

Regresión Armónica Dinámica (DHR)

Marcos Bujosa

10/02/2023

Puede encontrar la última versión de este material en

<http://www.ucm.es/fundamentos-analisis-economico2/marcos-bujosa>



Marcos Bujosa. Copyright © 2004–2023

Algunos derechos reservados. Esta obra está bajo una licencia de Creative Commons Reconocimiento-CompartirIgual 4.0 Internacional. Para ver una copia de esta licencia, visite <http://creativecommons.org/licenses/by-sa/4.0/> o envíe una carta a Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

Índice

LECCIÓN 1: Introducción a la Regresión Armónica Dinámica	2
<i>Transparencias de la Lección 1</i>	2
1 Introducción	2
2 Espectros	4
2.1 Análisis en el ámbito del tiempo: las autocovarianzas	4
2.2 Análisis en el ámbito de la frecuencia: el espectro	6
2.3 Espectro de un modelo $ARMA(p, q)$	9
2.3.1 El espectro de la serie de Líneas Aéreas	11
3 Método DHR	11
3.1 Modelo teórico de los componentes DHR	12
3.1.1 Tendencia y componente estacional	13
3.1.2 Modelo DHR completo	14
3.2 Análisis espectral del modelo DHR	16
3.3 Identificación modelos DHR	18
3.4 Estimación de las varianzas (optimización de los NVR)	19
3.5 Estimación de los componentes	20

LECCIÓN 1: Introducción a la Regresión Armónica Dinámica

Lección 1

1 Introducción

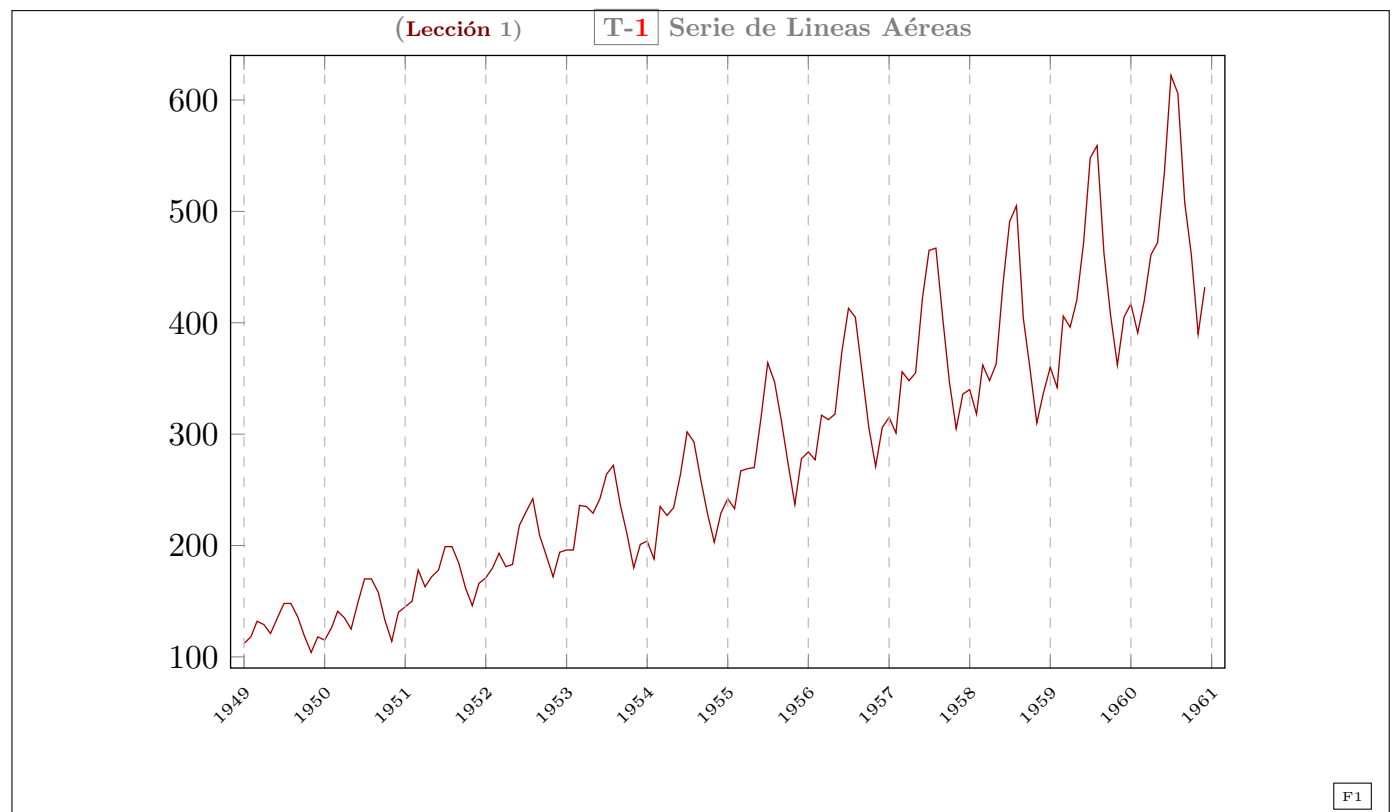
Monthly Airline Passenger Numbers 1949-1960

Description

The classic Box & Jenkins airline data. Monthly totals of international airline passengers (in thousands), 1949 to 1960.

Source

Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. (1976) Time Series Analysis, Forecasting and Control. Third Edition. Holden-Day. Series G.



