capítulos 3 y 5 asumen que el proceso observado es una serie de tiempo de retornos, los datos disponibles –en forma de series de precios, o índices– deben ser convertidos a retornos. La conversión en este trabajo fue realizada usando composición continua⁵, es decir, los precios p_t y retornos r_t están relacionados mediante:

$$r_t = \log\left(\frac{p_{t+1}}{p_t}\right) \quad (6.38)$$

En la Figura 6.11 se puede ver la serie de precios de cierre diario del índice *NASDAQ Composite*⁶, y sus respectivos retornos utilizando (6.38) para los días hábiles de los últimos 4 años. En esta figura, se ven los primeros indicios del comienzo de la recesión de fines del 2008. La muestra 600 representa Septiembre 4, 2008.

⁴*National Association of Securities Dealers Automated Quotations*

⁵*Continuous compounding.*

⁶Obtenida de <http://finance.yahoo.com>

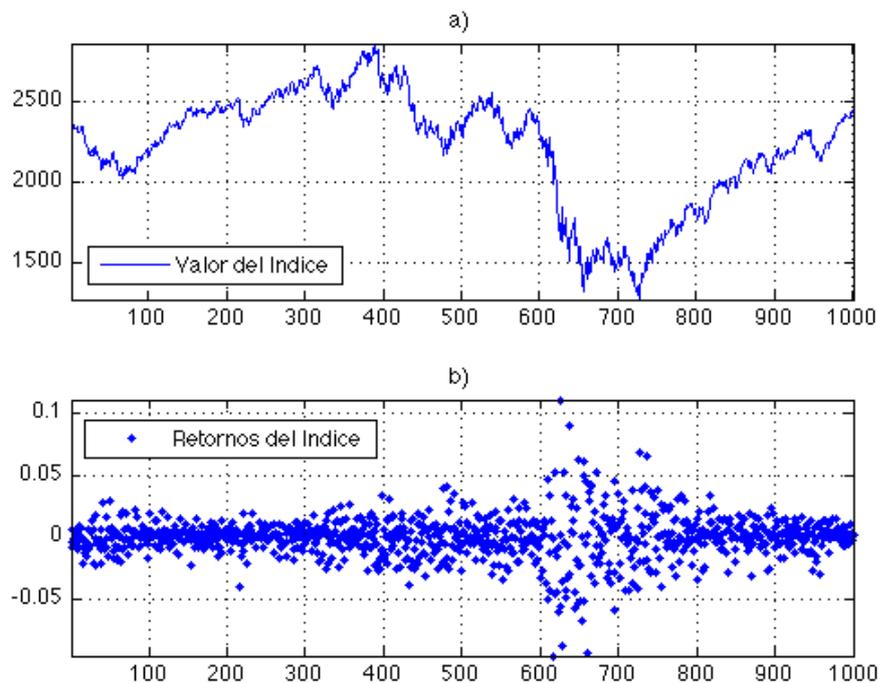


Figura 6.11: Índice *NASDAQ Composite* y sus retornos entre Abril 18, 2006 y Abril 8, 2010.

Además, para justificar el uso de modelos autoregresivos para la volatilidad (o la varianza), es necesario verificar que la serie de INC presenta *clustering* de volatilidad. Para verificar esto, considerése la autocorrelación de los retornos y de los cuadrados de éstos.

En la Figura 6.12 se puede ver que si bien los retornos no presentan una autocorrelación significativa, los cuadrados de éstos sí lo hacen, lo que motiva el uso de modelos autoregresivos para la varianza. Se debe notar además que la autocorrelación de los cuadrados de los retornos presenta un lento decaimiento, lo que es usual en series financieras y representa cierto grado de no estacionariedad/markovianidad común en las series de retornos financieros.

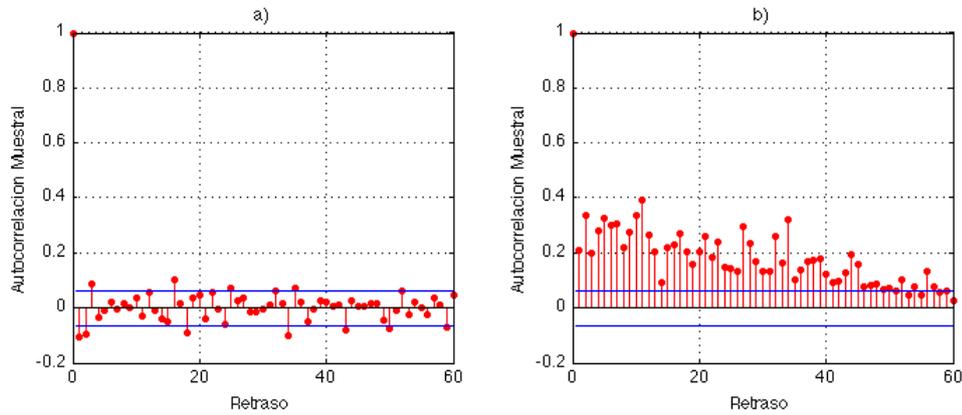


Figura 6.12: Funciones de autocorrelación muestral para **a)** el proceso de retornos y **b)** el proceso de los cuadrados de los retornos del índice *NASDAQ Composite* entre Abril 18, 2006 y Abril 8, 2010.

Continuando con la justificación de la utilización de modelos de volatilidad estocástica (i.e. uGARCH y log-VE), se deben verificar los estadísticos del proceso de retornos. Es posible estimar los momentos de orden superior (3er y 4to) con las medidas estándar, o convencionales, con la finalidad de tomar en cuenta toda la información contenida en los *outliers*⁷, i.e.:

$$Skewness \simeq \frac{1}{M} \sum_{t=1}^M \left(\frac{r_t - \hat{r}}{\hat{s}} \right)^3 \quad (6.39)$$

$$Kurtosis \simeq \frac{1}{M} \sum_{t=1}^M \left(\frac{r_t - \hat{r}}{\hat{s}} \right)^4 - 3 \quad (6.40)$$

donde $t = 1$ corresponde a Abril 18, 2006; el instante $M = 1000$ corresponde a Abril 8, 2010; r_t es el retorno del índice en el instante t , y tanto la media muestral \hat{r} , como la desviación estándar muestral \hat{s} están dadas por:

$$\hat{r} = \frac{1}{M} \sum_{t=1}^M r_t \quad (6.41)$$

$$\hat{s} = \frac{1}{M} \sum_{t=1}^M (r_t - \hat{r})^2 \quad (6.42)$$

⁷O realizaciones aisladas del proceso

Para una revisión sobre distintas medidas robustas de los 3er y 4to momentos, referirse a (Kim & White 2004).

Tabla 6.2: Mediana y 4 primeros momentos para los retornos del índice *NASDAQ Composite* entre Abril 18, 2006 y Abril 8, 2010

Mediana	Media	Varianza	<i>Skewness</i>	<i>Kurtosis</i>
0.0010	3.3665e-05	3.0177e-04	-0.1381	8.9040

Con estas consideraciones, en la Tabla 6.2 se presentan la mediana y los 4 primeros momentos muestrales para los retornos de índice dentro del período mencionados. Para verificar si estos valores efectivamente representan un exceso de 3er y 4to momento, considérense las estimaciones de los errores estándar de *skewness* (ses) y *kurtosis* (sek) presentadas en (Tabachnick & Fidell 1996), definidas por:

$$ses = \sqrt{\frac{6}{M}} = 0,0775 \quad (6.43)$$

$$sek = \sqrt{\frac{24}{M}} = 0,1549 \quad (6.44)$$

Además, basándose en el criterio de que para que exista efectivamente presencia de 3er y 4to momentos, las aproximaciones de éstos (Tabla 6.2) deben mayores que 2 veces sus respectivos errores estándar. Con esto, es posible afirmar que dichos valores sugieren que la serie de retornos considerada efectivamente presenta alta *kurtosis*. En cambio la presencia de asimetría no es tan evidente, pues la *skewness* muestral está dentro del rango considerado como error de aproximación.

Finalmente, en la Figura 6.13 se muestra un histograma de la serie de retornos y una distribución normal de igual media y varianza. En esta figura se puede ver claramente que la asimetría de los retornos no es considerable, en cambio las altas frecuencias en torno a la media evidencian la presencia de *excess kurtosis* con respecto a las distribución normal.

