

Procesos de Larga Memoria Localmente Estacionarios

Por
Ricardo Alonso Olea Ortega

Tesis sometida como requisito para optar al grado de

Doctor en Estadística

Pontificia Universidad Católica de Chile,

Santiago, Chile

Agosto 2008

PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE

DEPARTAMENTO DE ESTADÍSTICA

Los que abajo firman certifican haber leído y recomiendan a la Facultad de Matemáticas la aceptación de la Tesis titulada “**Procesos de Larga Memoria Localmente Estacionarios**” de **Ricardo Olea Ortega** como requerimiento para optar al grado de **Doctor en Estadística**.

29 de Agosto 2008.

Supervisor de Tesis: _____

Wilfredo Palma

Comité Examinador: _____

Reinaldo Arellano

Manuel Galea

Mauricio Zevallos

PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE

Fecha: 29 de Agosto 2008

Autor: **Ricardo Olea Ortega**

Título: **Procesos de Larga Memoria**

Localmente Estacionarios

Departamento de: **Estadística**

Grado: **Doctor** Convocación: **Agosto** Año: **2008**

Se le concede permiso a la para hacer circular y copiar, con propósitos no comerciales, el título ante dicho para los requerimientos de individuos y/o instituciones.

Firma del Autor

EL AUTOR SE RESERVA LOS DERECHOS DE OTRAS PUBLICACIONES, Y NI LA TESIS NI EXTRACTOS EXTENSOS DE ELLA, PUEDEN SER IMPRESOS O REPRODUCIDOS SIN EL PERMISO ESCRITO DEL AUTOR.

EL AUTOR ATESTIGUA QUE EL PERMISO SE HA OBTENIDO PARA EL USO DE CUALQUIER MATERIAL COPYRIGHTED QUE APAREZCA EN ESTA TESIS (CON EXCEPCIÓN DE LOS BREVES EXTRACTOS QUE REQUIEREN SOLAMENTE EL RECONOCIMIENTO APROPIADO EN LA ESCRITURA DEL ESTUDIANTE) Y QUE TODO USO ESTÉ RECONOCIDO CLARAMENTE.

*A mi madre, padre, hermanos y a mi amada
esposa Patricia.*

Índice general

1. Procesos Localmente Estacionarios	1
1.1. Introducción	1
1.2. Proceso Gaussiano Localmente Estacionario	4
1.2.1. Ruido Fraccionario Localmente Estacionario (LS-FN)	5
1.2.2. Proceso LS-ARFIMA	6
2. Métodos de Simulación	8
2.1. Expansión $MA(\infty)$ truncada	9
2.2. Algoritmo de Innovación	10
2.3. Aproximación Algoritmo de Innovación paquete “fracdiff”	10
2.4. Comparación	11
2.4.1. Simulación LS-FN	11
2.4.2. Varianza Teórica vs. Muestral	13
3. Estimador de Whittle	15

3.1. Definición	15
3.2. Teoría Asintótica	16
3.2.1. Supuestos	17
3.2.2. Teoremas y Principales Resultados	19
3.2.3. Ejemplos	20
3.3. Simulación	25
3.3.1. Caso Lineal	26
3.3.2. Caso Armonico	31
4. Estimador de Kalman	35
4.1. Representación de Espacio Estado	35
4.2. Estimación	37
4.3. Simulación	39
5. Ilustración	41
5.1. Espeleotema	43
5.2. Anillos de árbol	50
6. Conclusiones	56
A. Apéndice Técnico	58
A.1. Proposiciones	59
A.2. Lemas	82

A.3. Teoremas	93
B. Programas	99
B.1. Simuladores	99
B.1.1. lsfm.trunc.sim	99
B.1.2. lsfm.innov.sim y kapp	100
B.1.3. lsfm.fracdiff.sim	100
B.2. Estimador Whittle	102
B.2.1. lsfm.whittle	102
B.2.2. whittle.taper.loglik.new.1	105
B.2.3. taper	106
B.2.4. fn.density	106
B.2.5. var.nonstat.3	106
B.2.6. var.beta y var.alpha	107
B.2.7. whittle.nonstat.loglik.case.1 hasta whittle.nonstat.loglik.case.12	108
B.3. Estimador Kalman	116
B.3.1. lsfm.kalman	116
B.3.2. lsfm.kalman.loglik	116
B.3.3. lsfm.kalman.loglik.print	117

Índice de cuadros

3.1. Estimación de Whittle: Caso tendencia lineal (3.9) con tamaño muestral $T = 512$, bloques de tamaño $N = 128$ y saltos $S = 64$	26
3.2. Estimación de Whittle: Caso tendencia lineal (3.9) con tamaño muestral $T = 512$, bloques de tamaño $N = 256$ y saltos $S = 128$	27
3.3. Estimación de Whittle: Caso tendencia lineal (3.9) con tamaño muestral $T = 1024$, bloques de tamaño $N = 128$ y saltos $S = 64$	27
3.4. Estimación de Whittle: Caso tendencia lineal (3.9) con tamaño muestral $T = 1024$, bloques de tamaño $N = 256$ y saltos $S = 128$	28
3.5. Estimación de Whittle: Caso tendencia lineal (3.10) con tamaño muestral $T = 512$, bloques de tamaño $N = 128$ y saltos $S = 64$	28
3.6. Estimación de Whittle: Caso tendencia lineal (3.10) con tamaño muestral $T = 512$, bloques de tamaño $N = 256$ y saltos $S = 128$	29
3.7. Estimación de Whittle: Caso tendencia lineal (3.10) con tamaño muestral $T = 1024$, bloques de tamaño $N = 128$ y saltos $S = 64$	30

3.8. Estimación de Whittle: Caso tendencia lineal (3.10) con tamaño muestral $T = 1024$, bloques de tamaño $N = 256$ y saltos $S = 128$	31
3.9. Estimación de Whittle: Caso armónico (3.11) con tamaño muestral $T = 512$, bloques de tamaño $N = 128$ y saltos $S = 64$	32
3.10. Estimación de Whittle: Caso armónico (3.11) con tamaño muestral $T = 512$, bloques de tamaño $N = 256$ y saltos $S = 128$	32
3.11. Estimación de Whittle: Caso armónico (3.11) con tamaño muestral $T = 1024$, bloques de tamaño $N = 128$ y saltos $S = 64$	33
3.12. Estimación de Whittle: Caso armónico (3.11) con tamaño muestral $T = 1024$, bloques de tamaño $N = 256$ y saltos $S = 128$	33
4.1. Estimación Método de Kalman para datos faltantes. 1000 repeticiones de tamaño $T = 1024$ y $m = 20$	40
5.1. Depósitos de Minerales: Modelo ajustado	46
5.2. Anillos de árbol: Modelo ajustado	53

Índice de figuras

2.1. (a) Truncación, (b) Innovación y (c) fracdiff	11
2.2. (a) Simulación $t : 1, \dots, 100$, (b) simulación $t : 901, \dots, 100$	12
2.3. Varianza teórica en el tiempo vs. Muestral a partir de simulaciones	13
5.1. Depósitos de Minerales.	44
5.2. Depósitos de Minerales. (a) ACF muestral, (b) Var-Plot.	44
5.3. Depósitos de Minerales. ACF muestral: (a) Observaciones 1 a 500, (b) observaciones 1075 a 1575, y (c) observaciones 2150 a 2650.	45
5.4. Depósitos de Minerales. (a) Estimación del parámetro de larga memoria. (b) Las estimaciones de la varianza del ruido. En ambos paneles la línea continua representa el modelo ARFIMA localmente estacionario, la línea quebrada horizontal indica el modelo ARFIMA estacionario y los puntos representan el suavizamiento por kernel.	46

5.5. Depósitos de Minerales: Análisis Residual. (a) Residuos obtenidos del ajuste del modelo, (b) ACF muestral, (c) PACF muestral, (d) test de Ljung-Box.	47
5.6. Depósitos de Minerales: Periodograma vs. Densidad espectral teórica.	49
5.7. Anillos de árbol.	51
5.8. Anillos de árbol: (a) ACF muestral, (b) Var-Plot.	52
5.9. Anillos de árbol. ACF muestral: (a) Observaciones 1 a 500, (b) observaciones 750 a 1250, y (c) observaciones 1490 a 1990.	52
5.10. Anillos de árbol. a) Estimación del parámetro de larga memoria. (b) Estimación de la varianza del ruido. En ambos paneles la línea continua representa el modelo ARFIMA localmente estacionarios, la línea quebrada horizontal indica el modelo ARFIMA estacionario y los puntos representan el suavizamiento por kernel.	53
5.11. Anillos de árbol: Análisis Residual. (a) Residuos del modelo ajustado, (b) ACF muestral, (c) PACF muestral, (d) test de Ljung-Box.	54
5.12. Anillos de árbol: Periodograma vs. Densidad espectral teórica.	55

Capítulo 1

Procesos Localmente Estacionarios

1.1. Introducción

A pesar que la estacionaridad es una hipótesis muy atractiva para la teoría del análisis de series de tiempo, en la práctica la mayoría de estas no se ajustan a tal condición. Como consecuencia de esto, se han propuestos varios enfoques a esta situación en la literatura. Entre estas metodologías propuestas, la diferenciación y la estimación de tendencias son opciones comúnmente usadas. Otros enfoques incluyen por ejemplo, técnicas de la evolución del espectro discutidas por Priestley (1965). En el mismo estilo, durante las últimas décadas una serie de nuevos enfoques se han propuesto con respecto a la variación del espectro en el tiempo. Una de estas metodologías, llamada procesos localmente estacionarias desarrollada por Dahlhaus (1997), ha sido amplia-

mente discutida recientemente en la literatura de series de tiempo, ver por ejemplo Dahlhaus and Giraitis (1998), Dahlhaus (2000), von Sachs and MacGibbon (2000), Guo et al. (2003), Sakiyama and Taniguchi (2003), Genton and Perrin (2004), Orbe et al. (2005), Dahlhaus and Polonik (2006), Van Bellegem and Dahlhaus (2006), Chandler and Polonik (2006), y Fryzlewicz et al. (2006) entre muchos otros. Este enfoque permite a procesos estocásticos que no son estacionarios considerarlos como si lo fuesen, suponiendo que a un nivel local el modelo varía lo suficientemente suave en el tiempo para considerarlo aproximadamente estacionario. La teoría de procesos localmente estacionarios discutida por Dahlhaus (1997) sólo se aplica a series de tiempo de corta memoria, como lo son los procesos autoregresivos (AR) que varían en el tiempo. Desafortunadamente, esta teoría no puede aplicarse en el análisis de muchas series de tiempo que exhiben no estacionaridad y fuerte dependencia, que es el caso de las series que se discuten en este trabajo. Con el fin de tratar este tipo de series temporales, se introducirá una clase de procesos localmente estacionarios fuertemente dependientes y se desarrollarán técnicas de estimación espectral. Adicionalmente se propondrá una estimación mediante filtro de Kalman cuya teoría asintótica queda propuesta para futuros trabajos. En particular estos modelos incluyen parámetros que se les permite evolucionar en el tiempo. Como se analizará más adelante en los ejemplos, permitir que el parámetro de memoria varíe suavemente ofrece una gran flexibilidad para el análisis dinámico de una serie de tiempo.

Siguiendo con lo propuesto por Dahlhaus (1997), con el fin de ajustar un modelo paramétrico para series localmente estacionarias se utilizó una extensión del estimador propuesto por Whittle (1953). Una gran ventaja de este procedimiento de estimación es su eficiencia computacional, ya que sólo requiere el cálculo del periodograma por medio de la transformada de Fourier (FFT). Se demostró en este trabajo que el estimador de Whittle es consistente, normal y eficiente.

Una novedad del método estadístico utilizado es que permite la detección y el modelamiento de algunos tipos de no estacionaridades que puede ser difícil de manejar con otras técnicas simples. Por ejemplo, los cambios en la varianza de una serie temporal puede ser detectada por simple inspección de los datos. Sin embargo, las variaciones en la estructura de la dependencia de los datos son mucho más difíciles de detectar y modelar. Como quedará de manifiesto en los ejemplos que se analizarán más adelante, la metodología propuesta es útil para detectar varios tipos de no estacionaridades en los datos, incluidos cambios en la varianza en la serie, en la varianza del ruido, o la estructura de autocorrelación de los datos.

1.2. Proceso Gaussiano Localmente Estacionario

Un Proceso Gaussiano localmente estacionario de media cero puede ser definido por la representación

$$Y_{t,T} = \int_{-\pi}^{\pi} A_{t,T}^0(\lambda) e^{i\lambda t} d\xi(\lambda), \quad (1.1)$$

para $t = 1, \dots, T$ con función de transferencia A^0 . Y además,

- $\xi(\lambda)$ sea un movimiento Browniano en $[-\pi, \pi]$.
- y existe una función A , 2π -periódica con preimagen en $(0, 1] \times \mathbb{R}$ e imagen en \mathbb{C} , con $A(u, -\lambda) = \overline{A(u, \lambda)}$ y

$$\sup_{t,\lambda} \left| A_{t,T}^0(\lambda) - A\left(\frac{t}{T}, \lambda\right) \right| \leq \frac{K}{T}, \quad (1.2)$$

para todo T con K una constante positiva.

La suavidad de $A(\lambda)$ en t/T garantiza que este tipo de procesos tenga un comportamiento localmente estacionario.

Un ejemplo de esta clase de procesos localmente estacionario esta dada por una expansión infinita de medias móviles:

$$Y_{t,T} = \sigma \left(\frac{t}{T} \right) \sum_{j=0}^{\infty} \psi_j \left(\frac{t}{T} \right) \varepsilon_t, \quad (1.3)$$

donde $\{\varepsilon_t\}$ es una secuencia Gaussiana de media cero y varianza uno, y

$$\sum_{j=0}^{\infty} \{\psi_j(u)\}^2 < \infty, \quad \forall u \in [0, 1].$$

En este caso, la función de transferencia del proceso (1.3) está dada por $A_{t,T}^0(\lambda) = \sum_{j=0}^{\infty} \psi_j\left(\frac{t}{T}\right) e^{-i\lambda j} = A\left(\frac{t}{T}, \lambda\right)$, la cual satisface trivialmente (1.2). El modelo definido en (1.3) generaliza la expansión de Wold usual para procesos lineales estacionarios permitiendo que los coeficientes de la expansión infinita de medias móviles varíen lentamente en el tiempo.

1.2.1. Ruido Fraccionario Localmente Estacionario (LS-FN)

Un caso particular de (1.3) es la versión generalizada de un proceso de ruido fraccionario en tiempo discreto

$$Y_{t,T} = \sigma \left(\frac{t}{T}\right) (1 - B)^{-d(t/T)} \varepsilon_t = \sigma \left(\frac{t}{T}\right) \sum_{j=0}^{\infty} \eta_j \left(\frac{t}{T}\right) \varepsilon_t, \quad (1.4)$$

para $t = 1, \dots, T$, con $\{\varepsilon_t\}$ una secuencia de ruidos blancos Gaussianos de media cero y varianza uno. Los coeficientes de la expansión infinita de medias móviles $\{\eta_j\}$ están dados por

$$\eta_j \left(\frac{t}{T}\right) = \frac{\Gamma[j + d(t/T)]}{\Gamma(j+1) \Gamma[d(t/T)]}, \quad (1.5)$$

donde $\Gamma(\cdot)$ es la función Gamma y $d(\cdot)$ es el coeficiente de larga memoria que varía lentamente en el tiempo. A este proceso lo llamaremos ruido fraccionario localmente estacionario el cual lo denotaremos como LS-FN(d).

1.2.2. Proceso LS-ARFIMA

Una extensión del proceso definido en la Sección 1.2.1 es

$$\Phi(t/T, B) Y_{t,T} = \Theta(t/T, B) (1 - B)^{-d(t/T)} \sigma(t/T) \varepsilon_t, \quad (1.6)$$

para $t = 1, \dots, T$, donde $\Phi(u, B) = 1 + \phi_1(u) B + \dots + \phi_P(u) B^P$ es una componente autoregresiva, $\theta(u, B) = 1 + \theta_1(u) B + \dots + \theta_Q(u) B^Q$ es una componente de medias móviles, $d(u)$ es un parámetro de larga memoria, $\sigma(u)$ es un factor que determina la varianza para $u \in [0, 1]$ y $\{\varepsilon_t\}$ es una secuencia de ruido blanco Gaussiana de media cero y varianza uno. Esta clase de modelos es una extensión del proceso autoregresivo fraccionario integrado de medias móviles (ARFIMA), donde $\Phi(u, B)$, $\Theta(u, B)$ y $d(u)$ no dependen de u . Por simplicidad, el modelo (1.6) lo llamaremos ARFIMA localmente estacionario (LS-ARFIMA). Además, la evolución de los parámetros del modelo (1.6) puede ser especificada en términos de una clase general de funciones. Por ejemplo, sea $\{g_j(u)\}$, con $j = 1, 2, \dots$, una base de un espacio de funciones que varían lentamente y $\theta(u)$ un parámetro del modelo (1.6). Entonces podemos escribir este parámetro en términos de la base $\{g_j(u)\}$ como sigue

$$\theta(u) = \sum_{j=0}^k \alpha_j g_j(u), \quad (1.7)$$

para valores k y $\alpha_0, \dots, \alpha_k$ desconocidos. Para este caso, la estimación del parámetro $\theta(u)$ involucra determinar previamente k y luego estimar los coeficientes α 's. Un ejemplo de esta aproximación sería una clase de polinomios generadas por la base

$\{g_j(u) = u^j\}$, expansiones de Fourier generada por la base $\{g_j(u) = e^{iu_j}\}$ y familias por wavelets, tales como los sistemas Haar o Daubechies. Extensiones de estos casos podría también ser consideradas. Por ejemplo, las bases también puede incluir parámetros como en el caso $\{g_j(u) = e^{iu\beta_j}\}$, donde $\{\beta_j\}$ son valores desconocidos.

Capítulo 2

Métodos de Simulación

En este capítulo analizaremos el problema de simulación de datos proveniente de un proceso LS-FN. Se proponen tres métodos: (1) Método Expansión $MA(\infty)$ Truncada, ver ejemplo Bardet et al. (2003) p.590, (2) Algoritmo de Innovación, ver ejemplo en Brockwell and Davis (1991) p.172 y (3) una aproximación al algoritmo de innovación utilizando el paquete “fracdiff” implementado para el software estadístico *R*.

2.1. Expansión $\text{MA}(\infty)$ truncada

Sea $Y_{t,T}$ para $t \in \{1, 2, \dots, T\}$ un proceso lineal localmente estacionario de fuerte dependencia con representación

$$Y_{t,T} = \sigma\left(\frac{t}{T}\right) \sum_{j=0}^{\infty} \eta_j\left(\frac{t}{T}\right) \varepsilon_{t-j}, \quad \text{con} \quad \sum_{j=0}^{\infty} \eta_j^2\left(\frac{t}{T}\right) < \infty,$$

donde $\sigma(t/T)$ es un factor de escala para el proceso y $\{\varepsilon_t\}$ es una secuencia independiente e idénticamente distribuida con media cero y varianza uno. Consideremos el caso particular de un proceso LS-FN con $d(t/T) \in (-1/2, 1/2)$, el cual tiene representación

$$Y_{t,T} = \sigma\left(\frac{t}{T}\right) \sum_{j=0}^{\infty} \frac{\Gamma[j + d(t/T)]}{\Gamma(j+1)\Gamma[d(t/T)]} \varepsilon_{t-j}, \quad (2.1)$$

con $\sigma(u)$ y $d(u)$ parámetros construidos en términos de la base $\{g_i(u)\}$ definida en (1.7) para $u \in [0, 1]$.

Como

$$\frac{\Gamma[j + d(u)]}{\Gamma(j+1)\Gamma[d(u)]} \sim j^{d(u)-1},$$

cuando $k \rightarrow \infty$, una aproximación natural del proceso $Y_{t,T}$ es

$$Y_{t,T} = \sigma\left(\frac{t}{T}\right) \sum_{j=0}^M \frac{\Gamma[j + d(t/T)]}{\Gamma(j+1)\Gamma[d(t/T)]} \varepsilon_{t-j}, \quad (2.2)$$

para cada $t \in \{1, 2, \dots, T\}$ y M la truncación establecida.

2.2. Algoritmo de Innovación

Una muestra de un proceso LS-FN definido en (1.4)-(1.5) puede ser generada por medio del algoritmo de innovación, ver ejemplo Brockwell and Davis (1991) p.172. En esta aplicación la matriz de varianza-covarianza de el proceso, $\kappa(s, t)$, esta dada por

$$\begin{aligned}\kappa(s, t) &= E(Y_{s,T} Y_{t,T}) \\ &= \sigma\left(\frac{s}{T}\right) \sigma\left(\frac{t}{T}\right) \sum_{j=0}^{\infty} \eta_{j+s-t}\left(\frac{s}{T}\right) \eta_j\left(\frac{t}{T}\right) \\ &= \sigma\left(\frac{s}{T}\right) \sigma\left(\frac{t}{T}\right) \frac{\Gamma[1 - d(\frac{s}{T}) - d(\frac{t}{T})] \Gamma[s - t + d(\frac{s}{T})]}{\Gamma[1 - +d(\frac{s}{T})] \Gamma[d(\frac{s}{T})] \Gamma[s - t + 1 - d(\frac{t}{T})]},\end{aligned}$$

para $s, t = 1, \dots, T$, $s \geq t$, donde los coeficientes η_j están dados por (1.5).

2.3. Aproximación Algoritmo de Innovación paquete “fracdiff”

Otra alternativa al Algoritmo de Innovación es generar primero “ T ” series de tiempo $\{X_{i,t,T}\}$ de tamaño T a partir de las mismas innovaciones $\{\varepsilon_t\}$ normales con media cero y varianza uno, pero con parámetro de larga memoria $d(i/T)$ siendo una función que varia suavemente para $i = 1, \dots, T$ con la forma descrita en (1.7). Luego, la serie de tiempo de larga dependencia localmente estacionaria se construye considerando la observación 1 de la primera serie generada, la observación 2 de la segunda y

así sucesivamente, por ejemplo para el tiempo t la observación sería

$$Y_{t,T} = X_{t,T}[d(t/T)]$$

2.4. Comparación

2.4.1. Simulación LS-FN

La Figura 2.1 muestra un ejemplo de simulación de una serie de tamaño 1000 correspondiente a un proceso LS-FN, con parámetro de larga memoria definido por

$$d(u) = 0.15 + 0.20 \cdot u, \quad u \in [0, 1],$$

y factor de escala $\sigma(u) = 1$ para todo $u \in [0, 1]$ a partir de una secuencia de ruido

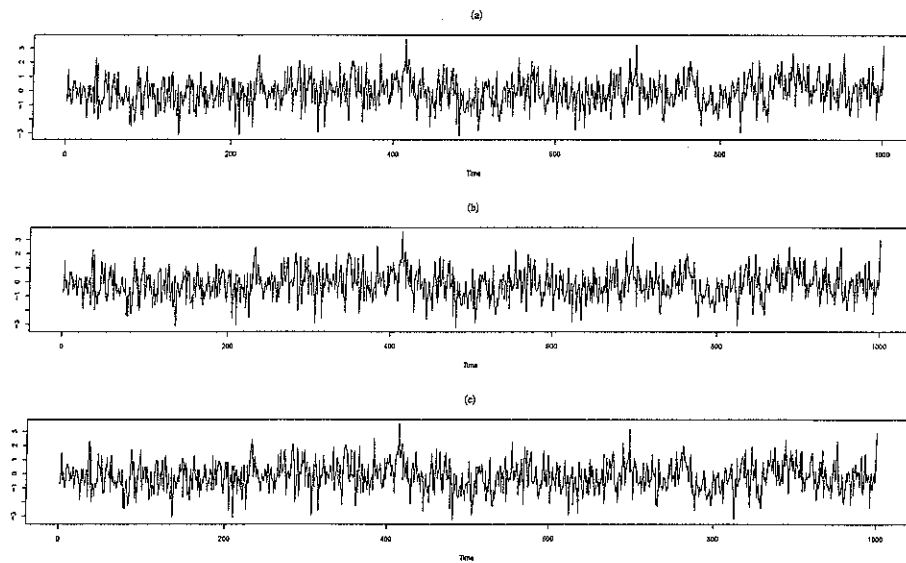


Figura 2.1: (a) Truncación, (b) Innovación y (c) fracdiff

blanco independiente e idénticamente distribuida normal de media cero y varianza uno. Los paneles (a)-(c) de esta figura muestran las series resultantes de los tres métodos. Para observar de mejor manera el resultado final de estos métodos en la Figura 2.2 se muestra un acercamiento de los primeros y últimos cien valores simulados en (a) y (b) respectivamente. Se puede apreciar que los tres métodos entregan como resultado final series de tiempo muy parecidas sobre todo para los primeros valores simulados, pero a medida que el tiempo pasa el método de truncación presentado como una línea continua se aparta de los otros dos métodos en el tiempo. Otra diferencia importante entre estos métodos de simulación es el tiempo que demoran, siendo el

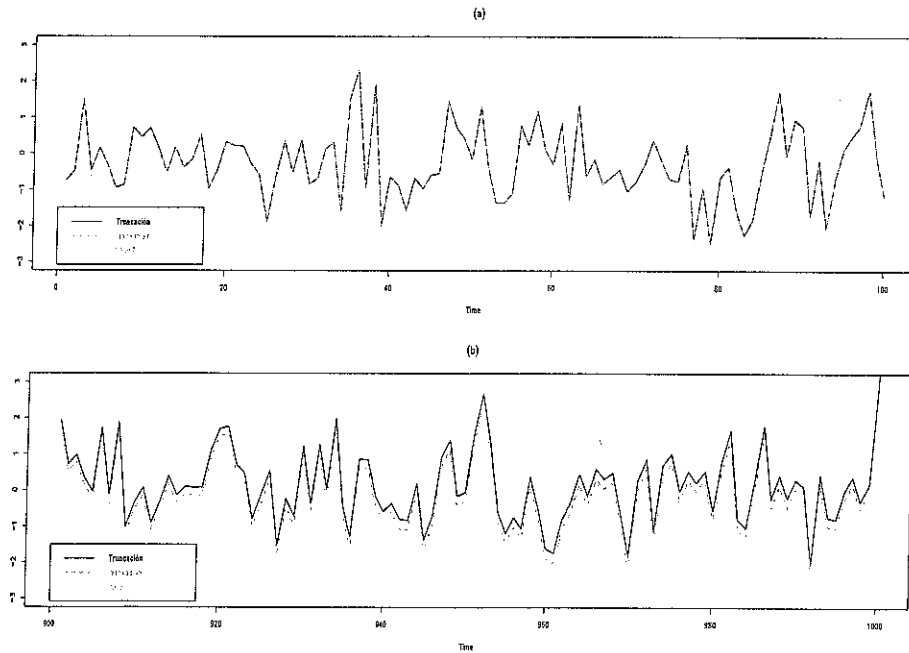


Figura 2.2: (a) Simulación $t : 1, \dots, 100$, (b) simulación $t : 901, \dots, 1000$.

último propuesto considerablemente más lento que los dos primeros.

2.4.2. Varianza Teórica vs. Muestral

El efecto de la diferencia en los tres métodos descritos anteriormente en la simulación final lograda se puede observar al analizar la varianza teórica esperada de los procesos simulados versus un suavizamiento por spline cúbica implementado en R, ver por ejemplo Hastie and Tibshirani (1990) y Chambers and Hastie (1992). Para obtener esta comparación se procedió primero a simular por los tres métodos 500 series de tamaño $T = 500$ de un proceso LS-FN a partir de las mismas innovaciones normales de media cero y varianza uno, donde el parámetro de larga memoria fue definido como $d(u) = 0.15 + 0.2 \cdot u$, con $u \in [0, 1]$ y el factor de escala para la varianza $\sigma^2(u)$ constante e igual a uno. Luego para cada uno de los 500 “tiempos” se

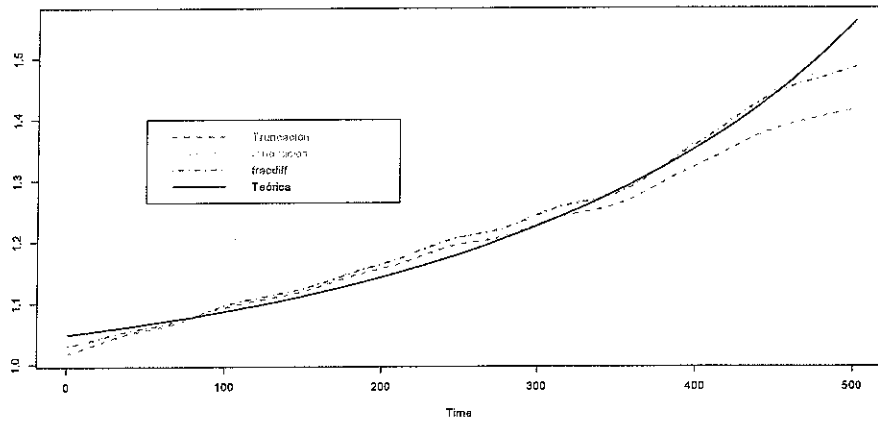


Figura 2.3: Varianza teórica en el tiempo vs. Muestral a partir de simulaciones

calculó la varianza de las 200 repeticiones y finalmente se aplica el suavizamiento por spline cúbica. La Figura 2.3 muestra en línea continua la varianza teórica del proceso LS-FN en el tiempo $u = t/T$, con $u \in [0, 1]$ para $t = 1, \dots, T$

$$\sigma_Y^2(u) = \frac{\Gamma[1 - 2d(u)]}{\Gamma[1 - d(u)]^2},$$

y en líneas entre-cortadas el suavizamiento de la variabilidad muestral de los tres métodos por kernel. Se observa que el método de truncación no logra capturar todo el efecto de la variabilidad teórica, mientras que los otros dos en alguna medida logran un mejor resultado y son parecidos entre sí.

Capítulo 3

Estimador de Whittle

3.1. Definición

Considerando una muestra $\{Y_{t,T}\}$, para $t = 1, \dots, T$, del proceso (1.1) podemos estimar el vector de parámetros del modelo, denotado por θ , maximizando la función log-Verosimilitud de Whittle

$$L_T(\theta) = -\frac{1}{4\pi} \frac{1}{M} \int_{-\pi}^{\pi} \sum_{j=1}^M \left\{ \log f_{\theta}(u_j, \lambda) + \frac{I_N(u_j, \lambda)}{f_{\theta}(u_j, \lambda)} \right\} d\lambda, \quad (3.1)$$

donde $f_{\theta}(u, \lambda) = |A(u, \lambda)|^2$ corresponde a la densidad espectral del proceso, e $I_N(u, \lambda) = \frac{|D_N(u, \lambda)|^2}{2\pi H_{2,N}(0)}$ es el periodograma con *taper* donde

$$D_N(u, \lambda) = \sum_{s=0}^{N-1} h\left(\frac{s}{N}\right) Y_{[uT]-N/2+s+1,T} e^{-i\lambda s},$$
$$H_{k,N} = \sum_{s=0}^{N-1} \left[h\left(\frac{s}{N}\right) \right]^2 e^{-i\lambda s},$$

$T = S(M-1) + N$, $u_j = t_j/T$, $t_j = S(j-1) + N/2$, $j = 1, \dots, M$ y $h(\cdot)$ es el *taper* de los datos. La muestra $\{Y_{t,T}\}$ es subdividida en M bloques de tamaño N y saltos S entre bloque y bloque. Por ejemplo, consideremos una serie de tiempo con $T = 652$ observaciones y $M = 100$ bloques de tamaño $N = 256$ con salto $S = 4$, entonces los bloques resultantes son: $(Y_{1,626}, \dots, Y_{256,625})$, $(Y_{5,626}, \dots, Y_{260,625})$, \dots , $(Y_{397,626}, \dots, Y_{652,625})$.

