

UNIVERSIDAD DE CONCEPCIÓN
DIRECCIÓN DE POSTGRADO
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
PROGRAMA MAGÍSTER EN ESTADÍSTICA



**ESTIMACIÓN ADAPTATIVA DE LA VOLATILIDAD EN UNA SERIE DE
TIEMPO FINANCIERA**

Tesis presentada a la Facultad de Ciencias Físicas y Matemáticas de la
Universidad de Concepción para optar al grado académico de Magíster en
Estadística.

POR: MARIANA CECILIA LAVÍN ROSAS.
Profesor Guía: Dra. Nora Serdyukova.

Enero 2022

CONCEPCIÓN-CHILE

©2022, Estimación Adaptativa de la Volatilidad en una Serie de Tiempo Financiera.
Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o
procedimiento, incluyendo la cita bibliográfica del documento.





UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS

ESTIMACIÓN ADAPTATIVA DE LA VOLATILIDAD EN UNA SERIE DE TIEMPO FINANCIERA.

Tesis presentada a la Facultad de Ciencias Físicas y Matemáticas de la Universidad de
Concepción para optar al grado académico de Magíster en Estadística.



Profesor Guía:

Dra. Nora Serdyukova

Firma:.....

Comisión:

Mag. Erwin Kracht Rojas

Firma:.....

Mag. Sebastian Niklitschek Soto

Firma:.....

Memorante:

Mariana Lavín Rosas

Firma:.....

DEDICATORIA



Dedicado a mi Hija por ser la luz de mi vida y mi inspiración.

AGRADECIMIENTOS

A Dios.

Por estar a mi lado, acompañando cada uno de mis pasos, por darme fortaleza y calma en los momentos de angustia.

A mi Hija.

Querida Josefa, tú me acompañaste en este camino casi desde el primer día. Gracias por llegar a mi vida, enseñarme lo que es el amor incondicional, la paciencia y la ternura, en tus ojos veo todo lo que quiero en esta vida. Espero ser un referente para ti y que veas que las mujeres podemos hacer todo lo que nos proponemos.



A mis Profesores.

Agradezco a cada uno de los profesores que contribuyeron en mi formación. Al profesor Arturo Mora, por su trato amable y su buena disposición. A la profesora Luisa Rivas, por lo dinámica de sus clases y por su exigencia en los informes. A la profesora María Paz Casanova, por su comprensión, al profesor Sebastián Niklitschek, por introducirme al estudio de las Series de Tiempo y Finalmente y no por eso menos importante, a la profesora Nora Serdyukova, por todas las asignaturas impartidas, por ser mi profesora guía, por exigirme para que pudiese realizar un buen trabajo y tenerme paciencia, por conducirme en este nuevo mundo.

A mi Pareja.

Querido Daniel, te agradezco por apoyarme en este proceso, por comprender que no tuviese tiempo, por respetar mis opciones y cuidar a nuestra pequeña para que yo me pudiese concentrar. Gracias por darme una palabra de apoyo y aliento cuando lo necesité.



RESUMEN

El presente trabajo de tesis de Magíster en Estadística de la Universidad de Concepción tiene como objetivo proponer una nueva forma de modelar y predecir la volatilidad multivariada. Está sustentado en la importancia para el mundo de las finanzas, debido a la fuerte relación entre volatilidad, rendimiento y riesgo de una inversión.

Los métodos paramétricos clásicos aunque muy utilizados en muchas ocasiones no son adecuados, ya que en su formulación contienen restricciones que a menudo son violadas por los datos empíricos. El modelo propuesto, utiliza como metodología el procedimiento adaptativo local de punto de cambio, donde el estimador se va adaptando a los datos, mediante el uso de la Divergencia de Kullback-Leibler. Se definen todos los conceptos necesarios para validar el uso del Estimador, Propiedades del Estimador, Condiciones de Propagación, Calidad de la Estimación mediante Desigualdades de Oráculo, Estabilidad, etc.

Esta investigación no sólo es teórica, también se aplica a Datos Reales de Tipo de Cambio de Divisas y se compara con técnicas clásicas, los modelos de Correlación Dinámica Condicional, donde se demostró que el modelo adaptativo fue muy superior, sobretodo en dos aspectos, tiempo invertido en la predicción y calidad de la estimación, donde resalta el horizonte de tiempo de 1 día con el 97,54% de los casos favorables para el estimador adaptativo y a 3 días con el 53,95% de los casos.

ABSTRACT

This thesis work of Master in Statistics the University of Concepcion aims to propose a novel method procedure for change point detection estimator adaptative to the data nonparametric statistics. It is based on the importance for the world of finance, due to the strong relationship between volatility, return and risk of an investment.

The classical parametric methods, although widely used, are often inadequate, since their formulation contains restrictions that are often violated by the empirical data. The proposed model uses as a methodology the local adaptive procedure of change point, where the estimator is adapted to the data, through the use of the Kullback-Leibler divergence. Properties of the estimator, propagation conditions, estimation quality by means of estimation inequalities, stability, etc., are defined.

This research is not only theoretical, it is also applied to Real Foreign Exchange Rate Data and is compared with classical techniques, the Conditional Dynamic Correlation models, where it was demonstrated that the adaptive model was far superior, especially in two aspects, time invested in the prediction and quality of the estimation, where the time horizon of 1 day stands out with 97,54 % of the favorable cases for the adaptive estimator and 3 days with 53,95 % of the cases.

keywords Volatility ; Multivariate Adaptative Model; Non-Parametric Statistics; Financial Time Series.



Índice general

Agradecimientos	II
Resumen	IV
Abstract	v
Índice de Tablas	IX
Índice de Figuras	XI
1. Introducción	1
2. Marco Teórico	4
2.1. Conceptos Generales	4
2.2. Definición de Volatilidad	8
2.3. Estimadores de la Volatilidad	9
2.3.1. Estimadores Puntuales	10
2.3.2. Estimadores Paramétricos	11
2.3.3. Estimadores No Paramétricos	17
2.3.4. Técnicas de Filtrado: Suavización Kernel	17



2.4.	Divergencia de Kullback-Leibler	26
2.5.	Razón de Verosimilitud	27
2.6.	Descripción de la problemática	29
2.6.1.	Caso Particular	33
3.	Objetivos y Actividades	38
3.1.	Objetivo General	38
3.2.	Objetivos Específicos	38
4.	Materiales y Métodos	40
4.1.	Técnicas de Filtrado Adaptativo	40
4.1.1.	Modelamiento de Volatilidad mediante Retornos Logaritmicos. Trans- formación de Potencia	40
4.1.2.	Estimador LAVE	42
4.1.3.	Estimación Adaptativa con Homogeneidad Local	43
4.1.4.	Estimador LCP	45
4.2.	Ajuste GARCH(P,Q)+ARMA(p,d,q)	49
5.	Modelo de Volatilidad Multivariante Multiescala	52
5.1.	Supuesto Paramétrico Local	52
5.2.	Riesgo de estimación en el caso no paramétrico	59
5.3.	Propiedades Teóricas	61
5.4.	Comportamiento bajo la condición de Pequeño Sesgo de Modelado	65
5.5.	Estabilidad después de verificada las Condiciones de Propagación	67
5.6.	Modelo de Regimen Cambiante y Calidad en la Detección de Puntos de Cambio	71
6.	Aplicación a la Estimación y Predicción en los Tipos de Cambio de Divisas	74
6.1.	Análisis Descriptivo de Datos	74

6.1.1. Aplicación LCP	85
6.1.2. Aplicación GARCH	87
6.1.3. Aplicación del Modelo DCC	103
6.2. Resultados de Ajuste por métodos DCC y LCP	106
6.2.1. Cálculo del MSqE	106
7. Resultados de Álgebra Lineal y Notaciones	115
8. Anexos	123
9. Conclusiones	128
10. Futuros Estudios	131



Índice de figuras

2.1. Descomposición del Riesgo Total	6
4.1. Intervalos Involucrados en el Estimador LAVE	44
4.2. Intervalos Involucrados en el Procedimiento LCP	48
4.3. Intervalos ℓ_k y J_k del procedimiento LCP	48
6.1. Retornos Logarítmicos Libra Esterlina.	77
6.2. Retornos Logarítmicos Yen Japonés.	77
6.3. Retornos Logarítmicos Franco Suizo.	78
6.4. Retornos Logarítmicos Corona Sueca.	78
6.5. Retornos Logarítmicos Dólar Neozelandés.	79
6.6. Retornos Logarítmicos Corona Danesa.	79
6.7. Retornos Logarítmicos Corona Noruega.	80
6.8. Retornos Logarítmicos Dólar Australiano.	80
6.9. Histograma Retornos Logarítmicos Libra Esterlina.	81
6.10. Histograma Retornos Logarítmicos Dólar Canadiense.	82
6.11. Histograma Retornos Logarítmicos Franco Suizo	82
6.12. Histograma Retornos Logarítmicos Corona Sueca.	83
6.13. Histograma Retornos Logarítmicos Dólar Neozelandés.	83
6.14. Histograma Retornos Logarítmicos Corona Danesa.	84
6.15. Histograma Retornos Logarítmicos Corona Noruega.	84
6.16. Histograma Retornos Logarítmicos Dólar Australiano.	85
6.17. Covarianza Estimada Series de Retornos Logarítmicos Libra Esterlina, Dólar Australiano [1001:1251].	86
6.18. Covarianza Estimada Series Retornos Logarítmicos Corona Sueca, Yen Japonés [501:1500].	87
6.19. Gráfico de Autocorrelación Serie Retornos Logarítmicos Corona Sueca [501:1500].	88
6.20. Gráfico de Autocorrelación Parcial Serie Retornos Logarítmicos Corona Sueca [501:1500].	89
6.21. Resumen Calidad de Ajuste ARMA(2,3) Serie Retornos Corona Sueca [501:1500].	90
6.22. Residuales Modelo ARMA(2,3) Serie Retornos Corona Sueca [501:1500].	91
6.23. Residuales Cuadrados Modelo ARMA(2,3), Serie Retornos Corona Sueca [501:1500].	91

6.24. ACF Residuales Cuadrados Estandarizados GARCH(1,1).	95
6.25. PACF Residuales Cuadrados Estandarizados GARCH(1,1).	96
6.26. Gráfico de Autocorrelación Serie Retornos Yen Japonés [501:1500].	96
6.27. Gráfico de Autocorrelación Parcial Serie Retornos Yen Japonés [501:1500].	97
6.28. Resumen Ajuste ARMA(2,2), Serie Retornos Yen Japonés [501:1500].	98
6.29. Residuales Ajuste ARMA(2,2), Serie Retornos Yen Japonés [501:1500].	98
6.30. Residuales Cuadrados Ajuste ARMA(2,2), Serie Retornos Yen Japonés [501:1500].	99
6.31. ACF Residuales Cuadrados Estandarizados Retornos Logarítmicos Yen Japonés [501:1500] GARCH(1,1).	102
6.32. PACF Residuales Cuadrados Estandarizados Retornos Logarítmicos Yen Japonés [501:1500] GARCH(1,1).	102
6.33. Ajuste DCC Serie de Retornos Logarítmicos Yen Japonés y Corona Sueca [501:1500].	106
6.34. Datos Reales Serie de Retornos Logarítmicos Yen Japonés y Corona Sueca [1251:1500].	112
6.35. Ajuste LCP Serie de Retornos Logarítmicos Yen Japonés y Corona Sueca [1251:1500].	113
6.36. Ajuste DCC Serie de Retornos Logarítmicos Yen Japonés y Corona Sueca [1251:1500].	113



Índice de cuadros

6.1. Resumen Estadísticos de Tendencia Central y Posición Datos Tipos de Cambio.	75
6.2. Resumen Estadísticos de Dispersión y Forma Datos de Tipo de Cambio. . . .	75
6.3. Resumen Estadísticos de Tendencia Central y Posición para las Series de Retornos.	76
6.4. Resumen Estadísticos de Dispersión y Forma para las Series de Retornos. . .	76
6.5. Test de Lilliefors para las Series de Retornos Logarítmicos.	85
6.6. Ajuste ARMA de las Series de Retornos Logarítmicos.	87
6.7. Error de Raíz Cuadrática Media para la Serie Bivariada de los Retornos del Yen Japonés con las 8 combinaciones restantes de los Tipos de Cambio con respecto al Dólar Estadounidense.	108
6.8. Error de Raíz Cuadrática Media para la Serie Bivariada de los Retornos del Franco Suizo con las 7 combinaciones restantes de los Tipos de Cambio con respecto al Dólar Estadounidense.	109
6.9. Error de Raíz Cuadrática Media para la Serie Bivariada de los Retornos del Corona Noruega con las 6 combinaciones restantes de los Tipos de Cambio con respecto al Dólar Estadounidense.	109
6.10. Error de Raíz Cuadrática Media para la Serie Bivariada de los Retornos del Corona Danesa con las 5 combinaciones restantes de los Tipos de Cambio con respecto al Dólar Estadounidense.	110
6.11. Error de Raíz Cuadrática Media para la Serie Bivariada de los Retornos del Dólar Neozelandés con las 4 combinaciones restantes de los Tipos de Cambio con respecto al Dólar Estadounidense.	110
6.12. Error de Raíz Cuadrática Media para la Serie Bivariada de los Retornos del Dólar Australiano con las 3 combinaciones restantes de los Tipos de Cambio con respecto al Dólar Estadounidense.	111
6.13. Error de Raíz Cuadrática Media para la Serie Bivariada de los Retornos del Dólar Canadiense con las 2 combinaciones restantes de los Tipos de Cambio con respecto al Dólar Estadounidense.	111
6.14. Error de Raíz Cuadrática Media para la Serie Bivariada de los Retornos de la Libra Esterlina con la Corona Sueca con respecto al Dólar Estadounidense. .	111

Capítulo 1

Introducción

El interés por el concepto de la volatilidad está sustentado por la relación inherente con el riesgo, asociado a un activo financiero, en el mercado bursátil.

Desde el trabajo presentado por Robert Engle en 1982, donde se propone a los modelos ARCH para su formulación, se han desarrollado bastante modificaciones en el sentido paramétrico, por nombrar algunos, los modelos GARCH, EGARCH, TGARCH, IGARCH, combinaciones de éstos, entre otros. Sin embargo, algunas razones que han llevado a buscar nuevos modelos y estimadores paramétricos, es que éstos, no logran describir adecuadamente todas las características de la volatilidad, siendo bastante restrictivos en los supuestos que permiten aplicar el modelo.

El objetivo fundamental del presente trabajo de tesis de Magíster en Estadística de la Universidad de Concepción es proponer un nuevo modelo no paramétrico para estimar la volatilidad utilizando una filtración adaptativa multivariada en el entorno de series de tiempo financieras con aplicación a datos reales, verificando su ajuste en comparación a modelos paramétricos clásicos.

La naturaleza del estudio es teórico-práctico, ya que en primera instancia se desarrolla la teoría necesaria para el planteamiento del nuevo estimador, la demostración de los teoremas y propiedades que validan su formulación. Posteriormente, se aplica a datos reales de Tipos de Cambio de Divisas en el contexto bivariado.

La contribución de este trabajo al conocimiento, es poder contar con un nuevo modelo no paramétrico teórico multivariado, que describa comportamientos financieros, también puede ser aplicado para definir carteras de inversión eficientes, estudiar correlaciones.

El texto está organizado de la siguiente forma:

- Capítulo 2: Marco Teórico. Se definen conceptos claves como rendimiento, volatilidad y riesgo financiero. Se clasifican los tipos de riesgo. Se definen las características empíricas de la volatilidad. Además se exponen los tipos de estimadores, mencionando algunos ejemplos, haciendo énfasis en los de la familia ARCH y los basados en la función Kernel. Concluyendo el capítulo con la definición de la divergencia de Kullback-Leibler, sus propiedades, descripción de la problemática y describiendo un caso particular.
- Capítulo 3: Objetivos. Se establece el objetivo principal y los objetivos específicos de esta investigación.
- Capítulo 4: Metodología. Se expone la metodología, haciendo referencia a los estimadores LAVE y LCP, cuya característica en común es que sustentan procedimientos adaptativos de estimación.
- Capítulo 5: Modelo de Volatilidad Multivariado Multiescala. Se describe el modelo y las Condiciones de Propagación. Además de estudiar el caso no paramétrico, calcular el riesgo de estimación en esta circunstancia. Por otro lado, se definen propiedades

teóricas y se demuestran los teoremas pertinentes. Se expresa una forma explícita para los valores críticos y se manifiesta la estabilidad del estimador. Finalmente, se menciona una aplicación a un modelo de volatilidad con régimen cambiante.

- Capítulo 6: Aplicación a la Estimación y Predicción a los Tipos de Cambio de Divisas. Se realiza el análisis descriptivo de los datos. Se procede a ajustar los modelos Dynamic Conditional Correlation (DCC) y se aplica el procedimiento Adaptativo Local Multivariante Multiescala, mediante el paquete AWS. Finalmente, se compara el rendimiento de ambos estimadores.
- Capítulo 7: Anexos.
- Capítulo 8: Conclusiones.
- Capítulo 9: Futuros Estudios.



Capítulo 2

Marco Teórico

2.1. Conceptos Generales



La globalización de los mercados y las herramientas tecnológicas que existen hoy en día permiten que el trabajo de los inversionistas sea más exhaustivo, debido a que cuentan con más instrumentos para hacer frente al análisis del mercado bursátil.

Una inversión es cualquier instrumento en el que se depositen fondos con el objetivo de conservar o aumentar el capital de una empresa o particular. El rendimiento de una inversión o acción es la ganancia o pérdida obtenida durante un determinado tiempo. Esto ocurre de dos formas; mediante ingresos corrientes (pagos periódicos producto de intereses) o ganancias de capital por vender a un mayor precio el bien comprado. (Gitman y Joehnk, 2009) [17].

En el artículo Forecasting Volatility, Figlewski, (1997) [12] menciona:

“En un mercado eficiente los movimientos de los precios de los activos pueden ser

descrito por:

$$R_t = \frac{S_t - S_{t-1}}{S_{t-1}} = \frac{S_t}{S_{t-1}} - 1 \quad \forall t = 1, \dots, T$$

”(p.3).

Donde S_t y S_{t-1} denota precio de los activos en un tiempo t , $t - 1$ respectivamente. Esta expresión define la rentabilidad obtenida. A su vez, se puede definir la rentabilidad continua aplicando un logaritmo a la expresión anterior $R_t = \ln \frac{S_t}{S_{t-1}} = \ln S_t - \ln S_{t-1}$, lo que se conoce en la literatura como retornos logarítmicos.

Al hablar de rendimiento es inherente mencionar el concepto de riesgo, es decir la posibilidad de que el rendimiento real de un activo no sea lo esperado. En la práctica sólo es motivo de preocupación, estudio y control el riesgo de pérdida. Gitman y Zutter (2009) [16] plantea:



“El riesgo asociado con determinado instrumento de inversión puede ser resultado de una combinación de posibles causas. Un inversionista prudente toma en cuenta cómo las principales causas de riesgo podrían afectar los posibles instrumentos de inversión.” (p.143)

Los principales riesgos se pueden clasificar en:

- **Riesgo de Inversión de un Activo Individual:** Como se mencionó anteriormente, este riesgo surge por la posibilidad de que el riesgo objetivo sea distinto de lo esperado. El riesgo debe ser tratado como una variable aleatoria debido a la incertidumbre. El rendimiento esperado, es el rendimiento promedio que se espera obtener si la inversión pudiese hacerse muchas veces bajo las mismas condiciones. Sus causas suelen ser económicas, legales o propias de la industria, podemos nombrar, PIB, tasa de interés bancaria, inflación, tributación, etc.
- **Riesgo de Portafolio:** El riesgo en que incurre la empresa está determinado por todos

CAPÍTULO 2. MARCO TEÓRICO

los activos que posee. El objetivo es obtener un portafolio eficiente que logre minimizar el riesgo. Una estrategia consiste en diversificar la cartera de activos.

- **Riesgo Sistemático:** Es el generado por la diversificación de la cartera incurriendo en mercados internacionales. Como ejemplo, fluctuación de las divisas, crisis económicas, guerras, pandemias, etc.

El Riesgo diversificable a través de una óptima gestión administrativa y financiera puede eliminarse casi por completo, el único riesgo relevante es entonces el no diversificable o sistemático. (Gitman y Zutter, 2009) [16], (Van Horne y Wachowicz, 2010) [37].

Se puede resumir el Riesgo total mediante el siguiente gráfico:

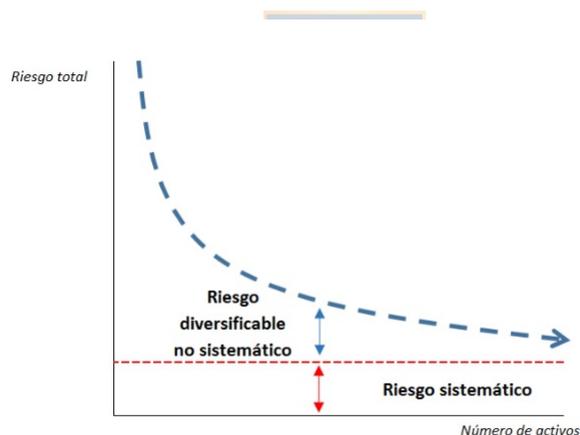


Figura 2.1: Descomposición del Riesgo Total

El riesgo no es algo observable o tangible en el mercado, por lo cual se acostumbra utilizar a la varianza o más bien, la desviación estándar de la serie temporal de rentabilidades de un activo como medición de este importante factor en la toma de decisiones. Por lo tanto, se busca estudiar, modelar y predecir la desviación estándar en algún horizonte de tiempo.

La relación entre el rendimiento, el riesgo y la volatilidad es inmediata. Como se men-

CAPÍTULO 2. MARCO TEÓRICO

cionó anteriormente, el rendimiento es la posible ganancia o pérdida que resulta de una inversión. Cabe mencionar que las acciones más rentables son también las más riesgosas, por ende, la serie temporal tenderá a poseer una mayor varianza y por consiguiente una mayor volatilidad.

La postura frente al riesgo dependerá entre otras cosas de la actitud que tenga el inversionista. En general, ésta se puede clasificar en tres categorías; aversión, indiferencia y amante del riesgo.

- Aversión : Es un inversionista al cuál no le gusta el riesgo, por ende exige un mayor rendimiento a cambio de exponerse.
- Indiferencia : No exige un mayor rendimiento a cambio de exponerse a un mayor riesgo.
- Amante del riesgo: Acepta menos rendimiento a cambio de mayor riesgo.



La mayoría de los inversionistas tienen una postura de aversión. Tratan de minimizar el riesgo para determinado nivel de rendimiento o maximizar el rendimiento para un determinado nivel de riesgo.

Por lo tanto, estudiar la volatilidad es muy importante para las finanzas, ya que además de ayudarnos a estudiar el rendimiento de un activo individual, también es utilizado para seleccionar carteras de inversión eficientes mediante la correlación, que como sabemos está definida utilizando a la desviación estándar individual y la covarianza de las inversiones. (Gitman y Joenk, 2009) [17].

2.2. Definición de Volatilidad

La importancia de estudiar la volatilidad radica en el hecho de ser uno de los indicadores de riesgo más importantes utilizados por los inversionistas para decidir si invertir en un determinado bien. (Novales, 2017) [27]. En general los inversores buscan la mayor rentabilidad posible con el menor riesgo.

En Márquez, M. (2002) [22] se formula una definición de volatilidad que tiene sentido en el contexto de este trabajo:

“La volatilidad es una medida de la intensidad de los cambios aleatorios o impredecibles en la rentabilidad o en el precio de un título; en la representación gráfica de una serie histórica de rendimientos se asocia con las fluctuaciones del rendimiento tanto en que se consideren en valor absoluto como en desviaciones al rededor de un valor medio”. (p. 21).

Existen características comunes a la volatilidad histórica de una serie financiera que hacen creer que es posible modelar y predecir esta variable. En términos generales se puede mencionar:

- **Leptocurtosis:** La fluctuación de los precios se desviará de la distribución lognormal. El valor de curtosis es mayor que 3, es decir, es más puntiaguda que una normal. Como consecuencia la distribución de la volatilidad en términos coloquiales tendrá colas más pesadas, es decir existe una baja probabilidad de que un evento extremo ocurra, pero cuando ocurra causará grandes consecuencias, por ejemplo pérdidas significativas en el precio del activo. (Fama, E., 2010) [10], (Figlewsky, 2004) [12].
- **Asimetría:** Presenta una asimetría negativa, es decir la volatilidad aumenta más cuando hay bajas en los precios y a su vez la volatilidad disminuye cuando los precios suben. (García et al., 2019) [18], (Engle, 1993) [9].

- Reversión a la media: Si un activo presenta rentabilidades más bajas en un mercado bursátil el comportamiento sistemático es obtener rentabilidades más altas en los períodos siguientes que compensarán las pérdidas. A largo plazo el comportamiento será más moderado. (Figlewsky, 2004) [12].
- Correlación en serie: En general la volatilidad depende de valores pasados, es decir varianza condicionada al pasado.(Francq, y Zakoian, 2010) [15],(Engle, 1982) [6].
- Discontinuidad de saltos en los precios: Esporádicamente ocurren valores altos de volatilidad en momentos concretos. (Figlewsky, 1997) [12].
- Presencia de conglomerados: Si la volatilidad es alta sigue siendolo en el siguiente periodo. (Engle, R., 1982) [6]. (Engle, 1993) [8].
- Movimientos vinculados: Al analizar una serie financiera de un mismo concepto en distintos mercados, se puede observar como los movimientos importantes afectan a todos los mercados donde interviene el bien.(Marquéz, 2002) [22].

2.3. Estimadores de la Volatilidad

La volatilidad puede estimarse mediante un valor puntual o utilizando una medida serial. El estimador es puntual cuando se considera un conjunto de rentabilidades observadas durante algún tiempo y se calcula la desviación estándar. Esta medida sólo permite encontrar un número asociado pero no conocer el comportamiento histórico de la variable, necesaria para predecir y modelar. Por esta razón, es menester considerar una medida proveniente de una serie de tiempo, ahora bien, el estimador puede ser paramétrico o no paramétrico.

Para una mejor comprensión del tema tratado en ésta tesis se van a definir a grandes

rasgos algunos de los estimadores utilizados sin ánimo de hacer un exhaustivo estudio de ellos.

2.3.1. Estimadores Puntuales

Sean R_t , $t = 1, \dots, k$, las rentabilidades observadas en un tiempo determinado. Los siguientes son estimadores puntuales de la volatilidad:

Desviación Estándar:

$$\sigma_t = \sqrt{\frac{1}{k} \sum_{j=0}^k r_{t-j}^2}$$

con $r_{t-j} = R_{t-j} - \mu$ donde μ es el retorno esperado.

En Robles, F. (2002) [30], se menciona que esta medida pondera de la misma forma las observaciones próximas y lejanas de los rendimientos, además de que el parámetro k se escoge de manera arbitraria de acuerdo a la apreciación personal de los inversionistas.

Suavizado Exponencial:

$$\sigma_t = \sqrt{(1 - \alpha)r_t^2 + \alpha\sigma_{t-1}^2}$$

con $0 < \alpha < 1$,

Este modelo soluciona la problemática anterior dando más importancia a retornos más cercanos en el tiempo en el cuál se está midiendo la volatilidad y quitando el hecho de tener que decidir el valor de k . (Jordan, 1996). [24].

2.3.2. Estimadores Paramétricos

Entre los estimadores paramétricos que asumen un modelo estadístico, consideramos importante mencionar los de la familia ARCH (Autoregressive Conditional Heteroskedasticity), postulado por Engle, (1982). [6].

Para predecir la volatilidad se acostumbraba introducir una variable exógena, entonces, $y_t = x_t \epsilon_t$ define un proceso gaussiano donde para cada t fijo ϵ_t tiene distribución Normal Estándar.

Debido a lo restrictivo y complejo de esta situación, Engle propone introducir una variable endógena para estimar la volatilidad. El modelo que establece es el siguiente:

$$y_t = \sigma_t \epsilon_t \quad \text{donde} \quad \mathbb{E}(\epsilon_t) = 0 \quad \text{y} \quad \text{Var}(\epsilon_t) = 1.$$



La varianza condicionada del proceso está dada por:

$$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + \alpha_2 y_{t-2}^2 + \dots + \alpha_q y_{t-q}^2 = \alpha_0 + \sum_{i=1}^q \alpha_i y_{t-i}^2$$

En conclusión, Engle propone que la varianza condicional actual cambie a través de una función lineal de los retardos al cuadrado. Si las innovaciones son normales entonces la variable aleatoria y_t también será normal y través del estimador máximo verosímil se puede estimar el valor de σ_t .

Para que este modelo sea válido se deben establecer condiciones de regularidad:

- $\alpha_0 > 0$

- $\alpha_i \geq 0 \quad i = 1, \dots, q$
- $\sum_{i=1}^q \alpha_i \leq 1$

Éstas se imponen para que la varianza condicional sea positiva para todos los períodos y la serie estacionaria.

El estimador, al ser una combinación lineal de parámetros α_i con los rezagos al cuadrado y_{t-i}^2 , permite darle menos importancia a rezagos lejanos, modificando el peso de ellos en la volatilidad condicionada actual. Además los valores grandes en la innovación afectan más a la varianza condicional debido a que están expresadas al cuadrado en el estimador, por lo que una gran perturbación tiene la tendencia a preservar este comportamiento. Es decir, es capaz de describir el comportamiento de la volatilidad por grupos. Además logra describir la Leptocurtosis característica de la dinámica de la volatilidad.

Si bien, este modelo revolucionó a los economistas, presenta algunas deficiencias que fueron abordadas sistemáticamente. Por nombrar algunas, el modelo asume que los shock positivos afectan del mismo modo a la volatilidad que los negativos, es decir no modela la asimetría característica de la volatilidad, (Fisher, 1976) [13], (Christie, 1982) [5], (Tsay, 2005) [36]. Otra desventaja es que, se necesita muchos parámetros para ajustar y por consiguiente muchos retardos, generando un número elevado de iteraciones lo que puede suponer pérdidas de precisión.

Modelos GARCH

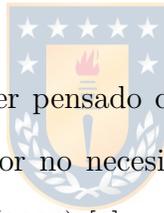
En 1986 Bollerslev introduce una modificación al modelo ARCH, los modelos GARCH (Generalized Autoregressive Conditional Heteroskedasticity) [3]. Este nuevo modelo considera la posibilidad de que la volatilidad pueda depender de los valores pasados de la varianza

condicional.

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i y_{t-i}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2$$

Este modelo también necesita establecer condiciones de regularidad para resguardar la estacionaridad y positividad. En efecto,

- $\alpha_0 > 0$
- $\alpha_i \geq 0 \quad i = 1, \dots, q$
- $\beta_i \geq 0 \quad i = 1, \dots, p$
- $\sum_{i=1}^q \alpha_i + \sum_{i=1}^p \beta_i \leq 1$



El modelo GARCH(p,q) puede ser pensado como un modelo ARCH(∞). Esta generalización supera al modelo anterior por no necesitar tantos parámetros para capturar la dinámica de la volatilidad. Bollerslev, (1986) [3], muestra un claro ejemplo donde se utilizó un modelo GARCH(1,1) versus los modelos ARCH(8) y ARCH(4), obteniendo un mejor ajuste por el modelo GARCH(1,1), además de tener una estructura lag más representativa. (p.321-322)

Los modelos GARCH(p,q) son bastante utilizados en econometría, en particular el modelo GARCH(1,1) ya que con sólo dos parámetros permite una descripción parsinómica de la volatilidad.

Algunas limitación de los modelos ARCH y GARCH es que las restricciones de no negatividad de los parámetros son difíciles de conseguir en la práctica. Además, se ha encontrado una correlación negativa entre los rendimientos actuales y la volatilidad futura. (Nelson, 1991) [26], no logran captar la persistencia de la varianza condicional (Bollerslev, y

Engle, 1993) [8], (Engle, 1986) [7].

Modelos EGARCH

Fue propuesto por Nelson, (1991) [26], en un afán de describir las características de la volatilidad que no logran ser captadas por los modelos GARCH, en particular, el efecto de apalancamiento o asimetría. A través de estudios empíricos se ha notado que las buenas y malas noticias afectan de distinta forma a la volatilidad. En general los shocks negativos generan un mayor impacto que los shocks positivos incluso considerando shocks de la misma magnitud.

El modelo está dado por:

$$y_t = \varepsilon_t \sigma_t$$

$$\ln(\sigma_t^2) = \alpha_t + \sum_{k=1}^{\infty} \beta_k g(\varepsilon_{t-k})$$

tanto $\{\alpha_t\}$ como $\{\beta_k\}$ son sucesiones de números reales no estocásticas.

σ_t^2 está definida como la varianza condicional de y_t , por ende, es no negativa. Al aplicar un logaritmo natural a σ_t^2 , el valor obtenido puede ser positivo o negativo. Para mantener la positividad necesaria en el postulado, Nelson, [26] propone que $\ln(\sigma_t^2)$ sea una función lineal de los rezagos utilizando una función “ g ” apropiada.

Esta innovación a la hora de imponer la positividad le da más flexibilidad al modelo ya que los parámetros estimados puede ser tanto positivos, negativos o cero. Esta es una de las principales diferencias con el modelo desarrollado por Bollerslev, donde la positividad se obtiene restringiendo el valor de los parámetros a valores no negativos.

Para explicar como aborda la asimetría se debe estudiar a la función “ g ” la cual está definida como una función del signo y la magnitud de y_t . Es decir, una función lineal de ε_t

y $|\varepsilon_t|$. En efecto,

$$g(\varepsilon_t) = \theta\varepsilon_t + \gamma(|\varepsilon_t| - \mathbb{E}(|\varepsilon_t|))$$

Observamos que si ε_t es no negativo $g(\varepsilon_t)$ es una recta con pendiente $\theta + \gamma$, por el contrario si ε_t es negativo $g(\varepsilon_t)$ es una recta con pendiente $\theta - \gamma$. A raíz de esto σ_t^2 puede responder de manera asimétrica.

Para una representación más parsinómica, el proceso se puede modelar con un $ARMA(p, q)$. En contraposición la formulación anterior es capaz de captar los procesos de memoria larga.

$$\ln(\sigma_t^2) = \alpha_0 + \sum_{j=1}^q (\alpha_j \ln \sigma_{t-j}^2) + \sum_{i=1}^p (\beta_i g(\varepsilon_{t-i}))$$



Modelos TARCH

Zakoian, (1992) [39], establece una modificación a los modelos ARCH que logra captar la asimetría en el comportamiento de la volatilidad, al igual que el modelo desarrollado por Nelson, en 1991 [26]. Instaure una forma de introducir la asimetría utilizando una función que dependa de las partes positivas y negativas de las innovaciones. Formalmente,

$$y_t = \sigma_t \varepsilon_t$$

$$\sigma_t = \alpha_0 + \sum_{i=1}^q (\alpha_i^+ y_{t-i}^+) - \alpha_i^- y_{t-i}^- + \sum_{j=1}^p (\beta_j \sigma_{t-j})$$

con ε_t iid, $\mathbb{E}(\varepsilon_t) = 0$, $\text{Var}(\varepsilon_t) = 1$, ε_t independiente de y_{t-1} para todo t .

y_t denota un proceso en tiempo discreto, y_{t-1} representa el conjunto de información conocida al tiempo t , $y_t^+ = \max(y_t, 0)$ $y_t^- = \min(y_t, 0)$ son la parte positiva y negativa de y_t .

$(\alpha_i^+)_{i=1,\dots,q}, (\alpha_i^-)_{i=1,\dots,q}$ y $(\beta_j^+)_{j=1,\dots,p}$ son secuencias de escalares.

Una ventaja mencionada por el autor, aparte del hecho de que permite que los shock positivos y negativos de igual tamaño tengan diferentes impactos en la volatilidad, es que no necesita establecer restricciones de positividad de los parámetros, lo que es muy positivo en referencia a la inferencia numérica. Sin embargo en cuanto a las propiedades probabilísticas, el estudio de vuelve complejo. Por lo mencionado anteriormente, se establecen de igual modo restricciones de positividad:

$$\alpha_0 > 0, \alpha_i^+ \geq 0, \alpha_i^- \geq 0, \beta_i \geq 0$$

para todo t .

Si bien, TARCH logra captar asimetrías al igual que EGARCH, estos difieren en el hecho de que TARCH es formulado como un modelo aditivo, donde la volatilidad es una función de innovaciones no normalizadas a diferencia de EGARCH. Aún más EGARCH impone una estructura constante en todos los retardos, TARCH, permite que los distintos retardos puedan otorgar contribuciones opuestas a la asimetría. Como última observación EGARCH sólo utiliza ecuaciones no lineales para cualquier función de ε , lo que tiene como consecuencia que los métodos de inferencia en dos pasos de mínimos cuadrados no sean aplicables. Por el contrario TARCH mantiene esta propiedad.

En palabras del mismo autor se menciona como conclusión que la mejoría en el modelo, surge por la forma de modelar la asimetría y el hecho de utilizar la desviación estándar condicional en vez de la varianza condicional.

Según el análisis desarrollado en el estudio señalado al comparar los modelos $ARCH(5)$, $TARCH(5)$, $GARCH(1, 1)$, $STGARCH(1, 1)$, $TGARCH(1, 1)$ y $EGARCH(1, 1)$ se obser-

va que los estimadores que describen mejor el comportamiento de los datos son los modelos asimétricos. El modelo $TARCH(5)$ posee un menor exceso de curtosis que los demás, lo que mejoraría la calidad de la estimación.

2.3.3. Estimadores No Paramétricos

En la sección anterior vimos que para estimar la volatilidad se pueden utilizar supuestos paramétricos del modelo, lo que conlleva a una estimación de los parámetros que la caracterizan. Si bien, estos métodos son muy utilizados, a su vez son bastante rígidos, no toleran desviaciones de las hipótesis que permiten plantear el modelo paramétrico.