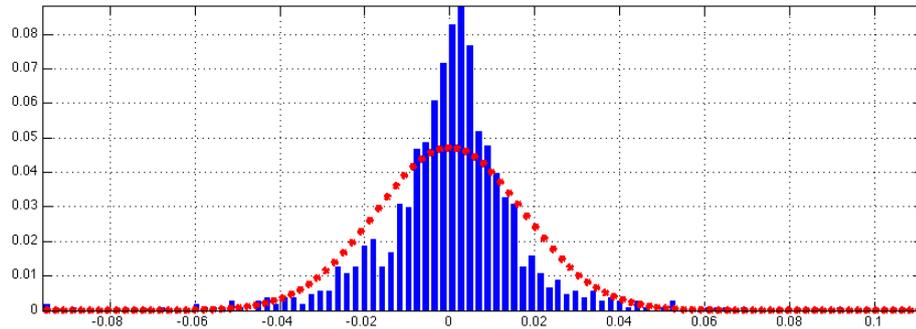


Figura 6.13: Histograma de los retornos del índice *NASDAQ Composite* entre Abril 18, 2006 y Abril 8, 2010. En rojo se presenta una distribución normal de igual media y varianza.

Con los resultados presentados en esta sección, es posible justificar la utilización de los modelos propuestos en los Capítulos 3 y 5, en conjunto con las estructuras de estimación de estado presentadas en el Capítulo 4. A continuación, se implementan dichos enfoques en el contexto del filtrado y la predicción de la volatilidad del índice *NASDAQ Composite*.

6.3.2. Filtrado

Para evaluar los resultados de la utilización de los modelos log-VE y uGARCH en el contexto del filtrado de la volatilidad estocástica de INC, se consideró su valor de cierre diario desde Julio 21, 2008 hasta Julio 17, 2009 (250 muestras correspondientes sólo a días hábiles, compuestas continuamente, y denotadas por $r_{1:250}$), obtenido desde <http://finance.yahoo.com>. La motivación para considerar este intervalo de tiempo es evaluar la respuesta de las estructuras de filtrado frente a cambios significativos en el índice, ya que el instante $t = 50$ del período considerado representa Septiembre 30, 2008; es decir, la fecha aproximada del inicio de la crisis 2008-2010 que comenzó en EUA⁸

6.3.2.1. log-VE

De la serie de retornos $r_{1:250}$, se consideró el subconjunto correspondiente a las primeras 80 muestras, i.e. $r_{1:80}$, para identificar un modelo GARCH(1,1), y posteriormente ajustar los

⁸De acuerdo al *National Bureau of Economic Research*.

parámetros y condiciones iniciales de los modelos log-VE para los filtros EKF y PF mediante ML, de la misma forma que se hizo en la Sección 6.2.

En la Figura 6.14 se presentan el índice y sus retornos, además de todas las estimaciones realizadas. En la Subfigura c) se puede ver la caída del índice en $t = 50$, mientras que d) muestra cómo los retornos tienen cada vez mayor varianza a partir de este instante. En cuanto a los resultados de los filtros, tanto PF como EKF identifican el cambio en la volatilidad de los retornos, pues ambas estimaciones de la volatilidad se asemejan a la señal suavizada, a pesar de que la estimación de EKF parece ser mucho más sesgada.

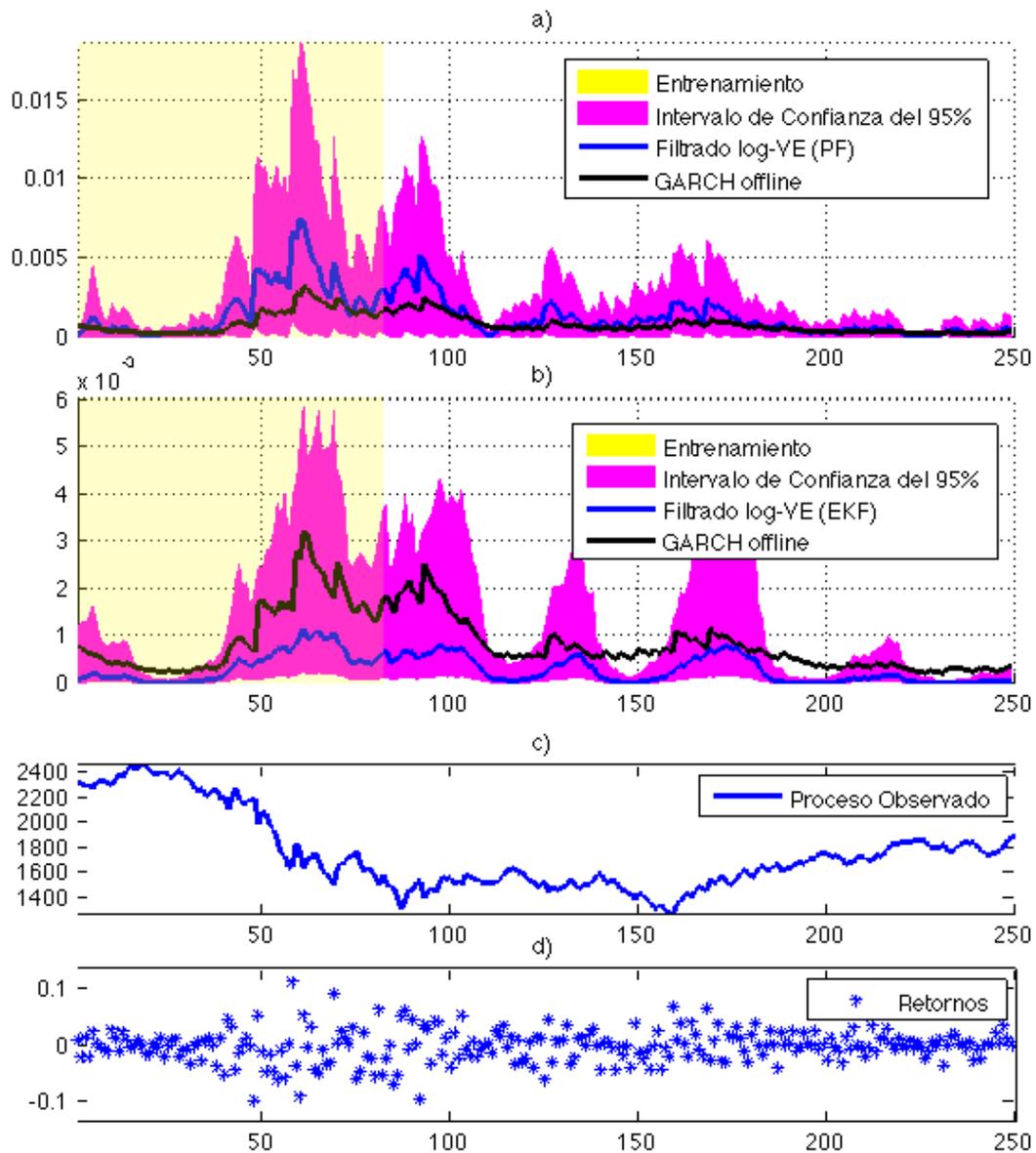


Figura 6.14: Resultados del filtrado de la volatilidad estocástica de los retornos del índice *NASDAQ Composite* $r_{1:250}$ durante el período Julio 21, 2008 - Julio 17, 2009 usando el modelo log-VE. **a)** Negro: Estimación suavizada de la volatilidad de índice por un modelo GARCH(1,1) y ML. Azul: Estimación de PF. Magenta: Estimación de PF del CI del 95%. **b)** Negro: Estimación suavizada de la volatilidad del índice por un modelo GARCH(1,1) y ML. Azul: Estimación de EKF. Magenta: Estimación de EKF del CI del 95%. **c)** Valores observados del índice. **d)** Retornos del índice usando composición continua.

En la Figura 6.15 se presentan las estimaciones de las fdp de la volatilidad al instante

$t = 76$, σ_{76}^2 (Noviembre 5, 2008) estimadas por ambos filtros, el conjunto de partículas, y la estimación suavizada para dicho instante. En esta figura se puede ver que a pesar de que EKF entrega una estimación más precisa, su densidad indica que es probable que la volatilidad tome valores que en realidad tienen verosimilitud nula (el soporte de la estimación de EKF de la fdp de σ_{76}^2 incluye valores negativos). Nótese que en la Figura 6.15 b), el conjunto de partículas fue truncado, ignorando el 0.02 % de las partículas.