Luego el espectro es estimado localmente por la media de las observaciones en el periodograma con taper para cada uno de los $M = 100$ bloques y promediados en forma (3.1). Finalmente, el estimador de Whittle de los parámetros del vector θ está dado por

$$\hat{\theta}_T = \arg \max L_T(\theta)$$

donde la maximización ocurre sobre el espacio paramétrico Θ .

3.2. Teoría Asintótica

Algunas propiedades para muestras grandes del estimador de Whittle definido en la sección anterior son establecidas a continuación. En particular se mostrará que los estimadores son consistentes, asintóticamente normales y eficientes. Estos resultados son establecidos en Teorema 3.2.1, Teorema 3.2.2 y Teorema 3.2.3 respectivamente. El comportamiento asintótico de los estimadores es estudiado asumiendo que la muestra $\{Y_{t,T}\}$ del proceso (1.1) es válido para $t = 1, \dots, T$. Así, el proceso es observado en los tiempos $t/T \in [0, 1]$. Cuando $T \rightarrow \infty$ más observaciones son agregadas y la

subdivisión comienza a ser cada vez más fina.

3.2.1. Supuestos

A continuación se establecen algunos supuestos que garantizan que los estimadores de Whittle cuasi-máximo verosimilitud cumplen algunas propiedades para muestras grandes como la consistencia, normalidad asintótica y eficiencia. El primer supuesto es sobre la variación de la densidad espectral del proceso en el tiempo. El segundo supuesto se relaciona con la función taper aplicada a los datos y el tercero se refiere al esquema de bloques. Se supone que el espacio paramétrico Θ es compacto. En el resto de este trabajo, K representa una constante positiva que puede cambiar de línea en línea.

A 1 *La densidad espectral del proceso (1.1) que varia en el tiempo es estrictamente positiva y satisface*

$$f_{\theta}(u, \lambda) \sim C_f(\theta, u) |\lambda|^{-2\alpha(\theta, u)},$$

cuando $|\lambda| \rightarrow 0$, donde $C_f(\theta, u) > 0$, $\alpha(\theta, u) \in (0, 1)$. Existe una una función $g(\lambda)$ integrable tal que $|\nabla_{\theta} \log f_{\theta}(u, \lambda)| \leq g(\lambda)$ para todo $\theta \in \Theta$, $u \in [0, 1]$ y $\lambda \in [-\pi, \pi]$. La función de transferencia $A(u, \lambda)$ es dos veces diferenciable con respecto a u y satisface

$$\int_{-\pi}^{\pi} A(u, \lambda) A(v, -\lambda) e^{ik\lambda} d\lambda \sim C(\theta, u, v) k^{\alpha(\theta, u) + \alpha(\theta, v) - 1},$$

cuando $k \rightarrow \infty$, donde $|C(\theta, u, v)| \leq K$ para $u, v \in [0, 1]$ y $\theta \in \Theta$. La función $f_{\theta}(u, \lambda)^{-1}$ es dos veces diferenciable con respecto a θ , u y λ .

A 2 El taper es una función $h(u)$ con $u \in [0, 1]$ simétrica en torno a $\frac{1}{2}$ y tiene derivadas acotadas.

A 3 El tamaño de la muestra T y sus subdivisiones N , S y M tienden a infinito satisfaciendo que $S/N \rightarrow 0$, $\sqrt{T} \log^2 N/N \rightarrow 0$, $\sqrt{T}/M \rightarrow 0$ y $N^3 \log^3 N/T^2 \rightarrow 0$.

Como ilustración de estos supuestos, consideremos la versión generalizada de un proceso de ruido fraccionario para tiempo discreto descrito en las ecuaciones (1.4)-(1.5). La densidad espectral de este proceso LS-FN(d) esta dada por

$$f_\theta(u, \lambda) = \frac{\sigma^2}{2\pi} \left(2 \sin \frac{\lambda}{2} \right)^{-2d_\theta(u)}. \quad (3.2)$$

Notemos que esta función es integrable en $\lambda \in [-\pi, \pi]$ para todos $u \in [0, 1]$ siempre y cuando $d_\theta(u) < \frac{1}{2}$ para todo $u \in [0, 1]$ y $\theta \in \Theta$. Por lo tanto, tenemos que $f_\theta(u, \lambda) \sim \frac{\sigma^2}{2\pi} |\lambda|^{-2d_\theta(u)}$, cuando $\lambda \rightarrow 0$. En el supuesto de que $|\nabla_\theta d_\theta(u)| \leq K$ tenemos que $|\nabla_\theta \log f_\theta(u, \lambda)| = |\nabla_\theta d_\theta(u)| \log \left(2 \sin \frac{\lambda}{2} \right)^2 \leq K \log |\lambda|$, que es una función integrable en $\lambda \in [-\pi, \pi]$. Además, de (1.5) la función de transferencia de este proceso cumple

$$\begin{aligned} \int_{-\pi}^{\pi} A(u, \lambda) A(v, -\lambda) \exp(ik\lambda) d\lambda &= \sum_{j=0}^{\infty} \eta_{j+k}(u) \eta_j(v) \\ &= \frac{\Gamma[1 - d(u) - d(v)] \Gamma[k + d(u)]}{\Gamma[1 - d(u)] \Gamma[d(u)] \Gamma[k + 1 - d(v)]}, \end{aligned}$$

para $k \geq 0$. Así, por la aproximación de Stirling's obtenemos

$$\int_{-\pi}^{\pi} A(u, \lambda) A(v, -\lambda) \exp(ik\lambda) d\lambda \sim \frac{\Gamma[1 - d(u) - d(v)]}{\Gamma[1 - d(u)] \Gamma[d(u)]} k^{d(u)+d(v)-1},$$

para $k \rightarrow \infty$. Por otra parte, un simple cálculo demuestra que $f_\theta(u, \lambda)^{-1}$ es dos veces diferenciable con respecto a u y λ mientras que $d_\theta(u)$ es dos veces diferenciable con respecto a u . Así, bajo estas condiciones la densidad espectral variante en el tiempo $f_\theta(u, \lambda)$ satisface el supuesto A1. Por otra parte, un ejemplo de taper para los datos que satisfaga A2 es la función campana-coseno

$$h(x) = \frac{1}{2}[1 - \cos(2\pi x)]. \quad (3.3)$$

Notar que si $S = \mathcal{O}(N^a)$ y $M = \mathcal{O}(N^b)$ entonces $T = \mathcal{O}(N^{a+b})$ para $a + b \geq 1$. De esta manera, se escogen los exponentes a y b tal que $a < 1$, $\frac{3}{2} < a + b < 2$ y $a < b$, luego el supuesto A3 también se cumple.

3.2.2. Teoremas y Principales Resultados

El análisis de las propiedades asintóticas del estimador de Whittle localmente estacionario que maximiza (1.6) se discuten más detalladamente en los siguientes tres teoremas. Las demostraciones se encuentran en el Apéndice A.

Definamos el estimador de Whittle $\hat{\theta}_T$ como el vector que minimiza (1.6).

Teorema 3.2.1 (*Consistencia*) Sea θ_0 el verdadero valor del parámetro θ . Asumiendo las suposiciones A1 - A3, el estimador $\hat{\theta}_T$ satisface

$$\hat{\theta}_T \rightarrow \theta_0,$$

en probabilidad, cuando $T \rightarrow \infty$.

Teorema 3.2.2 (Normalidad) Sea θ_0 el verdadero valor del parámetro θ . Si las suposiciones A1 - A3 se cumplen, entonces el estimador de Whittle $\hat{\theta}_T$ satisfacen el teorema central del límite

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \rightarrow N [0, \Gamma(\theta_0)^{-1}] ,$$

cuando $T \rightarrow \infty$, donde

$$\Gamma(\theta) = \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} [\nabla \log f_{\theta}(u, \lambda)] [\nabla \log f_{\theta}(u, \lambda)]' d\lambda du. \quad (3.4)$$

Teorema 3.2.3 Asumiendo que las condiciones A1 - A3 se cumplen, el estimador de Whittle $\hat{\theta}_T$ es asintóticamente eficiente.

3.2.3. Ejemplos

Con el fin de ilustrar la aplicación de los resultados teóricos establecidos en la sección anterior analizaremos casos en que la evolución del parámetro de larga memoria, la varianza del ruido, los componentes autorregresivos y de media móviles de un proceso LS-ARFIMA definido en (1.6) son especificadas por un polinomio o mediante componentes armónicas.

Ejemplo 3.2.1 Consideremos primero el caso $P = Q = 0$ en el modelo (1.6) donde

$d(u)$ y $\sigma(u)$ son especificadas por,

$$d(u) = \sum_{j=1}^p \alpha_j g_j(u), \quad \sigma(u) = \sum_{j=1}^q \beta_j h_j(u),$$

para $u \in [0, 1]$, donde $g_j(\cdot)$ y $h_j(\cdot)$ son funciones base tal como las especificadas en (1.7). En este caso el vector de parámetros es $(\alpha_0, \dots, \alpha_p, \beta_0, \dots, \beta_q)'$ y la matriz Γ dada por (3.4) puede ser escrita como

$$\Gamma = \begin{pmatrix} \Gamma_\alpha & 0 \\ 0 & \Gamma_\beta \end{pmatrix}, \quad (3.5)$$

donde

$$\Gamma_\alpha = \frac{\pi^2}{6} \left[\int_0^1 g_i(u) g_j(u) du \right]_{i,j=0,\dots,p}, \quad (3.6)$$

y

$$\Gamma_\beta = 2 \left[\int_0^1 \frac{h_i(u) h_j(u)}{\left[\sum_{j=1}^q \beta_j h_j(u) \right]^2} du \right]_{i,j=0,\dots,q}. \quad (3.7)$$

Ejemplo 3.2.2 Como un caso particular de los parámetros de la especificación del ejemplo anterior, considere el caso $P = Q = 0$ en el modelo (1.6) donde $d(u)$ y $\sigma(u)$ son especificados por polinomios,

$$d(u) = \alpha_0 + \alpha_1 u + \dots + \alpha_p u^p,$$

$$\sigma(u) = \beta_0 + \beta_1 u + \dots + \beta_q u^q,$$

para $u \in [0, 1]$. Similar al Ejemplo 3.2.1, en este caso el vector de parámetros es $(\alpha_0, \dots, \alpha_p, \beta_0, \dots, \beta_q)'$ y la matriz Γ dada por (3.4) puede escribirse como en (3.5) donde

$$\Gamma_\alpha = \left[\frac{\pi^2}{6(i+j+1)} \right]_{i,j=0,\dots,p},$$

y

$$\Gamma_\beta = 2 \left[\int_0^1 \frac{u^{i+j} du}{(\beta_0 + \beta_1 u + \dots + \beta_q u^q)^2} \right]_{i,j=0,\dots,q}.$$

Estas integrales pueden ser evaluadas por procedimientos de cálculo estándar, ver por ejemplo, página 64 de Gradshteyn and Ryzhik (2000) o por integración numérica.

Ejemplo 3.2.3 Siguiendo con el supuesto de $P = Q = 0$ en el modelo (1.6), considere que $d(u)$ y $\sigma(u)$ son definidos por expansiones armónicas

$$d(u) = \alpha_0 + \alpha_1 \cos(\lambda_1 u) + \dots + \alpha_p \cos(\lambda_p u),$$

$$\sigma(u) = \beta_0 + \beta_1 \cos(\omega_1 u) + \dots + \beta_q \cos(\omega_q u),$$

para $u \in [0, 1]$, donde $\lambda_0 = 0$, $\lambda_i^2 \neq \lambda_j^2$ para todo $i, j = 0, \dots, p$, $i \neq j$, $\omega_0 = 0$, y $\omega_i^2 \neq \omega_j^2$ para todo $i, j = 0, \dots, q$, $i \neq j$. Por simplicidad, los valores de las frecuencias $\{\lambda_j\}$ y $\{\omega_j\}$ son asumidos conocidos. Como en el Ejemplo 3.2.2, en este caso el vector de parámetros es $(\alpha_0, \dots, \alpha_p, \beta_0, \dots, \beta_q)'$ y la matriz Γ que aparece en (3.4) puede escribirse como (3.5) con

$$\Gamma_\alpha = \frac{\pi^2}{12} \left[\frac{\sin(\lambda_i - \lambda_j)}{\lambda_i - \lambda_j} + \frac{\sin(\lambda_i + \lambda_j)}{\lambda_i + \lambda_j} \right]_{i,j=0,\dots,p},$$

y

$$\Gamma_{\beta} = \frac{\pi^2}{12} \left[\frac{\sin(\omega_i - \omega_j)}{\omega_i - \omega_j} + \frac{\sin(\omega_i + \omega_j)}{\omega_i + \omega_j} \right]_{i,j=0,\dots,q}.$$

Ejemplo 3.2.4 *Consideremos ahora que $P = Q = 1$ en el modelo (1.6) donde $\sigma(u) = 1$ y $d(u)$, $\Phi(u, B)$, $\Theta(u, B)$ son especificadas*

$$d(u) = \alpha_1 u,$$

$$\Phi(u, B) = 1 + \phi(u)B, \quad \phi(u) = \alpha_2 u,$$

$$\Theta(u, B) = 1 + \theta(u)B, \quad \theta(u) = \alpha_3 u,$$

para $u \in [0, 1]$. En este caso, el vector de parámetros es $(\alpha_1, \alpha_2, \alpha_3)'$ y la matriz Γ de (3.4) puede escribirse como

$$\Gamma = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} \end{pmatrix},$$

donde

$$\gamma_{11} = \frac{1}{2\alpha_1^3} \log \frac{1 + \alpha_1}{1 - \alpha_1} - \frac{1}{\alpha_1^2}, \quad |\alpha_1| < 1,$$

$$\gamma_{12} = \frac{1}{(\alpha_1 \alpha_2)^{3/2}} g(\alpha_1 \alpha_2) - \frac{1}{\alpha_1 \alpha_2},$$

con $g(x) = \operatorname{arctanh}(\sqrt{x})$ para $x \in (0, 1)$ y $g(x) = \operatorname{arctan}(\sqrt{-x})$ para $x \in (-1, 0)$,

$$\begin{aligned}\gamma_{13} &= \frac{1}{2\alpha_1} \left\{ \left[\frac{1}{2} - \frac{1}{\alpha_1} \right] - \left[1 - \frac{1}{\alpha_1^2} \right] \log(1 + \alpha_1) \right\}, \\ \gamma_{22} &= \frac{1}{2\alpha_2^3} \log \frac{1 + \alpha_2}{1 - \alpha_2} - \frac{1}{\alpha_2^2}, \quad |\alpha_2| < 1, \\ \gamma_{23} &= \frac{1}{2\alpha_2} \left\{ \left[1 - \frac{1}{\alpha_2^2} \right] \log(1 + \alpha_2) - \left[\frac{1}{2} - \frac{1}{\alpha_2} \right] \right\}, \\ \gamma_{33} &= \frac{\pi^2}{18}.\end{aligned}$$

Ejemplo 3.2.5 Para un proceso de ruido fraccionario estacionario $FN(d)$, la varian-za asintótica del estimador máximo verosímil (MLE) del parámetro de larga memoria, \hat{d} , satisface

$$\lim_{T \rightarrow \infty} T \operatorname{Var}(\hat{d}) = \frac{6}{\pi^2}.$$

Por otro parte, supongamos que consideramos un proceso LS-FN proceso donde el parámetro de larga memoria varía por ejemplo de acuerdo a $d(u) = \alpha_0 + \alpha_1 u$. Para estimar $d(u)$, debemos estimar los dos parámetros α_0 y α_1 . Sea $\hat{\alpha}_0$ y $\hat{\alpha}_1$ los estimadores de Whittle, luego $\hat{d}(u) = \hat{\alpha}_0 + \hat{\alpha}_1 u$. Según el Teorema 3.2.2, la varianza asintótica de esta estimación cumple con

$$\lim_{T \rightarrow \infty} T \operatorname{Var}[\hat{d}(u)] = \frac{24}{\pi^2} (1 - 3u + 3u^2),$$

y luego integrando con respecto a u obtenemos

$$\lim_{T \rightarrow \infty} T \int_0^1 \operatorname{Var}[\hat{d}(u)] du = \frac{12}{\pi^2}.$$

Como se estiman dos parámetros, la varianza asintótica de $\widehat{d}(u)$ es dos veces la varianza asintótica de \widehat{d} . Este resultado pueden generalizarse para los casos en que tres o más coeficientes se calculan y para las tendencias más complejas, tal como se establece en el siguiente teorema.

Teorema 3.2.4 *Considere un proceso LS-FN (3.2) con parámetro de larga memoria variante en el tiempo $d_\beta(u) = \sum_{j=1}^p \beta_j g_j(u)$, donde $\{g_j(u)\}$ son funciones base tal como se definen en (1.7). Sea $\widehat{d}(u) = \sum_{j=1}^p \widehat{\beta}_j g_j(u)$ el estimador de $d_\beta(u)$ para $u \in [0, 1]$. Entonces, bajo los supuestos A1–A3 tenemos que*

$$\lim_{T \rightarrow \infty} T \int_0^1 \text{Var} \left[\widehat{d}(u) \right] du = \frac{6}{\pi^2} p. \quad (3.8)$$

Notar que de acuerdo al Teorema 3.2.4 el límite del promedio de las varianzas de $d(u)$ dado por (3.8) no depende de las funciones determinísticas $g_j(\cdot)$ para todo $j = 1, \dots, p$.

3.3. Simulación

Para analizar el desempeño de los estimadores de Whittle se simularon procesos LS-FN considerando primero una tendencia lineal para los parámetros de larga memoria y un factor de escala. Luego se evaluó el desempeño considerando una evolución del tipo armónica para el parámetro de larga memoria, con factor de escala fijo. Se consideraron muestras de tamaño $T = 512$ y $T = 1,024$ con subdivisiones $N = 128$,

$S = 64$ y $N = 256$, $S = 128$. Se utilizó como taper la campana-coseno definida en (3.3). Todos los resultados están basados en 1000 repeticiones.

3.3.1. Caso Lineal

Consideremos primero el caso donde

$$d(u) = \alpha_0 + \alpha_1 u \quad y \quad \sigma(u) = 1. \quad (3.9)$$

Los parámetros α_0 y α_1 de (3.9) pertenecen a un conjunto tal que $0 < d(u) < \frac{1}{2}$ para todo $u \in [0, 1]$. Los resultados para este conjunto de simulaciones se encuentran en las Tablas 3.1 - 3.4. En estas tablas, los promedios de las estimación de los parámetros, están representados en la tercera y cuarta columnas, los cuales están cerca de sus verdaderos valores. Por otra parte, la estimación de la desviación estándar se muestra en la séptima y octava columnas, que también están cerca de sus teórico en la quinta y sexta columnas, respectivamente. Notar que para un tamaño muestral $T = 512$ se logra un sesgo más pequeño con la combinación $N = 128$ y $S = 64$, mientras

Tabla 3.1: Estimación de Whittle: Caso tendencia lineal (3.9) con tamaño muestral $T = 512$, bloques de tamaño $N = 128$ y saltos $S = 64$.

Parámetros		Estimates		SD		Estimated SD	
α_0	α_1	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\sigma(\hat{\alpha}_0)$	$\sigma(\hat{\alpha}_1)$	$\hat{\sigma}(\hat{\alpha}_0)$	$\hat{\sigma}(\hat{\alpha}_1)$
0.15	0.25	0.141	0.251	0.069	0.119	0.093	0.153
0.15	0.30	0.157	0.283	0.069	0.119	0.089	0.130
0.20	0.25	0.206	0.235	0.069	0.119	0.094	0.137

Tabla 3.2: Estimación de Whittle: Caso tendencia lineal (3.9) con tamaño muestral $T = 512$, bloques de tamaño $N = 256$ y saltos $S = 128$.

Parámetros		Estimación		SD		SD Estimación	
α_0	α_1	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\sigma(\hat{\alpha}_0)$	$\sigma(\hat{\alpha}_1)$	$\hat{\sigma}(\hat{\alpha}_0)$	$\hat{\sigma}(\hat{\alpha}_1)$
0.15	0.25	0.156	0.234	0.069	0.119	0.103	0.175
0.15	0.30	0.173	0.263	0.069	0.119	0.093	0.146
0.20	0.25	0.222	0.214	0.069	0.119	0.100	0.155

que para un tamaño de muestra $T = 1024$ los mejores resultados se lograron con la combinación $N = 256$ y $S = 128$. Las desviaciones estándar es en ambos casos son muy parecidas.

Tabla 3.3: Estimación de Whittle: Caso tendencia lineal (3.9) con tamaño muestral $T = 1024$, bloques de tamaño $N = 128$ y saltos $S = 64$.

Parámetros		Estimación		SD		SD Estimación	
α_0	α_1	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\sigma(\hat{\alpha}_0)$	$\sigma(\hat{\alpha}_1)$	$\hat{\sigma}(\hat{\alpha}_0)$	$\hat{\sigma}(\hat{\alpha}_1)$
0.10	0.20	0.078	0.213	0.049	0.084	0.060	0.106
0.15	0.25	0.128	0.281	0.049	0.084	0.065	0.109
0.20	0.20	0.186	0.223	0.049	0.084	0.066	0.111
0.20	0.25	0.200	0.256	0.049	0.084	0.064	0.090
0.25	0.20	0.257	0.196	0.049	0.084	0.063	0.091

Para el caso

$$d(u) = \alpha_0 + \alpha_1 u \quad \text{y} \quad \sigma(u) = \beta_0 + \beta_1 u. \quad (3.10)$$

Los parámetros α_0 , α_1 , β_0 y β_1 de (3.10) pertenecen a un conjunto compacto tales que $0 < d(u) < \frac{1}{2}$ y $\sigma(u) > 0$ para todo $u \in [0, 1]$. Los resultados para este conjunto de

Tabla 3.4: Estimación de Whittle: Caso tendencia lineal (3.9) con tamaño muestral $T = 1024$, bloques de tamaño $N = 256$ y saltos $S = 128$.

Parámetros		Estimación		SD		SD Estimación	
α_0	α_1	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\sigma(\hat{\alpha}_0)$	$\sigma(\hat{\alpha}_1)$	$\hat{\sigma}(\hat{\alpha}_0)$	$\hat{\sigma}(\hat{\alpha}_1)$
0.10	0.20	0.089	0.209	0.049	0.084	0.062	0.108
0.15	0.25	0.141	0.269	0.049	0.084	0.067	0.113
0.20	0.20	0.195	0.213	0.049	0.084	0.064	0.110
0.20	0.25	0.210	0.244	0.049	0.084	0.062	0.094
0.25	0.20	0.263	0.189	0.049	0.084	0.061	0.092

Tabla 3.5: Estimación de Whittle: Caso tendencia lineal (3.10) con tamaño muestral $T = 512$, bloques de tamaño $N = 128$ y saltos $S = 64$.

Parámetros		Estimación		SD		SD Estimación	
α_0	α_1	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\sigma(\hat{\alpha}_0)$	$\sigma(\hat{\alpha}_1)$	$\hat{\sigma}(\hat{\alpha}_0)$	$\hat{\sigma}(\hat{\alpha}_1)$
0.15	0.25	0.146	0.254	0.069	0.119	0.089	0.153
0.15	0.30	0.155	0.284	0.069	0.119	0.081	0.126
0.20	0.25	0.202	0.239	0.069	0.119	0.078	0.129
0.15	0.25	0.142	0.255	0.069	0.119	0.086	0.146
0.15	0.30	0.144	0.297	0.069	0.119	0.076	0.126
0.20	0.25	0.199	0.242	0.069	0.119	0.078	0.126
β_0	β_1	$\hat{\beta}_0$	$\hat{\beta}_1$	$\sigma(\hat{\beta}_0)$	$\sigma(\hat{\beta}_1)$	$\hat{\sigma}(\hat{\beta}_0)$	$\hat{\sigma}(\hat{\beta}_1)$
0.50	0.50	0.500	0.515	0.038	0.079	0.047	0.097
0.50	0.50	0.499	0.516	0.038	0.079	0.048	0.099
0.50	0.50	0.500	0.517	0.038	0.079	0.046	0.094
1.00	-0.50	1.000	-0.495	0.053	0.079	0.064	0.102
1.00	-0.50	1.001	-0.492	0.053	0.079	0.057	0.090
1.00	-0.50	1.002	-0.490	0.053	0.079	0.058	0.091

simulaciones se encuentran en la Tabla 3.5 a la Tabla 3.8. En estas tablas, los promedios de las estimación de los parámetros, están representados en la tercera y cuarta columnas, los cuales presentan un pequeño sesgo. Por otra parte, la estimación de la desviación estándar se muestra en la séptima y octava columnas, que también están cerca de sus valores teóricos mostrados en la quinta y sexta columnas, respectivamente. Notar que nuevamente para un tamaño muestral $T = 512$ se logró un menor sesgo en la combinación $N = 128$ y $S = 64$, en cambio para un tamaño muestra $T = 1024$ los mejores resultados se logran con la combinación $N = 256$ y $S = 128$.

Tabla 3.6: Estimación de Whittle: Caso tendencia lineal (3.10) con tamaño muestral $T = 512$, bloques de tamaño $N = 256$ y saltos $S = 128$.

Parámetros		Estimación		SD		SD Estimación	
α_0	α_1	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\sigma(\hat{\alpha}_0)$	$\sigma(\hat{\alpha}_1)$	$\hat{\sigma}(\hat{\alpha}_0)$	$\hat{\sigma}(\hat{\alpha}_1)$
0.15	0.25	0.161	0.235	0.069	0.119	0.097	0.174
0.15	0.30	0.171	0.269	0.069	0.119	0.087	0.144
0.20	0.25	0.214	0.227	0.069	0.119	0.081	0.147
0.15	0.25	0.155	0.235	0.069	0.119	0.098	0.173
0.15	0.30	0.153	0.281	0.069	0.119	0.086	0.149
0.20	0.25	0.209	0.225	0.069	0.119	0.088	0.154
β_0	β_1	$\hat{\beta}_0$	$\hat{\beta}_1$	$\sigma(\hat{\beta}_0)$	$\sigma(\hat{\beta}_1)$	$\hat{\sigma}(\hat{\beta}_0)$	$\hat{\sigma}(\hat{\beta}_1)$
0.5	0.5	0.502	0.508	0.038	0.079	0.062	0.128
0.5	0.5	0.500	0.509	0.038	0.079	0.061	0.126
0.5	0.5	0.502	0.508	0.038	0.079	0.059	0.117
1.0	-0.5	1.000	-0.498	0.053	0.079	0.078	0.130
1.0	-0.5	0.999	-0.492	0.053	0.079	0.068	0.114
1.0	-0.5	1.000	-0.491	0.053	0.079	0.069	0.116

Tabla 3.7: Estimación de Whittle: Caso tendencia lineal (3.10) con tamaño muestral $T = 1024$, bloques de tamaño $N = 128$ y saltos $S = 64$.

Parámetros		Estimación		SD		SD Estimación	
α_0	α_1	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\sigma(\hat{\alpha}_0)$	$\sigma(\hat{\alpha}_1)$	$\hat{\sigma}(\hat{\alpha}_0)$	$\hat{\sigma}(\hat{\alpha}_1)$
0.10	0.20	0.080	0.213	0.049	0.084	0.060	0.112
0.15	0.25	0.133	0.277	0.049	0.084	0.066	0.109
0.20	0.20	0.187	0.221	0.049	0.084	0.067	0.113
0.20	0.25	0.197	0.258	0.049	0.084	0.057	0.092
0.25	0.20	0.252	0.204	0.049	0.084	0.057	0.091
0.10	0.20	0.078	0.215	0.049	0.084	0.059	0.114
0.15	0.25	0.132	0.275	0.049	0.084	0.067	0.115
0.20	0.20	0.193	0.214	0.049	0.084	0.067	0.114
0.20	0.25	0.195	0.262	0.049	0.084	0.058	0.088
0.25	0.20	0.254	0.198	0.049	0.084	0.055	0.089
β_0	β_1	$\hat{\beta}_0$	$\hat{\beta}_1$	$\sigma(\hat{\beta}_0)$	$\sigma(\hat{\beta}_1)$	$\hat{\sigma}(\hat{\beta}_0)$	$\hat{\sigma}(\hat{\beta}_1)$
0.50	0.50	0.498	0.509	0.027	0.056	0.030	0.064
0.50	0.50	0.499	0.518	0.027	0.056	0.032	0.066
0.50	0.50	0.499	0.518	0.027	0.056	0.031	0.064
0.50	0.50	0.500	0.521	0.027	0.056	0.030	0.062
0.50	0.50	0.506	0.513	0.027	0.056	0.030	0.060
1.00	-0.50	0.998	-0.494	0.038	0.056	0.042	0.067
1.00	-0.50	1.001	-0.492	0.038	0.056	0.044	0.067
1.00	-0.50	1.004	-0.495	0.038	0.056	0.043	0.066
1.00	-0.50	1.003	-0.491	0.038	0.056	0.041	0.062
1.00	-0.50	1.008	-0.498	0.038	0.056	0.040	0.062

Tabla 3.8: Estimación de Whittle: Caso tendencia lineal (3.10) con tamaño muestral $T = 1024$, bloques de tamaño $N = 256$ y saltos $S = 128$.