(Lección 1) T-2 Modelo de componentes no observables vs ARIMA

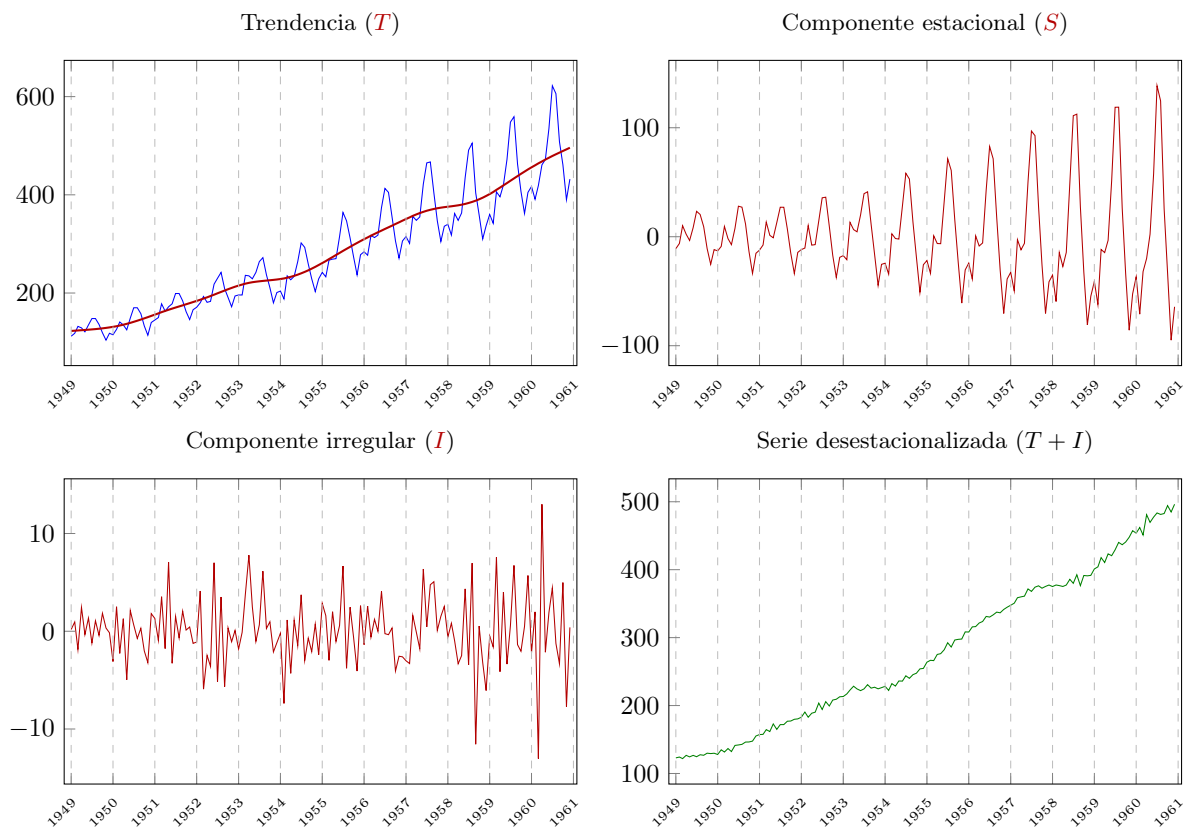
Enfoque ARIMA

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

Enfoque de componentes no observables

$$y_t = T_t + S_t + e_t$$

F2



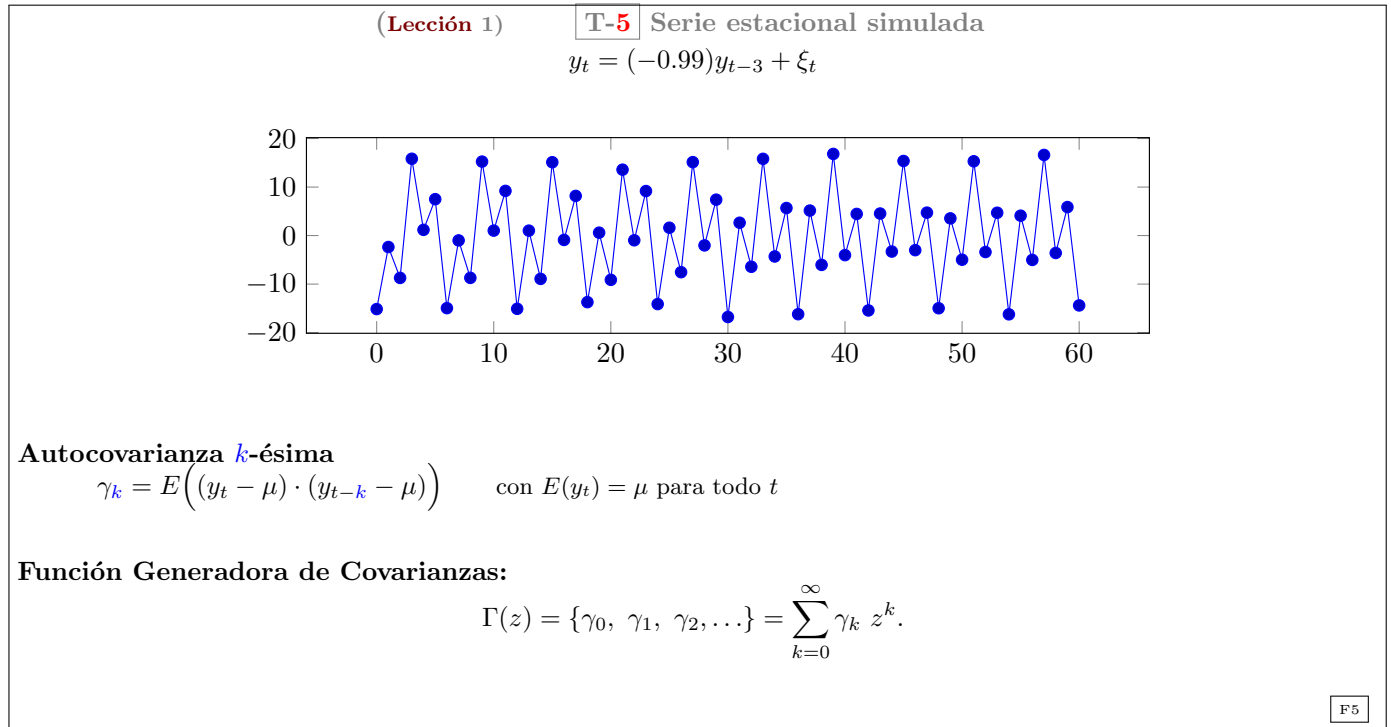
F3

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.

F4

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 cuál 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.¹

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

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:

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

donde $E(y_t) = \mu$ para todo t .

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 como²

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

donde z^k en $\gamma_k z^k$ indica que la autocovarianza γ_k aparece en la posición k -ésima; es decir, z^k , denota la posición en la secuencia del elemento al que multiplica (comenzando el índice de posiciones en 0).

¹Es una serie simulada que roza la no estacionariedad. Su modelo es $y_t = -0.99y_{t-3} + \psi_t$. 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 he escogido porque su comportamiento es muy claro a simple vista.

²En realidad esto es sólo la mitad de la sucesión. La sucesión completa es $\Gamma(z) = \dots, \gamma_{-2}, \gamma_{-1}, \gamma_0, \gamma_1, \gamma_2, \dots = \sum_{k=-\infty}^{\infty} \gamma_k z^k$. Además la sucesión es simétrica respecto a γ_0 , de manera que $\gamma_{-j} = \gamma_j$.

En la expresión anterior la letra griega Σ no es un sumatorio. La expresión $\Sigma \gamma_k z^k$ es mera notación y por tanto puede denotar una sucesión convergente o no.³ Así, la transcripción de la siguiente expresión

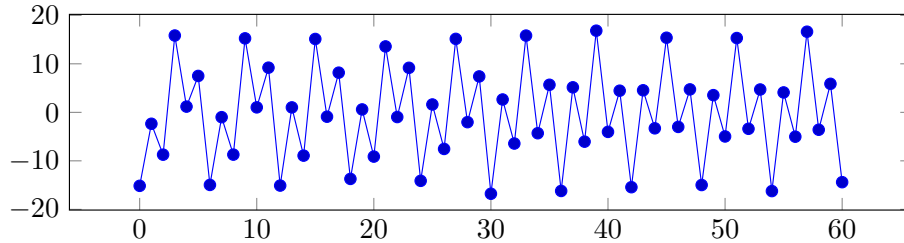
$$5z^3 + (-7)z^4 + 2z^6 + 1z^7$$

sería la sucesión $0, 0, 0, 5, -7, 0, 2, 1, 0, 0, \dots$ (donde la primera posición corresponde a z^0).

(Lección 1)

T-6 Estimación de las autocovarianzas para la serie simulada

$$y_t = (-0.99)y_{t-3} + \xi_t$$



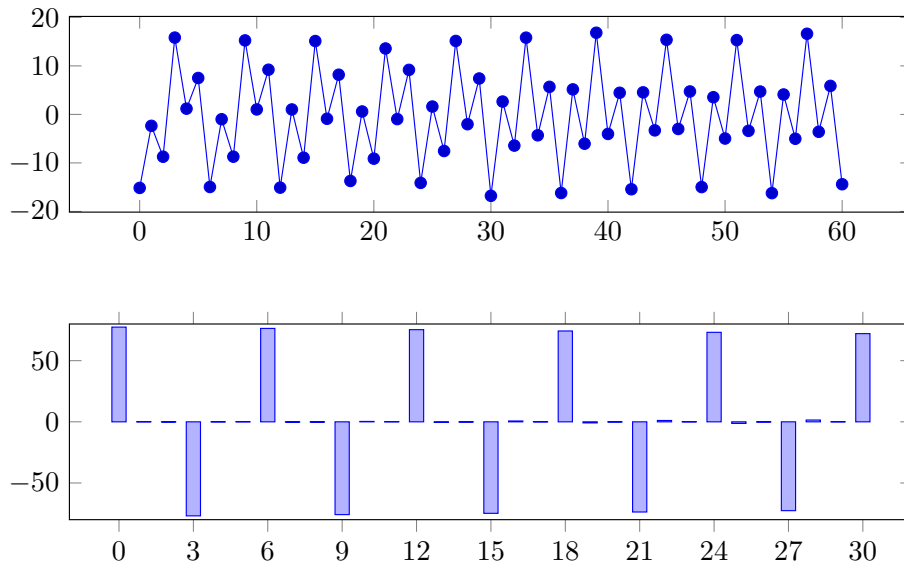
$$\begin{aligned} \widehat{\Gamma}(z) &= \{\widehat{\gamma}_0, \widehat{\gamma}_1, \widehat{\gamma}_2, \widehat{\gamma}_3, \widehat{\gamma}_4, \widehat{\gamma}_5, \widehat{\gamma}_6 \dots\} \\ &= \{45.3, -3.6, 3.6, -44.9, 3.8, -3.5, 44.4, -3.9, 3.4 \dots\} \\ &= 45.3 \cdot z^0 - 3.6 \cdot z^1 + 3.6 \cdot z^2 - 44.9 \cdot z^3 + 3.8 \cdot z^4 \dots \\ &= \sum_{k=0}^{\infty} \widehat{\gamma}_k z^k \end{aligned}$$

F6

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.