Para subsanar esta falta de flexibilidad se utilizan métodos no paramétricos. Vamos a describir algunos de los métodos no paramétricos para su mejor comprensión, en particular los derivados del Estimador Kernel.



2.3.4. Técnicas de Filtrado: Suavización Kernel

Para conocer cabalmente el comportamiento de una variable aleatoria es necesario conocer las funciones que la caracterizan, a saber, la función de densidad y distribución, con el objetivo de poder describir, realizar cálculos, inferencias, etc.

En general conocer la función de distribución y densidad de una v.a. no es factible, y tal supuesto suele ser bastante restrictivo. No obstante, se pueden utilizar métodos de regularización o suavización de datos que permiten hacer frente a la ausencia de conocimiento de la distribución de los datos. Existen métodos de regularización globales, por ejemplo métodos penalizados o proyector. (Flynn et al, 2013), [14]

En el presente estudio nos enfocamos en métodos locales de suavización. Una de las poderosas herramientas de localización es el estimador Kernel.

Estimador por histograma

Sean x_1, x_2, \dots, x_n observaciones independientes de una v.a. X . Consideremos el intervalo donde se encuentran las observaciones formando una partición de éstos. En efecto:

Sea $a_0 < a_1 < \dots < a_m$ una sucesión tal que $a_0 = \min_{i=1, \dots, n} x_i$ y $a_m = \max_{i=1, \dots, n} x_i$,

$B_j = [a_{j-1}, a_j]$, $j = 1, \dots, m$, intervalos, n_j la cantidad de observaciones perteneciente a B_j y f_j la frecuencia relativa de dicho intervalo, entonces:

$$\sum_{i=1}^n I_{B_j}(x_i) f_j = \frac{n_j}{n} = \frac{1}{n} \sum_{i=1}^n I_{B_j}(x_i).$$

Donde I es la función Indicatriz, es decir $I_A(x) = \begin{cases} 0 & \text{si } x \notin A \\ 1 & \text{si } x \in A \end{cases}$.

Para cada intervalo se dibuja un rectángulo que tiene por base B_j y altura c_j , tal que el área obtenida sea f_j . Sea x el punto donde se quiere estimar la densidad, luego el estimador de $f(x)$ es la altura mencionada, c_j .

Sin pérdida de generalidad, podemos utilizar intervalos de la misma longitud, es decir $h = a_{j+1} - a_j$ con h fijo. Al hacer variar $x \in \mathbb{R}$ obtenemos el estimador por histograma de la función de densidad. (Zambom, 2012) [40].

$$\hat{f}(x) = \sum_{i=1}^n f_i \frac{1}{h} I_{B_j}$$

Este estimador no es suave ya que utiliza la unión de funciones constantes a intervalos, por lo tanto no es continuo. Es importante notar que su forma y resultado está muy afectado por el ancho del intervalo seleccionado. Si h es pequeño la estimación tendrá poco sesgo y mucha varianza. Si h es muy grande la estimación tendrá mucho sesgo y poca varianza.

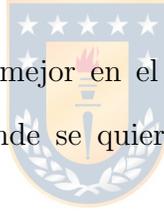
A continuación veremos que el mismo efecto de bias variance tradeoff es inherente a los estimadores locales en general. Por lo tanto el objetivo será encontrar una longitud h que equilibre el sesgo y la varianza.

Polígonos de Frecuencia

Corrige la falta de continuidad del estimador por histograma. Tomemos en cuenta la partición anterior a_0, \dots, a_m , pero ahora tomaremos en consideración el punto medio de cada uno de los intervalos, se unirán los dos puntos medios de los intervalos contiguos mediante una recta. Este proceso se llama interpolación de los datos. (Scott, 1985) [32].

Estimador Kernel

El estimador por histograma es mejor en el centro del intervalo, razón por la cual podemos considerar a x , el punto donde se quiere estimar la densidad como centro del intervalo $[x - h, x + h]$



$$\hat{f}_U(x) = \frac{1}{nh} \sum_{i=1}^n I_{[x-h, x+h]}(x_i) = \frac{1}{nh} \sum_{i=1}^n \frac{1}{2} I_{[-1,1]}\left(\frac{x - x_i}{h}\right)$$

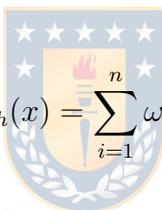
Luego hacemos variar x en todo \mathbb{R} , el estimador definido anteriormente se puede ver como una generalización del estimador por histograma.

Notamos que la función indicatriz que se está utilizando en la construcción del estimador no es suave por lo que podemos utilizar otras que cumplan con esta condición. A la función K se le llama Kernel o núcleo de la estimación.

Las funciones de suavización más utilizadas son:

Núcleo	$K(t)$	Rango
Epanechnikov	$\frac{3}{4}(1 - t^2)$	$ t < 1$
Gauss	$\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}$	$ t < \infty$
Triángular	$1 - t $	$ t < 1$
Rectangular	$\frac{1}{2}$	$ t < 1$
Biweight	$\frac{15}{16}(1 - t^2)^2$	$ t < 1$
Triweight	$\frac{35}{32}(1 - t^2)^3$	$ t < 1$
Arco Coseno	$\frac{\pi}{4} \cos(\frac{\pi}{2}t)$	$ t < 1$

Llevemos este concepto a un modelo de regresión. Sea $y_i = f(x_i) + \epsilon_i$ con $i = 1, \dots, n$. La idea es estimar la función de regresión real mediante: (Rosenblatt,1956) [31],(Parzen,1962) [29].



$$\hat{f}_{n,h}(x) = \sum_{i=1}^n \omega_{hi} y_i$$

donde $\omega_{hi} = \frac{K(\frac{x_i - x}{h})}{\sum_{i=1}^n K(\frac{x_i - x}{h})}$, K es una función núcleo.

Al igual que el estimador por histograma, al ser una generalización, el estimador kernel se ve afectado por la elección del ancho de banda. Si la longitud es muy pequeña se produce casi una interpolación de los datos por la presencia de mucho detalles y gran dispersión. Por el contrario si el ancho de banda es muy grande se obtiene una estimación muy suavizada y con gran sesgo. Por eso se debe buscar seleccionar el ancho de banda de manera óptima equilibrando el sesgo y la varianza. (Wasserman, 2006) [38].

La elección del ancho de banda se puede realizar utilizando los siguientes procedimientos: (Wasserman, 2006) [38].

- Reglas Básadas en Distribuciones Paramétricas.

- Reglas de Validación cruzada.
- Métodos Plug-In.
- Bootstrap Suavizado.

Características de las funciones núcleo:

La característica esencial de la función kernel que garantiza la consistencia del estimador es:

$$\int_{-\infty}^{+\infty} K(x) = 1$$

Habitualmente, salvo en las fronteras del intervalo de estimación, se trabaja con núcleos simétricos, esto es, $K(x) = K(-x)$.



Además se puede considerar un núcleo de orden s , es decir

$$\int_{-\infty}^{+\infty} x^i K(x) = 0 \quad i = 1, \dots, s - 1$$

$$\int_{-\infty}^{+\infty} x^s K(x) \neq 0$$

Otros supuestos deseables y postulados por Nadaraya, (1989) [25] son, por ejemplo:

$$|SupK(x)| < \infty$$

$$\int_{-\infty}^{+\infty} x^s |K(x)| < \infty$$

Se define primero el Sesgo y la Varianza de un estimador.

$$Sesgo\hat{f}(x) = \mathbb{E}[\hat{f}(x)] - f(x), \quad Var\hat{f}(x) = \mathbb{E}[\hat{f}(x) - \mathbb{E}(\hat{f}(x))]^2$$

Como medida del error de estimación se utiliza el MSE (Error Cuadrático Medio), el MISE (Error Cuadrático Medio Integrado), criterio de Kullback-Leibler, Normas L_p . (Zambom y Dias, 2012) [40]

$$MSE\hat{f}(x) = \mathbb{E}(\hat{f}(x) - f(x))^2 = \text{Var}\hat{f}(x) + (\text{Sesgo}\hat{f}(x))^2$$

$$MISE\hat{f}(x) = \mathbb{E}\left[\int (\hat{f}(x) - f(x))^2 dx\right]$$

$$L_p\hat{f}(x) = \left[\int |\hat{f}(x) - f(x)|^p\right]^{\frac{1}{p}}$$

En lo que resta de la sección estudiaremos la forma de estimar un modelo de regresión. Consideremos el modelo dado por $Y = f(x) + \varepsilon$, tal que $\mathbb{E}(\varepsilon) = 0$, entonces tenemos un Modelo de Regresión Media, ya que el valor esperado de la variable Y es efectivamente la función de regresión.



Ésta se dice paramétrica si $f(x) = f(x, \theta)$ donde $\theta = (\theta_1, \theta_2, \dots, \theta_p) \in \mathbb{R}^p$.

Al considerar n observaciones el modelo anterior queda definido por $y_i = f(x_i) + \varepsilon_i$ para $i \in 1, \dots, n$.

Sean $\varepsilon_i = y_i - f(x_i, \theta)$ errores del modelado y $\tilde{\theta}$ un estimador de θ , dicha estimación nos conducirá a estimadores de los errores individuales $\tilde{\varepsilon}_i = y_i - f(x_i, \tilde{\theta})$, lo que se conoce como residuales.

El objetivo es encontrar el estimador ideal, que haga que la distribución empírica P_n de los residuales imite de la mejor forma las características del error.

Se debe escoger una función ψ que minimize la esperanza de la siguiente suma:

$$\sum_{i=1}^n \psi(y_i - f(x_i, \theta))$$

, a saber:

$$\theta^* = \operatorname{argmin}_{\theta} \mathbb{E} \sum_{i=1}^n \psi(y_i - f(x_i, \theta))$$

Restringiendo a un conjunto paramétrico Θ el M-Estimador será:

$$\tilde{\theta} = \operatorname{argmin}_{\theta \in \Theta} \sum_{i=1}^n \psi(y_i - f(x_i, \theta))$$

Si $\mathbb{E}(\varepsilon_i) = 0$, $\mathbb{E}(\varepsilon_i^2) = \sigma_i < \infty$ y $\psi(u) = u^2$, al calcular la esperanza del error al cuadrado, llegaremos al estimador de mínimos cuadrados.

$$\tilde{\theta}_{LSE} = \operatorname{argmin}_{\theta \in \Theta} \sum_{i=1}^n (y_i - f(x_i, \theta))^2$$



En este problema se puede aplicar el enfoque de máxima verosimilitud, sea ρ la función de densidad de los errores ε_i que se suponen independientes e idénticamente distribuidos, al calcular la función de verosimilitud $\Pi \rho(y_i - f(x_i, \theta))$, y definiendo $l(y_i - f(x_i, \theta)) = \log \rho(y_i - f(x_i, \theta))$, maximizando obtenemos una solución al problema.

$$\tilde{\theta} = \operatorname{argmax}_{\theta \in \Theta} \sum_{i=1}^n l(y_i - f(x_i, \theta))$$

De manera análoga, si falla alguna o ambas de las suposiciones que permiten aplicar el enfoque de máxima verosimilitud, se puede operar mediante el enfoque de cuasi-máxima verosimilitud que nos conducirá al estimador de mínimos cuadrados en el caso de errores gaussianos, (Spokoiny y Dickhaus, 2015) [35].

Todos los métodos mencionados son enfoques locales, puesto que aplican regularización de los datos mediante su localización en la proximidad del punto de estimación. En su contraparte no paramétrica global tenemos estimación de proyección, métodos a trozos y

splines en todas sus variantes.

Estimación por la proyección

Sea $Y_i = f(x_i) + \varepsilon_i$, donde f es una función no paramétrica, para poder caracterizarla se considera un sistema de funciones ortogonales tal que:

$$f(x) = \sum_{j=1}^{\infty} \theta_j \psi_j(x).$$

Truncando la suma anterior en una cantidad “ p ” de coeficientes se puede aplicar el enfoque de cuasi-máxima verosimilitud, asumiendo que el modelo está mal especificado.

En conclusión, la función desconocida no paramétrica se aproxima mediante alguna expansión, ya sea polinomios ortogonales, funciones wavelets, trigonométricas, etc, para aplicar el enfoque de cuasi-máxima verosimilitud haciendo un supuesto distribucional sobre los errores del modelo minimizando el sesgo o residuo. (Spokoiny y Dickhaus, 2015) [35].

Métodos a trozos

Sea A_1, \dots, A_m una partición disjunta de X , si asumimos que f es constante a trozos tenemos la siguiente representación:

$$f(x) \approx \sum_{i=1}^m \theta_i \mathbb{1}(x_i \in A_i)$$

Para cada k , θ_k es estimado por:

$$\tilde{\theta}_k = \operatorname{argmax}_{\theta_k} \sum_{x_i \in A_k} l(y_i - \theta_k) \quad k = 1, \dots, m$$

Por lo tanto el estimador será:

$$\tilde{\theta} = \operatorname{argmax}_{\theta} L(\theta) = \operatorname{argmax}_{\theta=\theta_k} \sum_{k=1}^m \sum_{x_i \in A_k} l(y_i - \theta_k)$$

En el caso de la estimación lineal o polinomial a trozos lo único que cambia es la forma de la función que aproxima la función de regresión desconocida. (Spokoiny y Dickhaus, 2015), [35].

Splines

Este método se implementa para lograr que la aproximación de la función cumpla con una propiedad deseable, la continuidad. Sean $t_0 < t_1 < \dots < t_k$ nodos, un spline consiste en aproximar f por un polinomio de grado q en cada intervalo generado por los nodos preestablecidos. Dado un nodo, los polinomios y sus respectivas derivadas deben tener límites laterales iguales. La representación es la siguiente:

$$f(x) \approx \sum_{j=0}^q \alpha_j x^j + \sum_{k=1}^{m-1} \theta_k \phi_k(x)$$

donde $\phi_k(x) = (x - t_k)_+^q$ con $k = 1, \dots, m - 1$.

Smoothing Spline

Corresponde a un método penalizado donde se debe optimizar $L_{\lambda}(f) = L(f) - \lambda R_q(f) = \sigma^{-2} \sum |y_i - f(x_i)|^2 - \lambda \int_{\mathbb{X}} |f^q(x)|^2 dx$, donde λ es un valor propio de Lagrange. (Silverman y Green, 2019) [19].

Regresión Generalizada

En este caso se supone que las observaciones individuales y_i pertenecen a una familia paramétrica y que sólo el parámetro depende de los puntos de diseño, es decir $y_i \sim P_{f(x_i)}$.

Suponiendo un modelo lineal generalizado, es decir, $f(x_i) = \psi_i^T \theta$ y una parametrización canónica de la familia exponencial llegamos a la función de verosimilitud:

$$L(\theta) = S^T \theta - A(\theta)$$

con $S = \sum y_i \psi_i$ y $A(\theta) = \sum d(\psi_i^T \theta)$.

Esta función es cóncava, por lo tanto existe una solución y además es única. Para encontrarla se acostumbra utilizar el método de Newton-Rapson, hasta alguna convergencia indicada. (Spokoiny y Dickhause, 2015) [35]

Dado un estimador inicial $\theta^{(0)}$, el estimador $k+1$ -ésimo:

$$\theta^{(k+1)} = \theta^{(k)} + B(\theta^{(k)})^{-1} (S - \nabla A(\theta^{(k)})).$$

Donde ∇ es el operador vectorial diferencial.



2.4. Divergencia de Kullback-Leibler

Sean f y g dos funciones de distribución de una variable aleatoria, se define a la divergencia de Kullback-Leibler o también llamada entropía relativa como,

$$\int f(x) \log \left(\frac{f(x)}{g(x)} \right) dx.$$

Propiedades de la Divergencia de Kullback-Leibler, (Spokoiny y Dickhaus, 2015) [35].

- La K-L es no negativa.

Utilizando la desigualdad de Jensen, $\varphi(\int_{\Omega} g d\mu) \leq \int_{\Omega} \varphi g d\mu$, con φ convexa.

$$\begin{aligned} - \int f(x) \log \left(\frac{f(x)}{g(x)} \right) dx &= \int f(x) \log \left(\frac{g(x)}{f(x)} \right) dx \\ &\leq \log \left(\int f(x) \frac{g(x)}{f(x)} dx \right) = \log(1) = 0 \end{aligned}$$

- Es nula sí y solo sí $f(x) = g(x)$.
- No es simétrica, por lo tanto no se trata de una distancia.

2.5. Razón de Verosimilitud

Sea Y_1, Y_2, \dots, Y_n variables aleatorias independientes e idénticamente distribuidas con función de densidad marginal $f(y, \theta)$.



La función de verosimilitud se define por:

$$l(\theta, y_1, y_2, \dots, y_n) = \prod_{i=1}^n f(y_i, \theta)$$

De manera análoga se define la razón de verosimilitud:

$$\frac{l(\theta, y_1, y_2, \dots, y_n)}{l(\theta^*, y_1, y_2, \dots, y_n)}$$

Sin embargo, se acostumbra por comodidad aplicar un logaritmo a la expresión anterior $L(\theta, y_1, \dots, y_n) = \log(l(\theta, y_1, \dots, y_n))$, por lo tanto la razón de verosimilitud se puede expresar por:

$$L(\theta, y_1, y_2, \dots, y_n) - L(\theta^*, y_1, y_2, \dots, y_n)$$

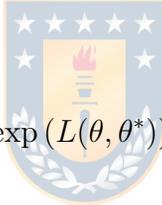
Razón de Verosimilitud Ajustada

Corresponde a la razón de verosimilitud donde se utiliza el Estimador Máximo Verosimil
 $\tilde{\theta} = \text{máx}_{\theta^* \in \Theta} l(\theta^*, y_1, y_2, \dots, y_n)$.

$$L(\theta, y_1, y_2, \dots, y_n) - L(\tilde{\theta}, y_1, y_2, \dots, y_n)$$

Propiedades de la Razón de Verosimilitud

Teorema 2.5.1: Bajo el supuesto paramétrico, es decir, $Y_i \sim f_{\theta^*}$ donde Y_i son variables aleatorias independientes e idénticamente distribuidas con $\theta^* \in \Theta$:



$$\mathbb{E}_{\theta^*} \exp(L(\theta, \theta^*)) = 1$$

Demostración:

$$\begin{aligned} \mathbb{E}_{\theta^*} \exp(L(\theta, \theta^*)) &= \int_{\mathbb{R}^n} \prod_{i=1}^n \frac{f_{\theta}(y_i)}{f_{\theta^*}(y_i)} f(y_1, y_2, \dots, y_n, \theta^*) dy_1 dy_2 \dots dy_n \\ &= \int_{\mathbb{R}^n} \prod_{i=1}^n \frac{f_{\theta}(y_i)}{f_{\theta^*}(y_i)} \prod_{i=1}^n f_{\theta^*}(y_i) dy_i \\ &= \int_{\mathbb{R}^n} \prod_{i=1}^n \frac{f_{\theta}(y_i)}{f_{\theta^*}(y_i)} f_{\theta^*}(y_i) dy_i \\ &= \int_{\mathbb{R}^n} \prod_{i=1}^n f_{\theta}(y_i) dy_i \\ &= \prod_{i=1}^n \int_{\mathbb{R}} f_{\theta}(y_i) dy_i = 1 \end{aligned}$$

2.6. Descripción de la problemática

Se observa (X_i, Y_i) , $i = 1, \dots, n$, donde X_1, X_2, \dots, X_n puntos de diseño aleatorio o determinístico e Y_1, Y_2, \dots, Y_n variables de respuesta. Entonces $Y_i = f(X_i) + \varepsilon_i$ con $i = 1, \dots, n$ representa un modelo de regresión.

Suponiendo que $\mathbb{E}(\varepsilon_i) = 0$, entonces se puede concluir que $\mathbb{E}(Y_i|X_i) = f(X_i)$.

Por lo tanto, el modelo de regresión anterior queda expresado mediante la siguiente igualdad:

$$Y_i = \mathbb{E}(Y_i|X_i) + \varepsilon_i \quad i = 1, \dots, n$$

es decir, un modelo de regresión media. (Wasserman 2015) [38], (Serdyukova, 2012) [33]. Para estimar la función de regresión se pueden utilizar métodos paramétricos o no paramétricos.



En este contexto, sean R_1, R_2, \dots, R_{t-1} variables aleatorias que definen la rentabilidad en el precio de un activo en un determinado tiempo, y sean $i = 1, 2, \dots, t$ puntos de diseño. Por consiguiente, existe un proceso de precios de un activo S_t , tal que $R_t = \log(\frac{S_t}{S_{t-1}})$. [12].

Para lograr predecir la rentabilidad en el precio de un activo, se considera un modelo con ruido multiplicativo:

$$R_t = \sigma_t \varepsilon_t$$

donde σ_t es un proceso de volatilidad y $\varepsilon_t \sim N(0, 1)$ v.a iid. Dicho proceso es medible con respecto a \mathcal{F}_{t-1} , donde $\mathcal{F}_{t-1} = \sigma(R_1, R_2, \dots, R_{t-1})$ es σ álgebra del pasado generado por el proceso R .

En términos simples, esto significa que la volatilidad en un tiempo "t" queda determinada por la información de las rentabilidades pasadas.

Sin pérdida de generalidad podemos aplicar una transformación de potencia al modelo anterior:

$$Y_t = R_t^2 = \sigma_t^2 \varepsilon_t^2$$

Ahora bien, el problema radica en estimar el proceso de volatilidad, $\theta_t = \sigma_t^2$, a partir de los datos observados Y_1, Y_2, \dots, Y_{t-1} . Según lo acotado con anterioridad, corresponde a estimar la función de regresión del modelo. Bajo el enfoque de adaptación local, se puede considerar que $\theta_t \approx \theta$ en un intervalo suficientemente pequeño, cuya longitud debe ser determinada a partir de los datos, lo que nos conducirá al estimador Nadaraya-Watson para cada tal intervalo. En efecto, se supone que existe un intervalo I , de longitud N_I que debe ser determinado a partir de los datos, tal que, θ_t puede ser una constante o bien aproximado por una constante dentro de I .



Como método de estimación se utilizará el enfoque de cuasi máxima verosimilitud localizado para un modelo mal especificado, ya que se está asumiendo que ε_i son i.i.d y se distribuyen normal estándar. Bajo esta suposición y restringiendo al intervalo I ,

$$L_I(\theta) = \sum_{t \in I} l(Y_t, \theta) = \sum_{t \in I} l(Y_t, \theta) \mathbb{1}_{\{t \in I\}}$$

Utilizando la función de densidad de la normal, $l(y, \theta) = -\frac{1}{2} \log(2\pi\theta) - \frac{y}{2\theta}$. El estimador cuasi máximo verosimil está dado por:

$$\tilde{\theta}_I = \operatorname{argmax}_{\theta} L_I(\theta) = \frac{1}{N_I} \sum_{t \in I} Y_t$$

Para que quede clara la conexión con la estimación no paramétrica notamos que $\mathbb{1}_{\{t \in I\}}$

actúa como función kernel lo que nos lleva al estimador Nadaraya-Watson ya mencionado.

$$\tilde{\theta}_I = \frac{1}{N_I} \sum_{t \in I} Y_t = \frac{\sum_t Y_t \mathbb{1}_{\{t \in I\}}}{\sum_t \mathbb{1}_{\{t \in I\}}}.$$

Para medir la exactitud de la estimación se utiliza la razón de verosimilitud, la que queda definida de manera proporcional a la divergencia de Kullback-Leibler, bajo condiciones de normalidad, el estimador máximo verosimil $\tilde{\theta}$ de θ cumple

$$L_I(\tilde{\theta}, \theta) = L_I(\tilde{\theta}) - L(\theta) = N_I K_I(\tilde{\theta}, \theta)$$

Akaike(1973) [2] nos entrega una forma de extender el principio de máxima verosimilitud. Considerando la maximización de la log-verosimilitud esperada podemos definir lo siguiente. En efecto, para un conjunto de estimadores $\tilde{\theta}$ del parámetro θ , donde $X \sim f(\cdot, \theta)$, tenemos que

$$\mathbb{E} \left(\log f(X, \tilde{\theta}) \right) = \mathbb{E} \int f(x, \theta) \log f(x, \tilde{\theta}) dx.$$

La esperanza es con respecto a la distribución de $\tilde{\theta}$.

Para medir la calidad de la estimación se utiliza la siguiente relación, mostrando la conexión con la divergencia de Kullback-Leibler:

$$\begin{aligned} \mathbb{E} \left(\log \frac{f(X, \tilde{\theta})}{f(X, \theta)} \right) &= \mathbb{E} \int f(x, \theta) \log \frac{f(x, \tilde{\theta})}{f(x, \theta)} dx \\ &= -\mathbb{E} K(\theta, \tilde{\theta}). \end{aligned}$$

Se definen las funciones de pérdida y de riesgo como siguen:

$$W(\theta, \tilde{\theta}) = -2 \int f(x, \theta) \log \frac{f(x, \tilde{\theta})}{f(x, \theta)} dx = 2K(\theta, \tilde{\theta})$$

$$R(\theta, \tilde{\theta}) = \mathbb{E} \left(W(\theta, \tilde{\theta}) \right)$$

Al considerar una muestra aleatoria simple de tamaño N ,

$$-\frac{2}{N} \sum_{i=1}^N \log \left(\frac{f(X_i, \tilde{\theta})}{f(X_i, \theta)} \right)$$

es un estimador consistente de $W(\theta, \hat{\theta})$. Por lo tanto el estimador que maximiza la relación anterior minimizará la función de pérdida. Al considerar un conjunto de estimadores de θ , se escogerá el que genere un mínimo riesgo. A esto se le conoce como extensión del principio de máxima verosimilitud.

Para resumir, Akaike (1973) [2] demuestra que bajo condiciones de regularidad, maximizar la esperanza de la log-verosimilitud corresponde a minimizar la divergencia de Kullback Leibler, esto es muy importante porque nos permite trabajar con el modelo mal especificado, por lo tanto el método de cuasi-verosimilitud nos lleva al mejor estimador del parámetro.

Más aún, Akaike propuso considerar una malla anillada de los parámetros $\theta_{k,k+1}, \theta_{k,k+2}, \dots, \theta_{k,\varkappa}$ correspondientes a $f(x, \theta_k)$, $k = 0, 1, \dots, \varkappa$, sus respectivos estimadores, seleccionando el "mejor" que minimiza el riesgo respectivo. Esa idea puede ser considerada como primera aproximación al concepto de los métodos locales adaptativos.

Para generalizar lo mencionado anteriormente, este estudio propone extender a un modelo normal multivariante, donde $R_t = \sum_t^{\frac{1}{2}} \varepsilon_t$ con \sum_t una matriz simétrica de la volatilidad.

Dicha matriz puede contener funciones constantes o no.

2.6.1. Caso Particular

Consideremos el modelo de rentabilidad de un activo:

$$R_t = \sigma_t \varepsilon_t.$$

Supongamos que el ruido se distribuye normal estándar, entonces el modelo de rentabilidad tendrá una distribución normal con media 0 y varianza σ_t^2 .

Si se supone que la volatilidad dentro de un intervalo, que luego se definirá a partir de los datos, puede ser bien aproximada por una constante, es decir, $\sigma_t^2 = \theta$ con $t \in I$ y se aplica una transformación al modelo anterior:

$$Y_t = R_t^2.$$

Lo que antes correspondía a estimar la varianza del modelo ahora corresponde a estimar la media del proceso transformado. En efecto,

$$\mathbb{E}_\theta(Y) = \mathbb{E}_\theta(R_t^2) = \mathbb{E}_\theta((\sigma_t \varepsilon_t)^2) = \theta.$$

Si se utiliza el logaritmo de la función de densidad de la distribución normal, $l(y, \theta) = \frac{1}{2} \log(2\pi\theta) - \frac{y}{2\theta}$. Luego la divergencia de Kullback-Leibler por definición viene dada mediante la expresión:

$$K(\theta, \theta') = \mathbb{E}_\theta \left(l(Y, \theta) - l(Y, \theta') \right).$$

Tomando en cuenta que

$$\begin{aligned}\mathbb{E}_\theta(l(Y, \theta)) &= \mathbb{E}_\theta\left(-\frac{1}{2}\log(2\pi\theta) - \frac{Y}{2\theta}\right) \\ &= -\frac{1}{2}\log(2\pi\theta) - \frac{\theta}{2\theta} \\ &= -\frac{1}{2}\log(2\pi\theta) - \frac{1}{2}.\end{aligned}$$

Análogamente,

$$\mathbb{E}_\theta(l(Y, \theta')) = -\frac{1}{2}\log(2\pi\theta') - \frac{\theta}{2\theta'}.$$

Por lo tanto,



$$K(\theta', \theta) = -\frac{1}{2}\left\{\log\left(\frac{\theta'}{\theta}\right) + 1 - \frac{\theta'}{\theta}\right\}.$$

El estimador máximo verosímil de θ en el intervalo I será $\tilde{\theta}_I = \operatorname{argmax}_\theta L_I(\theta)$. Teniendo presente que, bajo condiciones de normalidad, $L_I(\theta'_I, \theta) = N_I K(\theta'_I, \theta)$. Entonces,

$$\tilde{\theta}_I = \operatorname{argmin}_\theta K(\theta'_I, \theta), \quad t \in I.$$

Por ende, para cada intervalo I ,

$$L_I(\tilde{\theta}_I, \theta) = N_I K(\tilde{\theta}_I, \theta).$$

Spokoiny y Dickhaus, (2005) [35] mencionan propiedades locales de la divergencia de

Kullback Leibler que son de vital importancia. Bajo condiciones de regularidad, se tiene que:

- $K(\theta, \theta')|_{\theta'=\theta} = 0$
- $\frac{\partial}{\partial \theta'} K(\theta, \theta')|_{\theta'=\theta} = 0$
- $\frac{\partial^2}{\partial \theta'^2} K(\theta, \theta')|_{\theta'=\theta} = \mathbb{F}(\theta)$

Por lo tanto, la divergencia de Kullback- Leibler bajo se puede aproximar localmente

$$K(\theta, \theta') \approx \frac{1}{2} \mathbb{F}(\theta) |\theta - \theta'|^2,$$

donde \mathbb{F} es la cantidad de información de Fisher.

En el caso de la familia exponencial de distribuciones, consideremos parametrización canónica de una familia exponencial, tal que la función de densidad $p(y, \theta) = \frac{dP_v(y)}{d\mu_0} = p(y) \exp(yv - d(v))$, donde $p(y)$ es una función no negativa y $d(v)$ es una función convexa dos veces diferenciable. Entonces (véase Spokoiny y Dickhaus (2005)):

$$\begin{aligned} \mathbb{E}_v Y &= d'(v), \\ \text{Var}_v Y &= d''(v) = \mathbb{F}(v) \end{aligned}$$

donde $\mathbb{F}(v) = \mathbb{E}_v |l'(Y, v)|^2$ es la información de Fisher.

Además,

$$\begin{aligned}
 K(v, v_1) &= \mathbb{E}_v(l(Y, v, v_1)) \\
 &= \int \log\left(\frac{p(y, v)}{p(y, v_1)}\right) P_v(dy) \\
 &= \int \{y[v - v_1] - [d(v) - d(v_1)]\} P_v(dy) \\
 &= d'(v)(v - v_1) - [d(v) - d(v_1)] \\
 &= \frac{1}{2}d''(\check{v})(v_1 - v)^2
 \end{aligned}$$

donde \check{v} es un punto entre v y v_1 . La última igualdad se debe al Teorema de Taylor.

Por otro lado $\mathbb{P}_{v^*}(L(v, v^*) > z) \leq e^{-z}$. En efecto por la desigualdad de Markov

$$\mathbb{P}_{v^*}(e^{L(v, v^*)} > e^z) \leq \frac{\mathbb{E}_{v^*} \exp L(v, v^*)}{e^z} \leq \frac{1}{e^z} \leq e^{-z}.$$

Spokoiny (2009) establece el siguiente teorema: Sea $f(t) = \theta^*$, $\forall t \in I$, si las innovaciones son normal estándar e iid, para cualquier $z > 0$

$$\mathbb{P}_{\theta^*}(L_I(\tilde{\theta}_I, \theta^*) \geq z) = \mathbb{P}_{\theta^*}(N_I K_I(\tilde{\theta}_I, \theta^*) \geq z) \leq 2e^{-z}$$

Es decir tenemos una desigualdad de concentración para el estimador $\tilde{\theta}_I$, con lo que se puede definir una función de perdida de estimación para el parámetro, para cada intervalo I .

Teorema (Spokoiny(2009)): Sea R_t v.a i.i.d con $\mathcal{N}(0, \theta^*)$.Entonces, para todo $r \geq 0$ y todo intervalo I

$$\mathbb{E}_{\theta^*}|L_I(\tilde{\theta}_I, \theta^*)|^r = \mathbb{E}_{\theta^*}|N_I K_I(\tilde{\theta}_I, \theta^*)|^r \leq \tau_r$$

con $\tau_r = 2r \int_{z \geq 0} z^{r-1} e^{-z} dz = 2r\Gamma(r)$.

Si X es una variable aleatoria no negativa, entonces

$$\mathbb{E}(X^r) \leq \int_0^{+\infty} t^{r-1} d\mathbb{P}(X \geq t)$$

$$\begin{aligned} \mathbb{E}(X^r) &= \lim_{b \rightarrow +\infty} \int_0^b t d\mathbb{P}(X \leq t) \\ &= \lim_{b \rightarrow +\infty} b^r \mathbb{P}(X \leq b) - \int_0^b r t^{r-1} \mathbb{P}(X \leq b) dt \\ &= \lim_{b \rightarrow +\infty} b^r \mathbb{P}(X \leq b) - b^r + b^r - \int_0^b r t^{r-1} \mathbb{P}(X \leq b) dt \\ &= \lim_{b \rightarrow +\infty} -b^r (1 - \mathbb{P}(X \leq b)) + \int_0^b r t^{r-1} dt - \int_0^b r t^{r-1} \mathbb{P}(X \leq b) dt \\ &= \lim_{b \rightarrow +\infty} -b^r (1 - \mathbb{P}(X \leq b)) + \int_0^b r t^{r-1} (1 - \mathbb{P}(X \leq b)) dt \\ &= \int_0^{+\infty} r t^{r-1} (1 - \mathbb{P}(X \leq t)) dt \\ &= \int_0^{+\infty} r t^{r-1} \mathbb{P}(X \geq t) dt \\ &= \int_0^{+\infty} r t^{r-1} d\mathbb{P}(X \geq t) \end{aligned}$$

Capítulo 3

Objetivos y Actividades

3.1. Objetivo General



Proponer un nuevo modelo para estimar la volatilidad utilizando una filtración adaptativa multivariada en el entorno de series de tiempo financieras con aplicación a datos reales, verificando su ajuste en comparación a modelos paramétricos clásicos.

3.2. Objetivos Específicos

- Identificar los principales modelos paramétricos utilizados para estimar la volatilidad en series de tiempo financieras.
- Formular de manera teórica el nuevo modelo multivariado local adaptativo.
- Describir las propiedades teóricas del nuevo modelo.
- Comparar el modelo adaptativo multivariado con modelos paramétricos de Correlación

CAPÍTULO 3. OBJETIVOS Y ACTIVIDADES

Dinámica Condicional.



Capítulo 4

Materiales y Métodos



4.1. Técnicas de Filtrado Adaptativo

4.1.1. Modelamiento de Volatilidad mediante Retornos Logarítmicos. Transformación de Potencia

Sea S_t con $t = 1, \dots, t^*$, el precio de un activo, los retornos logarítmicos están dados por:

$$R_t = \ln S_t - \ln S_{t-1}. \quad (4.1)$$

Si bien, es de conocimiento común que los retornos logarítmicos no gozan de la normalidad, gracias a la elección adaptativa del intervalo de estimación este proceso se puede

modelar mediante heterocedasticidad condicional:

$$R_t = \sigma_t \varepsilon_t, \quad \varepsilon_t \sim N(0, 1). \quad (4.2)$$

Donde las innovaciones son independientes e idénticamente distribuidas normales estándar y σ_t denota la volatilidad variable en el tiempo. Se asume que el proceso de volatilidad σ_t es medible con respecto a la filtración \mathcal{F}_{t-1} generada por las $t - 1$ observaciones anteriores de R_t .

$$\sigma_t \sim \mathcal{F}_{t-1} \quad \mathcal{F}_{t-1} = \sigma(R_1, R_2, \dots, R_{t-1})$$

En Mercurio y Spokoiny, (2004) [23] se aplica una transformación al modelo de volatilidad (4.2) para generar un problema similar al de regresión.

Transformación de Potencia 2.

$$\log R_t^2 = \log \sigma_t^2 + C + V\zeta_t \quad (4.3)$$

Donde $C = \mathbb{E}(\log \varepsilon_t^2)$, $V^2 = \text{Var}(\log \varepsilon_t^2)$ y $\zeta_t = V^{-1}(\log \varepsilon_t^2 - C)$.

Es la transformación más natural para obtener un modelo de regresión aditivo, con función de regresión $f(t) = \log(\sigma_t^2) + C$, y ruido homogéneo $V\zeta_t$. Véase Gouriéroux C., (1997). [20]

Claramente, el proceso R_t transformado es condicionalmente gaussiano y $\mathbb{E}(R_t^2 | \mathcal{F}_{t-1}) = \sigma_t^2$. Mercurio D. y Spokoiny W., , (2004) [23], mencionan que el problema de esta representación es que la distribución de los errores ζ_t , es muy sesgada, es decir da pesos muy altos a valores pequeños de ε_t .

Transformación de Potencia γ .