Parámetros		Estimación		SD		SD Estimación	
α_0	α_1	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\sigma(\hat{\alpha}_0)$	$\sigma(\hat{\alpha}_1)$	$\hat{\sigma}(\hat{\alpha}_0)$	$\hat{\sigma}(\hat{\alpha}_1)$
0.10	0.20	0.092	0.207	0.049	0.084	0.062	0.115
0.15	0.25	0.145	0.266	0.049	0.084	0.067	0.109
0.20	0.20	0.196	0.212	0.049	0.084	0.065	0.114
0.20	0.25	0.203	0.253	0.049	0.084	0.055	0.092
0.25	0.20	0.260	0.192	0.049	0.084	0.055	0.094
0.10	0.20	0.091	0.207	0.049	0.084	0.062	0.117
0.15	0.25	0.144	0.261	0.049	0.084	0.065	0.115
0.20	0.20	0.200	0.206	0.049	0.084	0.065	0.112
0.20	0.25	0.203	0.252	0.049	0.084	0.057	0.090
0.25	0.20	0.257	0.193	0.049	0.084	0.054	0.092
β_0	β_1	$\hat{\beta}_0$	$\hat{\beta}_1$	$\sigma(\hat{\beta}_0)$	$\sigma(\hat{\beta}_1)$	$\hat{\sigma}(\hat{\beta}_0)$	$\hat{\sigma}(\hat{\beta}_1)$
0.50	0.50	0.500	0.505	0.027	0.056	0.034	0.071
0.50	0.50	0.501	0.509	0.027	0.056	0.035	0.071
0.50	0.50	0.499	0.512	0.027	0.056	0.034	0.069
0.50	0.50	0.499	0.515	0.027	0.056	0.032	0.066
0.50	0.50	0.504	0.508	0.027	0.056	0.033	0.067
1.00	-0.50	0.999	-0.496	0.038	0.056	0.044	0.072
1.00	-0.50	1.000	-0.494	0.038	0.056	0.046	0.071
1.00	-0.50	1.001	-0.495	0.038	0.056	0.045	0.072
1.00	-0.50	1.002	-0.495	0.038	0.056	0.044	0.068
1.00	-0.50	1.004	-0.498	0.038	0.056	0.043	0.068

3.3.2. Caso Armonico

Consideremos la siguiente estructura

$$d(u) = \alpha_0 + \alpha_1 \cos(2\pi u) \quad y \quad \sigma(u) = 1. \quad (3.11)$$

Análogamente para el caso de tendencia lineal, los parámetros α_0 y α_1 pertenecen a un conjunto compacto tal que $0 < d(u) < \frac{1}{2}$ para todo $u \in [0, 1]$. Desde la

Tabla 3.9: Estimación de Whittle: Caso armónico (3.11) con tamaño muestral $T = 512$, bloques de tamaño $N = 128$ y saltos $S = 64$.

Parámetros		Estimación		SD		SD Estimación	
α_0	α_1	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\sigma(\hat{\alpha}_0)$	$\sigma(\hat{\alpha}_1)$	$\hat{\sigma}(\hat{\alpha}_0)$	$\hat{\sigma}(\hat{\alpha}_1)$
0.15	0.10	0.139	0.097	0.034	0.049	0.038	0.049
0.20	0.10	0.184	0.099	0.034	0.049	0.047	0.057
0.20	0.15	0.192	0.148	0.034	0.049	0.044	0.049
0.25	0.10	0.240	0.108	0.034	0.049	0.048	0.062
0.25	0.15	0.242	0.156	0.034	0.049	0.046	0.053
0.25	0.20	0.249	0.194	0.034	0.049	0.042	0.040
0.30	0.10	0.295	0.104	0.034	0.049	0.047	0.059
0.30	0.15	0.297	0.155	0.034	0.049	0.041	0.047
0.35	0.10	0.345	0.105	0.034	0.049	0.042	0.055

Tabla 3.10: Estimación de Whittle: Caso armónico (3.11) con tamaño muestral $T = 512$, bloques de tamaño $N = 256$ y saltos $S = 128$.

Parámetros		Estimación		SD		SD Estimación	
α_0	α_1	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\sigma(\hat{\alpha}_0)$	$\sigma(\hat{\alpha}_1)$	$\hat{\sigma}(\hat{\alpha}_0)$	$\hat{\sigma}(\hat{\alpha}_1)$
0.15	0.10	0.147	0.090	0.034	0.049	0.040	0.046
0.20	0.10	0.194	0.090	0.034	0.049	0.047	0.052
0.20	0.15	0.204	0.139	0.034	0.049	0.046	0.048
0.25	0.10	0.249	0.096	0.034	0.049	0.047	0.055
0.25	0.15	0.255	0.142	0.034	0.049	0.045	0.052
0.25	0.20	0.267	0.180	0.034	0.049	0.045	0.040
0.30	0.10	0.304	0.093	0.034	0.049	0.048	0.054
0.30	0.15	0.311	0.140	0.034	0.049	0.043	0.046
0.35	0.10	0.356	0.094	0.034	0.049	0.042	0.051

Tabla 3.11: Estimación de Whittle: Caso armónico (3.11) con tamaño muestral $T = 1024$, bloques de tamaño $N = 128$ y saltos $S = 64$.

Parámetros		Estimación		SD		SD Estimación	
α_0	α_1	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\sigma(\hat{\alpha}_0)$	$\sigma(\hat{\alpha}_1)$	$\hat{\sigma}(\hat{\alpha}_0)$	$\hat{\sigma}(\hat{\alpha}_1)$
0.15	0.10	0.134	0.102	0.024	0.034	0.027	0.037
0.20	0.10	0.185	0.109	0.024	0.034	0.032	0.042
0.20	0.15	0.190	0.156	0.024	0.034	0.030	0.035
0.25	0.10	0.241	0.113	0.024	0.034	0.034	0.044
0.25	0.15	0.242	0.165	0.024	0.034	0.032	0.041
0.25	0.20	0.246	0.207	0.024	0.034	0.031	0.030
0.30	0.10	0.299	0.111	0.024	0.034	0.033	0.044
0.30	0.15	0.297	0.164	0.024	0.034	0.031	0.036
0.35	0.10	0.353	0.106	0.024	0.034	0.029	0.038

Tabla 3.9 a la Tabla 3.12 se presentan los resultados para el caso armónico (3.11) considerando tamaños muestrales $T = 512$ y $T = 1024$. Se aprecia en estas tablas

Tabla 3.12: Estimación de Whittle: Caso armónico (3.11) con tamaño muestral $T = 1024$, bloques de tamaño $N = 256$ y saltos $S = 128$.

Parámetros		Estimación		SD		SD Estimación	
α_0	α_1	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\sigma(\hat{\alpha}_0)$	$\sigma(\hat{\alpha}_1)$	$\hat{\sigma}(\hat{\alpha}_0)$	$\hat{\sigma}(\hat{\alpha}_1)$
0.15	0.10	0.143	0.099	0.024	0.034	0.029	0.035
0.20	0.10	0.195	0.104	0.024	0.034	0.031	0.039
0.20	0.15	0.199	0.152	0.024	0.034	0.030	0.035
0.25	0.10	0.249	0.108	0.024	0.034	0.032	0.041
0.25	0.15	0.251	0.159	0.024	0.034	0.029	0.039
0.25	0.20	0.257	0.203	0.024	0.034	0.029	0.030
0.30	0.10	0.305	0.104	0.024	0.034	0.031	0.040
0.30	0.15	0.306	0.156	0.024	0.034	0.028	0.033
0.35	0.10	0.359	0.102	0.024	0.034	0.029	0.035

que la estimación de los parámetros presentan un pequeño sesgo y las desviaciones estándar estimadas son cercanas a sus valores teóricos.

En resumen, estos experimentos de Monte Carlo sugieren que el desempeño para muestra finita los estimadores de Whittle propuestos son buenos en términos de sesgo y desviaciones estándares. Esto, a pesar de que en muchas de estas simulaciones hemos comprobado la metodología para valores grandes del parámetro de larga memoria d , es decir, los valores son cercanos a 0.5. Por ejemplo, en la Tabla 3.5 la combinación $\alpha_0 = 0,20$ y $\alpha_1 = 0,25$, tiene un promedio de $d(u)$ igual a 0,325 con valor máximo igual a 0.45.

Capítulo 4

Estimador de Kalman

Una alternativa a los estimadores de Whittle propuestos en el capítulo anterior son los estimadores obtenidas a partir de la técnica de Filtro de Kalman. Esta técnica tiene la ventaja de poder trabajar en presencia de datos faltantes y además entrega de forma automática algunos subproductos, por ejemplo los residuos del modelo, sus valores ajustados y la log-verosimilitud.

4.1. Representación de Espacio Estado

Consideremos $Y_{t,T}$ con $t \in \{1, 2, \dots, T\}$ un proceso lineal localmente estacionario de fuerte dependencia, con representación de medias móviles

$$Y_{t,T} = \sigma \left(\frac{t}{T} \right) \sum_{j=0}^{\infty} \eta_j \left(\frac{t}{T} \right) \varepsilon_{t-j}, \quad \text{con} \quad \sum_{j=0}^{\infty} \eta_j^2 \left(\frac{t}{T} \right) < \infty, \quad (4.1)$$

donde $\sigma(t/T)$ es un factor para la varianza del proceso y $\{\varepsilon_t\}$ es una secuencia de ruido blanco normal independiente e idénticamente distribuida con media cero y varianza uno. Si truncamos la expresión (4.1) hasta la m -ésima componente, podemos aproximar $Y_{t,T}$ como

$$Y_{t,T} = \sigma\left(\frac{t}{T}\right) \sum_{j=0}^m \eta_j\left(\frac{t}{T}\right) \varepsilon_{t-j},$$

la cual tiene representación de espacio estado m -dimensional, ver ejemplo Brockwell and Davis (1991) p.469, dada por

$$Y_{t,T} = \sigma(t/T) \begin{bmatrix} \eta_m(t/T) & \eta_{m-1}(t/T) & \cdots & \eta_1(t/T) & 1 \end{bmatrix} X_{t,T} \quad (4.2)$$

donde

$$X_{t,T} = \begin{bmatrix} x_{t-m} \\ x_{t-m-1} \\ \vdots \\ x_t \end{bmatrix} \quad (4.3)$$

y

$$X_{t+1,T} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} X_{t,T} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \varepsilon_{t+1}, \quad (4.4)$$

para $t = 1, \dots, T-1$.

4.2. Estimación

A partir de la Proposición 12.2.2 de Brockwell and Davis (1991) p.467 que suponen una representación de la forma

$$X_{t+1,T} = F_{t,T} X_{t,T} + V_{t,T}, \quad (4.5)$$

$$Y_{t,T} = G_{t,T} X_{t,T} + W_{t,T}, \quad (4.6)$$

para $t = 1, \dots, T-1$, donde

$$E[U_{t,T}] = E \begin{bmatrix} V_{t,T} \\ W_{t,T} \end{bmatrix} = 0, \quad E[U_{t,T} U'_{t,T}] = \begin{bmatrix} Q_{t,T} & S_{t,T} \\ S'_{t,T} & R_{t,T} \end{bmatrix},$$

$X_{1,T}, U_{1,T}, U_{2,T}, \dots$, son no correlacionados e $Y_{0,T}$ satisface

$$Y_{0,T} \perp V_{t,T} \quad \text{e} \quad Y_{0,T} \perp W_{t,T} \quad \forall t \quad (4.7)$$

Entonces el predictor a un paso

$$\hat{X}_{t,T} = P_{t-1,T} X_{t,T},$$

y la matriz de covarianza del error de predicción

$$\Omega_{t,T} = E[(X_{t,T} - \hat{X}_{t,T})(X_{t,T} - \hat{X}_{t,T})'],$$

son determinadas únicamente por las condiciones iniciales,

$$\hat{X}_{1,T} = P(X_{1,T} | Y_{0,T}), \quad \Omega_{1,T} = E[X_{1,T} X'_{1,T}],$$

y las recursiones para $t = 1, \dots, T - 1$,

$$\begin{aligned}
 \Delta_{t,T} &= G_{t,T} \Omega_{t,T} G'_{t,T} + R_{t,T}, \\
 \Theta_{t,T} &= F_{t,T} \Omega_{t,T} G'_{t,T} + S_{t,T}, \\
 \Omega_{t+1,T} &= F_{t,T} \Omega_{t,T} F'_{t,T} + Q_{t,T} - \Theta_{t,T} \Delta_{t,T}^{-1} \Theta'_{t,T}, \\
 \hat{X}_{t+1,T} &= F_{t,T} \hat{X}_{t,T} + \Theta_{t,T} \Delta_{t,T}^{-1} (Y_{t,T} - G_{t,T} \hat{X}_{t,T}),
 \end{aligned} \tag{4.8}$$

donde $\Delta_{t,T}^{-1}$ es la inversa de $\Delta_{t,T}$.

Para el caso de datos faltantes la modificación de las ecuaciones de Kalman, Palma and Chan (1997), son:

$$\Omega_{t+1,T} = \begin{cases} F_{t,T} \Omega_{t,T} F'_{t,T} + Q_{t,T} - \Theta_{t,T} \Delta_{t,T}^{-1} \Theta'_{t,T} & \text{si obs en } t \text{ es conocida} \\ F_{t,T} \Omega_{t,T} F'_{t,T} + Q_{t,T} & \text{si obs en } t \text{ es missing} \end{cases} \tag{4.9}$$

$$\hat{X}_{t+1,T} = \begin{cases} F_{t,T} \hat{X}_{t,T} + \Theta_{t,T} \Delta_{t,T}^{-1} (Y_{t,T} - G_{t,T} \hat{X}_{t,T}) & \text{si obs en } t \text{ es conocida} \\ F_{t,T} \hat{X}_{t,T} & \text{si obs en } t \text{ es missing} \end{cases} \tag{4.10}$$

Para (4.2) a (4.4) se tiene que

$$V_{t,T} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \varepsilon_{t+1}, \quad W_{t,T} = 0 \quad \text{y} \quad F_{t,T} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}, \tag{4.11}$$

para todo $t = 1, \dots, T$.

Ahora si consideremos una muestra $\{Y_{t,T}\}$, para $t = 1, \dots, T$, del proceso (1.6) podemos estimar el vector de parámetros del modelo, denotado por θ , maximizando la función log-Verosimilitud

$$L_T(\theta) = \frac{1}{2} \left\{ \sum_{t=1}^T \log \Delta_{t,T}(\theta) + \sum_{t=1}^T \frac{[Y_{t,T} - \hat{Y}_{t,T}(\theta)]^2}{\Delta_{t,T}(\theta)} \right\} \quad (4.12)$$

Finalmente, el estimador de kalman de los parámetros del vector θ está dado por

$$\hat{\theta}_T = \arg \max L_T(\theta)$$

donde la maximización ocurre sobre el espacio paramétrico Θ .

4.3. Simulación

Para analizar el desempeño de los estimadores de Kalman se simularon 1000 repeticiones de tamaño $T = 1024$ de un procesos LS-FN donde los parámetros $d(u)$ y $\sigma(u)$ son definido como en el caso lineal (3.9). Se procedió además a eliminar aleatoriamente un 10 % y 20 % de los datos para comparar el efecto real en la estimación de este método cuando se está en presencia de datos faltantes. La Tabla 4.1 muestra que las estimaciones obtenidas por el Método de Kalman con una truncación $m = 20$ para diferentes niveles de valores faltantes no difieren mayormente, pero se observa que las desviaciones estándares estimadas aumentan levemente a medida que el nivel de valores faltantes aumenta. Esto es una gran ventaja para el estudio de series de

tiempo que presenten esta característica.

Tabla 4.1: Estimación Método de Kalman para datos faltantes. 1000 repeticiones de tamaño $T = 1024$ y $m = 20$.

Parámetros			Estimación		SD Estimación	
% NA	α_0	α_1	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\sigma}(\hat{\alpha}_0)$	$\hat{\sigma}(\hat{\alpha}_1)$
0 %	0.15	0.25	0.1619	0.2247	0.0495	0.0809
10 %	0.15	0.25	0.1617	0.2248	0.0523	0.0854
20 %	0.15	0.25	0.1642	0.2222	0.0566	0.0922
% NA	β_0	β_1	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\sigma}(\hat{\beta}_0)$	$\hat{\sigma}(\hat{\beta}_1)$
0 %	1.00	0.00	0.9969	0.0073	0.0458	0.0781
10 %	1.00	0.00	0.9462	0.0065	0.0465	0.0799
20 %	1.00	0.00	0.8907	0.0086	0.0464	0.0823

Notar que en la Tabla 3.4 la combinación $\alpha_0 = 0.15$ y $\alpha_1 = 0.25$ con $\sigma(u) = 1$ los estimadores de Whittle presentan un menor sesgo, pero los estimadores de Kalman presentan una desviación estándar estimada considerablemente menor incluso para un nivel de valores faltantes del 20 %.

Capítulo 5

Ilustración

En este capítulo se ilustrará la aplicación de la metodología de larga memoria localmente estacionaria propuesta en el Capítulo 3 a través del análisis de dos series de tiempo reales. El primer ejemplo consiste en una serie temporal sobre depósitos de minerales formados en cavernas (Espeleotema). Un ejemplo de estas formaciones son las estalactitas y estalacmitas. El segundo ejemplo corresponde a los anchos anuales de los anillos de un árbol. Una motivación para el estudio de estas dos series de tiempo es el análisis de variables climáticas, ya que el espesor de los depósitos de las capas y los anchos de anillos de árboles son utilizados con frecuencia en muchas disciplinas como proxies climáticos, ver por ejemplo Tan et al. (2003) y sus referencias.

En las aplicaciones discutidas en este capítulo se utilizará el modelo LS-ARFIMA

$(0, d, 0)$ - (p, q) descritos en la Sección 1.2.2, donde el parámetro de larga memoria y el factor de escala para la variabilidad del proceso son explicados por un polinomio de orden p y q respectivamente,

$$\begin{aligned} d(u) &= \alpha_0 + \alpha_1 u + \cdots + \alpha_p u^p, \\ \sigma^2(u) &= (\beta_0 + \beta_1 u + \cdots + \beta_q u^q)^2. \end{aligned}$$

En este caso, el vector de parámetros es $\theta = (\alpha_0, \alpha_1, \dots, \alpha_p, \beta_0, \beta_1, \dots, \beta_q)$. Los modelos son seleccionados por el Criterio de Información de Akaike (AIC) y por el análisis de significancia de los parámetros involucrados. El AIC está dado por

$$\text{AIC}(k) = -2 L_T(\hat{\theta}) + 2(p + q + 2),$$

donde L_T es la función log-verosimilitud de Whittle definida en (3.1). Notar que esta clase de modelos LS-ARFIMA contienen, como caso particular, la familia de procesos ARFIMA estacionarios. Por lo tanto, observando el AIC y la significancia de los parámetros, es posible tener una idea acerca de si un modelo estacionario o un modelo localmente estacionario se ajusta mejor a los datos. Los dos ejemplos de la vida real presentados a continuación son tomados sólo como ejemplos de aplicación de la metodología desarrollada en este trabajo. En estos dos casos, los modelos localmente estacionarios parecen ajustarse mejor a las series de tiempo que los modelos estacionarios. Sin embargo, de ninguna manera estamos estableciendo que esos son los únicos modelos apropiados. Por otra parte, esta clase de procesos ofrecen una flex-

ibilidad y un enfoque parsimonioso para el problema. Es flexible porque permite el ajuste de una gran clase de parámetros que evolucionan según un patrón que depende de la elección adecuada de las funciones de base. Por otro lado, se trata de un enfoque parsimonioso ya que la selección de un modelo por AIC busca el modelo más simple que ajuste de mejor manera a los datos observados. Por otra parte, la estructura de un LS-ARFIMA permite evaluar las hipótesis estadísticas sobre los modelos propuestos por medio de las propiedades establecidas en los Teorema 3.2.1 a Teorema 3.2.3. En particular, esta estructura permite la comparación estadística del modelo ARFIMA estacionario y localmente estacionario, ya que están anidados en una gran clase de procesos definidos por (1.6).

5.1. Espeleotema

La Figura 5.1 muestra un serie de tiempo de 2650 años de registros de depósitos en las cavidades en una cueva, lo que se conoce como espeleotema. Este conjunto de datos contiene el espesor de las capas de estalagmitas de la cueva Shihua, Beijing, China Shihua Cave, a partir de 665 aC a 1985 dC, ver Tan et al. (2003) para más detalles.

La estructura de autocorrelación de esta serie se explora en la Figura 5.2. El gráfico (a) muestra la función de autocorrelación (ACF) empírica de la serie, mientras que en (b) se muestra el gráfico de varianzas, ver (Beran, 1994, p.92-94) para más detalles.

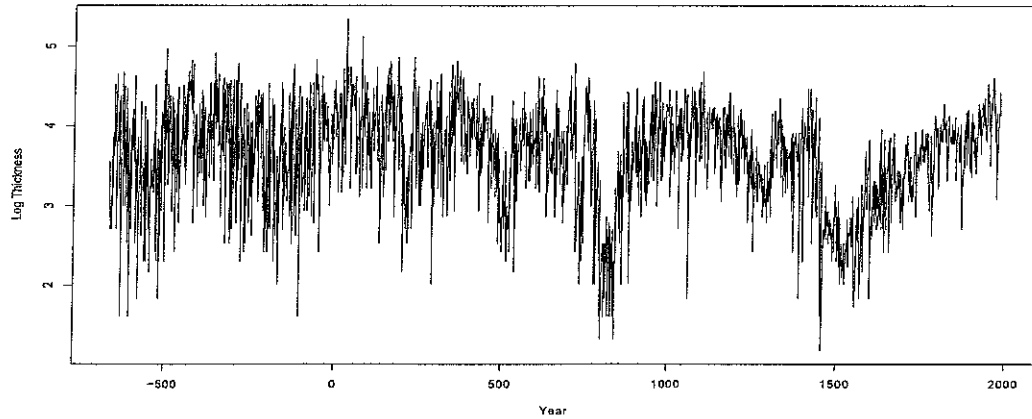


Figura 5.1: Depósitos de Minerales.

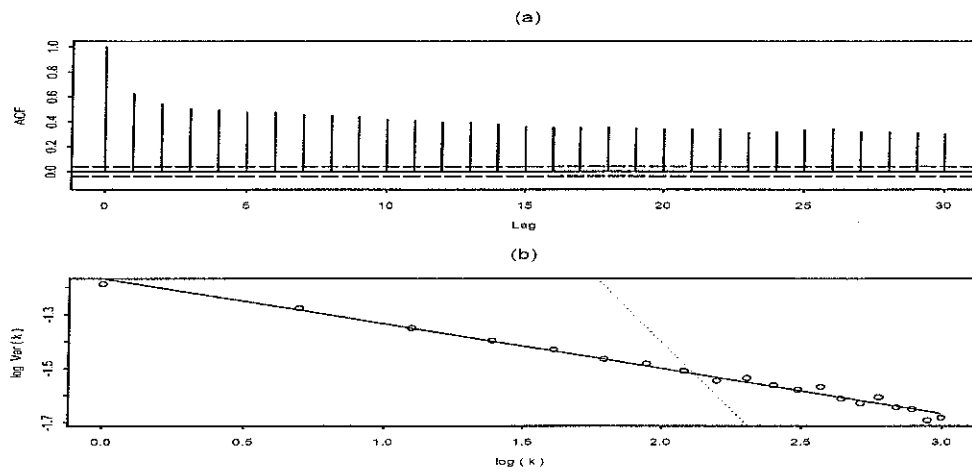


Figura 5.2: Depósitos de Minerales. (a) ACF muestral, (b) Var-Plot.

En este último gráfico, la línea punteada representa el comportamiento esperado de la varianza muestral promedio de un bloques de k observaciones para el caso de corta memoria. Por otro lado, la línea continua representa el comportamiento esperado de la varianza de un proceso de larga memoria. En ambos gráficos, esta serie muestra

una dependencia de largo plazo. Sin embargo, un vistazo más de cerca a la ACF de los datos revela que el grado de dependencia no parece ser constante en el tiempo. En particular, el gráfico (a) de la Figura 5.3 muestra que el ACF muestral de las primeras 500 observaciones, $\{Y_1, \dots, Y_{500}\}$, el (b) reporta el ACF muestral de 500 observaciones de la parte media de la serie, $\{Y_{1075}, \dots, Y_{1574}\}$, mientras que el (c) reporta el ACF muestral de las últimas 500 observaciones $\{Y_{2151}, \dots, Y_{2650}\}$. Notar que el nivel de dependencia en las autocorrelaciones es diferente en cada tramo de los datos observados. La autocorrelación parece ser cada vez más fuerte en el tiempo. Por otra parte, notar que la varianza de la serie decrece en el tiempo (Figura 5.2 panel (b)). Siendo consistente con el análisis exploratorio, se propone un ajuste localmente estacionario a esta serie de tiempo, ya que esta clase de procesos podría captar la no estacionaridad

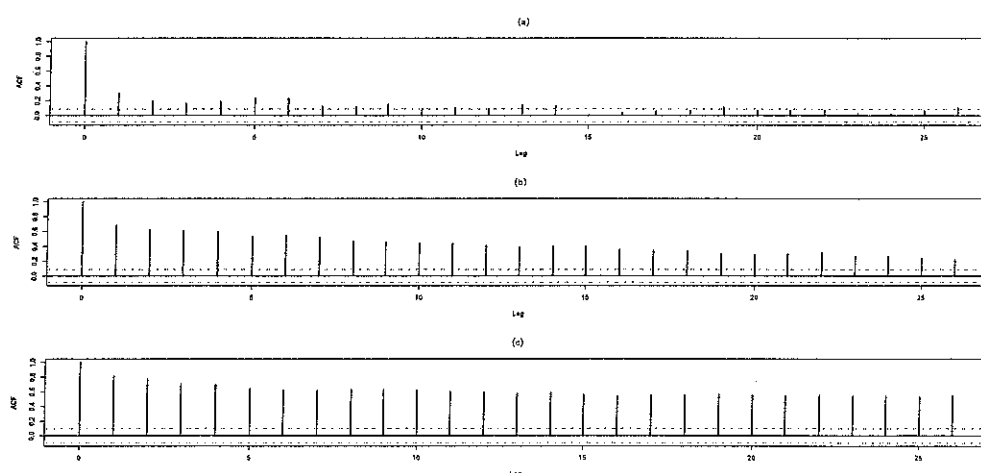


Figura 5.3: Depósitos de Minerales. ACF muestral: (a) Observaciones 1 a 500, (b) observaciones 1075 a 1575, y (c) observaciones 2150 a 2650.

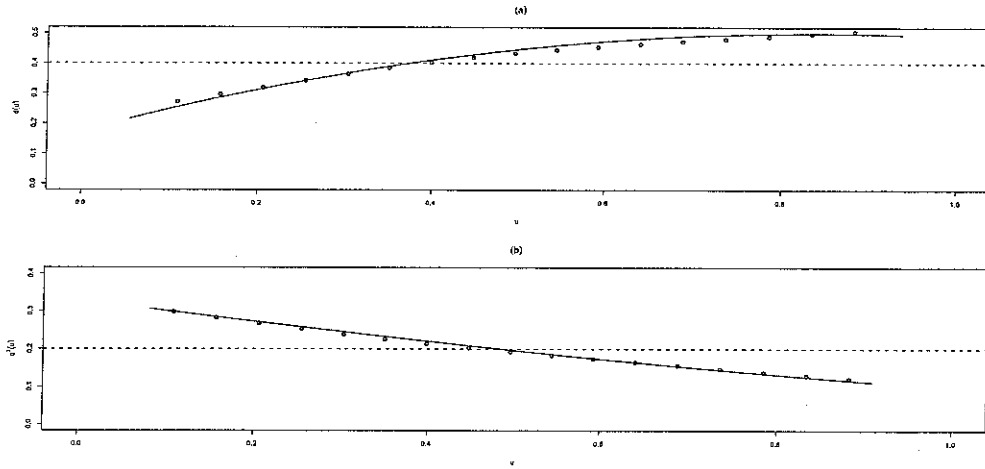


Figura 5.4: Depósitos de Minerales. (a) Estimación del parámetro de larga memoria. (b) Las estimaciones de la varianza del ruido. En ambos paneles la línea continua representa el modelo ARFIMA localmente estacionario, la línea quebrada horizontal indica el modelo ARFIMA estacionario y los puntos representan el suavizamiento por kernel.

Tabla 5.1: Depósitos de Minerales: Modelo ajustado

Parámetros	Estimación	SD	z-value
α_0	0.1763	0.0454	3.8796
α_1	0.7730	0.2099	3.6832
α_2	-0.4615	0.2032	-2.2709
β_0	0.5753	0.0137	42.007
β_1	-0.2650	0.0206	-12.853

de la varianza, así como la no estacionaridad del parámetro Hurst, todo esto dentro de un modelo simplificado. El modelo seleccionado por AIC y según significancia de los parámetros involucrados es el LS-ARFIMA(0, d , 0)-(2, 1) con valores estimados

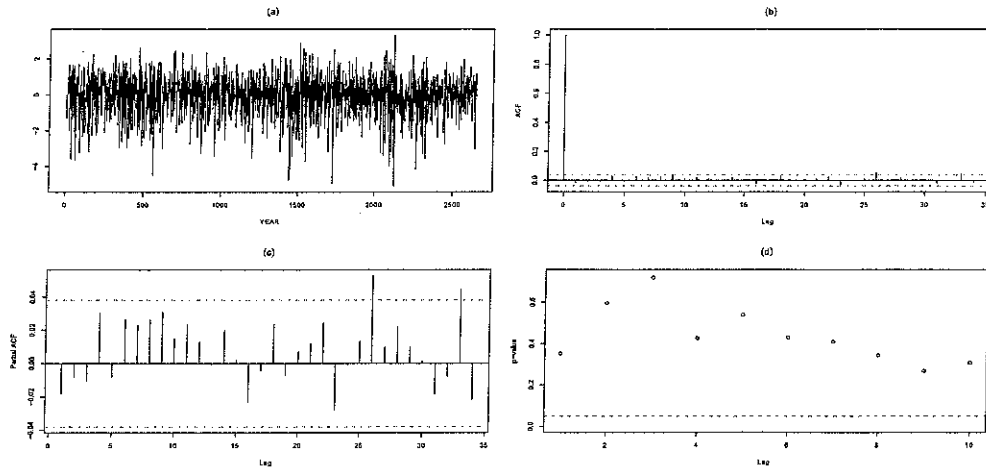


Figura 5.5: Depósitos de Minerales: Análisis Residual. (a) Residuos obtenidos del ajuste del modelo, (b) ACF muestral, (c) PACF muestral, (d) test de Ljung-Box.

en la Tabla 5.1. Tener en cuenta que, según la cuarta columna de esta tabla, todos los parámetros de este modelo son estadísticamente significativa al 5 %. La evolución del parámetro de larga memoria $d(u)$ y el factor de escala $\sigma^2(u)$ son representados en la Figura 5.4 parte (a) y (b) respectivamente. Para efectos de comparación, estos gráficos muestran tres métodos de estimación: el proceso localmente estacionario LS-ARFIMA, el proceso estacionario ARFIMA y un suavizamiento de una estimación por kernel. El modelo LS-ARFIMA fue ajustado por medio del paquete implementado para el software computacional *R* que se encuentra en el Apéndice B. El modelo ARFIMA estacionario fue ajustado usando el método Haslett-Raftery en *R*. Por otra parte, la estimación suavizada por Kernel es una aproximación heurística que permite una simple y directa visión sobre la estructura dinámica de los parámetros, ver

por ejemplo Görg (2007). Esta aproximación consiste en estimar los valores de los parámetros d y σ por bloques de N observaciones desplazados cada vez en S valores, y luego suavizar los valores de los parámetros por medio de un procedimiento tipo kernel. En este caso, se utilizó el método de suavizamiento de spline cúbica implementado en R , ver por ejemplo Hastie and Tibshirani (1990) y Chambers and Hastie (1992).

