

**Dpto. de Economía Cuantitativa**  
Universidad Complutense de Madrid  
**Introducción a la**  
**Regresión Armónica Dinámica (DHR)**

Marcos Bujosa  
Antonio García-Ferrer  
Peter C. Young

Seminario perteneciente al *Máster en Análisis Económico y Economía Financiera* (MAEF) del ICAE.

© 2004–2007 Marcos Bujosa [marcos.bujosa@ccee.ucm.es](mailto:marcos.bujosa@ccee.ucm.es)  
Actualizado el: 13 de marzo de 2007

Versión 1.1

Copyright © 2004–2007 Marcos Bujosa [marcos.bujosa@ccee.ucm.es](mailto:marcos.bujosa@ccee.ucm.es)



Algunos derechos reservados. Esta obra está bajo una licencia Reconocimiento-CompartirIgual de Creative Commons. Para ver una copia de esta licencia, visite <http://creativecommons.org/licenses/by-sa/2.5/es/deed.es> o envíe una carta a Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

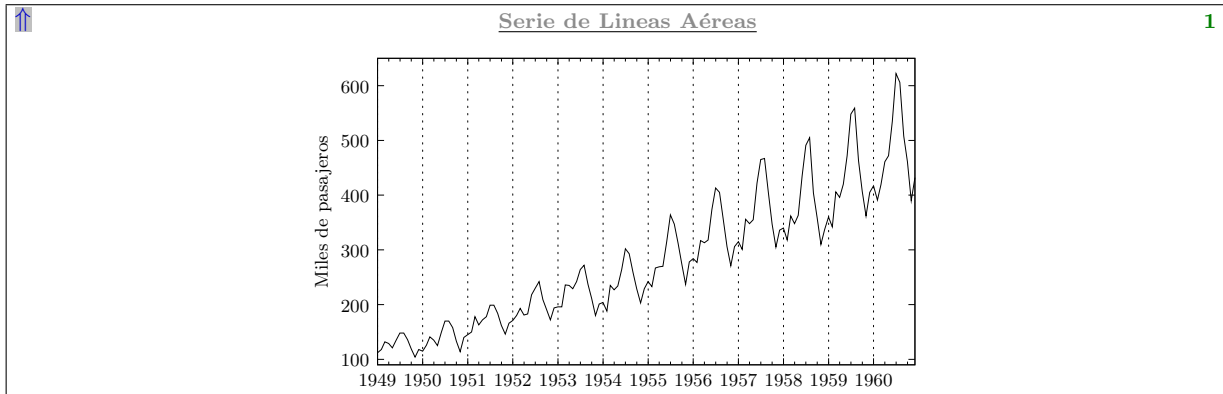
Puede encontrar la última versión de este material en:

<http://www.ucm.es/info/ecocuan/mbb/index.html#LDHR>

## Índice

<b>Índice</b>	<b>1</b>
<b>1. Introducción</b>	<b>2</b>
<b>2. Breve introducción al análisis espectral</b>	<b>3</b>
2.1. Análisis en el ámbito del tiempo: las autocovarianzas . . . . .	3
2.2. Análisis en el ámbito de la frecuencia: el espectro . . . . .	4
2.3. El espectro de una serie bien conocida . . . . .	7
<b>3. Método DHR</b>	<b>8</b>
3.1. Modelo teórico de los componentes DHR . . . . .	8
3.1.1. Tendencia y componente estacional . . . . .	9
3.1.2. Modelo DHR completo . . . . .	10
3.2. Análisis espectral del modelo DHR . . . . .	12
3.3. Estimación de las varianzas (optimización de los NVR) . . . . .	13
3.4. Estimación de los componentes . . . . .	14
<b>A. Identificación del modelo y estimación de los hiper-parámetros: el algoritmo LDHR</b>	<b>17</b>
A.1. Identificación de modelo DHR . . . . .	18
<b>B. Índice de Trasparencias</b>	<b>18</b>
<b>C. Bibliografía</b>	<b>19</b>

1. Introducción

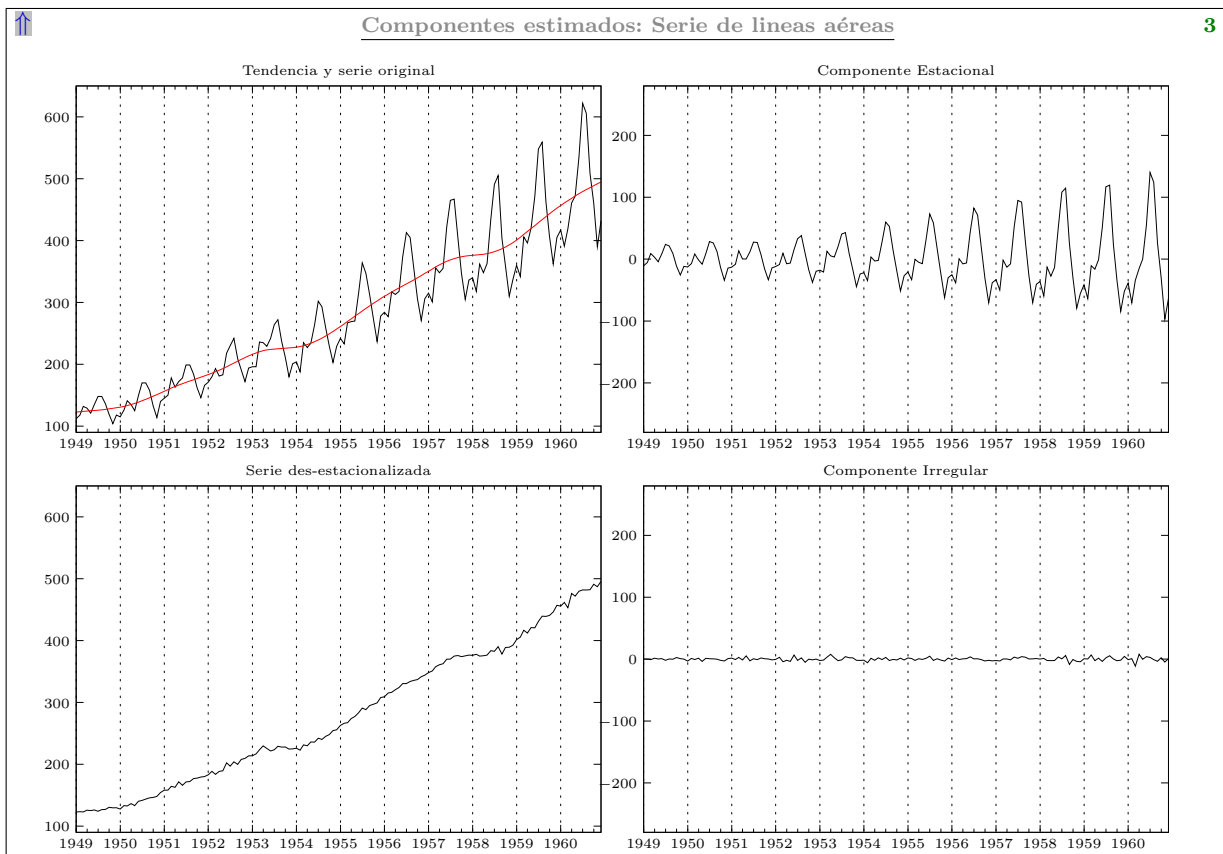


Modelo de componentes no observables 2

**Enfoque ARIMA**

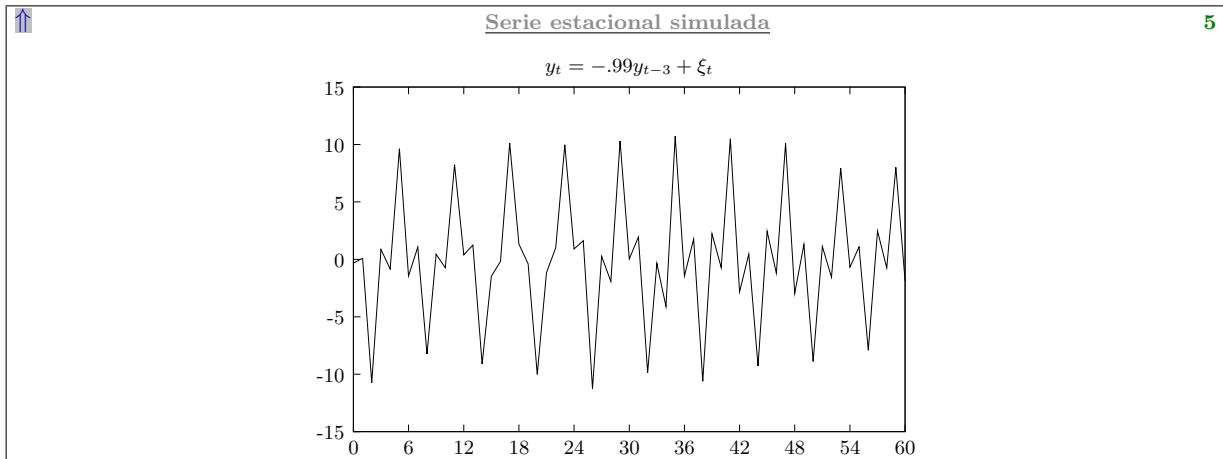
$$\phi(L)y_t = \theta(L)\xi_t, \quad \xi_t \sim \text{r.b.}(0, \sigma^2).$$

**Enfoque de componentes no observables**

$$y_t = T_t + S_t + e_t$$


- Pasos a seguir 4
1. **Identificar** el tipo de modelo para cada componente
  2. **Estimar** los hiper-parámetros de los modelos identificados
  3. “Filtrar” los componentes  
(estimar el valor esperado de cada componente en cada instante)
- Regresión Armónica Dinámica Lineal (LDHR)* se ocupa de los dos primeros pasos.

## 2. Breve introducción al análisis espectral



Para describir una serie temporal nos fijamos fundamentalmente en su nivel y como varía este nivel en relación a cómo varió en el pasado. Sabiendo esto, y conociendo cual ha sido su comportamiento reciente, podemos aventurarnos a pronosticar valores futuros para la serie temporal.

La serie de la figura sigue una pauta muy clara y nos servirá para ilustrar las ideas<sup>1</sup>.

1. El nivel de la serie en  $t$  es aproximadamente igual a menos (-) el nivel de la serie en  $t - 3$ .
2. El nivel de la serie en  $t$  es aproximadamente igual al nivel de la serie en  $t - 6$ .
3. El nivel de la serie en  $t$  es aproximadamente igual a menos (-) el nivel de la serie en  $t - 9$ .
4. ...

Autocovarianza y Función Generadora de Covarianzas 6

**Autocovarianza  $k$ -ésima;** relación entre:  
desviación respecto a la esperanza en  $t$  con desviación que hubo en  $t - k$ .

$$\gamma_k = E\left((y_t - \mu) \cdot (y_{t-k} - \mu)\right)$$

**Función Generadora de Covarianzas:** es la sucesión de autocovarianzas:

$$\Gamma(z) = \gamma_0, \gamma_1, \gamma_2, \dots = \sum_{k=0}^{\infty} \gamma_k z^k.$$

### 2.1. Análisis en el ámbito del tiempo: las autocovarianzas

Las autocovarianzas nos informan precisamente de cómo varía el nivel de la serie en relación a cómo varió en su pasado:

$$\gamma_t(k) = E((y_t - Ey_t)(y_{t-k} - Ey_{t-k})).$$

La sucesión de autocovarianzas para los distintos retardos nos proporciona una descripción de cómo evoluciona la serie. Este tipo de descripción se denomina análisis en el ámbito del tiempo.

