

Análisis espectral y ajuste estacional

18 de octubre de 2012

1. Introducción

El análisis espectral de series temporales es una herramienta muy útil para el ajuste estacional. Aunque no hay una definición precisa de estacionalidad, sí hay un acuerdo general sobre al menos algunos fenómenos que constituyen estacionalidad. En particular, hay estacionalidad en una serie x_t cuando ésta contiene componentes periódicas, cuyo periodo coincide con un tiempo de referencia¹. En el tipo de indicadores coyunturales que publican los organismos estadísticos este tiempo de referencia es el año. De esta forma, para datos mensuales las componentes periódicas de periodo 12 (o alguno de sus divisores) se consideran estacionales. Igualmente ocurre en el caso trimestral con las componentes periódicas de periodo 4 (o su único divisor, 2).

Precisamente la representación espectral de un proceso es una herramienta especialmente adecuada para sacar a la luz componentes periódicas, ya que consiste en expresar el proceso como una suma de sucesiones periódicas, concretamente trigonométricas. Esta representación, por tanto, nos permite centrar la atención en las frecuencias que más nos interesen.

Como consecuencia de lo anterior, el análisis espectral se puede emplear en la fase de diagnóstico del proceso de ajuste estacional, para comprobar la ausencia de componentes estacionales en la serie ajustada. Los programas que se usan para hacer el ajuste estacional (como TRAMO/SEATS) incluyen una representación del espectro entre las herramientas de diagnóstico.

Para interpretar correctamente el espectro conviene conocer, entre otras cosas:

¹Estrictamente hablando, cuando lo divide.

- Las distintas variantes de representación del espectro (unidades de frecuencia, escala de la densidad espectral, etc.)
- Dónde están las frecuencias asociadas con la estacionalidad.
- Cómo son los espectros de ciertos procesos, en particular de los ARMA.
- Cómo afectan los filtros al espectro.
- Qué diferencias hay entre la representación espectral de procesos estacionarios e integrados.

En este documento pretendemos –de manera concisa– introducir primero los fundamentos de la representación espectral en general y después las cuestiones especialmente relevantes para el ajuste estacional.

2. El espectro

Una serie temporal puede entenderse como una sucesión de números. Las sucesiones de números reales se pueden escribir como combinaciones lineales de senos y cosenos (o exponenciales complejas). Esta representación se hace mediante la Transformada de Fourier Discreta (TFD). En 2.1 veremos cómo.

Sin embargo, la teoría moderna de series temporales se basa en la hipótesis de que una serie temporal es una realización de un proceso estocástico. Análogamente a las series temporales, los procesos estocásticos se pueden representar mediante combinaciones lineales (más exactamente, integrales) de funciones trigonométricas o exponenciales complejas, aunque ponderados por coeficientes aleatorios. Esta representación es la llamada representación espectral de un proceso estocástico, que presentamos en 2.3. La varianza de los coeficientes de la representación se llama densidad espectral. Veremos en 2.2 que la TFD de una serie temporal puede entenderse como una estimación de la densidad espectral.

Desgraciadamente, la representación espectral de un proceso estocástico requiere usar un concepto matemático complejo: una medida aleatoria. Por otra parte, es posible definir la densidad espectral a partir de la función de autocovarianza sin aludir a la representación espectral. Por tanto, quien lo desee puede omitir el apartado 2.3.

2.1. Representación de Fourier de una sucesión

Una sucesión finita x_1, \dots, x_T (que puede ser una serie temporal, es decir, una realización de un proceso estocástico) puede escribirse como suma de senos y cosenos de la forma siguiente:

$$x_t = \sum_{k=0}^n (a_k \cos \omega_k t + b_k \sin \omega_k t), \quad (1)$$

donde $\omega_k = 2\pi k/T$ son las *frecuencias naturales*. Cuando T es par, $n = T/2$. Entonces, $\sin \omega_0 t = \sin \omega_n t = 0$, mientras que $\cos \omega_0 t = 1$ y $\cos \omega_n t = (-1)^t$. Consecuentemente,

$$x_t = a_0 + \sum_{k=1}^{n-1} (a_k \cos \omega_k t + b_k \sin \omega_k t) + a_n (-1)^t. \quad (2)$$

Si T es impar entonces $n = (T - 1)/2$ y

$$x_t = a_0 + \sum_{k=1}^n (a_k \cos \omega_k t + b_k \sin \omega_k t).$$

Los coeficientes a_k y b_k nos dan la *amplitud* de las funciones sinusoidales. Se puede comprobar que, en estas dos formas, la representación es única viendo que $\{\cos \omega_k t, \sin \omega_k t\}_{k=1}^{n-1} \cap \{1, (-1)^t\}$ en el caso par y $\{\cos \omega_k t, \sin \omega_k t\}_{k=1}^n \cap \{1\}$ en el impar forman una base del espacio de las sucesiones finitas de T elementos.