En la Figura 5.4, la línea continua representa los valores estimados para los parámetros $d(u)$ y $\sigma^2(u)$ para $u \in [0, 1]$. La línea punteada representan estos valores para un proceso ARFIMA(0, d , 0) estacionarios, ajustado a los mismos datos asumiendo d y σ como constantes. En este caso, las estimaciones de estos parámetros son $\hat{d} = 0.4005$ y $\hat{\sigma} = 0.4474$. Por otra parte, los puntos representan la estimación suavizada por kernel. Como era de esperar del análisis exploratorio, el parámetro de fuerte dependencia aumenta en el tiempo, mientras que la varianza disminuye con el transcurrir del tiempo.

Un análisis de los residuos de este ajuste es presentado en la Figura 5.5. En (a) se muestran los residuos estandarizados obtenidos al ajustar el modelo LS-ARFIMA. En (b) y (c) se pueden observar el ACF y PACF muestral, respectivamente. Por otra parte, en (d) se grafican los valores- p del estadístico de Box-Ljung (portmanteau)

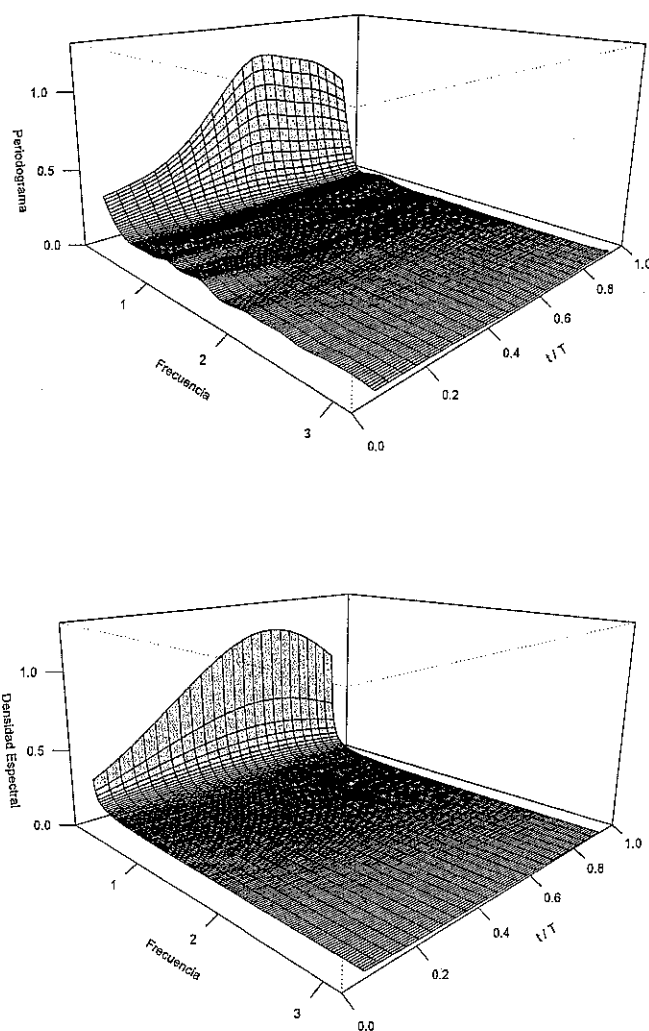


Figura 5.6: Depósitos de Minerales: Periodograma vs. Densidad espectral teórica.

para probar la blancura de estos residuos, donde la hipótesis que los residuos son ruidos blancos no es rechazada para todos los largos considerados, a un nivel del 5% de significancia. Las Figura 5.6 muestra la evolución del Periodograma y de la densidad espectral teórica del ajuste presentado en la Tabla 5.1 en el tiempo. Se

puede apreciar que los estimadores de Whittle logran representar de buena manera la estructura espectral de los datos.

5.2. Anillos de árbol

La Figura 5.7 muestra los anchos anuales de los anillos de un árbol *Pinus Longaeva* medidos en Mammoth Creek, Utah, a partir del año 0 al 1989. Estos datos, disponibles en el Centro Nacional de Datos Climáticos y la librería de datos de series de tiempo, son reportados por Graybill (1990). El ACF muestral de estos datos, está representada en la Figura 5.8 panel (a), la cual muestra una autocorrelación de largo alcance. Además, la varianza correspondiente, ver panel (b) de la misma figura, nos entrega una prueba más del presencia de larga memoria. Sin embargo, una mirada más atenta a la estructura de autocorrelación de la muestra, ver la Figura 5.9, sugiere que la fuerte dependencia decae lo largo del tiempo. De hecho, los valores de la ACF muestral de las primeras 500 observaciones, ver (a), son más fuertes que los correspondientes valores muestra del ACF para los otros dos períodos considerados, véase (b) (c).

Para tomar en cuenta esta pérdida de memoria en el tiempo se propone una clase de proceso LS-ARFIMA. El modelo elegido por la AIC y significancia de los parámetros es el LS-ARFIMA(0, d , 0)-(1, 2) con estimaciones reportadas en la Tabla 5.2. Notar que, según la cuarta columna de esta tabla, todos los parámetros de este modelo son estadísticamente significativos al 5 %.

Los paneles (a) y (b) de la Figura 5.10 muestran la evolución del parámetro de larga memoria $d(u)$ y la varianza $\sigma^2(u)$, respectivamente. Al igual que los datos ilustrados anteriormente se incluyen tres criterios de estimación: modelo LS-ARFIMA no estacionario, procesos ARFIMA estacionario y suavizado por kernel. En la Figura 5.10, la línea continua representa los valores estimados para los parámetros $d(u)$ y $\sigma^2(u)$ para $u \in [0, 1]$. Las líneas punteadas representan estos valores para un ajuste de un proceso ARFIMA(0, d , 0) estacionario, asumiendo d y σ constante. En este caso, las estimaciones de estos parámetros son $\hat{d} = 0.2415$ y $\hat{\sigma} = 0.3287$. Por otra parte, los puntos representan la estimación suavizada por kernel. Tal y como sugiere el análisis exploratorio, el modelo ajustado indica que el parámetro de larga dependencia decae en el tiempo. Por otro lado, la varianza parece tener una evolución más compleja.

Además, la Figura 5.11 exhibe cuatro gráficos que muestran la estructura de los

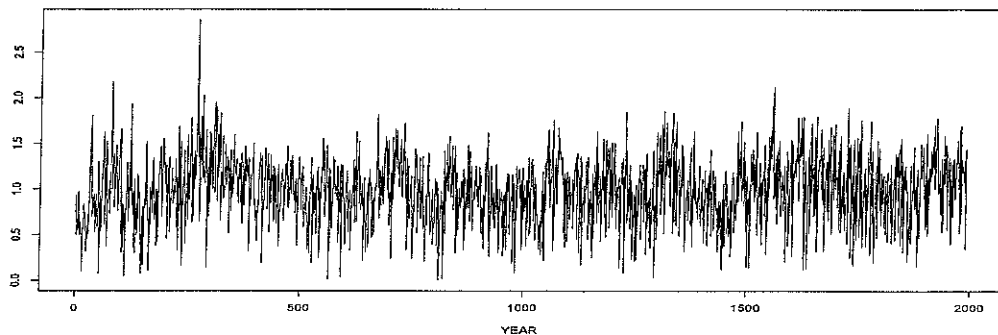


Figura 5.7: Anillos de árbol.

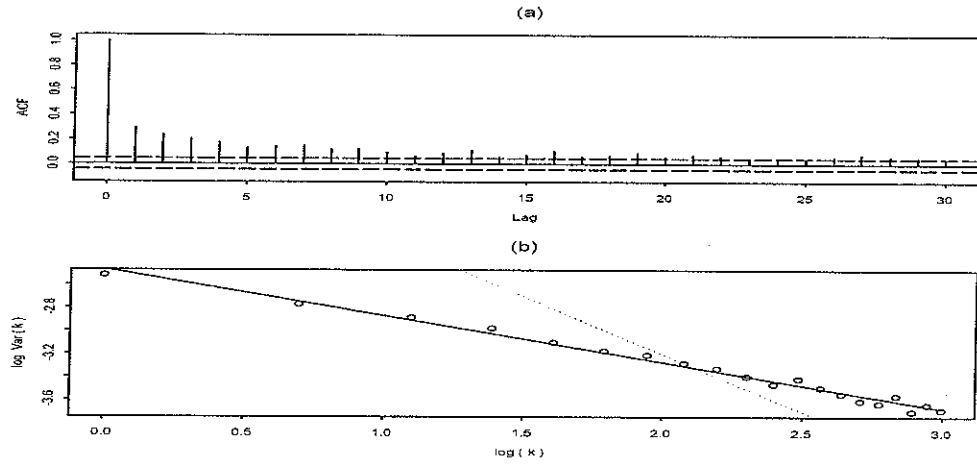


Figura 5.8: Anillos de árbol: (a) ACF muestral, (b) Var-Plot.

residuos. En (a) se muestran los residuos del ajuste del modelo LS-ARFIMA. En (b) y (c) se presentan los ACF y PACF muestrales, respectivamente, mientras que en

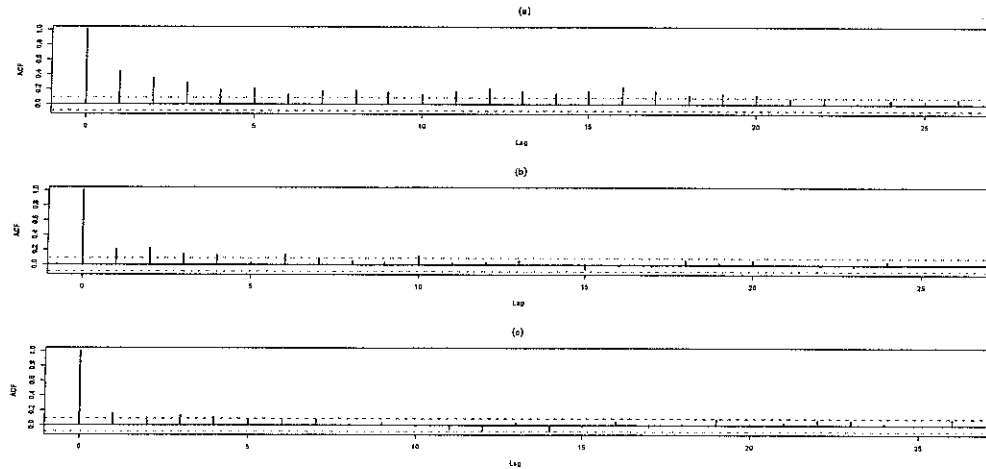


Figura 5.9: Anillos de árbol. ACF muestral: (a) Observaciones 1 a 500, (b) observaciones 750 a 1250, y (c) observaciones 1490 a 1990.

Tabla 5.2: Anillos de árbol: Modelo ajustado

Parámetros	Estimación	SD	z-value
α_0	0.3285	0.0349	9.3998
α_1	-0.1890	0.0605	-3.1223
β_0	0.3609	0.0162	22.2197
β_1	-0.3087	0.0753	-4.0983
β_2	0.3884	0.0749	5.1877

(d) se exhibe el estadístico Box-Ljung para la blancura. Por otra parte, la autocorrelación y la autocorrelación parcial se observa en (b) y (c) que no es significativa.

Esta conclusión es corroborada por el el test de Box-Ljung que indica la hipótesis de

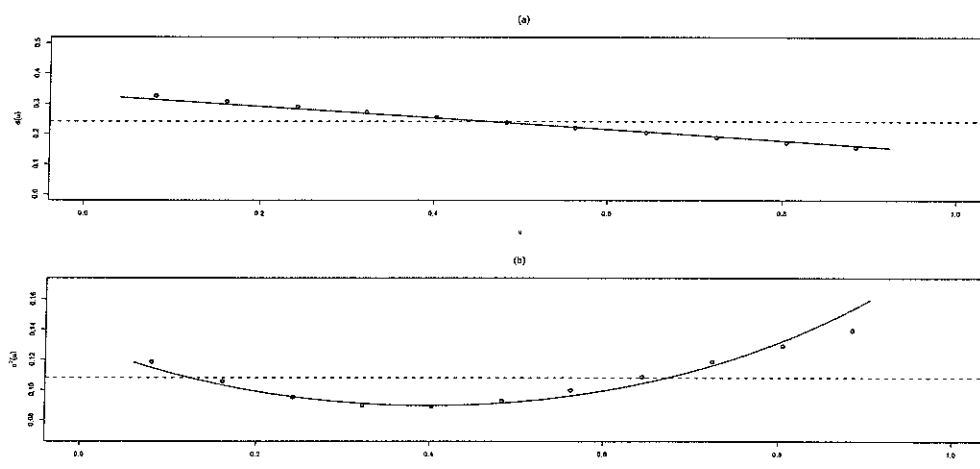


Figura 5.10: Anillos de árbol. a) Estimación del parámetro de larga memoria. (b) Estimación de la varianza del ruido. En ambos paneles la línea continua representa el modelo ARFIMA localmente estacionarios, la línea quebrada horizontal indica el modelo ARFIMA estacionario y los puntos representan el suavizamiento por kernel.

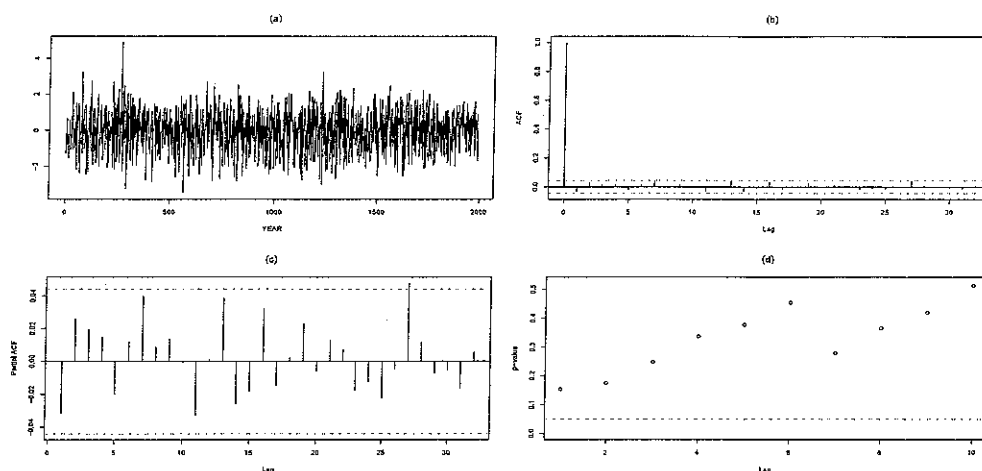


Figura 5.11: Anillos de árbol: Análisis Residual. (a) Residuos del modelo ajustado, (b) ACF muestral, (c) PACF muestral, (d) test de Ljung-Box.

ruido blanco de estos residuos no es rechazada para todos los rezagos considerados a un nivel del 5 % de significancia.

En las dos aplicaciones de la vida real examinadas anteriormente, las series de tiempo parecen exhibir una estructura de larga dependencia y heterocedasticidad dinámica en el tiempo, para lo cual no es adecuado ajustar por un proceso ARFIMA estacionario. Por otra parte, que la estructura de dependencia y la evolución del parámetro de escala son ajustados de mejor manera por un modelo ARFIMA localmente estacionario. Por lo tanto, estos ejemplos ponen de manifiesto que los procesos de larga memoria localmente estacionarios pueden ser muy útiles para modelar series temporales de datos con estructuras complejas variantes en el tiempo. Las Figura

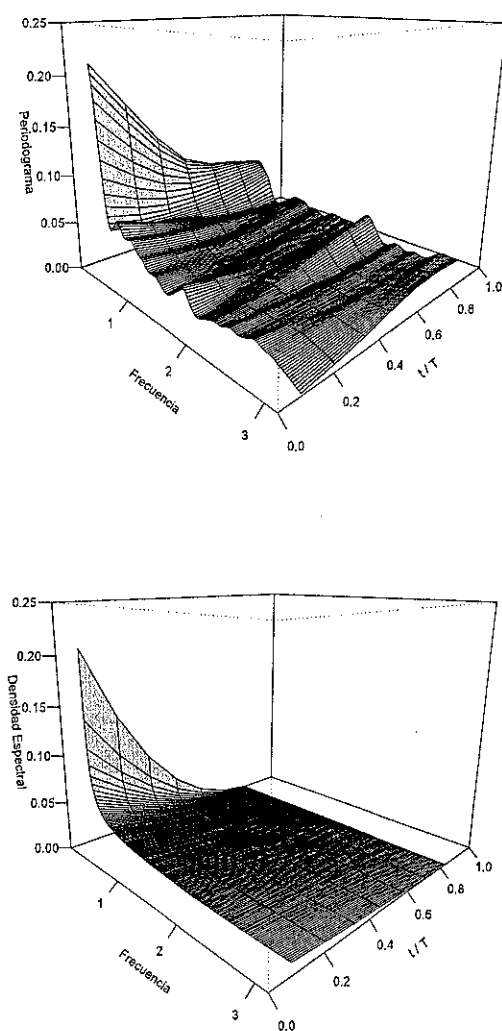


Figura 5.12: Anillos de árbol: Periodograma vs. Densidad espectral teórica.

5.12 muestra la evolución del Periodograma y de la densidad espectral teórica del ajuste presentado en la Tabla 5.2. Se puede apreciar que los estimadores de Whittle nuevamente logran representar de buena manera la estructura espectral de los datos.

Capítulo 6

Conclusiones

Una nueva clase de procesos de larga memoria se han introducido en este trabajo, que es capaz de modelar series de tiempo que exhiben no estacionaridad, heteroscedasticidad y fuerte dependencia dinámica en el tiempo. La metodología de estimación de Whittle que se propone es computacionalmente eficiente y se ha demostrado que estos estimadores poseen propiedades asintóticas muy deseables, como son la consistencia, normalidad y eficiencia. Por otra parte, varias simulaciones de Monte Carlo indican que las estimaciones realizadas son muy buenas incluso para tamaños muestrales relativamente pequeños en el contexto de larga memoria. Además, las aplicaciones de estas metodologías a los dos ejemplos de la vida real muestran que la nueva metodología propuesta en este documento ofrece una valiosa herramienta estadística para modelar estructuras de fuerte dependencia variantes en el tiempo. Por último,

nos gustaría señalar algunos temas interesantes que podrían valer la pena explorar en futuras investigaciones. Por ejemplo, es importante investigar el efecto en la estimación del parámetro para una media variante en el tiempo. Esto es de interés práctico en algunas aplicaciones, se requiere un detallado análisis de la validez de la simulación teórica y resultados obtenidos para el caso de media cero. Otros temas relevantes son los análisis del problema de predicción en el contexto de procesos de fuerte dependencia localmente estacionaria y la estimación de procesos localmente estacionarios con parámetro de larga memoria $d(u)$ que queden fuera del intervalo $(0, \frac{1}{2})$ para algún $u \in [0, 1]$. Finalmente, queda planteado el problema del desarrollo de la teoría asintótica de los estimadores de Kalman los cuales de acuerdo a las simulaciones mostradas en el Capítulo 4 parecen ser insesgadas en presencia de datos faltantes y ser más eficientes que los estimadores de Whittle.

Apéndice A

Apéndice Técnico

A continuación se presentan las demostraciones de los Teoremas propuestos en el Capítulo 3 junto con las proposiciones y lemas necesarios para ello. Esté apéndice se encuentra en inglés ya que estas demostraciones están en proceso de ser sometidas.

The theoretical results discussed in Section 3.2.2 which establish the asymptotic behavior of the Whittle estimators are proved in this appendix. Consider the function $\phi : [0, 1] \times [-\pi, \pi] \rightarrow \mathbb{R}$ and define the functional operator

$$J(\phi) = \int_0^1 \int_{-\pi}^{\pi} \phi(u, \lambda) f(u, \lambda) d\lambda du, \quad (\text{A.1})$$

where f is the spectral density of the process described by (1.1). Define the sample

version of $J(\cdot)$ as

$$J_T(\phi) = \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) I_N(u_j, \lambda) d\lambda, \quad (\text{A.2})$$

where M and u_j , $j = 1, \dots, M$ are given in Section 2. Furthermore, define the matrix

$$Q(u) = \left(\int_{-\pi}^{\pi} \phi(u, \lambda) e^{i\lambda(s-t)d\lambda} \right)_{s,t=1,\dots,N}, \quad (\text{A.3})$$

and the block-diagonal matrix $Q(\phi) = \text{diag}(Q(u_1), \dots, Q(u_M))$.

This appendix is structured in three parts. The first part is devoted to proving three fundamental propositions about the operator (A.2), while the second part presents and proves eight technical results summarized in Lemma 1 to Lemma 8. The third part of the appendix is devoted to the proofs of Theorem 1 to Theorem 4. In what follows, K is always a positive constant that could vary from line to line.

A.1. Proposiciones

Proposición A.1.1 *Let $f(u, \lambda)$ be a time-varying spectral density satisfying assumption A1 and assume that the function $\phi(u, \lambda)$ appearing in (A.1) is symmetric in λ and twice differentiable function with respect to u . Let $\widehat{f}(u, k)$ and $\widehat{\phi}(u, k)$ be their Fourier coefficients, respectively. If there is a positive constant K such that*

$$|\widehat{f}(u, k) \widehat{\phi}(u, k)| \leq K \left(\frac{\log k}{k^2} \right)$$

for all $u \in [0, 1]$ and $k > 1$, then, under assumptions A2 and A3 we have that

$$E[J_T(\phi)] = J(\phi) + \mathcal{O}\left(\frac{\log^2 N}{N}\right) + \mathcal{O}\left(\frac{1}{M}\right).$$

Demostración A.1.1 From the definition (A.2) we can write

$$\begin{aligned} E[J_T(\phi)] &= \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) E[I_N(u_j, \lambda)] d\lambda \\ &= \frac{1}{2\pi M H_{2,N}(0) N} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) E[|D_N(u_j, \lambda)|^2] d\lambda \\ &= \frac{1}{2\pi M H_{2,N}(0)} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) \sum_{t, s=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) c(u_j, t, s) e^{i\lambda(s-t)} d\lambda, \end{aligned}$$

where

$$c(u, t, s) = E\left(Y_{[uT]-N/2+t+1, T} Y_{[uT]-N/2+s+1, T}\right).$$

Thus,

$$\begin{aligned} E[J_T(\phi)] &= \frac{1}{2\pi M H_{2,N}(0)} \sum_{j=1}^M \sum_{t, s=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) c(u_j, t, s) \int_{-\pi}^{\pi} \phi(u_j, \lambda) e^{i\lambda(s-t)} d\lambda \\ &= \frac{1}{2\pi M H_{2,N}(0)} \sum_{j=1}^M \sum_{t, s=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) c(u_j, t, s) \widehat{\phi}(u_j, s-t) \\ &= \frac{1}{2\pi M H_{2,N}(0)} \sum_{j=1}^M \sum_{t, s=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N} + \frac{k}{N}\right) c(u_j, t, t+k) \widehat{\phi}(u_j, k) (2 - \delta_k), \end{aligned}$$

where $\delta_k = 1$ for $k = 0$ and $\delta_k = 0$ for $k \neq 0$. By assumption A2 and Taylor's theorem

$$h\left(\frac{t}{N} + \frac{k}{N}\right) = h\left(\frac{t}{N}\right) + h'(\xi_{t,k,N}) \frac{k}{N},$$

for some $\xi_{t,k,N} \in \left(\frac{t}{N}, \frac{t+k}{N}\right)$, for $k \geq 0$. Thus,

$$\begin{aligned} E[J_T(\phi)] &= \frac{1}{2\pi M H_{2,N}(0)} \sum_{j,t,k} h^2\left(\frac{t}{N}\right) c(u_j, t, t+k) \widehat{\phi}(u_j, k) (2 - \delta_k) \\ &+ \frac{1}{2\pi M H_{2,N}(0)} \sum_{j,t,k} h\left(\frac{t}{N}\right) h'(\xi_{t,k,N}) c(u_j, t, t+k) \widehat{\phi}(u_j, k) (2 - \delta_k). \end{aligned} \quad (\text{A.4})$$

Under assumption A1 we can expand $c(u, t, t+k)$ by Taylor's theorem as

$$\begin{aligned} c(u, t, t+k) &= \widehat{f}(u, k) + \widehat{f}(u, k) \varphi_1(u, k) d'(u) \left(\frac{t+1-N/2}{T}\right) \\ &+ \widehat{f}(u, k) \varphi_2(u, k) d'(u) \left(\frac{k}{T}\right) + R(u, t, k, N, T) \end{aligned}$$

where the remainder term is given by

$$R(u, t, k, N, T) = \mathcal{O} \left\{ \widehat{f}(u, k) \left[\left(\frac{k}{T}\right)^2 + \left(\frac{t}{T}\right)^2 \right] \log^2 k \right\},$$

with $|\varphi_1(u, k)| \leq K \log k$ and $\varphi_2(u, k) \leq K \log k$ for $k > 1$. Consequently, we can

write

$$\begin{aligned} &\sum_{h=0}^{N-t} c(u_j, t, t+k) \widehat{\phi}(u_j, k) (2 - \delta_k) \\ &= \sum_{k=0}^{N-t} \widehat{f}(u_j, k) \widehat{\phi}(u_j, k) (2 - \delta_k) \\ &+ \sum_{k=0}^{N-t} \widehat{f}(u_j, k) \widehat{\phi}(u_j, k) \varphi_1(u_j, k) (2 - \delta_k) d'(u_j) \left(\frac{t+1-n/2}{T}\right) \\ &+ \sum_{k=0}^{N-t} \widehat{f}(u_j, k) \widehat{\phi}(u_j, k) \varphi_2(u_j, k) (2 - \delta_k) d'(u_j) \frac{k}{T} \\ &+ \sum_{k=0}^{N-t} R(u_j, t, k, N, T) \widehat{\phi}(u_j, k) (2 - \delta_k). \end{aligned} \quad (\text{A.5})$$

Since by assumption $|\widehat{f}(u, k)\widehat{\phi}(u, k)| \leq K \log k/k^2$, for $k > 1$, uniformly in $u \in [0, 1]$,

we conclude that there is a limit $A(u) < \infty$ such that

$$A(u) = \lim_{N \rightarrow \infty} \sum_{k=0}^N \widehat{f}(u, k)\widehat{\phi}(u, k)(2 - \delta_k).$$

Consequently,

$$\begin{aligned} & \sum_{t=1}^{N-1} h^2 \left(\frac{t}{N} \right) \sum_{k=0}^{N-t} \widehat{f}(u_j, k)\widehat{\phi}(u_j, k)(2 - \delta_k) \\ &= \sum_{t=1}^{N-1} h^2 \left(\frac{t}{N} \right) \sum_{k=0}^{N-1} \widehat{f}(u_j, k)\widehat{\phi}(u_j, k)(2 - \delta_k) - \sum_{t=1}^{N-1} h^2 \left(\frac{t}{N} \right) \sum_{k=N-t+1}^{N-1} \widehat{f}(u_j, k)\widehat{\phi}(u_j, k)(2 - \delta_k) \\ &= \sum_{t=1}^{N-1} h^2 \left(\frac{t}{N} \right) \left[A(u_j) - \sum_{k=N}^{\infty} \widehat{f}(u_j, k)\widehat{\phi}(u_j, k)(2 - \delta_k) \right] + \mathcal{O}(\log^2 N), \end{aligned}$$

by Lemma A.2.7. Hence,

$$\begin{aligned} & \sum_{t=1}^{N-1} h^2 \left(\frac{t}{N} \right) \sum_{k=0}^{N-t} \widehat{f}(u_j, k)\widehat{\phi}(u_j, k)(2 - \delta_k) = A(u_j) \sum_{t=1}^{N-1} h^2 \left(\frac{t}{N} \right) \\ &= \sum_{t=1}^{N-1} h^2 \left(\frac{t}{N} \right) \sum_{k=N}^{\infty} \widehat{f}(u_j, k)\widehat{\phi}(u_j, k)(2 - \delta_k) + \mathcal{O}(\log^2 N). \end{aligned}$$

But,

$$\left| \sum_{k=N}^{\infty} \widehat{f}(u_j, k)\widehat{\phi}(u_j, k)(2 - \delta_k) \right| < K \sum_{k=N}^{\infty} \frac{\log k}{k^2} = \mathcal{O} \left(\frac{\log N}{N} \right),$$

and consequently,

$$\left| \sum_{t=1}^{N-1} h^2 \left(\frac{t}{N} \right) \sum_{k=N}^{\infty} \widehat{f}(u_j, k)\widehat{\phi}(u_j, k)(2 - \delta_k) \right| = \mathcal{O}(\log N).$$

Therefore,

$$\sum_{t=1}^{N-1} h^2 \left(\frac{t}{N} \right) \sum_{k=0}^{N-t} \widehat{f}(u_j, k)\widehat{\phi}(u_j, k)(2 - \delta_k) = A(u_j) \sum_{t=1}^{N-1} h^2 \left(\frac{t}{N} \right) + \mathcal{O}(\log^2 N).$$

On the other hand, by analyzing the term involving the second summand of (A.5) we get

$$\begin{aligned}
& \sum_{t=0}^{N-1} h^2 \left(\frac{t}{N} \right) \sum_{k=0}^{N-t} \varphi_1(u_j, k) \widehat{f}(u_j, k) \widehat{\phi}(u_j, k) (2 - \delta_k) \left(\frac{t+1-N/2}{T} \right) d'(u_j) \\
&= \sum_{t=0}^{N-1} h^2 \left(\frac{t}{N} \right) \left[\sum_{k=0}^{N-1} \varphi_1(u_j, k) \widehat{f}(u_j, k) \widehat{\phi}(u_j, k) (2 - \delta_k) \right] \left(\frac{t+1-N/2}{T} \right) d'(u_j) \\
&+ \sum_{t=1}^{N-1} h^2 \left(\frac{t}{N} \right) \sum_{k=N-t+1}^{N-1} \varphi_1(u_j, k) \widehat{\phi}(u_j, k) \widehat{f}(u_j, k) \left(\frac{t+1-N/2}{T} \right) d'(u_j) \\
&= \left[\sum_{t=0}^{N-1} h^2 \left(\frac{t}{N} \right) \left(\frac{t+1-N/2}{T} \right) \right] \left[\sum_{k=0}^{N-1} \varphi_1(u_j, k) \widehat{f}(u_j, k) \widehat{\phi}(u_j, k) (2 - \delta_k) \right] d'(u_j) \\
&+ \mathcal{O} \left(\frac{N \log^2 N}{T} \right),
\end{aligned}$$

by Lemma A.2.8. Now, since $h(\cdot)$ is symmetric around $1/2$, we have

$$\sum_{t=0}^{N-1} h^2 \left(\frac{t}{N} \right) \left(\frac{t+1-N/2}{T} \right) = \mathcal{O} \left(\frac{1}{T} \right).$$

Besides, $|\sum_{k=0}^{N-1} \varphi_1(u_j, k) \widehat{f}(u_j, k) \widehat{\phi}(u_j, k) (2 - \delta_k)| \leq K \sum_{k=1}^N \frac{(\log k)^2}{k^2} < \infty$, and $|d'(u_j)| <$

K , $j = 1, \dots, M$, by assumption A2. Consequently,

$$\sum_{t=0}^{N-1} h^2 \left(\frac{t}{N} \right) \sum_{k=0}^{N-t} \varphi_1(u_j, k) \widehat{f}(u_j, k) \widehat{\phi}(u_j, k) (2 - \delta_k) \left(\frac{t+1-N/2}{T} \right) d'(u_j) = \mathcal{O} \left(\frac{N}{T} \log^2 N \right).$$

The third term of (A.5) can be bounded as follows:

$$\sum_{k=0}^{N-t} \varphi_2(u_j, k) \widehat{f}(u_j, k) \widehat{\phi}(u_j, k) k \leq K \sum_{k=1}^N \frac{\log k}{k} \leq K \log^2 N,$$

and then

$$\left| \sum_{t=?}^{N-1} h^2 \left(\frac{t}{N} \right) \sum_{k=0}^{N-t} \varphi_2(u_j, k) \widehat{f}(u_j, k) \widehat{\phi}(u_j, k) \frac{k}{T} d'(u_j) \right| \leq K \frac{N}{T} \log^2 N.$$

The last term of (A.5) can be bounded as follows:

$$\left| \sum_{k=0}^{N-t} R(u_j, t, k, N, T) \widehat{\phi}(u_j, k) (2 - \delta_k) \right| \leq K \left(\frac{N}{T} \right) \log^2 N \sum_{k=1}^N \frac{1}{k^2} \leq K \log^2 N \left(\frac{N}{T} \right)^2,$$

and then

$$\left| \sum_{t=0}^{N-1} \sum_{k=0}^{N-t} h^2 \left(\frac{t}{N} \right) R(u_j, k, N, T) \widehat{\phi}(u_j, k) (2 - \delta_k) \right| \leq K \frac{N^3}{T^2} \log^2 N.$$

Note that by assumption A3, the term above converges to zero as $N, T \rightarrow \infty$. Therefore, the first term in (A.4) can be written as

$$\frac{1}{2\pi M H_{2,N}(0)} \sum_{j,t,k} h^2 \left(\frac{t}{N} \right) c(u_j, t, t+h) \widehat{\phi}(h) (2 - \delta_k) = \frac{1}{2\pi M} \sum_{j=1}^M A(u_j) + \mathcal{O} \left(\frac{\log^2 N}{N} \right).$$

Now, by Lemma A.2.1 we can write

$$A(u) = 2\pi \int_{-\pi}^{\pi} \phi(u, \omega) f(u, \omega) d\omega$$

and then

$$\frac{1}{2\pi M} \sum_{j=1}^M A(u_j) = J(\phi) + \mathcal{O} \left(\frac{1}{M} \right). \quad (\text{A.6})$$

On the other hand, the second term in (A.4) can be bounded as follows:

$$\begin{aligned} & |c(u_j, t, t+k) \widehat{\phi}(u_j, k) (2 - \delta_k)| \\ & \leq 2 \left\{ \widehat{f}(u_j, k) \widehat{\phi}(u_j, k) + \frac{N}{T} |\widehat{f}(u_j, k) \widehat{\phi}(u_j, k) \varphi_1(u_j, k)| \right. \\ & \quad + \frac{N}{T} |\widehat{f}(u_j, k) \widehat{\phi}(u_j, k) \varphi_2(u_j, k)| \\ & \quad \left. + \left(\frac{N}{T} \right)^2 \log^2 N |\widehat{f}(u_j, k) \widehat{\phi}(u_j, k)| \right\}. \end{aligned}$$

Since $|\varphi_i(u_j, k)| \leq K \log k$ for $i = 1, 2$, $j = 1, \dots, M$ and $k > 1$, we conclude that

$$|c(u_j, t, t+k) \widehat{\phi}(u_j, k)(2 - \delta_k)| \leq \frac{N \log k}{T k^2}.$$

Therefore, since $|h'(u)| \leq K$ for $u \in [0, 1]$ we have

$$\left| \sum_{k=0}^{N-t} c(u_j, t, t+k) \widehat{\phi}(u_j, k) \frac{k}{N} h'(\xi_{t,k,N}) \right| \leq \frac{K}{T} \sum_{k=1}^N \frac{\log k}{k} \leq K \frac{\log^2 N}{T}.$$

Consequently,

$$\left| \sum_{t=1}^{N-1} h\left(\frac{t}{N}\right) \sum_{k=0}^{N-t} c(u_j, t, t+k) \widehat{\phi}(u_j, k)(2 - \delta_k) \frac{k}{N} h'(\xi_{t,k,N}) \right| \leq KN \frac{\log^2 N}{T}.$$

Hence, the second term of (A.4) is bounded by $K(\log^2 N)/T$. From this and (A.6), the required result is obtained.