Carrol, R. y Ruppert, D. (1988) [4], proponen una transformación en la potencia γ de (4.2).

$$|R_t|^\gamma = C_\gamma \sigma_t^\gamma + D_\gamma \sigma_t^\gamma \zeta_t. \quad (4.4)$$

$$\mathbb{E}(|R_t|^\gamma / \mathcal{F}_{t-1}) = \sigma_t^\gamma \mathbb{E}(|\varepsilon|^\gamma / \mathcal{F}_{t-1}) = C_\gamma \sigma_t^\gamma,$$

$$\text{Var}(|R_t|^\gamma / \mathcal{F}_{t-1}) = \mathbb{E}(|R_t|^\gamma - C_\gamma \sigma_t^\gamma / \mathcal{F}_{t-1})^2 = \sigma_t^{2\gamma} \mathbb{E}(|\varepsilon|^\gamma - C_\gamma)^2 = \sigma_t^{2\gamma} D_\gamma^2. \text{ Con } C_\gamma = \mathbb{E}|\varepsilon|^\gamma \text{ y } D_\gamma^2 = \text{Var}|\varepsilon|^\gamma.$$

Estimar σ_t es equivalente a estimar $\theta_t = C_\gamma \sigma_t^\gamma$ que corresponde a la media del proceso transformado, con ruido aditivo $D_\gamma \sigma_t^\gamma \zeta_t$.

4.1.2. Estimador LAVE



Mercurio, D. y Spokoiny, V. (2004) [23] introduce el estimador LAVE (Locally adaptive volatility estimate). La importancia de este enfoque es que no asume una estructura paramétrica del modelo, afirma que la volatilidad desconocida puede ser aproximada por una constante de manera adaptativa en un intervalo de tiempo homogéneo. Esto quiere decir, la volatilidad se aproxima por una constante en un intervalo cuya longitud se elige a partir de los datos, esto es, el estimador se adapta automáticamente a los datos. Se establece un modelo sencillo pero donde el parámetro puede variar en el tiempo, lo que le aporta flexibilidad. La característica de que el método se compruebe y adapte regularmente a los datos, lleva a los siguientes objetivos, encontrar el mayor intervalo donde la volatilidad puede ser bien aproximada mediante una constante y estimar el valor de la volatilidad local. En este contexto, el estimador LAVE se puede entender como una técnica de filtrado que permite seleccionar el mayor intervalo de homogeneidad.

4.1.3. Estimación Adaptativa con Homogeneidad Local

El supuesto de intervalo de tiempo homogéneo significa que:

Para todo t^* , existe $I = [t^* - m, t^*]$ tal que σ_t es aproximadamente constante y R_t satisface la ecuación con tendencia $\theta_I = C_\gamma \sigma_I^\gamma$.

El objetivo es determinar el mayor intervalo I , donde σ_t puede ser bien aproximado por una constante y se puede aplicar el modelo de regresión (4.4) cuyo parámetro θ_I puede obtenerse promediando dentro del intervalo I :

$$\tilde{\theta}_I = \frac{1}{|I|} \sum_{t \in I} |R_t|^\gamma.$$

Para el caso específico de $\gamma = 2$ dicho estimador $\tilde{\theta}_I$ coincide con el estimador máximo verosimil de la volatilidad σ_t^2 . Véase (Fan et all, 2003) [11].

Del supuesto de tiempo homogéneo: [23]

$$\begin{aligned} \tilde{\theta}_I &= \frac{C_\gamma}{|I|} \sum_{t \in I} \sigma_t^\gamma + \frac{D_\gamma}{|I|} \sum_{t \in I} \sigma_t^\gamma \zeta_t \\ &= \frac{1}{|I|} \sum_{t \in I} \theta_t + \frac{S_\gamma}{|I|} \sum_{t \in I} \theta_t \zeta_t \end{aligned}$$

Donde $S_\gamma = \frac{D_\gamma}{C_\gamma}$. Luego $\mathbb{E}[\tilde{\theta}_I] = \frac{1}{|I|} \mathbb{E}(\sum_{t \in I} \theta_t)$. Además $V_I^2 = \text{Var}(\tilde{\theta}_I) = \frac{S_\gamma^2}{|I|^2} \sum_{t \in I} \theta_t^2$.

Notemos que gracias al supuesto de homogeneidad local el proceso θ_t es cercano a θ_{t^*} para todo $t \in I$.

Descripción del proceso de selección del intervalo de Homogeneidad

Mercurio, D. y Spokoiny, V.(2004) [23] consideran una familia de subintervalos de I de la forma $J = [t^* - m', t^*]$, donde $m' < m$ y su complemento $J^c = [t^* - m, t^* - m'[,$ luego se comparan los dos estimadores $\tilde{\theta}_J$ y $\tilde{\theta}_{J^c}$.

Si $|\tilde{\theta}_{J^c} - \tilde{\theta}_J| > \lambda \sqrt{\tilde{V}_J^2 + \tilde{V}_{J^c}^2}$ será rechazada la hipótesis de homogeneidad y el procedimiento para. En caso contrario se continuará iterando eligiendo un intervalo mayor.

Sea \hat{I} el último intervalo no rechazado, luego el estimador LAVE será $\hat{\theta}_{t^*} = \tilde{\theta}_{\hat{I}}$

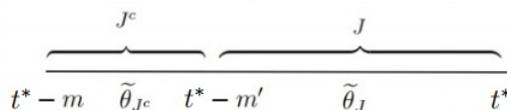


Figura 4.1: Intervalos Involucrados en el Estimador LAVE



Estimador LAVE para Detección de Puntos de Cambio

Consideramos una secuencia de tiempos de parada dada por: $\tau_1 < \tau_2 < \dots <$ respecto a la filtración \mathcal{F}_t y por los valores $\sigma_1, \sigma_2, \dots$, donde cada σ_k es medible respecto a la filtración \mathcal{F}_{τ_k} . Este modelo son un caso específico importante del modelo 4.2. Se define $\sigma_t = \sigma_k$ para $\tau_k < t \leq \tau_{k+1}$ y σ_t es constante para $t < \tau_1$. En este contexto buscamos el mayor intervalo posible que no contenga un punto de cambio. También se basa en la observación de los valores medios de las observaciones y_t dentro de los intervalos $[t^* - m, t^* - m']$ y en cada subintervalo $J = [t^* - m', t^*]$. El proceso adaptativo difiere un poco del anterior. Cuando se quiere detectar un punto de cambio se debe cumplir una probabilidad prescrita de una "falsa alarma", es decir detectar un punto de cambio es condiciones de homogeneidad. El procedimiento adaptativo nos lleva entonces a un entorno de pruebas múltiples.

Teorema (Mercurio y Spokoiny): Sea $I = [t^* - m, t^*]$ un intervalo de homogeneidad,

$\sigma_t = \sigma_{t^*}$ para todo $t \in I$, entonces:

$$P(\text{Rechazar } I) \leq \sum_{I \in I} \sum_{J \in J(I)} 6 \exp \frac{-\lambda^2}{2a_\gamma (1 + \lambda s_\gamma |J|^{-\frac{1}{2}})^2}$$

Este Teorema nos entrega un límite superior para esta probabilidad de falsa alarma. El valor λ se encuentra mediante simulaciones Monte Carlo bajo la hipótesis nula de que no haya cambios estructurales.

La calidad de la detección de los puntos de cambio se mide por el tiempo medio entre la aparición del punto de cambio y su correspondiente detección. Se considera el caso de estimación en un punto t^* inmediatamente después de un punto de cambio t^*_{PC} , por lo tanto el intervalo I será $I = [t^*_{PC}, t^*]$, se menciona que la estimación será tan buena como tener información previa de la localización de los puntos de cambio.



4.1.4. Estimador LCP

Spokoiny, (2009) [34], propone una nueva forma de determinar el intervalo de homogeneidad utilizando el análisis de punto de cambio local. Se consideran las condiciones de propagación con el objetivo de minimizar el sesgo de manera óptima. La calidad de la estimación se mide mediante Desigualdades de Oráculo.

El constructo de la estimación de la volatilidad utiliza el mayor intervalo posible, aceptado por el procedimiento, que no contenga ningún punto de cambio.

Este enfoque presenta ventajas comparado con el estimador LAVE, en cuanto a su simplicidad, sensibilidad a identificar cambios estructurales y la forma de definir la probabilidad de una falsa alarma. Una falsa alarma ocurre al rechazar la hipótesis de no existencia de un punto de cambio en condiciones de homogeneidad. Además el proceso puede extenderse al modelo de volatilidad múltiple.

El procedimiento no asume ninguna información a priori de la distribución subyacente, ni siquiera de manera local. Debido a esto, hay un precio a pagar por el desconocimiento de la distribución dada por la desigualdad oráculo, que será mencionada más adelante.

Caso Particular

Se explicará un caso particular del modelo de rentabilidad con el fin de poder extender metodología de Spokoyny, 2009, [34]. Suponiendo homogeneidad del modelo, es decir $\sigma_t^2 = \theta$, el modelo de retornos cuadráticos se puede escribir como sigue

$$Y_t = R_t^2 = \theta \varepsilon_t^2$$

considerando innovaciones normales, el parámetro se puede estimar utilizando el estimador máximo verosímil, por pertenecer a la familia exponencial. En el caso de que las innovaciones no sean normales a través de la divergencia de Kullback-Leibler se puede seguir aplicando el estimador máximo verosímil aún sin saber que forma funcional tiene la distribución cuyo procedimiento será llamado Cuasi-Verosimilitud.

Procedimiento Adaptativo

Consideramos el supuesto paramétrico local:

Para todo t^* , existe $[t^* - m, t^*[$ donde $\sigma_t = Cte$.

Sea $N_1 < N_2 < \dots < N_K$ una sucesión de números. Donde N_k es un parámetro de escala, es decir, describe la longitud del intervalo de tiempo histórico examinado en el paso k.

Se define la familia de intervalos anidados ℓ_k , $k = 1, \dots, K$ con $\ell_1 \subset \ell_2 \subset \dots \subset \ell_K$ donde

$$\ell_k = [t^* - N_k, t^*[$$

El procedimiento comienza testeando en el intervalo más pequeño, ℓ_1 , la hipótesis nula de homogeneidad contra la hipótesis alternativa de existencia de un punto de cambio.

Si la hipótesis no es rechazada se considera un intervalo más grande para volver a testear continuando este proceso hasta que se detecta un punto de cambio. El intervalo aceptado ℓ_k será el más grande donde no se determina la presencia de un punto de cambio local.

Si por el contrario la hipótesis es rechazada en algún ℓ_k y se detecta un punto de cambio, entonces el intervalo estimado de homogeneidad se define como el último intervalo aceptado ℓ_{k-1} . Finalmente el estimador de la volatilidad será $\hat{f}(t^*) = \hat{\theta} = \tilde{\theta}_{\ell_{k-1}}$.

Test de Homogeneidad

Sean $I = [t', t'']$ y J intervalos de prueba tal que $J \subset I$. Cada punto $\tau \in J$ divide el intervalo I en dos partes:

$$I' = [t', \tau] \text{ e } I'' = [\tau, t''].$$

La hipótesis alternativa significa que $f(t) = \theta''$ para todo $t \in I''$ e $f(t) = \theta'$ para todo $t \in I'$ con $\theta'' \neq \theta'$.

Esta hipótesis conduce a determinar la log-verosimilitud $L_{I''}(\theta'') + L_{I'}(\theta')$.

Por lo tanto el estadístico de prueba es:

$$T_{I,\tau} = \max_{\theta'', \theta'} \left\{ L_{I''}(\theta'') + L_{I'}(\theta') \right\} - \max_{\theta} L_I(\theta)$$

Definiendo $L(\theta, \theta') = L(\theta) - L(\theta')$, obtenemos:

$$\begin{aligned} &= \min_{\theta} \max_{\theta'', \theta'} \left\{ L_{I''}(\theta'', \theta) + L_{I'}(\theta', \theta) \right\} \\ &= N_{I''} K(\tilde{\theta}_{I''}, \tilde{\theta}_I) + N_{I'} K(\tilde{\theta}_{I'}, \tilde{\theta}_I) \end{aligned}$$

El test de punto de cambio para el intervalo J que Spokoiny (2009) define como el máximo del estadístico de prueba $T_{I,J} = \max_{\tau \in J} T_{I,\tau}$.

Se considera un valor crítico \mathfrak{z} calibrado mediante simulaciones Monte Carlo bajo hipótesis nula. La hipótesis de homogeneidad se rechaza cuando $T_{I,J} \geq \mathfrak{z}$ y en este caso $T_{I,J} = \max_{\tau \in J} T_{I,\tau}$.

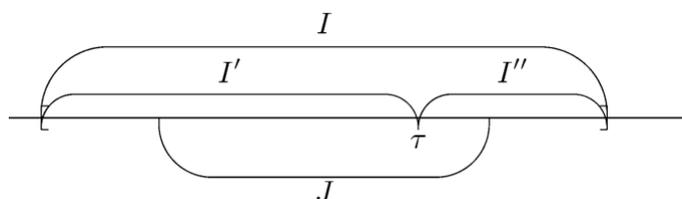


Figura 4.2: Intervalos Involucrados en el Procedimiento LCP

Considerando el caso de los intervalos anidados, el procedimiento resulta en un test multiescala y la siguiente figura ejemplifica la forma de proceder.

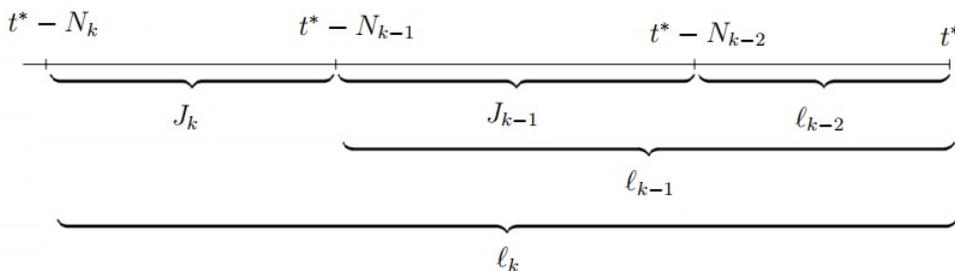


Figura 4.3: Intervalos ℓ_k y J_k del procedimiento LCP

El intervalo ℓ_0 es aceptado. En cada paso k , y en cada punto del intervalo $J_k = \ell_k - \ell_{k-1}$ se realiza el test de punto de cambio. ℓ_k es aceptado si el intervalo anterior lo es y el estadístico de prueba definido por $T_k = T_{\ell_{k+1}, J_k}$ no excede al valor crítico \mathfrak{z}_k . El evento $\{\ell_k \text{ es rechazado}\}$ si el estadístico excede para algún $\ell \leq k$. Para cada k el índice κ_k al mayor intervalo aceptado

despues de los primeros k pasos y $\hat{\theta}_k = \hat{\theta}_{\kappa_k}$ al correspondiente estimador.

$$\kappa = \max\{k \leq K : T_\nu \leq \mathfrak{z}_\nu, \nu = 1, \dots, k\} \quad \hat{\theta} = \hat{\theta}_\kappa. \quad (4.5)$$

4.2. Ajuste GARCH(P,Q)+ARMA(p,d,q)

Un modelo ARMA(p,q) está dado por:

$$Y_t = \varepsilon_t + \sum_{i=1}^p \phi_i Y_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i}.$$

Este modelo no ajusta bien a la volatilidad ya que asume que la varianza es constante, que como hemos visto no es adecuado en el contexto de series de tiempo financieras. El modelo GARCH(P,Q) mejora el ajuste.



$$Y_t = \sigma_t \varepsilon_t$$

tal que,

$$\sigma_t = \sqrt{\omega + \sum_{i=1}^p \alpha_i Y_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2}.$$

En términos generales la serie de tiempo Y_t modelado por un GARCH(P,Q), se puede pensar como el ruido de un modelo ARIMA(p,d,q) de Y_t^2 . En efecto,

Si $\eta_t = Y_t^2 - \sigma_t^2$ entonces, $\mathbb{E}(\eta_t / F_{t-1}) = 0$. Por otro lado,

$$v = \frac{\omega}{1 - \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i)}$$

$$Y_t^2 - v = \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) (Y_{t-i}^2 - v) - \sum_{j=1}^q \beta_j \eta_{t-j} + \eta_t.$$

Luego Y_t^2 es un proceso $ARMA(max(p, q), q)$ del proceso con media $\mu = v$.

Por condiciones de estacionaridad se pide que:

$$\sum_{i=1}^{max(p,q)} (\alpha_i + \beta_i) < 1$$

En este modelado hay dos tipos de residuos, el residuo común \hat{Y}_t dado por la diferencia entre Y_t y la esperanza condicional de $\mathbb{E}(Y_t/Y_1, \dots, Y_{t-1})$. Un residuo estandarizado $\hat{\varepsilon}_t$ que se obtiene mediante la división del \hat{Y}_t con el estimador de la desviación estándar $\hat{\sigma}_t$. Si el modelo ajusta bien, ni $\hat{\varepsilon}_t$ ni $\hat{\varepsilon}_t^2$ deberán presentar autocorrelación serial.

Explicaremos el procedimiento empírico para realizar el ajuste, en el software R project.

Comenzamos convirtiendo los datos en series de tiempo y calculamos los retornos logarítmicos.

$$R_t = \log\left(\frac{S_t}{S_{t-1}}\right) = \log(S_t) - \log(S_{t-1}).$$

Se grafican las series tiempos de los retornos con el fin de observar el comportamiento general de la serie de tiempo, posteriormente se analizan los gráficos de autocorrelación y autocorrelación parcial para ajustar manualmente los parámetros del modelo $ARIMA(p,d,q)$.

Se prueba con los posibles modelos, utilizando una distribución Normal o de Student para los residuales. Para posteriormente comparar los distintos modelos aplicando un test de significancia de los parámetros, además de un test de Ljung a los residuales para observar si el modelado es capaz de eliminar la autocorrelación de la serie de retornos. Si esto funciona bien, se aplica un test de Ljung a los residuales al cuadrado, si se rechaza la hipótesis de autocorrelación en los residuales al cuadrado y al aplicar la prueba ARCH se acepta la presencia de un componente GARCH en los residuales al cuadrado se procede con el ajuste

GARCH.

Una vez realizado el ajuste GARCH, se aplican las pruebas que validan el modelo. Comenzando con la significancia de los parámetros, test de Ljung para los residuales estandarizados y residuales cuadrados estandarizados, un test ARCH de los residuales, finalmente una prueba de bondad de ajuste para los residuales estandarizados. Una vez verificada estas condiciones se procede a aplicar el modelo DCC.



Capítulo 5

Modelo de Volatilidad Multivariante

Multiescala



Como generalización al modelo. Sea $R_t = \Sigma_t^{\frac{1}{2}} \varepsilon_t$ un vector de proceso de precios de un activo en tiempo discreto, con $\varepsilon_t \sim N_d(0, \mathbf{1}_d)$, donde Σ_t es la Matriz de Volatilidad simétrica de tamaño $d \times d$. Consideremos que Σ_t es medible con respecto a la σ -álgebra generada por las observaciones hasta el instante $t - 1$.

5.1. Supuesto Paramétrico Local

Para el proceso anterior, vamos a suponer que la matriz de volatilidad puede ser aproximada por una matriz constante $\Sigma_t = \Sigma$ en algún intervalo I.

El parámetro Σ puede ser estimado mediante el método de máxima verosimilitud.

$$\tilde{\Sigma}_I = \operatorname{argmax}_{\Sigma \succ 0} L_I(\Sigma) = \operatorname{argmax}_{\Sigma \succ 0} \sum_{t \in I} l(R_t, \Sigma)$$

donde $l(R_t, \Sigma) = -\frac{d}{2} \log(2\pi) - \frac{1}{2} \log |\Sigma| - \frac{1}{2} R_t^\top \Sigma^{-1} R_t$ es la función de log verosimilitud.

La función de log-verosimilitud local está dada por:

$$L_I(\Sigma) = -\frac{dN_I}{2} \log(2\pi) - \frac{N_I}{2} \log(|\Sigma|) - \frac{1}{2} \sum_{t \in I} R_t^\top \Sigma^{-1} R_t$$

donde N_I es el número de observaciones en I.

Teorema 5.1: Para cada intervalo I , el estimador máximo verosimil $\tilde{\Sigma} = \operatorname{argmax}_{\Sigma} L(\Sigma)$ viene dado por:

$$\tilde{\Sigma}_I = N_I^{-1} \sum_{t \in I} R_t R_t^\top$$

Demostración:

Sea $L_I(\Sigma) = -\frac{dN_I}{2} \log(2\pi) - \frac{N_I}{2} \log(|\Sigma|) - \frac{1}{2} \sum_{t \in I} R_t^\top \Sigma^{-1} R_t$ donde N_I es el número de observaciones en I, la función de log-verosimilitud local.

Por lo tanto,

$$\begin{aligned} \frac{\partial L_I(\Sigma)}{\partial \Sigma} &= -\frac{N_I}{2} \Sigma^{-\top} + \frac{1}{2} \sum_{t \in I} \Sigma^{-\top} R_t R_t^\top \Sigma^{-\top} \\ \frac{\partial L_I(\Sigma)}{\partial \Sigma} = 0 &\iff -\frac{N_I}{2} \Sigma^{-\top} + \frac{1}{2} \sum_{t \in I} \Sigma^{-\top} R_t R_t^\top \Sigma^{-\top} = 0 \\ N_I I &= \sum_{t \in I} \Sigma^{-\top} R_t R_t^\top \Sigma^{-\top} \Sigma^\top \\ N_I I &= \Sigma^{-\top} \sum_{t \in I} R_t R_t^\top \end{aligned}$$

El siguiente estimador satisface dicha ecuación:

$$\tilde{\Sigma}_I = N_I^{-1} \sum_{t \in I} R_t R_t^\top$$

Por el Lema 7.2 : dicho estimador es donde se alcanza el máximo de la función $L_I(\Sigma)$.

Más aún, para cada $\Sigma \succ 0$ la razón de log-verosimilitud $L_I(\tilde{\Sigma}_I, \Sigma') = \max_{\Sigma} L_I(\Sigma, \Sigma') = \max_{\Sigma} (L_I(\Sigma) - L_I(\Sigma'))$ satisface

$$L_I(\tilde{\Sigma}_I, \Sigma') = N_I K(\tilde{\Sigma}_I, \Sigma')$$

donde

$$K(\Sigma, \Sigma') = -\frac{1}{2} \left\{ \log |\Sigma| - \log |\Sigma'| - \text{tr}(\Sigma \Sigma'^{-1}) + d \right\}$$



Demostración:

$$\begin{aligned} \mathbb{E}_{\Sigma} [l(R_t, \Sigma)] &= -\frac{d}{2} \log 2\pi - \frac{1}{2} \log |\Sigma| - \frac{1}{2} \text{tr}(\Sigma \Sigma^{-1}) \\ &= -\frac{d}{2} \log 2\pi - \frac{1}{2} \log |\Sigma| - \frac{d}{2} \end{aligned}$$

De manera análoga:

$$\mathbb{E}_{\Sigma} [l(R_t, \Sigma')] = -\frac{d}{2} \log 2\pi - \frac{1}{2} \log |\Sigma'| - \frac{1}{2} \text{tr}(\Sigma \Sigma'^{-1})$$

Restando ambas expresiones tenemos:

$$K(\Sigma, \Sigma') = -\frac{1}{2} \left\{ \log |\Sigma| - \log |\Sigma'| - \text{tr}(\Sigma \Sigma'^{-1}) + d \right\}$$

Por otro lado,

$$\begin{aligned}
 \frac{1}{N_I} L_I(\Sigma, \Sigma') &= \frac{1}{N_I} \left\{ -\frac{N_I}{2} \log |\Sigma| - \frac{1}{2} \sum_{t \in I} R_t^\top \Sigma^{-1} R_t + \frac{N_I}{2} \log |\Sigma'| + \frac{1}{2} \sum_{t \in I} R_t^\top \Sigma'^{-1} R_t \right\} \\
 &= -\frac{1}{2} \left\{ \log |\Sigma| + \frac{1}{N_I} \sum_{t \in I} R_t^\top \Sigma^{-1} R_t - \log |\Sigma'| - \frac{1}{N_I} \sum_{t \in I} R_t^\top \Sigma'^{-1} R_t \right\} \\
 &= -\frac{1}{2} \left\{ \log |\Sigma| - \log |\Sigma'| + \frac{1}{N_I} \sum_{t \in I} \text{tr}(\Sigma^{-1} R_t R_t^\top) - \frac{1}{N_I} \sum_{t \in I} \text{tr}(\Sigma'^{-1} R_t R_t^\top) \right\} \\
 &= -\frac{1}{2} \left\{ \log |\Sigma| - \log |\Sigma'| + \mathbb{E}_\Sigma \text{tr}(\Sigma^{-1} R_t R_t^\top) - \mathbb{E}_{\Sigma'} \text{tr}(\Sigma'^{-1} R_t R_t^\top) \right\} \\
 &= -\frac{1}{2} \left\{ \log |\Sigma| - \log |\Sigma'| + \text{tr}(\mathbb{E}_\Sigma(\Sigma^{-1} R_t R_t^\top)) - \text{tr}(\mathbb{E}_{\Sigma'}(\Sigma'^{-1} R_t R_t^\top)) \right\} \\
 &= -\frac{1}{2} \left\{ \log |\Sigma| - \log |\Sigma'| + \text{tr}((\Sigma^{-1} \Sigma) - \text{tr}(\Sigma'^{-1} \Sigma)) \right\} \\
 &= -\frac{1}{2} \left\{ \log |\Sigma| - \log |\Sigma'| + d - \text{tr}(\Sigma'^{-1} \Sigma) \right\} \\
 &= K(\Sigma, \Sigma')
 \end{aligned}$$



Por lo tanto, $L_I(\tilde{\Sigma}_I, \Sigma') = N_I K(\tilde{\Sigma}_I, \Sigma')$

La razón para considerar innovaciones normales está justificada en la literatura financiera, aunque sabemos que en general las colas de ε_t son más pesadas y que según lo descrito en el marco teórico, en general difiere de una distribución normal.

Sin embargo, incluso si las innovaciones no son normales, se puede aplicar el enfoque de ajuste bajo cuasi máxima verosimilitud.

Teorema 5.2: Sea $\varepsilon_t \sim N_d(0, \mathbb{1}_d)$ independientes y existe $f(t) = \Sigma^*$ para $t \in I$. Entonces para todo $\mathfrak{z} > 0$ se verifica que:

$$\mathbb{P}_{\Sigma^*} \left(L_I(\tilde{\Sigma}_I, \Sigma^*) > \mathfrak{z} \right) = \mathbb{P}_{\Sigma^*} \left(N_I K(\tilde{\Sigma}_I, \Sigma^*) > \mathfrak{z} \right) \leq 2e^{-\mathfrak{z}}.$$

Demostración:

Si $I = I' \cup I''$, entonces $\{L_I(\tilde{\Sigma}_I, \Sigma^*) > \mathfrak{z}\} \subseteq \{L_{I'}(\tilde{\Sigma}_{I'}, \Sigma^*) > \mathfrak{z}\} \cup \{L_{I''}(\tilde{\Sigma}_{I''}, \Sigma^*) > \mathfrak{z}\}$.

$$\begin{aligned}
 \mathbb{P}_{\Sigma^*} \left(L_I(\tilde{\Sigma}_I, \Sigma^*) > \mathfrak{z} \right) &= \mathbb{P}_{\Sigma^*} \left(e^{(L_I(\tilde{\Sigma}_I, \Sigma^*) > \mathfrak{z})} \right) \\
 &\leq \mathbb{P}_{\Sigma^*} \left(e^{(L_{I'}(\tilde{\Sigma}_{I'}, \Sigma^*) > \mathfrak{z})} \right) + \mathbb{P}_{\Sigma^*} \left(e^{(L_{I''}(\tilde{\Sigma}_{I''}, \Sigma^*) > \mathfrak{z})} \right) \\
 &\leq \frac{\mathbb{E}_{\Sigma^*} \left[e^{L_{I'}(\tilde{\Sigma}_{I'}, \Sigma^*)} \right] + \mathbb{E}_{\Sigma^*} \left[e^{L_{I''}(\tilde{\Sigma}_{I''}, \Sigma^*)} \right]}{e^{\mathfrak{z}}} \\
 &\leq 2e^{-\mathfrak{z}}
 \end{aligned}$$

Por el Lema 7.3 podemos concluir que $\mathbb{E}_{\Sigma^*} \left[e^{L_{I'}(\tilde{\Sigma}_{I'}, \Sigma^*)} \right] = 1$.

Este Teorema es vital para el sustento del presente trabajo. En palabras simples nos dice que nuestro estimador al pertenecer a la familia exponencial multivariada, posee una cota superior para la probabilidad de superar el valor crítico \mathfrak{z} .

Podemos notar que en el límite esta probabilidad tiene a cero a medida que \mathfrak{z} se vuelve más grande.

Finalmente, puede ser utilizado para construir intervalos de confianza para nuestro parámetro, como se muestra en el siguiente Teorema. Con intervalo de confianza para nuestro parámetro nos referimos a definir bandas de confianza para cada elemento en la matriz de volatilidad a estimar.

Teorema 5.3: Si $\mathfrak{z}_\alpha \geq \log(2\alpha^{-1})$, entonces:

$$\xi_\alpha = \left\{ \Sigma : N_I K(\tilde{\Sigma}_I, \Sigma) \leq \mathfrak{z}_\alpha \right\}$$

es un intervalo de confianza para el parámetro Σ^* .

El teorema anterior, establece que con probabilidad alta, la divergencia de Kullback-Leibler $K(\tilde{\Sigma}_I, \Sigma^*)$ está dominada por $\frac{\mathfrak{z}}{N_I}$, para \mathfrak{z} suficientemente grande.

Teorema 5.4 Bajo las mismas condiciones del Teorema 5.2,

$$\mathbb{E}_{\Sigma^*} \left| L_I(\tilde{\Sigma}_I, \Sigma^*) \right|^r \leq \mathcal{D}(r, d)$$

donde $\mathcal{D}(r, d) = (2r\tau(r))^d$ con $\tau(r) = \int_{\mathfrak{z} \geq 0} \mathfrak{z}^{r-1} e^{-\mathfrak{z}}$



Demostración:

Sea $X = (X_1, X_2, \dots, X_d)$ un vector d-dimensional de variables aleatorias, donde cada $X_i > 0$.

$$\begin{aligned} \mathbb{E}[X^r] &= \mathbb{E}[(X_1^r, X_2^r, \dots, X_d^r)] \\ &\leq (-1)^d \int_0^{+\infty} \int_0^{+\infty} \dots \int_0^{+\infty} t_1^r t_2^r \dots t_d^r d\mathbb{P}(x_1 \geq t_1, x_2 \geq t_2, \dots, x_d \geq t_d) \\ &\leq (-1)^d \int_0^{+\infty} \int_0^{+\infty} \dots \int_0^{+\infty} t_1^r t_2^r \dots t_d^r d\mathbb{P}(x_1 \geq t_1) d\mathbb{P}(x_2 \geq t_2) \dots d\mathbb{P}(x_d \geq t_d) \\ &\leq (-1)^{d-1} \int_0^{+\infty} \int_0^{+\infty} \dots \left(\int_0^{+\infty} r t_1^{r-1} \mathbb{P}(x_1 \geq t_1) dt_1 \right) t_2^r t_3^r \dots t_d^r d\mathbb{P}(x_2 \geq t_2) \dots d\mathbb{P}(x_d \geq t_d) dt_2 \dots dt_d \\ &\leq \int_0^{+\infty} r t_1^{r-1} \mathbb{P}(x_1 \geq t_1) dt_1 \int_0^{+\infty} r t_2^{r-1} \mathbb{P}(x_2 \geq t_2) dt_2 \dots \int_0^{+\infty} r t_d^{r-1} \mathbb{P}(x_d \geq t_d) dt_d \end{aligned}$$

Teniendo en cuenta que $\int_0^{+\infty} r z^{r-1} \mathbb{P}_{\theta_1^*} \left(L_I(\tilde{\theta}_{1I}, \theta_1^*) \geq z \right) dz \leq 2r \int_0^{+\infty} z^{r-1} e^{-z} dz$.

Considerando ambas condiciones, se tiene que:

$$\mathbb{E}_{\Sigma^*} \left| L_I(\tilde{\Sigma}_I, \Sigma^*) \right|^r \leq [2r\Gamma(r)]^d$$

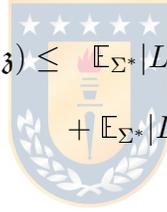
Podemos notar que la esperanza de la función de pérdida de potencia está acotada por una constante que depende de r y de la dimensión. Al hacer crecer mucho esta dimensión se pueden generar problemas computacionales al respecto, lo que es conocido en inglés como “*Curse of Dimensionality*.”

Teorema 5.5: Sea R_t v.a idénticamente distribuidas con $\mathcal{N}(0, \Sigma^*)$ e independientes con $r > 0$ e I un intervalo arbitrario, $\mathfrak{z} \geq 1$ y l otro intervalo arbitrario dado.

$$\mathbb{E}_{\Sigma^*} |L_l(\tilde{\Sigma}_l, \Sigma^*)|^r \mathbf{1}(L_I(\tilde{\Sigma}_I, \Sigma^*) \geq \mathfrak{z}) \leq 4\mathcal{D}(r, d)e^{-\mathfrak{z}} \mathfrak{z}^r$$

Demostración:

$$\begin{aligned} \mathbb{E}_{\Sigma^*} |L_l(\tilde{\Sigma}_l, \Sigma^*)|^r \mathbf{1}(L_I(\tilde{\Sigma}_I, \Sigma^*) \geq \mathfrak{z}) &\leq \mathbb{E}_{\Sigma^*} |L_I(\tilde{\Sigma}_I, \Sigma^*)|^r \mathbf{1}(L_I(\tilde{\Sigma}_I, \Sigma^*) \geq \mathfrak{z}) \\ &\quad + \mathbb{E}_{\Sigma^*} |L_l(\tilde{\Sigma}_l, \Sigma^*)|^r \mathbf{1}(L_I(\tilde{\Sigma}_I, \Sigma^*) \geq \mathfrak{z}) \end{aligned}$$



$$\mathbb{E}_{\Sigma^*} |L_I(\tilde{\Sigma}_I, \Sigma^*)|^r \mathbf{1}(L_I(\tilde{\Sigma}_I, \Sigma^*) \geq \mathfrak{z}) = \begin{cases} \mathbb{E}_{\Sigma^*} |L_I(\tilde{\Sigma}_I, \Sigma^*)|^r & \text{si } L_I(\tilde{\Sigma}_I, \Sigma^*) \geq \mathfrak{z} \\ 0 & \text{si } L_I(\tilde{\Sigma}_I, \Sigma^*) < \mathfrak{z} \end{cases}$$

Pero $\mathbb{E}_{\Sigma^*} |L_I(\tilde{\Sigma}_I, \Sigma^*)|^r \leq \mathcal{D}(r, d)$.

Acotando por la esperanza de la función indicatriz:

$$\begin{aligned} \mathbb{E}_{\Sigma^*} |L_I(\tilde{\Sigma}_I, \Sigma^*)|^r \mathbf{1}(L_I(\tilde{\Sigma}_I, \Sigma^*) \geq \mathfrak{z}) &\leq [2r\Gamma(r)]^d \mathbb{P}(L_I(\tilde{\Sigma}_I, \Sigma^*) \geq \mathfrak{z}) \\ &\leq \mathcal{D}(r, d) 2e^{-\mathfrak{z}} \end{aligned}$$

Aplicando esto a nuestra desigualdad:

$$\begin{aligned}
 \mathbb{E}_{\Sigma^*} |L_l(\tilde{\Sigma}_l, \Sigma^*)|^r \mathbf{1}(L_I(\tilde{\Sigma}_I, \Sigma^*) \geq \mathfrak{z}) &\leq \mathbb{E}_{\Sigma^*} |L_I(\tilde{\Sigma}_I, \Sigma^*)|^r \mathbf{1}(L_I(\tilde{\Sigma}_I, \Sigma^*) \geq \mathfrak{z}) \\
 &\quad + \mathbb{E}_{\Sigma^*} |L_l(\tilde{\Sigma}_l, \Sigma^*)|^r \mathbf{1}(L_I(\tilde{\Sigma}_I, \Sigma^*) \geq \mathfrak{z}) \\
 &\leq 2\mathcal{D}(r, d)e^{-\mathfrak{z}} \mathfrak{z}^r + 2\mathcal{D}(r, d)e^{-\mathfrak{z}} \mathfrak{z}^r \\
 &\leq 4\mathcal{D}(r, d)e^{-\mathfrak{z}} \mathfrak{z}^r
 \end{aligned}$$

5.2. Riesgo de estimación en el caso no paramétrico

Sea $R_t = \Sigma_t^{\frac{1}{2}} \varepsilon_t$, donde $\Sigma_t = \begin{pmatrix} f_{11}(t) & f_{12}(t) & \cdots & f_{1d}(t) \\ f_{12}(t) & f_{22}(t) & \cdots & f_{2d}(t) \\ \vdots & \vdots & \ddots & \vdots \\ f_{1d}(t) & f_{2d}(t) & \cdots & f_{dd}(t) \end{pmatrix}$, una matriz simétrica con

funciones no necesariamente constantes.

Sea $Z_\Sigma = \frac{d\mathbb{P}}{d\mathbb{P}_\Sigma}$, la razón de verosimilitud de la medida verdadera subyacente \mathbb{P} con respecto a la medida paramétrica \mathbb{P}_Σ , correspondiente al parámetro $f(\cdot) = \Sigma$.

Al restringir esta medida al intervalo I:

$$\log Z_{I, \Sigma} = \sum_{t \in I} \log \left(\frac{p(R_t, \Sigma_t)}{p(R_t, \Sigma)} \right).$$

Esta definición nos conducirá a determinar la calidad del ajuste utilizando la divergencia de Kullback- Leibler.

$$\Delta_I(\Sigma) = \sum_{t \in I} K(\Sigma_t, \Sigma)$$

El siguiente Teorema nos permite extender desde el concepto de Riesgo Parámetro a la situación No Paramétrica.