Para cada retardo hay una autocovarianza. Llamamos *función generadora de covarianzas* a la sucesión de autocovarianzas y la denotamos como <sup>2</sup>:

$$\Gamma(z) = \gamma_0, \gamma_1, \gamma_2, \dots = \sum_{k=0}^{\infty} \gamma_k z^k$$

donde  $z^k$  en  $\gamma_k z^k$  indica que la autocovarianza  $\gamma_k$  aparece en la posición  $k$ -ésima; es decir,  $z^k$  es un índice de posición. En la expresión anterior la letra griega  $\sum$  no es un sumatorio. La expresión  $\sum b_j z^j$  es mera notación y por tanto puede denotar una sucesión convergente o no (aunque la función generadora de autocovarianzas es siempre convergente, i.e., sumable).

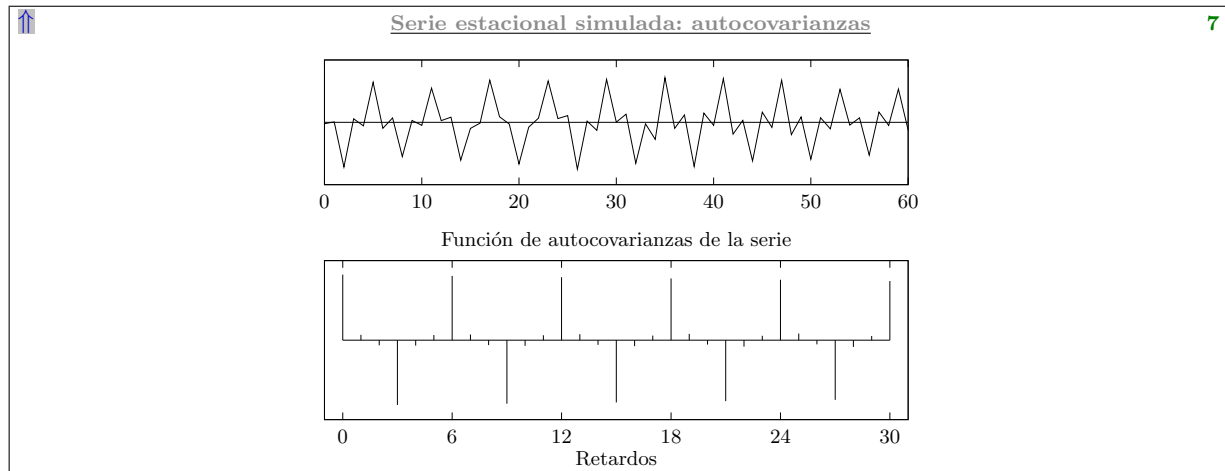
<sup>1</sup>Es una serie simulada que roza la no estacionariedad. Su modelo es  $y_t = -.99y_{t-3} + \xi_t$ . Tiene tres raíces de módulo 0.99666. Una de las raíces es real y negativa (con ciclos asociados de periodo 2), y las otras son un par de raíces complejas conjugadas (con ciclos asociados de periodo 6). La hemos escogido porque su comportamiento es muy claro a simple vista.

<sup>2</sup>En realidad aquí mostramos sólo la mitad de la sucesión para simplificar. La sucesión completa es simétrica respecto a  $\gamma_0$  (es decir  $\gamma_{-j} = \gamma_j$ ):  $\Gamma(z) = \dots, \gamma_{-2}, \gamma_{-1}, \gamma_0, \gamma_1, \gamma_2, \dots = \sum_{k=-\infty}^{\infty} \gamma_k z^k$ .

*Observación.* La transcripción de la siguiente expresión

$$5z^3 + (-7)z^4 + 2z^6 + 1z^7$$

es la sucesión 0, 0, 5, -7, 0, 2, 1, 0, 0, ...



**Ejemplo 1.** Veámoslo con la serie del ejemplo. Hay una señal muy clara para algunos retardos concretos. En este caso son los múltiplos de 3. La sucesión de autocovarianzas claramente nos indica que hay una fuerte relación negativa entre el valor en  $t$  y el valor en  $t - 3$  (y consecuentemente también  $t + 3$ ), y una fuerte relación positiva entre el valor en  $t$  y el valor en  $t - 6$  (ó  $t + 6$ ). Hay otros retardos importantes, pero debe subrayarse que su influencia decrece a medida que el retardo aumenta:

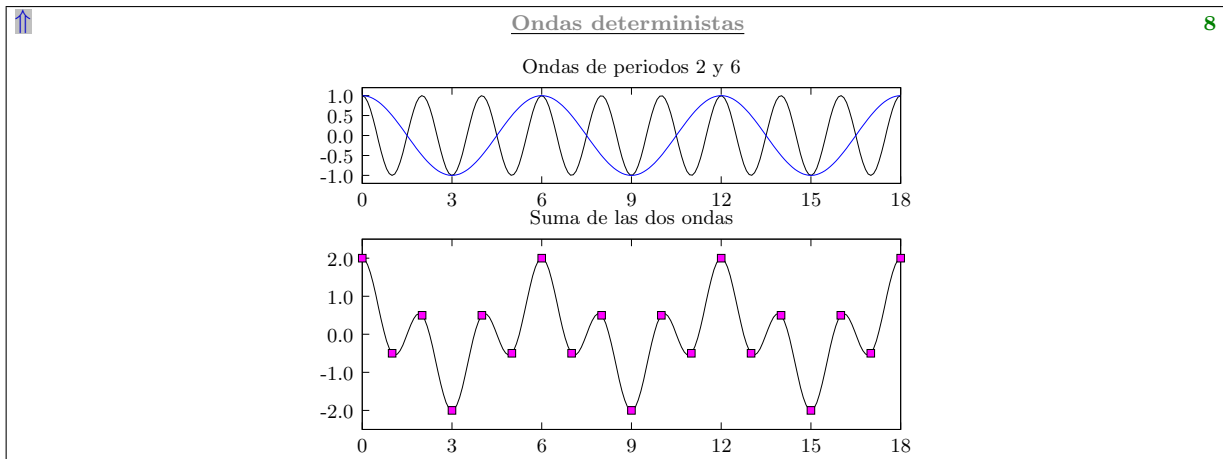
$$\begin{aligned} \widehat{\Gamma}(z) &= \mathbf{45.3}, 3.6, -3.6, \mathbf{-44.9}, -3.8, 3.5, \mathbf{44.4}, 3.9, -3.4, \mathbf{-44.0}, -4.0, 3.4, \mathbf{43.5}, 4.1, \dots \\ &= \mathbf{45.3} \cdot z^0 + 3.6 \cdot z^1 - 3.6 \cdot z^2 - \mathbf{44.9} \cdot z^3 - 3.8 \cdot z^4 + 3.5 \cdot z^5 + \mathbf{44.4} \cdot z^6 + \dots \\ &= \sum_{k=0}^{\infty} \widehat{\gamma}_k z^k \end{aligned}$$

## 2.2. Análisis en el ámbito de la frecuencia: el espectro

Ahora introduciremos un tipo de análisis distinto. Para motivar este nuevo enfoque vamos a imaginar un experimento. Supongamos que tenemos una boya sobre la superficie del mar, y que medimos el nivel de la boya cada  $x$  segundos. Pensemos que la serie del ejemplo anterior (transparencias 0.5 y 0.7) es la sucesión de las mediciones del nivel de la boya.

Podemos describir la variabilidad en el nivel de la boya en función de las olas. En el mar hay simultáneamente olas de diferentes frecuencias (o longitudes). Podemos entender las mareas como una enorme ola que se repite cada varias horas, y que obviamente afectará al nivel de nuestra boya. Si suponemos que el mar está “movido”, habrá grandes olas con una frecuencia de, quizá, 15 ó 20 segundos. Seguramente haya también pequeñas olas con una frecuencia elevada (unos pocos segundos). El nivel de la boya será función de la magnitud y la frecuencia de estas olas.

El análisis espectral trata de descomponer la variabilidad de la serie indicando que parte de la variabilidad es debida a oscilaciones en torno a ciertas frecuencias o periodicidades (en el ejemplo, debida a cada tipo de ola, es decir, mareas, grandes olas, olas menores, etc.).



En la figura superior mostramos dos ondas (deterministas) de periodos 2 y 6, y en la figura inferior la agregación de ambas ondas<sup>3</sup>. Los puntos cuadrados representan las mediciones del nivel en los instantes de tiempo 0, 1, 2, ...

Nótese el parecido entre la disposición de los cuadrados y los datos de la serie del ejemplo (trasparencias 0.5 y 0.7) *sobre todo al comienzo de la 2ª mitad*. Obviamente estas ondas describen un movimiento determinista mientras la serie temporal anterior es estocástica; sin embargo, la similitud en el movimiento parece indicar que la varianza de la serie estacional simulada es debida principalmente a las oscilaciones (*estocásticas*) de periodos 2 y 6 (es decir, a oscilaciones de frecuencias  $2\pi/2$  y  $2\pi/6$  radianes) de manera parecida al comportamiento determinista la figura anterior.

El análisis en el ámbito de la frecuencia, o *análisis espectral*, describe la varianza de un proceso estocástico como la suma de las contribuciones a la varianza debidas a las oscilaciones en cada una de las frecuencias del rango  $[-\pi, \pi]$  ( ó  $[0, \pi]$ ); es decir, oscilaciones de periodos comprendidos entre 2 e infinito.

Esta descripción (como suma de los efectos de ondas superpuestas) se resume en una función, que llamamos *espectro* y que muestra las contribuciones a la varianza de las oscilaciones en cada frecuencia.

Función Generadora de Covarianzas y Espectro 9

$$\Gamma(z) = \sum_{k=0}^{\infty} \gamma_k z^k.$$

Espectro es Transformada de Fourier de  $\Gamma(z)$ , es decir:

*sustituir  $z^k$  por  $\cos(k\omega)$ , e interpretar  $\sum$  como sumatorio:*

$$f(\omega) = \sum_{k=0}^{\infty} \gamma_k \cos(k\omega); \quad -\pi \leq \omega \leq \pi.$$

(suma de funciones coseno)

Nótese que en esta suma ponderada, la importancia de la contribución de las ondas de frecuencia  $k\omega$  viene determinada precisamente por la covarianza  $k$ -ésima  $\gamma_k$ .

Con la función inversa de la transformada de Fourier recuperamos las autocovarianzas

$$\gamma_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\omega) e^{-i\omega k} d\omega.$$

Para el caso especial donde  $k = 0$ , tenemos:

$$\sigma^2 \equiv \gamma_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\omega) d\omega;$$

es decir, *la integral del espectro es igual a la varianza total del proceso (la suma de las contribuciones a la varianza en las distintas frecuencias)*. Por tanto, podemos interpretar el espectro como una descomposición de la varianza del proceso (o de la serie temporal) en las distintas frecuencias.