Por otra parte, los coeficientes a_k y b_k se pueden obtener como

$$a_k = \frac{2}{T} \sum_{t=1}^T x_t \cos \omega_k t \quad b_k = \frac{2}{T} \sum_{t=1}^T x_t \sin \omega_k t.$$

En ambos casos, las frecuencias que aparecen en la representación cumplen $\omega_k \leq \pi$. El por qué necesitamos solo frecuencias hasta π en lugar de recorrer todo el intervalo $[0, 2\pi)$ es consecuencia de lo siguiente. En la figura 1 representamos las funciones $\cos 2\pi 2t/T$ y $\cos 2\pi 6t/T$ para $T = 8$. Como se puede ver, las dos funciones coinciden en los valores enteros del eje de abscisas. Igualmente ocurre para cualquier par de frecuencias $(\omega_k, 2\pi - \omega_k)$. Esto quiere decir que tendríamos que tener valores de la variable que corresponden a instantes t no enteros para poder distinguir las dos funciones.

Podemos expresar esta idea de otra forma: la mayor frecuencia que se puede observar en una serie x_t es la que corresponde a medio ciclo por unidad de tiempo. Esta frecuencia se llama frecuencia de Nyquist.

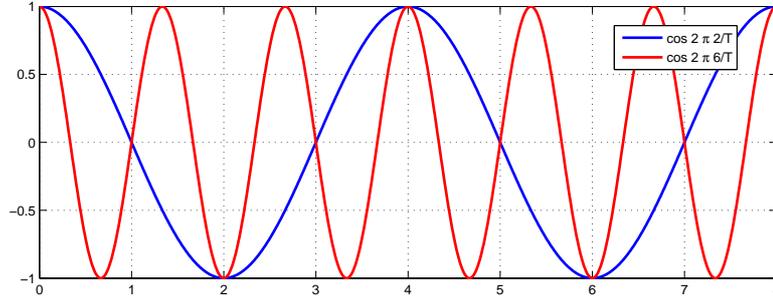


Figura 1: Frecuencias mayores de π no son observables.

De esta forma, una sucesión se puede representar como un conjunto de pares (t, x_t) o (con cierta pérdida de información) como un conjunto de pares $(\omega_k, a_k^2 + b_k^2)$. La primera representación se llama *en el dominio del tiempo* y la segunda *en el dominio de la frecuencia*. A menudo, por analogía con las señales físicas se dice que $a_k^2 + b_k^2$ es la *energía* contenida en la frecuencia ω_k .

En la figura 2 representamos la sucesión $x_t = \cos \omega_1 t + (1/2) \sin \omega_2 t$, donde $T = 12$, $\omega_1 = 2\pi 2/T$ y $\omega_2 = 2\pi 4/T$. En el panel superior representamos la sucesión en el dominio del tiempo, en el segundo, mediante un diagrama de barras, la sucesión en el dominio de la frecuencia (representación espectral) y en el tercero, representamos las componentes sinusoidales que corresponden a las dos frecuencias que aparecen en la representación espectral.

La representación (2) se puede expresar también mediante exponenciales complejas. Para ello sustituimos $\cos \omega_k t$ y $\sin \omega_k t$ por

$$\frac{e^{i\omega_k t} + e^{-i\omega_k t}}{2} \quad \frac{e^{i\omega_k t} - e^{-i\omega_k t}}{2i}$$

respectivamente. De ahí obtenemos

$$x_t = a_0 + \sum_{k=1}^{n-1} \left(\frac{a_k - ib_k}{2} e^{i\omega_k t} + \frac{a_k + ib_k}{2} e^{-i\omega_k t} \right) + a_n (-1)^t.$$