Teorema 5.6: Sea $\Sigma^* \in \Theta$ y $\Delta^* \geq 0$ tal que se tiene:

$$\mathbb{E}\Delta_I(\Sigma^*) \leq \Delta^*.$$

Entonces para todo estimador $\tilde{\Sigma}$ medible con respecto a \mathcal{F}_I ,

$$\mathbb{E} \log \left\{ 1 + \frac{|L_I(\tilde{\Sigma}, \Sigma^*)|^r}{\mathcal{R}(\tilde{\Sigma}, \Sigma^*)} \right\} \leq \Delta^* + 1,$$

con $\mathcal{R}(\tilde{\Sigma}, \Sigma^*) = \mathbb{E}_{\Sigma^*} \left| L_I(\tilde{\Sigma}, \Sigma^*) \right|^r$ función de riesgo paramétrico.

Bajo el Lema 2.6 de Spokoiny (2009, pp 1411), tenemos que:

$$\begin{aligned} \mathbb{E} \log \left\{ 1 + \frac{|L_I(\tilde{\Sigma}, \Sigma^*)|^r}{\mathcal{R}(\tilde{\Sigma}, \Sigma^*)} \right\} &\leq \mathbb{E}_{\Sigma^*} (Z_{I, \Sigma^*} \log Z_{I, \Sigma^*}) + \mathbb{E} \frac{|L_I(\tilde{\Sigma}, \Sigma^*)|^r}{\mathcal{R}(\tilde{\Sigma}, \Sigma^*)} \\ &\leq \mathbb{E}(\log Z_{I, \Sigma^*}) + 1 \\ &\leq \mathbb{E} \sum_{t \in I} \log \frac{p(R_t, \Sigma_t)}{p(R_t, \Sigma^*)} + 1 \\ &\leq \mathbb{E} \sum_{t \in I} \mathbb{E} \left\{ \log \frac{p(R_t, \Sigma_t)}{p(R_t, \Sigma)} \middle| \mathcal{F}_{t-1} \right\} + 1 \\ &\leq \mathbb{E}\Delta_I(\Sigma^*) + 1 \\ &\leq \Delta^* + 1 \end{aligned}$$

Finalmente, el riesgo de Estimación en la situación no paramétrica se obtiene mediante

la extensión del riesgo bajo hipótesis nula, es decir, riesgo paramétrico.

$$\mathbb{E} \left| L_I(\tilde{\Sigma}_{t,I}, \Sigma) \right|^r = \left| N_I \mathbb{E} K(\tilde{\Sigma}_{t,I}, \Sigma) \right|^r$$

Por lo comentado anteriormente, si $\Delta_I(\Sigma)$ es pequeño, el Riesgo No Paramétrico estará acotado.

5.3. Propiedades Teóricas

Vamos a establecer dos condiciones que deben satisfacer los intervalos l_k . La primera condición es con respecto al crecimiento de la longitud N_k de l_k , debe ser de orden exponencial.



Condiciones de Regularidad

Vamos a suponer que para algunas constantes \mathbf{u}_0, \mathbf{u} con $0 < \mathbf{u}_0 \leq \mathbf{u} < 1$ se verifica que:

$$\mathbf{u}_0 \leq \frac{N_{k-1}}{N_k} \leq \mathbf{u}$$

Para el parámetro vamos a suponer que existe una constante \mathbf{a} tal que:

$$\mathbf{a}^2 \mathbf{1}_d \preceq \Sigma^{-\frac{1}{2}} \Sigma_0 \Sigma^{-\frac{1}{2}} \preceq \mathbf{a}^{-2} \mathbf{1}_d$$

Comenzaremos el análisis enunciado dos Lemmas que nos serán útiles.

Lema 5.1 : Sean I e J intervalos de prueba y $T_{I,J} = \max_{\tau \in J} T_{I,\tau}$ el estadístico de prue-

ba, con $T_{I,\tau} = \max_{\Sigma'', \Sigma'} \{L_{I''}(\Sigma'') + L_{I'}(\Sigma')\} - \max_{\Sigma} L_I(\Sigma) = N_{I''}K(\tilde{\Sigma}_{I''}, \tilde{\Sigma}_I) + N_{I'}K(\tilde{\Sigma}_{I'}, \tilde{\Sigma}_I)$.

Bajo condiciones regularidad, se tiene que:

$$\mathbb{E}_{\Sigma^*} |N_l K(\tilde{\Sigma}_l, \Sigma^*)|^r \mathbf{1}(T_{I,J} > \mathfrak{z}) \leq 2N_J \mathcal{D}(r, d) \mathfrak{z}^r e^{-\frac{\mathfrak{z}}{2}}$$

Donde $\mathcal{D}(r, d)$ es la constante del Teorema 5.4.

Demostración:

$$\begin{aligned} & \mathbb{E}_{\Sigma^*} \left| N_l K(\tilde{\Sigma}_l, \Sigma^*) \right|^r \mathbf{1}(T_{I,\tau} > \mathfrak{z}) \\ & \leq \mathbb{E}_{\Sigma^*} \left| N_l K(\tilde{\Sigma}_l, \Sigma^*) \right|^r \left\{ \mathbf{1} \left(N_{I'} K(\tilde{\Sigma}_{I'}, \Sigma^*) > \frac{\mathfrak{z}}{2} \right) + \mathbf{1} \left(N_{I''} K(\tilde{\Sigma}_{I''}, \Sigma^*) > \frac{\mathfrak{z}}{2} \right) \right\} \\ & \leq 2\mathcal{D}(r, d) \mathfrak{z}^r e^{-\frac{\mathfrak{z}}{2}} \end{aligned}$$

Por $\mathbf{1}(T_{I,J} > \mathfrak{z}) \leq \sum_{\tau \in J} \mathbf{1}(T_{I,\tau} > \mathfrak{z})$, se obtiene la desigualdad.

Lema 5.2 : Bajo condiciones de regularidad, para cada secuencia $\Sigma_0, \Sigma_1, \dots, \Sigma_m \in \Theta$,

$$\begin{aligned} K^{\frac{1}{2}}(\Sigma_1, \Sigma_2) & \leq \mathfrak{a} \left\{ K^{\frac{1}{2}}(\Sigma_1, \Sigma_0) + K^{\frac{1}{2}}(\Sigma_2, \Sigma_0) \right\} \\ K^{\frac{1}{2}}(\Sigma_0, \Sigma_m) & \leq \mathfrak{a} \left\{ K^{\frac{1}{2}}(\Sigma_0, \Sigma_1) + \dots + K^{\frac{1}{2}}(\Sigma_{m-1}, \Sigma_m) \right\} \end{aligned}$$

La demostración de este teorema es similar a Polzehl y Spokoiny (2006).

Teorema 5.7 : Considerando las condiciones de regularidad, sea $\Sigma_t = \Sigma^*$ para todo $t \in l_k$ $k \leq K$. Entonces existe una constante $\mathfrak{a}_{0,d}$ que depende sólo de d y \mathbf{u}_0, \mathbf{u} tal que la

elección de

$$\mathfrak{z}_k = \mathfrak{a}_{0,d} \log(K) + 2 \log(\alpha^{-1}) + 2r \log\left(\frac{N_K}{N_k}\right) + 2 \log(N_k)$$

Para todo $k \leq K$.

En efecto,

Para \mathfrak{z}_k con $k > 1$. Considerando los conjuntos $\mathcal{A}_k = \{T_k \leq \mathfrak{z}_k\}$, $\mathcal{A}^k = \mathcal{A}_1 \cap \dots \cap \mathcal{A}_k$ y considerando la desigualdad $(a + b)^{2r} \leq 2^{(2r-1)+}(a^{2r} + b^{2r})$, tenemos que:

$$\mathbb{E}_{\Sigma^*} \left| N_k K(\tilde{\Sigma}_k, \hat{\Sigma}_l) \right|^r \leq 2^{(2r-1)+} \mathfrak{a}^{2r} \mathcal{D}(r, d) \sum_{l=1}^k \left[1 + \left(\frac{N_k}{N_l} \right)^r \right] 4N_l \mathfrak{z}_l^r e^{-\frac{\mathfrak{z}_l}{2}}$$



$$\begin{aligned} \left(K^{\frac{1}{2}}(\tilde{\Sigma}_k, \tilde{\Sigma}_l) \right)^{2r} &\leq \left(\mathfrak{a} K^{\frac{1}{2}}(\tilde{\Sigma}_k, \Sigma^*) + \mathfrak{a} K^{\frac{1}{2}}(\tilde{\Sigma}_l, \Sigma^*) \right)^{2r} \\ &\leq 2^{(2r-1)+} \left[\mathfrak{a}^{2r} K^r(\tilde{\Sigma}_k, \Sigma^*) + \mathfrak{a}^{2r} K^r(\tilde{\Sigma}_l, \Sigma^*) \right] \end{aligned}$$

Aplicando esperanza con respecto a Σ^* y restringiendo al conjunto $\mathfrak{B}_l = \{T_1 \leq \mathfrak{z}_1, \dots, T_{l-1} \leq \mathfrak{z}_{l-1}, T_l > \mathfrak{z}_l\}$ tenemos:

$$\begin{aligned} \mathbb{E}_{\Sigma^*} \left(K^r(\tilde{\Sigma}_k, \tilde{\Sigma}_l) \right) \mathbf{1}(\mathfrak{B}_l) &\leq 2^{(2r-1)+} \left[\mathfrak{a}^{2r} \mathbb{E}_{\Sigma^*} K^r(\tilde{\Sigma}_k, \Sigma^*) + \mathfrak{a}^{2r} \mathbb{E}_{\Sigma^*} K^r(\tilde{\Sigma}_l, \Sigma^*) \right] \mathbf{1}(\mathfrak{B}_l) \\ &\leq 2^{(2r-1)+} \left[\mathfrak{a}^{2r} \mathbb{E}_{\Sigma^*} K^r(\tilde{\Sigma}_k, \Sigma^*) + \mathfrak{a}^{2r} \mathbb{E}_{\Sigma^*} K^r(\tilde{\Sigma}_l, \Sigma^*) \right] \mathbf{1}(T_{l+1}, \mathcal{J}_l > \mathfrak{z}_l) \end{aligned}$$

Por el lemma anterior:

$$\mathbb{E}_{\Sigma^*} K^r(\tilde{\Sigma}_k, \Sigma^*) \mathbf{1}(T_{l+1}, \mathcal{J}_l > \mathfrak{z}_l) \leq \frac{2N_{\mathcal{J}_l} \mathcal{D}(r, d) \mathfrak{z}_l^r e^{-\frac{\mathfrak{z}_l}{2}}}{N_k^r}$$

Análogamente:

$$\mathbb{E}_{\Sigma^*} K^r(\tilde{\Sigma}_l, \Sigma^*) \mathbf{1}(T_{l+1}, \mathcal{J}_l > \mathfrak{z}_l) \leq \frac{2N_{\mathcal{J}_l} \mathcal{D}(r, d) \mathfrak{z}_l^r e^{-\frac{\mathfrak{z}_l}{2}}}{N_l^r}$$

Pero $N_{\mathcal{J}_l} = N_l$, reemplazando en (*) tenemos una cota superior para la esperanza.

$$\mathbb{E}_{\Sigma^*} \left(K^r(\tilde{\Sigma}_k, \tilde{\Sigma}_l) \right) \mathbf{1}(\mathfrak{B}_l) \leq 2^{(2r-1)+} \mathfrak{a}^{2r} 2N_l \mathcal{D}(r, d) \mathfrak{z}_l^r e^{-\frac{\mathfrak{z}_l}{2}} \left[\frac{1}{N_k^r} + \frac{1}{N_l^r} \right]$$

Para $k \leq K$, tenemos:

$$\begin{aligned} \mathbb{E}_{\Sigma^*} \left| N_k K(\tilde{\Sigma}_k, \hat{\Sigma}_k) \right|^r &\leq N_k^r \mathbb{E}_{\Sigma^*} \sum_{l=1}^k K^r(\tilde{\Sigma}_k, \tilde{\Sigma}_l) \mathbf{1}(\mathfrak{B}_l) \\ &\leq 2^{(2r-1)+} \mathfrak{a}^{2r} \mathcal{D}(r, d) \sum_{l=1}^k \left[1 + \left(\frac{N_k}{N_l} \right)^r \right] 4N_l \mathfrak{z}_l^r e^{-\frac{\mathfrak{z}_l}{2}} \end{aligned}$$

Tomando $\mathfrak{z}_k = \mathfrak{a}_{0,d} \log(K) + 2 \log\left(\frac{N_k}{\alpha}\right) + 2r \log\left(\frac{N_K}{N_k}\right)$ para $k \leq K$, logramos que la desigualdad anterior este acotada por una constante.

En particular podemos tomar $\mathfrak{z}_l = \mathfrak{a}_{0,d} \log(k) + 2 \log\left(\frac{N_l}{\alpha}\right) + 2r \log\left(\frac{N_k}{N_l}\right)$. De manera asintótica

podemos no considerar el valor $\mathfrak{z}l^r$, ya que crece más lento que la exponencial.

$$\begin{aligned}
 \mathbb{E}_{\Sigma^*} \left| N_k K(\tilde{\Sigma}_k, \hat{\Sigma}_k) \right|^r &\leq 2^{(2r-1)+} \mathfrak{a}^{2r} \mathcal{D}(r, d) \sum_{l=1}^k \left[1 + \left(\frac{N_k}{N_l} \right)^r \right] 4N_l k^{\frac{-\mathfrak{a}_{0,d}}{2}} \left(\frac{N_k}{N_l} \right)^{-r} \left(\frac{N_l}{\alpha} \right)^{-1} \\
 &\leq 2^{(2r-1)+} \mathfrak{a}^{2r} \mathcal{D}(r, d) \alpha k^{\frac{-\mathfrak{a}_{0,d}}{2}} \sum_{l=1}^k \left[\left(\frac{N_l}{N_k} \right)^r + 1 \right] \\
 &\leq 2^{(2r-1)+} \mathfrak{a}^{2r} \mathcal{D}(r, d) \alpha k^{\frac{-\mathfrak{a}_{0,d}}{2}} \sum_{l=1}^k [\mathbf{u}^{(k-l)} + 1]^r \\
 &\leq 2^{(2r-1)+} \mathfrak{a}^{2r} \mathcal{D}(r, d) \alpha k^{\frac{-\mathfrak{a}_{0,d}}{2}} \mathbf{u}^{-r(k-1)} \sum_{l=1}^k [\mathbf{u}^{(k-1)} + \mathbf{u}^{(l-1)}]^r \\
 &\leq 2^{(2r-1)+} \mathfrak{a}^{2r} \mathcal{D}(r, d) \alpha k^{\frac{-\mathfrak{a}_{0,d}}{2}} \mathbf{u}^{-r(k-1)} 2^r \sum_{l=1}^{+\infty} [\mathbf{u}^{r(l-1)}]
 \end{aligned}$$



La serie considerada en la desigualdad anterior es convergente al ser una serie geométrica con razón menor que 1. Por lo tanto podemos mostrar que la esperanza está acotada por una constante. Estos valores críticos son útiles para la justificación del procedimiento y es importante que sean finitos. No obstante para valores más ajustados en la práctica no serán utilizados ya que en general se utiliza simulación Monte Carlo para su calibración.

5.4. Comportamiento bajo la condición de Pequeño Sesgo de Modelado

En esta sección extendemos lo anterior para el caso donde la suposición paramétrica no se cumple exactamente, pero su desviación del modelo paramétrico es pequeña.

Consideremos el paso k del algoritmo, entonces:

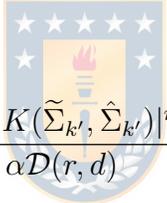
$$\Delta_k(\Sigma) = \Delta_{l_{k+1}}(\Sigma) = \sum_{l_{k+1}} K(\Sigma, \Sigma_t)$$

Donde l_{k+1} es el intervalo utilizado para testear la hipótesis de punto de cambio dentro de \mathcal{J}_k . Podemos notar que el sesgo de modelación $\Delta_k(\Sigma)$ incrementa junto con k .

Si existe algún número k' tal que $\Delta_{k'}(\Sigma)$ es pequeño, entonces es pequeño para todo $k \leq k'$.

Sea $\hat{\Sigma}_{k'}$ el estimador obtenido luego de los primeros k' pasos del procedimiento, entonces se tiene el siguiente resultado:

Teorema 5.9: Bajo condiciones de regularidad. Si Σ^* y k' tal que $\mathbb{E}_{\Delta_{k'}}(\Sigma^*) \leq \Delta^*$, para algún $\Delta^* \geq 0$, entonces:



$$\mathbb{E} \log \left\{ 1 + \frac{|N_{k'} K(\tilde{\Sigma}_{k'}, \hat{\Sigma}_{k'})|^r}{\alpha \mathcal{D}(r, d)} \right\} \leq 1 + \Delta^*$$

$$\mathbb{E} \log \left\{ 1 + \frac{|N_{k'} K(\tilde{\Sigma}_{k'}, \Sigma^*)|^r}{\mathcal{D}(r, d)} \right\} \leq 1 + \Delta^*$$

Por el lemma 2.6 Spokoiny(2009), se sabe que si ζ es una variable aleatoria tal que $\mathbb{E}_0 \zeta < \infty$ entonces:

$$\mathbb{E} \log(1 + \zeta) \leq \Delta + \mathbb{E}_0 \zeta$$

Tomando $\zeta = \frac{|N_{k'}K(\tilde{\Sigma}_{k'}, \hat{\Sigma}_{k'})|^r}{\alpha\mathcal{D}(r,d)}$ y teniendo en consideración las condiciones de propagación,

$$\begin{aligned} \mathbb{E} \log \left\{ 1 + \frac{|N_{k'}K(\tilde{\Sigma}_{k'}, \hat{\Sigma}_{k'})|^r}{\alpha\mathcal{D}(r,d)} \right\} &\leq \Delta^* + \mathbb{E} \frac{|N_{k'}K(\tilde{\Sigma}_{k^*}, \hat{\Sigma}_{k'})|^r}{\alpha\mathcal{D}(r,d)} \\ &\leq \Delta^* + \frac{\mathbb{E} |N_{k'}K(\tilde{\Sigma}_{k'}, \hat{\Sigma}_{k'})|^r}{\alpha\mathcal{D}(r,d)} \\ &\leq \Delta^* + \frac{\alpha\mathcal{D}(r,d)}{\alpha\mathcal{D}(r,d)} \\ &\leq \Delta^* + 1 \end{aligned}$$

Análogamente, considerando $\zeta = \frac{|N_{k'}K(\tilde{\Sigma}_{k'}, \Sigma^*)|^r}{\mathcal{D}(r,d)}$ y el Teorema 5.6, se obtiene el resultado.

Este Teorema nos permite acotar el Riesgo de Estimación cuando el modelo se aleja del supuesto paramétrico generando un sesgo aceptable en la modelación.



5.5. Estabilidad después de verificada las Condiciones de Propagación

Si el sesgo del modelado es pequeño, por ende acotado, es decir $\Delta_{k^*}(\Sigma^*) \leq \Delta^*$, tenemos que el procedimiento funciona bastante bien. Al generalizar al caso Σ_t con funciones no necesariamente constantes y al tener en cuenta que el riesgo de estimación crece de manera proporcional a la cantidad de retornos agregados, se debe asegurar que el estimador final $\hat{\Sigma}$ no varía demasiado en pasos posteriores del procedimiento, teniendo en cuenta la estimación agregada $\hat{\Sigma}_k$ luego de cumplida las condiciones de propagación, cuando la divergencia de Kullback-Leibler se hace grande.

Teorema 5.10: Para algún $k \leq K$ y l_k el intervalo aceptado por el procedimiento adaptativo y $\widehat{\Sigma}_k = \widetilde{\Sigma}_k$, entonces,

$$N_k K(\widehat{\Sigma}_k, \widehat{\Sigma}_{k+1}) \leq \mathfrak{z}_k$$

Más aún bajo las condiciones de regularidad, para todo k' tal que $k \leq k' \leq K$,

$$N_k K(\widehat{\Sigma}_k, \widehat{\Sigma}_{k'}) \leq \mathfrak{a}^2 C_u^2 \overline{\mathfrak{z}_k}$$

Donde $C_u = (\mathfrak{u}^{-\frac{1}{2}} - 1)^{-1}$ y $\overline{\mathfrak{z}_k} = \max_{l \geq k} \mathfrak{z}_l$.

Demostración:



Si l_{k+1} es rechazado, la desigualdad se obtiene trivialmente. Si l_{k+1} es aceptado, entonces $\widehat{\Sigma}_k = \widetilde{\Sigma}_k$ y $\widehat{\Sigma}_{k+1} = \widetilde{\Sigma}_{k+1}$.

Decir que l_k sea aceptado significa que $T_k = T_{l_{k+1}, \mathcal{J}_k} \leq \mathfrak{z}_k$, en particular $T_{l_{k+1}, \tau} \leq \mathfrak{z}_k$ con $\tau = t^\diamond - N_k$.

Por otro lado,

$$T_{l_{k+1}, \tau} = N_{\mathcal{J}_{k+1}} K(\widetilde{\Sigma}_{\mathcal{J}_{k+1}}, \widetilde{\Sigma}_{l_{k+1}}) + N_{l_k} K(\widetilde{\Sigma}_{l_k}, \widetilde{\Sigma}_{l_{k+1}}) \leq \mathfrak{z}_k,$$

pero, $N_{l_k} K(\widetilde{\Sigma}_{l_k}, \widetilde{\Sigma}_{l_{k+1}}) = N_k K(\widetilde{\Sigma}_k, \widetilde{\Sigma}_{k+1})$.

De esta desigualdad podemos concluir que: $N_k K(\widetilde{\Sigma}_k, \widetilde{\Sigma}_{k+1}) \leq \mathfrak{z}_k$

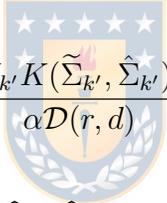
Por un lema anterior, sabemos que:

$$\begin{aligned}
 K^{\frac{1}{2}}(\widehat{\Sigma}_k, \widehat{\Sigma}_{k'}) &\leq \mathfrak{a} \sum_{j=k}^{k'-1} K^{\frac{1}{2}}(\widehat{\Sigma}_j, \widehat{\Sigma}_{j+1}) \\
 &\leq \mathfrak{a} \sum_{j=k}^{k'-1} \left(\frac{\mathfrak{z}_k}{N_k} \right)^{\frac{1}{2}} \\
 &\leq \mathfrak{a} \left(\frac{\bar{\mathfrak{z}}_k}{N_k} \right)^{\frac{1}{2}} \sum_{j=k}^{k'-1} (u^{j-k})^{\frac{1}{2}}
 \end{aligned}$$

Acotando por la serie geométrica, tenemos lo pedido.

Teorema 5.11: Sea $\mathbb{E}\Delta_k(\Sigma) \leq \Delta^*$ para algún $\Sigma \in \Theta$ y $k \leq k'$. Entonces:

$$\mathbb{E} \log \left\{ 1 + \frac{|N_{k'} K(\widetilde{\Sigma}_{k'}, \widehat{\Sigma}_{k'})|^r}{\alpha \mathcal{D}(r, d)} \right\} \leq 1 + \Delta^*$$



$$N_k K(\widehat{\Sigma}_{k'}, \widehat{\Sigma}) \leq \mathfrak{a}^2 C_u^2 \bar{\mathfrak{z}}_{k'}$$

Este Teorema describe la propiedad oráculo de nuestra estimación y justifica el uso del procedimiento descrito. Nos dice que no existe una diferencia significativa entre el estimador seleccionado $\widehat{\Sigma}$ y el estimador oráculo $\widetilde{\Sigma}_{k^*}$.

Teorema 5.12: Bajo condiciones de regularidad, para algunos k^* , Σ^* y Δ^* , tal que $\mathbb{E}_{\Delta_{k^*}}(\Sigma^*) \leq \Delta^*$ entonces,

$$\mathbb{E} \log \left(1 + N_{k^*}^{\frac{1}{2}} \frac{K^{\frac{1}{2}}(\widehat{\Sigma}, \Sigma^*)}{\alpha \mathcal{D}(\frac{1}{2}, d)} \right) \leq \log \left(1 + \frac{C_u \sqrt{\bar{\mathfrak{z}}_{k^*}}}{\mathcal{D}(\frac{1}{2}, d)} \right) + \Delta^* + \alpha + 1.$$

Demostración:

$$\mathbf{a}^{-1} \left| N_{k^*} K(\widehat{\Sigma}^*, \Sigma^*) \right|^{\frac{1}{2}} \leq \left| N_{k^*} K(\widehat{\Sigma}_{k^*}^*, \widehat{\Sigma}^*) \right|^{\frac{1}{2}} + \left| N_{k^*} K(\widetilde{\Sigma}_{k^*}^*, \widehat{\Sigma}_{k^*}^*) \right|^{\frac{1}{2}} + \left| N_{k^*} K(\widetilde{\Sigma}_{k^*}^*, \Sigma^*) \right|^{\frac{1}{2}}$$

Por el teorema 5.11, $\left| N_{k^*} K(\widehat{\Sigma}_{k^*}^*, \widehat{\Sigma}^*) \right|^{\frac{1}{2}} \leq \mathbf{a} C_u \sqrt{\widehat{\mathfrak{z}}_{k^*}} \leq C_u \sqrt{\widehat{\mathfrak{z}}_{k^*}}, \quad (\mathbf{a} \leq 1),$

$$\mathbf{a}^{-1} \left| N_{k^*} K(\widehat{\Sigma}^*, \Sigma^*) \right|^{\frac{1}{2}} \leq C_u \sqrt{\widehat{\mathfrak{z}}_{k^*}} + \left| N_{k^*} K(\widetilde{\Sigma}_{k^*}^*, \widehat{\Sigma}_{k^*}^*) \right|^{\frac{1}{2}} + \left| N_{k^*} K(\widetilde{\Sigma}_{k^*}^*, \Sigma^*) \right|^{\frac{1}{2}}$$

Por lo tanto,

$$\log \left(1 + \frac{|N_{k^*} K(\widehat{\Sigma}^*, \Sigma^*)|^{\frac{1}{2}}}{\mathbf{a} \mathcal{D}(\frac{1}{2}, d)} \right) \leq \log \left(1 + \frac{C_u \sqrt{\widehat{\mathfrak{z}}_{k^*}}}{\mathcal{D}(\frac{1}{2}, d)} + \frac{|N_{k^*} K(\widetilde{\Sigma}_{k^*}^*, \widehat{\Sigma}_{k^*}^*)|^{\frac{1}{2}}}{\mathcal{D}(\frac{1}{2}, d)} + \frac{|N_{k^*} K(\widetilde{\Sigma}_{k^*}^*, \Sigma^*)|^{\frac{1}{2}}}{\mathcal{D}(\frac{1}{2}, d)} \right).$$

Considerando la desigualdad elemental, $\log(1 + a + b) \leq \log(1 + a) + \log(1 + b)$ tenemos:

$$\begin{aligned} & \log \left(1 + \frac{|N_{k^*} K(\widehat{\Sigma}^*, \Sigma^*)|^{\frac{1}{2}}}{\mathbf{a} \mathcal{D}(\frac{1}{2}, d)} \right) \\ & \leq \log \left(1 + \frac{C_u \sqrt{\widehat{\mathfrak{z}}_{k^*}}}{\mathcal{D}(\frac{1}{2}, d)} \right) + \log \left(1 + \frac{|N_{k^*} K(\widetilde{\Sigma}_{k^*}^*, \widehat{\Sigma}_{k^*}^*)|^{\frac{1}{2}}}{\mathcal{D}(\frac{1}{2}, d)} + \frac{|N_{k^*} K(\widetilde{\Sigma}_{k^*}^*, \Sigma^*)|^{\frac{1}{2}}}{\mathcal{D}(\frac{1}{2}, d)} \right) \end{aligned}$$

Aplicando la esperanza a la desigualdad anterior:

$$\begin{aligned} & \mathbb{E} \log \left(1 + \frac{\mathbf{a}^{-1} N_{k^*} K(\widehat{\Sigma}^*, \Sigma^*)}{\mathcal{D}(\frac{1}{2}, d)} \right) \\ & \leq \log \left(1 + \frac{C_u \sqrt{\widehat{\mathfrak{z}}_{k^*}}}{\mathcal{D}(\frac{1}{2}, d)} \right) + \mathbb{E} \log \left(1 + \frac{|N_{k^*} K(\widetilde{\Sigma}_{k^*}^*, \widehat{\Sigma}_{k^*}^*)|^{\frac{1}{2}}}{\mathcal{D}(\frac{1}{2}, d)} + \frac{|N_{k^*} K(\widetilde{\Sigma}_{k^*}^*, \Sigma^*)|^{\frac{1}{2}}}{\mathcal{D}(\frac{1}{2}, d)} \right) \end{aligned}$$

Por el Teorema 5.4 y el Lema 2.6 de Spokoiny (2009), tenemos:

$$\begin{aligned}
 \mathbb{E} \log \left(1 + \frac{|N_{k^*} K(\tilde{\Sigma}_{k^*}^*, \hat{\Sigma}_{k^*}^*)|^{\frac{1}{2}}}{\mathcal{D}(\frac{1}{2}, d)} + \frac{|N_{k^*} K(\tilde{\Sigma}_{k^*}^*, \Sigma^*)|^{\frac{1}{2}}}{\mathcal{D}(\frac{1}{2}, d)} \right) &\leq \Delta^* + \mathbb{E}_{\Sigma^*} \left(1 + \frac{|N_{k^*} K(\tilde{\Sigma}_{k^*}^*, \hat{\Sigma}_{k^*}^*)|^{\frac{1}{2}}}{\mathcal{D}(\frac{1}{2}, d)} \right) \\
 &\quad + \mathbb{E}_{\Sigma^*} \left(\frac{|N_{k^*} K(\tilde{\Sigma}_{k^*}^*, \Sigma^*)|^{\frac{1}{2}}}{\mathcal{D}(\frac{1}{2}, d)} \right) \\
 &\leq \Delta^* + \alpha \frac{\mathcal{D}(\frac{1}{2}, d)}{\mathcal{D}(\frac{1}{2}, d)} + \frac{\mathcal{D}(\frac{1}{2}, d)}{\mathcal{D}(\frac{1}{2}, d)} \\
 &\leq \Delta^* + \alpha + 1
 \end{aligned}$$

Por lo tanto, $\mathbb{E} \log \left(1 + \frac{|N_{k^*} K(\hat{\Sigma}^*, \Sigma^*)|^{\frac{1}{2}}}{\alpha \mathcal{D}(\frac{1}{2}, d)} \right) \leq \log \left(1 + \frac{C_u \sqrt{\delta_{k^*}}}{\mathcal{D}(\frac{1}{2}, d)} \right) + \Delta^* + \alpha + 1.$

5.6. Modelo de Regimen Cambiante y Calidad en la Detección de Puntos de Cambio

En nuestro estudio buscamos el mayor intervalo de manera adaptativa donde Σ es una matriz simétrica constante. Sin embargo, se acepta facilmente que la volatilidad puede estimarse por un Modelo de Regimen Cambiante, es decir a través de un Momento de Markov o más conocido como Tiempo de Parada, ya que este proceso estocástico la probabilidad de un evento depende inmediatamente de los eventos anteriores, por ende no hay independencia en el proceso y dado $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F})_{t \in T}, \mathbb{P})$ un espacio de Probabilidad, y ω una variable aleatoria, que será un tiempo de parada si se cumple lo siguiente: $\omega : \Omega \rightarrow T, \quad \omega(\beta) \leq t \in \mathcal{F}_t, \quad t \in T.$

Por ende, se supone que existe una secuencia ν_1, ν_2, \dots donde Σ_j^* es medible con respecto al campo generado por las observaciones contenidas en dicho Momento de Markov, es decir, $\Sigma_t^2 = f(t) = \Sigma_t^*$ con $\nu_j \leq t < \nu_{j+1}$. En este contexto, el procedimiento propuesto nos permite encontrar con alta probabilidad la ubicación de los puntos de cambio, como si se conociera

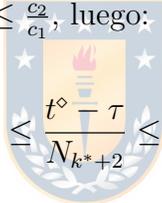
a priori.

Sea ν un punto de cambio y t^\diamond un punto posterior. Tomando intervalos de la forma $l_k = [t^\diamond - N_k, t^\diamond[$, se debe escoger el intervalo más grande.

Sea l_{k^*} un intervalo que no contenga a dicho punto de cambio ν y l_{k^*+1} el primer intervalo rechazado por contener a ν . Supongamos que existe un único punto de cambio en $\mathcal{J}_{k^*+1} = [t^\diamond - N_{k^*+1}, t^\diamond - N_{k^*}[$ y en l_{k^*+2} .

Por otro lado, sea τ un punto, tal que $I' = [t^\diamond - N_{k^*+2}, \tau]$, y para $t \in I'$, $f(t) = \Sigma'$, a su vez, $I'' = [\tau, t^\diamond[$ con $t \in I''$, $f(t) = \Sigma''$.

Sea $c_1 = \frac{N_{k^*}}{N_{k^*+2}}$ y $c_2 = \frac{N_{k^*+1}}{N_{k^*+2}}$, constantes menores a 1. Considerando las condiciones de regularidad, se tiene que $\mathbf{u}_0^2 \leq c_1$, $\mathbf{u}^{-1} \leq \frac{c_2}{c_1}$, luego:

$$c_1 \leq \frac{t^\diamond - \tau}{N_{k^*+2}} \leq c_2$$


Vamos a definir la siguiente medida respecto a la distancia entre Σ' y Σ'' ,

$$d^2(\Sigma', \Sigma'') = \inf_{\Sigma \geq 0} \inf_{c \in [c_1, c_2]} \{(1-c)K(\Sigma', \Sigma) + cK(\Sigma'', \Sigma)\}.$$

Si $d^2(\Sigma', \Sigma'')$ es grande entonces T_{k^*+1} también lo será, por ende podemos rechazar l_{k^*+1} .

Teorema 5.13: Si para algún $\mathfrak{z} > 0$, $N_{k^*+2}d^2(\Sigma', \Sigma'') \geq 2\mathfrak{a}^2(\mathfrak{z}_{k^*+1} + z)m$ entonces:

$$\mathbb{P}(l_{k^*+1} \text{ no sea rechazado}) \leq 4e^{-\mathfrak{z}}.$$

Demostración:

Sea $\mathfrak{z} > 0$, consideremos $\mathcal{A}(\mathfrak{z}) = \mathbf{1} \left\{ N_{I'} K(\tilde{\Sigma}_{I'}, \Sigma') \leq \mathfrak{z}, N_{I''} K(\tilde{\Sigma}_{I''}, \Sigma'') \leq \mathfrak{z} \right\}$, por el teorema (5.2), $\mathbb{P}(\mathcal{A}(\mathfrak{z})) \geq 1 - 2e^{-\mathfrak{z}} - 2e^{-\mathfrak{z}} = 1 - 4e^{-\mathfrak{z}}$.

Por otro lado, recordemos que $T_{I,\tau} = N_{I''} K(\tilde{\Sigma}_{I''}, \tilde{\Sigma}_I) + N_{I'} K(\tilde{\Sigma}_{I'}, \tilde{\Sigma}_I)$.

$$\begin{aligned} K^{\frac{1}{2}}(\Sigma', \tilde{\Sigma}_I) &\leq \mathfrak{a} K^{\frac{1}{2}}(\tilde{\Sigma}_{I'}, \Sigma') + \mathfrak{a} K^{\frac{1}{2}}(\tilde{\Sigma}_{I'}, \tilde{\Sigma}_I) \\ &\leq \mathfrak{a} \left(\frac{\mathfrak{z}}{N_{I'}} \right)^{\frac{1}{2}} + \mathfrak{a} K^{\frac{1}{2}}(\tilde{\Sigma}_{I'}, \tilde{\Sigma}_I) \end{aligned}$$

Luego, $K(\tilde{\Sigma}_{I'}, \tilde{\Sigma}_I) \geq (2\mathfrak{a}^2)^{-1} K^{\frac{1}{2}}(\Sigma', \tilde{\Sigma}_I) - \frac{\mathfrak{z}}{N_{I'}}$.

De manera análoga $K(\tilde{\Sigma}_{I''}, \tilde{\Sigma}_I) \geq (2\mathfrak{a}^2)^{-1} K^{\frac{1}{2}}(\Sigma'', \tilde{\Sigma}_I) - \frac{\mathfrak{z}}{N_{I''}}$. Aplicando esto al estadístico de prueba $T_{I,\nu}$,

$$\begin{aligned} T_{I,\nu} &= N_{I''} K(\tilde{\Sigma}_{I''}, \tilde{\Sigma}_I) + N_{I'} K(\tilde{\Sigma}_{I'}, \tilde{\Sigma}_I) \\ &\geq (2\mathfrak{a}^2)^{-1} \left\{ N_{I''} K(\Sigma'', \tilde{\Sigma}_I) + N_{I'} K(\Sigma', \tilde{\Sigma}_I) \right\} - \mathfrak{z} \\ &\geq (2\mathfrak{a}^2)^{-1} N_{k^*+2} \left\{ c K(\Sigma', \tilde{\Sigma}_I) + (1-c) K(\Sigma'', \tilde{\Sigma}_I) \right\} - \mathfrak{z} \end{aligned}$$

Tomando $c = \frac{N_{I'}}{N_{k^*+2}}$, tenemos que:

$$T_{I,\nu} \geq (2\mathfrak{a}^2)^{-1} N_{k^*+2} d^2(\Sigma', \Sigma'') - \mathfrak{z}$$

Bajo la hipótesis del teorema, obtenemos:

$$T_{I,\nu} \geq (2\mathfrak{a}^2)^{-1} 2\mathfrak{a}^2 (\mathfrak{z}^{k^*+1} + \mathfrak{z}) * m - \mathfrak{z}$$

Si $m = 1$, obtenemos la conclusión.