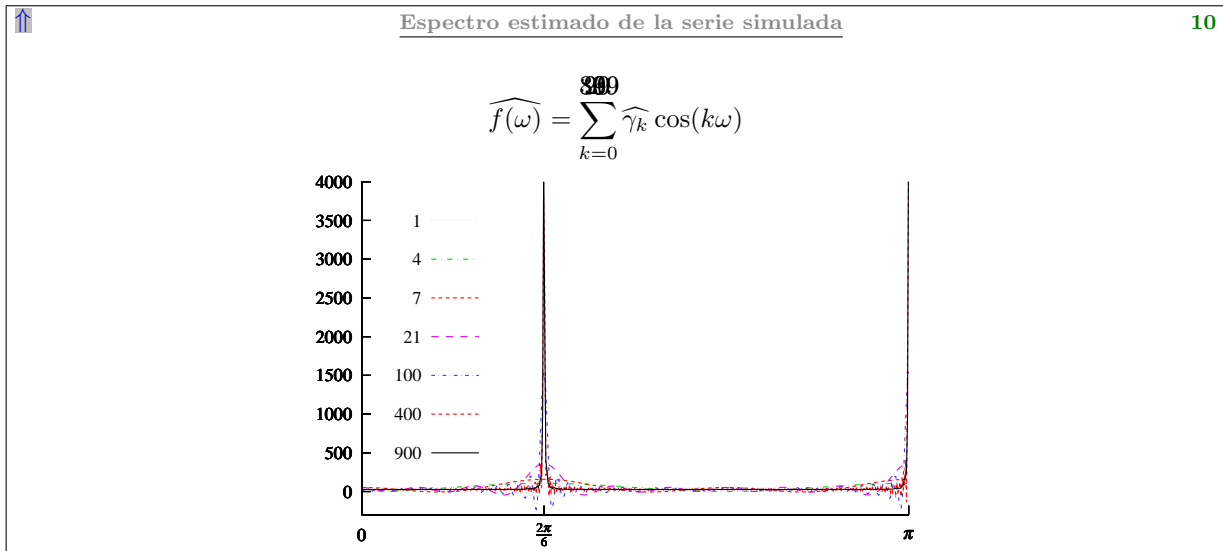
<sup>3</sup>si también hubiera una pequeña oscilación aleatoria, el aspecto sería muy parecido al ejemplo de serie estacional de la transparencia 0.5

**Ejemplo 2.** Para la serie del ejemplo, la estimación del espectro sería

$$\widehat{f(\omega)} = \sum_{k=0}^{\infty} \widehat{\gamma}_k \cos(k\omega)$$

$$= 45.3 + 3.6 \cos(\omega) - 3.6 \cos(2\omega) - 44.9 \cos(3\omega) - 3.8 \cos(4\omega) + 3.5 \cos(5\omega) + 44.4 \cos(6\omega) \dots$$

Si se tiene una serie temporal suficientemente larga, y si se añaden suficientes covarianzas (si el sumatorio incorpora suficientes términos), la función obtenida tendrá un aspecto parecido al de la transparencia 0.11.



El espectro de la serie del ejemplo indica claramente que las contribuciones a la varianza más importantes se encuentran muy concentradas en torno a ciclos (*estocásticos*) de periodos 2 y 6 (es decir, entorno a las frecuencias  $2\pi/2$  y  $2\pi/6$ ). Esto explica la semejanza entre las ondas deterministas y la serie temporal de los ejemplos anteriores.

- Deducir en la pizarra el espectro de un proceso de ruido blanco.
- Comentar la etimología de la expresión “ruido blanco”.
- Explicar que si un proceso no es ruido blanco, se puede inferir su dinámica (extraer información).
- Definir lo que es un filtro (relacionar con un cristal de colores, o con un ecualizador de una cadena musical). Relacionar con los modelos ARMA.

**Resumiendo:** El espectro y la función generadora de covarianzas nos aportan exactamente la misma información, pero expresada de forma diferente. Como ya hemos visto, el espectro se define a partir de la función generadora de covarianzas  $\Gamma(z)$  sustituyendo  $z^k$  por  $\cos(k\omega)$  e interpretando la sucesión como una suma de funciones de  $\omega$ :

$$f(\omega) \equiv \sum_{k=0}^{\infty} \gamma_k \cos(k\omega).$$

Pero lo relevante es la sucesión de números  $\gamma_k$ ;  $k = 0, 1, \dots$ ; que es compartida tanto por la función generadora de covarianzas como por el espectro, lo que explica que ambas funciones nos aporten exactamente la misma información.

**Nota 1 (Relación entre las autocovarianzas y el espectro en un proceso ARMA).** Sea  $y_t$  el proceso estocástico *estacionario* que verifica

$$\phi(L)y_t = \theta(L)\xi_t, \quad \xi_t \sim \text{r.b.}(0, \sigma^2).$$

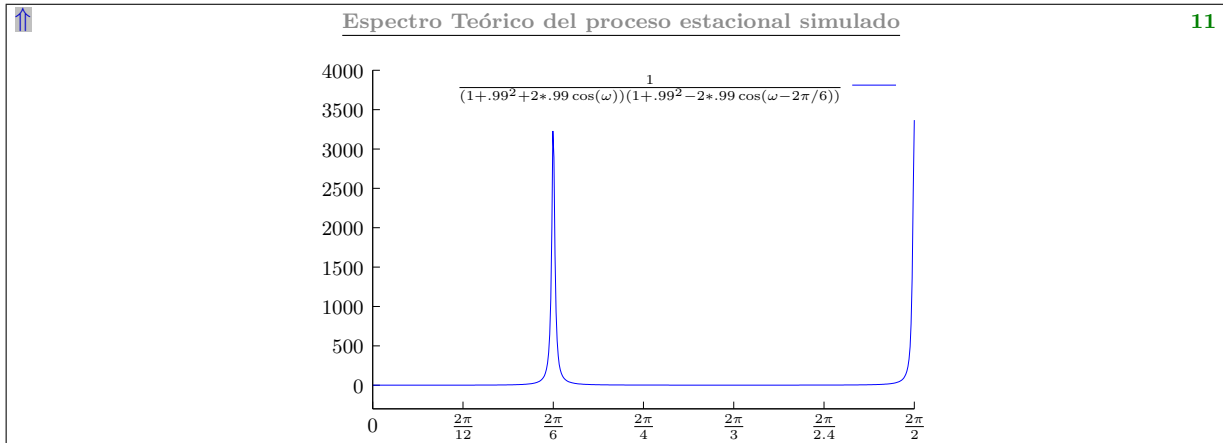
$y_t$  tiene como *función generadora de covarianzas* la siguiente sucesión:

$$\Gamma_y(z) = \sigma_\xi^2 \frac{\theta(z) * \theta(z^{-1})}{\phi(z) * \phi(z^{-1})};$$

y su *espectro* es (sustituyendo  $z$  por  $e^{-i\omega}$ )

$$f_y(\omega) = \mathcal{F}(\Gamma_y(z)) = \sigma_\xi^2 \frac{\theta(e^{-i\omega})\theta(e^{i\omega})}{\phi(e^{-i\omega})\phi(e^{i\omega})};$$

que tiene una expansión como suma de cosenos y productos de cosenos, ya que  $e^{-ik\omega} + e^{ik\omega} = 2 \cos(k\omega)$ .



Por tanto, una forma de estimar el espectro es ajustar un modelo AR, y de la estimación paramétrica

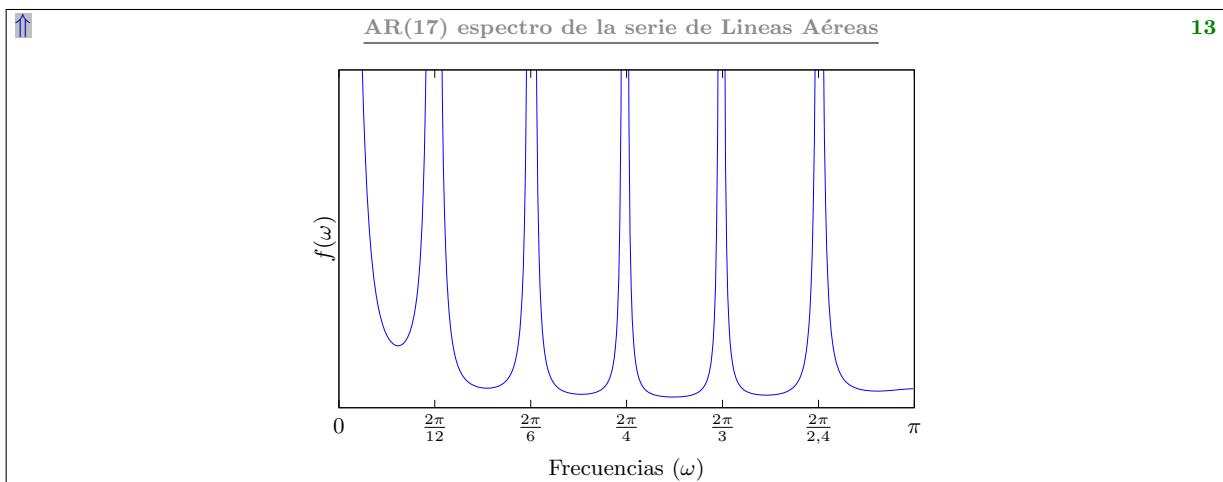
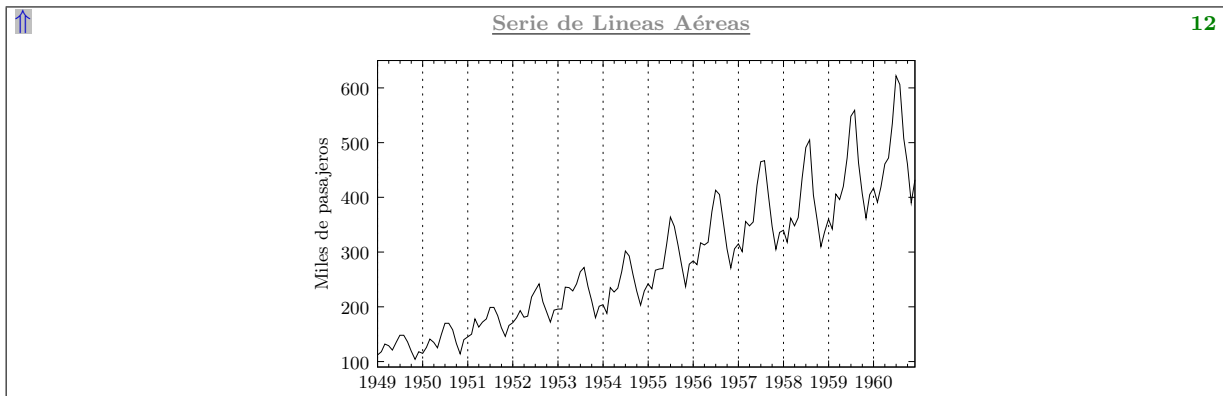
$$\widehat{\phi}(L)y_t = \xi_t, \quad \xi_t \sim \text{r.b.}(0, \widehat{\sigma}_\xi^2),$$

obtener la función

$$\widehat{f}_y(\omega) = \frac{\widehat{\sigma}_\xi^2}{\widehat{\phi}(e^{-i\omega})\widehat{\phi}(e^{i\omega})}.$$

Otra forma es estimar el periodograma (para más información se puede consultar los manuales [Box et al., 1994](#); [Fuller, 1996](#); [Priestley, 1981](#); [Percival y Walden, 1993](#); [Shumway y Stoffer, 2000](#))

### 2.3. El espectro de una serie bien conocida



Nótese que, en la figura anterior, no todos los picos del espectro tienen la misma anchura; lo cual nos indica que las contribuciones a la varianza en distintas frecuencias son diferentes (unas frecuencias (o periodicidades) tienen mayor importancia que otras).