Proposición A.1.2 *Let $f(u, \lambda)$ be a time-varying spectral density satisfying assumption A1. Let $\phi, \varphi : [0, 1] \times [-\pi, \pi] \rightarrow \mathbb{R}$ be two functions such that $\phi(u, \lambda)$ and $\varphi(u, \lambda)$ are symmetric in λ , twice differentiable with respect to u and their Fourier coefficients satisfy $|\widehat{\phi}(u, k)|, |\widehat{\varphi}(u, k)| \leq K|k|^{-2\alpha(\theta, u)-1}$ for $u \in [0, 1]$ and $k > 1$. If assumptions A2 and A3 hold, then*

$$\lim_{T \rightarrow \infty} T \text{Cov}[J_T(\phi), J_T(\varphi)] = 4\pi \int_0^1 \int_{-\pi}^{\pi} \phi(u, \lambda) \varphi(u, \lambda) f(u, \lambda)^2 d\lambda du.$$

Demostración A.1.2 *We can write*

$$\begin{aligned} & T \text{Cov}[J_T(\phi), J_T(\varphi)] \\ &= \frac{T}{M^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi(u_j, \lambda) \varphi(u_k, \mu) \text{Cov}[I_N(u_j, \lambda), I_N(u_k, \mu)] d\lambda d\mu. \end{aligned}$$

But,

$$\begin{aligned}
& \text{Cov}[I_N(u_j, \lambda), I_N(u_k, \mu)] \\
&= \frac{1}{[2\pi H_{2,N}(0)]^2} \text{Cov}(|D_N(u_j, \lambda)|^2, |D_N(u_k, \mu)|^2) \\
&= \frac{1}{[2\pi H_{2,N}(0)]^2} \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t)+i\mu(m-p)} \\
&\quad \times \text{Cov}\left(X_{[u_j T] - \frac{N}{2} + s + 1, T}, X_{[u_j T] - \frac{N}{2} + t + 1, T}, X_{[u_k T] - \frac{N}{2} + p + 1, T}, X_{[u_k T] - \frac{N}{2} + m + 1, T}\right).
\end{aligned}$$

Now, an application of Theorem 2.3.2 of Brillinger (1981) yields

$$\begin{aligned}
& \text{Cov}[I_N(u_j, \lambda), I_N(u_k, \mu)] \\
&= \frac{1}{[2\pi H_{2,N}(0)]^2} \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t)+i\mu(m-p)} \\
&\quad \times \left\{ \text{Cov}\left(X_{[u_j T] - \frac{N}{2} + t + 1, T}, X_{[u_k T] - \frac{N}{2} + m + 1, T}\right) \right. \\
&\quad \times \text{Cov}\left(X_{[u_j T] - \frac{N}{2} + s + 1, T}, X_{[u_k T] - \frac{N}{2} + p + 1, T}\right) \\
&\quad \left. + \text{Cov}\left(X_{[u_j T] - \frac{N}{2} + t + 1, T}, X_{[u_k T] - \frac{N}{2} + p + 1, T}\right) \right. \\
&\quad \left. \times \text{Cov}\left(X_{[u_j T] - \frac{N}{2} + s + 1, T}, X_{[u_k T] - \frac{N}{2} + m + 1, T}\right) \right\} \\
&= \frac{1}{[2\pi H_{2,N}(0)]^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} H_N\left(A_{t_j - \frac{N}{2} + 1 + \cdot, T}^0(x) h\left(\frac{\cdot}{N}\right), \lambda - x\right) \\
&\quad \times H_N\left(A_{t_k - \frac{N}{2} + 1 + \cdot, T}^0(x) h\left(\frac{\cdot}{N}\right), x - \mu\right)
\end{aligned}$$

$$\begin{aligned}
& \times H_N \left(A_{t_j - \frac{N}{2} + 1+, T}^0(y) h\left(\frac{\cdot}{N}\right), -y - \lambda \right) \\
& \times H_N \left(A_{t_k - \frac{N}{2} + 1+, T}^0(y) h\left(\frac{\cdot}{N}\right), y + \mu \right) e^{i\lambda(s-t) + i\mu(m-p)} dx dy \\
& + \frac{1}{[2\pi H_{2,N}(0)]^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} H_N \left(A_{t_j - \frac{N}{2} + 1+, T}^0(x) h\left(\frac{\cdot}{N}\right), \lambda - x \right) \\
& \times H_N \left(A_{t_k - \frac{N}{2} + 1+, T}^0(x) h\left(\frac{\cdot}{N}\right), x + \mu \right) \\
& \times H_N \left(A_{t_j - \frac{N}{2} + 1+, T}^0(y) h\left(\frac{\cdot}{N}\right), -y - \lambda \right) \\
& \times H_N \left(A_{t_k - \frac{N}{2} + 1+, T}^0(y) h\left(\frac{\cdot}{N}\right), y - \mu \right) e^{i\lambda(s-t) + i\mu(m-p)} dx dy.
\end{aligned}$$

Thus,

$$TCov(J_T(\phi), J_T(\varphi)) = \frac{T}{[2\pi M H_{2,N}(0)]^2} [B_N^{(1)} + B_N^{(2)}], \quad (\text{A.7})$$

where

$$\begin{aligned}
B_N^{(1)} &= \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi(u_j, \lambda) \varphi(u_k, \mu) \\
& \times H_N \left(A_{t_j - \frac{N}{2} + 1+, T}^0(x) h\left(\frac{\cdot}{N}\right), \lambda - x \right) \\
& \times H_N \left(\overline{A_{t_k - \frac{N}{2} + 1+, T}^0(x)} h\left(\frac{\cdot}{N}\right), x - \mu \right) \\
& \times H_N \left(A_{t_j - \frac{N}{2} + 1+, T}^0(y) h\left(\frac{\cdot}{N}\right), -y - \lambda \right) \\
& \times H_N \left(\overline{A_{t_k - \frac{N}{2} + 1+, T}^0(y)} h\left(\frac{\cdot}{N}\right), y + \mu \right) e^{i(x+y)(t_j - t_k)} dx dy d\mu d\lambda,
\end{aligned}$$

and,

$$\begin{aligned}
B_N^{(2)} &= \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi(u_j, \lambda) \varphi(u_k, \mu) \\
&\quad \times H_N \left(A_{t_j - \frac{N}{2} + 1 +, T}^0(x) h\left(\frac{\cdot}{N}\right), \lambda - x \right) \\
&\quad \times H_N \left(\overline{A_{t_k - \frac{N}{2} + 1 +, T}^0(x)} h\left(\frac{\cdot}{N}\right), x + \mu \right) \\
&\quad \times H_N \left(A_{t_j - \frac{N}{2} + 1 +, T}^0(y) h\left(\frac{\cdot}{N}\right), -y - \lambda \right) \\
&\quad \times H_N \left(\overline{A_{t_k - \frac{N}{2} + 1 +, T}^0(y)} h\left(\frac{\cdot}{N}\right), y - \mu \right) e^{i(x+y)(t_j-t_k)} dx dy d\mu d\lambda.
\end{aligned}$$

The term $B_N^{(1)}$ can be written as follows:

$$\begin{aligned}
B_N^{(1)} &= \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi(u_j, \lambda) \varphi(u_k, \mu) \\
&\quad \times A(u_j, x) A(u_k, -x) A(u_j, y) A(u_k, -y) \\
&\quad \times H_N(\lambda - x) H_N(x - \mu) H_N(\mu + y) H_N(-y - \lambda) \\
&\quad \times e^{i(x+y)(t_j-t_k)} dx dy d\lambda d\mu + R_N \\
&= \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi(u_j, x) A(u_j, x) A(u_k, -x) \\
&\quad \times \varphi(u_k, y) A(u_j, y) A(u_k, -y) \\
&\quad \times H_N(\lambda - x) H_N(x - \mu) H_N(\mu + y) H_N(-y - \lambda) \\
&\quad \times e^{i(x+y)(t_j-t_k)} dx dy d\lambda d\mu + \Delta_N^{(1)} + \Delta_N^{(2)} + R_N, \tag{A.8}
\end{aligned}$$

with

$$\begin{aligned}
\Delta_N^{(1)} &= \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [\phi(u_j, \lambda) - \phi(u_j, x)] \varphi(u_k, \mu) \\
&\quad \times A(u_j, x) A(u_k, -x) A(u_j, y) A(u_k, -y) \\
&\quad \times H_N(\lambda - x) H_N(x - \mu) H_N(\mu + y) H_N(-y - \lambda) \\
&\quad \times e^{i(x+y)(t_j-t_k)} dx dy d\lambda d\mu,
\end{aligned}$$

$$\begin{aligned}
\Delta_N^{(2)} &= \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi(u_j, x) [\varphi(u_k, \mu) - \varphi(u_k, y)] \\
&\quad \times A(u_j, x) A(u_k, -x) A(u_j, y) A(u_k, -y) \\
&\quad \times H_N(\lambda - x) H_N(x - \mu) H_N(\mu + y) H_N(-y - \lambda) \\
&\quad \times e^{i(x+y)(t_j-t_k)} dx dy d\lambda d\mu,
\end{aligned}$$

and by Lemma A.2.2 the remainder term R_N can be bounded as follows,

$$\begin{aligned}
|R_N| &\leq \frac{N}{T} \left| \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi(u_j, \lambda) \varphi(u_k, \mu) A(u_j, x) A(u_k, -x) \right. \\
&\quad \times A(u_j, y) y^{-d(u_k)} L_N(y + \mu) H_N(\lambda - x) H_N(x - \mu) \\
&\quad \left. \times H_N(-y - \lambda) e^{i(x+y)(t_j-t_k)} dx dy d\lambda d\mu \right|. \tag{A.9}
\end{aligned}$$

By integrating with respect to μ the term $\Delta_N^{(1)}$ can be written as,

$$\begin{aligned}\Delta_N^{(1)} &= \sum_{j,k=1}^M \sum_{t,s=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) \widehat{\varphi}(u_k, t-s) \\ &\quad \times \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [\phi(u_j, \lambda) - \phi(u_j, x)] A(u_j, x) A(u_k, -x) \\ &\quad \times A(u_j, y) A(u_k, -y) H_N(\lambda - x) H_N(-y - \lambda) \\ &\quad \times e^{i(x+y)(t_j-t_k)-ixt-iy s} dx dy d\lambda,\end{aligned}$$

and by integrating with respect to y we get

$$\begin{aligned}\Delta_N^{(1)} &= \sum_{j,k=1}^M \sum_{t,s,p=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) \widehat{\varphi}(u_k, t-s) \\ &\quad \times \widehat{f}(u_j, u_k, t_j - t_k - s + p) \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [\phi(u_j, \lambda) - \phi(u_j, x)] \\ &\quad \times A(u_j, x) A(u_k, -x) H_N(\lambda - x) e^{ix(t_j-t_k-t)+i\lambda p} dx d\lambda \\ &= \sum_{j,k=1}^M \sum_{t,s,p=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) \widehat{\varphi}(u_k, t-s) \\ &\quad \times \widehat{f}(u_j, u_k, t_j - t_k - s + p) \varepsilon_N(u_j, u_k, p, t_j - t_k - t),\end{aligned}$$

where the terms $\widehat{f}(u, v, k)$ and $\varepsilon_N(r)$ as given by

$$\widehat{f}(u, v, k) = \int_{-\pi}^{\pi} A(u, \lambda) A(v, -\lambda) e^{i\lambda k} d\lambda,$$

and

$$\begin{aligned}
\varepsilon_N(u_j, u_k, p, r) &= \sum_{m=0}^{N-1} h\left(\frac{m}{N}\right) \widehat{\phi}(u_j, p-m) \int_{-\pi}^{\pi} A(u_j, x) A(u_k, -x) e^{ix(r+m)} dx \\
&\quad - 2\pi h\left(\frac{p}{m}\right) \int_{-\pi}^{\pi} \phi(u_j, x) A(u_j, x) A(u_k, x) e^{ix(r+p)} dx \\
&= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi(u_j, \lambda) A(u_j, x) A(u_k, -x) e^{i(p\lambda+rx)} \sum_{m=0}^{N-1} h\left(\frac{m}{N}\right) e^{im(x-\lambda)} d\lambda dx \\
&\quad - 2\pi \int_{-\pi}^{\pi} h\left(\frac{m}{N}\right) \phi(u_j, x) A(u_j, x) A(u_k, -x) e^{i(r+p)x} dx.
\end{aligned}$$

But $h\left(\frac{m}{M}\right) = h\left(\frac{p}{M}\right) + h'(\xi_{p,m}) \frac{m-p}{N}$ for some $\xi_{p,m} \in [0, 1]$. Thus,

$$\begin{aligned}
\varepsilon_N(u_j, u_k, p, r) &= h\left(\frac{p}{N}\right) \left\{ \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi(u_j, \lambda) A(u_j, x) A(u_k, -x) e^{i(p\lambda+rx)} \right. \\
&\quad \times \sum_{m=0}^{N-1} e^{im(x-\lambda)} d\lambda dx - 2\pi \int_{-\pi}^{\pi} \phi(u_j, x) A(u_j, x) A(u_k, -x) e^{i(r+p)x} dx \Big\} \\
&\quad + \sum_{m=0}^{N-1} h'(\xi_{p,m}) \frac{m-p}{N} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi(u_j, \lambda) A(u_j, x) A(u_k, -x) \\
&\quad \times e^{im(x-\lambda)+i(p\lambda+rx)} d\lambda dx \\
&= h\left(\frac{p}{N}\right) \varepsilon_N^{(1)}(u_j, u_k, p, r) + \varepsilon_N^{(2)}(u_j, u_k, p, r),
\end{aligned}$$

where the term $\varepsilon_N^{(1)}(u_j, u_k, p, r)$ is given by

$$\varepsilon_N^{(1)}(u_j, u_k, p, r) = \int_{-\pi}^{\pi} g(\omega) e^{ir\omega} \sum_{m=0}^{N-1} e^{im\omega} d\omega - 2\pi g(0),$$

with

$$g(\omega) = \int_{-\pi}^{\pi} \phi(u_j, \lambda) A(u_j, \omega + \lambda) A(u_k, -\omega - \lambda) e^{i(p+r)\lambda} d\lambda.$$

Observe that by Lemma A.2.1, for every u_j, u_k, p, r , $\varepsilon_N^{(1)}(u_j, u_k, p, r) \rightarrow 0$ as $N \rightarrow \infty$, consequently we can write

$$\varepsilon_N^{(1)}(u_j, u_k, p, r) = \int_{-\pi}^{\pi} g(\omega) e^{i r \omega} \sum_{m=N}^{\infty} e^{i m \omega} d\omega = \sum_{m=N}^{\infty} \hat{\phi}(u_j, p-m) \hat{f}(u_j, u_k, r+m).$$

On the other hand, by assumption A1, $|\hat{f}(u_j, u_k, r+m)| \leq K|r+m|^{d(u_j)+d(u_k)-1}$. Thus, the term $\varepsilon_N^{(2)}(u_j, u_k, p, r)$ is bounded by

$$\begin{aligned} \left| \varepsilon_N^{(2)}(u_j, u_k, p, r) \right| &\leq \frac{K}{N} \sum_{m=0}^{N-1} |m-p| \hat{\phi}(u_j, m-p) \hat{f}(u_j, u_k, r+m) \\ &\leq \frac{K}{N} \sum_{m=0}^{N-1} |m-p|^{-2d(u_j)} |r+m|^{d(u_j)+d(u_k)-1} \\ &\leq K \left\{ \sum_{m=0}^{N-1} \left| \frac{m}{N} - \frac{p}{N} \right| \left| \frac{r}{N} - \frac{m}{N} \right|^{d(u_j)+d(u_k)-1} \frac{1}{N} \right\} N^{d(u_k)-d(u_j)-1} \\ &\leq K \int_0^1 \left| x - \frac{p}{N} \right|^{-2d(u_j)} \left| \frac{r}{N} + x \right|^{d(u_j)+d(u_k)-1} dx N^{d(u_k)-d(u_j)-1} \\ &\leq K N^{d(u_k)-d(u_j)-1} \left(\frac{r}{N} \right)^{d(u_j)+d(u_k)-1} \int_0^1 \left| x - \frac{p}{N} \right|^{-2d(u_j)} dx \\ &\leq K N^{-2d(u_j)} r^{d(u_j)+d(u_k)-1}. \end{aligned}$$

Thus,

$$\begin{aligned} \varepsilon_N(u_j, u_k, p, r) &= h\left(\frac{p}{N}\right) \sum_{m=N}^{\infty} \hat{\phi}(u_j, p-m) \hat{f}(u_j, u_k, r+m) \\ &\quad + \mathcal{O}\left(N^{-2d(u_j)} r^{d(u_j)+d(u_k)-1}\right). \end{aligned}$$

Hence $\Delta_N^{(1)}$ can be written as

$$\begin{aligned} \Delta_N^{(1)} &= \sum_{j,k=1}^M \sum_{t,s,p=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) \widehat{\varphi}(u_k, t-s) \widehat{f}(u_j, u_k, t_j - t_k - s + p) \\ &\quad \times \left\{ h\left(\frac{p}{N}\right) \sum_{m=N+1}^{\infty} \widehat{\phi}(u_j, p-m) \widehat{f}(u_j, u_k, t_j - t_k - t + m) \right. \\ &\quad \left. + O\left(N^{-2d(u_j)} |t_j - t_k - t|^{d(u_j)+d(u_k)-1}\right) \right\} \\ &\equiv \Delta_N^{(1.1)} + \Delta_N^{(1.2)}, \end{aligned}$$

say. Therefore, $\left|\Delta_N^{(1)}\right| \leq \left|\Delta_N^{(1.1)}\right| + \left|\Delta_N^{(1.2)}\right|$.

Observe that $\varphi(u, \lambda) \sim C|\lambda|^{2d(u)}$ as $\lambda \rightarrow 0$. Thus, $\varphi(u, 0) = \sum_{k=-\infty}^{\infty} \widehat{\varphi(u, k)} = 0$.

aaa

Note that the term $\Delta_N^{(1.1)}$ can be bounded as follows:

$$\begin{aligned} \left|\Delta_N^{(1.1)}\right| &\leq \sum_{j,h=1}^M \sum_{t,s,p=0}^{N-1} |\widehat{\varphi}(u_k, t-s)| \left|\widehat{f}(u_j, u_k, t_j - t_k - s + p)\right| \\ &\quad \times \sum_{m=N+1}^{\infty} \left|\widehat{\phi}(u_j, p-m)\right| \left|\widehat{f}(u_j, u_k, t_j - t_k - t + m)\right|. \end{aligned}$$

Since,

$$\sum_{t,s=1}^N |\widehat{\varphi}(u_k, t-s)| \leq K N^{1-2d(u_k)},$$

we conclude that,

$$\begin{aligned}
\left| \Delta_N^{(1.1)} \right| &\leq K \sum_{j,k=1, j \neq k}^M \sum_{p=0}^{N-1} N^{1-2d(u_k)} |S(j-k) + p|^{d(u_j)+d(u_k)-1} \\
&\quad \times \sum_{m=N+1}^{\infty} |p-m|^{-2d(u_j)-1} |S(j-k) + m|^{d(u_j)+d(u_k)-1} \\
&\leq K \sum_{j,k=1, j \neq k}^M \sum_{p=0}^{N-1} \left| \frac{S}{N}(j-k) + \frac{p}{N} \right|^{d(u_j)+d(u_k)-1} \frac{1}{N} \sum_{m=N+1}^{\infty} \left| \frac{p}{N} - \frac{m}{N} \right|^{-2d(u_j)-1} \\
&\quad \times \left| \frac{S}{N}(j-k) + \frac{m}{N} \right|^{d(u_j)+d(u_k)-1} \frac{1}{N} \\
&\leq K \sum_{j,k=1, j \neq k}^M \int_0^1 \int_1^{\infty} \left| \frac{S}{N}(j-k) + x \right|^{d(u_j)+d(u_k)-1} |x-y|^{-2d(u_j)-1} \\
&\quad \times \left| \frac{S}{N}(j-k) + y \right|^{d(u_j)+d(u_k)-1} dy dx.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\left| \Delta_N^{(1.1)} \right| &\leq K \sum_{j,k=1, j \neq k}^M \left| \frac{S}{N}(j-k) \right|^{d(u_j)+d(u_k)-1} \\
&\quad \times \int_0^1 \int_1^{\infty} |x-y|^{-2d(u_j)-1} \left| \frac{S}{N}(j-k) + y \right|^{d(u_j)+d(u_k)-1} dy dx.
\end{aligned}$$

Since the integrands in the above expression are all positive, an application of Tonelli's

theorem yields

$$\begin{aligned}
\left| \Delta_N^{(1,1)} \right| &\leq K \sum_{j,k=1, j \neq k}^M \left| \frac{S}{N}(j-k) \right|^{d(u_j)+d(u_k)-1} \\
&\quad \times \int_1^\infty \int_0^1 |x-y|^{-2d(u_j)-1} \left| \frac{S}{N}(j-k) + y \right|^{d(u_j)+d(u_k)-1} dx dy \\
&\leq K \sum_{j,k=1, j \neq k}^M \left| \frac{S}{N}(j-k) \right|^{d(u_j)+d(u_k)-1} \\
&\quad \times \int_1^\infty [(y-1)^{-2d(u_j)} - y^{-2d(u_j)}] \left| \frac{S}{N}(j-k) + y \right|^{d(u_j)+d(u_k)-1} dy
\end{aligned}$$

Then, by Lemma A.2.3 we conclude that

$$\begin{aligned}
\left| \Delta_N^{(1,1)} \right| &\leq K \sum_{j,k=1, j \neq k}^M \left| \frac{S}{N}(j-k) \right|^{2d(u_j)+2d(u_k)-2} \\
&\leq K \left[\sum_{\substack{j,k=1, j \neq k \\ d(u_j)+d(u_k) \leq \frac{1}{2}}}^M \left| \frac{S}{N}(j-k) \right|^{2d(u_j)+2d(u_k)-2} + \sum_{\substack{j,k=1, j \neq k \\ d(u_j)+d(u_k) > \frac{1}{2}}}^M \left| \frac{S}{N}(j-k) \right|^{2d(u_j)+2d(u_k)-2} \right].
\end{aligned}$$

For the first summand above we have the upper bound

$$\sum_{j,k=1, j \neq k}^M |j-k|^{-1} \left(\frac{N}{S} \right)^2 \leq K \left(\frac{N}{S} \right)^2 M \log M,$$

while the second summand can be bounded as follows:

$$\begin{aligned}
&\sum_{\substack{j,k=1, j \neq k \\ d(u_j)+d(u_k) > \frac{1}{2}}}^M \left| \frac{SM}{N} \left(\frac{j}{M} - \frac{k}{M} \right) \right|^{2d(u_j)+2d(u_k)-2} \\
&\leq K \left(\frac{T}{N} \right)^{-\varepsilon} M^2 \sum_{j,k=1}^M \left| \frac{j}{M} - \frac{k}{M} \right|^{\varepsilon-1} \frac{1}{M^2} \\
&\leq K \left(\frac{T}{N} \right)^{-\varepsilon} M^2 \int_0^1 \int_0^1 |x-y|^{\varepsilon-1} dx dy \leq K \left(\frac{T}{N} \right)^{-\varepsilon} M^2 \leq KM^2.
\end{aligned}$$

Thus,

$$\left| \Delta_N^{(1,1)} \right| \leq K \left(\frac{N}{S} \right)^2 M \log M + M^2. \quad (\text{A.10})$$

A similar bound can be found for $\left| \Delta_N^{(1,2)} \right|$ and consequently for $\left| \Delta_N^{(1)} \right|$. Furthermore, an analogous argument yields a similar bound for the term $\left| \Delta_N^{(2)} \right|$ appearing in (A.8). Now, we focus on obtaining an upper bound for the remaining R_N from (A.11). By integrating that expression with respect to λ we get

$$\begin{aligned} |R_N| \leq \frac{N}{T} & \left| \sum_{j,k=1}^M \sum_{t,s=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) \widehat{\phi}(u_j, s-t) \right. \\ & \times \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \varphi(u_k, \mu) A(u_j, x) A(u_k, -x) A(u_j, y) y^{-d(u_k)} \\ & \left. \times L_N(y + \mu) H(x - \mu) e^{ix(t_j - t_k + t) + iy(t_j - t_k + s)} dx dy d\mu \right|, \end{aligned}$$

where the function $L_N(\cdot)$ is defined as

$$L_N(x) = \begin{cases} N & |x| \leq 1/N, \\ 1/|x| & 1/N < |x| \leq \pi. \end{cases}$$

Hence,

$$\begin{aligned} |R_N| \leq \frac{N}{T} & \left| \sum_{j,k=1}^M \sum_{t,s,p=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) \widehat{\phi}(u_j, s-t) \widehat{f}(u_j, u_k, t_j - t_k + t - p) \right. \\ & \left. \times \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \varphi(u_k, \mu) A(u_j, y) y^{-d(u_k)} L_N(\mu + y) e^{iy(t_j - t_k + s) + ip\mu} dy d\mu \right| \end{aligned}$$

$$\begin{aligned}
&\leq K \frac{N}{T} \sum_{j,k=1}^M \sum_{t,s,p=0}^{N-1} \left| \widehat{\phi}(u_j, s-t) \right| \left| \widehat{f}(u_j, u_k, t_j - t_k + t - p) \right| \\
&\quad \times \left| \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \varphi(u_k, \mu) A(u_j, y) y^{-d(u_k)} L_N(\mu + y) e^{iy(t_j - t_k + s) + ip\mu} dy d\mu \right| \\
&\leq K \frac{N}{T} \sum_{j,k=1}^M \sum_{t,s,p=0}^{N-1} \left| \widehat{\phi}(u_j, s-t) \right| \left| \widehat{f}(u_j, u_k, t_j - t_k + t - p) \right| \\
&\quad \times \left| \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} L_N(\mu + y) y^{-d(u_j) - d(u_k)} dy d\mu \right| \\
&\leq K \frac{N \log N}{T} \sum_{j,k=1}^M \sum_{t,s,p=0}^{N-1} |s-t|^{-2d(u_j)-1} |S(j-k) + t-p|^{d(u_j)+d(u_k)-1} \\
&\leq K \frac{N^2 \log N}{T} \sum_{j,k=1}^M \sum_{t,s,p=0}^{N-1} |s-t|^{-2d(u_j)-1} S^{d(u_j)+d(u_k)-1} |j-k|^{d(u_j)+d(u_k)-1} \\
&\leq K \frac{N^3 \log N}{T} M^2 \sum (SM)^{d(u_j)+d(u_k)-1} \left| \frac{j}{M} - \frac{k}{M} \right|^{d(u_j)+d(u_k)-1} \frac{1}{M^2}.
\end{aligned}$$