Además, las estimaciones adaptativas de los parámetros –presentadas en la Figura 6.16– muestran (i) variaciones muy pequeñas para μ ($< 5\%$), (ii) consistencia por parte de ambos filtros para ϕ , y (iii) pequeñas variaciones para σ^n ($< 5\%$).

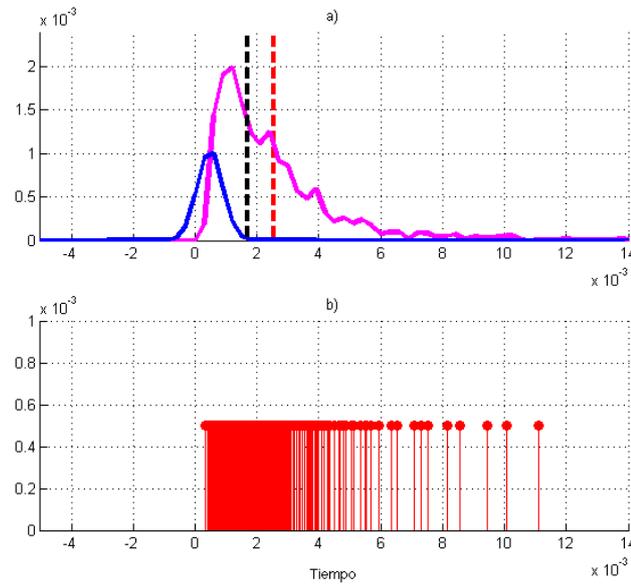


Figura 6.15: Densidades estimadas para la volatilidad del índice *NASDAQ Composite* en Noviembre 5, 2008 usando el modelo log-VE. **a)** Magenta: Versión continua de la estimación de la fdp de σ_{76}^2 estimada mediante PF, Azul: Estimación de la fdp de σ_{76}^2 por EKF, Negro: Estimación suavizada de σ_{76}^2 por GARCH(1,1) y ML, Rojo: Valor esperado de σ_{76}^2 estimado por PF. **b)** Partículas correspondientes a la estimación de σ_{76}^2 de PF.

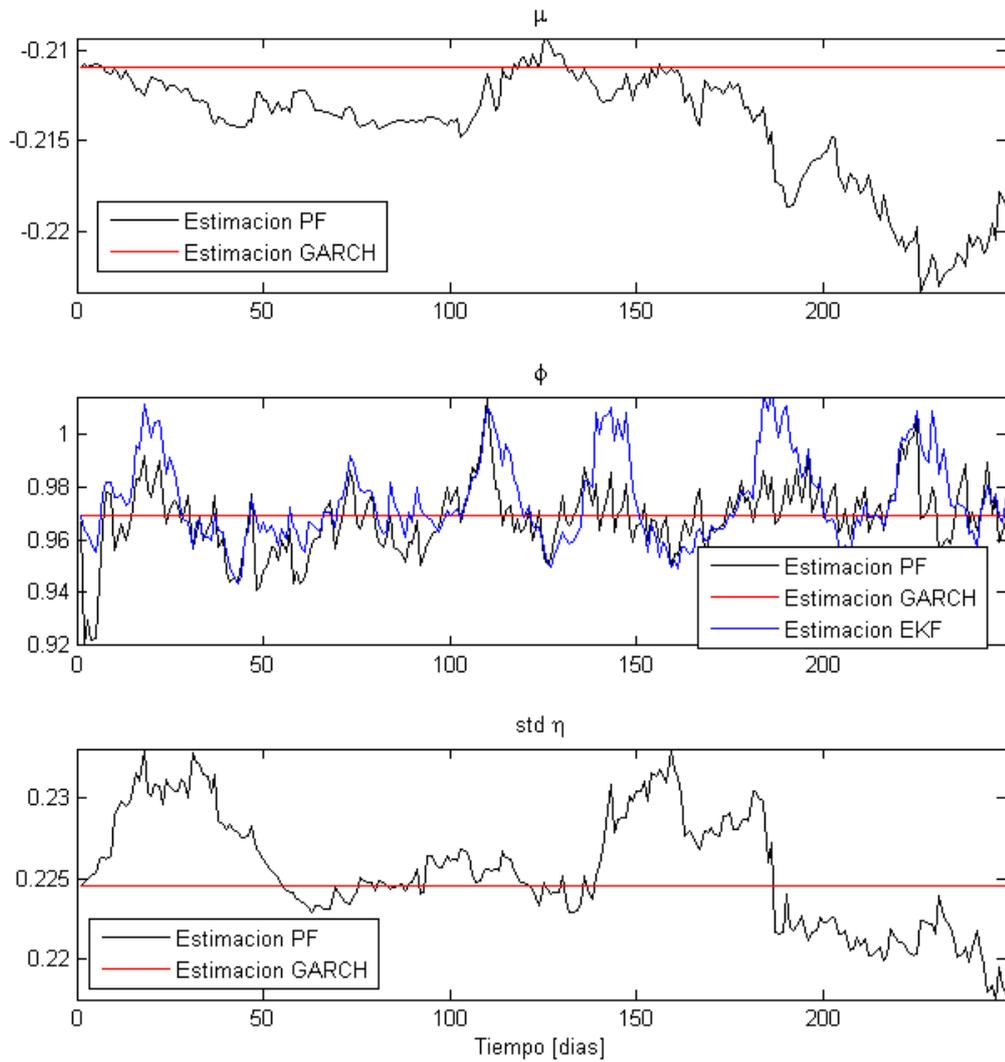


Figura 6.16: Estimaciones de los parámetros γ , ϕ y σ^η del modelo log-VE para los retornos del índice *NASDAQ Composite*. Rojo: Estimación de máxima verosimilitud para el modelo GARCH usando $r_{1:80}$ (offline). Azul: Estimación online EKF. Negro: Estimación online PF.

6.3.2.2. uGARCH

Para implementar el filtrado de la volatilidad del INC usando en modelo uGARCH, se utilizó el subconjunto r_{80} para identificar un modelo GARCH(1,1) mediante ML y estimar los parámetros y condiciones iniciales de un modelo uGARCH (se aprovecharon las similitudes

entre los modelos GARCH y uGARCH de la misma forma que se hizo en la Sección 6.2).

En la Figura 6.17 se presentan el índice y sus retornos, además de todas las estimaciones realizadas. En esta figura se puede ver la estructura basada en EKF no es capaz de seguir la estimación suavizada, en cambio PF sigue la estimación suavizada de forma muy exacta y sin caer es una estimación imprecisa del CI. Esta figura representa la habilidad del conjunto PF-uGARCH para seguir cambios abruptos en los procesos de volatilidad estocástica, tal como se vió en simulación en la Figura 6.4.