Si ahora extendemos a_k y b_k a $k = T/2 + 1, \dots, T - 1$ mediante $a_k = a_{T-k}$ y $b_k = -b_{T-k}$ entonces

$$x_t = \sum_{k=1}^T z_k e^{i\omega_k t},$$

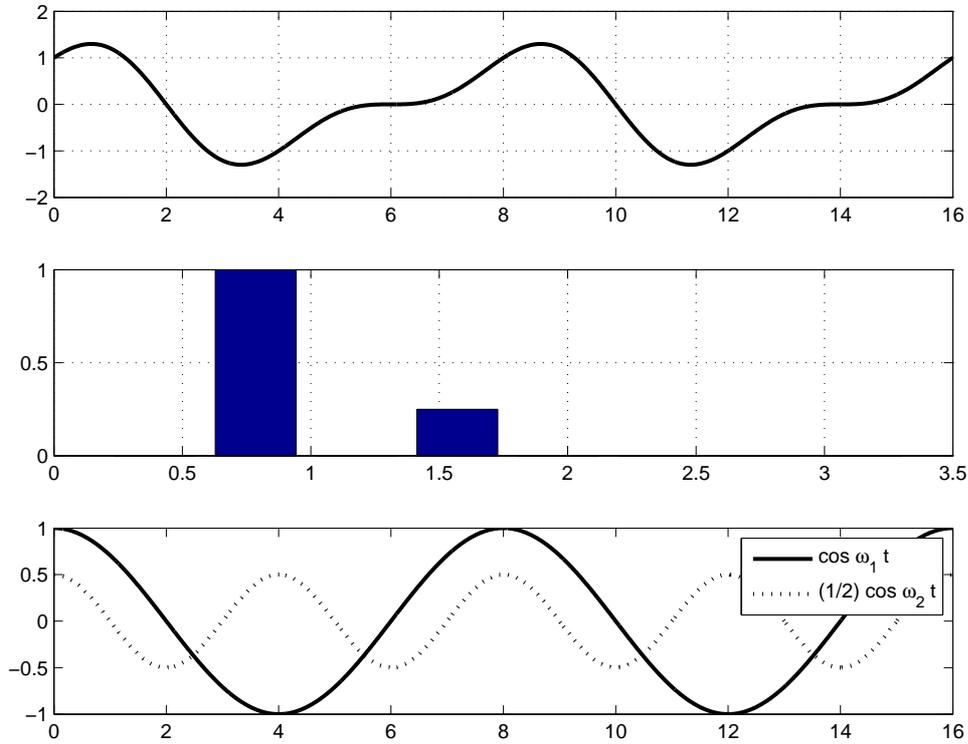


Figura 2: Una sucesión que es combinación de dos funciones sinusoidales.

donde $z_k = a_k - ib_k$, siendo $b_0 = b_{T/2} = 0$.

Los coeficientes z_k se pueden obtener como

$$z_k = \frac{1}{T} \sum_{t=1}^T x_t e^{-i\omega_k t}.$$

2.1.1. Representación de una sucesión periódica

Supongamos ahora que una sucesión x_t de longitud T es periódica de periodo s , donde $s < T$. Es decir, que $x_{t+s} = x_t$ y por tanto, el conjunto de valores x_1, \dots, x_s se repite sucesivamente. Podemos considerar estos valores x_1, \dots, x_s como una sucesión en sí misma y aplicar la representación espectral. Suponiendo para simplificar que s sea par, podemos escribir

$$x_t = a_0 + \sum_{k=1}^{s/2-1} (a_k \cos 2\pi kt/s + b_k \sin 2\pi kt/s) + a_n (-1)^t. \quad (3)$$

Puesto que todos los senos y cosenos que aparecen son periódicos de periodo s , entonces el lado derecho de la ecuación es también periódico. Esto significa que (3) no solo se cumple para $t = 1, \dots, s$, sino también para $t = s+1, \dots, T$. Consecuentemente: **en la representación de Fourier de una sucesión periódica de periodo s aparecen solamente las $s/2 + 1$ frecuencias $0, 2\pi/s, \dots, \pi$.**

Como ejemplo, consideramos el caso de $T = 32$ y $s = 4$ en la figura 3 representamos una sucesión periódica de periodo s y en la figura 4 una sucesión no periódica. Como vemos, en el primer caso la representación espectral solo contiene las frecuencias $2\pi/4$ y $4\pi/4$. Por el contrario, en el segundo caso, aparecen muchas otras frecuencias.