### 3. Método DHR

Los modelos estructurales de series temporales son modelos formulados en términos de algunos componentes de interés. Debido a que estos componentes no son observados directamente, dichos modelos también se denominan “modelos de componentes no observables”.

El método DHR es un algoritmo que descompone las series temporales en varios componentes no observables en función de un modelo estructural que denominamos modelo DHR. La descomposición de cada serie temporal depende de varios hiper-parámetros que han de ser estimados previamente.

Siguiendo Ng y Young (1990), podemos descomponer una serie temporal  $y_t$  en:

$$y_t = T_t + S_t + e_t; \quad t = 0, 1, 2, \dots,$$

donde  $T_t$  es la tendencia,  $S_t$  el componente estacional y  $e_t$  el componente irregular. La *tendencia* ( $T_t$ ) es la señal asociada a las bajas frecuencias, y recoge el comportamiento o evolución a largo plazo. El *componente estacional* ( $S_t$ ) recoge las oscilaciones cíclicas correspondientes al periodo estacional y sus armónicos. El *componente irregular* ( $e_t$ ) recoge todas aquellas oscilaciones no sistemáticas no incluidas en los componentes anteriores. En general supondremos que el componente irregular es un proceso estocástico gaussiano  $\{e_t\}_{t=0}^{\infty}$  de ruido blanco con media cero y con varianza  $\sigma_e^2$  (es decir,  $\{e_t\} \sim \text{r.b. } N(0, \sigma_e^2)$ ).

↑
Modelo básico
14

$$y_t = T_t + S_t + e_t; \quad t = 0, 1, 2, \dots, \tag{3.1}$$

- *Tendencia* (o ciclo-tendencia): ( $T_t$ )
- *Componente estacional*: ( $S_t$ )
- *Componente irregular*: ( $e_t$ );  $\{e_t\} \sim \text{r.b. } N(0, \sigma_e^2)$

#### 3.1. Modelo teórico de los componentes DHR

Los componentes de *Tendencia* (o ciclo-tendencia  $T_t$ ) y *Componente estacional* ( $S_t$ ) están, a su vez, generados por un agregado de sub-componentes DHR que denotamos individualmente por  $s_t^{p_j}$ .

Cada componente DHR,  $s_t^{p_j}$ , recoge las oscilaciones de  $y_t$  en torno a la frecuencia  $\omega_j$  (es decir, oscilaciones de periodo  $p_j$ ) según el modelo

$$s_t^{p_j} = a_{j_t} \cos(\omega_j t) + b_{j_t} \sin(\omega_j t),$$

donde  $\omega_j$  es la frecuencia en ciclos por unidad de tiempo de cada componente, y  $p_j \equiv 2\pi/\omega_j$  indica el periodo correspondiente<sup>4</sup>. Las series  $a_{j_t}$  y  $b_{j_t}$  recogen la variación en la amplitud de las oscilaciones a lo largo del tiempo; y son realizaciones de dos procesos estocásticos.

Dichos procesos estocásticos  $\{a_j\}$  y  $\{b_j\}$ , que modulan la amplitud de las oscilaciones, siguen modelos AR(1) o AR(2) (generalmente no estacionarios) idénticos e incorrelados entre sí:

$$\left. \begin{aligned} (1 - \alpha_j B)(1 - \beta_j B)a_{j_t} &= \xi_{j_{t-1}} \\ (1 - \alpha_j B)(1 - \beta_j B)b_{j_t} &= \xi_{j_{t-1}} \end{aligned} \right\} 0 \leq \alpha_j, \beta_j \leq 1, \{\xi_j\} \sim \text{w.n. } N(0, \sigma_j^2);$$

El modelo presenta los siguientes casos particulares (Young, 1984):

- Proceso autorregresivo (AR):  $\alpha < 1; \quad \beta < 1.$
- Paseo aleatorio simple (RW):  $\alpha = 0; \quad \beta = 1.$
- Paseo aleatorio suavizado (SRW):  $0 < \alpha < 1; \quad \beta = 1.$
- Paseo aleatorio integrado (IRW):  $\alpha = 1; \quad \beta = 1.$

Éstos son los casos que emplearemos para describir la evolución de los parámetros  $a_{j_t}$  y  $b_{j_t}$ . Estos modelos pertenecen a la familia de *paseos aleatorios generalizados*. De este modo se pueden describir series que son no estacionarias estacionalmente (véase Young, 1994).

Cada componente DHR  $s_t^{p_j}$  tiene una representación ARMA equivalente (ver tabla 1 en la página 11):

$$\varphi_j(B)s_t^{p_j} = \theta_j(B)\xi_{j_t}.$$

<sup>4</sup>para una serie mensual como la de “Líneas Aéreas”,  $P_j = \infty, 12, 6, 4, 3, 2.4, 2$  y  $\omega_j = 0, \frac{\pi}{12}, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}, \frac{\pi}{2.4}, \frac{\pi}{2}$ ,

↑
Modelo teórico de los sub-componentes DHR
15

$$s_t^{p_j} = a_{j_t} \cos(\omega_j t) + b_{j_t} \sin(\omega_j t), \quad (3.2)$$

donde  $\omega_j = \frac{2\pi}{p_j}$  es la frecuencia y  $p_j$  es el periodo.  
 $a_{j_t}$  y  $b_{j_t}$  modulan amplitud oscilaciones  
 (procesos AR(1) o AR(2) no estacionarios con parámetros conocidos, excepto varianzas).

### 3.1.1. Tendencia y componente estacional

Los componentes DHR se agregan en función de las frecuencias  $\omega_j$  para formar los componentes  $T_t$  y  $S_t$ .

La tendencia ( $T_t$ ) es el componente DHR que recoge las oscilaciones en torno a la frecuencia cero (periodo infinito), y el componente estacional ( $S_t$ ) es la agregación de los componentes DHR asociados a las frecuencias estacionales (asociadas a los periodos 12, 6, 4, 3, 2.4, y 2 en el caso de series mensuales; o periodos 4 y 2 en el caso de series trimestrales).

**Tendencia** La tendencia está constituida por el componente DHR  $s_t^\infty$ , que corresponde a la frecuencia  $\omega_0 = 0$ . Por lo tanto,  $T_t \equiv s_t^\infty$ , y la Ecuación (3.2) pasa a ser:

$$T_t = a_{0t} \cos(0 \cdot t) + b_{0t} \sin(0 \cdot t) = a_{0t};$$

es decir, sigue el mismo proceso estocástico que  $\{a_0\}$ . Así pues,  $T_t$  es un proceso de la forma

$$(1 - \alpha L)(1 - \beta L)T_t = \xi_{0t} \quad 0 \leq \alpha \leq 1, \beta = 1,$$

donde  $\xi_{0t} \sim \text{r.b.}(0, \sigma_0^2)$  es un proceso no correlado con el componente irregular  $e_t$  de la Ecuación (3.1) ni con el resto de componentes.

↑
Modelo de la Tendencia
16

$$T_t = a_{0t} \cos(0 \cdot t) + b_{0t} \sin(0 \cdot t) = a_{0t}; \quad (3.3)$$

$$(1 - \alpha L)(1 - L)T_t = \xi_{0t-1}.$$

El gráfico muestra una curva azul que comienza en un punto muy alto y estrecho en el eje vertical (representando una alta densidad de probabilidad) cerca de omega=0. A medida que omega aumenta, la curva se aplana y se extiende horizontalmente, acercándose al eje horizontal (representando una densidad de probabilidad cercana a cero). El eje horizontal está etiquetado como 'Frecuencias (omega)' y tiene marcas en 0 y pi.

**Componente estacional completo** El componente estacional  $S_t$  está constituido por la suma de los componentes DHR correspondientes a la frecuencia estacional y sus armónicos fundamentales:

$$S_t = \sum_{j=1}^R s_t^{p_j} = \sum_{j=1}^R [a_{j_t} \cos(\omega_j t) + b_{j_t} \sin(\omega_j t)], \quad (3.4)$$

donde  $R$  representa el número de armónicos fundamentales de la frecuencia estacional.

Desde ahora nos referimos a los sumandos de (3.4) como componentes estacionales, y llamamos a  $S_t$  componente estacional *completo*.

Llamamos *componentes estacionales intermedios* al conjunto de componentes estacionales de periodo distinto a 2. Bujosa et al. (2006) demuestran que los componentes estacionales intermedios DHR tienen una representación ARMA equivalente. En concreto, son modelos ARMA(2,1) cuando los parámetros  $a_j$

y  $b_j$  siguen un proceso RW, o ARMA(4,2) cuando siguen un proceso del tipo SRW o IRW<sup>5</sup>.

↑
Modelo para cada componente estacional DHR
17

$$s_t^{pj} = a_{jt} \cos(\omega_j t) + b_{jt} \sin(\omega_j t) \quad (3.5)$$

**Componente Nyquist** El componente Nyquist,  $s^2$ , es el componente estacional de periodo 2 y está asociado a la frecuencia  $\omega_R = \pi$ . Sustituyendo  $\omega_j = \pi$  en la Ecuación (3.2) tenemos:

$$s_t^2 = a_{Rt} \cos(\pi \cdot t) + b_{Rt} \sin(\pi \cdot t) = a_{Rt} \cdot (-1)^t.$$

Esto equivale a un proceso de tipo

$$(1 + \alpha L)(1 + \beta L)s_t^2 = \xi_{Rt} \quad 0 \leq \alpha \leq 1, \beta = 1.$$

Así pues, el modelo del componente Nyquist es análogo al modelo de la tendencia, pero con raíces reales negativas.

Un resumen de los modelos ARMA de los componentes DHR aparece en la Tabla 1 en la página siguiente.

### 3.1.2. Modelo DHR completo

Si sustituimos  $T_t$  y  $S_T$  en la Ecuación (3.1) por las expresiones (3.3), y (3.4), obtenemos

$$y_t = \sum_{j=0}^R [a_{jt} \cos(\omega_j t) + b_{jt} \sin(\omega_j t)] + e_t,$$

donde  $j = 0$  corresponde a la tendencia y  $j = R$  corresponde al componente Nyquist.

↑
Modelo DHR completo
18

$$y_t = \sum_{j=0}^R [a_{jt} \cos(\omega_j t) + b_{jt} \sin(\omega_j t)] + e_t, \quad (3.6)$$

donde  $j = 0$  corresponde a la tendencia.

Parámetros

- Conocidos:  $\omega_j$
- “Supuestamente conocido”: modelo AR de procesos  $\{a_j\}$  y  $\{b_j\}$
- **Desconocidos:** Varianzas innovaciones de los procesos  $\{a_j\}$  y  $\{b_j\}$ .

En la transparencia 0.21 se muestran los pseudo-espectros de los componentes de un modelo DHR completo, con una tendencia IRW y seis componentes estacionales RW centrados en las frecuencias correspondientes a los periodos  $12/j$ ,  $j = 1, 2, \dots, 6$ .