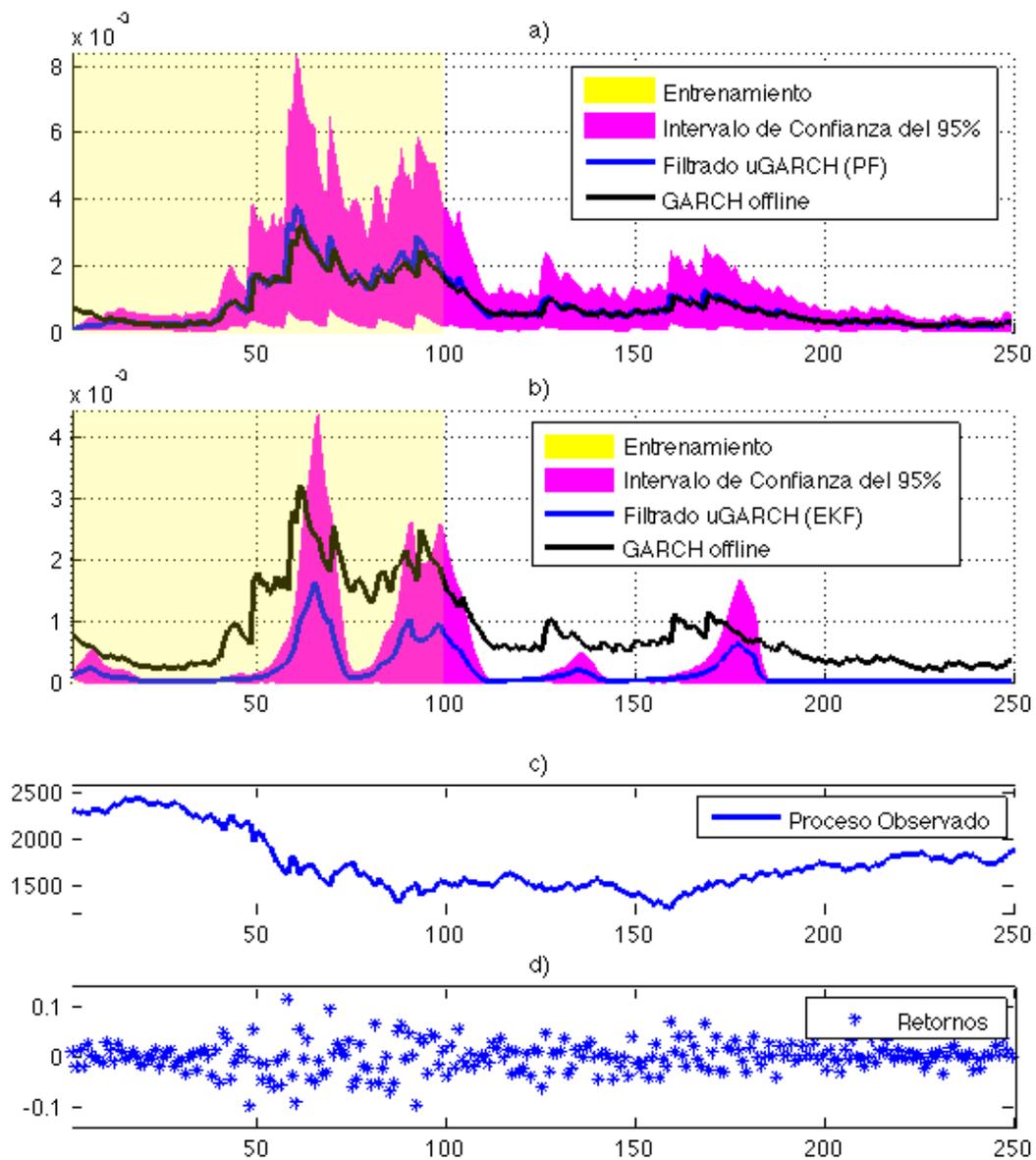


Figura 6.17: Resultados del filtrado de la volatilidad estocástica de los retornos del índice *NASDAQ Composite* $r_{1:250}$ durante el período Julio 21, 2008 - Julio 17, 2009 usando el modelo uGARCH. **a)** Negro: Estimación suavizada de la volatilidad de índice por un modelo GARCH(1,1) y ML. Azul: Estimación de PF. Magenta: Estimación de PF del Intervalo de confianza del 95%. **b)** Negro: Estimación suavizada de la volatilidad de índice por un modelo GARCH(1,1) y ML. Azul: Estimación de EKF. Magenta: Estimación de EKF del Intervalo de confianza del 95%. **c)** Valores observados del índice. **d)** Retornos del índice usando composición continua.

En la Figura 6.18 se presentan las densidades estimadas por ambos filtros para la vola-

tilidad al instante $t = 76$, σ_{76}^2 (Noviembre 5, 2008), además del conjunto de partículas y de la estimación suavizada para este instante. En esta Figura se puede ver que, igualmente que los resultados de la utilización del modelo log-VE mostrado en la Figura 6.15, la estimación de PF es más exacta que la de EKF, de hecho, el valor esperado de σ_{76}^2 estimado por PF es muy similar a la estimación suavizada de la volatilidad. Sin embargo, la verosimilitud de que la volatilidad tome valores negativos –estimada por EKF– es mucho menor que la reportada al usar el modelo log-VE.

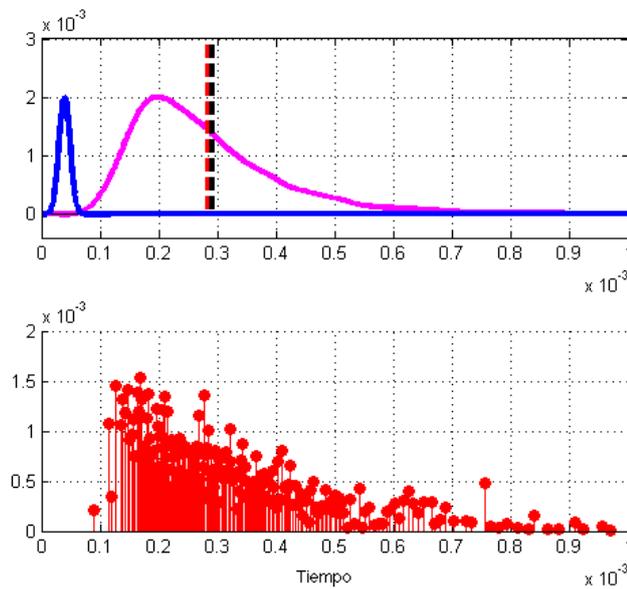


Figura 6.18: Densidades estimadas para la volatilidad del índice *NASDAQ Composite* en Noviembre 5, 2008 usando el modelo uGARCH. **a)** Magenta: Versión continua de la estimación de la fdp de σ_{76}^2 estimada mediante PF, Azul: Estimación de la fdp de σ_{76}^2 por EKF, Negro: Estimación suavizada de σ_{76}^2 por GARCH(1,1), Rojo: Valor esperado de σ_{76}^2 estimado por PF. **b)** Partículas correspondientes a la estimación de σ_{76}^2 de PF.

La Figura 6.19 presenta la evolución de la estimación de los parámetros del modelo uGARCH, esta figura muestra claramente el problema de observabilidad de los parámetros α y β generado al estimar los parámetros como se explicó en el Capítulo 5. Las estimaciones de ambos parámetros realizadas por EKF son iguales salvo una constante, pues en realidad el parámetro que se estimó fue $\alpha + \beta$. En cuanto a la estimación de los parámetros realizada

por PF, a pesar de presentar pequeñas variaciones, la caída de la estimación de α_t a partir de $t = 100$ ciclos explica la estabilización de INC después de su caída en Septiembre 30, 2008.

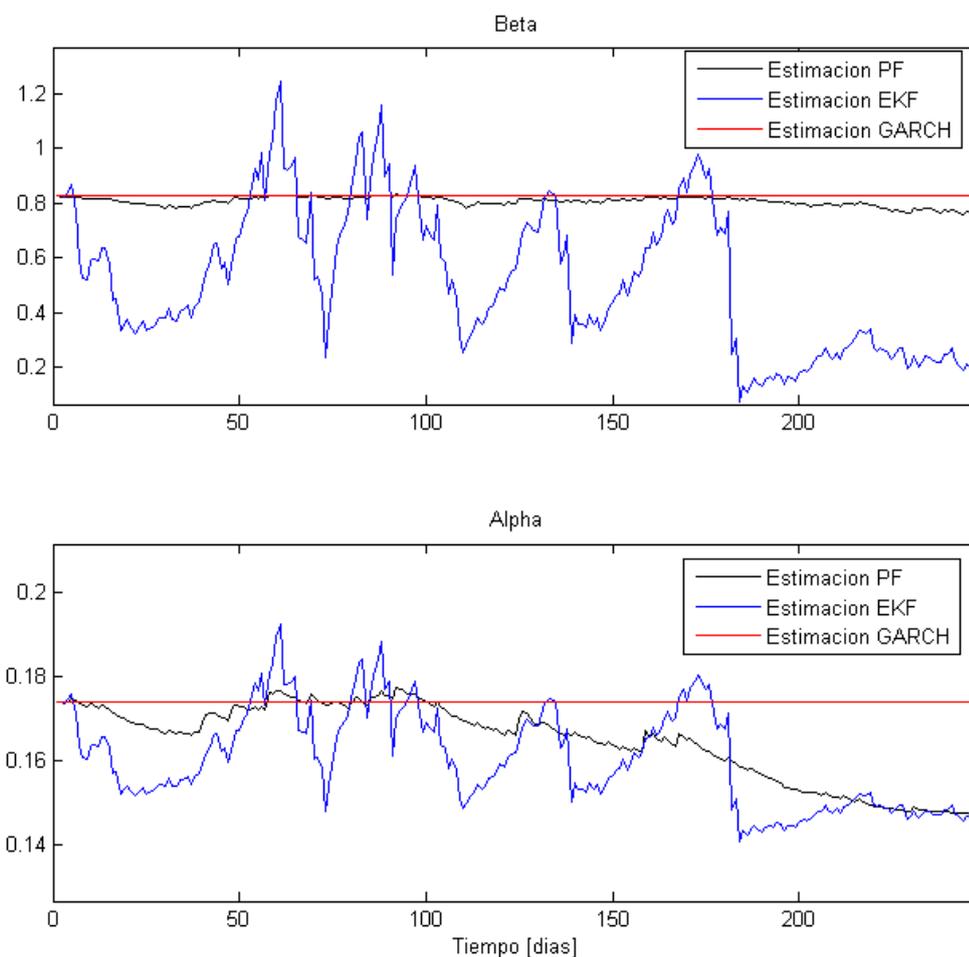


Figura 6.19: Estimaciones de los parámetros α y β del modelo uGARCH para los retornos del índice *NASDAQ Composite*. Rojo: Estimación de máxima verosimilitud de los parámetros del modelo GARCH usando $r_{1:80}$ (offline). Azul: Estimación online EKF. Negro: Estimación online PF.