(Lección 1)

T-7 Serie estacional simulada



F7

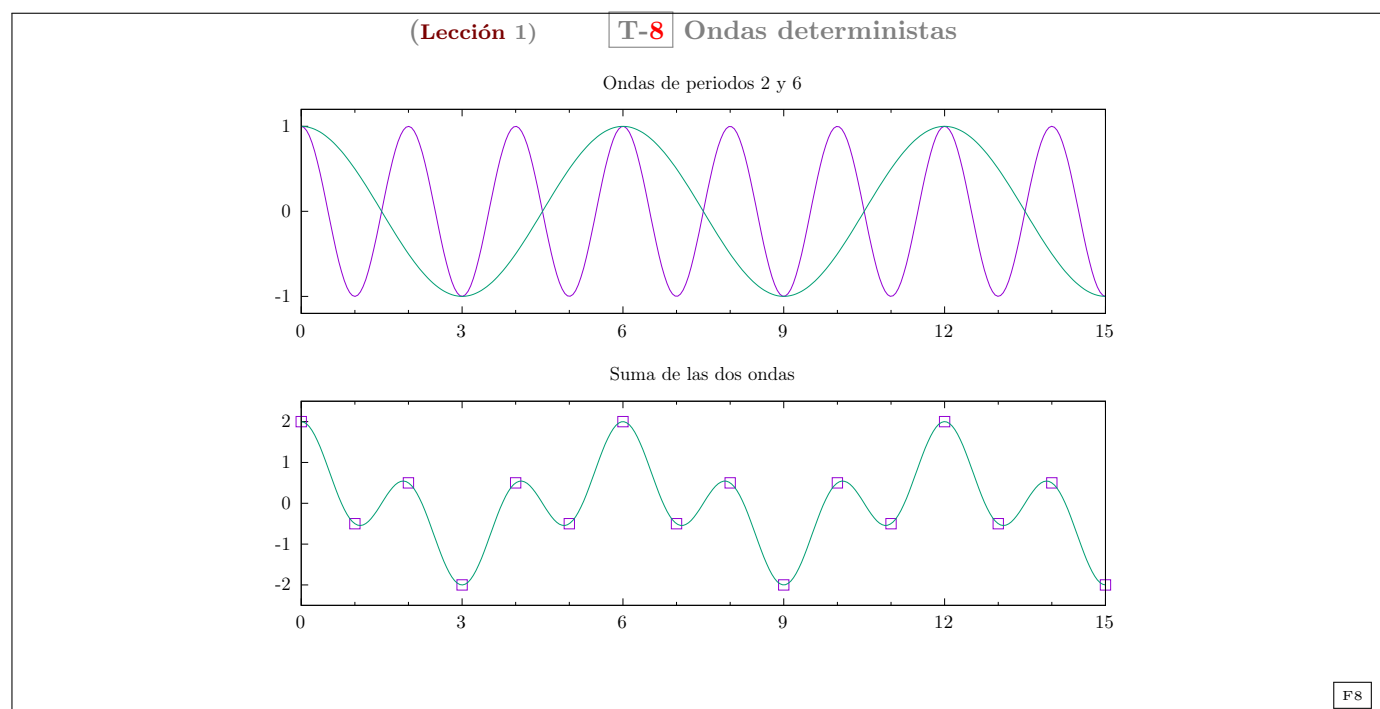
³Aunque como las sucesiones de autocovarianzas siempre son de cuadrado sumable, un poco más adelante podremos aplicar la Transformada de Fourier de dicha secuencia de autocovarianzas.

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 (trasparencias 5 y 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 planetaria 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 (pocos segundos). El nivel de la boya en cada momento será función de la suma de magnitudes de estas olas de diferentes frecuencias.

El análisis espectral trata de descomponer la variabilidad de una 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.⁴ 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 5 y 7) sobre todo al comienzo de la 2 a 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

⁴Si también hubiera una pequeña oscilación aleatoria, el aspecto sería muy parecido al ejemplo de serie estacional de la transparencia 5.

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

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

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

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

De la inversa de la transformada de Fourier, $\gamma_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i\omega k} f(\omega) d\omega$, tenemos que

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

F9

Nótese que en esta suma ponderada, la importancia de la contribución al espectro del coseno de $k\omega$ viene determinada precisamente por la covarianza k -ésima γ_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 ω .

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

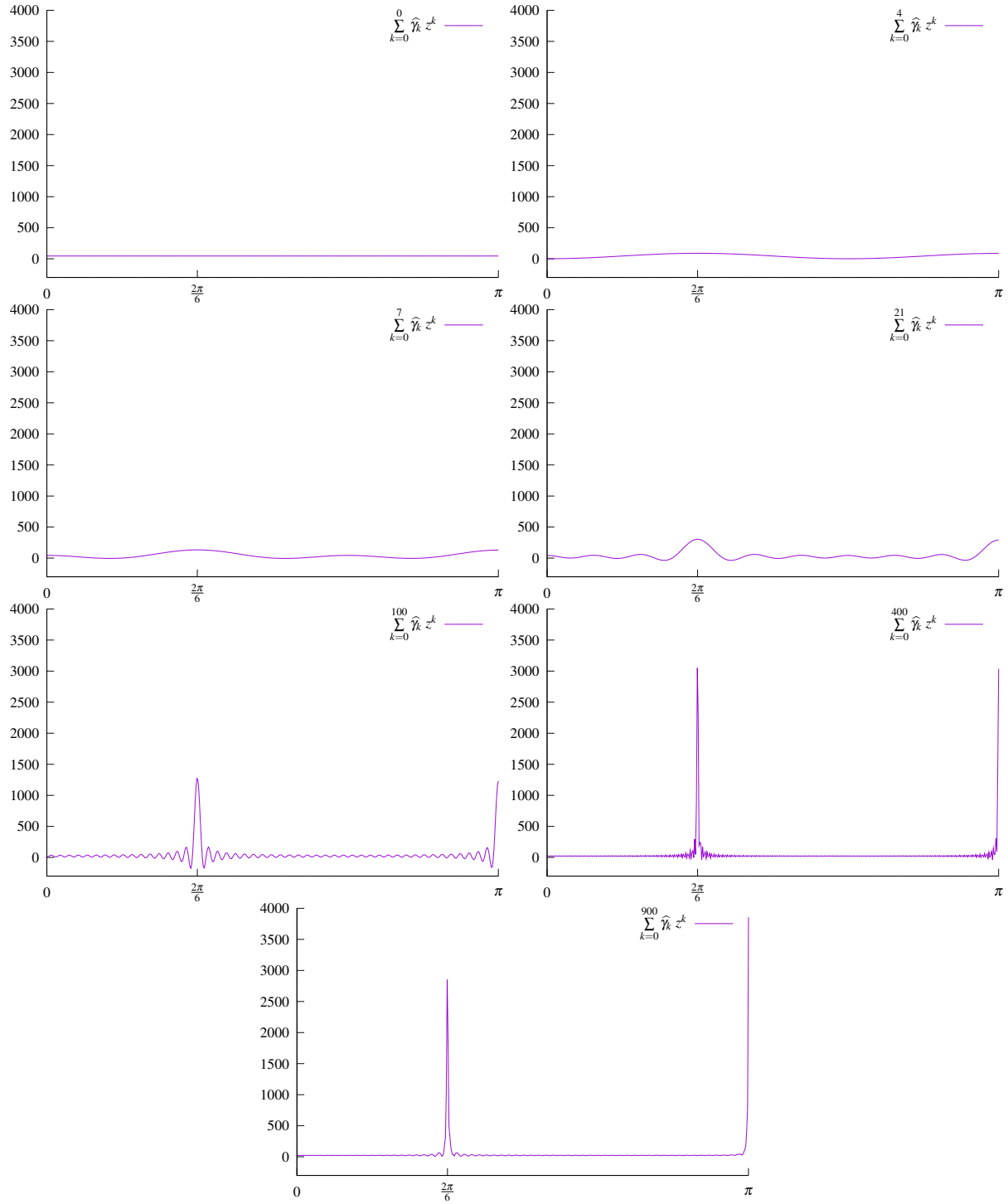
$$\begin{aligned} \widehat{f(\omega)} &= \sum_{k=0}^{\infty} \widehat{\gamma}_k \cos(k\omega) \\ &= \mathbf{45.3} + 3.6 \cos(\omega) - 3.6 \cos(2\omega) - \mathbf{44.9} \cos(3\omega) - 3.8 \cos(4\omega) + 3.5 \cos(5\omega) + \mathbf{44.4} \cos(6\omega) \dots \end{aligned}$$

Si se tiene una serie temporal suficientemente larga, y si se añaden suficientes covarianzas (si el sumatorio incorpora suficientes términos), la función obtenida adquirirá un aspecto parecido al de la última figura de la Transparencia 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 $\frac{2\pi}{2}$ y $\frac{2\pi}{6}$). Esto explica la semejanza entre las ondas deterministas y la serie temporal de los ejemplos anteriores.

$$\begin{aligned} \widehat{f(\omega)} &= \mathbf{45.3} + 3.6 \cos(1\omega) - 3.6 \cos(2\omega) \\ &\quad - \mathbf{44.9} \cos(3\omega) - 3.8 \cos(4\omega) + 3.5 \cos(5\omega) + \dots \end{aligned}$$

F10



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 ω :

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

También es posible obtener la función generadora de covarianzas, $\Gamma(z)$, a partir del espectro, $f(\omega)$, mediante la inversa de la Transformada de Fourier. Por tanto, la sucesión de números γ_k ; $k = 0, 1, \dots$; (compartida tanto por la función generadora de covarianzas como por el espectro) lo que explica que ambas funciones contengan la misma información (aunque expresada de manera diferente).