Capítulo 6

Aplicación a la Estimación y Predicción en los Tipos de Cambio de Divisas



6.1. Análisis Descriptivo de Datos

Las observaciones estudiadas se obtuvieron de la página web de la Reserva Federal EE.UU. y corresponden a 2583 observaciones de tipo de cambio del dólar estadounidense a 9 monedas distintas, desde el 1 de Enero de 1990 a el 7 de Abril de 2000. A saber, Dólar Australiano (AUD), Libra Esterlina (GBP), Dólar Canadiense (CAD), Corona Danesa (DKK), Yen Japonés (JPY), Corona Noruega (NOK), Dólar Neozelandés (NZD), Corona Sueca (SEK) y Franco Suizo (CHF).

En las siguientes tablas están contenidos los resúmenes de las variables en cuánto a posición, dispersión y forma de la distribución de los datos en las series de tiempo de Tipo

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

de Cambio.

Moneda	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Mod.
AUD	0.5550	0.6710	0.7385	0.7222	0.7701	0.8351	0.7405
GBP	1.417	1.554	1.620	1.639	1.683	2.003	1.6135 1.6210 1.6230
CAD	1.120	1.198	1.360	1.330	1.404	1.577	1.3745
DKK	5.344	5.917	6.382	6.351	6.745	7.819	6.455 6.460 6.575 6.600 6.940
JPY	81.12	105.96	118.08	118.50	129.88	159.90	123.9 124.9
NOK	5.510	6.413	6.773	6.874	7.420	8.490	6.505
NZD	0.4810	0.5405	0.5785	0.5893	0.6410	0.7148	0.534
SEK	5.088	6.217	7.350	7.143	7.915	8.851	7.915 7.975
CHF	1.117	1.292	1.430	1.397	1.490	1.681	1.4895

Cuadro 6.1: Resumen Estadísticos de Tendencia Central y Posición Datos Tipos de Cambio.

Moneda	Sd.	Var.	Skew.	Curtosis
AUD	0.0034	0.0586	-0.4944	-0.8464
GBP	0.0144	0.1198	1.0624	0.8004
CAD	0.1180	0.0139	-0.1999	-1.0371
DKK	0.5173	0.2676	0.1526	-0.5206
JPY	16.0869	258.7887	0.2159	-0.4202
NOK	0.6286	0.3952	0.2846	-0.7765
NZD	0.0602	0.0036	0.4260	-0.9974
SEK	0.9317	0.8680	-0.3157	-1.1354
CHF	0.1237	0.0153	-0.4460	-0.6872

Cuadro 6.2: Resumen Estadísticos de Dispersión y Forma Datos de Tipo de Cambio.

Observando las tablas, podemos notar que algunas de las series no son unimodales, el caso más extremo es el de la corona danesa que posee cinco modas, esto es un claro indicio de la no normalidad de los datos.

En cuanto a la dispersión, el Yen Japonés es el que tiene la varianza más alta, por ende es el tipo de cambio más volátil entre los estudiados. Todas las series son asimétricas. En cuanto a la curtosis todas a excepción de la Libra Esterlina son platicúrticas.

Las siguientes tablas resumen la información de las series de retornos logarítmicos, en

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

adelante el estudio se centró en dichas series. Se puede notar que todas las series son unimodales, que son todas leptocúrticas además de asimétricas pero esta vez con valores más cercanos a cero.

Moneda	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Mod.
RETAUD	-0.0444865	-0.0030509	0.0001366	-0.0001041	0.0029981	0.0482481	0
RETGBP	-3.286e-02	-2.940e-03	6.459e-05	-6.790e-06	3.160e-03	2.889e-02	0
RETCAD	-1.621e-02	-1.617e-03	7.068e-05	8.819e-05	1.720e-03	1.652e-02	0
RETDKK	-3.286e-02	-3.644e-03	1.597e-04	6.099e-05	3.573e-03	3.467e-02	0
RETJPY	-0.0563021	-0.0037492	0.0000000	-0.0001271	0.0039408	0.0323990	0
RETNOK	-4.492e-02	-3.405e-03	7.785e-05	9.493e-05	3.457e-03	5.969e-02	0
RETNZD	-4.064e-02	-2.692e-03	0.000e+00	-6.581e-05	2.614e-03	4.144e-02	0
RETSEK	-0.0459534	-0.0035632	0.0000000	0.0001267	0.0037883	0.0645385	0
RETCHEF	-3.899e-02	-4.185e-03	2.633e-04	1.481e-05	4.346e-03	3.112e-02	0

Cuadro 6.3: Resumen Estadísticos de Tendencia Central y Posición para las Series de Retornos.



Moneda	Sd.	Var.	Skew.	Curtosis
RETAUD	0.005649095	3.191227e-05	-0.1878116	5.854727
RETGBP	0.005941711	3.530393e-05	-0.2793477	2.79293
RETCAD	0.002992616	8.955753e-06	0.04277628	2.49934
RETDKK	0.00648215	4.201827e-05	-0.03752007	1.967052
RETJPY	0.007407423	5.486992e-05	-0.5856599	4.366011
RETNOK	0.006520444	4.251619e-05	0.313869	5.638011
RETNZD	0.005436945	2.956037e-05	0.06976292	6.991142
RETSEK	0.006793834	4.615617e-05	0.372656	6.6594
RETCHEF	0.007350001	5.402252e-05	-0.1868456	1.526278

Cuadro 6.4: Resumen Estadísticos de Dispersión y Forma para las Series de Retornos.

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

A continuación observamos las series de tiempo de los retornos logarítmicos de los tipos de cambio. Podemos ver gráficamente que las series no son estacionarias en varianza y que se presentan agrupaciones de volatilidad, pero esto debe ser probado.

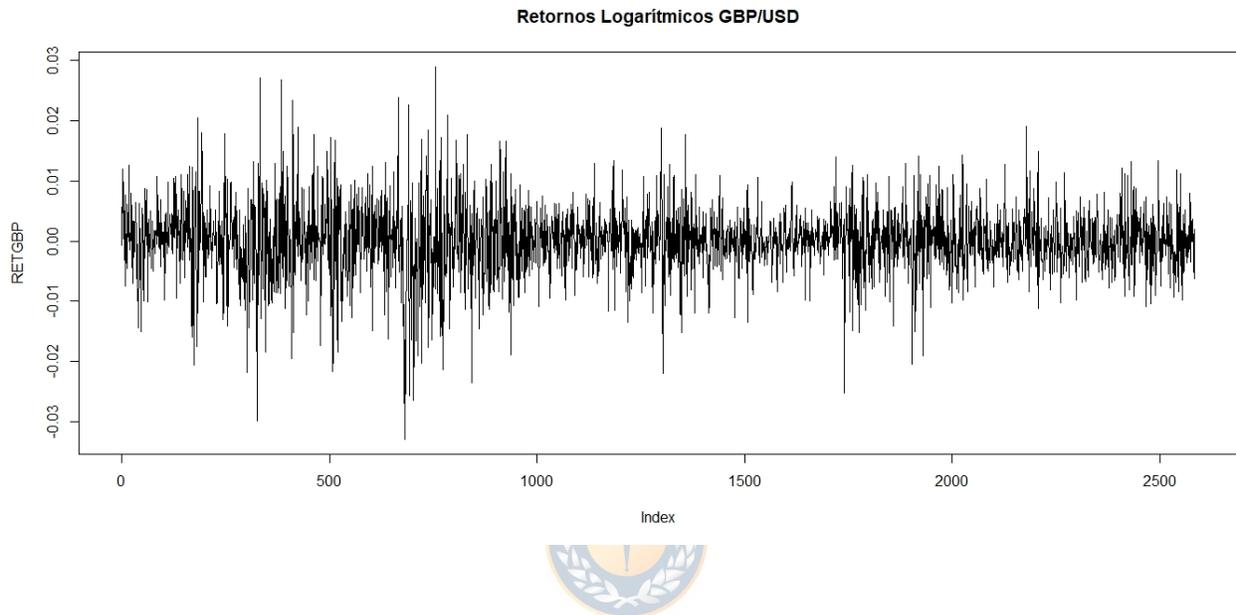


Figura 6.1: Retornos Logarítmicos Libra Esterlina.

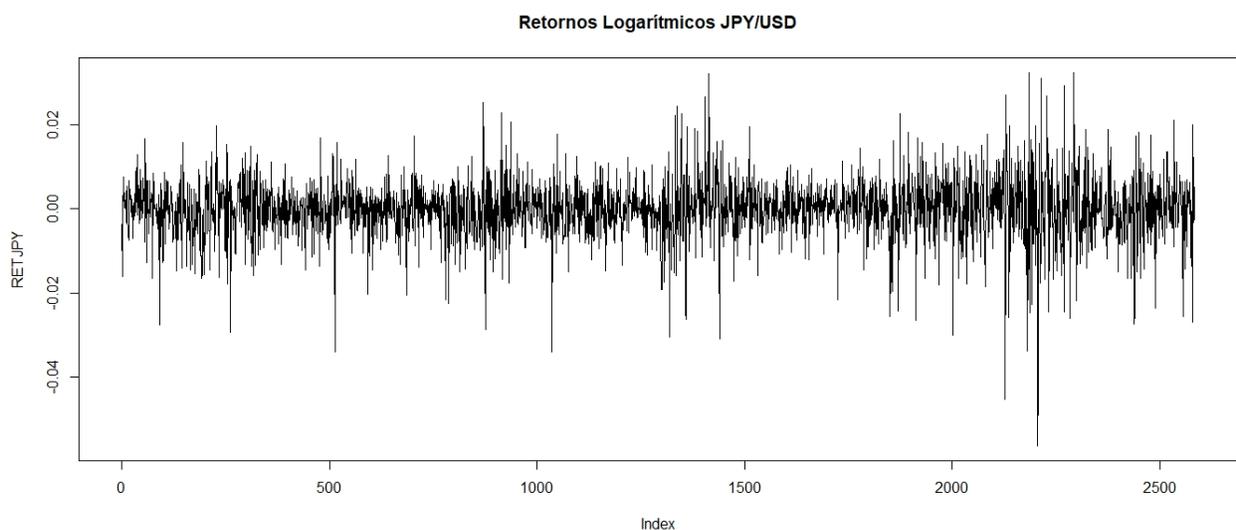


Figura 6.2: Retornos Logarítmicos Yen Japonés.

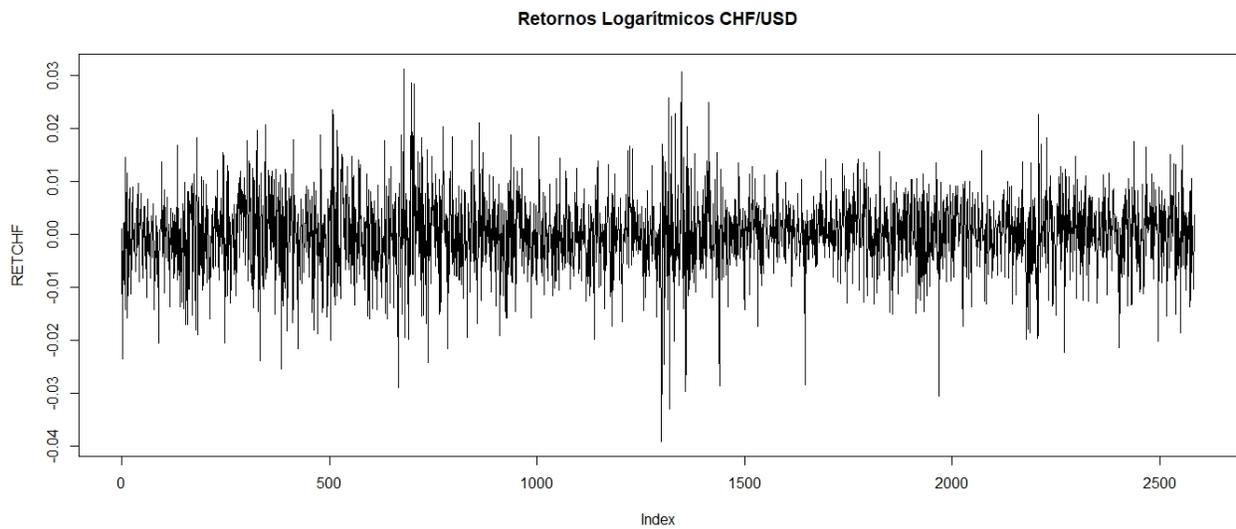


Figura 6.3: Retornos Logarítmicos Franco Suizo.

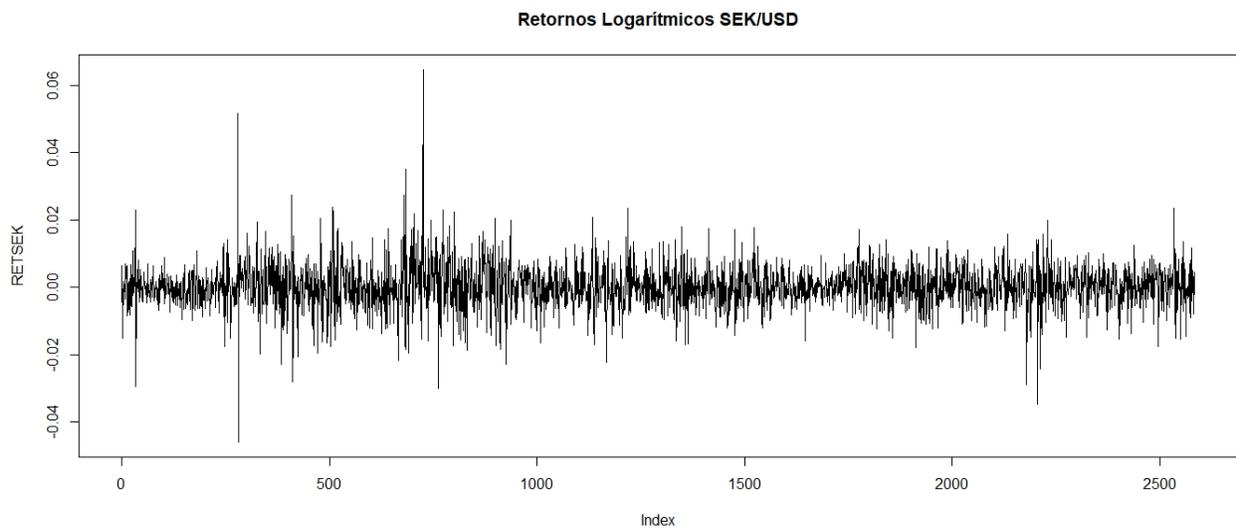


Figura 6.4: Retornos Logarítmicos Corona Sueca.

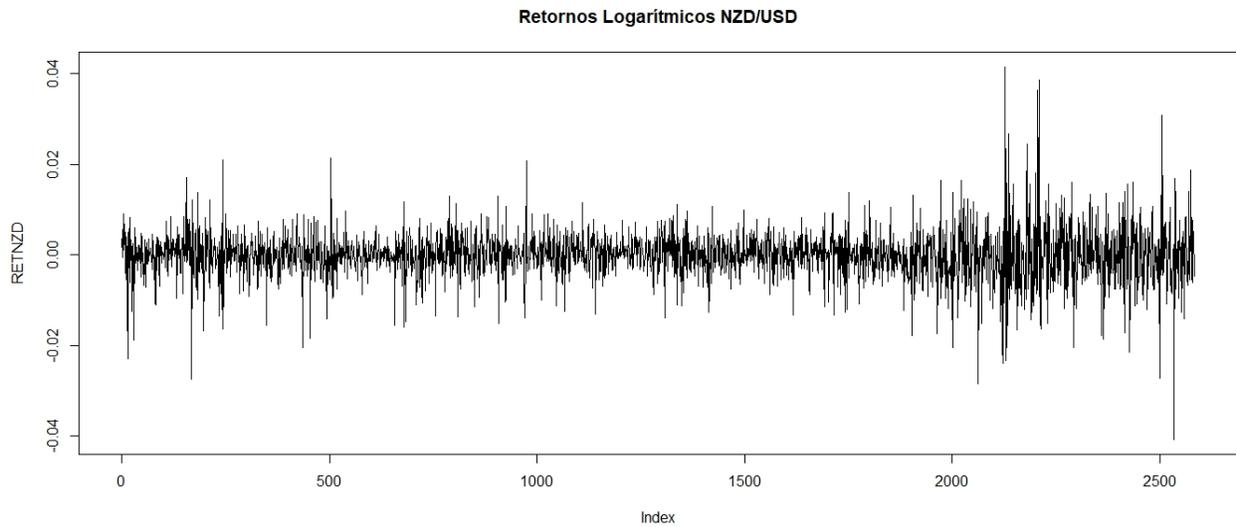


Figura 6.5: Retornos Logarítmicos Dólar Neozelandés.

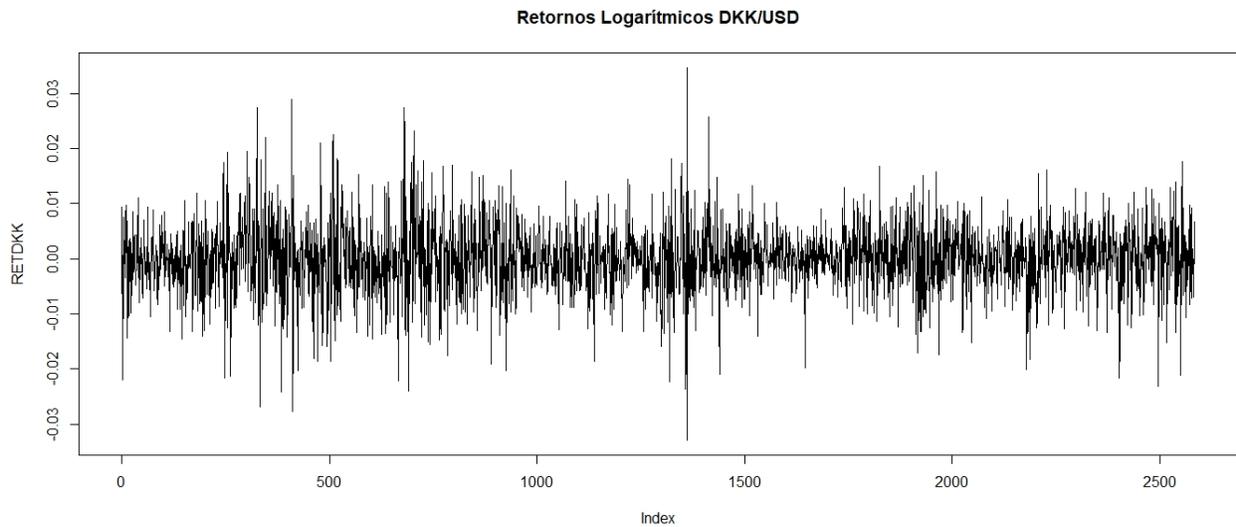


Figura 6.6: Retornos Logarítmicos Corona Danesa.

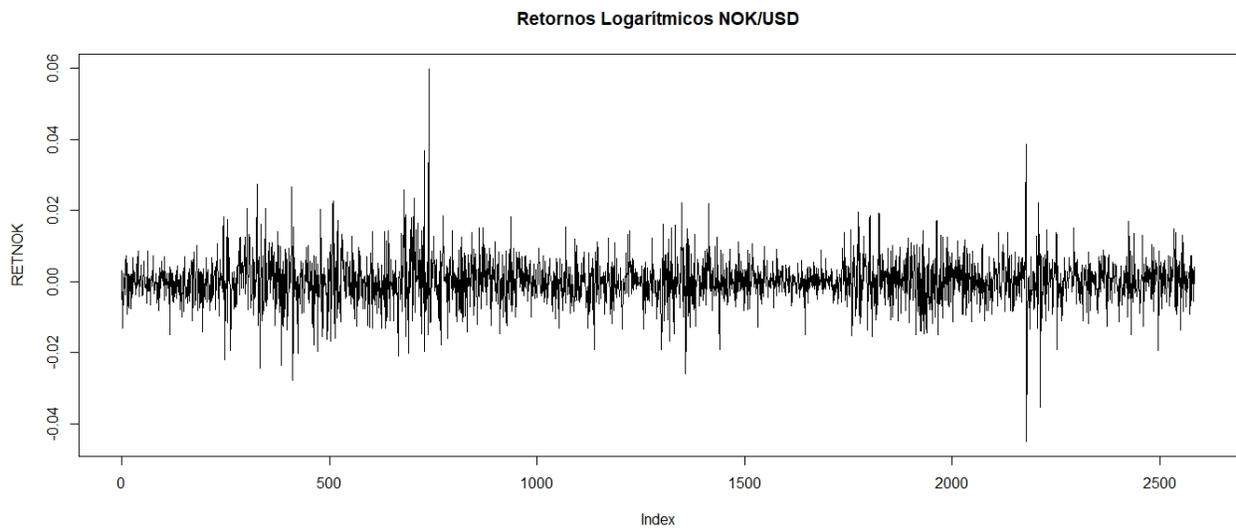


Figura 6.7: Retornos Logarítmicos Corona Noruega.

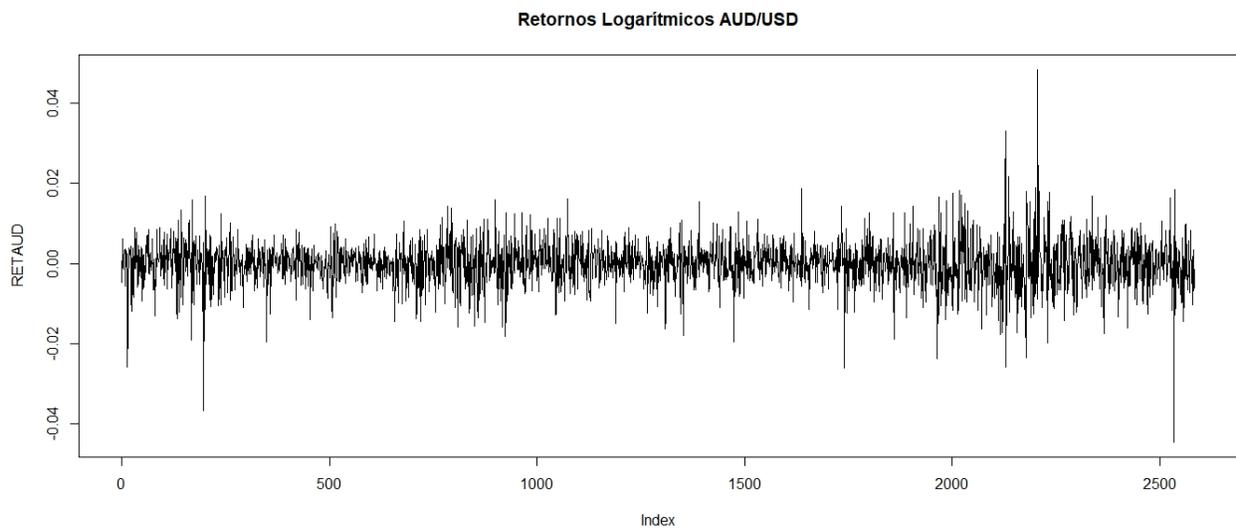


Figura 6.8: Retornos Logarítmicos Dólar Australiano.

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

Continuamos este estudio notando que la distribución de los datos no son normales, como comunmente ocurre en los datos financieros. Se realizó el histograma de los retornos, junto a sus Q-Q gráficos. Además se aplicó el test de Lilliefors para comprobar de manera numérica.

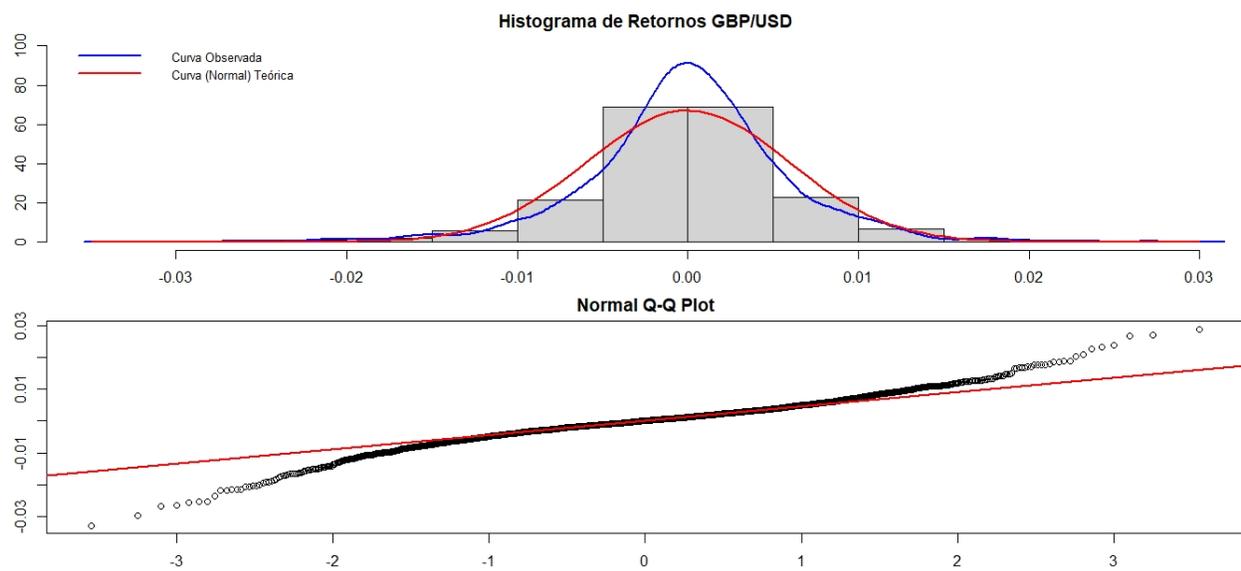


Figura 6.9: Histograma Retornos Logarítmicos Libra Esterlina.

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

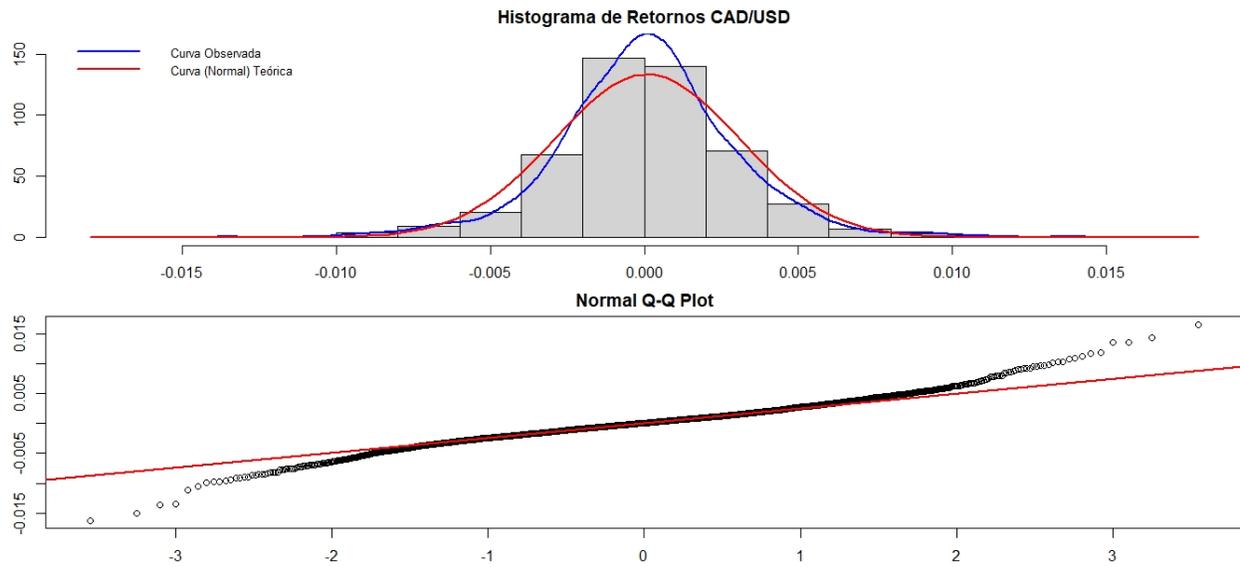


Figura 6.10: Histograma Retornos Logarítmicos Dólar Canadiense.

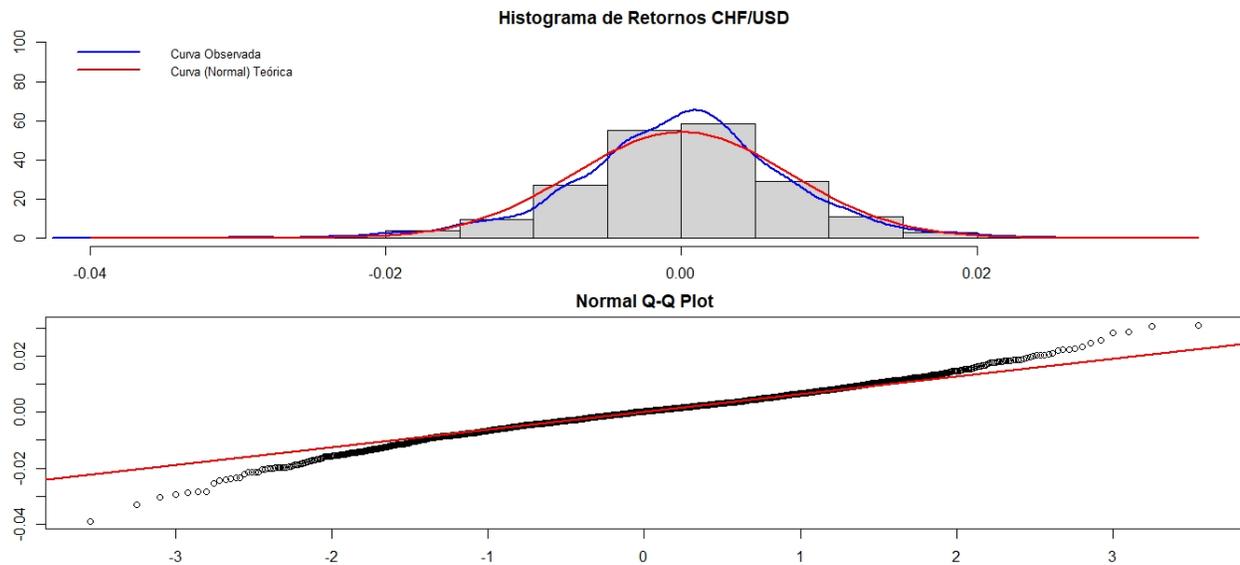


Figura 6.11: Histograma Retornos Logarítmicos Franco Suizo

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

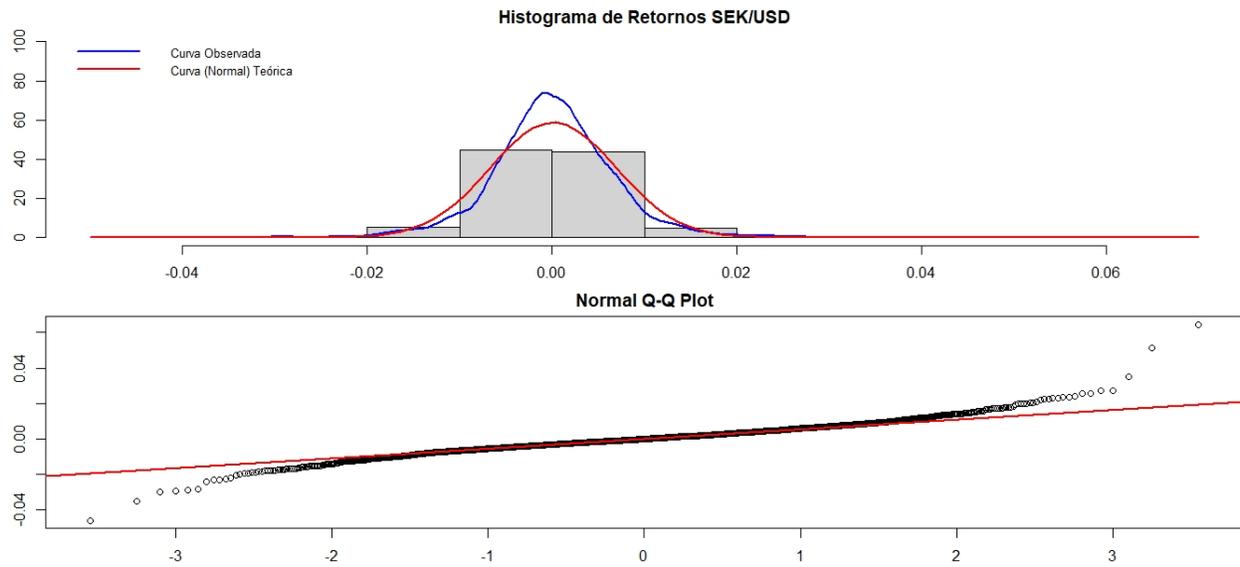


Figura 6.12: Histograma Retornos Logarítmicos Corona Sueca.

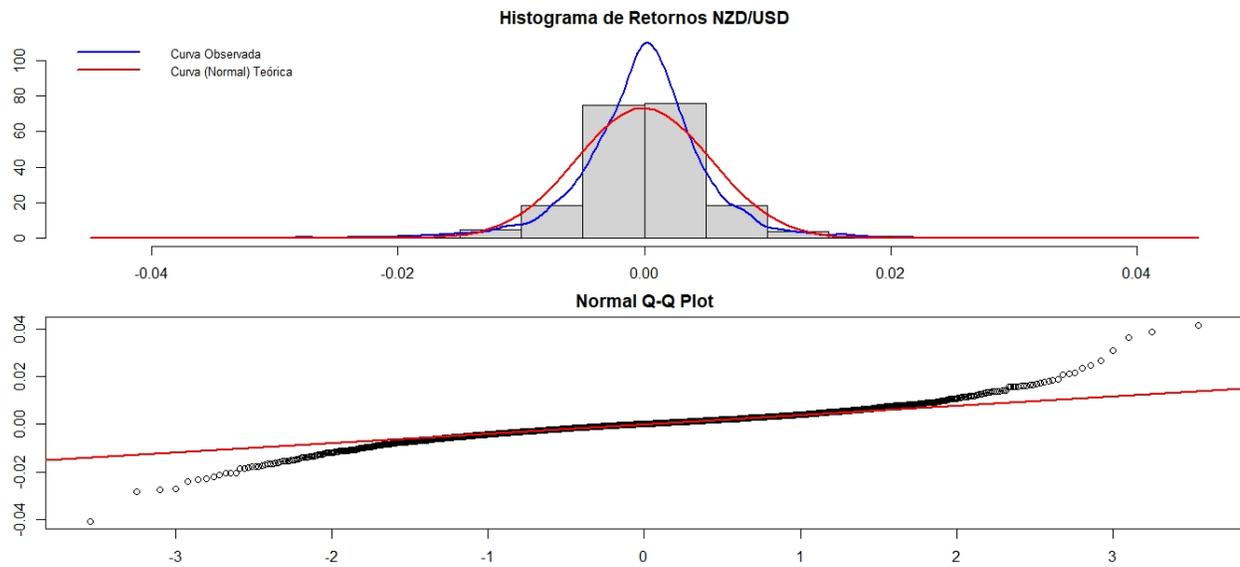


Figura 6.13: Histograma Retornos Logarítmicos Dólar Neozelandés.

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

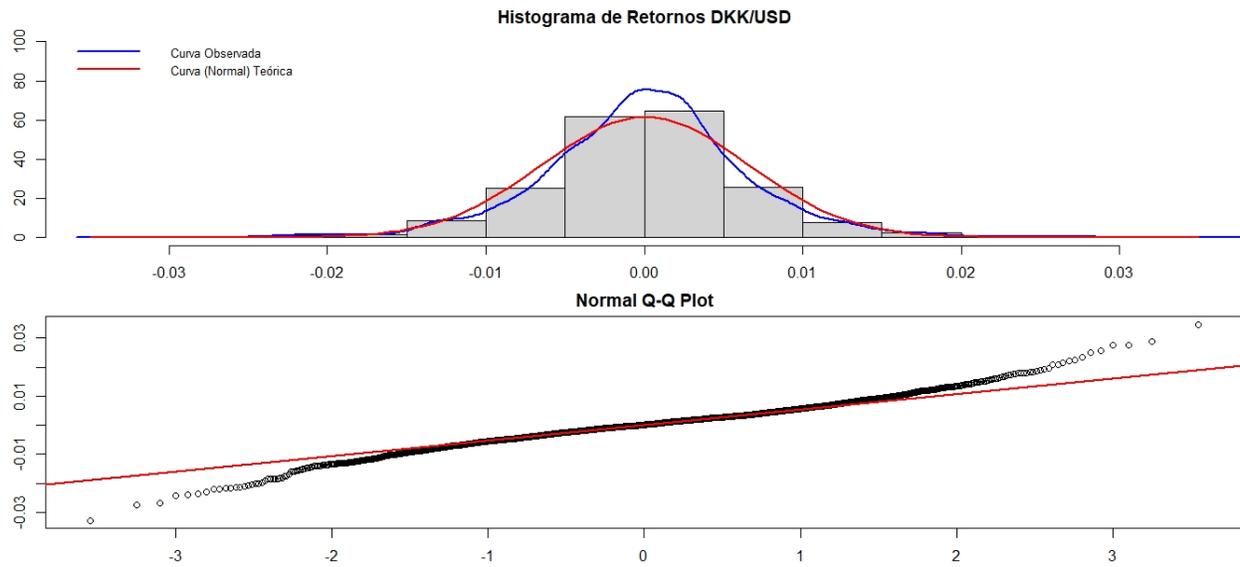


Figura 6.14: Histograma Retornos Logarítmicos Corona Danesa.