Los resultados presentados en esta sección dejan de manifiesto las ventajas del conjunto uGARCH-PF en aplicaciones de filtrado de la volatilidad estocástica de instrumentos financieros, en el sentido de entregar estimaciones precisas y exactas. Estas aseveraciones se respaldarán en la Sección 6.3.4.

A continuación se presentan los resultados de las estructuras presentadas para el problema de la predicción de la volatilidad estocástica de INC.

6.3.3. Predicción

Considerando el criterio de predicción explicado en la Sección 6.2.4 se han implementado rutinas de predicción de la volatilidad del índice *NASDAQ Composite*. Para esto, se consideró el proceso de retornos $r_{1:100}$, equivalente al período Julio 21, 2008 - Diciembre 15, 2008, donde el subconjunto $r_{1:50}$ se utilizó para identificar los parámetros y las condiciones iniciales de los modelos a utilizar para predecir la volatilidad al instante $t = 100$, σ_{100}^2 considerando muestras hasta $\tau \in [50, 100]$. Los resultados de esta rutina de predicción para cada uno de los modelos log-VE y uGARCH son presentados a continuación tal como se hizo en simulación (Sección 6.2.4).

6.3.3.1. log-VE

Al igual que en la implementación de filtrado usando el modelo log-VE, el modelo para predecir la volatilidad de INC fue ajustado considerando un modelo GARCH(1,1), identificado mediante ML, y la transformación $V_{1:50} = 2 \log(\sigma_{1:50})$.

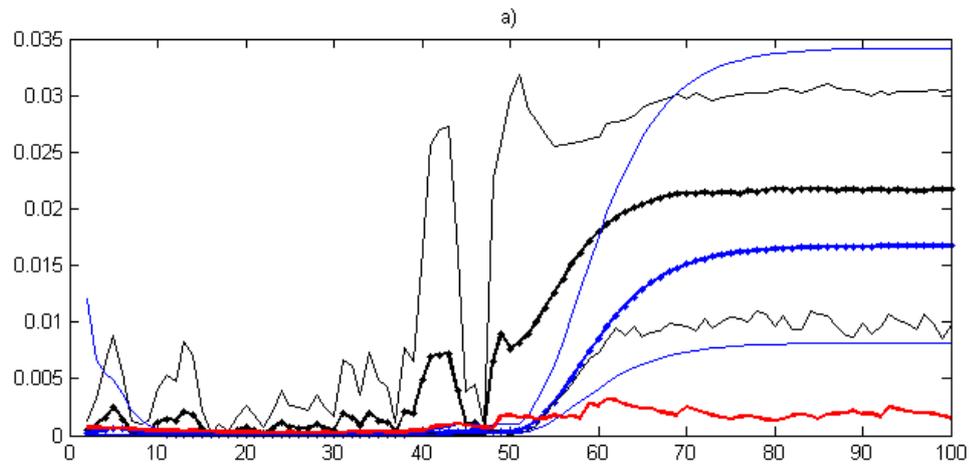


Figura 6.20: Predicción de la volatilidad estocástica del índice *NASDAQ Composite* usando el modelo log-VE para $t \in [50, 100]$. Rojo: Estimación suavizada de σ_t^2 con GARCH(1,1). Negro: Estimación de PF de media (grosso) y CI del 95 % (delgado). Azul: Estimación de EKF de media (grosso) y CI del 95 % (delgado). Para $t \leq 50$ la estimación corresponde a filtrado y para $t > 50$ la predicción considerando sólo la información hasta $t = 50$ (Abcisas = t).

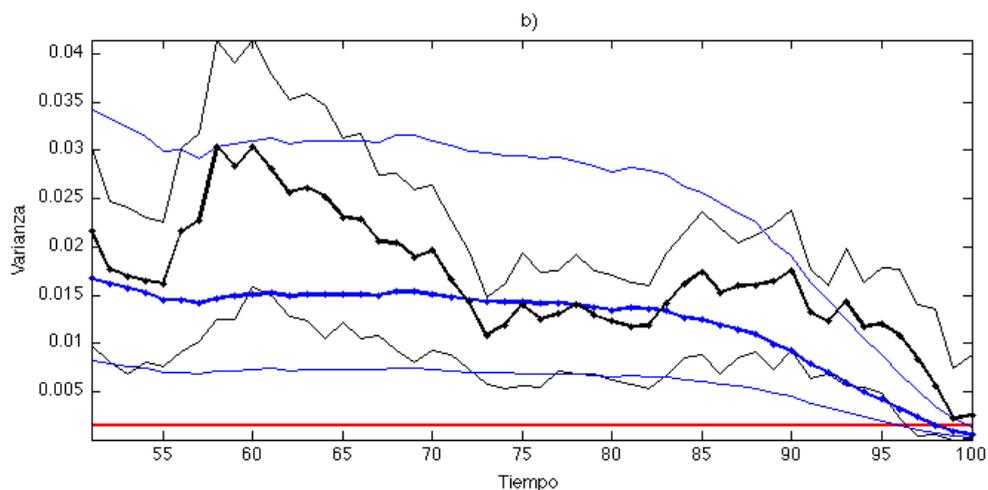


Figura 6.21: Predicción de la volatilidad estocástica del índice *NASDAQ Composite* usando el modelo log-VE para $t = 100$. Rojo: Estimación suavizada de la volatilidad estocástica en $t = 100$, σ_{100}^2 . Negro: Predicción del valor esperado (grosso) el CI del 95 % (delgado) de σ_{100}^2 usando muestras hasta $\tau \in [50, 100]$ usando PF. Azul: Predicción del valor esperado (grosso) el CI del 95 % (delgado) de σ_{100}^2 usando muestras hasta $\tau \in [50, 100]$ usando EKF. (Abcisas = τ)

En la Figura 6.20 se presentan los resultados de la predicción del proceso de volatilidad

de INC σ_t^2 y se puede ver cómo las predicciones tienden al punto de equilibrio a medida que el horizonte de predicción aumenta. De forma similar a los resultados de simulación, los puntos de equilibrio son distintos, pues los ambos enfoques usan distintos modelos dinámicos y criterios de predicción.

Con respecto de la mejora de las predicciones de σ_{100}^2 a medida que el horizonte de predicción disminuye, en la Figura 6.21 se ve que para ambas estructuras, EKF y PF, el valor esperado de las estimaciones se aproxima a σ_{100}^2 de forma similar a los resultados en simulación. De hecho, la dependencia en el tiempo del tamaño de los intervalos de confianza estimados por ambos filtros también muestra propiedades similares a los resultados simulados.

6.3.3.2. uGARCH

En el caso de la predicción de la volatilidad de INC usando el modelo uGARCH, también se ha ocupado la ventana de tiempo Julio 21, 2008 - Diciembre 15, 2008 (equivalente a $r_{1:100}$), y el subconjunto $r_{1:50}$ para ajustar los parámetros del modelo de estimación.

En la Figura 6.22 se puede ver que las predicciones igualmente tienden a su punto de equilibrio, el cual está muy alejado del valor real de la volatilidad, pues justamente en $t = 50$ hay un cambio considerable en el comportamiento de índice. Igual que en los resultados de simulación, en este caso los puntos de equilibrios son muy similares, esto es debido a que ambas estructuras de filtrado estiman adaptativamente los mismos parámetros del modelo

Además, de la Figura 6.23 se puede ver que si bien ambas estimaciones del valor esperado mejoran de forma monótona a medida que el horizonte de predicción disminuye, el EKF falla en asignar verosimilitud no nula a valores negativos para σ_{100}^2 . Es importante también notar que el CI del 95 % estimado por PF incluye a la estimación suavizada de σ_{100}^2 para horizontes de predicción menores a 6, mientras que el CI estimado por EKF lo hace para horizontes menores a 4.