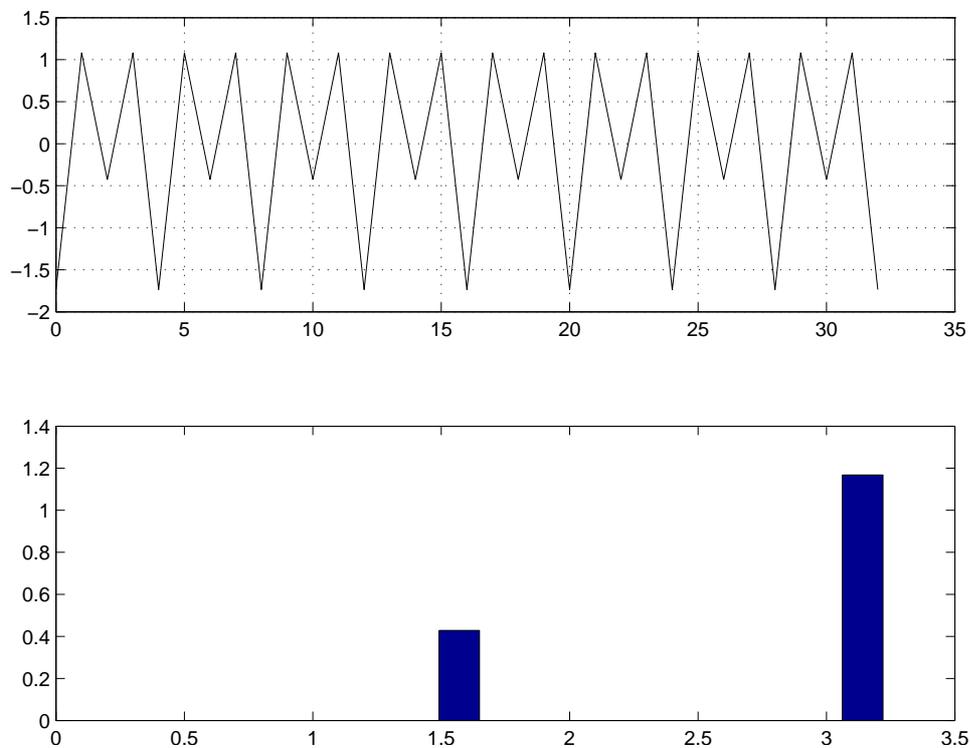


Figura 3: Una sucesión periódica y su representación espectral.

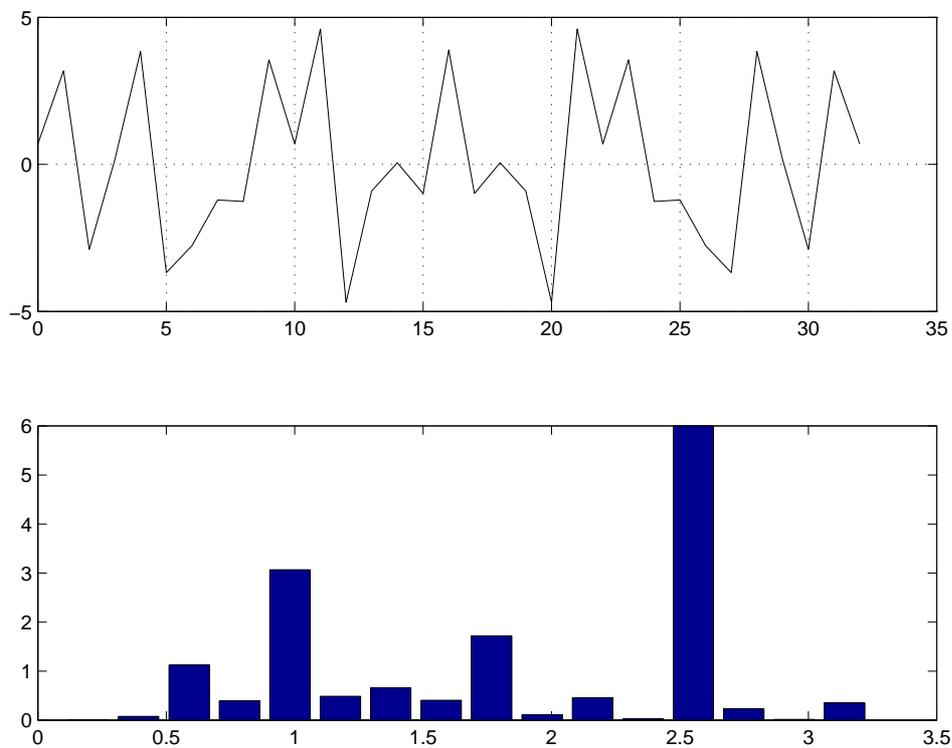


Figura 4: Una sucesión no periódica y su representación espectral.

2.1.2. Escalas

Uno de los aspectos donde hay más variedad en cuanto a la representación espectral es en las unidades que se emplean para expresar la frecuencia. Hasta ahora, para nosotros la frecuencia ω es el coeficiente que multiplica a t en las funciones trigonométricas esto significa que sus unidades son radianes/unidad de tiempo y sus valores recorren el intervalo $[0, \pi]$. Sin embargo, existen otras formas:

- (a) Normalizando de la forma $f = \omega/(2\pi)$, tenemos la frecuencia en ciclos/unidad de tiempo. El intervalo de representación es $[0, 1/2]$.
- (b) Menos habitual, pero también se puede encontrar la normalización $f = \omega/\pi$. Las unidades entonces serán medios ciclos por unidad de tiempo y el intervalo de representación es $[0, 1]$.