<sup>5</sup>Los componentes de tipo RW en la frecuencia  $\pi/2$  son AR(2) puros; y los componentes de tipo SRW e IRW en las frecuencias  $\pi/4$  y  $3\pi/4$  tienen una raíz cero en la parte MA (para más detalles véase Bujosa, 2000)

Modelo GRW	Tendencia $\omega_0 = 0$	Componentes estacionales $0 < \omega_j < \pi$	Componente Nyquist $\omega_j = \pi$
Modelo general ( $0 \leq \alpha, \beta \leq 1$ )	$(1 - (\alpha_0 + \beta_0)L + \alpha_0\beta_0L^2)T_t = \xi_{0t-1}$	$(\phi_j^\alpha(L) * \phi_j^\beta(L))s^{pj}_t = \left( \frac{\alpha_j\beta_j \cos(2\omega_j)}{\gamma_j^* \eta_j^*} \right) (1 - \theta_j^1 L - \theta_j^2 L^2) \xi_{jt-1}$	$(1 + (\alpha_R + \beta_R)L + \alpha_R\beta_R L^2)s_t^2 = \xi_{Rt-1}$
RW ( $\alpha=0, \beta=1$ )	$(1 - L)T_t = \xi_{0t-1}$	$\phi_j^\beta(L)s_t^{pj} = \frac{1}{\sqrt{1 + \sin(\omega_j)}} \left( 1 - \frac{\cos(\omega_j)}{1 + \sin(\omega_j)} L \right) \xi_{jt-1}$	$(1 + L)s_t^2 = \xi_{Rt-1}$
SRW ( $0 < \alpha < 1, \beta=1$ )	$(1 - (1 + \alpha_0)L + \alpha_0L^2)T_t = \xi_{0t-1}$	$(\phi_j^\alpha(L) * \phi_j^\beta(L))s^{pj}_t = \left( \frac{\alpha_j \cos(2\omega_j)}{\gamma_j^* \eta_j^*} \right) (1 - \theta_j^1 L - \theta_j^2 L^2) \xi_{jt-1}$	$(1 + (1 + \alpha_R)L + \alpha_R L^2)s_t^2 = \xi_{Rt-1}$
IRW ( $\alpha=\beta=1$ )	$(1 - 2L + L^2)T_t = \xi_{0t-1}$	$(\phi_j^\alpha(L) * \phi_j^\beta(L))s^{pj}_t = \left( \frac{\cos(2\omega_j)}{\gamma_j^* \eta_j^*} \right) (1 - \theta_j^1 L - \theta_j^2 L^2) \xi_{jt-1}$	$(1 + 2L + L^2)s_t^2 = \xi_{Rt-1}$

$$\begin{aligned} \phi_j^\alpha(L) &= [1 - \alpha e^{i\omega_j} L] * [1 - \alpha e^{-i\omega_j} L] = [1 - 2\alpha \cos(\omega_j)L + \alpha^2 L^2]; \\ \phi_j^\beta(L) &= [1 - \beta e^{i\omega_j} L] * [1 - \beta e^{-i\omega_j} L] = [1 - 2\beta \cos(\omega_j)L + \beta^2 L^2]; \\ \theta_j^1 &= \gamma_j^* + \eta_j^*; \theta_j^2 = -\gamma_j^* \eta_j^* \text{ (para más detalles véase Bujosa et al., 2006)}. \end{aligned}$$

Cuadro 1: Especificación ARMA de los componentes DHR

### 3.2. Análisis espectral del modelo DHR

El método de estimación de los hiper-parámetros  $NVR_j$  se lleva a cabo en el dominio de la frecuencia, y emplea las expresiones analíticas de los pseudo-espectros de los modelos de componentes DHR y del ar-espectro de la serie. Por este motivo realizamos a continuación un análisis espectral del modelo DHR.

**Pseudo-espectro de paseos aleatorios generalizados** Dado un paseo aleatorio suavizado:

$$(1 - \alpha L)(1 - L)n_t = \xi_t$$

su pseudo-espectro es:

$$\begin{aligned} f_n(\omega) &= \mathcal{FE}(\Lambda_n(z)) \\ &= \mathcal{FE}\left(\frac{\sigma_\xi^2}{(1-\alpha z)(1-z)(1-\alpha z^{-1})(1-z^{-1})}\right) = \frac{\sigma_\xi^2}{(1+\alpha^2-2\alpha\cos(\omega))(2-2\cos(\omega))}, \end{aligned} \quad (3.7)$$

donde  $\Lambda_n(z)$  es la función generadora de pseudo-covarianzas de  $n_t$ , y  $\mathcal{FE}$  es la transformada de Fourier extendida (véase Bujosa et al., 2002).

**Pseudo-espectro del modelo DHR** Puesto que los componentes DHR son modelos de la forma:

$$s_t^{p_j} = a_t \cos(\omega_j t) + b_t \sin(\omega_j t), \quad \omega_j \in (0, \pi);$$

sus pseudo-espectros son

$$f_{s^{p_j}}(\omega) = \frac{1}{2}[f_a(\omega - \omega_j) + f_a(\omega + \omega_j)], \quad \omega_j \in (0, \pi)$$

(para más detalles véase Bujosa, 2000); donde  $f_a(\omega)$  es el pseudo-espectro del proceso  $\{a_t\}$  que modula la amplitud del componente  $s^{p_j}$ .

Así pues, el pseudo-espectro del modelo DHR completo se puede expresar del siguiente modo:

$$f_{dhr}(\omega, \sigma^2) = \sum_{j=0}^R \sigma_{\omega_j}^2 f_{s^{p_j}}(\omega) + \sigma_e^2, \quad (3.8)$$

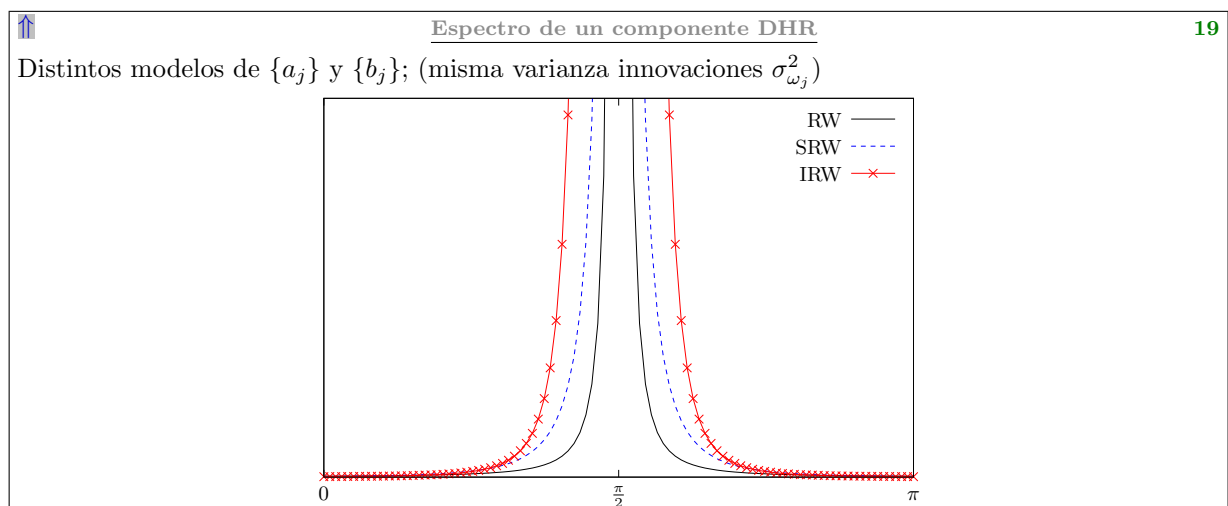
donde

$$\sigma^2 = [\sigma_e^2, \sigma_{\omega_0}^2, \sigma_{\omega_1}^2, \dots, \sigma_{\omega_R}^2]' = [\sigma_e^2, \sigma_{dhr}^2]';$$

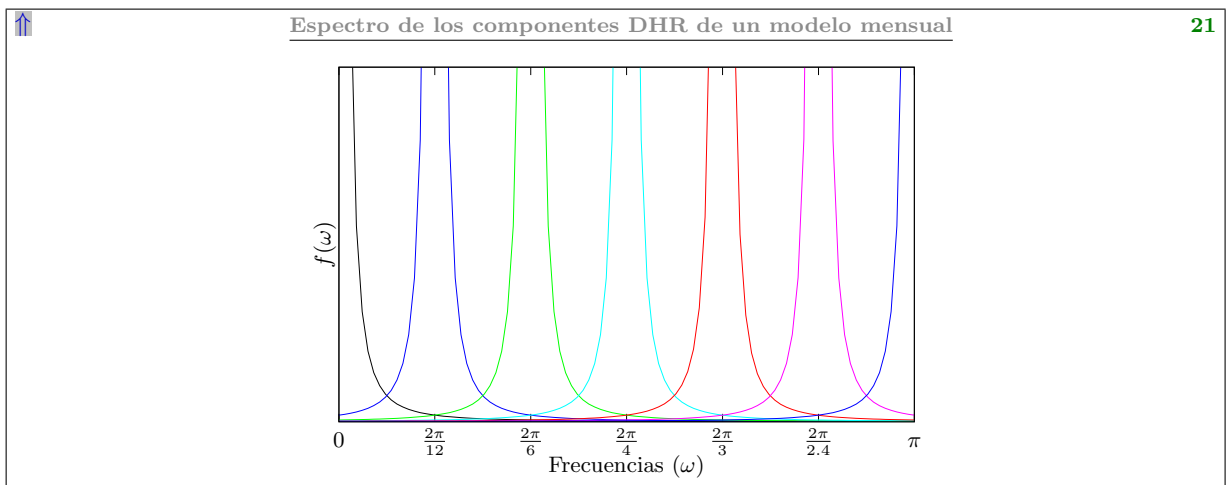
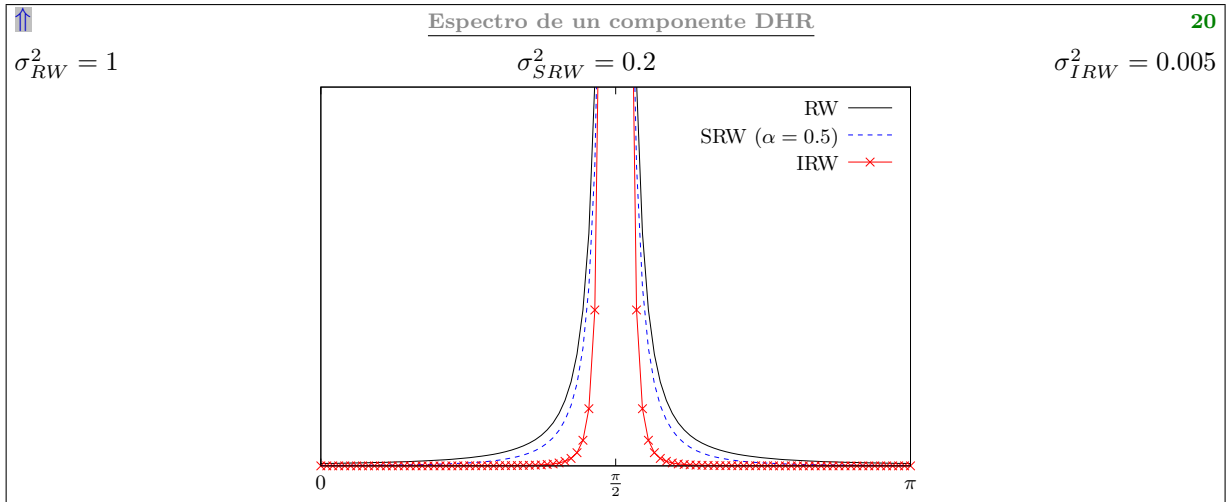
y donde  $\sigma_e^2$  es la varianza del componente irregular.