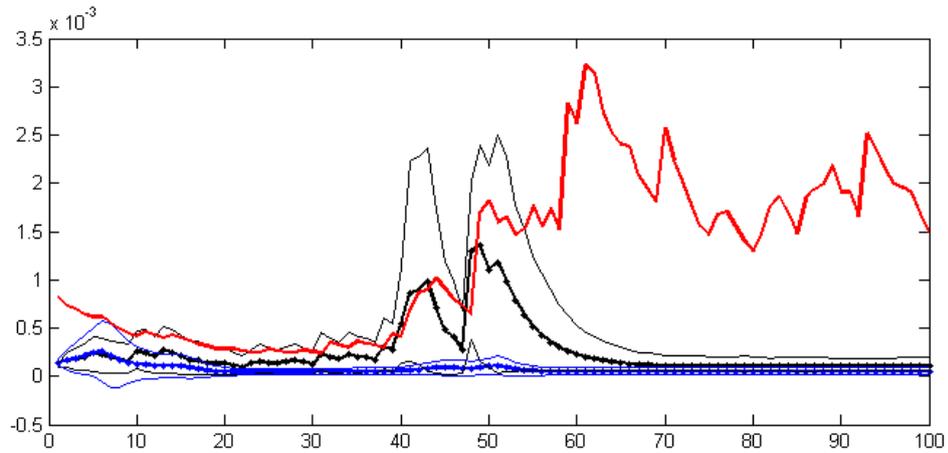


Figura 6.22: Predicción de la volatilidad estocástica del índice *NASDAQ Composite* usando el modelo uGARCH para $t \in [150, 200]$. Rojo: Estimación suavizada de σ_t^2 con GARCH(1,1). Negro: Estimación de PF de media (grosso) y CI del 95 % (delgado). Azul: Estimación de EKF de media (grosso) y CI del 95 % (delgado). Para $t \leq 150$ la estimación corresponde a filtrado y para $t > 150$ a predicción considerando sólo la información hasta $t = 150$ (Abciscas = t).

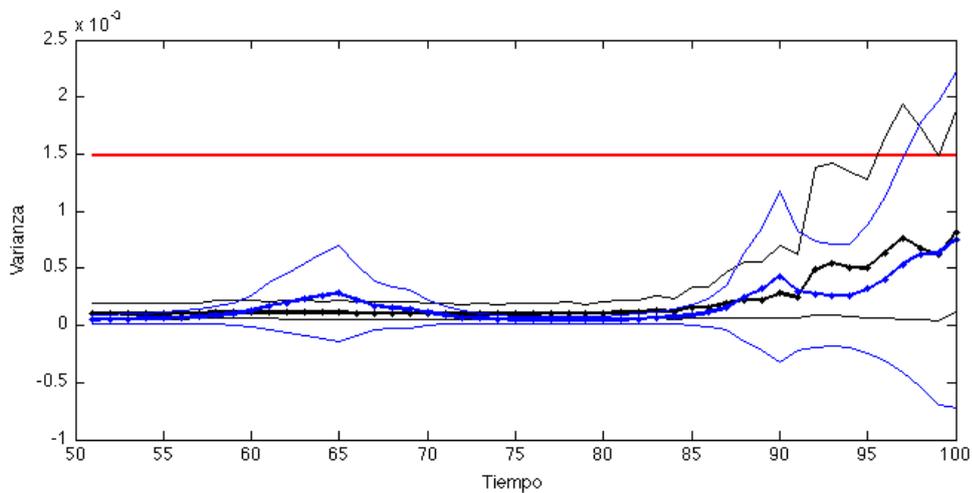


Figura 6.23: Predicción de la volatilidad estocástica del índice *NASDAQ Composite* usando el modelo uGARCH para $t = 200$. Rojo: Estimación suavizada de la volatilidad estocástica en $t = 200$, σ_{200}^2 . Negro: Predicción del valor esperado (grosso) y del CI del 95 % (delgado) de σ_{200}^2 usando muestras hasta $\tau \in [150, 200]$ usando PF. Azul: Predicción del valor esperado (grosso) y del CI del 95 % (delgado) de σ_{200}^2 usando muestras hasta $\tau \in [150, 200]$ usando EKF. (Abciscas = τ)

6.3.4. Evaluación

Tal como se hizo con los resultados de simulación, a continuación se presenta una evaluación de las estructuras propuestas para filtrar y predecir la volatilidad de INC usando los indicadores de desempeño presentados en la Sección 6.1.

En la Tabla 6.3 se presentan dichos índices. Es posible notar que los resultados de estimación de estado que consideran filtro de partículas, son mucho más exactos que sus correspondiente de EKF, de hecho, para el caso de modelo uGARCH, el índice de exactitud es un orden de magnitud inferior al de log-VE. Por otro lado, y al igual que los resultados de simulación, el EKF es la estructura que reporta estimaciones más precisas, aunque no muy alejadas de las presentadas por PF.

En cuanto a los resultados de predicción, la superioridad de uGARCH frente a log-VE es considerable. Esta diferencia es atribuible a que el subconjunto de muestras considerado es suficiente para ajustar un modelo uGARCH –en una rutina de predicción– pero no para ajustar un modelo log-VE.

Tabla 6.3: Índices de desempeño para resultados de filtrado y predicción del índice *NASDAQ Composite* usando los modelos log-VE y uGARCH, y los filtros PF y EKF.

Modelo	log-VE		uGARCH	
	PF	EKF	PF	EKF
I_{95}^{AC}	1.34	3.52	0.29	4.06
I_{95}^{PR}	8.695	7.92	4.21	2.53
I_{95}^{FR}	34.96	32.71	3.39	4.02

Cabe destacar que el conjunto uGARCH-PF es el que presenta las ventajas más evidentes para filtrar y predecir la volatilidad estocástica tanto en ambientes simulados como en el caso real visto en esta sección. Estos resultados son consistentes con la teoría presentada, pues se ha mostrado la habilidad de los modelos no Gaussianos para modelar series de tiempos de alta incertidumbre, y la capacidad de filtro de partículas para estimar el estado, e identificar adaptativamente los parámetros de modelos no lineales y no Gaussianos.

Capítulo 7

Conclusiones

En este trabajo se presenta, y evalúa, una estructura que considera estado latente, y parámetros variantes en el tiempo, para modelar la volatilidad de una serie de retornos financieros basada en la estructura del modelo determinístico GARCH(1,1). Esta estructura es comparada con el modelo logarítmico de volatilidad estocástica en los problemas de estimación y predicción de volatilidad, en ambientes simulados y con datos reales, utilizando dos estructuras de filtrado: filtro de partículas y filtro extendido de Kalman.

Mediante la introducción de índices de desempeño, fue posible cuantificar tanto la exactitud y precisión de las estructuras propuestas en el problema de estimación de estado, como también la consistencia de la predicción de intervalos de confianza. Basado en dichos índices, es posible afirmar que los resultados más exactos del problema de estimación –tanto en ambientes simulados como en la consideración de datos reales– son los correspondientes a la implementación del modelo uGARCH en conjunto con filtro de partículas. Igualmente, dicho conjunto presenta respuestas considerablemente precisas, en comparación a las otras estructuras consideradas. Es importante notar también que la implementación del modelo uGARCH con EKF también entrega estimaciones precisas, pero altamente sesgadas.

En cuanto a los resultados del problema de predicción, es posible afirmar que el conjunto uGARCH-PF –tanto en simulación como con datos reales– es el que presenta los resultados más consistentes en el sentido de entregar predicciones exactas del intervalo de confianza, que incluyen el valor real del proceso a predecir (o a una estimación suavizada de éste en el

caso de la implementación con datos reales).

Finalmente, se cree que considerando mejores estimaciones para los parámetros, y las condiciones iniciales de las cantidades variantes en el tiempo, los resultados del conjunto uGARCH-PF podrían mejorar. En este sentido se propone identificar estas cantidades mediante Quasi-Máxima Verosimilitud, o *Expectation Maximisation*, y analizar los resultados que dichas estimaciones producen para verificar si la incertidumbre relacionada a las estructuras de estimación se acerca a los niveles de incertidumbre inherentes al proceso, es decir, alcanzar índices de precisión en torno a 1.