Since by assumption A3, $T/N^2 \rightarrow 0$,

$$|R_N| \leq K \frac{N^3 M^2}{T^{2-d}} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |x-y|^{2d-1} dx dy \leq K N^3 M^2 T^{d-2}. \quad (\text{A.11})$$

Thus, from (A.10) and (A.11) we conclude

$$\begin{aligned}
\frac{T}{[2\pi M H_{2,N}(0)]^2} B_N^{(1)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi(u_j, x) \\
&\quad \times A(u_j, x) A(u_k, -x) \varphi(u_k, y) A(u_j, y) A(u_k, -y) \\
&\quad \times H_N(\lambda - x) H_N(x - \mu) H_N(\mu + y) H_N(-y - \lambda) \\
&\quad \times e^{i(x+y)(t_j - t_k)} dx dy d\lambda d\mu + C_N,
\end{aligned}$$

where

$$C_N = \mathcal{O} \left(\frac{\log M}{S} + \frac{T}{N^2} + N T^{d-1} \right).$$

Therefore, by assumption A3 we conclude that $C_N = o(1)$. By following successive decompositions as in (A.8) we replace $\varphi(u_k, y)$ by $\varphi(u_k, x)$, $A(u_k, -y)$ by $A(u_k, -x)$ and $A(u_j, y)$ by $A(u_j, x)$, respectively. Thus,

$$\begin{aligned} \frac{T}{[2\pi M H_{2,N}(0)]^2} B_N^{(1)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi(u_j, x) A(u_j, x) \\ &\quad \times A(u_k, -x) \varphi(u_k, x) A(u_j, x) A(u_k, -x) H_N(\lambda - x) \\ &\quad \times H_N(x - \mu) H_N(\mu + y) H_N(-y - \lambda) \\ &\quad \times e^{i(x+y)(t_j - t_k)} dx dy d\lambda d\mu + o(1). \end{aligned}$$

By integrating with respect to μ and λ we get

$$\begin{aligned} &\frac{T}{[2\pi M H_{2,N}(0)]^2} B_N^{(1)} \\ &= \frac{T}{[M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi(u_j, x) A(u_j, x) A(u_k, -x) \\ &\quad \times \varphi(u_k, x) A(u_j, x) A(u_k, -x) |H_{2,N}(x + y)|^2 \\ &\quad \times e^{i(x+y)(t_j - t_k)} dx dy + o(1) \\ &= \frac{T}{[M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [\phi(u_j, x) f(u_j, x)] [\varphi(u_k, x) f(u_k, x)] \\ &\quad \times |H_{2,N}(x + y)|^2 e^{i(x+y)[s(j-k)]} dx dy + o(1) \end{aligned}$$

$$\begin{aligned}
&= \frac{T}{[M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [\phi(u_j, x) f(u_j, x)] [\varphi(u_k, x) f(u_k, x)] \\
&\quad \times |H_{2N}(z)|^2 e^{iz[s(j-k)]} dx dz + o(1) \\
&= \frac{T}{[M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} [\phi(u_j, x) f(u_j, x)] [\varphi(u_k, x) f(u_k, x)] \\
&\quad \times \sum_{s,t=0}^{N-1} h^2\left(\frac{t}{N}\right) h^2\left(\frac{s}{N}\right) \int_{-\pi}^{\pi} e^{iz[s(j-k)+t-s]} dx dz + o(1) \\
&= \frac{2\pi T}{[M H_{2,N}(0)]^2} \sum_{s,t=0}^{N-1} \sum_{\substack{j,k=1 \\ S(j-k)=s-t}}^M h^2\left(\frac{t}{N}\right) h^2\left(\frac{s}{N}\right) \int_{-\pi}^{\pi} [\phi(u_j, x) f(u_j, x)] \\
&\quad \times [\varphi(u_k, x) f(u_k, x)] dx + o(1).
\end{aligned}$$

By assumption A3, for $S < N$ we can write

$$\begin{aligned}
\frac{T}{[2\pi M H_{2,N}(0)]^2} B_N^{(1)} &= \frac{2\pi T}{[M H_{2,N}(0)]^2} \sum_{t=0}^{N-1} \sum_{p=-\frac{t}{S}}^{\frac{N-t}{S}} h^2\left(\frac{t}{N}\right) h^2\left(\frac{t}{N} + \frac{pS}{N}\right) \\
&\quad \times \sum_{j=1}^{M-|p|} \int_{-\pi}^{\pi} [\phi(u_j, x) f(u_j, x)] [\phi(u_{j+p}, x) f(u_{j+p}, x)] dx + o(1).
\end{aligned}$$

Observe that by the assumptions of this proposition the products $\phi(u, x) f(u, x)$ and $\varphi(u, x) f(u, x)$ are differentiable with respect to u . Furthermore, note that by assumption A3, $\lim_{T,S \rightarrow \infty} \frac{S|p|}{T} = 0$ for any $|p| \leq \frac{N}{S}$. Consequently,

$$\begin{aligned}
\frac{S}{T} \sum_{j=1}^{M-|p|} \int [\phi(u_j, x) f(u_j, x)] [\varphi(u_{j+p}, x) f(u_{j+p}, x)] dx \\
\rightarrow \int_0^1 \int_{-\pi}^{\pi} \phi(u, x) \varphi(u, x) f(u, x)^2 dx,
\end{aligned}$$

for any $|p| < \frac{N}{S}$ as $M, N, S, T \rightarrow \infty$. On the other hand,

$$\begin{aligned} & \frac{2\pi T^2 N^2}{S^2 [M H_{2,N}(0)]^2} \sum_{t=0}^{N-1} \sum_{p=-\frac{t}{S}}^{\frac{N-t}{S}} h^2\left(\frac{t}{N}\right) h^2\left(\frac{t}{N} + \frac{pS}{N}\right) \frac{S}{N^2} \\ & \rightarrow 2\pi \int_0^1 \int_{-x}^{1-x} h^2(x) h^2(x+y) dx dy \left(\int_0^1 h^2(x) dx \right)^{-2} = 2\pi, \end{aligned}$$

as $M, N, S, T \rightarrow \infty$. Therefore, in this case

$$\frac{T}{[2\pi M H_{2,N}(0)]^2} B_N^{(1)} \rightarrow 2\pi \int_0^1 \int_{-\pi}^{\pi} \phi(u, x) \varphi(u, x) f(u, x)^2 dx du,$$

as $M, N, S, T \rightarrow \infty$. Similarly, we have that

$$\frac{T}{[2\pi M H_{2,N}(0)]^2} B_N^{(2)} \rightarrow 2\pi \int_0^1 \int_{-\pi}^{\pi} \phi(u, x) \varphi(u, x) f(u, x)^2 dx du,$$

as $M, N, S, T \rightarrow \infty$. Therefore, by virtue of (A.7) this proposition is proved.

Proposición A.1.3 Let $\text{cum}_p(\cdot)$ be the p -th order cumulant with $p \geq 3$. Then

$$T^{p/2} \text{cum}_p(J_T(\phi)) \rightarrow 0,$$

as $T \rightarrow \infty$.

Demostración A.1.3 Observe that $J_T(\phi)$ can be written as

$$J_T(\phi) = \frac{1}{2\pi M H_{2,N}(0)} Y' Q(\phi) Y,$$

where the block-diagonal matrix $Q(\phi)$ is defined in (A.3) and $Y \in \mathbb{R}^{NM}$ is a Gaussian random vector defined by $Y = (Y(u_1)', \dots, Y(u_M'))'$, $Y(u) = (Y_1(u), \dots, Y_N(u))$,

$Y_t(u) = h\left(\frac{t}{N}\right) Y_{[uT] - \frac{N}{2} + t + 1, T}$ with $Y_{[uT] - \frac{N}{2} + t + 1, T}$ satisfying (??). For simplicity, denote the matrix $Q(\phi)$ as Q . Since Y is Gaussian,

$$\text{cum}_p(J_T(\phi)) = \frac{2^{p-1}(p-1)!}{(2\pi M H_{2,N}(0))^p} \text{tr}(RQ)^p$$

where $R = \text{Var}(Y)$. Let $|A| = [\text{tr}(AA')]^{1/2}$ be the Euclidean norm of matrix A and let $\|A\| = \sup_{\|x\|=1} (Ax)'Ax$ be the spectral norm of A . Now, since $|\text{tr}(QB)| \leq |Q||B|$ and $|QB| \leq \|Q\||B|$ we get $|\text{tr}(RQ)^p| \leq \|RQ\|^{p-2} |RQ|^2$.

On the other hand, for fixed λ , decompose the function $\phi(\cdot, \lambda)$ as $\phi(\cdot, \lambda) = \phi_+(\cdot, \lambda) - \phi_-(\cdot, \lambda)$ where $\phi_+(\cdot, \lambda), \phi_-(\cdot, \lambda) \geq 0$. Thus, we can write $Q = Q(\phi) = Q(\phi_+ - \phi_-) = Q(\phi_+) - Q(\phi_-) = Q_+ - Q_-$, say. Now, by Lemma A.2.6 we conclude that

$$\|RQ\| = \|RQ_+ - RQ_-\| \leq \|RQ_+\| + \|RQ_-\| \leq K(MN^{1-2d}T^{2d-1}),$$

and by Proposition A.1.2 we have that $|RQ|^2 \leq K \frac{M^2 N^2}{T}$. Thus,

$$|\text{tr}(RQ)^p| \leq K(MN^{1-2d}T^{2d-1}) \frac{M^2 N^2}{T}.$$

Consequently,

$$|T^{p/2} \text{cum}_p(J_T(\phi))| \leq K M^{1-\frac{p}{2}} \left(\frac{N}{T}\right)^{(1-2d)(p-2)} \left(\frac{\sqrt{T}}{N}\right)^{p-2}.$$

Since $p \geq 3$ and by assumption A2, $N/T \rightarrow 0$ and $\sqrt{T}/N \rightarrow 0$ as $T, N \rightarrow \infty$, the required result is obtained.

A.2. Lemas

Lemma A.2.1 *Let $f(u, \lambda)$ be a time-varying spectral density satisfying assumption A1 and let $\phi : [0, 1] \times [-\pi, \pi] \rightarrow \mathbb{R}$ be a function such that $\phi(u, \lambda)$ is continuously differentiable in λ . Consider the function defined by*

$$g(u, \lambda) = \int_{-\pi}^{\pi} \phi(u, \lambda + \omega) f(u, \omega) d\omega,$$

and its Fourier coefficients

$$\widehat{g}(u, k) = \int_{-\pi}^{\pi} g(u, \lambda) e^{-ik\lambda} d\lambda.$$

Under assumptions A1, for every $u \in [0, 1]$ we have that

$$\lim_{n \rightarrow \infty} \sum_{k=-n}^n \widehat{g}(u, k) = 2\pi g(u, 0).$$

Demostración A.2.1 *By Fubini's theorem and finite sum we have that*

$$\begin{aligned} \sum_{k=-n}^n \widehat{g}(u, k) &= \int_{-\pi}^{\pi} g(u, \lambda) \sum_{k=-n}^n e^{-ik\lambda} d\lambda = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi(u, \lambda + \omega) f(u, \omega) \sum_{k=-n}^n e^{-ik\lambda} d\lambda d\omega, \\ &= \int_{-\pi}^{\pi} f(u, \omega) \sum_{k=-n}^n \left\{ \int_{-\pi}^{\pi} \phi(u, \lambda + \omega) e^{-ik(\lambda + \omega)} d\lambda \right\} e^{ik\omega} d\omega \\ &= 2\pi \int_{-\pi}^{\pi} f(u, \omega) \sum_{k=-n}^n \widehat{\phi}(u, \omega) e^{ik\omega} d\omega. \end{aligned}$$

Let $S_n(\omega) = \sum_{k=-n}^n \widehat{\phi}(u, k) e^{ik\omega}$, be the partial sum of the Fourier series of $\phi(u, \cdot)$,

where

$$\widehat{\phi}(u, k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi(u, \lambda) e^{ik\lambda} d\lambda.$$

Since $\phi(u, \cdot)$ has continuous derivative, $\phi(u, \cdot)$ is of bounded variation. Then, by the Dirichlet-Jordan criterion (Zygmund, 1959, Theorem 8.1(ii)), the Fourier series $S_n(\omega)$ converges uniformly to $\phi(u, \omega)$. Therefore, since $f(u, \cdot)$ is integrable we have

$$\sum_{k=-n}^n \widehat{g}(u, k) = 2\pi \int_{-\pi}^{\pi} f(u, \omega) S_n(\omega) d\omega \rightarrow 2\pi \int_{-\pi}^{\pi} f(u, \omega) \phi(u, \omega) d\omega = 2\pi g(u, 0),$$

as $n \rightarrow \infty$.

Lemma A.2.2 Consider the function $\phi : [0, 1] \times [-\pi, \pi] \rightarrow \mathbb{C}$, such that $\partial\phi(u, \gamma)/\partial u$ exists and

$$|\partial\phi(u, \gamma)/\partial u| \leq K|\gamma|^{-2d(u)},$$

where $0 \leq d(u) \leq d$ for all $u \in [0, 1]$. Then, for any $0 \leq t \leq N$ we have that

$$H_N \left[\phi \left(\frac{t}{T}, \gamma \right) h \left(\frac{t}{N} \right), \lambda \right] = \phi \left(\frac{t}{T}, \gamma \right) H_N(\lambda) + \mathcal{O} \left[\frac{N}{T} |\gamma|^{-2d} L_N(\lambda) \right].$$

Demostración A.2.2 Summation by parts yields

$$\begin{aligned} H_N \left[\phi \left(\frac{t}{T}, \gamma \right) h \left(\frac{t}{N} \right), \lambda \right] - \phi \left(\frac{t}{T}, \gamma \right) H_N(\lambda) &= \left[\phi \left(\frac{N-1}{T}, \gamma \right) - \phi \left(\frac{t}{T}, \gamma \right) \right] H_N \left[h \left(\frac{t}{N} \right), \lambda \right] \\ &\quad - \sum_{s=1}^{N-1} \left[\phi \left(\frac{s}{T}, \gamma \right) - \phi \left(\frac{s-1}{T}, \gamma \right) \right] H_N \left[h \left(\frac{s}{N} \right), \lambda \right]. \end{aligned}$$

But, by the mean value theorem

$$\left| \phi \left(\frac{s}{T}, \gamma \right) - \phi \left(\frac{s-1}{T}, \gamma \right) \right| \leq \sup_{0 \leq u \leq \frac{N}{T}} \left| \frac{\partial}{\partial u} \phi(u, \gamma) \right| \frac{1}{T}.$$

But, by assumption we have

$$\sup_{u \leq \frac{N}{T}} \left| \frac{\partial}{\partial u} \phi(u, \gamma) \right| \leq K \sup_{0 \leq u \leq 1} |\gamma|^{-2d(u)} \leq K |\gamma|^{-2d}.$$

Thus,

$$\left| \phi\left(\frac{s}{T}, \gamma\right) - \phi\left(\frac{s-1}{T}, \gamma\right) \right| \leq K \frac{|\gamma|^{-2d}}{T}.$$

Furthermore, by (Dahlhaus, 1988, Lemma 5.4) $|H_s(\lambda)| \leq K L_N(\lambda)$. Consequently,

$$\left| H_N \left[\phi\left(\frac{\cdot}{T}, \gamma\right) h\left(\frac{\cdot}{N}\right), \lambda \right] - \phi\left(\frac{t}{N}, \gamma\right) H_N(\lambda) \right| \leq K |\gamma|^{-2d} L_N(\lambda),$$

as required.

Lemma A.2.3 Consider $d_1, d_2 \in [0, 1/2]$ and for any $\ell \in \mathbb{Z}$ define the integral $I(\ell)$

as

$$I(\ell) = \int_1^\infty \left[(x-1)^{-2d_1} - x^{-2d_1} \right] |\ell + x|^{d_1+d_2-1} dx.$$

Then, $I(\ell) = \mathcal{O}(|\ell|^{d_1+d_2-1})$.

Demostración A.2.3 Assume first that $\ell < 0$ and let $n = -\ell > 0$. By defining the

positive function $g(x) = (x-1)^{-2d_1} - x^{-2d_1}$ for $x > 1$, we can write the integral $I(\ell)$

as

$$I(\ell) = \int_1^n g(x) (n-x)^{d_1+d_2-1} dx + \int_n^\infty g(x) (x-n)^{d_1+d_2-1} dx,$$

where $n \geq 1$. Therefore, we can write

$$\begin{aligned} & \int_1^n g(x) (n-x)^{d_1+d_2-1} dx \\ &= \int_1^n (x-1)^{-2d_1} (n-x)^{d_1+d_2-1} dx - \int_1^n x^{-2d_1} (n-x)^{d_1+d_2-1} dx \\ &= (n-1)^{d_2-d_1} B(1-2d_1, d_1+d_2) - \int_1^n x^{-2d_1} (n-x)^{d_1+d_2-1} dx \end{aligned}$$

$$\begin{aligned}
&= (n-1)^{d_2-d_1} B(1-2d_1, d_1+d_2) - \int_0^n x^{-2d_1} (n-x)^{d_1+d_2-1} dx \\
&\quad + \int_0^1 x^{-2d_1} (n-x)^{d_1+d_2-1} dx \\
&= (n-1)^{d_2-d_1} B(1-2d_1, d_1+d_2) - n^{d_2-d_1} B(1-2d_1, d_1+d_2) \\
&\quad + \int_0^1 x^{-2d_1} (n-x)^{d_1+d_2-1} dx \\
&= [(n-1)^{d_2-d_1} - n^{d_2-d_1}] B(1-2d_1, d_1+d_2) + \int_0^1 x^{-2d_1} (n-x)^{d_1+d_2-1} dx \\
&\leq K n^{d_2-d_1-1} + K n^{d_1+d_2-1} \leq O(n^{d_1+d_2-1}),
\end{aligned}$$

where $B(\cdot, \cdot)$ is the Beta function. On the other hand,

$$\int_n^\infty g(x) (x-n)^{d_1+d_2-1} dx \leq K \int_n^\infty x^{-2d_1-1} (x-n)^{d_1+d_2-1} dx \leq K n^{d_2-d_1-1}.$$

Therefore, $I(\ell) = O((- \ell)^{d_1+d_2-1})$, $\ell < 0$. Now, for $\ell \geq 0$ we have

$$I(\ell) \leq \ell^{d_1+d_2-1} \int_1^\infty g(x) dx,$$

but $0 \leq g(x) \leq K x^{-2d_1-1}$. Therefore,

$$\int_1^\infty g(x) dx \leq K \int_1^\infty x^{-2d_1-1} dx \leq K.$$

Hence, the result holds.

Lemma A.2.4 Let $\phi(u, \lambda)$ be a positive function, symmetric in λ , such that $\phi(u, \lambda) \geq C|\lambda|^{2d(u)}$, for $\lambda \in [-\pi, \pi]$, where $d(u)$ is a positive bounded function for $u \in [0, 1]$ and

$C > 0$. Let $Q(u)$ for $u \in [0, 1]$ be the matrix defined in (A.3). Then, there exists $K > 0$ such that

$$X'Q(u)^{-1}X \leq KX'XN^{2d(u)},$$

for all vector $X \in \mathbb{R}^N$.

Demostración A.2.4 Since ϕ is positive and symmetric in λ , there exist $Q(u)^{-1/2}$ such that $Q(u)^{-1} = Q(u)^{-1/2}Q(u)^{-1/2}$. Let $Y = Q(u)^{-1/2}X$. Then

$$X'X = Y'Q(u)Y = \int_{-\pi}^{\pi} \phi(u, \lambda) \left| \sum_{j=1}^N e^{i\lambda j} Y_j \right|^2 d\lambda \geq C \int_{-\pi}^{\pi} |\lambda|^{2d(u)} \left| \sum_{j=1}^N e^{i\lambda j} Y_j \right|^2 d\lambda.$$

Thus,

$$\frac{X'X}{Y'Y} \geq 2\pi C \frac{\int_{-\pi}^{\pi} |\lambda|^{2d(u)} \left| \sum e^{i\lambda j} Y_j \right|^2 d\lambda}{\int_{-\pi}^{\pi} \left| \sum e^{i\lambda j} Y_j \right|^2 d\lambda}.$$

Define $h^*(\lambda) = \frac{\left| \sum_{j=1}^N e^{i\lambda j} Y_j \right|^2}{\int_{-\pi}^{\pi} \left| \sum_{j=1}^N e^{i\lambda j} Y_j \right|^2 d\lambda}$. Then, this is a probability function over $[-\pi, \pi]$ satisfying $\int_{-\pi}^{\pi} h^*(\lambda) d\lambda = 1$ and $h^*(\lambda) \leq \frac{N}{2\pi}$. Consequently,

$$\frac{X'X}{Y'Y} \geq C \int_{-\pi}^{\pi} |\lambda|^{2d(u)} h^*(\lambda) d\lambda \geq C \inf_{h \in \mathcal{P}} \int_{-\pi}^{\pi} |\lambda|^{2d(u)} h(\lambda) d\lambda,$$

where $\mathcal{P} = \{h : \text{probability density in } [-\pi, \pi], h \leq N/2\pi\}$. Consequently,

$$\frac{X'X}{Y'Y} \geq C \frac{N}{\pi} \int_0^{\frac{\pi}{N}} \lambda^{2d(u)} d\lambda = \frac{\pi^{2d(u)} C}{2d(u) + 1} N^{-2d(u)}.$$

Hence $Y'Y = X'Q(u)^{-1}(\phi)X \leq KX'XN^{2d(u)}$ where K is the positive constant $K = \pi^{-2d(u)}[2d(u) + 1]C^{-1}$.

Lemma A.2.5 *Let $\phi(u, \lambda)$ be a positive function, symmetric in λ , such that $\phi(u, \lambda) \geq C|\lambda|^{2d(u)}$, for $\lambda \in [-\pi, \pi]$, where $d(u)$ is a positive bounded function for $u \in [0, 1]$ and $C > 0$. Let $Q(u)$ for $u \in [0, 1]$ be the matrix defined in (A.3). Then, there exists $K > 0$ such that*

$$|X'[Q(\phi)^{-1} - Q(\varphi)]X| \leq KX'XN^{2d+\frac{1}{2}},$$

where $\varphi(u, \cdot) = \phi(u, \cdot)^{-1}/4\pi^2$, $d = \sup d(u) < \infty$ and $X \in \mathbb{R}^{NM}$.

Demostración A.2.5 *By decomposing the vector $X \in \mathbb{R}^{NM}$ as $X = (X(u_1)', \dots, X(u_M)')'$, with $X(u) = (X_1(u), \dots, X_N(u))$ and u_j are the indices defined in (A.3), we can write*

$$X'[Q(\phi)^{-1} - Q(\varphi)]X = \sum_{j=1}^M X(u_j)'[Q(u_j)^{-1} - Q(\varphi_j)]X(u_j)$$

where $\varphi_j = \phi(u_j, \cdot)^{-1}/4\pi^2$. Note that $Q(u_j)^{-1}$ exists since $\phi(u, \lambda) > 0$. Furthermore, since ϕ is symmetric in λ , $Q(U_j)^{-1}$ can be decomposed as

$$Q(u_j)^{-1} = Q(u_j)^{-1/2}Q(u_j)^{-1/2}.$$

Therefore,

$$\begin{aligned} & X'[Q(\phi)^{-1} - Q(\varphi)]X \\ &= \sum_{j=1}^M [Q(u_j)^{-1/2}X(u_j)]'[I - Q(u_j)^{1/2}Q(\varphi_j)Q(u_j)^{1/2}][Q(u_j)^{-1/2}X(u_j)]. \end{aligned}$$

Hence,

$$|X'[Q(\phi)^{-1} - Q(\psi)]X| \leq \sum_{j=1}^M [X(u_j)'Q(u_j)^{-1}X(u_j)][|I - Q(u_j)^{1/2}Q(\varphi_j)Q(u_j)^{1/2}|].$$

By Lemma A.2.4, $X(u_j)'Q(u_j)^{-1}X(u_j) \leq KX(u_j)'X(u_j)N^{2d(u_j)}$. On the other hand by Theorem 1.a of Fox and Taqqu (1987) we have that

$$|I - Q(u_j)^{1/2}Q(\varphi_j)Q(u_j)^{1/2}|^2 = \text{tr}[I - 2Q(\varphi_j)Q(u_j) + Q(\varphi_j)^2Q(u_j)^2] = \mathcal{O}(N).$$

Therefore, $|X'[Q(\phi)^{-1} - Q(\varphi)]X| \leq KN^{2d+\frac{1}{2}} \sum_{j=1}^M X'(u_j)X(u_j)$ as required.

Lemma A.2.6 Let $\phi(u, \lambda)$ be a positive function, symmetric in λ , such that $\phi(u, \lambda) \geq C|\lambda|^{2d(u)}$, for $\lambda \in [-\pi, \pi]$, where $d(u)$ is a positive bounded function for $u \in [0, 1]$ and $C > 0$. Let $Q(\phi)$ be the block-diagonal matrix defined in (A.3). Then, there exists $K > 0$ such that

$$\sup_X \left| \frac{X'RX}{X'Q(\phi)^{-1}X} \right| \leq KMN^{1-2d}T^{2d-1}$$

where $d = \sup d(u) < \infty$ and $X \in \mathbb{R}^{NM}$.

Demostración A.2.6 Assume that the vector $X \in \mathbb{R}^{MN}$ is decomposed as $X' = (X'_1, \dots, X'_M)$ where $X'_i = (X_i(1), \dots, X_i(N)) \in \mathbb{R}^N$. Thus, $X'RX = \sum_{i,j=1}^M X'_i R_{ij} X_j$.

In turn, by (??)

$$\begin{aligned} R_{ij}(r, s) &= E(Y_{t_i-N/2+1+r, T} Y_{t_j-N/2+1+s, T}) \\ &= h\left(\frac{r}{N}\right) h\left(\frac{s}{N}\right) \int_{-\pi}^{\pi} e^{2\lambda(t_i-t_j+s-r)} A_{t_i-N/2+r+1}^{\circ}(\lambda) \overline{A_{t_j-N/2+s+1}^{\circ}(-\lambda)} d\lambda, \\ &= h\left(\frac{r}{N}\right) h\left(\frac{s}{N}\right) \int_{-\pi}^{\pi} e^{2\lambda(t_i-t_j+s-r)} A(u_i, \lambda) \overline{A(u_j, -\lambda)} d\lambda + \mathcal{O}\left(\frac{1}{T}\right) \\ &\equiv h\left(\frac{r}{N}\right) h\left(\frac{s}{N}\right) \widehat{f}(u_i, u_j, t_i - t_j + r - s) + \mathcal{O}\left(\frac{1}{T}\right). \end{aligned}$$

Therefore,

$$X_i' R_{ij} X_j = \sum_{r,s=1}^N h\left(\frac{r}{N}\right) h\left(\frac{s}{N}\right) \widehat{f}(u_i, u_j, t_i - t_j + r - s) X_i(r) X_j(s) + \mathcal{O}\left(\frac{N^2}{T}\right),$$

and

$$X' R X = \sum_{i,j=1}^M \sum_{r,s=1}^N h\left(\frac{r}{N}\right) h\left(\frac{s}{N}\right) \widehat{f}(u_i, u_j, t_i - t_j + r - s) X_i(r) X_j(s) + \mathcal{O}\left(\frac{M^2 N^2}{T}\right).$$

Thus,

$$\begin{aligned} |X' R X| &\leq \sum_{i,j=1}^M \sum_{r,s=1}^N h\left(\frac{r}{N}\right) h\left(\frac{s}{N}\right) |t_i - t_j|^{d(u_i)+d(u_j)-1} \left| \omega\left(\frac{i}{M}, \frac{r}{N}\right) \omega\left(\frac{i}{M}, \frac{r}{N}\right) \right| \\ &\quad + \mathcal{O}\left(\frac{M^2 N^2}{T}\right) \\ &\leq K M^2 N^2 T^{-1} \int_0^1 \int_0^1 \int_0^1 \int_0^1 h(x) h(y) |u - v|^{d(u)+d(v)-1} T^{d(u)+d(v)} \\ &\quad \times |\omega(u, x)| |\omega(v, y)| dx dy du dv + \mathcal{O}\left(\frac{M^2 N^2}{T}\right). \end{aligned} \quad (\text{A.12})$$

where $\omega : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ is an integrable function such that $X_i(r) = \omega(\frac{i}{M}, \frac{r}{N})$ for $i = 1, \dots, M$ and $r = 1, \dots, N$. On the other hand, by Lemma A.2.5 with $\delta = 2d + 1$ we have,

$$X' A^{-1} X = X' A(\varphi) X + \mathcal{O}(X' X N^\delta),$$

and,

$$X' A(\varphi) X = \sum_{j=1}^M X_j' A(\varphi) X_j = \sum_{j=1}^M \sum_{r,s=1}^N X_j(r) X_j(s) \widehat{\varphi}(u, j, r - s),$$

where $\varphi(\lambda) = [4\pi^2\phi(\lambda)]^{-1}$. Thus,

$$\begin{aligned} X'A(\varphi)X &= K \sum_{j=1}^M \sum_{r,s=1}^N \omega\left(\frac{j}{M}, \frac{r}{N}\right) \omega\left(\frac{j}{M}, \frac{s}{N}\right) \left|\frac{r}{N} - \frac{s}{N}\right|^{2d_j-1} \frac{1}{N^2 M} (MN^{2d_j+1}) \\ &\geq KMN \int_0^1 \int_0^1 \int_0^1 |\omega(u, x)| |\omega(u, y)| |x - y|^{2d(u)-1} N^{2d(u)} dx dy du. \end{aligned} \quad (\text{A.13})$$

Now, by combining (A.12) and (A.13) we get

$$\left| \frac{X'RX}{X'A^{-1}X} \right| \leq K \frac{MN}{T} B(N, T),$$

where

$$\begin{aligned} B(N, T) &= \\ &\int_0^1 \int_0^1 \int_0^1 \int_0^1 h(x)h(y) |u - v|^{d(u)+d(v)-1} T^{d(u)+d(v)} |\omega(u, x)| |\omega(v, y)| dx dy du dv \\ &\quad \times \left\{ \int_0^1 \int_0^1 \int_0^1 |\omega(u, x)| |\omega(u, y)| |x - y|^{2d(u)-1} N^{2d(u)} dx dy du \right\}^{-1}. \end{aligned}$$

Given that $0 \leq d(u), d(v) \leq d$, for all $u, v \in [0, 1]$ we conclude

$$B(N, T) \leq K \left(\frac{T}{N} \right)^{2d},$$

proving the result.