En la transparencia 0.20 mostramos los pseudo-espectros de tres modelos de componente estacional centrados en la frecuencia  $\omega_j = \pi/2$  para los casos: paseo aleatorio (RW), paseo aleatorio suavizado (SRW,  $\alpha = 1/2$ ) y paseo aleatorio integrado (IRW)<sup>6</sup>.

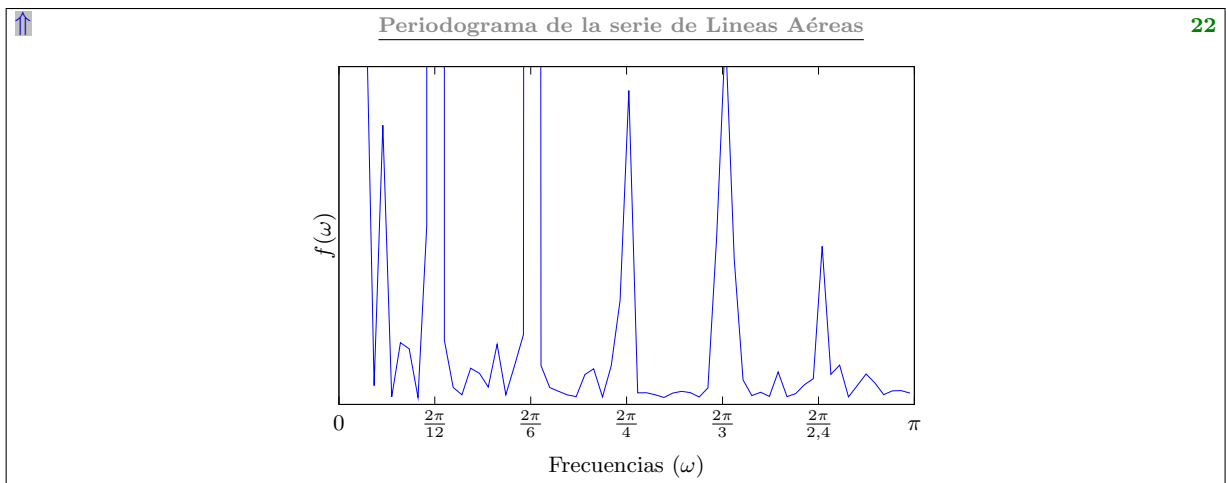
Como se puede apreciar en la figura, la elección del tipo de paseo aleatorio incide en la tasa con la que decrece la importancia de las frecuencias adyacentes a la principal. Un modelo IRW decrece a una tasa mayor que un modelo SRW o RW, por lo que el corte —entre las frecuencias que forman parte del componente y las que no— es más abrupto.

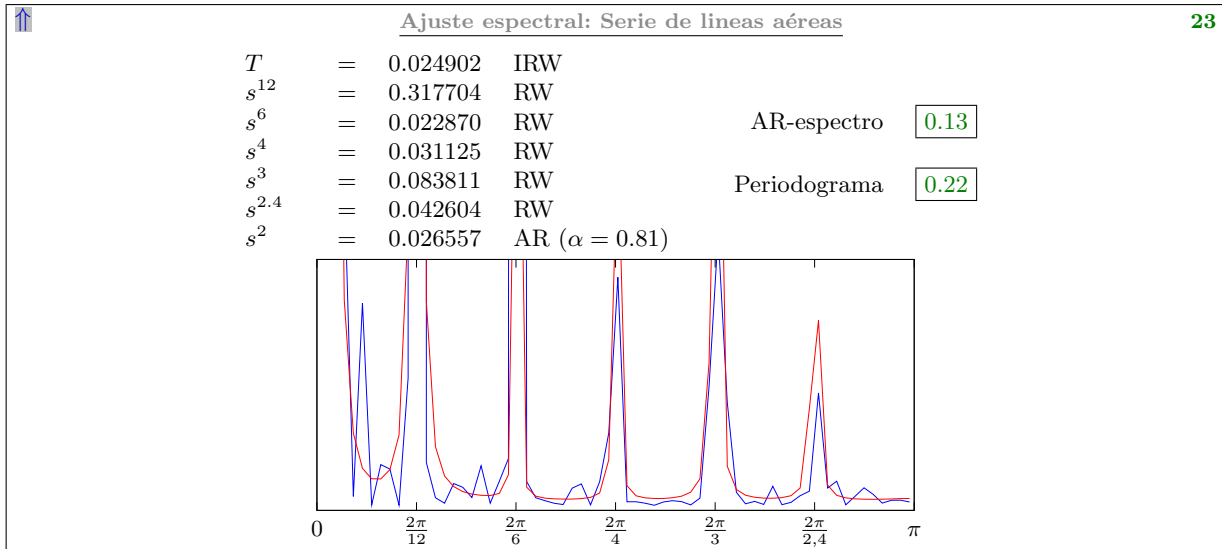


<sup>6</sup>En la figura sólo representamos el intervalo  $[0, \pi]$ , debido a que las funciones son simétricas respecto a la frecuencia 0.



**3.3. Estimación de las varianzas (optimización de los NVR)**





↑
Estimación de los hiper-parámetros
24

$$\min_{[\sigma^2] \in \mathbb{R}^{R+1}} \|f_y(\omega) - f_{dhr}(\omega, \sigma^2)\|,$$

donde

$$f_{dhr}(\omega, \sigma^2) = \sum_{j=1}^R \sigma_j^2 f_{s^{p_j}}(\omega) + \sigma_e^2.$$

y  $\sigma^2 = [\sigma_e^2, \sigma_{\omega_0}^2, \sigma_{\omega_1}^2, \dots, \sigma_{\omega_R}^2]'$

### 3.4. Estimación de los componentes

El algoritmo de filtrado propuesto en [Young et al. \(1999\)](#) para la estimación de los componentes DHR es el suavizado de intervalo fijo, y consta de dos fases: el filtrado hacia delante, que calcula la esperanza de cada componente en el instante  $t$  dada la información hasta  $t - 1$ ; y el suavizado hacia atrás, que nos proporciona la esperanza en el instante  $t$  dada la información de toda la muestra.

Los filtros *pasa banda* empleados están centrados en las frecuencias  $\omega_j$  correspondientes a cada uno de los componentes  $s^{p_j}$ . El parámetro  $\alpha_j$  determina la tasa de decaimiento del filtro alrededor de la frecuencia  $\omega_j$  (véase [Young et al., 1999](#)).

Los componentes estimados son *cuasi-ortogonales*, de manera que comparten muy poca información en común ([del Hoyo, 1994](#)).

**El vector de hiper-parámetros NVR** Una vez fijados los valores  $\omega_j$  y  $\alpha_j$ , lo único desconocido son los cocientes entre varianzas  $\sigma_{\omega_j}^2 / \sigma_e^2$ . Denominamos a estos cocientes hiper-parámetros  $NVR_j$ <sup>7</sup>, y llamamos vector de hiper-parámetros **NVR** a

$$\mathbf{NVR} = [NVR_0, NVR_1, \dots, NVR_R]'$$

donde  $NVR_j = \sigma_{\omega_j}^2 / \sigma_e^2$ ,  $j = 0, 1, \dots, R$ .

Así pues, para poder filtrar la serie  $y_t$ , necesitamos una estimación del vector de varianzas del modelo completo

$$\sigma^2 = [\sigma_e^2, \sigma_{\omega_0}^2, \sigma_{\omega_1}^2, \dots, \sigma_{\omega_R}^2]' \tag{3.9}$$

o una estimación del vector de hiper-parámetros **NVR**.

Dado el modelo DHR completo (número de componentes y valores  $\omega_j$  y  $\alpha_j$  para cada uno de ellos), el vector **NVR** determina el ancho de banda del filtro, es decir, la cantidad de varianza que recoge cada componente.

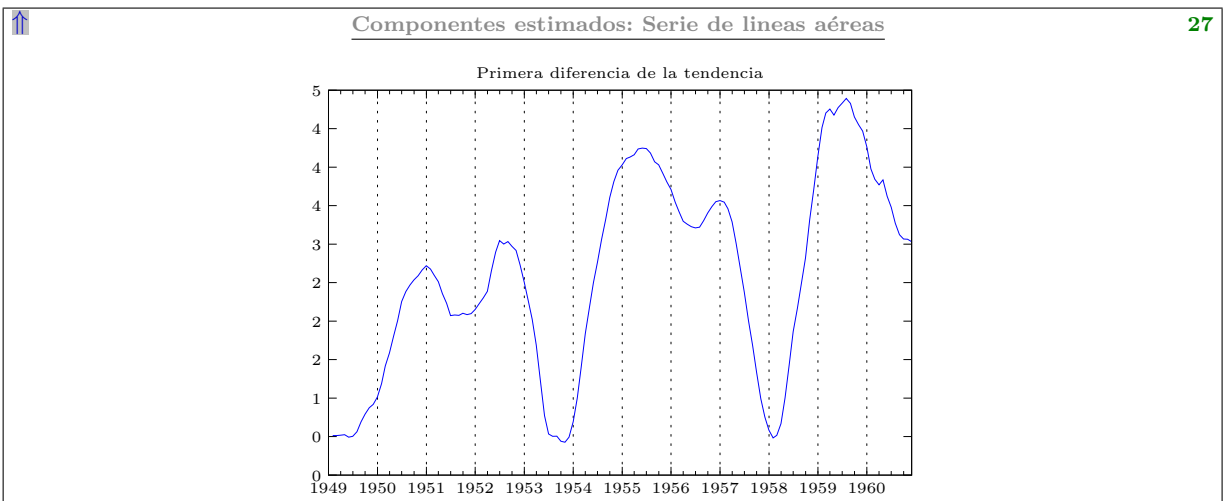
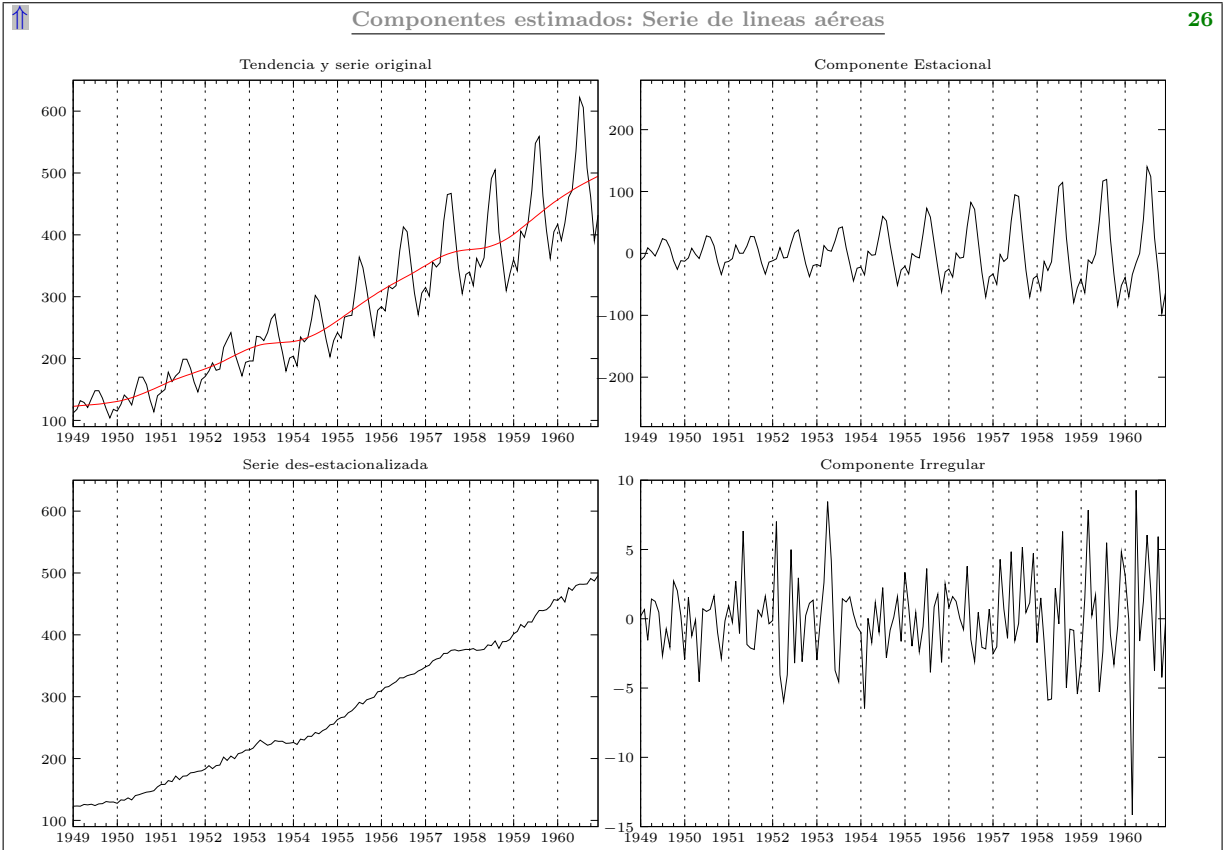
<sup>7</sup>del inglés “Noise Variance Ratio”.