- (c) En datos mensuales o trimestrales, hay otra unidad de tiempo además de la de t , que es el año. Entonces se emplea a menudo la normalización $f = \omega/(2\pi f_0) = s\omega/(2\pi)$, donde s es igual a 4 o a 12 según los datos sean trimestrales o mensuales y $f_0 = 1/s$ es la frecuencia que corresponde a un ciclo por año expresada en ciclos por unidad de tiempo. De esta forma, las unidades en las que expresamos la frecuencia son ciclos por año y el intervalo donde se mueven las frecuencias es $[0, s/2]$. Esta es la forma en la que representa las frecuencias el programa TRAMO/SEATS.

Esto tiene como consecuencia que las frecuencias en las que está concentrado el espectro de las sucesiones periódicas son distintas según la normalización. En (a) están en $1/s, 2/s, \dots, 1/2$; en (b) están en $2/s, 4/s, \dots, 1$ y en (c) están en $1, 2, \dots, s/2$. Esto tiene consecuencias para el análisis de la estacionalidad, como veremos más adelante.

El extremo derecho del intervalo es la frecuencia de Nyquist. Por tanto, para cada normalización esa frecuencia se expresa de manera distinta. Mientras que en radianes/unidad de tiempo (a) la máxima frecuencia observable es π , en ciclos por año (b) la máxima frecuencia se expresa de distinta manera según que los datos sean mensuales o trimestrales: para datos mensuales es 6 ciclos por año y para datos trimestrales, 2 ciclos por año.

La escala en la que se expresa la energía puede ser tanto lineal como logarítmica. En el segundo caso habitualmente se emplea el logaritmo decimal y como unidad 1/10 (decibelio o dB). Es decir, que la energía espectral en decibelios es $10 \log_{10}(a_k^2 + b_k^2)$.

2.1.3. Solapamiento o *aliasing*

Como hemos dicho antes, en el espectro no aparecen frecuencias mayores de π rad/unidad de tiempo o medio ciclo por unidad de tiempo. Sin embargo, cuando la serie que observamos está afectada por oscilaciones de frecuencias mayores, éstas se manifiestan indirectamente, apareciendo en frecuencias más bajas asociadas, que se llaman *alias*.

Este fenómeno aparece ocasionalmente en imágenes en movimiento. Tanto es así que se emplea la frase *efecto de rueda de carro*. Todos hemos visto alguna vez cómo la rueda de un carro en una película parece moverse a una velocidad más baja de lo que debería o incluso en la dirección contraria. Las imágenes cinematográficas son sucesiones de fotogramas a la frecuencia de 24 por segundo.

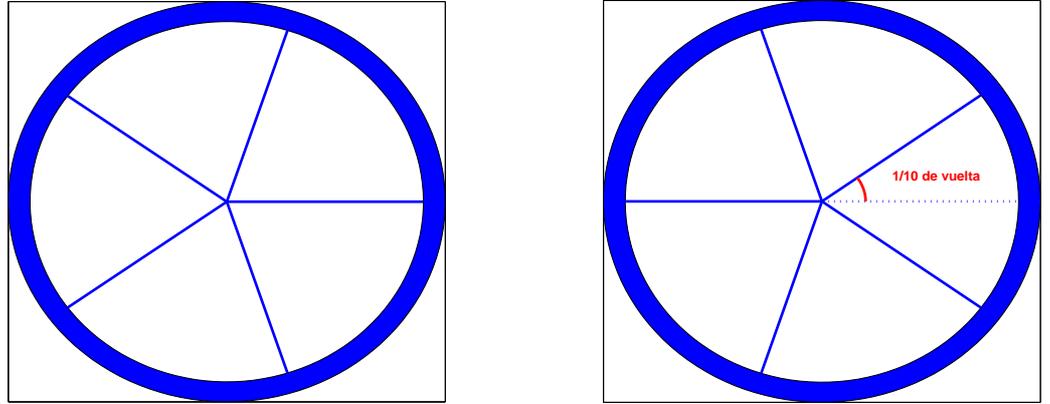


Figura 5: Solapamiento. Entre la imagen izquierda y la derecha ha pasado $1/24$ de segundo y la rueda ha girado 10 vueltas enteras y $1/10$ de vuelta, pero como no la observamos entre los dos fotogramas, nos parece que solo ha girado $1/10$ de vuelta.