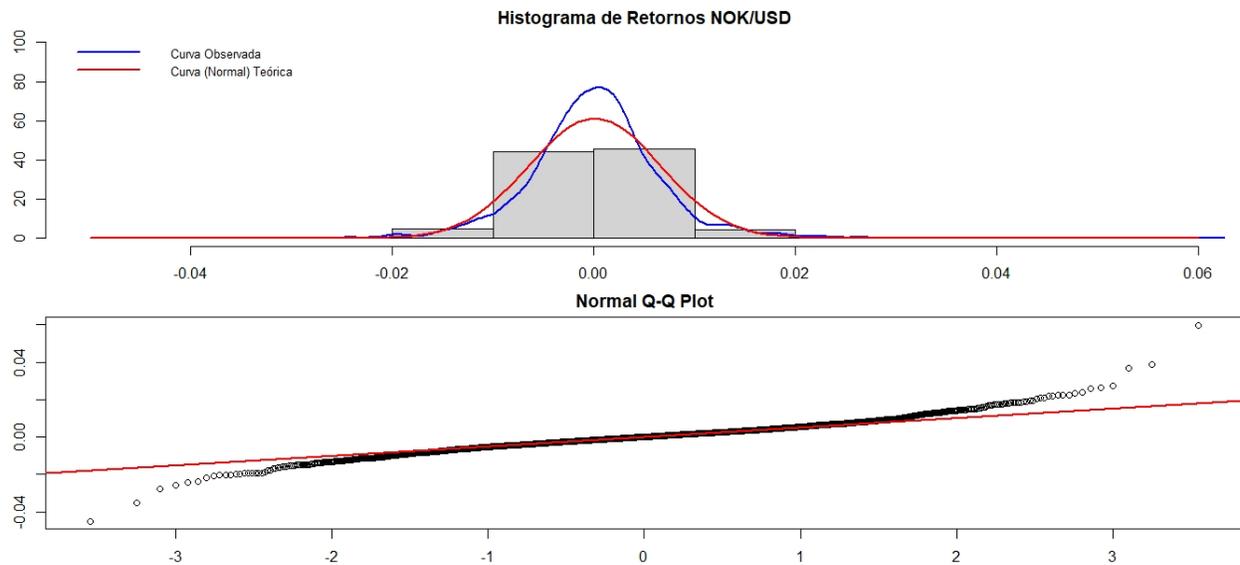


Figura 6.15: Histograma Retornos Logarítmicos Corona Noruega.

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

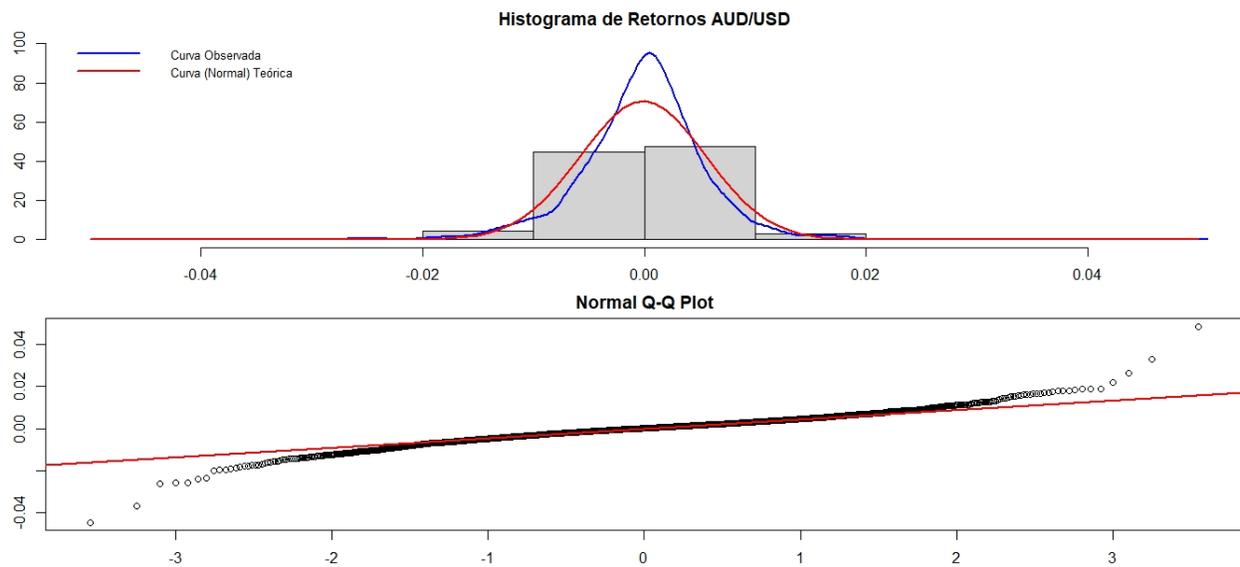


Figura 6.16: Histograma Retornos Logarítmicos Dólar Australiano.

Moneda	Estadístico	p-valor
RETGBP	0.065542	2.2e-16
RETJPY	0.066131	2.2e-16
RETNZD	0.071766	2.2e-16
RETCHEF	0.040038	2.818e-10
RETDKK	0.045491	1.866e-13
RETAUD	0.059419	2.2e-16
RETCAD	0.045861	1.092e-13
RETNOK	0.05582	2.2e-16
RETSEK	0.047879	5.443e-15

Cuadro 6.5: Test de Lilliefors para las Series de Retornos Logarítmicos.

6.1.1. Aplicación LCP

En el contexto multivariado se realizó un ajuste bivariado a los datos ya descritos, aplicando el paquete AWS: Adaptive Weights Smoothing desarrollado por Polzehl, J. (2021), que contiene funciones para estimar volatilidad adaptativa, además de herramientas técnicas

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

para análisis de imágenes.

Se aplica el paquete utilizando intervalos de 250 datos dejando los primeros 1000 como una muestra. El objetivo de escoger esta longitud de 250 datos es generar 6 intervalos y con esto analizar todos los datos disponibles. Por lo tanto, se aplica el enfoque en el intervalo de [1001,1251] datos, continuando con el intervalo [1251,1501], así sucesivamente... Observamos los gráficos para la Libra Esterlina y el Dólar Australiano, la Corona Sueca y el Yen Japonés, respectivamente. Notamos que la estimación adaptativa genera una curva más bien suave, pero que es capaz de notar los cambios bruscos en la volatilidad.

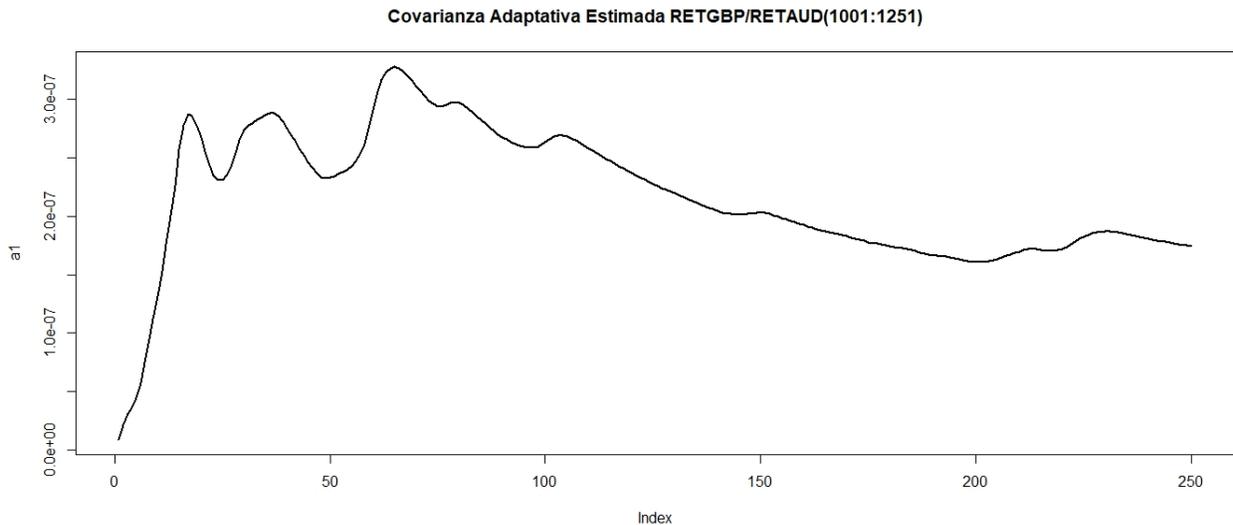


Figura 6.17: Covarianza Estimada Series de Retornos Logarítmicos Libra Esterlina, Dólar Australiano [1001:1251].

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

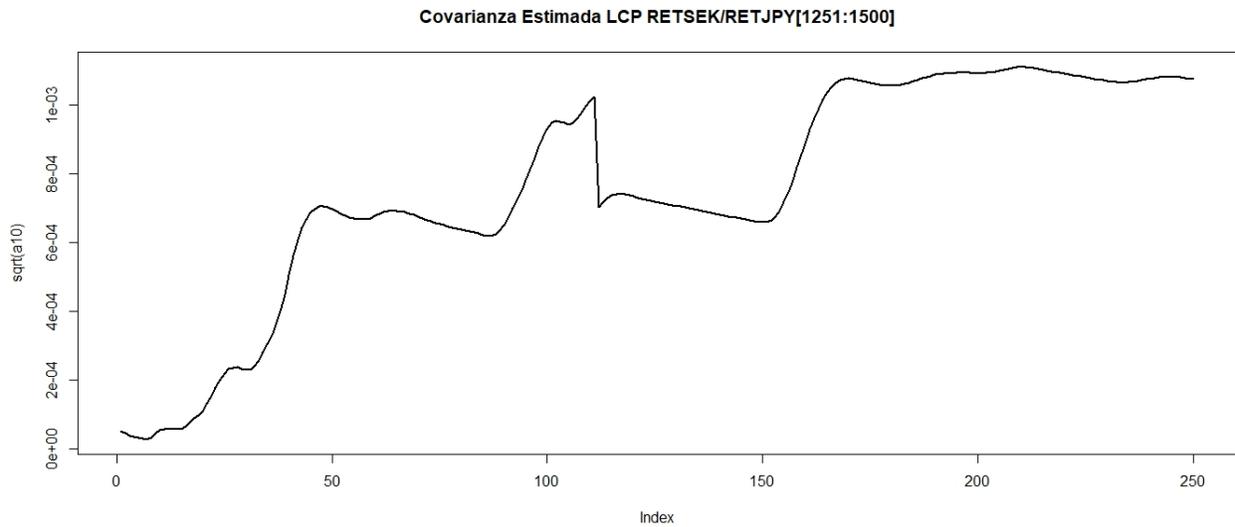


Figura 6.18: Covarianza Estimada Series Retornos Logarítmicos Corona Sueca, Yen Japonés [501:1500].



6.1.2. Aplicación GARCH

Se aplicó el modelo GARCH utilizando los 1000 datos anteriores para un mejor ajuste de los parámetros del modelo. La siguiente tabla resume los modelos utilizados en cada intervalo. Posteriormente se realizó un ajuste GARCH(1,1) para cada serie de tiempo.

Serie	[251:1250]	[501:1500]	[751:1750]	[1001:2000]	[1251:2250]	[1501:2500]
RETGBP	ARMA(2,3)	ARMA(1,1)	ARMA(1,1)	ARMA(1,1)	ARMA(1,1)	ARMA(2,2)
RETAUD	ARMA(2,2)	ARMA(4,2)	ARMA(1,2)	ARMA(1,1)	ARMA(2,2)	ARMA(2,2)
RETJPY	ARMA(2,2)	ARMA(2,2)	ARMA(2,2)	ARMA(2,2)	ARMA(2,2)	ARMA(2,2)
RETCHEF	ARMA(1,1)	ARMA(2,2)	ARMA(2,2)	ARMA(2,3)	ARMA(2,2)	ARMA(3,3)
RETSEK	ARMA(1,1)	ARMA(2,3)	ARMA(2,2)	ARMA(4,4)	ARMA(4,5)	ARMA(2,3)
RETDKK	ARMA(1,1)	ARMA(2,2)	ARMA(4,4)	ARMA(1,1,2)	ARMA(3,4)	ARMA(5,4)
RETNZD	ARMA(2,1)	ARMA(2,1)	ARMA(1,2)	ARMA(2,2)	ARMA(2,2)	ARMA(2,2)
RETCAD	ARMA(1,1,1)	ARMA(1,1)	ARMA(1,1)	ARMA(1,1)	ARMA(4,2)	ARMA(4,2)
RETNOK	ARMA(2,3)	ARMA(2,2)	ARMA(4,4)	ARMA(0,1,1)	ARMA(2,2)	ARMA(2,2)

Cuadro 6.6: Ajuste ARMA de las Series de Retornos Logarítmicos.

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

Explicaremos cada paso del ajuste para dos series en particular. Comenzamos con la serie SEK[501:1500] analizando sus gráficas de autocorrelación y autocorrelación parcial, donde podemos ver al rededor de 6 lag significativos en la ACF y 4 lag significativos en la PACF.

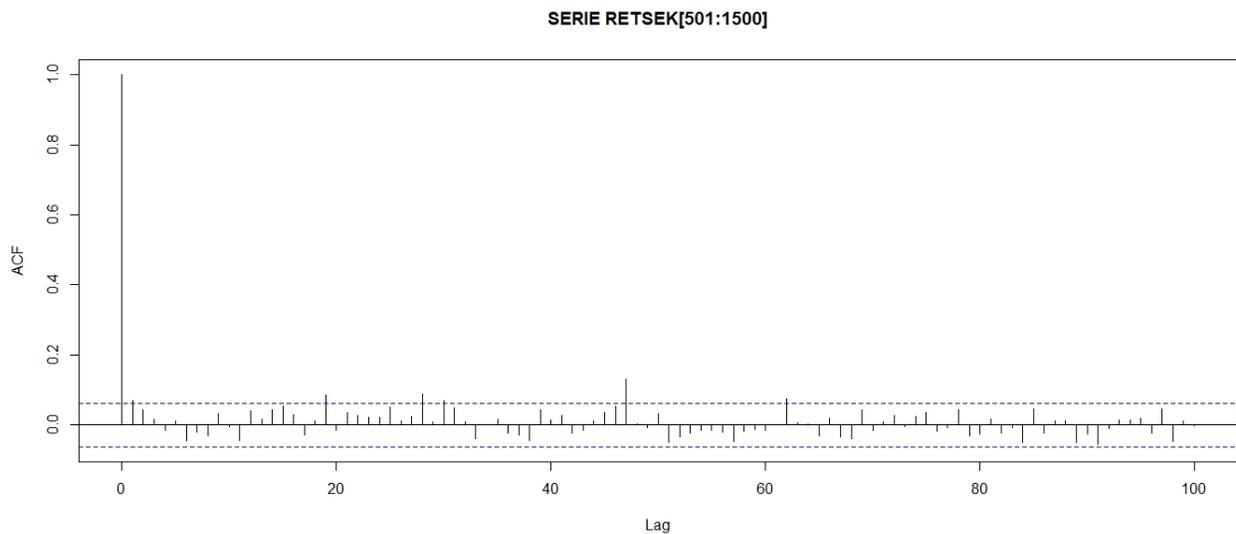


Figura 6.19: Gráfico de Autocorrelación Serie Retornos Logarítmicos Corona Sueca [501:1500].

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

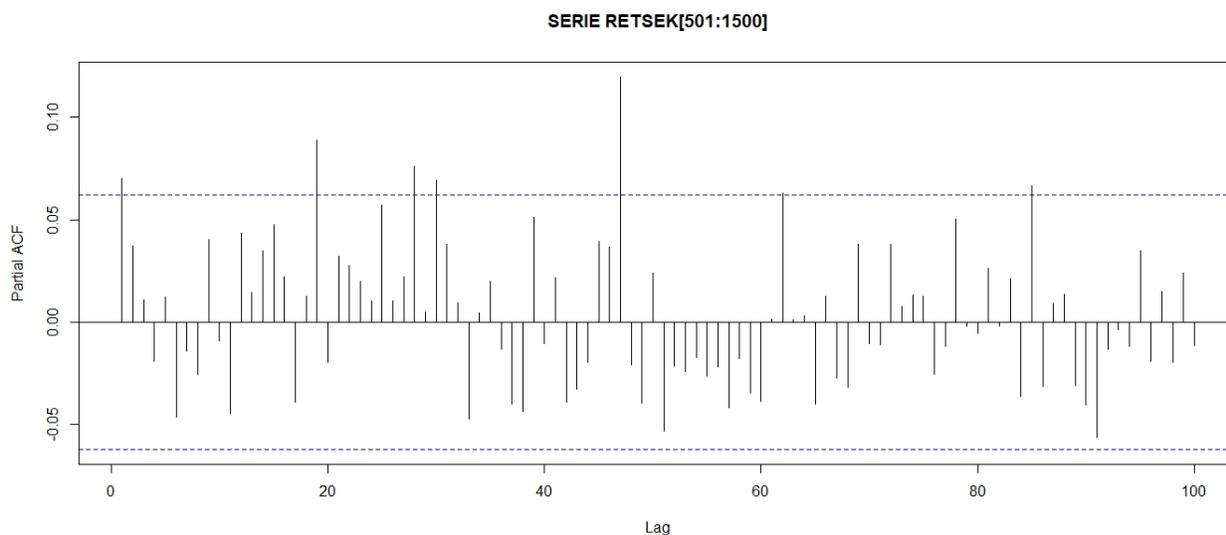


Figura 6.20: Gráfico de Autocorrelación Parcial Serie Retornos Logarítmicos Corona Sueca [501:1500].



Después de probar con varios modelos ARMA, ajustamos un modelo ARMA(2,3) a los datos.

```
> coefest (ARMASEK23)

z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ar1      0.73800129  0.01840847  40.0903 < 2.2e-16 ***
ar2     -0.95579246  0.01848804 -51.6979 < 2.2e-16 ***
ma1     -0.67917949  0.03570334 -19.0229 < 2.2e-16 ***
ma2      0.92357274  0.02881873  32.0477 < 2.2e-16 ***
ma3      0.08326569  0.03185549   2.6139  0.008953 **
intercept 0.00017257  0.00026546   0.6501  0.515643
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

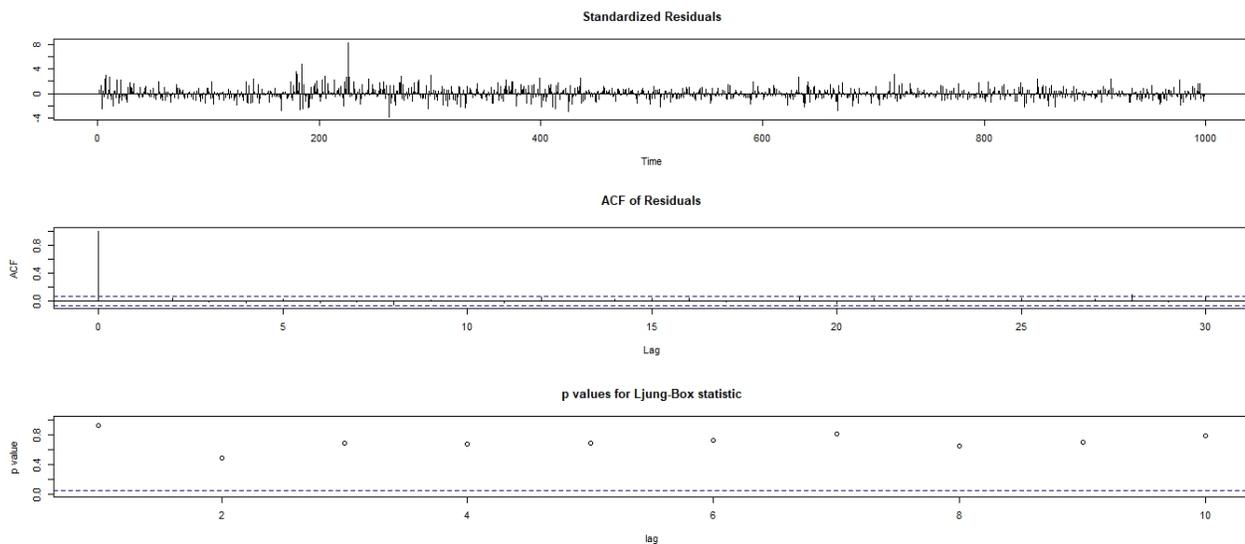


Figura 6.21: Resumen Calidad de Ajuste \star ARMA(2,3) Serie Retornos Corona Sueca [501:1500].



A simple vista el ajuste de los parámetros parece funcionar, aplicando la prueba de Ljung a los residuales obtenemos que el modelo ajusta bien a los datos es decir los residuos tienen forma de ruido blanco. Sin embargo al analizar los residuales al cuadrado esta situación cambia, se rechaza la hipótesis nula del test de Ljung-Box y se concluye que la serie está correlacionada. Aplicamos una prueba ARCH, para comprobar que existen componentes autorregresivos en los residuales lo que nos hace sospechar que el modelo pueda ser ajustado mediante un modelo GARCH(p,q).

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

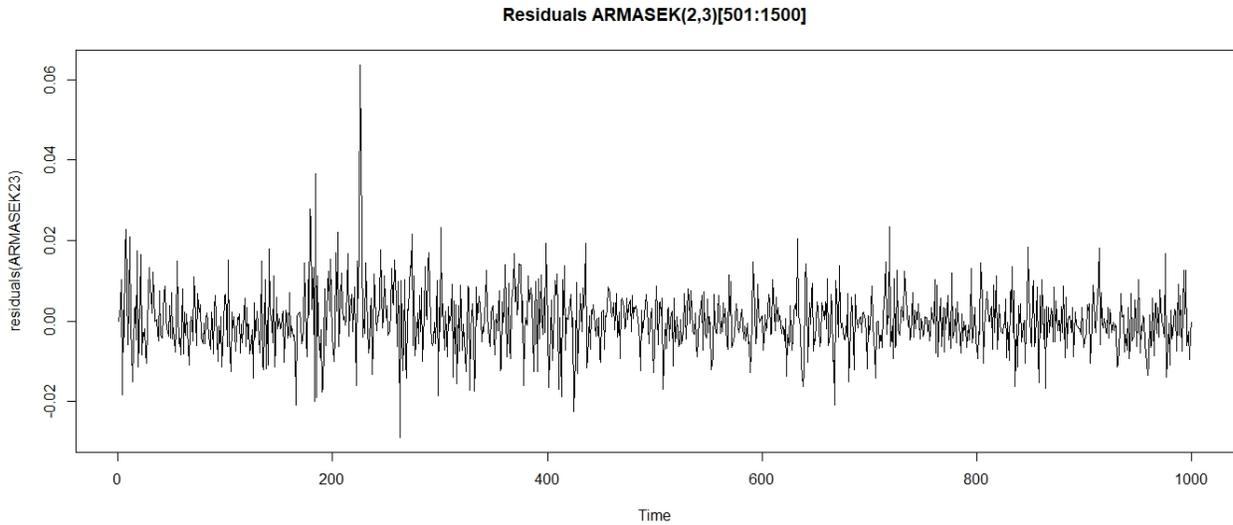


Figura 6.22: Residuales Modelo ARMA(2,3) Serie Retornos Corona Sueca [501:1500].

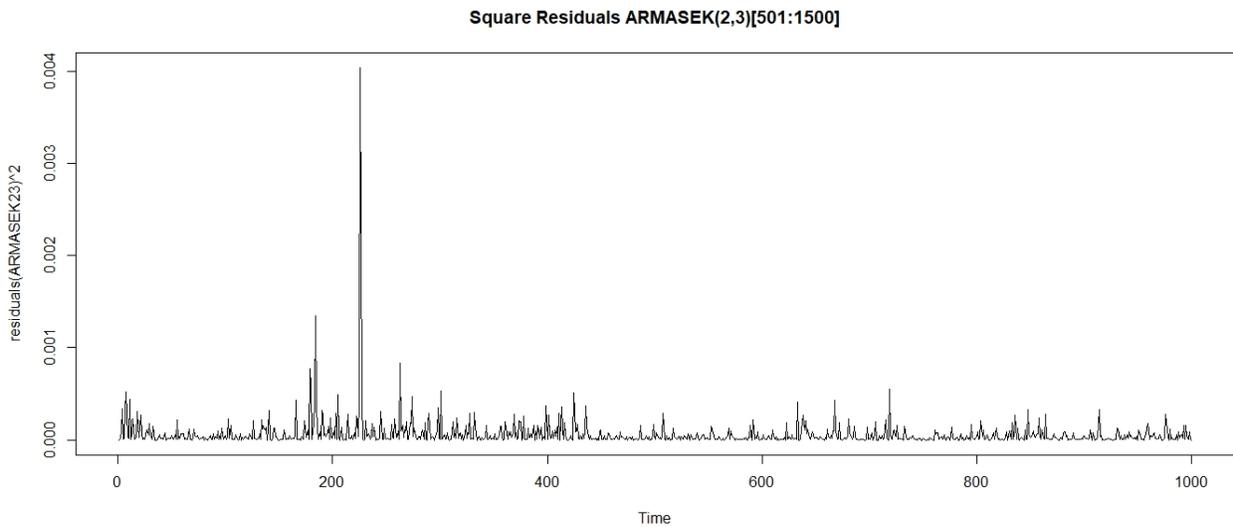


Figura 6.23: Residuales Cuadrados Modelo ARMA(2,3), Serie Retornos Corona Sueca [501:1500].

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

```
> ArchTest(residuals(ARMASEK23),lag=3)

      ARCH LM-test; Null hypothesis: no ARCH effects

data: residuals(ARMASEK23)
Chi-squared = 32.755, df = 3, p-value = 3.627e-07

> Box.test(residuals(ARMASEK23),type="Ljung-Box")

      Box-Ljung test

data: residuals(ARMASEK23)
X-squared = 0.01068, df = 1, p-value = 0.9177

> Box.test(residuals(ARMASEK23)^2,lag=1,type=c("Box-Pierce","Ljung-Box"),fitdf=0)

      Box-Pierce test

data: residuals(ARMASEK23)^2
X-squared = 29.762, df = 1, p-value = 4.884e-08
```

Por simplicidad utilizaremos un modelo GARCH(1,1) donde la media será modelada por un ARMA(2,3), y la distribución de los residuos será ajustada a través de una distribución t-student, en los parámetros estimados sólo resultan no significativos las constantes del modelo, lo que no será relevante para el ajuste.



CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

> FITGARCHSEK

```
*-----*
*           GARCH Model Fit           *
*-----*

Conditional Variance Dynamics
-----
GARCH Model      : sGARCH(1,1)
Mean Model       : ARFIMA(2,0,3)
Distribution      : std

Optimal Parameters
-----
      Estimate  Std. Error   t value  Pr(>|t|)
mu      -0.000249   0.000214   -1.16090 0.245681
ar1      0.707814   0.034983   20.23307 0.000000
ar2     -0.944234   0.014764  -63.95546 0.000000
ma1     -0.696463   0.032053  -21.72853 0.000000
ma2      0.933104   0.000834 1118.81608 0.000000
ma3      0.039281   0.002144   18.31730 0.000000
omega    0.000002   0.000002    0.82562 0.409018
alpha    0.059641   0.028817    2.06961 0.038489
beta1    0.908118   0.036464   24.90480 0.000000
shape    6.810502   1.456312    4.67654 0.000003

Robust Standard Errors:
      Estimate  Std. Error   t value  Pr(>|t|)
mu      -0.000249   0.000212   -1.17579 0.239678
ar1      0.707814   0.045656   15.50323 0.000000
ar2     -0.944234   0.017421  -54.20017 0.000000
ma1     -0.696463   0.042909  -16.23112 0.000000
ma2      0.933104   0.000793 1176.57027 0.000000
ma3      0.039281   0.000988   39.77469 0.000000
omega    0.000002   0.000014    0.12799 0.898160
alpha    0.059641   0.148415    0.40185 0.687793
beta1    0.908118   0.208246    4.36080 0.000013
shape    6.810502   2.895927    2.35175 0.018685

LogLikelihood : 3526.757

Information Criteria
-----
Akaike      -7.0335
Bayes      -6.9844
Shibata    -7.0337
Hannan-Quinn -7.0149
```

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

Weighted Ljung-Box Test on Standardized Residuals

```

-----
                    statistic p-value
Lag[1]                1.476  0.2244
Lag[2*(p+q)+(p+q)-1][14]  5.290  1.0000
Lag[4*(p+q)+(p+q)-1][24]  9.212  0.9029
d.o.f=5
H0 : No serial correlation

```

Weighted Ljung-Box Test on Standardized Squared Residuals

```

-----
                    statistic p-value
Lag[1]                3.716  0.05389
Lag[2*(p+q)+(p+q)-1][5]  5.410  0.12343
Lag[4*(p+q)+(p+q)-1][9]  7.258  0.17850
d.o.f=2

```

Weighted ARCH LM Tests

```

-----
          Statistic Shape Scale P-Value
ARCH Lag[3]    0.2701 0.500 2.000  0.6033
ARCH Lag[5]    2.2076 1.440 1.667  0.4272
ARCH Lag[7]    3.5271 2.315 1.543  0.4199

```

Nyblom stability test

Joint Statistic: 76.1554

Individual Statistics:

```

mu      0.24326
ar1     0.01815
ar2     0.06056
ma1     0.03525
ma2     0.10396
ma3     0.13553
omega   5.47086
alpha   0.34335
beta1   0.29351
shape   0.07347

```



Asymptotic Critical Values (10% 5% 1%)

Joint Statistic: 2.29 2.54 3.05

Individual Statistic: 0.35 0.47 0.75

Sign Bias Test

```

-----
          t-value   prob sig
Sign Bias      0.57250 0.56711
Negative Sign Bias 0.03397 0.97291
Positive Sign Bias 2.38216 0.01740 **
Joint Effect    6.29380 0.09816  *

```

Adjusted Pearson Goodness-of-Fit Test:

```

-----
group statistic p-value(g-1)
1    20     22.40     0.2648
2    30     31.70     0.3332
3    40     44.24     0.2600
4    50     54.40     0.2764

```

Al observar ahora los gráficos de autocorrelación y autocorrelación parcial de los residuales al cuadrado estandarizados y aplicar el test de Ljung-Box podemos notar que el ajuste GARCH(1,1) mejora bastante en comparación con el modelo ARMA(2,3).

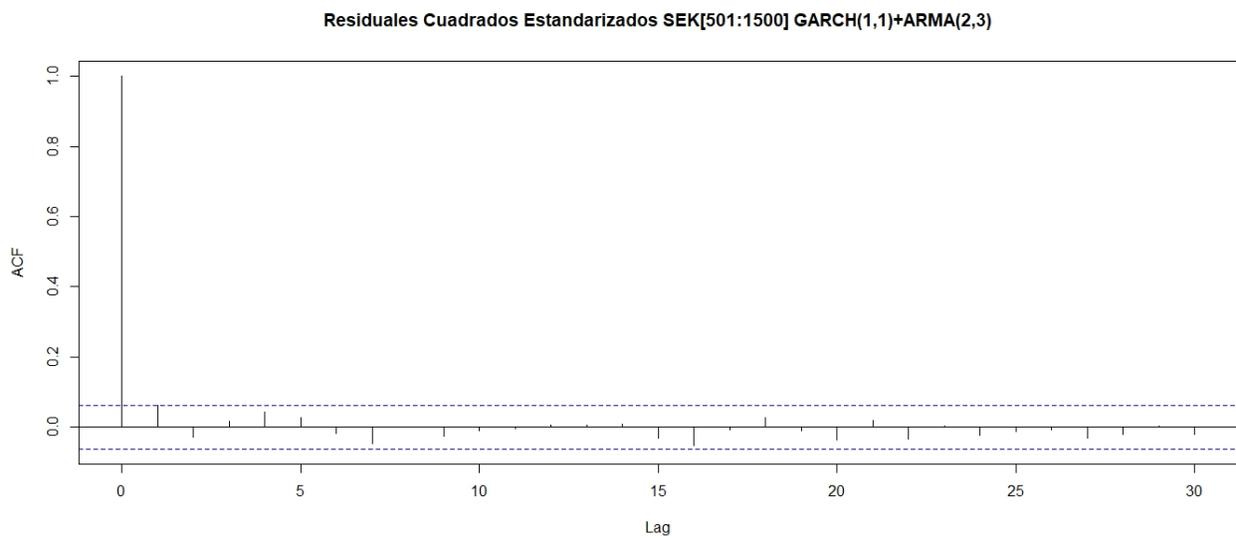


Figura 6.24: ACF Residuales Cuadrados Estandarizados GARCH(1,1).

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

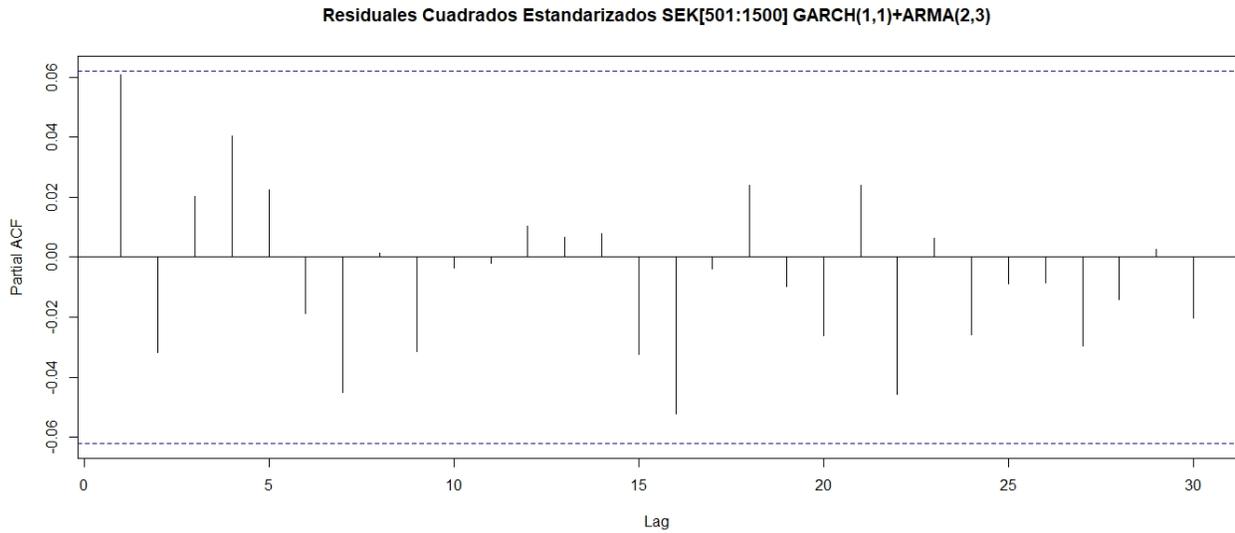


Figura 6.25: PACF Residuales Cuadrados Estandarizados GARCH(1,1).



De manera análoga se hace el estudio para la serie RETJPY.

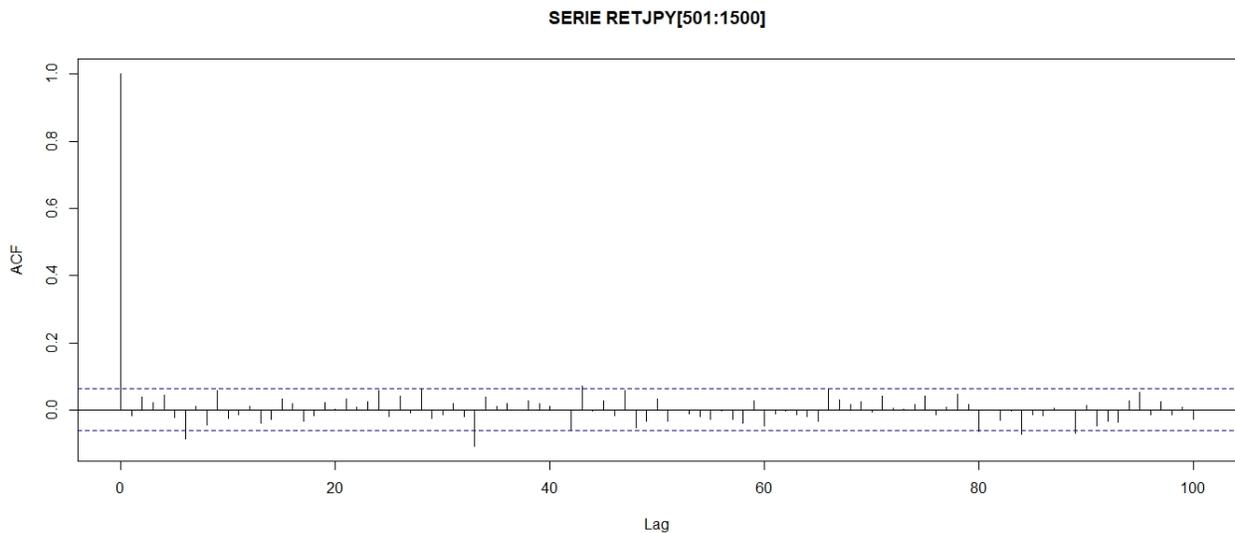


Figura 6.26: Gráfico de Autocorrelación Serie Retornos Yen Japonés [501:1500].

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

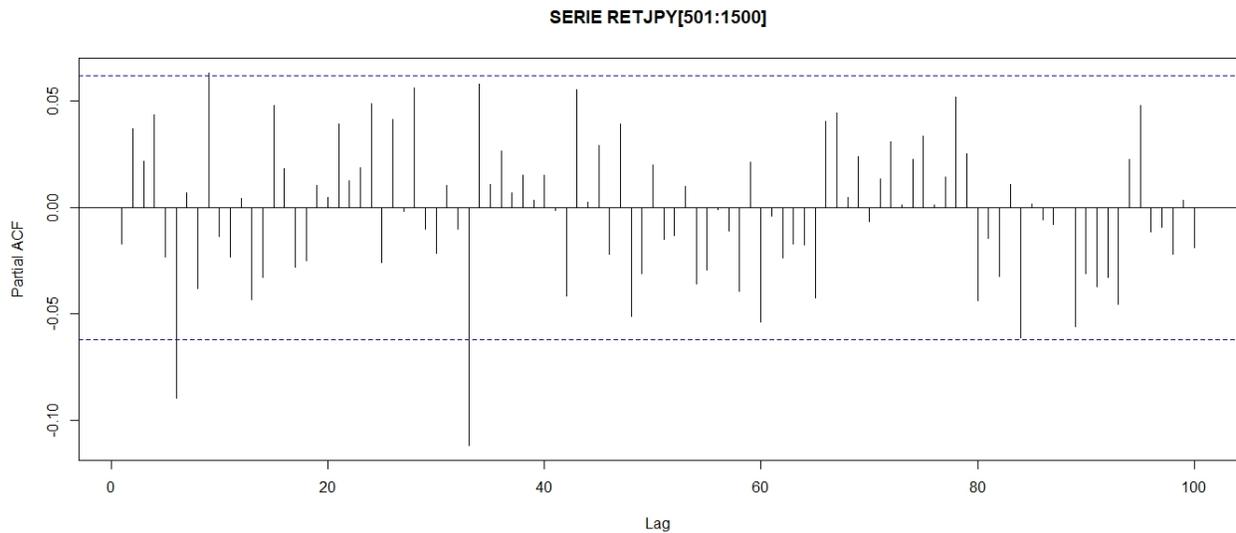


Figura 6.27: Gráfico de Autocorrelación Parcial Serie Retornos Yen Japonés [501:1500].



Ajustamos un modelo ARMA(2,2) a los datos.