Lemma A.2.7 Let $f(\lambda)$ and $\phi(\lambda)$ be two real-valued functions defined over $\lambda \in [-\pi, \pi]$ with Fourier coefficients $\widehat{f}(k)$ and $\widehat{\phi}(k)$, respectively, satisfying $|\widehat{f}(k)\widehat{\phi}(k)| \leq K/k^2$, for some positive constant K and $|k| > 0$. Let $C(N)$ be given by

$$C(N) = \sum_{t=1}^{N-1} h^2\left(\frac{t}{N}\right) \sum_{k=N-t}^{N-1} \widehat{f}(k)\widehat{\phi}(k)$$

with bounded data taper, $|h(u)| < K$, for all $u \in [0, 1]$. Then, there exists a positive constant K such that

$$|C(N)| \leq K \log^2 N.$$

Demostración A.2.7 Consider a real number $\alpha \in (0, 1]$ satisfying $1 - \alpha = \mathcal{O}(1/N)$.

Then, we can write

$$|C(N)| \leq K \log N \sum_{t=1}^{N-1} \sum_{k=N-t}^{N-1} \frac{1}{k^2} = K \log N \left\{ \sum_{t=1}^{\alpha N} \sum_{k=N-t}^{N-1} \frac{1}{k^2} + \sum_{t=\alpha N+1}^{N-1} \sum_{k=N-t}^{N-1} \frac{1}{k^2} \right\}.$$

For $\frac{t}{N} < \alpha < 1$ we have

$$\sum_{k=N-t}^{N-1} \frac{1}{k^2} = \frac{1}{N} \sum_{k=N-t}^{N-1} \left(\frac{N}{k} \right)^2 \frac{1}{N} \leq K \frac{1}{N} \int_{1-\frac{t}{N}}^1 \frac{dx}{x^2} \leq K \frac{1}{N} \frac{t/N}{1-t/N}.$$

Thus,

$$\begin{aligned} \sum_{t=1}^{\alpha N} \sum_{k=N-t}^{N-1} \frac{1}{k^2} &\leq \frac{K}{N} \sum_{t=1}^{\alpha N} \frac{t/N}{1-t/N} \leq K \int_0^{\alpha} \frac{x}{1-x} dx \\ &\leq K \int_0^{\alpha} \frac{dx}{1-x} = -K \log(1-\alpha) \end{aligned} \quad (\text{A.14})$$

Since $1 - \alpha = \mathcal{O}(\frac{1}{N})$, we have $-\log(1-\alpha) \sim \log N$. On the other hand, for $\alpha N + 1 \leq$

$t \leq N - 1$ we can write

$$\sum_{k=N-t}^{N-1} \frac{1}{k^2} \leq \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6},$$

Therefore,

$$\sum_{t=\alpha N+1}^{N-1} \sum_{k=N-t}^{N-1} \frac{1}{k^2} \leq \sum_{t=\alpha N+1}^{N-1} \frac{\pi^2}{6} \leq N(1-\alpha) \frac{\pi^2}{6} = \mathcal{O}(1). \quad (\text{A.15})$$

Finally, by combining inequalities (A.14) and (A.15) we obtain the result.

Lemma A.2.8 *Let $D(N, T)$ be defined by*

$$D(N, T) = \frac{1}{N} \sum_{t=1}^{N-1} \sum_{k=N-t+1}^{N-1} \frac{\varphi(k)}{k^2 - d^2} \left(\frac{t - N/2}{T} \right)$$

with $|\varphi(k)| < C \log N$ for all $0 \leq k \leq N$, $N > 1$, where C is a positive constant.

Then, there exists a constant $K > 0$ such that

$$|D(N, T)| \leq K \frac{\log^2 N}{T}.$$

Demostración A.2.8 *Note that for $k \geq 1$ we have $k^2 - d^2 \geq (k-1)^2$. Hence, for any $\alpha \in (0, 1]$ we can write*

$$\begin{aligned} |D(N, T)| &\leq C \frac{\log N}{N} \sum_{t=1}^{N-1} \sum_{k=N-t}^{N-1} \frac{1}{(k-1)^2} \left(\frac{t}{T} \right) \\ &= C \frac{\log N}{T} \sum_{t=1}^{N-1} \sum_{k=N-t}^{N-2} \frac{1}{k^2} \left(\frac{t}{N} \right) \\ &= C \frac{\log N}{T} \left\{ \sum_{t=2}^{\alpha N} \sum_{k=N-t}^{N-2} \frac{1}{k^2} \left(\frac{t}{N} \right) + \sum_{t=\alpha N+1}^{N-1} \sum_{k=N-t}^{N-2} \frac{1}{k^2} \left(\frac{t}{N} \right) \right\}. \end{aligned}$$

For $2 \leq t \leq \alpha N$, we have

$$\sum_{k=N-t}^{N-2} \frac{1}{k^2} = \frac{1}{N} \sum_{k=N-t}^{N-2} \left(\frac{N}{k} \right)^2 \frac{1}{N} \leq K \frac{1}{N} \int_{1-\frac{t}{N}}^1 x^{-2} dx \leq K \frac{1}{N} \left[\frac{(\frac{t}{N})}{1 - (\frac{t}{N})} \right],$$

where $\frac{t}{N} \leq \alpha < 1$. Hence,

$$\begin{aligned} \sum_{t=2}^{\alpha N} \sum_{k=N-t}^{N-2} \frac{1}{k^2} \left(\frac{t}{N} \right) &\leq K \frac{1}{N} \sum_{t=2}^{\alpha N} \frac{(\frac{t}{N})^2}{1 - (\frac{t}{N})^2} \\ &\leq K \int_0^\alpha \frac{x^2}{1-x} dx \leq K \int_0^\alpha \frac{1}{1-x} dx = -K \log(1-\alpha). \end{aligned}$$

Now, by choosing $\alpha = 1 - \mathcal{O}(\frac{1}{N})$ we have that $-\log(1 - \alpha) = \mathcal{O}(\log N)$. The second term can be bounded as follows. Since

$$\sum_{k=N-t}^{N-2} \frac{1}{k^2} \leq \sum_{k=1}^N \frac{1}{k^2} \leq \frac{\pi^2}{6},$$

we have

$$\sum_{t=\alpha N+1}^{N-1} \sum_{k=N-t}^{N-2} \frac{1}{k^2} \left(\frac{t}{N}\right) \leq \frac{\pi^2}{6} \sum_{t=\alpha N+1}^{N-1} \left(\frac{t}{N}\right) \leq KN \int_{\alpha}^1 x dx = \frac{K}{2} N(1 - \alpha^2).$$

But,

$$1 - \alpha^2 = 1 - \left[1 - \mathcal{O}\left(\frac{1}{N}\right)\right]^2 = \mathcal{O}\left(\frac{1}{N}\right),$$

so that

$$\left| \sum_{t=\alpha N+1}^{N-1} \sum_{k=N-t}^{N-2} \frac{1}{k^2} \frac{t}{N} \right| \leq K.$$

Consequently, $|D(N, T)| \leq K \frac{\log^2 N}{T}$, as required.

A.3. Teoremas

Demostración A.3.1 (Theorem 1). To prove the consistency of the Whittle estimator, it suffices to show that

$$\sup_{\theta} |\mathcal{L}_T(\theta) - \mathcal{L}(\theta)| \rightarrow 0,$$

in probability, as $T \rightarrow \infty$. Define $g_{\theta}(u, \lambda) = f_{\theta}(u, \lambda)^{-1}$. By assumption A1, $g_{\theta}(u, \lambda)$ is continuous in θ , λ and u . Thus, g_{θ} can be approximated by the Cesaro sum of its

Fourier series

$$g_{\theta}^{(L)}(u, \lambda) = \frac{1}{4\pi^2} \sum_{\ell=-L}^L \sum_{m=-L}^L \left(1 - \frac{|\ell|}{L}\right) \left(1 - \frac{|m|}{L}\right) \widehat{g}_{\theta}(\ell, m) \exp(-i2\pi u_j \ell - i\lambda m),$$

such that $\sup_{\theta} |g_{\theta}(u, \lambda) - g_{\theta}^{(L)}(u, \lambda)| < \varepsilon$, see for example Theorem 1.5 (ii) of Körner (1988). Following Theorem 3.2 of Dahlhaus (1997) we can write

$$\begin{aligned} \sup_{\theta} |\mathcal{L}_T(\theta) - \mathcal{L}(\theta)| &\leq \mathcal{O}\left(\frac{1}{M}\right) + \frac{\varepsilon}{4\pi} \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} [I_N(u_j, \lambda) + f(u_j, \lambda)] d\lambda \\ &+ \frac{1}{16\pi^3} \sum_{\ell=-L}^L \sum_{m=-L}^L \left(1 - \frac{|\ell|}{L}\right) \left(1 - \frac{|m|}{L}\right) \sup_{\theta} |\widehat{g}_{\theta}(\ell, m)| \\ &\times \left| \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \exp(-i2\pi u_j \ell - i\lambda m) \{I_N(u_j, \lambda) - f(u_j, \lambda)\} d\lambda \right|, \end{aligned}$$

where

$$\widehat{g}_{\theta}(\ell, m) = \int_0^1 \int_{-\pi}^{\pi} g_{\theta}(u, \lambda) \exp(i2\pi u \ell + i\lambda m) du d\lambda.$$

Consequently, $|\widehat{g}_{\theta}(\ell, m)| \leq 2\pi \sup_{(\theta, u, \lambda)} |g_{\theta}(u, \lambda)|$. However, by assumption A1 $|g_{\theta}(u, \lambda)|$ is continuous in θ, u and λ . Thus, since the parameter space is compact we have that $|\widehat{g}_{\theta}(\ell, m)| \leq K$, for some positive constant K . Now, by defining for fixed $\ell, m = 1, \dots, L$, $\phi(u, \lambda) = \cos(2\pi u \ell) \cos(\lambda m)$ or $\phi(u, \lambda) = \sin(2\pi u \ell) \cos(\lambda m)$ in Proposition A.1.1 and $\phi(u, \lambda) = \varphi(u, \lambda) = \cos(2\pi u \ell) \cos(\lambda m)$ or $\phi(u, \lambda) = \varphi(u, \lambda) =$

$\sin(2\pi u\ell) \cos(\lambda m)$ in Proposition A.1.2, we deduce that

$$\begin{aligned}
& \frac{1}{16\pi^3} \sum_{\ell=-L}^L \sum_{m=-L}^L \left(1 - \frac{|\ell|}{L}\right) \left(1 - \frac{|m|}{L}\right) \sup_{\theta} |\widehat{g}_{\theta}(\ell, m)| \\
& \times \left| \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \exp(-i2\pi u_j \ell - i\lambda m) [I_N(u_j, \lambda) - f(u_j, \lambda)] d\lambda \right| \\
& \leq \frac{1}{16\pi^3} \sum_{\ell=-L}^L \sum_{m=-L}^L \left(1 - \frac{|\ell|}{L}\right) \left(1 - \frac{|m|}{L}\right) \sup_{\theta} |\widehat{g}_{\theta}(\ell, m)| \\
& \times \left\{ \left| \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \cos(2\pi u_j \ell) \cos(\lambda m) [I_N(u_j, \lambda) - f(u_j, \lambda)] d\lambda \right| \right. \\
& \quad \left. + \left| \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \sin(2\pi u_j \ell) \cos(\lambda m) [I_N(u_j, \lambda) - f(u_j, \lambda)] d\lambda \right| \right\} \rightarrow 0,
\end{aligned} \tag{A.16}$$

and,

$$\frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \{I_N(u_j, \lambda) + f(u_j, \lambda)\} d\lambda \rightarrow 2 \int_0^1 \int_{-\pi}^{\pi} f(u, \lambda) d\lambda du, \tag{A.17}$$

in probability, as $M \rightarrow \infty$. Now, from the limits (A.16) and (A.17), this theorem follows.

Demostración A.3.2 (Theorem 2) Let $\widehat{\theta}_T$ be the parameter value that minimizes the Whittle log-likelihood function $\mathcal{L}_T(\theta)$ given by (??) and let θ_0 be the true value of the parameter. By the mean value theorem there exists a vector $\bar{\theta}_T$ satisfying $\|\bar{\theta}_T - \theta_0\| \leq \|\widehat{\theta}_T - \theta_0\|$, such that

$$\nabla \mathcal{L}_T(\widehat{\theta}_T) - \nabla \mathcal{L}_T(\theta_0) = [\nabla^2 \mathcal{L}_T(\bar{\theta}_T)] (\widehat{\theta}_T - \theta_0).$$

Therefore, it suffices to show that (a) $\nabla^2 \mathcal{L}_T(\theta_0) \rightarrow \Gamma(\theta_0)$, as $T \rightarrow \infty$; (b) $\nabla^2 \mathcal{L}_T(\bar{\theta}_T) - \nabla^2 \mathcal{L}_T(\theta_0) \rightarrow 0$ in probability, as $T \rightarrow \infty$; and (c) $\sqrt{T} \nabla \mathcal{L}_T(\theta_0) \rightarrow N[0, \Gamma(\theta_0)]$, in distribution, as $T \rightarrow \infty$. To this end, observe that

$$\begin{aligned}
\nabla^2 \mathcal{L}_T(\theta) &= \frac{1}{4\pi} \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} [I_N(u_j, \lambda) - f_\theta(u_j, \lambda)] \nabla^2 f_\theta(u_j, \lambda)^{-1} \\
&\quad - \nabla f_\theta(u_j, \lambda) [\nabla f_\theta(u_j, \lambda)^{-1}]' d\lambda \\
&= \frac{1}{4\pi} \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) [I_N(u_j, \lambda) - f_\theta(u_j, \lambda)] \\
&\quad + \frac{1}{4\pi} \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \nabla \log f_\theta(u_j, \lambda) [\nabla \log f_\theta(u_j, \lambda)]' d\lambda \\
&= \frac{1}{4\pi} [J_T(\phi) - J(\phi)] + \Gamma(\theta) + \mathcal{O}\left(\frac{1}{M}\right),
\end{aligned}$$

where $\phi(u, \lambda) = \nabla^2 f_\theta(u, \lambda)^{-1}$. Hence, an application of Proposition A.1.1 and Proposition A.1.2 yields parts (a) and (b). On the other hand, part (c) can be proved by means of the cumulant method. That is, by showing that all the cumulants of $\sqrt{T} \nabla \mathcal{L}_T(\theta_0)$ converge to zero, excepting the second order cumulant. To this end, note that

$$\begin{aligned}
\nabla \mathcal{L}_T(\theta_0) &= \frac{1}{4\pi} \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} [I_N(u_j, \lambda) - f_{\theta_0}(u_j, \lambda)] \nabla f_{\theta_0}(u_j, \lambda)^{-1} d\lambda \\
&= \frac{1}{4\pi} J_T(\phi) - \frac{1}{4\pi} \sum_{j=1}^M \int_{-\pi}^{\pi} f_{\theta_0}(u_j, \lambda) \nabla f_{\theta_0}(u_j, \lambda)^{-1} d\lambda \\
&= \frac{1}{4\pi} [J_T(\phi) - J(\phi)] + \mathcal{O}\left(\frac{1}{M}\right),
\end{aligned}$$

where $\phi(u, \lambda) = \nabla f_{\theta_0}(u, \lambda)^{-1}$. Hence, by Proposition A.1.1 and assumption A3, the

first order cumulant of $\sqrt{T}\nabla\mathcal{L}_T(\theta_0)$ satisfies

$$\sqrt{T}E[\nabla\mathcal{L}_T(\theta_0)] = \mathcal{O}\left(\frac{\sqrt{T}\log^2 N}{N}\right) + \mathcal{O}\left(\frac{\sqrt{T}}{M}\right) \rightarrow 0,$$

as $T \rightarrow \infty$. Furthermore, by Proposition A.1.2 we have that the second order cumulant of $\sqrt{T}\nabla\mathcal{L}_T(\theta_0)$ can be written as

$$TCov[\nabla\mathcal{L}_T(\theta_0), \nabla\mathcal{L}_T(\theta_0)] = \frac{1}{16\pi^2}TCov[J_T(\phi), J_T(\phi)].$$

Therefore, by Proposition A.1.2 we have that,

$$\begin{aligned} & \lim_{T \rightarrow \infty} TCov[\nabla\mathcal{L}_T(\theta_0), \nabla\mathcal{L}_T(\theta_0)] \\ &= \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \nabla f_{\theta_0}(u, \lambda)^{-1} [\nabla f_{\theta_0}(u, \lambda)^{-1}]' f_{\theta_0}(u, \lambda)^2 d\lambda du \\ &= \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \nabla \log f_{\theta_0}(u, \lambda) [\nabla \log f_{\theta_0}(u, \lambda)]' d\lambda du = \Gamma(\theta_0). \end{aligned}$$

Finally, for $p > 2$, Proposition A.1.3 gives $T^{p/2}cum_p[\nabla\mathcal{L}_T(\theta_0)] \rightarrow 0$, as $T \rightarrow \infty$, proving part (c).

Demostración A.3.3 (Theorem 3) By observing that the Fisher information matrix evaluated at the true parameter, $\Gamma_T(\theta_0)$, is given by

$$\Gamma_T(\theta_0) = TCov[\nabla\mathcal{L}_T(\theta_0), \nabla\mathcal{L}_T(\theta_0)],$$

the result is an immediate consequence of Proposition A.1.2.

Demostración A.3.4 (*Theorem 4*). Let $V_p^{(T)} = \text{Var}(\widehat{\beta})$, then

$$\begin{aligned} \int_0^1 \text{Var}[\widehat{d}(u)] &= \int_0^1 \sum_{i=1}^p \sum_{j=1}^p g_i(u) V_{ij}^{(T)} g_j(u) d(u) \\ &= \sum_{i=1}^p \sum_{j=1}^p V_{ij}^{(T)} \int_0^1 g_i(u) g_j(u) d(u) = \sum_{i=1}^p \sum_{j=1}^p V_{ij}^{(T)} b_{ij}, \end{aligned}$$

where

$$b_{ij} = \int_0^1 g_i(u) g_j(u) d(u) = b_{ji}.$$

Therefore, by Theorem 3.2.2

$$\lim_{T \rightarrow \infty} T \int_0^1 \text{Var}[\widehat{d}(u)] du = \sum_{i=1}^p \sum_{j=1}^p \lim_{T \rightarrow \infty} V_{ij}^{(T)} b_{ij} = \sum_{i=1}^p \sum_{j=1}^p a_{ij} b_{ij},$$

where $A = \Gamma^{-1}$ and $\Gamma = \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \frac{\partial}{\partial \beta_i} \log f(u, \lambda) \frac{\partial}{\partial \beta_j} \log f(u, \lambda) d\lambda du$. But, $\log f(u, \lambda) = \log(\sigma^2) - \log(2\pi) - d_\beta(u) \log |1 - e^{i\lambda}|^2$. Thus,

$$\frac{\partial}{\partial \beta_i} \log f(u, \lambda) = -g_i(u) \log |1 - e^{i\lambda}|^2.$$

Hence,

$$\Gamma_{ij} = \int_0^1 g_i(u) g_j(u) du \times \frac{1}{4\pi} \int_{-\pi}^{\pi} \left(\log |1 - e^{i\lambda}|^2 \right)^2 d\lambda = \frac{\pi^2}{6} b_{ij},$$

and then $\Gamma = \frac{\pi^2}{6} B$ and $A = \frac{6}{\pi^2} B^{-1}$. Consequently, since A and B are symmetric matrices

$$\lim_{T \rightarrow \infty} T \int_0^1 \text{Var}[\widehat{d}(u)] du = \text{tr}(AB) = \frac{6}{\pi^2} \text{tr}(I_p) = \frac{6p}{\pi^2}.$$

Apéndice B

Programas

A continuación se presentan los programas que se implementaron para simular y ajustar el caso lineal descrito en el Ejemplo 3.2.2.

B.1. Simuladores

B.1.1. lsfn.trunc.sim

```
lsfn.trunc.sim <- function(n = 1000, m = 100, alpha = numeric(), beta = numeric(), z = rnorm(n)){  
  z <- c(rep(0,m), z)  
  
  aux1 <- length(beta)  
  aux2 <- length(alpha)  
  u <- (1:n)/n  
  up <- matrix(NA, ncol = aux1, nrow = n)  
  uq <- matrix(NA, ncol = aux2, nrow = n)  
  
  for(j in 1:(aux1)){  
    up[,j] <- u^(j-1)  
  }  
  
  for(j in 1:(aux2)){  
    uq[,j] <- u^(j-1)  
  }  
  
  d <- up%*%beta  
  sigma <- uq%*%alpha  
  
  y <- vector("numeric")  
  k <- 1:m  
  
  for(t in 1:n){  
    Psi <- exp(lgamma(d[t]+k) - lgamma(k + 1)-lgamma(d[t]))  
    y[t] <- z[t + m] + sum( z[t:(m + t - 1)] * rev(Psi) )  
  }  
}
```

```

y <- as.vector(y*sigma)
return(y)
}

```

B.1.2. lsfm.innov.sim y kapp

```

kapp <- function(n, beta = numeric(), alpha = numeric()){

  aux1 <- length(beta)
  aux2 <- length(alpha)
  u <- (1:n)/n
  up <- matrix(NA, ncol = aux1, nrow = n)
  uq <- matrix(NA, ncol = aux2, nrow = n)

  for(j in 1:(aux1)){
    up[,j] <- u^(j-1)
  }

  for(j in 1:(aux2)){
    uq[,j] <- u^(j-1)
  }

  d <- up%*%beta
  sigma <- uq%*%alpha

  out = matrix(0,nc=n,nr=n)
  for (i in 1:n){
    j = 1:i
    a0 = gamma(1-d[i]-d[j])
    a1 = lgamma(i-j+d[i])
    a2 = lgamma(i-j+1-d[j])
    a3 = gamma(1-d[i])*gamma(d[i])
    out[i,j] = sigma[i] * sigma[j] * a0 * exp(a1-a2)/a3
  }
  out = out + t(out)
  diag(out) = diag(out)/2

  return(out)
}

lsfm.innov.sim <- function(n = 200, alpha = numeric(), beta = numeric(), z = rnorm(n)){
  a=kapp(n, beta = beta, alpha = alpha)
  b=chol(a)
  y=t(b)%*%z
  return(as.vector(y))
}

```

B.1.3. lsfm.fracdiff.sim

```

lsfm.fracdiff.sim <- function(n = 1000, ar = NULL, ma = NULL, beta = numeric(), alpha = numeric(),
p = length(ar), q = length(ma), z = rnorm(n + q)){
#####
## Requiere paquete: "fracdiff" ##
#####
  aux1 <- length(beta)
  aux2 <- length(alpha)
  u <- (1:n)/n

```

```

up <- matrix(NA, ncol = aux1, nrow = n)
uq <- matrix(NA, ncol = aux2, nrow = n)

for(j in 1:(aux1)){
  up[,j] <- u^(j-1)
}

for(j in 1:(aux2)){
  uq[,j] <- u^(j-1)
}

d <- up%*%beta
sigma <- uq%*%alpha

y <- vector("numeric")

for(i in 1:n){
  aux <- fracdiff.sim(n = n, ar = ar, ma = ma, d = d[i], innov = z)$series[i]
  y[i] <- sigma[i]*aux
}
return(y)
}

```


B.2. Estimador Whittle

B.2.1. lsfn.whittle

```
lsfn.whittle <- function(series1, N, S, model = list(p = 1, q = 1), start = list(beta.0 = 0.2, beta.1 = 0,
beta.2 = 0, beta.3 = 0, alpha.0 = 1, alpha.1 = 0, alpha.2 = 0), iter.max = 150, eval.max = 200){
  T <- length(series1)

  if(model$p==1 & model$q == 1){
    start = c(start$beta.0, start$beta.1, start$alpha.0, start$alpha.1)
    aux <- nlminb(start = start, whittle.nonstat.loglik.case.4,
series = series1, N = N, S = S,
control = list(iter.max = iter.max, eval.max = eval.max))
    par <- aux$par
    loglik <- -aux$objective
    aic <- -2*loglik + 2 * length(par)/T
    convergence <- aux$convergence

    summary <- cbind(par,
var.nonstat.3(length(series1), beta = par[1:2], alpha = par[3:4])$SD,
par/var.nonstat.3(length(series1), beta = par[1:2], alpha = par[3:4])$SD)
rownames(summary) <- c("beta.0", "beta.1", "alpha.0", "alpha.1")
colnames(summary) <- c("Estimate", "S.D.", "t value")
summary
suppressWarnings(return(par, loglik, aic, convergence, summary))
  }

  else{
    if(model$p==1 & model$q == 0){
      start = c(start$beta.0, start$beta.1, start$alpha.0)
      aux <- nlminb(start = start, whittle.nonstat.loglik.case.3, series = series1, N = N, S = S)
      par <- aux$par
      loglik <- -aux$objective
      aic <- -2*loglik + 2 * length(par)/T
      convergence <- aux$convergence

      summary <- cbind(par,
var.nonstat.3(length(series1), beta = par[1:2], alpha = par[3])$SD,
par/var.nonstat.3(length(series1), beta = par[1:2], alpha = par[3])$SD)
rownames(summary) <- c("beta.0", "beta.1", "alpha.0")
colnames(summary) <- c("Estimate", "S.D.", "t value")
summary
suppressWarnings(return(par, loglik, aic, convergence, summary))
    }

    else{
      if(model$p==0 & model$q == 1){
        start = c(start$beta.0, start$alpha.0, start$alpha.1)
        aux <- nlminb(start = start, whittle.nonstat.loglik.case.2, series = series1, N = N, S = S)
        par <- aux$par
        loglik <- -aux$objective
        aic <- -2*loglik + 2 * length(par)/T
        convergence <- aux$convergence

        summary <- cbind(par,
var.nonstat.3(length(series1), beta = par[1], alpha = par[2:3])$SD,
par/var.nonstat.3(length(series1), beta = par[1], alpha = par[2:3])$SD)
rownames(summary) <- c("beta.0", "alpha.0", "alpha.1")
colnames(summary) <- c("Estimate", "S.D.", "t value")
summary
suppressWarnings(return(par, loglik, aic, convergence, summary))
      }
    }
  }
}
```

```

}
else{
  if(model$p==0 & model$q == 0){
    start = c(start$beta.0, start$alpha.0)
    aux <- nlminb(start = start, whittle.nonstat.loglik.case.1, series = series1, N = N, S = S)
    par <- aux$par
    loglik <- -aux$objective
    aic <- -2*loglik + 2 * length(par)/T
    convergence <- aux$convergence

    summary <- cbind(par,
      var.nonstat.3(length(series1), beta = par[1], alpha = par[2])$SD,
      par/var.nonstat.3(length(series1), beta = par[1], alpha = par[2])$SD)
    rownames(summary) <- c("beta.0", "alpha.0")
    colnames(summary) <- c("Estimate", "S.D.", "t value")
    summary
    suppressWarnings(return(par, loglik, aic, convergence, summary))
  }
  else{
    if(model$p==0 & model$q == 2){
      start = c(start$beta.0, start$alpha.0, start$alpha.1, start$alpha.2)
      aux <- nlminb(start = start, whittle.nonstat.loglik.case.5, series = series1, N = N, S = S)
      par <- aux$par
      loglik <- -aux$objective
      aic <- -2*loglik + 2 * length(par)/T
      convergence <- aux$convergence

      summary <- cbind(par,
        var.nonstat.3(length(series1), beta = par[1], alpha = par[2:4])$SD,
        par/var.nonstat.3(length(series1), beta = par[1], alpha = par[2:4])$SD)
      rownames(summary) <- c("beta.0", "alpha.0", "alpha.1", "alpha.2")
      colnames(summary) <- c("Estimate", "S.D.", "t value")
      summary
      suppressWarnings(return(par, loglik, aic, convergence, summary))
    }
    else{
      if(model$p==1 & model$q == 2){
        start = c(start$beta.0, start$beta.1, start$alpha.0, start$alpha.1, start$alpha.2)
        aux <- nlminb(start = start, whittle.nonstat.loglik.case.6, series = series1, N = N, S = S)
        par <- aux$par
        loglik <- -aux$objective
        aic <- -2*loglik + 2 * length(par)/T
        convergence <- aux$convergence

        summary <- cbind(par,
          var.nonstat.3(length(series1), beta = par[1:2], alpha = par[3:5])$SD,
          par/var.nonstat.3(length(series1), beta = par[1:2], alpha = par[3:5])$SD)
        rownames(summary) <- c("beta.0", "beta.1", "alpha.0", "alpha.1", "alpha.2")
        colnames(summary) <- c("Estimate", "S.D.", "t value")
        summary
        suppressWarnings(return(par, loglik, aic, convergence, summary))
      }
      else{
        if(model$p==2 & model$q == 0){
          start = c(start$beta.0, start$beta.1, start$beta.2, start$alpha.0)
          aux <- nlminb(start = start, whittle.nonstat.loglik.case.7, series = series1, N = N, S = S)
          par <- aux$par
          loglik <- -aux$objective
          aic <- -2*loglik + 2 * length(par)/T
          convergence <- aux$convergence

```

```

summary <- cbind(par,
var.nonstat.3(length(series1), beta = par[1:3], alpha = par[4])$SD,
par/var.nonstat.3(length(series1), beta = par[1:3], alpha = par[4])$SD)
rownames(summary) <- c("beta.0", "beta.1", "beta.2", "alpha.0")
colnames(summary) <- c("Estimate", "S.D.", "t value")
summary
suppressWarnings(return(par, loglik, aic, convergence, summary))
}

elseif
if(model$p==2 & model$q == 1){
start = c(start$beta.0, start$beta.1, start$beta.2, start$alpha.0, start$alpha.1)
aux <- nlminb(start = start, whittle.nonstat.loglik.case.8, series = series1, N = N, S = S)
par <- aux$par
loglik <- -aux$objective
aic <- -2*loglik + 2 * length(par)/T
convergence <- aux$convergence

summary <- cbind(par,
var.nonstat.3(length(series1), beta = par[1:3], alpha = par[4:5])$SD,
par/var.nonstat.3(length(series1), beta = par[1:3], alpha = par[4:5])$SD)
rownames(summary) <- c("beta.0", "beta.1", "beta.2", "alpha.0", "alpha.1")
colnames(summary) <- c("Estimate", "S.D.", "t value")
summary
suppressWarnings(return(par, loglik, aic, convergence, summary))
}

elseif
if(model$p==2 & model$q == 2){
start = c(start$beta.0, start$beta.1, start$beta.2, start$alpha.0, start$alpha.1, start$alpha.2)
aux <- nlminb(start = start, whittle.nonstat.loglik.case.9, series = series1, N = N, S = S)
par <- aux$par
loglik <- -aux$objective
aic <- -2*loglik + 2 * length(par)/T
convergence <- aux$convergence

summary <- cbind(par,
var.nonstat.3(length(series1), beta = par[1:3], alpha = par[4:6])$SD,
par/var.nonstat.3(length(series1), beta = par[1:3], alpha = par[4:6])$SD)
rownames(summary) <- c("beta.0", "beta.1", "beta.2", "alpha.0", "alpha.1", "alpha.2")
colnames(summary) <- c("Estimate", "S.D.", "t value")
summary
suppressWarnings(return(par, loglik, aic, convergence, summary))
}

elseif
if(model$p==3 & model$q == 0){
start = c(start$beta.0, start$beta.1, start$beta.2, start$beta.3, start$alpha.0)
aux <- nlminb(start = start, whittle.nonstat.loglik.case.10, series = series1, N = N, S = S)
par <- aux$par
loglik <- -aux$objective
aic <- -2*loglik + 2 * length(par)/T
convergence <- aux$convergence

summary <- cbind(par,
var.nonstat.3(length(series1), beta = par[1:4], alpha = par[5])$SD,
par/var.nonstat.3(length(series1), beta = par[1:4], alpha = par[5])$SD)
rownames(summary) <- c("beta.0", "beta.1", "beta.2", "beta.3", "alpha.0")
colnames(summary) <- c("Estimate", "S.D.", "t value")
summary
suppressWarnings(return(par, loglik, aic, convergence, summary))
}

elseif
if(model$p==3 & model$q == 1){