Supongamos que una rueda de carro gira a una velocidad angular de $2\pi 24(10 + 1/10)$ rad/s. Esto significa que en $1/24$ s la rueda ha dado 10 vueltas y $1/10$ de vuelta (ver figura 5). Como nosotros no observamos el proceso entre dos fotogramas, nuestro cerebro interpreta que la rueda ha dado $1/10$ de vuelta solamente y por tanto que su velocidad angular es aproximadamente 100 veces menor de la que es en realidad. Si la velocidad angular es de $2\pi 24(10 - 1/10)$ rad/s, entonces la rueda parece girar lentamente en la dirección contraria al movimiento.

Con el espectro de una serie temporal ocurre lo mismo. Aquellas oscilaciones que se producen a una frecuencia mayor que medio ciclo por unidad de tiempo se atribuyen a una menor que se llama alias. En concreto, si $p = 1/f$ es el periodo de la oscilación expresado en las unidades de tiempo en las que está observada la serie, podemos calcular el periodo que corresponde a la frecuencia alias (ver Melis Maynar, 1992) como

$$p_a = \begin{cases} p_a = \frac{2}{2/p - [2/p]} & \text{mod}([2/p], 2) = 0 \\ p_a = \frac{2}{1 - 2/p + [2/p]} & \text{mod}([2/p], 2) \neq 0 \end{cases}$$

En el ejemplo de la rueda, la frecuencia expresada en fotogramas (o veinticuatroavos de segundo) es $10 + 1/10 = 101/10$, luego $p = 10/101$. Entonces, $\text{mod}([2/p], 2) = \text{mod}(20, 2) = 0$, luego

$$p_a = \frac{2}{2/p - [2/p]} = \frac{2}{202/10 - 20} = 10,$$

que corresponde justamente a la frecuencia que decíamos que se observa en la película.

2.2. De una serie temporal a un proceso estocástico

Vamos a ver cómo se traslada la representación espectral de series temporales al marco de procesos estocásticos. Supongamos que x_t es un proceso estacionario y su función de autocovarianza es γ_h . Si $\sum_h |\gamma_h|$ es convergente, entonces la serie

$$f(\omega) = \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} \gamma_h e^{ih\omega} \quad (4)$$

es absolutamente convergente uniformemente en $\omega \in \mathbb{R}$. La función $f(\cdot)$ se llama densidad espectral de x_t . Esta función es periódica de periodo 2π , luego basta considerarla en $(-\pi, \pi]$. Si la serie temporal es real, además $f(\omega)$ es par, luego podemos reducir el intervalo de estudio a $[0, \pi]$.

Por otra parte, podemos obtener la función de autocovarianza a partir de la densidad espectral:

$$\gamma_h = \int_{-\pi}^{\pi} e^{ih\omega} f(\omega) d\omega.$$

Por ejemplo, ver proposición 6.1 en Hamilton, 1994.

2.2.1. Estimación del espectro

A partir de la definición de densidad espectral (válida cuando esta existe, que es en la mayoría de los casos), se puede obtener de manera natural un estimador sustituyendo las covarianzas poblacionales por las muestrales de la forma

$$\hat{f}(\omega) = \frac{1}{2\pi} \sum_{|h| < r} \hat{\gamma}_h e^{ih\omega}, \quad (5)$$

con la limitación de que solo se puede incluir un número finito de covarianzas, aquí $2r - 1$.

2.2.2. El periodograma

Como máximo se puede fijar $r = T$ en (5) para una serie de longitud T , y en ese caso $\hat{f}(\omega)$ se llama periodograma². Por otra parte, **el periodograma coincide con la Transformada de Fourier Discreta de la serie temporal que vimos en 2.1.**

La expresión (5) nos puede hacer sospechar que las propiedades del periodograma pueden ser no muy buenas. La estimación de la h -ésima covarianza con T datos se obtiene promediando $T - h$ valores. Si incluimos covarianzas hasta $T - 1$, esto significa que la estimación de las primeras y las últimas se hará con muy pocos valores y será en general muy mala. En efecto, el periodograma como estimador de la densidad espectral es:

- **asintóticamente insesgado**, pero
- **inconsistente**, ya que la distribución de $\hat{f}(\omega)$ cuando $T \rightarrow \infty$, se distribuye aproximadamente como $f(\omega) \cdot (1/2)\chi_2^2$. Además para distintas frecuencias, los valores del periodograma son asintóticamente independientes, lo que explica la apariencia ruidosa del periodograma (ver figura 6).