```
> coeftest(ARMAJPY22)
```

```
z test of coefficients:
```

	Estimate	Std. Error	z value	Pr(> z)
ar1	9.8828e-01	1.0349e-01	9.5500	<2e-16 ***
ar2	-8.0063e-01	7.7692e-02	-10.3052	<2e-16 ***
ma1	-1.0138e+00	8.6464e-02	-11.7256	<2e-16 ***
ma2	8.6557e-01	6.3642e-02	13.6006	<2e-16 ***
intercept	-6.0643e-05	2.3541e-04	-0.2576	0.7967

```
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

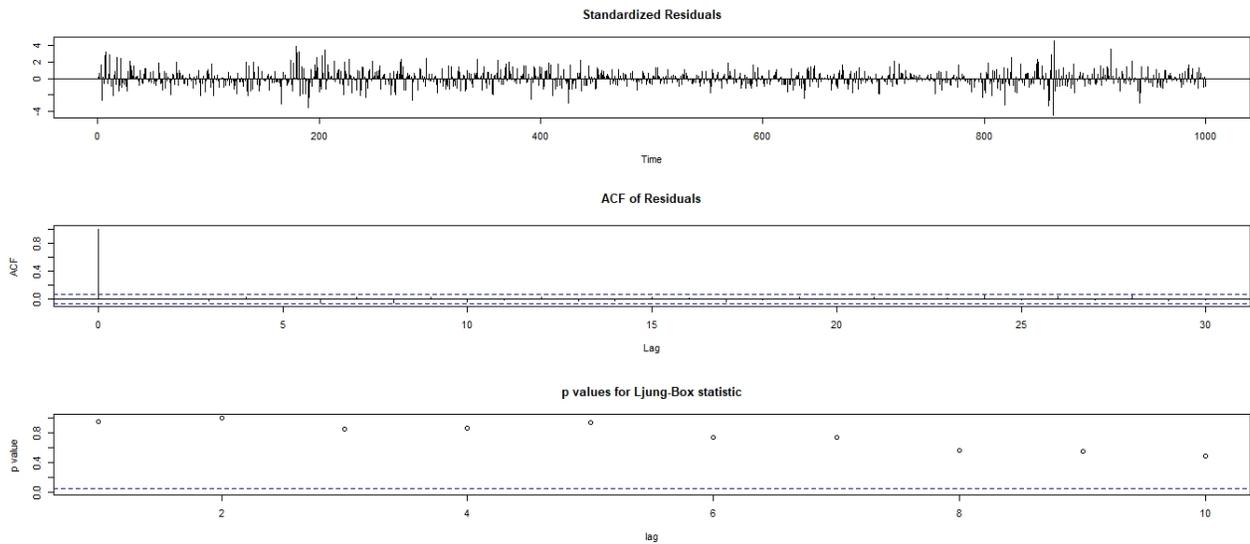


Figura 6.28: Resumen Ajuste ARMA(2,2), Serie Retornos Yen Japonés [501:1500].

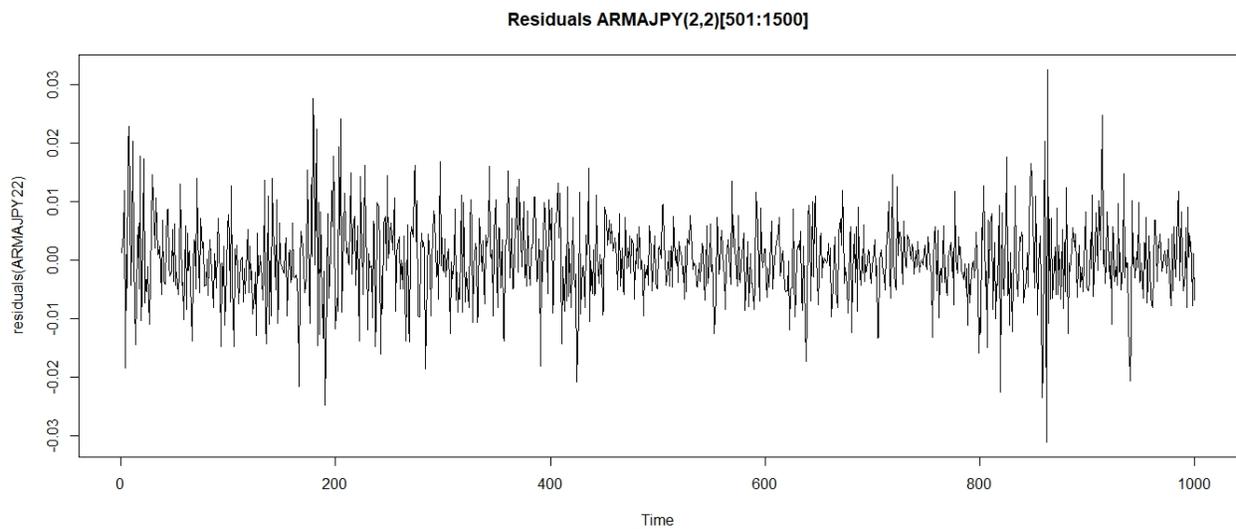


Figura 6.29: Residuales Ajuste ARMA(2,2), Serie Retornos Yen Japonés [501:1500].

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

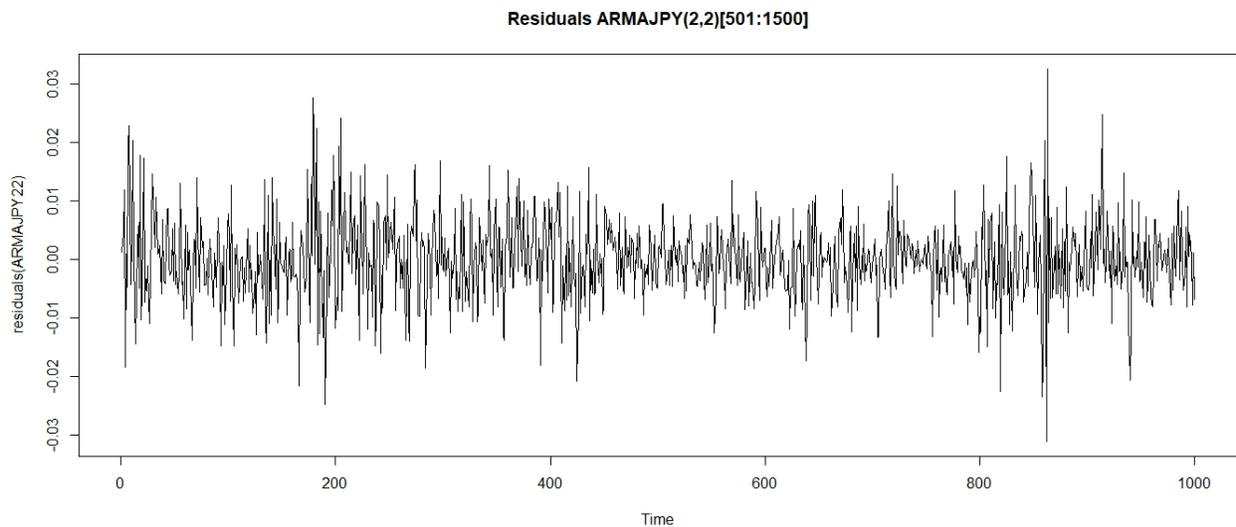


Figura 6.30: Residuales Cuadrados Ajuste ARMA(2,2), Serie Retornos Yen Japonés [501:1500].



```
> ArchTest(residuals(ARMAJPY22), lag=3)

      ARCH LM-test; Null hypothesis: no ARCH effects

data: residuals(ARMAJPY22)
Chi-squared = 86.493, df = 3, p-value < 2.2e-16

> Box.test(residuals(ARMAJPY22), type="Ljung-Box")

      Box-Ljung test

data: residuals(ARMAJPY22)
X-squared = 0.0057498, df = 1, p-value = 0.9396

> Box.test(residuals(ARMAJPY22)^2, lag=1, type=c("Box-Pierce", "Ljung-Box"), fitdf=0)

      Box-Pierce test

data: residuals(ARMAJPY22)^2
X-squared = 64.047, df = 1, p-value = 1.221e-15
```

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

```
> FITGARCHJPY
```

```
*-----*
*           GARCH Model Fit           *
*-----*

Conditional Variance Dynamics
-----
GARCH Model      : sGARCH(1,1)
Mean Model       : ARFIMA(2,0,2)
Distribution      : std

Optimal Parameters
-----
      Estimate Std. Error  t value Pr(>|t|)
mu      -0.000206   0.000195  -1.05707 0.290481
ar1      0.923431   0.094516   9.77008 0.000000
ar2     -0.842010   0.071504 -11.77566 0.000000
ma1     -0.959060   0.080206 -11.95747 0.000000
ma2      0.891667   0.059430  15.00355 0.000000
omega    0.000002   0.000003   0.57411 0.565895
alpha    0.089693   0.051044   1.75719 0.078885
beta1    0.875364   0.062367  14.03570 0.000000
shape    8.473275   2.626687   3.22584 0.001256

Robust Standard Errors:
      Estimate Std. Error  t value Pr(>|t|)
mu      -0.000206   0.000201  -1.025124 0.305305
ar1      0.923431   0.140814   6.557794 0.000000
ar2     -0.842010   0.071784 -11.729718 0.000000
ma1     -0.959060   0.120365  -7.967927 0.000000
ma2      0.891667   0.065079  13.701269 0.000000
omega    0.000002   0.000027   0.066304 0.947136
alpha    0.089693   0.397279   0.225769 0.821381
beta1    0.875364   0.508068   1.722927 0.084902
shape    8.473275  11.309036   0.749248 0.453708

LogLikelihood : 3598.179

Information Criteria
-----
Akaike      -7.1784
Bayes       -7.1342
Shibata     -7.1785
Hannan-Quinn -7.1616
```

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

Weighted Ljung-Box Test on Standardized Residuals

```

-----
                statistic p-value
Lag[1]                3.193 0.07394
Lag[2*(p+q)+(p+q)-1][11] 6.283 0.31031
Lag[4*(p+q)+(p+q)-1][19] 9.257 0.59476
d.o.f=4
H0 : No serial correlation
  
```

Weighted Ljung-Box Test on Standardized Squared Residuals

```

-----
                statistic p-value
Lag[1]                0.5488 0.4588
Lag[2*(p+q)+(p+q)-1][5] 1.7707 0.6735
Lag[4*(p+q)+(p+q)-1][9] 3.3284 0.7037
d.o.f=2
  
```

Weighted ARCH LM Tests

```

-----
                Statistic Shape Scale P-Value
ARCH Lag[3]    0.01593 0.500 2.000 0.8996
ARCH Lag[5]    0.12132 1.440 1.667 0.9827
ARCH Lag[7]    1.84306 2.315 1.543 0.7505
  
```

Nyblom stability test

Joint Statistic: 75.3005

Individual Statistics:

```

mu      0.07291
ar1     0.05902
ar2     0.06704
ma1     0.04682
ma2     0.06879
omega   5.96614
alpha   0.26997
beta1   0.38783
shape   0.53693
  
```



Asymptotic Critical Values (10% 5% 1%)

```

Joint Statistic:    2.1 2.32 2.82
Individual Statistic: 0.35 0.47 0.75
  
```

Sign Bias Test

```

-----
                t-value  prob sig
Sign Bias                0.7706 0.4411
Negative Sign Bias      0.5565 0.5780
Positive Sign Bias      0.4616 0.6444
Joint Effect            2.0197 0.5683
  
```

Adjusted Pearson Goodness-of-Fit Test:

```

-----
group statistic p-value(g-1)
1   20    11.84    0.8923
2   30    20.72    0.8691
3   40    28.00    0.9048
4   50    35.30    0.9292
  
```

Elapsed time : 0.413305

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

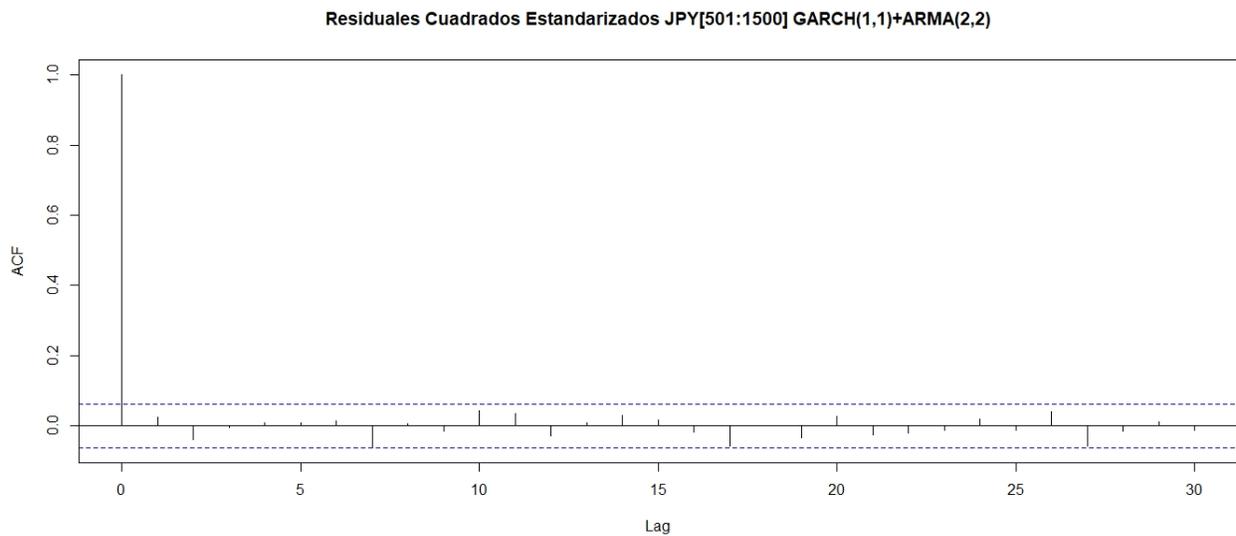


Figura 6.31: ACF Residuales Cuadrados Estandarizados Retornos Logarítmicos Yen Japonés [501:1500] GARCH(1,1).

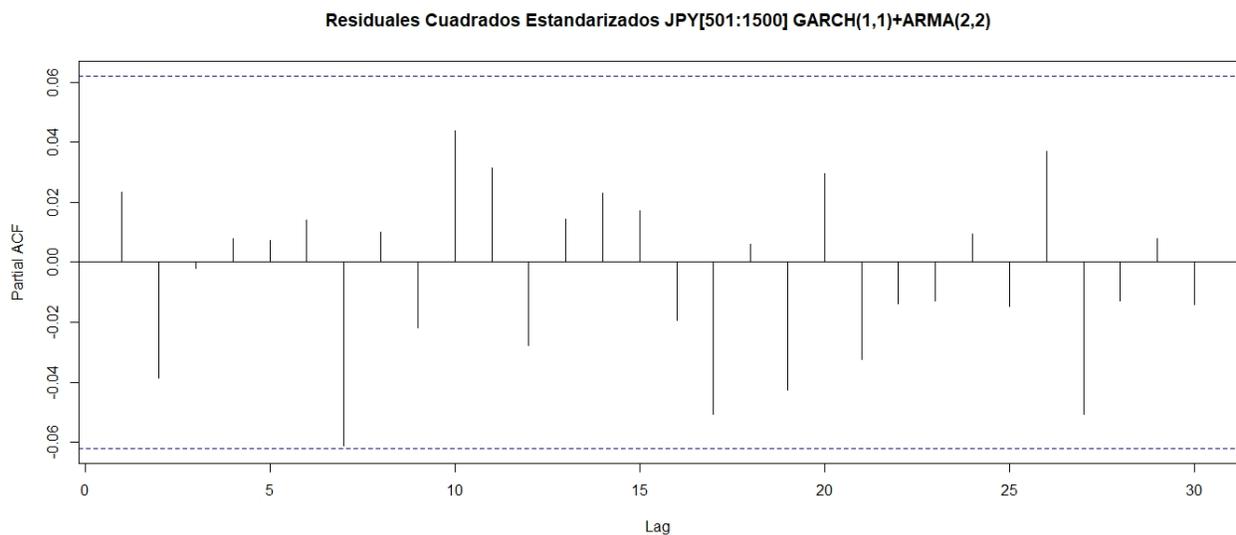


Figura 6.32: PACF Residuales Cuadrados Estandarizados Retornos Logarítmicos Yen Japonés [501:1500] GARCH(1,1).

6.1.3. Aplicación del Modelo DCC

En este apartado se realizó un ajuste bivariado siguiendo las técnicas clásicas donde destacamos al modelo DCC, que será descrito para una mejor comprensión de esta aplicación.

Modelos de Correlación Condicional.

En Almeida, D. et al,(2015), [1], se menciona la posibilidad de modelar la matriz de Covarianza Condicional Σ_t mediante el producto de la matriz de covarianzas y correlaciones condicionales de la siguiente forma:

$$\Sigma_t = \text{diag}(\Sigma_t)^{\frac{1}{2}} H_t \text{diag}(\Sigma_t)^{\frac{1}{2}}$$



Donde $\text{diag}(\Sigma_t) = \text{diag}(\sigma_{11,t}, \dots, \sigma_{nn,t})$ es la matriz diagonal cuyos elementos son las varianzas condicionales de cada serie temporal y H_t es una matriz que contiene las correlaciones condicionales de cada serie.

En este modelo se asume que las varianzas y correlaciones condicionales no están relacionadas, por lo que se pueden estimar primero las varianzas condicionales y luego las correlaciones.

El modelo se puede resumir en los siguientes pasos:

- Utilizar un modelo GARCH(p,q) para la varianza estimada de cada serie temporal,

$$\sigma_{i,t}^2 = \omega_i + \sum_{i=1}^q \alpha_i a_{i,t-i}^2 + \sum_{i=1}^p \beta_i \sigma_{i,t-i}^2.$$

- Construir la matriz diagonal que contiene estas varianzas,

$$D_t = \text{diag}\{\sigma_{1,t}, \dots, \sigma_{d,t}\}.$$

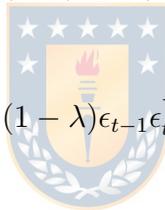
- Obtener las innovaciones escaladas,

$$\epsilon_t = D_t^{-1} \mathbf{a}_t.$$

, $\mathbf{a}_t = \Sigma_t^{\frac{1}{2}} \varepsilon_t$ con ε_t un vector de ruido blanco.

- Aplicar un modelo EWMA (Media Móvil Ponderada Exponencial) a las innovaciones escaladas,

$$Q_t = (1 - \lambda) \epsilon_{t-1} \epsilon_{t-1}^\top + \lambda Q_{t-1}.$$



- Obtener la matriz de Correlación Condicional,

$$H_t = \text{diag}\{Q_t\}^{-\frac{1}{2}} Q_t \text{diag}\{Q_t\}^{-\frac{1}{2}}.$$

- Obtener la matriz de Covarianza Condicional,

$$\Sigma_t = D_t H_t D_t.$$

Por lo tanto, para formular el modelo DCC, se debe tratar cada serie temporal de rentabilidades por separado y ajustar el modelo GARCH(P,Q) adecuado, para generar una matriz diagonal con dichos modelos. Posteriormente, se realiza el producto matricial entre el proceso vectorial de ruido blanco y inversa de la matriz Diagonal obtenida. Los componentes

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

de la matriz Q_t representan variables auxiliares que hacen el papel de covarianzas estimadas, se aplica un modelo recursivo para su obtención. Finalmente, se procede a reescalar la matriz Q_t para obtener matrices con 1 en la diagonal principal, la que será utilizada para estimar la matriz de volatilidad Σ_t .

```
> DCCSEKJPYGARCH11

*-----*
*           DCC GARCH Fit           *
*-----*

Distribution      : mvt
Model             : DCC(1,1)
No. Parameters    : 21
[VAR GARCH DCC UncQ] : [0+17+3+1]
No. Series        : 2
No. Obs.          : 1000
Log-Likelihood    : 7802.755
Av.Log-Likelihood : 7.8

-----
Optimal Parameters
-----

```

	Estimate	Std. Error	t value	Pr(> t)
[Asset_1].ar1	-0.959214	0.252987	-3.79156	0.000150
[Asset_1].ar2	-0.684478	0.123218	-5.55501	0.000000
[Asset_1].ma1	0.977977	0.252846	3.86787	0.000110
[Asset_1].ma2	0.662852	0.144294	4.59377	0.000004
[Asset_1].ma3	-0.056980	0.037457	-1.52121	0.128206
[Asset_1].omega	0.000001	0.000003	0.56743	0.570423
[Asset_1].alpha1	0.052237	0.025630	2.03811	0.041539
[Asset_1].beta1	0.926046	0.033052	28.01781	0.000000
[Asset_1].shape	7.009306	1.666935	4.20491	0.000026
[Asset_2].ar1	0.926745	0.106676	8.68743	0.000000
[Asset_2].ar2	-0.840904	0.070008	-12.01158	0.000000
[Asset_2].ma1	-0.961916	0.091096	-10.55941	0.000000
[Asset_2].ma2	0.891295	0.058830	15.15033	0.000000
[Asset_2].omega	0.000002	0.000006	0.28757	0.773675
[Asset_2].alpha1	0.089927	0.113693	0.79097	0.428963
[Asset_2].beta1	0.875239	0.136196	6.42631	0.000000
[Asset_2].shape	8.637216	3.606395	2.39497	0.016622
[Joint]dccal	0.028158	0.009424	2.98801	0.002808
[Joint]dccbl	0.946279	0.018321	51.65034	0.000000
[Joint]mshape	5.621092	0.535124	10.50428	0.000000

```
-----
Information Criteria
-----

Akaike      -15.564
Bayes       -15.460
Shibata     -15.564
Hannan-Quinn -15.524

Elapsed time : 2.851176
```

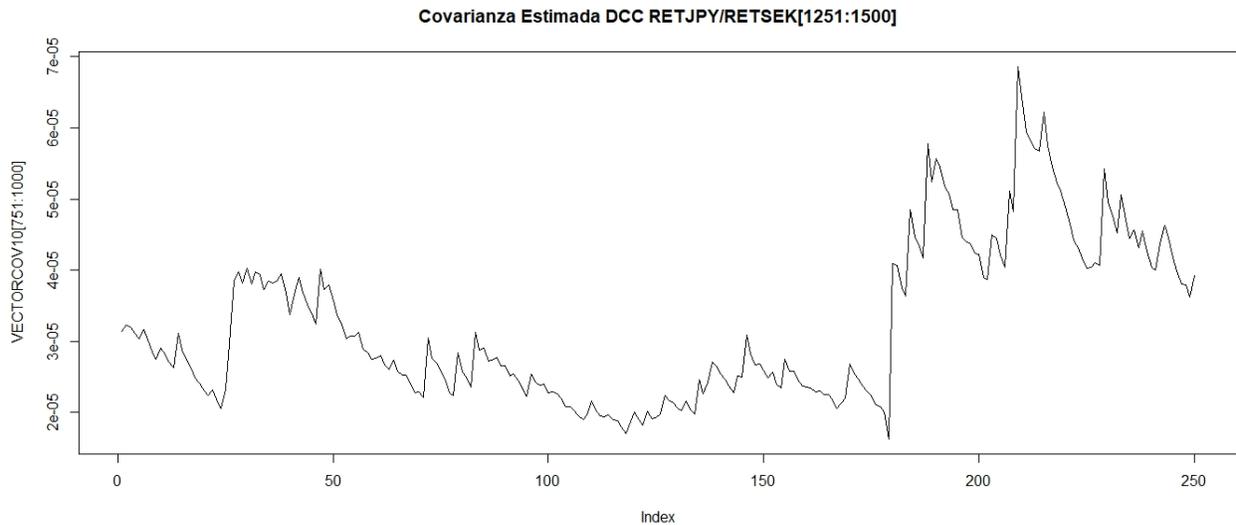


Figura 6.33: Ajuste DCC Serie de Retornos Logarítmicos Yen Japonés y Corona Sueca [501:1500].



Podemos observar que el ajuste DCC, genera una curva menos suave de la estimación del proceso de Volatilidad en la Serie Bivariada de los retornos de la Corona Sueca y el Yen Japonés, a simple vista se podría pensar que describe con mayor exactitud la fluctuación en el cambio de los precios de los activos. Sin embargo en la siguiente sección se comparan los errores de estimación mediante el Error de Raíz Cuadrático Medio, donde se concluye todo lo contrario. El procedimiento adaptativo ajusta mejor en la mayor parte de los casos.

6.2. Resultados de Ajuste por métodos DCC y LCP

6.2.1. Cálculo del MSqE

Sean $R_{1,t}$ y $R_{2,t}$ series de tiempo de retornos logarítmicos. Definimos $\bar{V}_{t,h} \equiv R_{1,t} \cdot R_{2,t}$ como la volatilidad bivariada realizada.

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

Bajo el supuesto de homogeneidad local la predicción constante $\hat{\sigma}_{\{1,2,t\}}^2$ de la volatilidad queda expresada mediante la siguiente expresión en un horizonte de tiempo pequeño h ,

$$V_{t,h}^{LCP} = \hat{\sigma}_{\{1,2,t\}}^2 \equiv h \cdot \text{cov}(\hat{R}_{1,t}, \hat{R}_{2,t}).$$

Sea

$$V_{t,h}^{DCC} \equiv \sum_{k=1}^h \text{cov}(R_{1,t+h/t}, R_{2,t+h/t})^{DCC}$$

la volatilidad estimada mediante en modelo DCC. Definimos el Error de Raíz Cuadrática Media:

$$MSqE = \frac{\sum_{t \in I} |V_{t,h}^{LCP} - \bar{V}_{t,h}|^{\frac{1}{2}}}{\sum_{t \in I} |V_{t,h}^{DCC} - \bar{V}_{t,h}|^{\frac{1}{2}}}$$

Las siguientes tablas contienen los resultados del MSqE dos a dos, considerando seis intervalos de 250 datos cada uno.



CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

SERIES	H	1001:1250	1251:1500	1501:1750	1751:2000	2001:2250	2251:2500
FRANCO SUIZO (CHF) , YEN JAPONÉS (JPY)	h=1	0.8234	0.7468	0.9003	0.8129	0.8387	0.8236
	h=3	1.0915	0.7925	1.0923	1.0993	1.0232	1.0771
	h=5	1.2699	0.8025	1.2159	1.2649	1.1454	1.2229
CORONA SUECA (SEK), YEN JAPONÉS (JPY)	h=1	0.7199	0.5620	0.8345	0.6878	0.7300	0.7135
	h=3	0.9812	0.6960	1.0742	0.9321	0.8954	0.9623
	h=5	1.0165	1.7435	1.2311	0.9969	0.9466	1.0493
CORONA NORUEGA (NOK), YEN JAPONÉS (JPY)	h=1	0.7528	0.6536	0.8982	0.9767	0.7670	0.8546
	h=3	0.9812	0.7767	1.1105	0.9683	0.9963	1.1271
	h=5	1.0620	0.8255	1.2231	0.9706	1.0732	1.2426
CORONA DANESA (DKK), YEN JAPONÉS (JPY)	h=1	0.8178	P.C.DCC.	0.9050	0.9839	0.8273	0.8387
	h=3	1.0784	P.C.DCC	1.1343	0.9743	1.0515	1.1538
	h=5	1.2319	P.C.DCC	1.2545	0.9754	1.2212	1.3290
DÓLAR NEOZELANDÉS (NZD), YEN JAPONÉS (JPY)	h=1	0.9805	0.9370	0.9834	0.9781	0.8189	0.9314
	h=3	1.0054	0.9668	1.0048	1.0118	0.8491	0.9724
	h=5	0.9970	0.9668	1.0246	1.0340	0.8769	0.9722
DÓLAR AUSTRALIANO (AUD), YEN JAPONÉS (JPY)	h=1	1.0092	0.9931	0.9909	0.9997	0.8633	0.9548
	h=3	1.0141	1.0074	0.9999	1.0036	0.8770	0.9847
	h=5	1.0350	1.0576	1.0202	1.0443	0.9050	1.0508
DÓLAR CANADIENSE (CAD), YEN JAPONÉS (JPY)	h=1	1.0002	0.8444	0.9663	0.9373	0.9855	0.9785
	h=3	1.0097	0.8293	0.9661	0.9352	0.9964	0.9751
	h=5	1.0166	0.8030	0.9594	0.9120	0.9958	0.9642
LIBRA ESTERLINA (GBP), YEN JAPONÉS (JPY)	h=1	0.8568	0.7714	0.9564	0.9201	0.9085	0.9225
	h=3	1.1030	0.9245	1.1209	1.1443	1.1290	1.1295
	h=5	1.2562	0.9769	1.2039	1.2265	1.2412	1.2427

Cuadro 6.7: Error de Raíz Cuadrática Media para la Serie Bivariada de los Retornos del Yen Japonés con las 8 combinaciones restantes de los Tipos de Cambio con respecto al Dólar Estadounidense.

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

SERIES	H	1001:1250	1251:1500	1501:1750	1751:2000	2001:2250	2251:2500
CORONA SUECA (NOK), FRANCO SUIZO (CHF)	h=1	0.7633	0.7224	0.8969	0.9838	0.8559	0.8910
	h=3	1.0232	0.7340	1.1397	0.9850	1.0683	1.1760
	h=5	1.1402	0.7290	1.2480	0.9844	1.1542	1.3269
CORONA DANESA (DKK), FRANCO SUIZO (CHF)	h=1	0.8209	0.7468	0.9042	0.9855	0.8447	0.8253
	h=3	1.0952	0.7968	1.0968	0.9676	1.0359	1.0814
	h=5	1.2635	0.8025	1.2074	0.9671	1.1564	1.2163
DÓLAR NEOZELANDÉS (NZD), FRANCO SUIZO (CHF)	h=1	0.9796	0.9048	0.9768	0.9838	0.8573	0.9535
	h=3	0.9933	0.8899	0.9941	0.9923	0.8805	0.9767
	h=5	0.9883	0.8868	1.0153	1.0201	0.9323	1.0180
DÓLAR AUSTRALIANO (AUD), FRANCO SUIZO (CHF)	h=1	0.9705	1.0074	0.9788	0.9885	0.9948	1.0603
	h=3	0.9656	0.9958	0.9749	0.9867	0.9934	1.0689
	h=5	0.9688	0.9939	0.9699	0.9875	1.0120	1.0862
DÓLAR CANADIENSE (CAD), FRANCO SUIZO (CHF)	h=1	0.9987	0.8216	0.9750	0.9451	0.9700	0.9705
	h=3	1.0018	0.7924	0.9755	0.9312	0.9516	0.9663
	h=5	1.0024	0.7304	0.9394	0.8835	0.9035	0.9269
CORONA SUECA (SEK), FRANCO SUIZO (CHF)	h=1	0.7712	0.7877	0.9142	0.8287	0.8422	0.8585
	h=3	1.0185	0.9566	1.1365	1.1390	1.0658	1.1829
	h=5	1.1458	0.9817	1.2388	1.2547	1.1608	1.3485
LIBRA ESTERLINA (GBP), FRANCO SUIZO (CHF)	h=1	0.8905	0.7796	0.9717	0.9758	0.9319	0.9355
	h=3	1.1720	0.9389	1.1601	1.1581	1.1577	1.1734
	h=5	1.2598	0.9754	1.2307	1.2031	1.2172	1.2347

Cuadro 6.8: Error de Raíz Cuadrática Media para la Serie Bivariada de los Retornos del Franco Suizo con las 7 combinaciones restantes de los Tipos de Cambio con respecto al Dólar Estadounidense.

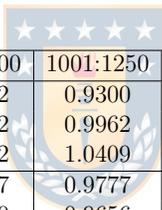
SERIES	H	1001:1250	1251:1500	1501:1750	1751:2000	2001:2250	2251:2500
CORONA DANESA (DKK), CORONA NORUEGA (NOK)	h=1	0.7495	0.6536	0.9046	0.9200	0.7749	0.8555
	h=3	0.9753	0.7767	1.1088	1.2024	1.0101	1.1280
	h=5	1.0547	0.8255	1.2198	1.3707	1.0923	1.2440
DÓLAR NEOZELANDÉS (NZD), CORONA NORUEGA (NOK)	h=1	0.9650	0.8831	0.9757	0.9991	0.7213	0.8367
	h=3	0.9984	0.9275	1.0054	1.0093	0.7173	0.8632
	h=5	0.9972	0.9302	1.0223	1.0124	0.7068	0.8825
DÓLAR AUSTRALIANO (AUD), CORONA NORUEGA (NOK)	h=1	0.9699	0.9834	0.9871	0.9959	0.8415	0.9050
	h=3	0.9831	0.9757	0.9786	0.9998	0.8823	0.9157
	h=5	0.9831	0.9802	0.9787	0.9932	0.8861	0.9425
DÓLAR CANADIENSE (CAD), CORONA NORUEGA (NOK)	h=1	1.0014	0.7999	0.9693	0.9985	0.9167	0.9717
	h=3	1.0118	0.7725	0.9490	1.0085	0.9164	0.9792
	h=5	1.0205	0.7262	0.9463	1.0138	0.8985	0.9860
CORONA SUECA (SEK), CORONA NORUEGA (NOK)	h=1	0.6766	0.7058	0.9039	0.9829	0.6720	0.7799
	h=3	0.8233	0.9151	1.1067	0.9923	0.7969	1.0152
	h=5	0.9029	1.0122	1.2055	1.0002	0.8571	1.1475
LIBRA ESTERLINA (GBP), CORONA NORUEGA (NOK)	h=1	0.8225	0.7562	0.9476	0.9840	0.8945	0.9773
	h=3	1.0724	0.9372	1.1164	0.9978	1.1220	1.1457
	h=5	1.2527	1.0309	1.2328	0.9876	1.2474	1.2371

Cuadro 6.9: Error de Raíz Cuadrática Media para la Serie Bivariada de los Retornos del Corona Noruega con las 6 combinaciones restantes de los Tipos de Cambio con respecto al Dólar Estadounidense.

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

SERIES	H	1001:1250	1251:1500	1501:1750	1751:2000	2001:2250	2251:2500
DÓLAR NEOZELANDÉS (NZD), CORONA DANESA (DKK)	h=1	0.9815	0.9370	0.9865	0.9972	0.8248	0.9360
	h=3	1.0061	0.9668	1.0034	1.0000	0.8564	0.9813
	h=5	0.9964	0.9686	1.0235	1.0177	0.8845	0.9817
DÓLAR AUSTRALIANO (AUD), CORONA DANESA (DKK)	h=1	0.9748	0.9798	0.9752	0.9925	0.9142	0.9740
	h=3	0.9766	0.9754	0.9864	0.9880	0.9433	0.9951
	h=5	1.0016	0.9713	0.9735	0.9922	0.9786	1.0103
DÓLAR CANADIENSE (CAD), CORONA DANESA (DKK)	h=1	1.0008	0.8447	0.9669	0.9938	0.9865	0.9843
	h=3	1.0090	0.8296	0.9667	1.0023	0.9933	0.9768
	h=5	1.0168	0.8034	0.9593	1.0044	0.9896	0.9667
CORONA SUECA (SEK), CORONA DANESA (DKK)	h=1	0.7312	0.7737	0.9212	1.0058	0.8058	0.8398
	h=3	0.9333	1.0182	1.1289	1.0137	1.0138	1.1599
	h=5	1.0358	1.1146	1.2339	1.0130	1.1164	1.2707
LIBRA ESTERLINA (GBP), CORONA DANESA (DKK)	h=1	0.8555	0.7713	0.9587	0.9902	0.9135	0.9235
	h=3	1.0989	0.9244	1.1192	0.9891	1.1293	1.1296
	h=5	1.2521	0.9769	1.1961	0.9797	1.2398	1.2401

Cuadro 6.10: Error de Raíz Cuadrática Media para la Serie Bivariada de los Retornos del Corona Danesa con las 5 combinaciones restantes de los Tipos de Cambio con respecto al Dólar Estadounidense.