Referencias

- Abramowitz, M. & Stegun, N. C. (1970). *Handbook of Mathematical Functions*, New York: Dover Publications. Inc.
- Alspach, D. L. & Sorenson, H. W. (1972). Nonlinear bayesian estimation using gaussian sum approximation, *IEEE Transactions on Automatic Control* **17**(4): 439–448.
- Anderson, B. D. & Moore, J. B. (1979). *Optimal Filtering*, Dover Publications.
- Andrieu, C., Doucet, A. & Punkskaya, E. (2001). *Sequential Monte Carlo Methods in Practice*, Springer, chapter Sequential Monte Carlo Methods for Optimal Filtering.
- Bachelier, L. (1900). *Théorie de la Spéculation*, PhD thesis, La Sorbonne.
- Balakrishnan, N. & Chen, W. (1999). *Handbook of Tables for Order Statistics from Lognormal Distributions with Applications*, Springer.
- Bar-Shalom, Y., Li, X. R. & Kirubarajan, T. (2001). *Estimation with Application to Tracking and Navigation*, New York: John Wiley & Sons.
- Berzuini, C., Best, N. G., Gilks, W. R. & Larizza, C. (1997). Dynamic conditional independence models and markov chain monte carlo methods, *Journal of the American Statistical Association* **92**: 1403–1412.
- Black, F. & Scholes, M. (1973). The pricing of options and corporate liabilities, *Journal of Political Economy* **81**(3): 637.

- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* **31**: 307–327.
- Carnero, M. A., Pena, D. & Ruiz, E. (2004). Persistence and kurtosis in garch and stochastic volatility models, *JOURNAL OF FINANCIAL ECONOMETRICS* **2**(2): 319–342.
- Christie, A. (1982). The stochastic behavior of common stock variances, *Journal of Financial Economics* **10**: 407–432.
- Chui, C. K. & Chen, G. (1991). *Kalman Filtering with Real-Time Applications*, 2nd ed edn, New York: Springer-Verlag.
- Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices, *Econometrica* **41**(1): 135–155.
- Del Moral, P. (2004). *Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications*, Probability and its Applications, Springer, New York.
- Doucet, A., de Freitas, N. & Gordon, N. (2001). *Sequential Monte Carlos Methods in Practice*, Statistics for engineering and information science, New York: Springer-Verlag.
- Doucet, A., Godsill, S. & Andrieu, C. (2000). On sequential monte carlo sampling methods for bayesian filtering, *Statistics and Computing* **10**: 197–208.
- Doucet, A. & Johansen, A. (2008). A tutorial on particle filtering and smoothing: fifteen years later, *Technical report*, Department of Statistics, University of British Columbia.
- Engle, R. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of the united kingdom inflation, *Econometrica* **50**: 987–1007.
- Franses, P. H., van der Leij, M. & Paap, R. (2008). A simple test for garch against a stochastic volatility model, *Journal of Financial Econometrics* **6**(3): 291–306.
- Gallop, R. J. (2007). Nonlinear programming for maximum likelihood estimation, NESUG, Baltimore, Maryland.

- Geweke, J. F. (1989). Bayesian inference in econometric models using monte carlo integration, *Econometrica* **57**: 1317–1340.
- Ghysels, E., Harvey, A. & Renault, E. (1996). *Statistical Models in Finance*, North-Holland, Amsterdam, chapter Stochastic volatility, pp. 119–191.
- Gordon, N. J., Salmond, D. J. & Smith, A. F. M. (2002). Novel approach to nonlinear/non-gaussian bayesian state estimation, *Radar and Signal Processing, IEE Proceedings F* **140**(2): 107–113.
- Grimmett, G. & Welsh, D. (1986). *Probability: An Introduction*, Oxford Science Publications.
- Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators, *Econometrica* **50**(4): 1029–54.
URL: <http://ideas.repec.org/a/ecm/emetrp/v50y1982i4p1029-54.html>
- Harvey, A. C., Ruiz, E. & Shephard, N. (1994). Multivariate stochastic variance models, *Review of Economic Studies* **61**(2): 247–264.
- Haug, A. (2005). A tutorial on bayesian estimation and tracking techniques applicable to nonlinear and non-gaussian processes, *Technical report*, MITRE.
- Hull, J. C. (2008). *Options, Futures, and Other Derivatives*, 7th edn, Prentice Hall.
- Joanes, D. N. & Gill, C. A. (1998). Comparing measures of sample skewness and kurtosis, *Journal of the Royal Statistical Society: Series D (The Statistician)* **47**(1): 183 – 189.
- Johannes, M. S., Polson, N. G. & Stroud, J. R. (2009). Optimal filtering of jump-diffusions: Extracting latent states from asset prices, *Review of Financial Studies* **22**(7): 2559–2599.
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems, *Journal of Basic Engineering* **82**(1): 35–45.
- Kalman, R. E. & Bucy, R. S. (1961). New results in linear filtering and prediction theory, *Transactions of the ASME. Series D, Journal of Basic Engineering* **83**: 95–107.

- Karatzas, I. & Shreve, S. (1988). *Brownian Motion and Stochastic Calculus*, Vol. 113 of *Graduate Texts in Mathematics*, Springer-Verlag, New York, Heidelberg & Berlin.
- Karatzas, I. & Shreve, S. (1998). *Methods of Mathematical Finance*, Vol. 39 of *Applications of Mathematics*, Springer-Verlag, New York, Heidelberg, and Berlin.
- Kim, T.-H. & White, H. (2004). On more robust estimation of skewness and kurtosis, *Finance Research Letters* **1**(1): 56 – 73.
- Kitagawa, G. (1998). Self-organizing state space model, *Journal of the American Statistical Association* **93**(443): 1203–1215.
- Kitagawa, G. & Sato, S. (2001). *Sequential Monte Carlo Methods in Practice*, Springer, chapter Monte Carlo Smoothing and Self-Organising State-Space Model.
- Kong, A., Liu, J. S. & Wong, W. H. (1994). Sequential imputations and bayesian missing data problems, *Journal of the American Statistical Association* **89**(425): 278–288.
URL: <http://dx.doi.org/10.2307/2291224>
- Krishnan, V. (2005). *Nonlinear Filtering and Smoothing: An Introduction to Martingales, Stochastic Integrals and Estimation*, Dover.
- Lehar, A., Scheicher, M. & Schittenkopf, C. (2002). Garch vs. stochastic volatility: Option pricing and risk management, *Journal of Banking & Finance* **26**(2-3): 323–345.
URL: <http://ideas.repec.org/a/eee/jbfin/v26y2002i2-3p323-345.html>
- Li, Q. & Racine, J. S. (2007). *Nonparametric Econometrics: Theory and Practice*, Princeton University Press.
- Liu, J. S. (1996). Metropolized independent sampling with comparisons to rejection sampling and importance sampling, *Statistics and Computing* **6**(2): 113–119.
URL: <http://dx.doi.org/10.1007/BF00162521>

- Liu, J. & West, M. (2001). *Sequential Monte Carlo Methods in Practice*, Springer, chapter Combined Parameter and State Estimation in Simulation-Based Filtering.
- Mandelbrot, B. (1963). The variation of certain speculative prices, *Journal of Business* **36**: 394–419.
- Mandelbrot, B. B. & Hudson, R. L. (2004). *The (mis)behaviour of markets : a fractal view of risk, ruin and reward*, Basic Books.
- Merton, R. C. (1973). Theory of rational option pricing, *The Bell Journal of Economics and Management Science* **4**(1): 141–183.
URL: <http://www.jstor.org/stable/3003143>
- Oksendal, B. (2007). *Stochastic Differential Equations: An Introduction with Applications (Universitext)*, 6th edn, Springer.
- Oksendal, B. & Sulem, A. (2004). *Applied Stochastic Control of Jump Diffusions*, Springer.
- Orchard, M. E. (2009). *On-Line Fault Diagnosis and Failure Prognosis Using Particle Filters: Theoretical Framework and Case Studies*, VDM Verlag.
- Orchard, M. E., Tobar, F. A. & Vachtsevanos, G. (2009). Outer feedback correction loops in particle filtering-based prognostic algorithms: Statistical performance comparison, *Studies in Informatics and Control* **18**(4).
- Papoulis, A. (1984). *Probability, Random Variables, and Stochastic Processes*, New York: McGraw-Hill.
- Rachev, S. T., Hsu, J. S. J., Bagasheva, B. S. & Fabozzi, F. J. (2008). *Bayesian Methods in Finance*, Frank J. Fabozzi Series, 1 edn, Wiley.
- Raggi, D. & Bordignon, S. (n.d.). Sequential monte carlo methods for stochastic volatility models with jumps. Preprint.
URL: <http://homes.stat.unipd.it/raggi/?page=Publications&lang=EN>