2.2.3. Otras formas de estimación espectral

Para suavizar las variaciones del periodograma, se puede emplear la convolución con una función adecuada –llamada *ventana*– en el espacio de la frecuencia. Una función así, digamos $W(\cdot)$, deberá tener su densidad concentrada en cero y cumplir que $\int W(\omega)d\omega = 1$. Entonces, podemos calcular la estimación suavizada como

$$\begin{aligned}\hat{f}_W(\omega) &= \int W(\lambda)\hat{f}(\omega - \lambda)d\lambda = \int W(\lambda)\frac{1}{2\pi} \sum_{|h|<r} \hat{\gamma}_h e^{ih(\omega-\lambda)}d\lambda = \\ &= \frac{1}{2\pi} \sum_{|h|<r} \hat{\gamma}_h e^{ih\omega} \int W(\lambda)e^{-ih\lambda}d\lambda = \frac{1}{2\pi} \sum_{|h|<r} w(h)\hat{\gamma}_h e^{ih\omega}.\end{aligned}$$

La última expresión la podemos entender como una modificación del periodograma en la que las distintas autocovarianzas están pesadas de acuerdo a la función $w(h)$. Como decíamos antes, cuando $|h|$ es grande, la estimación de la

²Para algunos autores, la palabra periodograma se reserva para este estimador en concreto, pero a veces se usa de manera más general, para modificaciones como las que describimos en 2.2.3.

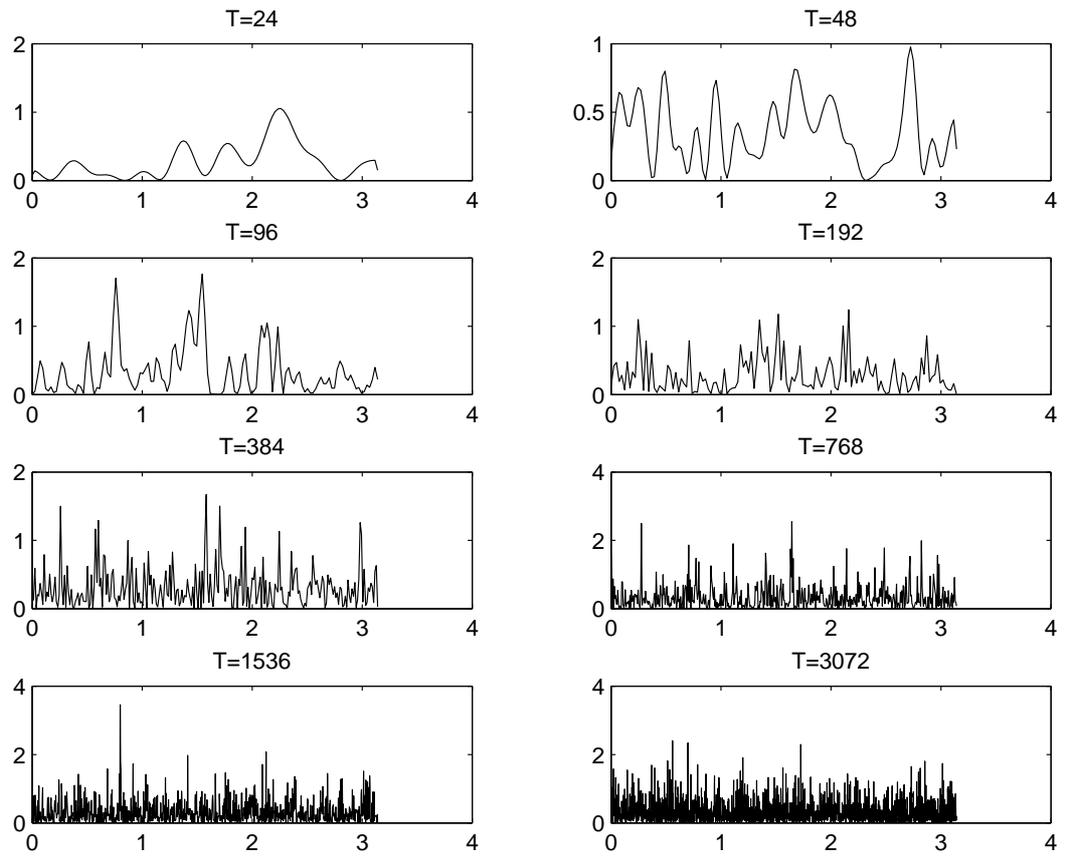


Figura 6: El periodograma no es un estimador consistente.

h -ésima autocovarianza, es mala. Por tanto, una buena función $w(h)$ debería pesar más las covarianzas para h pequeña.