Series	H	751:1000	1001:1250	1251:1500	1501:2000	2001:2250	2251:2500
DÓLAR AUSTRALIANO (AUD), DÓLAR NEOZELANDÉS (NZD)	h=1	0.8962	0.9300	0.9450	0.8664	0.5695	0.6664
	h=3	0.9002	0.9962	1.0185	0.9039	0.5260	0.6621
	h=5	0.9152	1.0409	1.0424	0.9533	0.5072	0.6816
DÓLAR CANADIENSE (CAD), DÓLAR NEOZELANDÉS (NZD)	h=1	0.9927	0.9777	0.9917	1.0050	0.8183	0.7743
	h=3	0.9890	0.9656	1.0130	1.0092	0.8358	0.7877
	h=5	0.9851	0.9510	1.0315	1.0188	0.8519	0.8056
LIBRA ESTERLINA (GBP), DÓLAR NEOZELANDÉS (NZD)	h=1	0.9684	0.9670	0.9760	0.9550	0.8925	0.9414
	h=3	1.0011	0.9889	1.0198	1.0097	0.9407	1.0126
	h=5	1.0137	0.9744	1.0324	1.0238	0.9699	1.0308
CORONA SUECA (SEK), DÓLAR NEOZELANDÉS (NZD)	h=1	0.9462	0.9661	0.9917	0.9794	0.7477	0.8650
	h=3	0.9964	1.0046	1.0217	1.0196	0.7864	0.9108
	h=5	1.0101	1.0253	1.0528	1.0473	0.7779	0.9180

Cuadro 6.11: Error de Raíz Cuadrática Media para la Serie Bivariada de los Retornos del Dólar Neozelandés con las 4 combinaciones restantes de los Tipos de Cambio con respecto al Dólar Estadounidense.

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

SERIES	H	1001:1250	1251:1500	1501:1750	1751:2000	2001:2250	2251:2500
DÓLAR CANADIENSE (CAD), DÓLAR AUSTRALIANO (AUD)	h=1	0.9972	0.9789	0.9866	0.9985	0.8727	0.7884
	h=3	0.9951	0.9880	1.0126	1.0199	0.9017	0.8299
	h=5	0.9926	0.9909	1.0297	1.0269	0.9444	0.8424
CORONA SUECA (SEK), DÓLAR AUSTRALIANO (AUD)	h=1	0.9524	1.0067	0.9977	1.0023	0.8585	0.8893
	h=3	0.9766	1.0252	0.9989	0.9950	0.9365	0.9344
	h=5	0.9980	1.0397	1.0033	0.9987	0.9505	0.9574
LIBRA ESTERLINA (GBP), DÓLAR AUSTRALIANO (AUD)	h=1	0.9789	0.9836	0.9958	0.9813	0.9485	0.9595
	h=3	1.0058	1.0060	1.0020	1.0171	1.0041	1.0219
	h=5	1.0130	1.0116	1.0043	1.0313	1.0213	1.0382

Cuadro 6.12: Error de Raíz Cuadrática Media para la Serie Bivariada de los Retornos del Dólar Australiano con las 3 combinaciones restantes de los Tipos de Cambio con respecto al Dólar Estadounidense.

SERIES	H	1001:1250	1251:1500	1501:1750	1751:2000	2001:2250	2251:2500
CORONA SUECA (SEK), DÓLAR CANADIENSE (CAD)	h=1	1.0015	0.9455	1.0185	0.9860	0.8973	0.9442
	h=3	1.0031	0.9138	0.9888	0.9531	0.9283	0.9611
	h=5	1.0166	0.9102	0.9635	0.9492	0.9194	1.0056
LIBRA ESTERLINA (GBP), DÓLAR CANADIENSE (CAD)	h=1	1.0009	0.9676	1.0022	1.0047	1.0254	1.0061
	h=3	0.9983	0.9535	0.9940	1.0024	1.0353	0.9992
	h=5	0.9975	0.9495	0.9927	1.0106	1.0349	1.0048

Cuadro 6.13: Error de Raíz Cuadrática Media para la Serie Bivariada de los Retornos del Dólar Canadiense con las 2 combinaciones restantes de los Tipos de Cambio con respecto al Dólar Estadounidense.

SERIES	H	1001:1250	1251:1500	1501:1750	1751:2000	2001:2250	2251:2500
CORONA SUECA (SEK), LIBRA ESTERLINA (GBP)	h=1	0.8075	0.8296	0.9681	0.9132	0.8842	0.9370
	h=3	1.0775	1.0416	1.1588	1.1654	1.1444	1.1797
	h=5	1.1787	1.1483	1.2538	1.2805	1.2692	1.2936

Cuadro 6.14: Error de Raíz Cuadrática Media para la Serie Bivariada de los Retornos de la Libra Esterlina con la Corona Sueca con respecto al Dólar Estadounidense.

Lo primero que queremos notar es que en el caso de modelos DCC para las series de retornos bivariadas JPY y DKK el método DCC generó un problema de convergencia a pesar de haber ajustado los modelos ARMA y GARCH paso a paso verificando todas las condiciones necesarias en su ajuste, por otro lado el modelo LCP multiescala no generó ningún conflicto en su implementación.

En cuánto al rendimiento de los métodos mencionamos que con un retorno agregado el método LCP es muy superior al DCC, siendo mejor en un 97, 54 % de los casos, considerando

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

3 retornos el rendimiento del LCP sigue siendo superior 53,95 %, pero en el caso de 5 retornos el ajuste DCC es superior con un 55,18 % de los casos.

En el intervalo de [1251 : 1500] datos el ajuste LCP es muy superior donde por ejemplo en el caso JPY/SEK se obtuvo un error medio absoluto de 0.5620, 0.6960 y 0.7435 para 1,3 y 5 retornos agregados respectivamente.

En el intervalo de los datos [1001:1250] el rendimiento que obtuvo LCP fue de un 53,70 %, en el segundo intervalo [1251:1500] es de 82,85 %, en el tercer intervalo es de 50,93 %, en el cuarto intervalo es de 59,35 %, en el quinto de un 71,29 %, finalmente en el intervalo de [2251:2500] es de un 61,11 %.

En los siguientes gráficos tenemos la distribución de la covarianza real, la estimada mediante el método LCP y el modelo DCC ajustado.

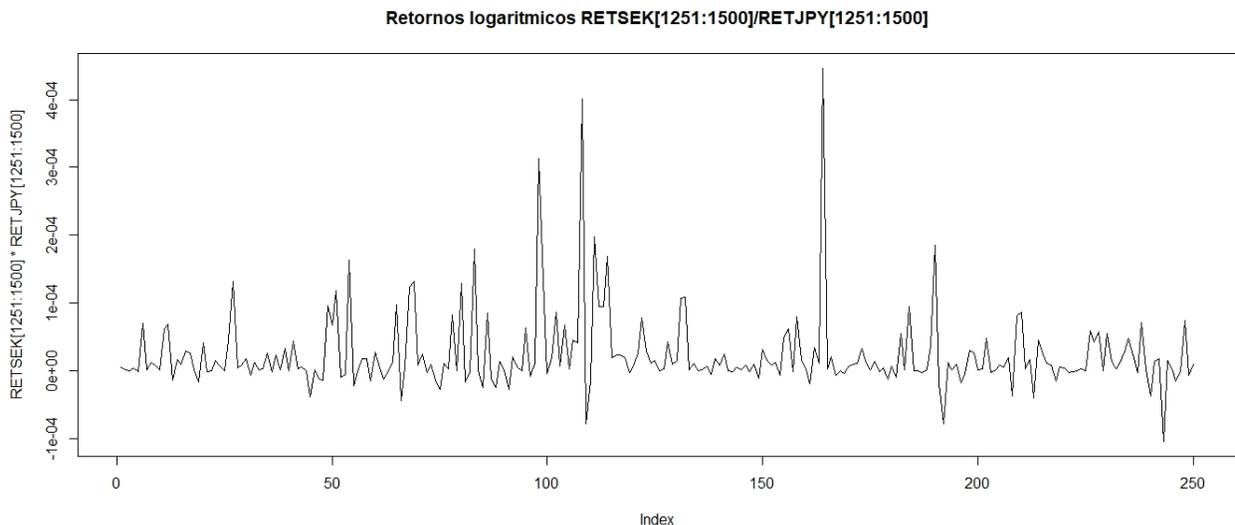


Figura 6.34: Datos Reales Serie de Retornos Logarítmicos Yen Japonés y Corona Sueca [1251:1500].

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

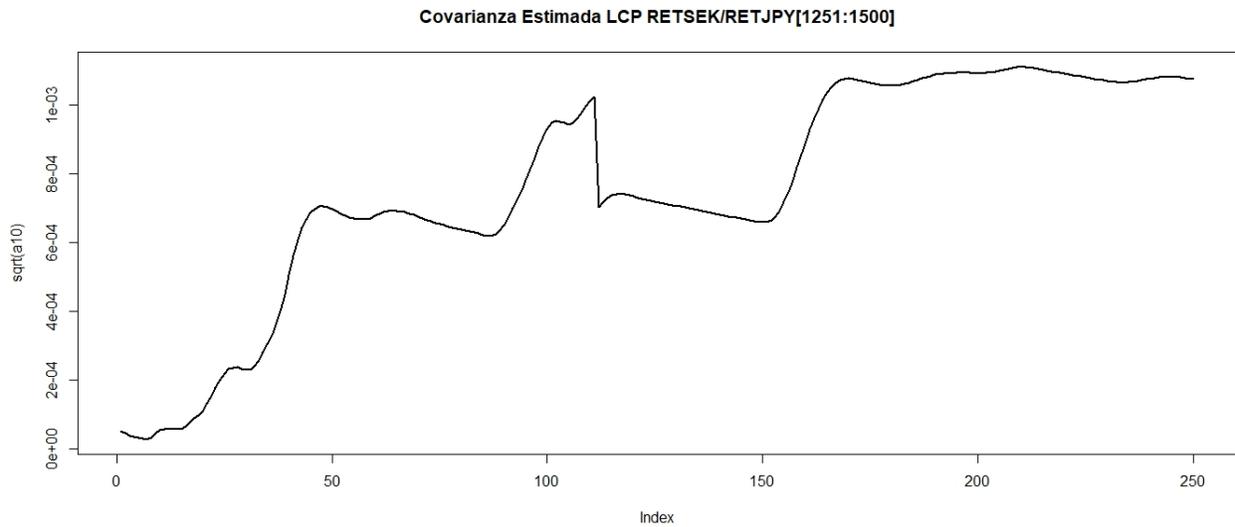


Figura 6.35: Ajuste LCP Serie de Retornos Logarítmicos Yen Japonés y Corona Sueca [1251:1500].

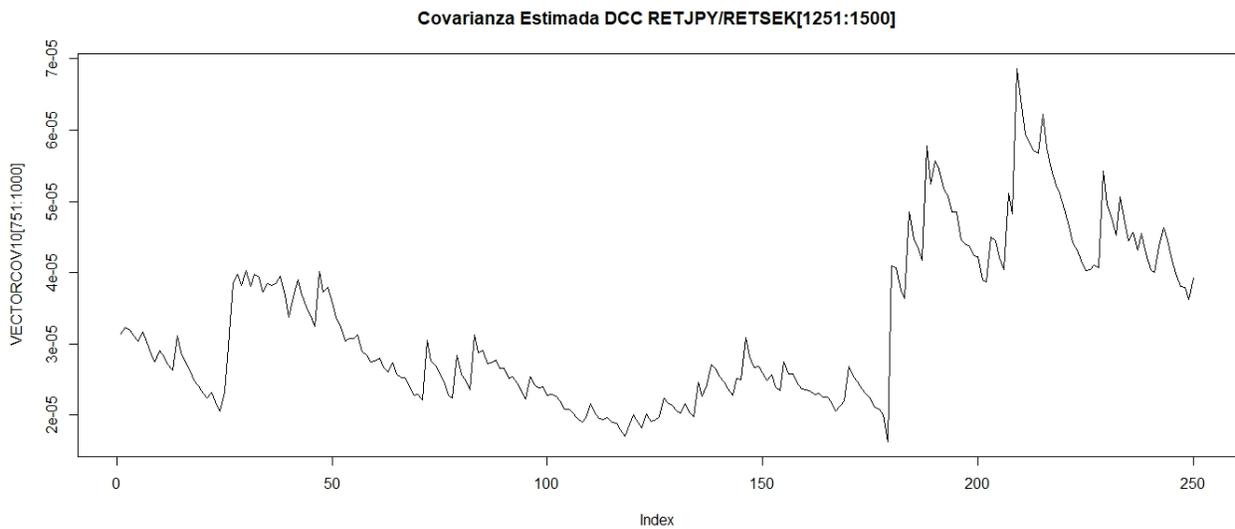


Figura 6.36: Ajuste DCC Serie de Retornos Logarítmicos Yen Japonés y Corona Sueca [1251:1500].

Al comparar los dos gráficos obtenidos de la estimación de la volatilidad, se puede

CAPÍTULO 6. APLICACIÓN A LA ESTIMACIÓN Y PREDICCIÓN EN LOS TIPOS DE CAMBIO DE DIVISAS

observar que el método DCC genera una estimación con una mayor fluctuación, pareciera ser más sensible a las perturbaciones del mercado, sin embargo ya vimos que el ajuste es mejor en el método local adaptativo según lo estudiado en el error cuadrático medio. El método LCP, realiza un ajuste más suave, pero aún así percibe las fluctuaciones o cambios fuertes en el proceso de la volatilidad.

Queremos mencionar que el procedimiento adaptativo desarrollado se puede aplicar en el contexto del proceso de exportación, ya que se buscará asegurar el mejor tipo de cambio, es decir el con menor volatilidad para que las transacciones sean menos riesgosas.



Capítulo 7

Resultados de Álgebra Lineal y

Notaciones



Sea $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$ una matriz cuadrada de tamaño $n \times n$.

Matriz Traspuesta

Se define como matriz traspuesta al resultado de reordenar la matriz original mediante el cambio de filas por columnas y las columnas por filas en una nueva matriz.

$$A^{\top} = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \cdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix}$$

Propiedades de la Traspuesta

1. $(A^{\top})^{\top} = A$
2. $(A + B)^{\top} = A^{\top} + B^{\top}$
3. $(cA)^{\top} = cA^{\top}$
4. $(AB)^{\top} = B^{\top}A^{\top}$



Determinante de una matriz

Se define el determinante de la matriz A con notación $|A|$, de manera inductiva:

Si $n = 1$, $|A| = a_{11}$.

Si $n = 2$, $|A| = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}$

Si $n = 3$ el determinante será $a_{11}|A_{11}| - a_{12}|A_{12}| + a_{13}|A_{13}|$ donde $|A_{ij}|$ corresponde al determinante que queda de eliminar la fila i y la columna j de la matriz.

De manera general, podemos llegar a la siguiente expresión algebraica:

$$|A| = \sum_{j=1}^n (-1)^{1+j} a_{1j} |A_{1j}|.$$

Esto corresponde al desarrollo del determinante mediante la columna j

De manera análoga se puede desarrollar mediante la fila i , obteniendo:

$$|A| = \sum_{i=1}^n (-1)^{i+1} a_{i1} |A_{i1}|.$$

Propiedades del Determinante

1. $|A|^T = |A|$
2. $|A \cdot B| = |A| \cdot |B|$
3. $|A|^{-1} = \frac{1}{|A|}$
4. El determinante de una matriz con alguna fila o columna de ceros es 0
5. Si una matriz tiene filas o columnas linealmente dependientes, entonces su determinante es 0.
6. Si se cambia el orden de n filas o columnas, el determinante cambia de signo si n es impar.



Traza de una Matriz

Dada una matriz Cuadrada A de tamaño $n \times n$, la traza de la Matriz se define como la suma de los elementos de la diagonal principal.

$$tr(A) = a_{11} + a_{22} + \dots + a_{nn}.$$

Propiedades de la traza de una matriz

1. $tr(A + B) = tr(A) + tr(B)$
2. $tr(A^T) = tr(A)$

3. $tr(AB) = tr(BA)$

4. $tr(ABC) = tr(CAB) = tr(BCA)$

Derivada de una función escalar con respecto a una matriz

Sea y un escalar y X una matriz de orden $p \times q$:

$$\frac{\partial y}{\partial X} = \begin{pmatrix} \frac{\partial y}{\partial x_{11}} & \frac{\partial y}{\partial x_{21}} & \cdots & \frac{\partial y}{\partial x_{p1}} \\ \frac{\partial y}{\partial x_{12}} & \frac{\partial y}{\partial x_{22}} & \cdots & \frac{\partial y}{\partial x_{p2}} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial y}{\partial x_{1q}} & \frac{\partial y}{\partial x_{2q}} & \cdots & \frac{\partial y}{\partial x_{pq}} \end{pmatrix}$$



Algunas propiedades importantes de la derivada de una matriz

1. $\frac{\partial}{\partial X}(a^T X^{-T} b) = -X^{-t} a b^T X^{-T}$

2. $\frac{\partial}{\partial X} \log |X| = X^{-T}$

Lema 7.1 : Si A una matriz cuadrada e invertible entonces $|A|' = |A|tr(A)$.

Demostración:

Sean f y g las funciones definidas por:

$$g : M_{d \times d}(\mathbb{R}) \rightarrow \mathbb{R}$$

$$A \mapsto g(A) = |A|$$

$$f : \mathbb{R} \rightarrow \mathbb{R}$$

$$x \mapsto f(x) = |xI_d + B|$$

Del álgebra lineal sabemos que la función f define el polinomio característico de $-B$, observamos que dicho polinomio es mónico, por ende su coeficiente principal es 1 y el coeficiente que acompaña al término $n - 1$ es $tr(B)$.

De manera general, para cualquier matriz C invertible, su polinomio característico se puede escribir de la siguiente forma:



$$P_C(x) = \sum_{k=0}^n x^{n-k} (-1)^k tr(\bigwedge^k C)$$

donde,

$$tr(\bigwedge^k C) = \frac{1}{k!} \begin{vmatrix} tr(C) & k-1 & 0 & \dots \\ tr(C^2) & tr(C) & k-2 & \dots \\ \vdots & \vdots & \ddots & \vdots \\ tr(C^k) & tr(C^{k-1}) & \dots & tr(C) \end{vmatrix}.$$

Reemplazando $x = 1$ en la función f , tenemos que, $f(1) = 1 + tr(B) + \mathcal{O}(\|B\|^2)$.

Por otro lado, $|A + H| = |A| |I_d + A^{-1}H|$, luego

$$|A + H| = |A| + |A|tr(A^{-1}H) + \mathcal{O}(\|H\|^2).$$

En particular, concluimos que $|A'| = |A|tr(A^{-1})$.

Lema 7.2 : Si D es una matriz definida positiva el máximo de la función:

$$f(G) = -N\log|G| - tr(G^{-1}D)$$

con respecto a las matrices definidas positivas G existe y ocurre en $G = \frac{1}{N}D$

Demostración:

Sea $D = EE^T$ y $H = EH^{-1}E^T$ entonces $G = EH^{-1}E^T$.

Por lo tanto, $|G| = |EH^{-1}E^T| = |E||H^{-1}||E^T| = |H^{-1}||EE^T| = |H^{-1}||D| = \frac{|D|}{|H|}$.

Además, $tr(G^{-1}D) = tr(G^{-1}EE^T) = tr(E^T G^{-1}E) = tr(H)$.

Luego la función a maximizar es:

$$f = -N\log\left(\frac{|D|}{|H|}\right) - tr(H) = -N\log(|D|) + N\log|H| - tr(H)$$

Sea $H = TT^T$ donde T una matriz triangular inferior (Descomposición de Cholesky).

$$\begin{aligned} f &= -N\log(|D|) + N\log|T|^2 - tr(TT^T) \\ &= -N\log(|D|) + \sum_{i=1}^p (N\log t_{ii}^2 - t_{ii}^2) - \sum_{i>j} t_{ij}^2 \end{aligned}$$

El máximo de esta función ocurre cuando $t_{ii}^2 = N$ y $t_{ij} = 0$, $i \neq j$.

Veamos que efectivamente es un máximo,

Sea $x \in \mathbb{R}$, $f(x) = N \log x^2 - x^2$, cuya derivada es $f'(x) = \frac{N2x}{x^2} - 2x$ entonces el punto crítico se alcanza en $N = x^2$.

La segunda derivada de la función $f''(x) = -N/x^2 - 2$ es siempre negativa, por ende $N = x^2$ es un punto de máximo.

Lema 7.3 : Sea Y_1, Y_2, \dots, Y_n una muestra aleatoria simple de vectores independientes e idénticamente distribuidos tal que $Y_i \sim f_{\Sigma^*}$, entonces:

$$\mathbb{E}_{\Sigma^*} [e^{L(\Sigma, \Sigma^*)}] = 1$$

Donde $L(\Sigma, \Sigma^*)$ es la razón de log-verosimilitud.

Demostración:



$$\begin{aligned} \mathbb{E}_{\Sigma^*} [e^{L(\Sigma, \Sigma^*)}] &= \int_{\mathbb{R}^n} \prod_{i=1}^n \frac{f_{\Sigma}(Y_i)}{f_{\Sigma^*}(Y_i)} \cdot \prod_{i=1}^n f_{\Sigma^*}(Y_i) \partial Y_i \\ &= \int_{\mathbb{R}^n} \prod_{i=1}^n f_{\Sigma}(Y_i) \partial Y_i \\ &= \prod_{i=1}^n \int_{\mathbb{R}^n} f_{\Sigma}(Y_i) \partial Y_i \\ &= 1 \end{aligned}$$

Lema 7.4 : La derivada de la esperanza de la divergencia de Kullback-Leibler en un entorno del parámetro θ es cero.

Demostración:

$$\begin{aligned} &\frac{\delta}{\delta \theta} \mathbb{E}_{\theta} [l(y, \theta) - l(y, \theta')] \\ &= \mathbb{E}_{\theta} \left[\frac{\delta}{\delta \theta} (l(y, \theta) - l(y, \theta')) \right] \end{aligned}$$

$$\begin{aligned} &= \mathbb{E}_\theta\left[\frac{1}{p(y, \theta)} p'(y, \theta)\right] \\ &= \int \frac{1}{p(y, \theta)} p'(y, \theta) p(y, \theta) \delta\mu_0(y) \\ &= \int p'(y, \theta) \delta\mu_0(y) \\ &= \frac{\delta}{\delta\theta} \left(\int p(y, \theta) \delta\mu_0(y) \right) \\ &= \frac{\delta}{\delta\theta} (1) \\ &= 0 \end{aligned}$$



Capítulo 8

Anexos



Código R para la Estimación LCP mediante el paquete AWS

```
library(aws)
library(awsMethods)
library(tseries)
Datos=read.table(file.choose(),h=T, sep=" ", dec = ".", na.strings = "NA",check.names = TRUE)
attach(Datos)
Date=Datos[,1]
AUD=Datos[,2]
ts.AUD=na.exclude(AUD)
RETAUD=diff(log(ts.AUD))
GBP=Datos[,4]
ts.GBP=na.exclude(GBP)
RETGBP=diff(log(ts.GBP))
x1=awsdata(aws(data.frame(RETGBP[1001:1251],RETAUD[1001:1251])), "theta")
```

```
a1=c()
for(i in 2:251)a1[i-1] =(cov(x1[1:i,1:2,1])[2])
```

Código R para el Ajuste Bi-GARCH

Carga de Librerías Necesarias:

```
library(tseries)
library(stats)
library(rugarch)
library(rmgarch)
library(parallel)
library(forecast)
library(lmtest)
library(zoo)
```



Carga de Datos en el Software R-Project:

```
Datos=read.table(file.choose(),h=T, sep="," , dec = ".", na.strings = "NA",check.names = TRUE)
attach(Datos)
Date=Datos[,1]
```

Gráficos de Autocorrelación y Autocorrelación Parcial:

```
JPY=Datos[,6]
ts.JPY=na.exclude(JPY)
RETJPY=diff(log(ts.JPY))
acf(RETJPY[501:1500],main="Retornos Logaritmicos JPY/USD",lag.max=100)
pacf(RETJPY[501:1500],main="Retornos Logaritmicos JPY/USD",lag.max=100)
SEK=Datos[,9]
```

```
ts.SEK=na.exclude(SEK)
RETSEK=diff(log(ts.SEK))
acf(RETSEK[501:1500],main="Retornos Logaritmicos SEK/USD",lag.max=100)
pacf(RETSEK[501:1500],main="Retornos Logaritmicos SEK/USD",lag.max=100)
```

Ajuste Modelo ARMA:

```
ARMAJPY22=arima(RETJPY[501:1500],order=c(2,0,2))
coeftest(ARMAJPY22)
tsdiag(ARMAJPY22)
Box.test(residuals(ARMAJPY22),type="Ljung-Box",lag=4)
plot(residuals(ARMAJPY22),main="Residuals ARMAJPY(2,2)[501:1500]")
plot(residuals(ARMAJPY22)^2, main="Square Residuals ARMAJPY(2,2)[501:1500]")
ARMASEK23=arima(RETJPY[501:1500],order=c(2,0,3))
coeftest(ARMASEK23)
tsdiag(ARMASEK23)
Box.test(residuals(ARMASEK23),type="Ljung-Box",lag=4)
plot(residuals(ARMASEK23),main="Residuals ARMASEK(2,3)[501:1500]")
plot(residuals(ARMASEK23)^2, main="Square Residuals ARMASEK(2,3)[501:1500]")
```

Aplicación ARCH-TEST:

```
ArchTest(residuals(ARMAJPY22),lag=3)
ArchTest(residuals(ARMASEK23),lag=3)
```

Ajuste Modelo GARCH(p,q):

```
GARCHJPY=ugarchspec(mean.model=list(armaOrder=c(2,2)),
variance.model=list(garchOrder=c(1,1)),distribution.model="std")
FITGARCHJPY=ugarchfit(data=RETJPY[501:1500],spec=GARCHJPY)
GARCHSEK=ugarchspec(mean.model=list(armaOrder=c(2,3)),
variance.model=list(garchOrder=c(1,1)),distribution.model="std")
FITGARCHSEK=ugarchfit(data=RETSEK[501:1500],spec=GARCHSEK)
```

Código R para ajuste DCC

```
MODELJPYj=ugarchspec(mean.model = list(armaOrder = c(2,2), include.mean=F),
variance.model = list(garchOrder = c(1,1),model = "sGARCH"),distribution.model="std")
MODELJPY=ugarchspec(mean.model = list(armaOrder = c(2,2), include.mean=F),
variance.model = list(garchOrder = c(1,1),model = "sGARCH"),distribution.model="std")
DATOSSEKJPY=cbind(RETSEK[501:1500],RETJPY[501:1500])
uspec = c(MODELSEK,MODELJPY)
DCCSEKJPY=dccspec(uspec=multispec(uspec),dccOrder=c(1,1),distribution="mvt")
DCCSEKJPYGARCH11=dccfit(DCCSEKJPY, data = DATOSCHFJPY)
VECTORCOV=(DCCSEKJPYGARCH11@mfit$H[1:1,2:2,]
plot(VECTORCOV[751:1000], type="l",main= "Covarianza Estimada DCC")
```

Caso Particular Volatilidad Si $\Sigma =$

$$\begin{pmatrix} \theta_1 & 0 & \dots & 0 \\ 0 & \theta_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \theta_d \end{pmatrix}$$

Por otro lado,

$$|\Sigma|' = \begin{vmatrix} 1 & 0 & \dots & 0 \\ 0 & \theta_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \theta_d \end{vmatrix} + \begin{vmatrix} \theta_1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \theta_d \end{vmatrix} + \dots + \begin{vmatrix} \theta_1 & 0 & \dots & 0 \\ 0 & \theta_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{vmatrix}$$

luego,

$$\frac{|\Sigma|'}{|\Sigma|} = \frac{1}{\theta_1} + \frac{1}{\theta_2} + \dots + \frac{1}{\theta_d} = \text{tr}(\Sigma^{-1})$$



Capítulo 9

Conclusiones



A lo largo de este estudio se señaló la importancia de la relación entre riesgo financiero, rendimiento y volatilidad para poder tomar decisiones por parte de los agentes de inversión. Los dos principales cuantificadores del riesgo son la desviación estándar y el coeficiente de correlación entre las variables.

El coeficiente de correlación se utiliza en general para definir una cartera de inversión eficiente, mediante el proceso de diversificación lo que genera una minimización del riesgo. Para esto se deben escoger acciones con correlación negativa.

Por otro lado, la desviación estándar o la volatilidad de la serie de retornos es el principal indicador de riesgo, por esta razón se busca modelar y predecir su comportamiento en un horizonte de tiempo definido. Es por esto que el procedimiento Adaptativo Local multivariado es un enfoque interesante, debido a que se adapta automáticamente a los datos, no se requieren supuestos restrictivos de manera teórica que no puedan ser comprobados por datos reales.

CAPÍTULO 9. CONCLUSIONES

En el caso práctico presentado con series de tiempo bivariadas de tipos de cambio de nueve monedas al Dólar Estadounidense se realizaron predicciones con horizontes de 1, 3 y 5 retornos. El método ajustó con gran exactitud sobretodo para 1 y 3 retornos, no se generó ningún problema de convergencia.

En este contexto, el estudio de la volatilidad bivariada de los tipos de cambio mediante el método el procedimiento local adaptativo(LCP) es más conviene que el procedimiento clásico de Dinamic Conditional Correlation (DCC) en varios aspectos. Destacó su simplicidad en la implementación numérica ya que mediante el paquete de R, Adaptive Weights Smoothing, desarrollado por Joerg Polzehl, es muy sencilla de llevar a cabo. Por otro lado, ajustar el modelo DCC fue una ardua tarea, ya que primero se ajustó un modelo ARMA para la media de cada retorno logaritmico de los tipos de cambio y luego el correspondiente GARCH(1,1), posteriormente se tuvo que ajustar el DCC para cada par bivariado. Lo que aumentó considerablemente el tiempo de análisis. En el estudio se realizaron predicciones utilizando intervalos de 250 datos, en específico se realizó el procedimiento y ajuste 6 veces para cada par de datos bivariados, donde además de la no optimización del tiempo y recursos computacionales, algunos modelos cambiaban de manera drástica al avanzar en el tiempo, por ejemplo en el dólar australiano que pasa de ser modelado por un GARCH(1,1)+ARMA(2,2), un GARH(1,1)+ARMA(4,2) a un GARCH(1,1)+ARMA(1,2), etc.

Otra desventaja en el DCC es la necesidad en algunas ocasiones de tener que diferenciar alguna de las series que componen, esa situación nos puede llevar a una pérdida de precisión en la estimación. En cuánto al cálculo del MSqE en el horizonte de tiempo de 1 día, el enfoque LCP fue superior al DCC en el 97,54 % de los casos, al igual que en un horizonte de 3 días, donde fue superior en un 53,95 %. Sin embargo, al considerar 5 días el rendimiento del enfoque LCP disminuye a un 44,19 %.

CAPÍTULO 9. CONCLUSIONES

Cabe destacar que ésta desventaja en el horizonte de 5 días es muy fácil de solucionar aplicando el método LCP nuevamente teniendo la certeza que una predicción a un día es realmente óptima, mencionamos que éste es el precio a pagar por la falta de supuestos acerca de los datos.



Capítulo 10

Futuros Estudios



En el proyecto se desarrolló la teoría necesaria para aplicar el modelo local adaptativo para datos multivariados y en la aplicación se abordó el ajuste para datos bivariados con ayuda del paquete AWS, que está diseñado para el análisis de datos en una, dos y tres dimensiones, por lo tanto si se quiere estimar volatilidad para datos de una dimensión mayor se deberá desarrollar el código para este menester, lo que quedará abierto para una futura investigación.

También puede ser de interés para los investigadores aplicar el método LCP multivariado para seleccionar carteras de inversión eficientes utilizando la correlación serial entre los distintos retornos.

Una última línea de investigación que nos gustaría mencionar es comparar el método LCP multivariado con otro método no paramétrico muy en boga hoy en día; el uso de las Redes Neuronales Artificiales.

Bibliografía

- [1] Almeida, D. Hotta, L. and y Ruiz, E. (2015). *MGARCH MODELS: Tradeoff Between Feasibility and Flexibility*. [Manuscrito presentado para publicación]. Departamento de Estadística. Universidad Carlos III de Madrid.
- [2] Akaike, H. (1973). *Information Theory and an Extension of the Maximum Likelihood Principle*. Second International Symposium on Information Theory, (Tsahkadsor, 1971), Budapest, 267-281.
- [3] Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*. 31(3), 307-327.
- [4] Carrol, R. and Ruppert, D. (1988). *Transformation and Weighting in Regression*. New York: Chapman and Hall.
- [5] Christie, A. (1982). The Stochastic Behavior of Common Stock Variance-Value, Leverage and Interest Rate Effects. *Journal of Financial Economics*, 10(4),407-432.
- [6] Engle, R. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of Unites Kingdom Inflation. *Econometrica*, 50(4),987-1007.
- [7] Engle, R. and Bollerslev, T. (1986). Modelling the Persistence of Conditional Variances. With comments and a reply by the authors. *Econometric Reviews*, 5(1),1-87.
- [8] Bollerslev, T. and Engle, R. (1993). Common Persistence in Conditional Variances. *Econométrica*, 61(1),167-186.
- [9] Engle, R. and NG, K. (1993). Measuring and Testing the Impact of News on Volatility. *Journal of Finance*, 48(5),1749-1778.

BIBLIOGRAFÍA

- [10] Fama, E. and French, K. (2010). Luck versus skill in the cross-section of mutual fund returns. *The Journal of Finance*, 65(5),1915-1947.
- [11] Fan, J., Jiang, J., Zhang, C. and Zhou, Z. (2003). Time-dependent diffusion models for term structure dynamics. Statistical applications in financial econometrics. *Satist Sinica*, 13,965-992.
- [12] Figlewski, S. (1997). Forecasting Volatility. *Financial markets, Institutions and Instruments*, 6(2),1-88.
- [13] Fisher, B. (1976). The Pricing of Commodity Contracts. *Journal of Financial Economics*, 3(1-2), 167-179.
- [14] Flynn, C., Hurvich, C. and Simonof, J. (2013). Efficiency for Regularization Parameter Selection in Penalized Likelihood Estimation of Misspecified. *Journal of the American Statistical Association*, 108(503),1031.1043.
- [15] Francq, C. and Zakoian, J. (2010). *GARCH models, Structure, Statistical inference and Financial Applications*. Chichester:John Wiley & Sons.
- [16] Gitman, L. and Zutter, C. (2009). *Principio de Administración Financiera*. Pearson.
- [17] Gitman, L y Joehnk, M, (2009). *Fundamentos de Inversiones*. Pearson.
- [18] Gutiérrez, R., Ortiz, E. and y García, O. (2019). Long-term effects of the asymmetry and persistence of the prediction of volatility: Evidence for the equity markets of Latin America. *Contaduría y Administración*, 62(4),1081-1099.
- [19] Green, J. and Silvermann, W. (2019). *Nonparametric Regression and Generalized Linear Models, A ROUGHNESS PENALTY APPROACH*. London: Chapman & Hall.
- [20] Gouriéroux, C. (1997). *ARCH Models and Finanacial Applications*. New York: Springer-Verlag.
- [21] Kwan, C., Li, W. y Ng, K. (2005). Kwan, CK, Li, WK y Ng, K. Un modelo GARCH de umbral multivariante con correlaciones variables en el tiempo. *Revisiones econométricas*, 24,1-33.
- [22] Marquéz M. (2002). *Modelo SETAR aplicado a la volatilidad de la rentabilidad de las acciones: algoritmos para su identificación*. Trabajo de Grado Doctoral, Universidad Politécnica de Cataluña. Cataluña.
- [23] Mercurio, D y y Spokoiny, V. (2004). Statistical Inference for Time-Inhomogeneous

BIBLIOGRAFÍA

- Volatility Models. *Annals of Statistics*, 32(2),577-602.
- [24] *Morgan, J.* (1996). *RiskMetrics, Technical Document*. New York: Morgan Guaranty Trust Company.
- [25] *Nadaraya, E.* (1989). *Nonparametric Estimation of Probability Densities and Regression Curves*. Dordrecht: Kluwer Academic Publishers.
- [26] *Nelson, D.* (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, 59(2),347-370.
- [27] *Novalés A.* (2017). *Midiendo el Riesgo en Mercados Financieros*. Departamento de Economía Cuantitativa. Universidad Complutense.
- [28] *Pagan, A. and Schwert, W.* (1990). Alternative Models for Conditional Stock Volatility. *Journal of Econometrics*, 45,267-290.
- [29] *Parzen, É.* (1962). On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, 33,1065-1076.
- [30] *Robles, M.* (2002). Medidas de la Volatilidad. *Revista Española de Financiación y Contabilidad*, 31(114),1073-1110.
- [31] *Rosenblatt, M.* (1956). Remarks on some nonparametric estimates of a density function. *Annals of Mathematical Statistics*, 27,832-837.
- [32] *Scott, D.* (1985). Frequency polygons: theory and application. *Journal of the American Statistical Association*, 80(390), 348-354.
- [33] *Serdyukova, N.* (2012) Spatial Adaptation in Heterocedastic Regression: Propagation Approach. *Electronic Journal of Statistics*, 6,861- 907.
- [34] *Spokoiny, V.* (2009). Multiscale Local Change Point Detection with Applications to Value-At-Risk. *Annals of Statistics*, 37(3),1405-1436.
- [35] *Spokoiny V. and Dickhaus T.* (2015). *Regression Estimate, Basic of modern mathematical statistics*. Heidelberg: Springer.
- [36] *Tsay, R.* (2005). *Analysis of Financial Time Series*. Hoboken: John Wiley & Sons.
- [37] *Van Horne J. and Wachowicz J.* (2010). *Fundamentos de Administración Financiera*. Madrid: Pearson.

BIBLIOGRAFÍA

- [38] Wasserman, L. (2006). *All of Nonparametric Statistics*. New York: Springer.
- [39] Zakoian, J-M. (1994). Threshold Heterocedasticity Models. *Journal of Economic Dynamics and Control*, 18(5),931-955.
- [40] Zambom, A. and Dias, R. (2012)..^A *Review of Kernel Density Estimation with Applications to Econometrics.*”,arXiv, 2017, arxiv.org/pdf/1212.2812v1.pdf. Accessed 12 June 2012. Preprint.



Bibliografía