2.3 Espectro de un modelo $ARMA(p, q)$

Sea y_t un proceso estocástico que verifica

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

Llamamos *espectro* del proceso y_t a

$$f_y(\omega) = \sigma_\xi^2 \frac{\mathcal{F}(\theta(z) * \theta(z^{-1}))}{\mathcal{F}(\phi(z) * \phi(z^{-1}))} = \sigma_\xi^2 \frac{\theta(e^{-i\omega})\theta(e^{i\omega})}{\phi(e^{-i\omega})\phi(e^{i\omega})}.$$

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 $e^{-ik\omega} + e^{ik\omega} = 2\cos(k\omega)$.

(Lección 1)

T-12

Espectro de un proceso ARMA

Sea y_t un proceso estocástico *estacionario* que verifica

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

La *función generadora de covarianzas* de y_t es siguiente sucesión:

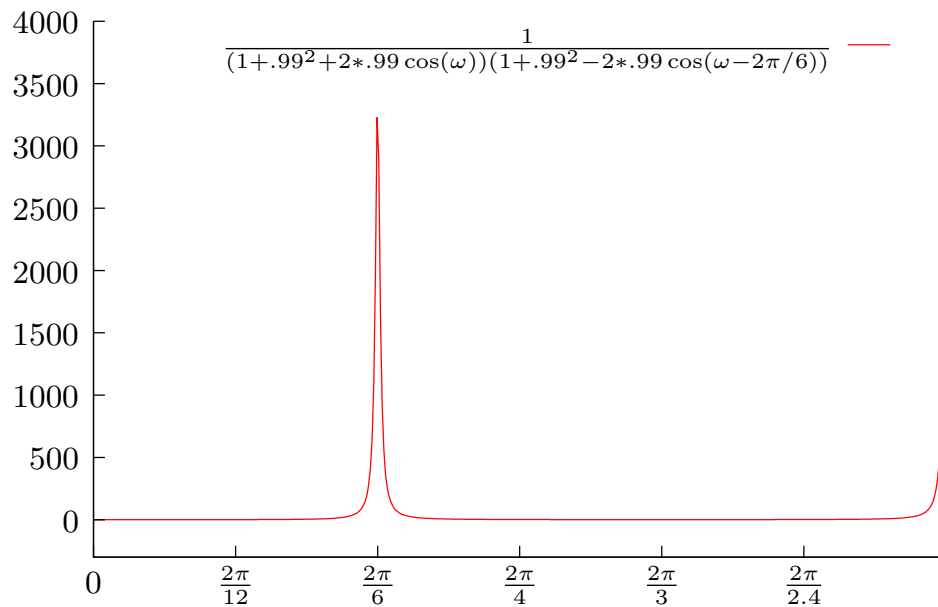
$$\Gamma_y(z) = \sigma_\xi^2 \frac{\theta(z) * \theta(z^{-1})}{\phi(z) * \phi(z^{-1})};$$

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

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

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

F12



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

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

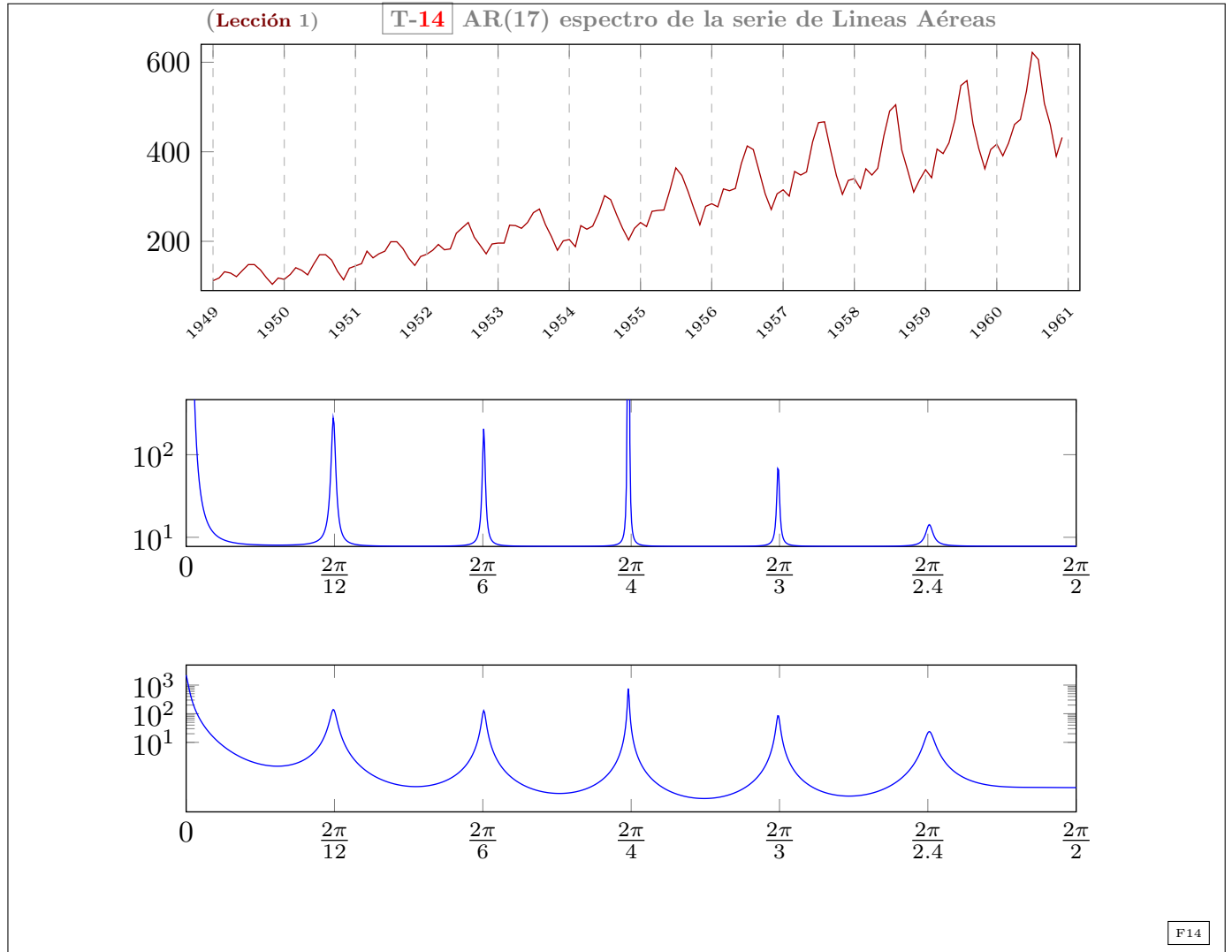
obtener la función

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

- 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.

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

2.3.1 El espectro de la serie de Lineas Aéreas



Nótese que, en la figura anterior, no todos los picos del espectro tienen la misma magnitud; 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 and 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 σ_e^2 (es decir, $\{e_t\} \sim \text{r.b. } N(0, \sigma_e^2)$).

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

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

F15

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 ω_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 ω_j es la frecuencia en ciclos por unidad de tiempo de cada componente, y $p_j \equiv 2\pi/\omega_j$ indica el periodo correspondiente⁵. 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 si:

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

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

- Proceso auto-regresivo (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 on page 15):

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

$$s_t^{p_j} = a_{j_t} \cos(\omega_j t) + b_{j_t} \sin(\omega_j t), \quad (2)$$

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).

F16

⁵para una serie mensual como la de “Líneas Aéreas”, $P_j = \infty, 12, 6, 4, 3, 2.4, 2$ y $\omega_j = 0, \frac{2\pi}{12}, \frac{2\pi}{6}, \frac{2\pi}{4}, \frac{2\pi}{3}, \frac{2\pi}{2.4}, \frac{2\pi}{2}$.

3.1.1 Tendencia y componente estacional

Los componentes DHR se agregan en función de las frecuencias ω_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).