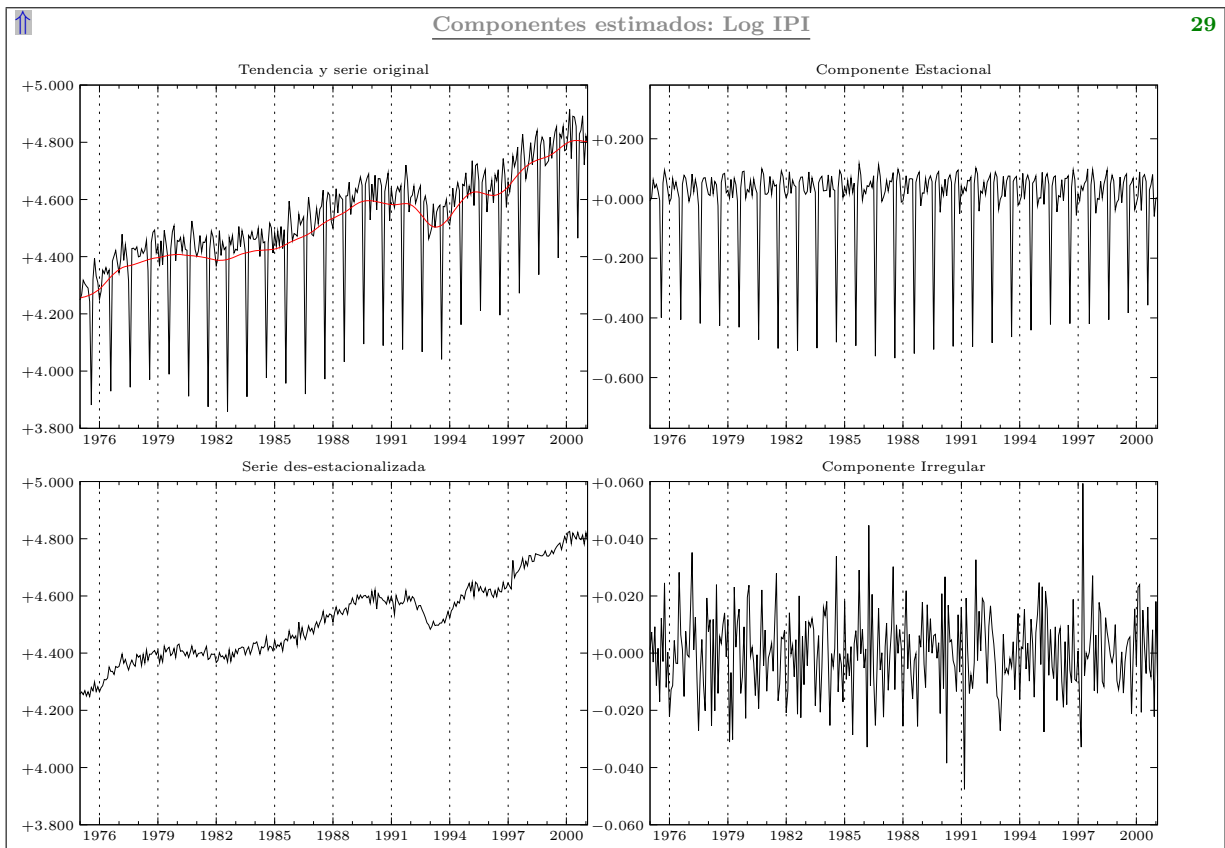
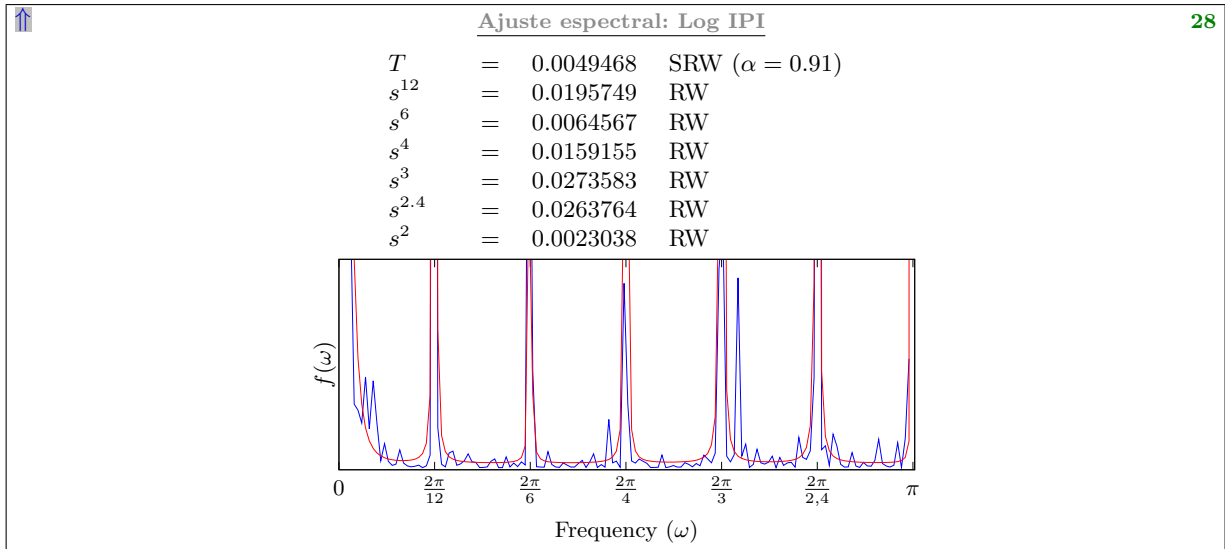
Estimación de los componentes  $s_t^{p_j}$

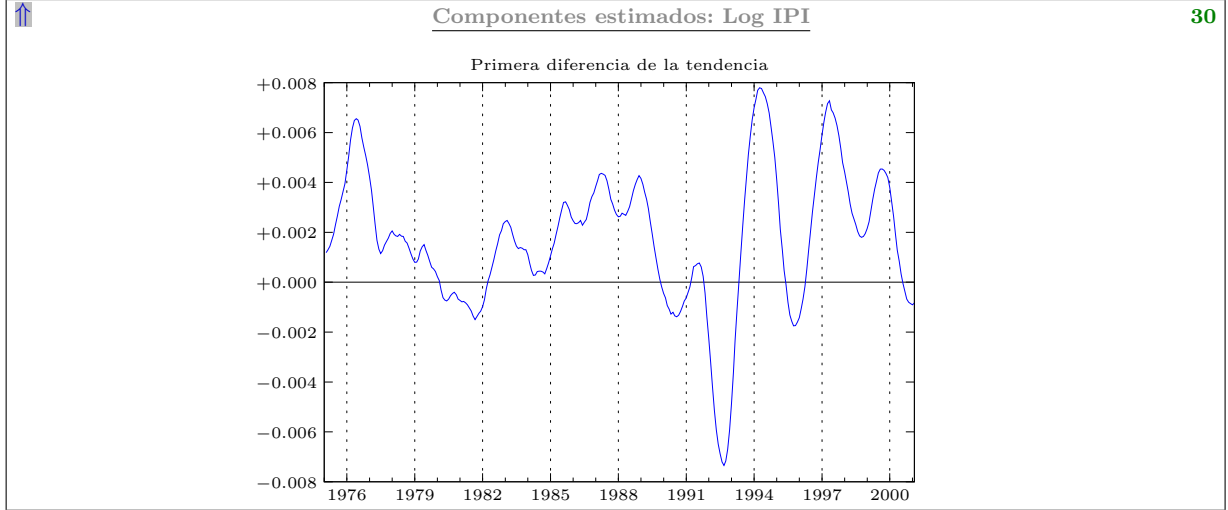
Una vez estimadas las varianzas desconocidas (*hiper-parámetros*)

- Formulación del modelo en Espacio de los Estados
- Filtro de Kalman (suavizado de intervalo fijo)

$$NVR_j = \frac{\sigma_{\omega_j}^2}{\sigma_e^2}$$







### A. Identificación del modelo y estimación de los hiper-parámetros: el algoritmo LDHR

El algoritmo LDHR<sup>8</sup> propuesto por Bujosa et al. (2006) identifica el modelo DHR y estima sus hiper-parámetros. Este algoritmo se basa en la caracterización ARMA de los componentes DHR. Los hiper-parámetros del modelo son las varianzas del vector  $\sigma^2 = [\sigma_e^2, \sigma_{\omega_0}^2, \sigma_{\omega_1}^2, \dots, \sigma_{\omega_R}^2]'$  de la Ecuación (3.8).