- Rao, C. (1965). *Linear Statistical Inference and its Applications*, Wiley.
- Ristic, B., Arumlampalam, S. & Gordon, N. (2004). *Beyond the Kalman filter: particle filters for tracking applications.*, Artech House.
- Rubin, D. (1988). *Bayesian Statistics*, Vol. 3, Oxford University Press, chapter Using the SIR algorithm to simulate posterior distributions.
- Ruiz, E. (1994). Quasi-maximum likelihood estimation of stochastic volatility models, *Journal of Econometrics* **63**(1): 289 – 306.
- Sandmann, G. & Koopman, S. J. (1996). Maximum likelihood estimation of stochastic volatility models, *FMG Discussion Papers dp248*, Financial Markets Group.
URL: <http://ideas.repec.org/p/fmg/fmgdps/dp248.html>
- Schmalensee, R. & Trippi, R. R. (1978). Common stock volatility expectations implied by option premia, *The Journal of Finance* **33**(1): 129–147.
URL: <http://www.jstor.org/stable/2326355>
- Schön, T. B., Wills, A. & Ninness, B. (2006). Maximum likelihood nonlinear system estimation, Vol. 14th, IFAC Symposium on System Identification, Newcastle, Australia.
- Shephard, N. (1996). *Time Series Models in Econometrics, Finance and Other Fields*, Chapman & Hall, London, chapter Statistical aspects of ARCH and stochastic volatility, pp. 1–67.
- Simon, D. (2006). *Optimal State Estimation. Kalman, H-Infinity, and Nonlinear Approaches*, Wiley-Interscience.
- Stratonovich., R. L. (1959). On the theory of optimal non-linear filtration of random functions, *Teor. Veroyatnost. i Primenen* **4**: 223–225.
- Tabachnick, B. G. & Fidell, L. S. (1996). *Using multivariate statistics*, 3rd ed. edn, New York: Harper Collins.

- Tauchen, G. E. & Pitts, M. (1983). The price variability-volume relationship on speculative markets, *Econometrica* **51**: 485–505.
- Taylor, S. (1986). *Modelling Financial Time Series*, 1 edn, John Wiley & Sons.
- Tobar, F. A. & Orchard, M. E. (2009). Efectos de lazos externos de corrección en algoritmos de pronóstico de fallas basados en filtro de partículas, *Anales del Instituto de Ingenieros de Chile*.
- West, M., Harrison, P. J. & Migon, H. S. (1985). Dynamic generalized linear models and bayesian forecasting (with discussion), *Journal of the American Statistical Association* **80**(389): 73–83.
- URL:** <http://www.jstor.org/stable/2288042>

Apéndices

A . Códigos

```
1 T=500; // Cantidad de datos a generar
2  $\tau=150$ ; // Cantidad de datos a considerar para ajustar modelo GARCH
3  $\sigma_0 = 0,002, C = 9 \cdot 10^{-4}, K = 10^{-6}, \beta = 0,6$ , y  $\alpha = 0,2$ ;
4  $r_0 \sim \mathcal{N}(0, \sigma_0)$ ;
5 for  $i = 1 : \text{floor}(T/2)$  do
6      $\sigma_i^2 = K + \sigma_{i-1}^2 * \beta + \alpha * (r_{i-1} - C)^2$ ;
7      $r_i \sim \mathcal{N}(0, \sigma_i)$ ;
8 end
9  $\beta = 0,85, \alpha = 0,1$ ; // Cambiar valor de  $\alpha$  y  $\beta$ 
10 for  $i = \text{floor}(T/2) + 1 : T$  do
11      $\sigma_i^2 = K + \sigma_{i-1}^2 * \beta + \alpha * (r_{i-1} - C)^2$ ;
12      $r_i \sim \mathcal{N}(0, \sigma_i)$ ;
13 end
14  $[\bar{\sigma}_t, \bar{C}, \bar{K}, \bar{\beta}, \bar{\alpha}] = \text{garchfit}(r_{1:T})$ ; // GARCH Toolbox
```

Algoritmo 1: Código de la generación de datos y ajuste de modelo correspondiente a la Sección 6.2 - modelo uGARCH

```

1 T=500; // Cantidad de datos a generar
2  $\tau=150$ ; // Cantidad de datos a considerar para ajustar modelo GARCH
3  $V_0 = -9,5, \mu = -1, \phi = 0,85$ , y  $seta = 0,1$ ;
4  $r_0 \sim \mathcal{N}(0, \exp(V_0/2))$ ;
5 for  $i = 1 : \text{floor}(T/2)$  do
6    $V_i = \mu + \phi V_{i-1} + \eta_i, (\eta_i \sim \mathcal{N}(0, seta))$ ;
7    $r_i \sim \mathcal{N}(0, \sigma_i)$ ;
8 end
9  $\phi = 0,8, seta = 0,2$ ; // Cambiar valor de  $\phi$  y  $seta$ 
10 for  $i = \text{floor}(T/2) + 1 : T$  do
11    $V_i = \mu + \phi V_{i-1} + \eta_i, (\eta_i \sim \mathcal{N}(0, seta))$ ;
12    $r_i \sim \mathcal{N}(0, \exp(V_i/2))$ ;
13 end
14  $[\bar{\sigma}_t, \bar{C}, \bar{K}, \bar{\beta}, \bar{\alpha}] = \text{garchfit}(r_{1:T})$ ; // GARCH Toolbox
15  $[\bar{V}_t, \bar{\mu}, \bar{\phi}, \bar{seta}] = \text{LeastSquares}([r_{1:T}, 2 \log(\bar{\sigma}_t)])$ 

```

Algoritmo 2: Código de la generación de datos y ajuste de modelos correspondiente a la Sección 6.2 - modelo log-VE

```

1 N=2000; // Número de partículas
2  $R_{th} = 0,6$ ; // Umbral de remuestreo
3 for  $j=1:T$  do // Horizonte de estimación
4   for  $i=1:N$  do // Para todas las partículas
5      $\alpha_{j+1}^{(i)} \sim \mathcal{N}(\alpha_j^{(i)}, 0, 1\alpha_0)$ ;
6      $\beta_{j+1}^{(i)} \sim \mathcal{N}(\beta_j^{(i)}, 0, 1\beta_0)$ ;
7      $\sigma_{j+1}^{2(i)} = K + \beta_j \sigma_j^{2(i)} + \alpha_j \sigma_j^{2(i)} \eta_j^2, (\eta_j \sim \mathcal{N}(0, 1))$ ;
8      $W_{j+1}^{(i)} = W_j^{(i)} p(r_{j+1} | \sigma_{j+1}^{2(i)})$ ;
9   end
10   $W_{j+1}^{(1:N)} = \frac{W_{j+1}^{(1:N)}}{\sum_{i=1}^N W_{j+1}^{(i)}}$ ; // Normalización de pesos
11  if  $\frac{1}{\sum_{i=1}^N W_{j+1}^{2(i)}} < R_{th}$  then // Remuestreo
12    for  $i=1:N$  do // Para todas las partículas
13      Elegir  $[\hat{\sigma}_{j+1}^{2(i)}, \hat{\alpha}_{j+1}^{(i)}, \hat{\beta}_{j+1}^{(i)}] = [\sigma_{j+1}^{2(i)}, \alpha_{j+1}^{(i)}, \beta_{j+1}^{(i)}]$  con probabilidad  $W_{j+1}^{(i)}$ 
14    end
15     $[\sigma_{j+1}^{2(1:N)}, \alpha_{j+1}^{(1:N)}, \beta_{j+1}^{(1:N)}] = [\hat{\sigma}_{j+1}^{2(1:N)}, \hat{\alpha}_{j+1}^{(1:N)}, \hat{\beta}_{j+1}^{(1:N)}]$ 
16  end
17 end

```

Algoritmo 3: Código de las rutinas de filtrado de las Secciones 6.2 y 6.3. Para la rutina de predicción sólo se debe reemplazar la línea 8 por $W_{j+1}^{(i)} = W_j^{(i)}$