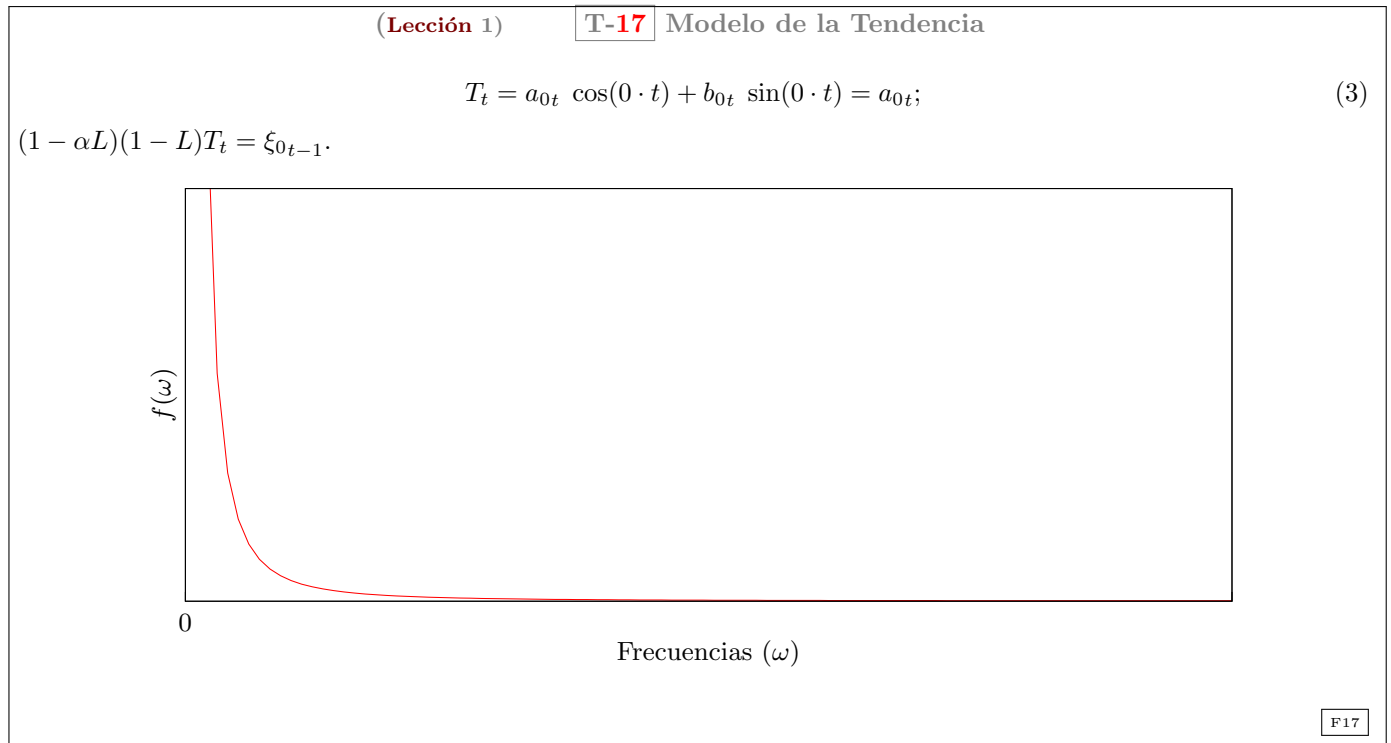
Tendencia La tendencia está constituida por el componente DHR s_t^∞ , que corresponde a la frecuencia $\omega_0 = 0$. Por lo tanto, $T_t \equiv s_t^\infty$, y la Ecuación (2) pasa a ser:

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

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 (1) ni con el resto de componentes.



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:

$$S_t = \sum_{j=1}^R s_t^{p_j} = \sum_{j=1}^R [a_{jt} \cos(\omega_j t) + b_{jt} \sin(\omega_j t)], \quad (4)$$

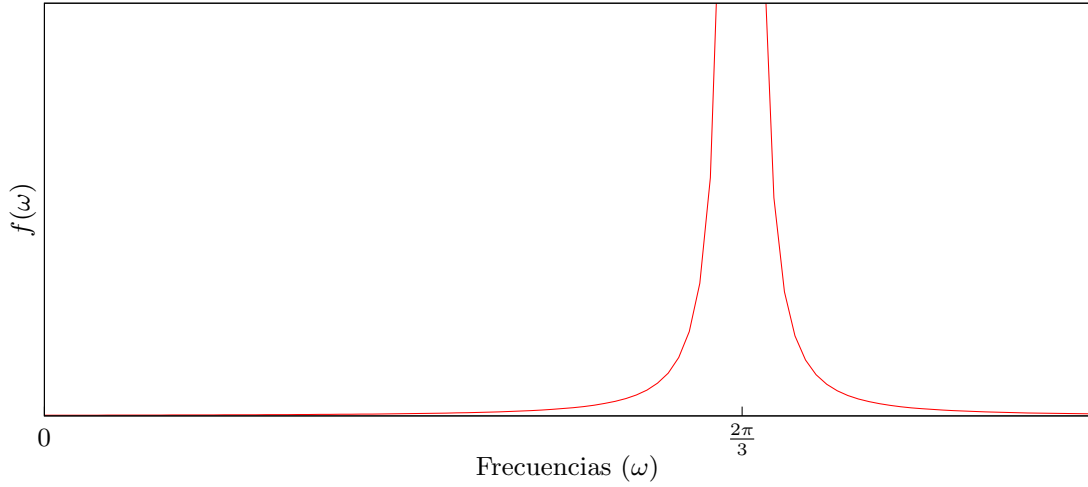
donde R representa el número de armónicos fundamentales de la frecuencia estacional.

Desde ahora nos referimos a los sumandos de (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. (2007) 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 IRW⁶.

⁶Los 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

$$s_t^{pj} = a_{jt} \cos(\omega_j t) + b_{jt} \sin(\omega_j t) \quad (5)$$



F18

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 (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 on the next page.

3.1.2 Modelo DHR completo

Si sustituimos T_t y S_T en la Ecuación (1) por las expresiones (3), y (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.

$$y_t = \sum_{j=0}^R [a_{jt} \cos(\omega_j t) + b_{jt} \sin(\omega_j t)] + e_t, \quad (6)$$

donde $j = 0$ corresponde a la tendencia.

Parámetros

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

F19

$3\pi/4$ 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 ($0 \leq \alpha, \beta \leq 1$)	$(1 - (\alpha_0 + \beta_0)L + \alpha_0\beta_0L^2)T_t = \xi_{0t-1}$	$(\phi_j^\alpha(L) * \phi_j^\beta(L))s^{p_j}_t = \left(\sqrt{\frac{\alpha_j\beta_j \cos(2\omega_j)}{\gamma_j^* \eta_j^*}} \right) (1 - \theta_j^1 L - \theta_j^2 L^2) \xi_{jt-1}$	$(1 + (\alpha_R + \beta_R)L + \alpha_R\beta_R L^2)s_t^2 = \xi_{Rt-1}$
RW ($\alpha=0, \beta=1$)	$(1 - L)T_t = \xi_{0t-1}$	$\phi_j^\beta(L)s_t^{p_j} = \frac{1}{\sqrt{1 + \sin(\omega_j)}} \left(1 - \frac{\cos(\omega_j)}{1 + \sin(\omega_j)} L \right) \xi_{jt-1}$	$(1 + L)s_t^2 = \xi_{Rt-1}$
SRW ($0 < \alpha < 1, \beta=1$)	$(1 - (1 + \alpha_0)L + \alpha_0L^2)T_t = \xi_{0t-1}$	$(\phi_j^\alpha(L) * \phi_j^\beta(L))s^{p_j}_t = \left(\sqrt{\frac{\alpha_j \cos(2\omega_j)}{\gamma_j^* \eta_j^*}} \right) (1 - \theta_j^1 L - \theta_j^2 L^2) \xi_{jt-1}$	$(1 + (1 + \alpha_R)L + \alpha_R L^2)s_t^2 = \xi_{Rt-1}$
IRW ($\alpha=\beta=1$)	$(1 - 2L + L^2)T_t = \xi_{0t-1}$	$(\phi_j^\alpha(L) * \phi_j^\beta(L))s^{p_j}_t = \left(\sqrt{\frac{\cos(2\omega_j)}{\gamma_j^* \eta_j^*}} \right) (1 - \theta_j^1 L - \theta_j^2 L^2) \xi_{jt-1}$	$(1 + 2L + L^2)s_t^2 = \xi_{Rt-1}$

$$\begin{aligned}
\phi_j^\alpha(L) &= [1 - \alpha e^{i\omega_j} L] * [1 - \alpha e^{-i\omega_j} L] = [1 - 2\alpha \cos(\omega_j) L + \alpha^2 L^2]; \\
\phi_j^\beta(L) &= [1 - \beta e^{i\omega_j} L] * [1 - \beta e^{-i\omega_j} L] = [1 - 2\beta \cos(\omega_j) L + \beta^2 L^2]; \\
\theta_j^1 &= \gamma_j^* + \eta_j^*; \quad \theta_j^2 = -\gamma_j^* \eta_j^* \quad (\text{para más detalles véase Bujosa et al., 2007}).
\end{aligned}$$

Table 1: Especificación ARMA de los componentes DHR

En la transparencia 22 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 $12/j$, $j = 1, 2, \dots, 6$.