**Estimación de los hiper-parámetros** En el algoritmo LDHR la estimación de dichos hiper-parámetros se realiza por mínimos cuadrados ordinarios (MCO). El problema de optimización consiste en minimizar la distancia entre el espectro del modelo DHR,  $f_{dhr}(\omega, \sigma^2)$ , y el espectro de la serie a descomponer,  $f_y(\omega)$ , es decir,

$$\min_{[\sigma^2] \in \mathbb{R}^{R+1}} \|f_y(\omega) - f_{dhr}(\omega, \sigma^2)\|,$$

donde  $f_y(\omega)$  se sustituye por el ar-espectro estimado para la serie  $y_t$ , y  $f_{dhr}(\omega, \sigma^2)$  es la suma de los pseudo-espectros de los modelos de los componentes DHR más el espectro del componente irregular (Ecuación (3.8)). Previamente, es necesaria una transformación del problema de optimización que elimine los polos de los pseudo-espectros. Para lograrlo nos basamos en la caracterización ARMA de los componentes DHR. Esta transformación consiste en multiplicar las funciones  $f_y(\omega)$  y  $f_{dhr}(\omega, \sigma^2)$  por la transformada de Fourier del polinomio  $\Psi(\omega)$ , que contiene todas las raíces de módulo uno de las partes autorregresivas de los componentes DHR; de esta manera el problema de optimización resulta ser:

$$\min_{\sigma^2 \in \mathbb{R}^{R+2}} \|\Psi(\omega) \cdot [f_y(\omega) - f_{dhr}(\omega, \sigma^2)]\|,$$

donde los picos espectrales de los componentes DHR están alineados con los picos espectrales del ar-espectro estimado asociados a las frecuencias estacionales y a la tendencia.

Como refinamiento a este procedimiento, el algoritmo LDHR realiza la estimación de  $\sigma^2$  en dos etapas. En la primera se estiman por MCO las  $R+1$  varianzas  $\sigma_{dhr}^2 = [\sigma_{\omega_0}^2, \sigma_{\omega_1}^2, \dots, \sigma_{\omega_R}^2]'$ , añadiendo componentes adicionales que ajusten en la primera regresión los picos espurios del ar-espectro que no están asociados a los componentes DHR:

$$\min_{[\sigma_{dhr}^2, \sigma_{ac}^2] \in \mathbb{R}^{k+1}} \left\| \Psi(\omega) \cdot \left[ f_y(\omega) - \sum_{j=0}^R \sigma_j^2 S_j(\omega) - f_{ac}(\omega, \sigma_{ac}^2) \right] \right\|,$$

donde

$$f_{ac}(\omega, \sigma_{ac}^2) = \sum_{h=R+1}^k \sigma_h^2 S_h(\omega), \quad \sigma_{ac}^2 = [\sigma_{R+1}^2, \sigma_{R+2}^2, \dots, \sigma_k^2];$$

con ello se evita la influencia de dichos picos espurios en la estimación de las  $R+1$  varianzas asociadas a los componentes DHR. En la segunda etapa se estima la varianza del componente irregular  $\sigma_e^2$  minimizando

$$\min_{\sigma_e^2 \in \mathbb{R}} \left\| \Psi(\omega) \cdot \left[ f_y(\omega) - \sum_{j=0}^R \hat{\sigma}_j^2 S_j(\omega) - \sigma_e^2 \right] \right\|$$

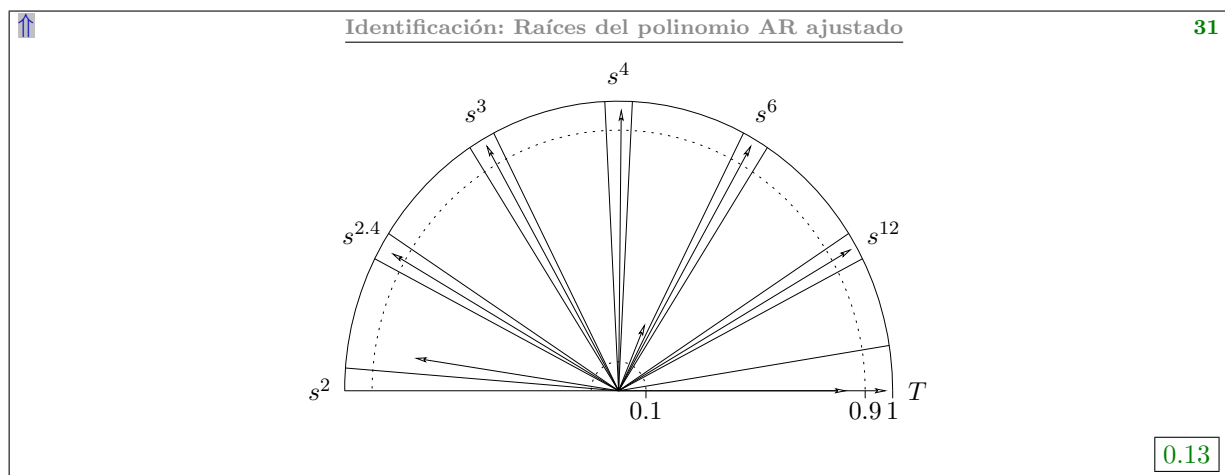
por MCO, y donde  $\hat{\sigma}_j^2$  son las  $R+1$  varianzas estimadas en la primera etapa.

<sup>8</sup>Linear Dynamic Harmonic Regression (Regresión armónica dinámica lineal)

### A.1. Identificación de modelo DHR

Debido a que la estimación de las varianzas se realiza por MCO sin restricciones, la bondad de las estimaciones depende en gran medida de una correcta identificación del modelo DHR a estimar. Por ello el algoritmo LDHR incorpora un procedimiento de identificación automático basado nuevamente en la caracterización ARMA de los componentes DHR.

Una buena identificación en este contexto supone que el contorno o perfil del pseudo-espectro del modelo identificado se asemeja al contorno del espectro de la serie temporal que se desea analizar. En nuestro caso empleamos el ar-espectro, cuyo contorno queda determinado por las raíces del polinomio autorregresivo. Una manera de obtener un modelo DHR con un pseudo-espectro parecido al ar-espectro estimado es escoger precisamente aquel modelo cuyas raíces autorregresivas aparecen en el polinomio del ar-espectro. Esta es la forma de proceder del algoritmo DHR; encuentra las raíces del ar-espectro asociadas a las frecuencias estacionales y de la tendencia, y entonces incorpora al modelo DHR aquellos componentes que tienen las mismas raíces autorregresivas.



Así pues, una vez estimado el ar-espectro la identificación es inmediata. No obstante, es necesario establecer previamente un procedimiento para decidir el orden del ar-espectro. En el algoritmo original del Young (1994) se sugiere seleccionar dicho orden mediante el Criterio de Información de Akaike. También aquí el algoritmo LDHR se aparta de la metodología original, proponiendo un criterio diferente. Un modelo DHR consiste en la suma de uno o más componentes DHR más un proceso de ruido blanco que denominamos componente irregular. Si suponemos que la serie temporal que analizamos se ajusta adecuadamente a un modelo DHR, al sustraer los espectros de los componentes DHR al espectro de dicha serie, el espectro resultante debería ser precisamente el correspondiente al del componente irregular y, por consiguiente, debería poseer un contorno similar al del espectro de un ruido blanco. Es decir, al sustraer el espectro de los componentes DHR, el espectro "residual" debería ser aproximadamente "blanco". Nuestro procedimiento de elección se basa en esta idea, y consiste en lo siguiente: primero se estiman ar-espectros de órdenes desde AR(14) hasta AR(36); a continuación se identifica y estima un modelo DHR para cada uno de los ar-espectro estimados; y finalmente se selecciona aquel modelo DHR cuyo espectro residual tiene un contorno más parecido al espectro de un proceso de ruido blanco.

## B. Índice de Transparencias

### Lista de Transparencias

- 1 Serie de Líneas Aéreas
- 2 Modelo de componentes no observables
- 3 Componentes estimados: Serie de líneas aéreas
- 4 Pasos a seguir
- 5 Serie estacional simulada
- 6 Autocovarianza y Función Generadora de Covarianzas
- 7 Serie estacional simulada: autocovarianzas
- 8 Ondas deterministas
- 9 Función Generadora de Covarianzas y Espectro
- 10 Espectro estimado de la serie simulada
- 11 Espectro Teórico del proceso estacional simulado
- 12 Serie de Líneas Aéreas

- 13 AR(17) espectro de la serie de Lineas Aéreas
- 14 Modelo básico
- 15 Modelo teórico de los sub-componentes DHR
- 16 Modelo de la Tendencia
- 17 Modelo para cada componente estacional DHR
- 18 Modelo DHR completo
- 19 Espectro de un componente DHR
- 20 Espectro de un componente DHR
- 21 Espectro de los componentes DHR de un modelo mensual
- 22 Periodograma de la serie de Lineas Aéreas
- 23 Ajuste espectral: Serie de lineas aéreas
- 24 Estimación de los hiper-parámetros
- 25 Estimación de los componentes  $s_t^{pj}$
- 26 Componentes estimados: Serie de lineas aéreas
- 27 Componentes estimados: Serie de lineas aéreas
- 28 Ajuste espectral: Log IPI
- 29 Componentes estimados: Log IPI
- 30 Componentes estimados: Log IPI
- 31 Identificación: Raíces del polinomio AR ajustado

### C. Bibliografía

- Box, G. E. P., Jenkins, G. M., y Reinsel, G. (1994). *Time Series Analysis, Forecasting and Control*. Prentice Hall, Englewood Cliffs, tercera ed. 7
- Bujosa, A., Bujosa, M., y García-Ferrer, A. (2002). A Note on the Pseudo-Spectra and the Pseudo-Covariance Generating Functions of ARMA Processes. Working paper 0203, Instituto Complutense de Análisis Económico, Universidad Complutense de Madrid. 12
- Bujosa, M. (2000). *Contribuciones al método de regresión armónica dinámica: Desarrollos teóricos y nuevos algoritmos*. Tesis Doctoral, Dpto. de Análisis Económico: Economía Cuantitativa, Universidad Autónoma de Madrid. 10, 12
- Bujosa, M., García-Ferrer, A., y Young, P. (2006). Linear Dynamic Harmonic Regression. *Computational Statistics and Data Analysis*, vol. En proceso de evaluación, -. 9, 11, 17
- Fuller, W. A. (1996). *Introduction to Statistical Time Series*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, New York, segunda ed. ISBN 0-471-55239-9. 7
- del Hoyo, J. (1994). Discussion Comments on ‘Time-variable Parameter and Trend Estimation in Non-Stationary Economic Time Series’. *Journal of Forecasting*, vol. 13, 211–214. 14
- Ng, C. y Young, P. (1990). Recursive estimation and forecasting of nonstationary time series. *Journal of Forecasting*, vol. 9, 173–204. 8
- Percival, D. B. y Walden, A. T. (1993). *Spectral Analysis for Physical Applications*. Cambridge University Press, Cambridge, UK. ISBN 0-521-43541-2. 7
- Priestley, M. P. (1981). *Spectral Analysis and Time Series*. Probability and Mathematical Statistics. Academic Press, London, primera ed. ISBN 0-12-564922-3. 7
- Shumway, R. H. y Stoffer, D. S. (2000). *Time Series Analysis and Its Applications*. Springer series in Statistics. Springer-Verlag, New York. ISBN 0-387-98950. 7
- Young, P. C. (1984). *Recursive Estimation and Time Series Analysis*. Communications and control engineering series. Springer-Verlag, Berlin, primera ed. 8
- Young, P. C. (1994). Time variable parameters and trend estimation in non-stationary economic time series. *Journal of Forecasting*, vol. 13, 179–210. 8, 18
- Young, P. C., Pedregal, D., y Tych, W. (1999). Dynamic Harmonic Regression. *Journal of Forecasting*, vol. 18, 369–394. 14