```



```

a <- (4 * a[1:m])/(pi * 3 * n)
#
#Fourier frequencies
#
w <- (pi * (1:m))/m
#
#Spectral Density:
#
b <- sigma^2*fn.density(lambda = w, d = d)
#
# Whittle Log-likelihood
#
loglik <- 2 * pi * ( sum(log(b)) + sum(a/b) )
return(loglik/n)
}

```

B.2.3. taper

```

taper <- function(x){
n <- length(x)
a <- 0.5*(1-cos(2*pi*(0:(n-1))/n))
return(x*a)
}

```

B.2.4. fn.density

```

fn.density <- function(lambda, d){
#
# Spectral density of an ARFIMA(0,d,0) process
#
a <- (2 * sin(lambda/2))^(2 * d)
a <- a/(2 * pi)
return(a)
}

```

B.2.5. var.nonstat.3

```

var.nonstat.3 <- function(T = 1, alpha = numeric(), beta = numeric(), subdivisions = 100){
#####
## Calcula varianza teorica. ##
## Case: ##
## d(u) = b0 + b1 * u + ... + bp * u^p ##
## sigma(u) = a0 + a1 * u + ... + ap * u^p ##
#####

if(length(alpha)==0){
if(length(beta)==0){
warning("Ingresar coeficientes para d(u) y sigma(u)")
}
else{
warning("Ingresar coeficientes de sigma(u)")
}
}
}

```

```

else{
  if(length(beta)==0){
    warning("Ingresar coeficientes para d(u)")
  }
  else{
    a <- length(beta)
    b <- length(alpha)
    VAR <- matrix(0, ncol = a + b, nrow = a + b)
    VAR[(1:a), (1:a)] <- var.beta(T = T, p = length(beta)-1)
    VAR[(a+1):(a+b), (a+1):(a+b)] <- var.alpha(alpha = alpha, subdivisions = subdivisions, T = T)
    SD <- sqrt(diag(VAR))
    suppressWarnings(return(VAR,SD))
  }
}

}

```

B.2.6. var.beta y var.alpha

```

var.beta <- function(T = 1, p = 0){
  M <- matrix(NA, nrow = p + 1, ncol = p + 1)
  for(i in 1:(p+1)){
    for(j in 1:(p+1)){
      M[i,j] <- pi^2/(6*((i-1)+(j-1)+1))
    }
  }
  solve(M)/T
}

var.alpha <- function(alpha = numeric(), subdivisions = 100, T = 1){
  q <- length(alpha)
  if(q==0){
    warning("Ingresar coeficientes")
  }
  else{
    f <- function(u, alpha, i, j){
      z <- vector("numeric")
      for(k in 1:length(u)){
        z[k] <- u[k]^(i+j)/(sum(u[k]^(0:(q-1))*alpha))^2
      }
      z
    }

    M <- matrix(NA, nrow = q, ncol = q)

    for(i in 0:(q-1)){
      for(j in 0:(q-1)){
        M[i+1,j+1] <- integrate(f, lower = 0, upper = 1, alpha = alpha,
          subdivisions = subdivisions, i = i, j = j)$value
      }
    }
    solve(2*M)/T
  }
}

```

B.2.7. whittle.nonstat.loglik.case.1 hasta whittle.nonstat.loglik.case.12

```

whittle.nonstat.loglik.case.1 <- function(x, series, N, S){
  ## Case 1: d(u) = beta.0, sigma(u) = alpha.0 ##
  T1 <- length(series)
  M <- trunc((T1 - N + S)/S)
  beta.0 <- x[1]
  beta.1 <- 0
  alpha.0 <- x[2]
  alpha.1 <- 0
  #
  # Verifying d in (0,1/2)
  #
  beta.min <- beta.0 - max(0, - beta.1)
  beta.max <- beta.0 - min(.49, .49 - beta.1)
  alpha.sum <- alpha.0+alpha.1

  if((beta.min < 0) || (beta.max >= 0) || (alpha.sum == 0) )
    loglik = 1000000000000.
  else {
    cc <- vector("numeric")
    for(j in 1:M) {
      d <- beta.0 + beta.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1)
      sigma <- alpha.0 + alpha.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1)
      i.ini <- S * (j - 1) + 1
      i.fin <- i.ini + N - 1
      series.j <- series[i.ini:i.fin]
      cc[j] <- whittle.taper.loglik.new.1(d = d, series = series.j, sigma = sigma)
    }
    loglik <- mean(cc)/(4*pi)
  }
  return(loglik)
}

whittle.nonstat.loglik.case.2 <- function(x, series, N, S){
  ## Case 2: d(u) = beta.0, sigma(u) = alpha.0 + alpha.1 * u##
  T1 <- length(series)
  M <- trunc((T1 - N + S)/S)
  beta.0 <- x[1]
  beta.1 <- 0
  alpha.0 <- x[2]
  alpha.1 <- x[3]
  #
  # Verifying d in (0,1/2)
  #
  beta.min <- beta.0 - max(0, - beta.1)
  beta.max <- beta.0 - min(.49, .49 - beta.1)
  alpha.sum <- alpha.0+alpha.1

  if((beta.min < 0) || (beta.max >= 0) || (alpha.sum == 0) )
    loglik = 1000000000000.
  else {
    cc <- vector("numeric")
    for(j in 1:M) {
      d <- beta.0 + beta.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1)
      sigma <- alpha.0 + alpha.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1)
      i.ini <- S * (j - 1) + 1
      i.fin <- i.ini + N - 1

```

```

series.j <- series[i.ini:i.fin]
cc[j] <- whittle.taper.loglik.new.1(d = d, series = series.j, sigma = sigma)
}
loglik <- mean(cc)/(4*pi)
}
return(loglik)
}

whittle.nonstat.loglik.case.3 <- function(x, series, N, S){
## Case 3: d(u) = beta.0 + beta.1 * u, sigma(u) = alpha.0##
T1 <- length(series)
M <- trunc((T1 - N + S)/S)
beta.0 <- x[1]
beta.1 <- x[2]
alpha.0 <- x[3]
alpha.1 <- 0
#
# Verifying d in (0,1/2)
#
beta.min <- beta.0 - max(0, - beta.1)
beta.max <- beta.0 - min(.49, .49 - beta.1)
alpha.sum <- alpha.0+alpha.1

if((beta.min < 0) || (beta.max >= 0) || (alpha.sum == 0) )
loglik = 1000000000000.
else {
cc <- vector("numeric")
for(j in 1:M) {
d <- beta.0 + beta.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1)
sigma <- alpha.0 + alpha.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1)
i.ini <- S * (j - 1) + 1
i.fin <- i.ini + N - 1
series.j <- series[i.ini:i.fin]
cc[j] <- whittle.taper.loglik.new.1(d = d, series = series.j, sigma = sigma)
}
loglik <- mean(cc)/(4*pi)
}
return(loglik)
}

whittle.nonstat.loglik.case.4 <- function(x, series, N, S){
## Case 4: d(u) = beta.0 + beta.1 * u, sigma(u) = alpha.0 + alpha.1 * u##
T1 <- length(series)
M <- trunc((T1 - N + S)/S)
beta.0 <- x[1]
beta.1 <- x[2]
alpha.0 <- x[3]
alpha.1 <- x[4]
#
# Verifying d in (0,1/2)
#
beta.min <- beta.0 - max(0, - beta.1)
beta.max <- beta.0 - min(.49, .49 - beta.1)
alpha.sum <- alpha.0+alpha.1

if((beta.min < 0) || (beta.max >= 0) || (alpha.sum == 0) )
loglik = 1000000000000.
else {
cc <- vector("numeric")
for(j in 1:M) {
d <- beta.0 + beta.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1)

```



```

sigma <- alpha.0 + alpha.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1)
i.ini <- S * (j - 1) + 1
i.fin <- i.ini + N - 1
series.j <- series[i.ini:i.fin]
cc[j] <- whittle.taper.loglik.new.1(d = d, series = series.j, sigma = sigma)
}
loglik <- mean(cc)/(4*pi)
}
return(loglik)
}

whittle.nonstat.loglik.case.5 <- function(x, series, N, S){
## Case 4: d(u) = beta.0, sigma(u) = alpha.0 + alpha.1 * u + alpha.2 * u^2##
T1 <- length(series)
M <- trunc((T1 - N + S)/S)
beta.0 <- x[1]
beta.1 <- 0
alpha.0 <- x[2]
alpha.1 <- x[3]
alpha.2 <- x[4]
#
# Verifying d in (0,1/2)
#
beta.min <- beta.0 - max(0, - beta.1)
beta.max <- beta.0 - min(.49, .49 - beta.1)
alpha.sum <- alpha.0+alpha.1+alpha.2
if((beta.min < 0) || (beta.max >= 0) || ((alpha.2 > 0) && (alpha.1^2 > 4 * alpha.0 * alpha.2))
|| ((alpha.2 < 0) && ( (alpha.0 < 0) || (alpha.sum < 0) ) ))
loglik = 100000000000.
else {
cc <- vector("numeric")
for(j in 1:M) {
d <- beta.0 + beta.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1)
sigma <- alpha.0 + alpha.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1)
+ alpha.2 * ((S * (j - 1))/T1 + (0.5 * N)/T1)^2
i.ini <- S * (j - 1) + 1
i.fin <- i.ini + N - 1
series.j <- series[i.ini:i.fin]
cc[j] <- whittle.taper.loglik.new.1(d = d, series = series.j, sigma = sigma)
}
loglik <- mean(cc)/(4*pi)
}
return(loglik)
}

whittle.nonstat.loglik.case.6 <- function(x, series, N, S){
## Case 4: d(u) = beta.0 + beta.1 * u, sigma(u) = alpha.0 + alpha.1 * u + alpha.2 * u^2##
T1 <- length(series)
M <- trunc((T1 - N + S)/S)
beta.0 <- x[1]
beta.1 <- x[2]
alpha.0 <- x[3]
alpha.1 <- x[4]
alpha.2 <- x[5]
#
# Verifying d in (0,1/2)
#
beta.min <- beta.0 - max(0, - beta.1)
beta.max <- beta.0 - min(.49, .49 - beta.1)
alpha.sum <- alpha.0+alpha.1+alpha.2

```

```

if((beta.min < 0) || (beta.max >= 0) || ((alpha.2 > 0) && (alpha.1^2 > 4 * alpha.0 * alpha.2))
|| (alpha.2 < 0) && ( (alpha.0 < 0) || (alpha.sum < 0) ) )
loglik = 100000000000.
else {
cc <- vector("numeric")
for(j in 1:M) {
d <- beta.0 + beta.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1)
sigma <- alpha.0 + alpha.1 * ((S * (j - 1))/T1
+ (0.5 * N)/T1) + alpha.2 * ((S * (j - 1))/T1 + (0.5 * N)/T1)^2
i.ini <- S * (j - 1) + 1
i.fin <- i.ini + N - 1
series.j <- series[i.ini:i.fin]
cc[j] <- whittle.taper.loglik.new.1(d = d, series = series.j, sigma = sigma)
}
loglik <- mean(cc)/(4*pi)
}
return(loglik)
}

whittle.nonstat.loglik.case.7 <- function(x, series, N, S){
## Case 4: d(u) = beta.0 + beta.1 * u + beta.2 * u^2, sigma(u) = alpha.0##
T1 <- length(series)
M <- trunc((T1 - N + S)/S)
beta.0 <- x[1]
beta.1 <- x[2]
beta.2 <- x[3]
alpha.0 <- x[4]
alpha.1 <- 0
alpha.2 <- 0
u <- seq(0,1,0.001)
d <- beta.0 + beta.1 * u + beta.2 * u^2
s <- alpha.0 + alpha.1 * u + alpha.2 * u^2
d.max <- max(d)
d.min <- min(d)
s.max <- max(s)
s.min <- min(s)

if(d.max >= 0.5 || d.min <= 0 || s.min <= 0){
loglik = 100000000000.
}
else {
cc <- vector("numeric")
for(j in 1:M) {
d <- beta.0 + beta.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1) + beta.2 * ((S * (j - 1))/T1 + (0.5 * N)/T1)^2
sigma <- alpha.0 + alpha.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1) + alpha.2 * ((S * (j - 1))/T1 + (0.5 * N)/T1)^2
i.ini <- S * (j - 1) + 1
i.fin <- i.ini + N - 1
series.j <- series[i.ini:i.fin]
cc[j] <- whittle.taper.loglik.new.1(d = d, series = series.j, sigma = sigma)
}
loglik <- mean(cc)/(4*pi)
}
return(loglik)
}

whittle.nonstat.loglik.case.8 <- function(x, series, N, S){
## Case 4: d(u) = beta.0 + beta.1 * u + beta.2 * u^2, sigma(u) = alpha.0 + alpha.1 * u##
T1 <- length(series)
M <- trunc((T1 - N + S)/S)
beta.0 <- x[1]

```

```

beta.1 <- x[2]
beta.2 <- x[3]
alpha.0 <- x[4]
alpha.1 <- x[5]
alpha.2 <- 0
u <- seq(0,1,0.001)
d <- beta.0 + beta.1 * u + beta.2 * u^2
s <- alpha.0 + alpha.1 * u + alpha.2 * u^2
d.max <- max(d)
d.min <- min(d)
s.max <- max(s)
s.min <- min(s)

if(d.max >= 0.5 || d.min <= 0 || s.min <= 0){
  loglik = 100000000000.
}
else {
  cc <- vector("numeric")
  for(j in 1:M) {
    d <- beta.0 + beta.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1) + beta.2 * ((S * (j - 1))/T1 + (0.5 * N)/T1)^2
    sigma <- alpha.0 + alpha.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1) + alpha.2 * ((S * (j - 1))/T1 + (0.5 * N)/T1)^2
    i.ini <- S * (j - 1) + 1
    i.fin <- i.ini + N - 1
    series.j <- series[i.ini:i.fin]
    cc[j] <- whittle.taper.loglik.new.1(d = d, series = series.j, sigma = sigma)
  }
  loglik <- mean(cc)/(4*pi)
}
return(loglik)
}

whittle.nonstat.loglik.case.9 <- function(x, series, N, S){
  ## Case 4: d(u) = beta.0 + beta.1 * u + beta.2 * u^2, sigma(u) = alpha.0 + alpha.1 * u + alpha.2 * u^2##
  T1 <- length(series)
  M <- trunc((T1 - N + S)/S)
  beta.0 <- x[1]
  beta.1 <- x[2]
  beta.2 <- x[3]
  alpha.0 <- x[4]
  alpha.1 <- x[5]
  alpha.2 <- x[6]
  #
  # Verifying d in (0,1/2) and sigma in (0, +infty)
  #
  u <- seq(0,1,0.001)
  d <- beta.0 + beta.1 * u + beta.2 * u^2
  s <- alpha.0 + alpha.1 * u + alpha.2 * u^2
  d.max <- max(d)
  d.min <- min(d)
  s.max <- max(s)
  s.min <- min(s)

  if(d.max >= 0.5 || d.min <= 0 || s.min <= 0){
    loglik = 100000000000.
  }

  else {
    cc <- vector("numeric")
    for(j in 1:M) {
      d <- beta.0 + beta.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1) + beta.2 * ((S * (j - 1))/T1 + (0.5 * N)/T1)^2
      sigma <- alpha.0 + alpha.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1) + alpha.2 * ((S * (j - 1))/T1 + (0.5 * N)/T1)^2

```

```

i.ini <- S * (j - 1) + 1
i.fin <- i.ini + N - 1
series.j <- series[i.ini:i.fin]
cc[j] <- whittle.taper.loglik.new.1(d = d, series = series.j, sigma = sigma)
}
loglik <- mean(cc)/(4*pi)
}
return(loglik)
}

whittle.nonstat.loglik.case.10 <- function(x, series, N, S){
## Case 10: d(u) = beta.0 + beta.1 * u + beta.2 * u^2 + beta.3 * u^3, sigma(u) = alpha.0##
T1 <- length(series)
M <- trunc((T1 - N + S)/S)
beta.0 <- x[1]
beta.1 <- x[2]
beta.2 <- x[3]
beta.3 <- x[4]
alpha.0 <- x[5]
alpha.1 <- 0
alpha.2 <- 0
#
# Verifying d in (0,1/2) and sigma in (0, +infty)
#
u <- seq(0,1,0.001)
d <- beta.0 + beta.1 * u + beta.2 * u^2 + beta.3 * u^3
s <- alpha.0 + alpha.1 * u + alpha.2 * u^2
d.max <- max(d)
d.min <- min(d)
s.max <- max(s)
s.min <- min(s)

if(d.max >= 0.5 || d.min <= 0 || s.min <= 0){
loglik = 100000000000.
}

else {
cc <- vector("numeric")
for(j in 1:M) {
d <- beta.0 + beta.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1) + beta.2 * ((S * (j - 1))/T1 + (0.5 * N)/T1)^2
sigma <- alpha.0 + alpha.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1) + alpha.2 * ((S * (j - 1))/T1 + (0.5 * N)/T1)^2
i.ini <- S * (j - 1) + 1
i.fin <- i.ini + N - 1
series.j <- series[i.ini:i.fin]
cc[j] <- whittle.taper.loglik.new.1(d = d, series = series.j, sigma = sigma)
}
loglik <- mean(cc)/(4*pi)
}
return(loglik)
}

whittle.nonstat.loglik.case.11 <- function(x, series, N, S){
## Case 11: d(u) = beta.0 + beta.1 * u + beta.2 * u^2 + beta.3 * u^3,
## sigma(u) = alpha.0 + alpha.1 * u
T1 <- length(series)
M <- trunc((T1 - N + S)/S)
beta.0 <- x[1]
beta.1 <- x[2]
beta.2 <- x[3]
beta.3 <- x[4]
alpha.0 <- x[5]

```

```

alpha.1 <- x[6]
alpha.2 <- 0
#
# Verifying d in (0,1/2) and sigma in (0, +infty)
#
u <- seq(0,1,0.001)
d <- beta.0 + beta.1 * u + beta.2 * u^2 + beta.3 * u^3
s <- alpha.0 + alpha.1 * u + alpha.2 * u^2
d.max <- max(d)
d.min <- min(d)
s.max <- max(s)
s.min <- min(s)

if(d.max >= 0.5 || d.min <= 0 || s.min <= 0){
  loglik = 100000000000.
}

else {
  cc <- vector("numeric")
  for(j in 1:M) {
    d <- beta.0 + beta.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1) + beta.2 * ((S * (j - 1))/T1 + (0.5 * N)/T1)^2
    sigma <- alpha.0 + alpha.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1) + alpha.2 * ((S * (j - 1))/T1 + (0.5 * N)/T1)^2
    i.ini <- S * (j - 1) + 1
    i.fin <- i.ini + N - 1
    series.j <- series[i.ini:i.fin]
    cc[j] <- whittle.taper.loglik.new.1(d = d, series = series.j, sigma = sigma)
  }
  loglik <- mean(cc)/(4*pi)
}
return(loglik)
}

whittle.nonstat.loglik.case.12 <- function(x, series, N, S){
  ## Case 12: d(u) = beta.0 + beta.1 * u + beta.2 * u^2 + beta.3 * u^3,
  ## sigma(u) = alpha.0 + alpha.1 * u + alpha.2 * u^2
  T1 <- length(series)
  M <- trunc((T1 - N + S)/S)
  beta.0 <- x[1]
  beta.1 <- x[2]
  beta.2 <- x[3]
  beta.3 <- x[4]
  alpha.0 <- x[5]
  alpha.1 <- x[6]
  alpha.2 <- x[7]
  #
  # Verifying d in (0,1/2) and sigma in (0, +infty)
  #
  u <- seq(0,1,0.001)
  d <- beta.0 + beta.1 * u + beta.2 * u^2 + beta.3 * u^3
  s <- alpha.0 + alpha.1 * u + alpha.2 * u^2
  d.max <- max(d)
  d.min <- min(d)
  s.max <- max(s)
  s.min <- min(s)

  if(d.max >= 0.5 || d.min <= 0 || s.min <= 0){
    loglik = 100000000000.
  }

  else {
    cc <- vector("numeric")

```

```

for(j in 1:M) {
d <- beta.0 + beta.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1)
  + beta.2 * ((S * (j - 1))/T1 + (0.5 * N)/T1)^2
sigma <- alpha.0 + alpha.1 * ((S * (j - 1))/T1 + (0.5 * N)/T1)
  + alpha.2 * ((S * (j - 1))/T1 + (0.5 * N)/T1)^2
i.ini <- S * (j - 1) + 1
i.fin <- i.ini + N - 1
series.j <- series[i.ini:i.fin]
cc[j] <- whittle.taper.loglik.new.1(d = d, series = series.j, sigma = sigma)
}
loglik <- mean(cc)/(4*pi)
}
return(loglik)
}

```

B.3. Estimador Kalman

B.3.1. lsfn.kalman

```
lsfn.kalman <- function(series1, m = 10, start = list(beta.0 = 0.2, beta.1 = 0, alpha.0 = 1, alpha.1 = 0)){
  start = c(start$beta.0, start$beta.1, start$alpha.0, start$alpha.1)
  aux <- nlminb(start = start, lsfn.kalman.loglik, series = series1, m = m)

  aux1 <- lsfn.kalman.loglik.print(x = aux$par, series = series1, m = m)
  loglik <- aux1$loglik
  res <- aux1$res
  res.stand <- aux1$res.stand
  delta <- aux1$delta
  hat.y <- aux1$hat.y
  convergence <- aux$convergence
  par <- aux$par

  suppressWarnings(return(par, res, res.stand, delta, loglik, hat.y, convergence))
}
```

B.3.2. lsfn.kalman.loglik

```
lsfn.kalman.loglik <- function(x, series, m = 10, nsample = length(series)){

  beta.0 <- x[1]
  beta.1 <- x[2]
  alpha.0 <- x[3]
  alpha.1 <- x[4]

  M <- m+1

  #####
  # Verifying d in (0,1/2) #
  #####
  beta.min <- beta.0 - max(0, - beta.1)
  beta.max <- beta.0 - min(.49, .49 - beta.1)
  alpha.sum <- alpha.0+alpha.1

  if((beta.min < 0) || (beta.max >= 0) || (alpha.sum == 0) ){
    loglik = 100000000000.
  }

  else{

    d <- vector("numeric")
    d <- beta.0 + beta.1 * (1:nsample)/nsample

    sigma <- vector("numeric")
    sigma <- alpha.0 + alpha.1 * (1:nsample)/nsample
    sigma[nsample+1]<-sigma[nsample]

    Omega <- matrix(0, nrow = M, ncol = M)
    diag(Omega) <- sigma[1]^2

    X <- rep(0, M)

    delta <- vector("numeric") ## Vector de Delta's
```

```

hat.y <- vector("numeric") ## vector de Ajuste's

for(i in 1:nsample){

  g <- sigma[i]*rev(psi(d[i], m))
  aux <- Omega %*% g
  delta[i] <- g %*% aux
  Theta <- c(aux[2:M],0)
  aux <- matrix(0, nrow = M, ncol = M)
  aux[1:m,1:m] <- Omega[2:M,2:M]
  aux[M,M] <- 1
  if(is.na(series[i])){
    Omega <- aux
    hat.y[i]<-t(g)%*%X
    X <- c(X[2:M],0)
  }
  else{
    Omega <- aux - Theta %*% t(Theta)/delta[i]
    hat.y[i]<-t(g)%*%X
    aux <- c(X[2:M],0)
    X <- aux + Theta * (series[i]-hat.y[i]) / delta[i]
  }
}
loglik <- sum(log(delta)) + sum(na.exclude((series-hat.y)^2/delta))
}
return(loglik)
}

```

B.3.3. lsfn.kalman.loglik.print

```

lsfn.kalman.loglik.print <- function(x, series, m = 10, nsample = length(series)){

  M <- m+1

  beta.0 <- x[1]
  beta.1 <- x[2]
  alpha.0 <- x[3]
  alpha.1 <- x[4]

  #####
  # Verifying d in (0,1/2) #
  #####
  beta.min <- beta.0 - max(0, - beta.1)
  beta.max <- beta.0 - min(.49, .49 - beta.1)
  alpha.sum <- alpha.0+alpha.1

  if((beta.min < 0) || (beta.max >= 0) || (alpha.sum == 0) ){
    loglik = 100000000000.
  }

  else{

    d <- vector("numeric")
    d <- beta.0 + beta.1 * (1:nsample)/nsample

    sigma <- vector("numeric")
    sigma <- alpha.0 + alpha.1 * (1:nsample)/nsample
    sigma[(nsample+1)]<-sigma[nsample]
  }
}

```



```

Omega <- matrix(0, nrow = M, ncol = M)
diag(Omega) <- sigma[i]^2

X <- rep(0, M)

delta <- vector("numeric") ## Vector de Delta's
hat.y <- vector("numeric") ## vector de Ajuste's

for(i in 1:nsample){

  g <- sigma[i]*rev(psi(d[i], m))
  aux <- Omega %*% g
  delta[i] <- g %*% aux
  Theta <- c(aux[2:M],0)
  aux <- matrix(0, nrow = M, ncol = M)
  aux[1:m,1:m] <- Omega[2:M,2:M]
  aux[M,M] <- 1
  if(is.na(series[i])){
    Omega <- aux
    hat.y[i] <- t(g)%*%X
    X <- c(X[2:M],0)
  }
  else{
    Omega <- aux - Theta %*% t(Theta)/delta[i]
    hat.y[i] <- t(g)%*%X
    aux <- c(X[2:M],0)
    X <- aux + Theta * (series[i]-hat.y[i]) / delta[i]
  }
  loglik <- sum(log(delta)) + sum(na.exclude((series-hat.y)^2/delta))
  ## Residuos ##
  res <- (series-hat.y)
  ## Residuos Estandarizados ##
  res.stand <- res/sigma[1:nsample]
}
suppressWarnings(return(loglik, res, delta, hat.y, res.stand))
}

```

Bibliografía

- J. M. Bardet, G. Lang, G. Oppenheim, A. Philippe, and M. S. Taqqu. Generators of long-range dependent processes: a survey. In *Theory and applications of long-range dependence*, pages 579–623. Birkhäuser Boston, Boston, MA, 2003. 8
- J. Beran. Statistics for long-memory processes, 1994. 43
- D. R. Brillinger. *Time series*. Holden-Day Inc., Oakland, Calif., second edition, 1981. Data analysis and theory, Holden-Day Series in Time Series Analysis. 66
- P. J. Brockwell and R. A. Davis. *Time Series: Theory and Methods*. Springer, New York, second edition, 1991. 8, 10, 36, 37
- J. M. Chambers and T. J. Hastie. *Statistical Models in S*. Wadsworth & Brooks/Cole, Pacific Grove, Ca, 1992. 13, 48
- G. Chandler and W. Polonik. Discrimination of locally stationary time series based on the excess mass functional. *J. Amer. Statist. Assoc.*, 101(473):240–253, 2006. 2

- R. Dahlhaus. Small sample effects in time series analysis: a new asymptotic theory and a new estimate. *Ann. Statist.*, 16(2):808–841, 1988. 84
- R. Dahlhaus. Fitting time series models to nonstationary processes. *Ann. Statist.*, 25(1):1–37, 1997. 1, 2, 3, 94
- R. Dahlhaus. A likelihood approximation for locally stationary processes. *Ann. Statist.*, 28(6):1762–1794, 2000. 2
- R. Dahlhaus and L. Giraitis. On the optimal segment length for parameter estimates for locally stationary time series. *J. Time Ser. Anal.*, 19(6):629–655, 1998. 2
- R. Dahlhaus and W. Polonik. Nonparametric quasi-maximum likelihood estimation for Gaussian locally stationary processes. *Ann. Statist.*, 34(6):2790–2824, 2006. 2
- R. Fox and M. S. Taqqu. Central limit theorems for quadratic forms in random variables having long-range dependence. *Probab. Theory Related Fields*, 74(2):213–240, 1987. 88
- P. Fryzlewicz, T. Sapatinas, and S. Subba Rao. A Haar-Fisz technique for locally stationary volatility estimation. *Biometrika*, 93(3):687–704, 2006. 2
- M. Genton and O. Perrin. On a time deformation reducing nonstationary stochastic processes to local stationarity. *J. Appl. Probab.*, 41(1):236–249, 2004. 2

- M. G. Görg. Long memory versus structural breaks: A time-varying memory approach. Master Thesis, 2007. 48
- I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Academic, San Diego, CA, sixth edition, 2000. 22
- D. A. Graybill. Pinus longaeva tree ring data. Mammoth Creek, Utah, 1990. 50
- W. Guo, M. Dai, H. C. Ombao, and R. von Sachs. Smoothing spline ANOVA for time-dependent spectral analysis. *J. Amer. Statist. Assoc.*, 98(463):643–652, 2003. 2
- T. J. Hastie and R. J. Tibshirani. *Generalized additive models*, volume 43 of *Mono-graphs on Statistics and Applied Probability*. Chapman and Hall Ltd., London, 1990. 13, 48
- T. W. Körner. *Fourier analysis*. Cambridge University Press, Cambridge, 1988. ISBN 0-521-25120-6. 94
- S. Orbe, E. Ferreira, and J. Rodriguez-Poo. Nonparametric estimation of time varying parameters under shape restrictions. *J. Econometrics*, 126(1):53–77, 2005. 2
- W. Palma and N. H. Chan. Estimation and forecasting of long-memory processes with missing values. *J. Forecasting*, 16:395–410, 1997. 38

- M. B. Priestley. Evolutionary spectra and non-stationary processes. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 27:204–237, 1965. 1
- K. Sakiyama and M. Taniguchi. Testing composite hypotheses for locally stationary processes. *J. Time Ser. Anal.*, 24(4):483–504, 2003. 2
- M. Tan, T.S. Liu, J. Hou, X. Qin, H. Zhang, and T. Li. Cyclic rapid warming on centennial-scale revealed by a 2650-year stalagmite record of warm season temperature. *Geophysical Research Letters*, 30:191–194, 2003. 41, 43
- S. Van Bellegem and R. Dahlhaus. Semiparametric estimation by model selection for locally stationary processes. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 68(5):721–746, 2006. 2
- R. von Sachs and B. MacGibbon. Non-parametric curve estimation by wavelet thresholding with locally stationary errors. *Scand. J. Statist.*, 27(3):475–499, 2000. 2
- P. Whittle. Estimation and information in stationary time series. *Ark. Mat.*, 2:423–434, 1953. 3
- A. Zygmund. *Trigonometric Series. Vol. I.* Cambridge University Press, Cambridge, 1959. 83