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:

$$\begin{aligned} f_n(\omega) &= \mathcal{FE}(\Lambda_n(z)) \\ &= \mathcal{FE}\left(\frac{\sigma_\xi^2}{(1-\alpha z)(1-z)(1-\alpha z^{-1})(1-z^{-1})}\right) = \frac{\sigma_\xi^2}{(1+\alpha^2-2\alpha\cos(\omega))(2-2\cos(\omega))}, \end{aligned} \quad (7)$$

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., 2015).

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 et al., 2015); 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:

$$f_{dhr}(\omega, \sigma^2) = \sum_{j=0}^R \sigma_{\omega_j}^2 f_{s^{p_j}}(\omega) + \sigma_e^2, \quad (8)$$

donde

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

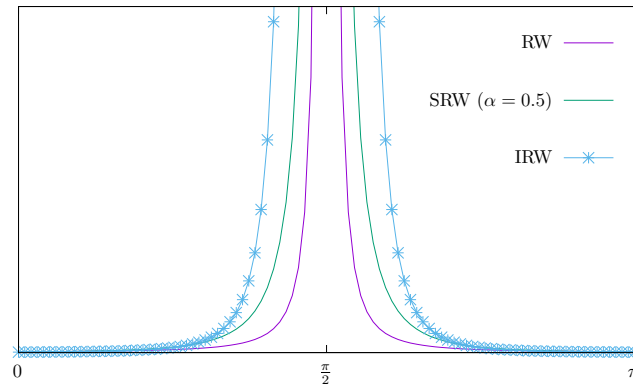
y donde σ_e^2 es la varianza del componente irregular.

En la transparencia 21 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)⁷.

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.

⁷En la figura sólo representamos el intervalo $[0, \pi]$, debido a que las funciones son simétricas respecto a la frecuencia 0.

Distintos modelos de $\{a_j\}$ y $\{b_j\}$; (misma varianza innovaciones $\sigma_{\omega_j}^2$)

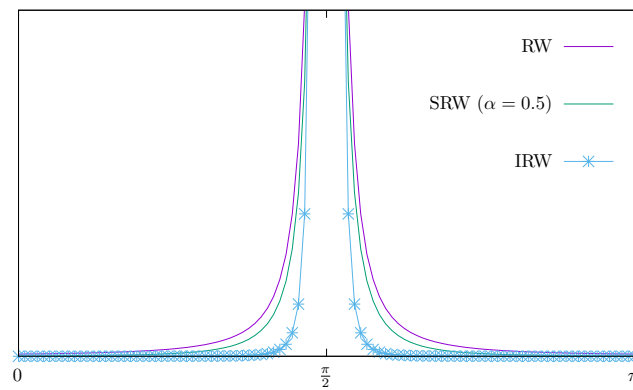


F20

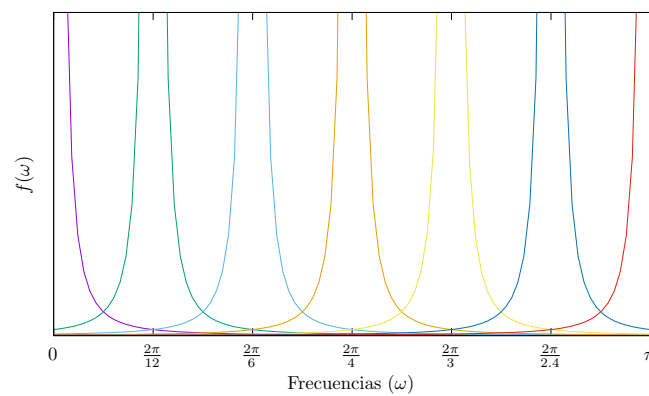
$$\sigma_{RW}^2 = 1$$

$$\sigma_{SRW}^2 = 0.2$$

$$\sigma_{IRW}^2 = 0.005$$

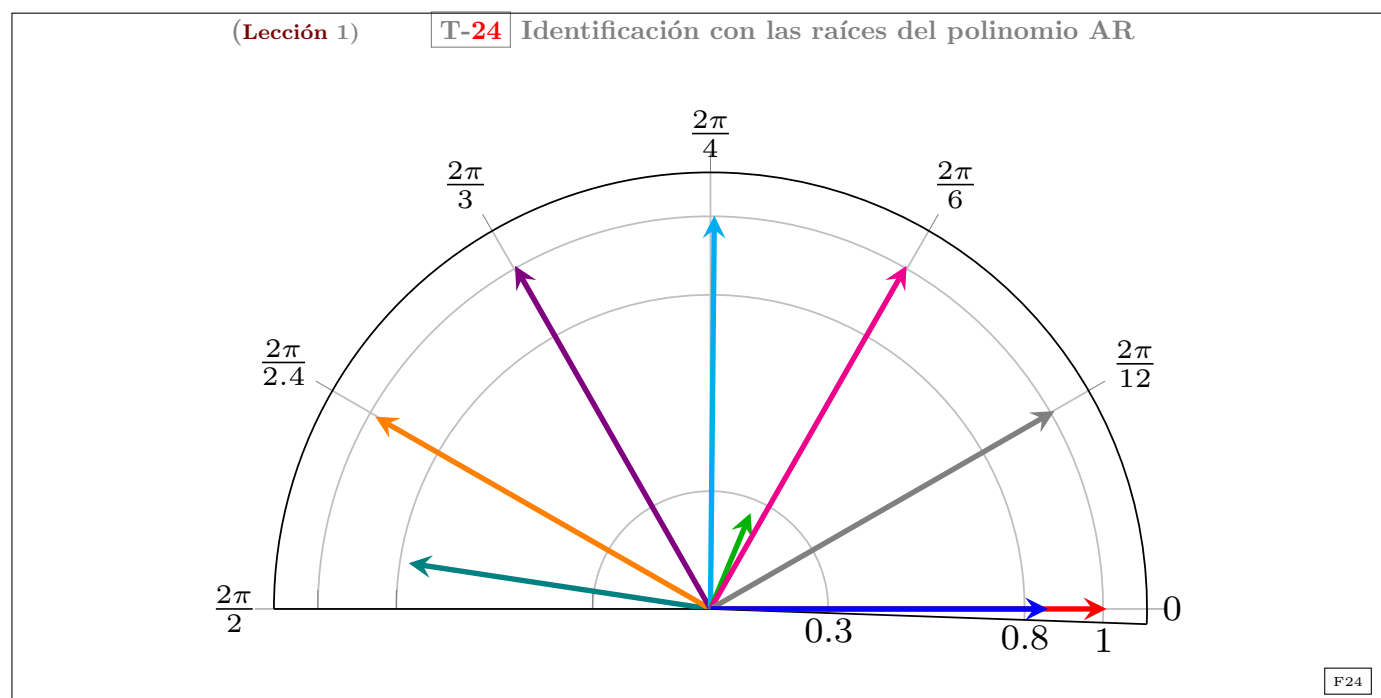
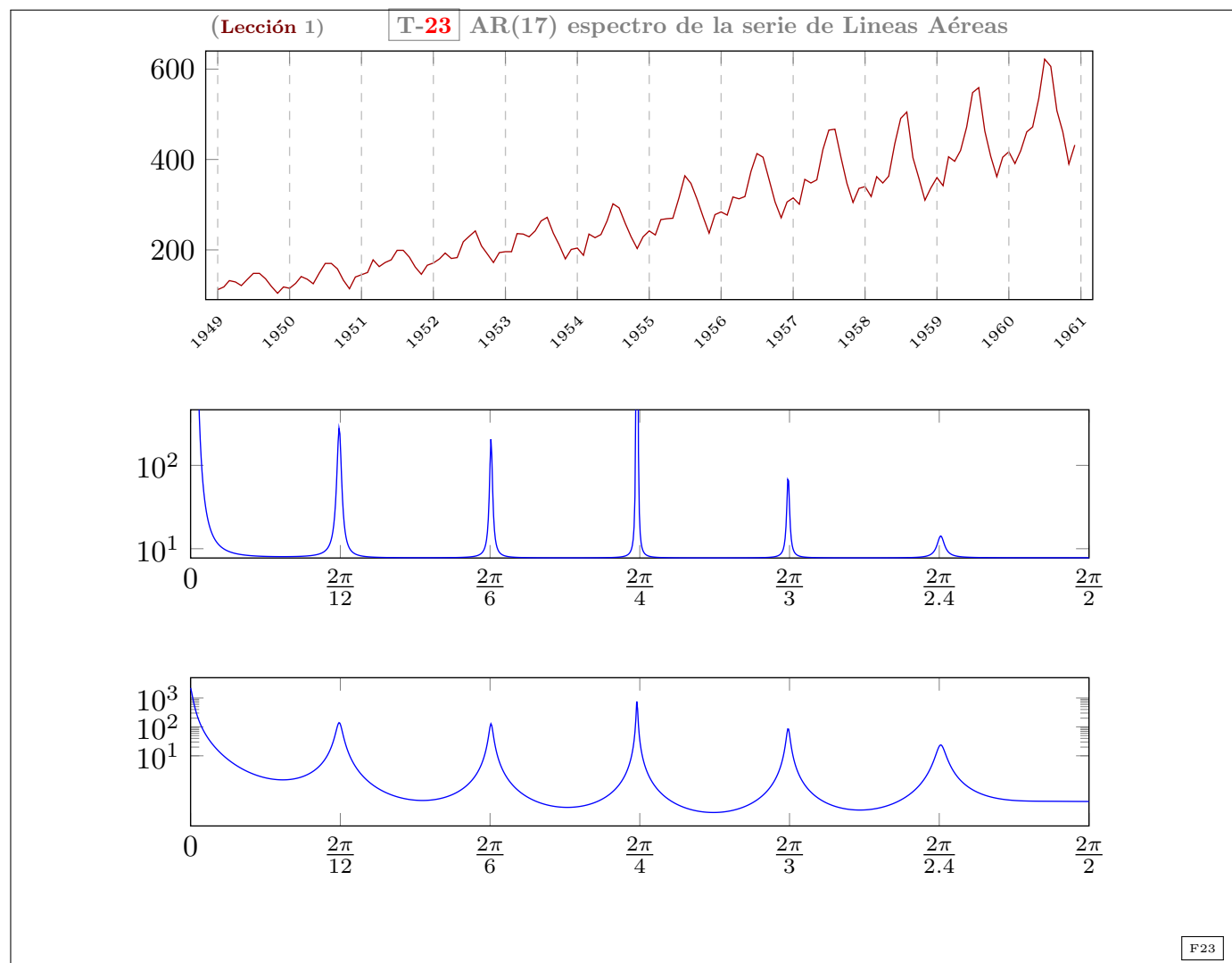


F21



F22

3.3 Identificación modelos DHR



$$\text{Modelo DHR No estacionario : } \begin{cases} T & \rightarrow SRW \\ S^{12} & \rightarrow RW \\ S^6 & \rightarrow RW \\ S^4 & \rightarrow RW \\ S^3 & \rightarrow RW \\ S^{2.4} & \rightarrow RW \end{cases}$$

3.4 Estimación de las varianzas (optimización de los NVR)

(Lección 1)

T-25 Estimación de los hiperparámetros σ^2

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

donde

$\widehat{f_y(\omega)}$ es una estimación del espectro de la serie y

donde el espectro del modelo DHR es

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

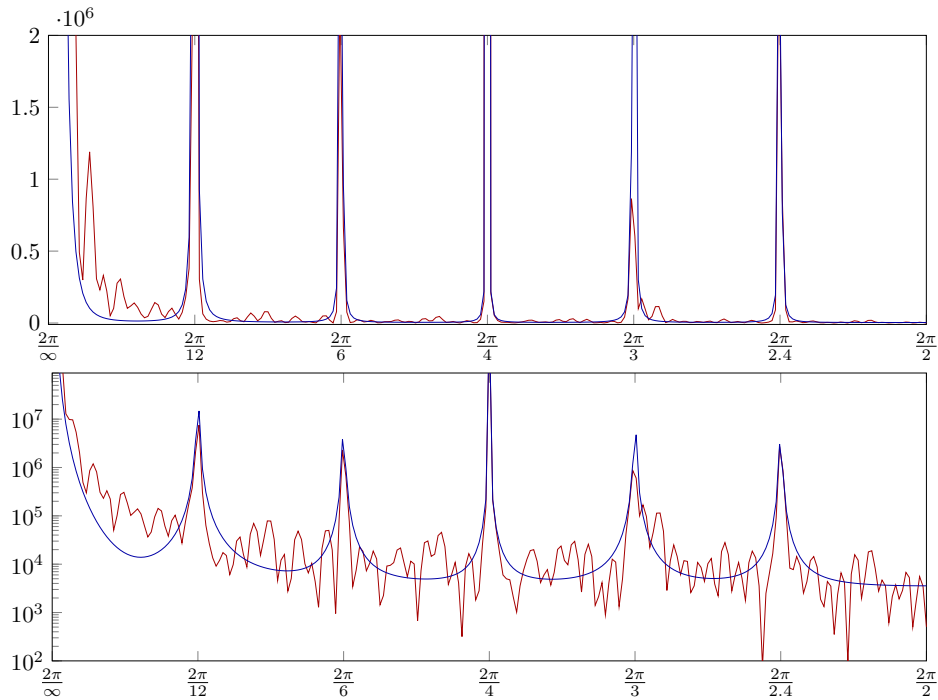
y donde

- $\sigma^2 = (\sigma_e^2, \sigma_{\omega_0}^2, \sigma_{\omega_1}^2, \dots, \sigma_{\omega_R}^2)$
- $f_{s^{p_j}}(\omega)$ es el espectro del componente j -ésimo.

F25

(Lección 1)

T-26 Ajuste del modelo DHR al periodograma



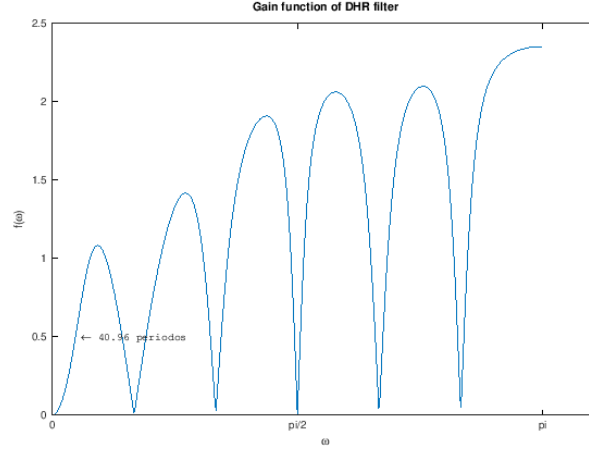
F26

Modelo DHR tiene una expresión ARIMA equivalente

$$y_t = \frac{\theta(L)}{\phi(L)} \epsilon_t \quad \text{donde} \quad \phi(L) = \varphi(L) * \Phi(L)$$

donde $\Phi(L)$ es el polinomio con las raíces de módulo uno.

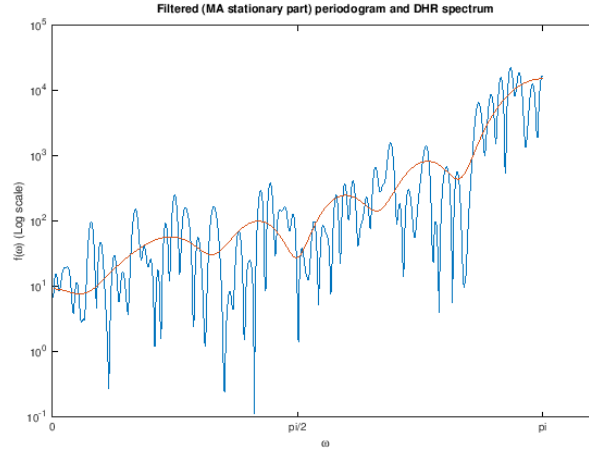
Sea $\Psi(\omega)$ la Transformada de Fourier de $\Phi(L)$



F27

Estimación por MCO

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



F28

3.5 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 ω_j correspondientes a cada uno de los componentes s^{p_j} . El parámetro α_j determina la tasa de decaimiento del filtro alrededor de la frecuencia ω_j (véase Young et al., 1999).

El vector de hiper-parámetros NVR Una vez fijados los valores ω_j y α_j , lo único desconocido son los cocientes entre varianzas $\sigma_{\omega_j}^2/\sigma_e^2$. Denominamos a estos cocientes hiper-parámetros NVR_j ⁸, 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

$$\boldsymbol{\sigma}^2 = [\sigma_e^2, \sigma_{\omega_0}^2, \sigma_{\omega_1}^2, \dots, \sigma_{\omega_R}^2]', \quad (9)$$

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

Dado el modelo DHR completo (número de componentes y valores ω_j y α_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.

(Lección 1)

T-29

Estimación de los componentes $s_t^{p_j}$

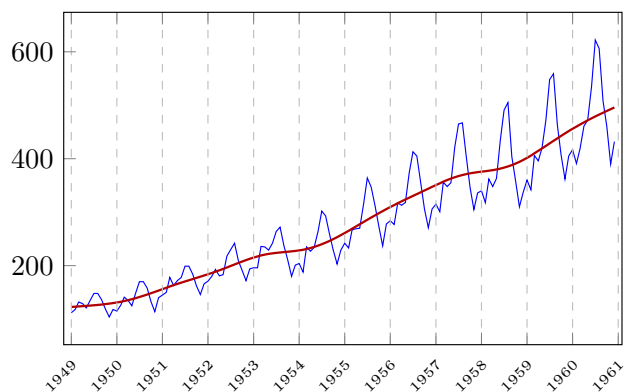
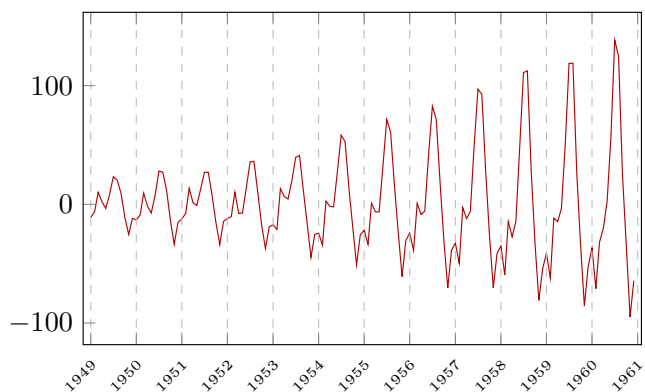
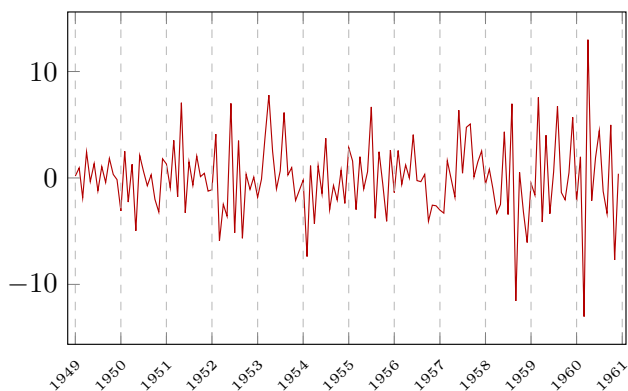
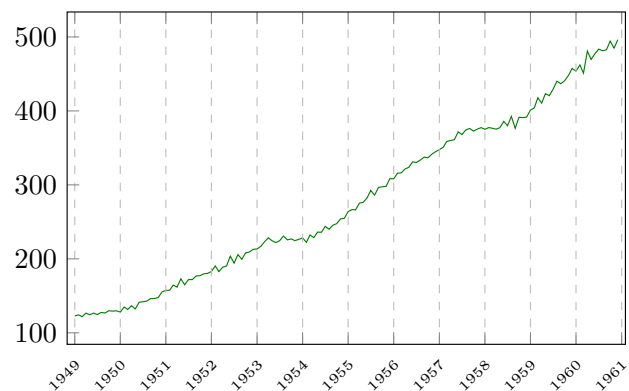
Una vez estimadas las varianzas desconocidas (*hiper-parámetros*)

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

F29

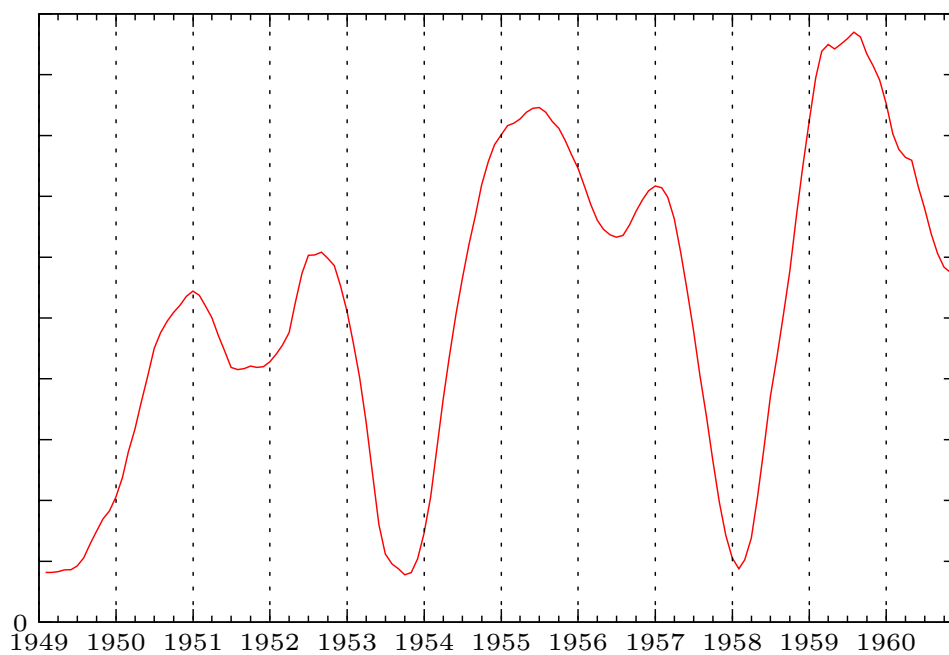
$$NVR_j = \frac{\sigma_{\omega_j}^2}{\sigma_e^2}$$

⁸del inglés “Noise Variance Ratio”.

Tendencia (T)Componente estacional (S)Componente irregular (I)Serie desestacionalizada ($T + I$)

F30

Primera diferencia de la tendencia



F31

References

- Box, G. E. P., Jenkins, G. M., and Reinsel, G. (1994). *Time Series Analysis, Forecasting and Control*. Prentice Hall, Englewood Cliffs, third ed.
- Bujosa, M. (2000). *Contribuciones al método de regresión armónica dinámica: Desarrollos teóricos y nuevos algoritmos*. Ph.D. thesis, Dpto. de Análisis Económico: Economía Cuantitativa, Universidad Autónoma de Madrid.
- Bujosa, M., Bujosa, A., and García-Ferrer, A. (2015). Mathematical framework for pseudo-spectra of linear stochastic difference equations. *IEEE Transactions on Signal Processing*, **63**(24), 6498–6509.
URL <http://eprints.ucm.es/20699/>
- Bujosa, M., García-Ferrer, A., and Young, P. C. (2007). Linear dynamic harmonic regression. *Comput. Stat. Data Anal.*, **52**(2), 999–1024. ISSN 0167-9473.
URL <http://www.sciencedirect.com/science/article/pii/S0167947307002666>
- Fuller, W. A. (1996). *Introduction to Statistical Time Series*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, New York, second ed. ISBN 0-471-55239-9.
- Ng, C. and Young, P. (1990). Recursive estimation and forecasting of nonstationary time series. *Journal of Forecasting*, **9**, 173–204.
- Percival, D. B. and Walden, A. T. (1993). *Spectral Analysis for Physical Applications*. Cambridge University Press, Cambridge, UK. ISBN 0-521-43541-2.
- Priestley, M. P. (1981). *Spectral Analysis and Time Series*. Probability and Mathematical Statistics. Academic Press, London, first ed. ISBN 0-12-564922-3.
- Shumway, R. H. and Stoffer, D. S. (2000). *Time Series Analysis and Its Applications*. Springer series in Statistics. Springer-Verlag, New York. ISBN 0-387-98950.
- Young, P. C. (1984). *Recursive Estimation and Time Series Analysis*. Communications and control engineering series. Springer-Verlag, Berlin, first ed.
- Young, P. C. (1994). Time variable parameters and trend estimation in non-stationary economic time series. *Journal of Forecasting*, **13**, 179–210.
- Young, P. C., Pedregal, D., and Tych, W. (1999). Dynamic harmonic regression. *Journal of Forecasting*, **18**, 369–394.