Ejemplo 1. Una posibilidad es excluir los retardos h más largos, es decir

$$w(h) = \begin{cases} 1 & |h| \leq r \\ 0 & |h| > r \end{cases}$$

Esto es equivalente a elegir

$$W(\omega) = \frac{1}{2\pi} \frac{\sin(r + 1/2)\lambda}{\sin \lambda/2}.$$

Ejemplo 2. También se emplea la ventana triangular, que para T par tiene la forma:

$$w(h) = \begin{cases} \frac{2h-1}{T} & 1 \leq h \leq \frac{T}{2} \\ \frac{2(T-h)+1}{T} & 1 + \frac{T}{2} \leq h \leq T \end{cases}$$

Hay muchas otras posibilidades para la función $W(\cdot)$. Otros métodos de estimación empleados son el de Welch y el *multitaper*.

2.3. Representación espectral de un proceso estocástico*

En este apartado introducimos conceptos adicionales que sirven para presentar el teorema de representación espectral que da sentido al análisis espectral, ya que permite interpretar la densidad espectral en términos del propio proceso y no solo de su función de autocovarianza.

No todos los procesos estacionarios cumplen la condición $\sum_h |\gamma_h| < \infty$ y en consecuencia, no todos tienen una densidad espectral. Sin embargo, todos tienen una *medida espectral*:

Teorema 1. Si x_t es estacionario, entonces existe una única medida F finita en $(-\pi, \pi]$ tal que

$$\gamma_h = \int_{-\pi}^{\pi} e^{ih\omega} dF(\omega).$$

Si la medida F tiene una densidad $f(\omega)$, entonces $f(\omega)$ es la densidad espectral de x_t .

Ejemplo 3. La densidad espectral de un proceso de ruido blanco es constante $f(\omega) = (2\pi)^{-1}\gamma_0$.

Ejemplo 4. Sea $x_t = A \cos \lambda t + B \sin \lambda t$, donde A, B son variables aleatorias incorreladas de media cero y varianza σ^2 . Entonces

$$\gamma_h = \sigma^2 (\cos \lambda t \cos \lambda(t+h) + \sin \lambda t \sin \lambda(t+h)) = \sigma^2 \cos h\lambda = \sigma^2 \frac{1}{2} (e^{ih\lambda} + e^{-ih\lambda})$$

y en consecuencia F es una medida discreta concentrada en dos puntos, de manera que $F(\{-\lambda\}) = F(\{\lambda\}) = \sigma^2/2$.

Definición 1. Una medida aleatoria con incrementos ortogonales Z es una colección $\{Z(B) : B \in \mathcal{B}\}$ de variables aleatorias complejas de media cero, indexada por los conjuntos de Borel en $(-\pi, \pi]$ definida en cierto espacio de probabilidad tal que para alguna medida de Borel finita μ en $(-\pi, \pi]$,

$$\text{cov}(Z(B_1), Z(B_2)) = \mu(B_1 \cap B_2).$$

Empleando esta definición podemos expresar el resultado principal de esta sección.

Teorema 2. Para todo proceso estacionario de media cero x_t con distribución espectral F hay una medida aleatoria con incrementos ortogonales Z relativa a F tal que $\{Z(B) : B \in \mathcal{B}\} \subset L^2(x_t : t \in \mathbb{Z})$ tal que

$$x_t = \int_{-\pi}^{\pi} e^{it\omega} dZ(\omega). \quad (6)$$

La definición de la integral estocástica $\int \cdot dZ(\omega)$ y una descripción más detallada de la representación espectral se puede encontrar en Koopmans (1995).

3. Filtros lineales y espectro

Uno de los puntos fuertes del análisis espectral de procesos estocásticos se manifiesta cuando empleamos filtros lineales. Un filtro lineal \mathcal{F} es una aplicación que transforma procesos de la forma

$$x_t \mapsto y_t = \mathcal{F}(B)x_t = \sum_{k \in \mathbb{Z}} a_k B^k x_t = \sum_{k \in \mathbb{Z}} a_k x_{t-k}, \quad (7)$$

donde $Bx_t = x_{t-1}$. En la terminología de series temporales esto también se llama media móvil.

Ejemplo 5. Una media móvil simétrica es un filtro lineal:

$$y_t = \frac{1}{3}x_{t-1} + \frac{1}{3}x_t + \frac{1}{3}x_{t+1}.$$

Ejemplo 6. Si $a_k = 0$ para $k < 0$, entonces \mathcal{F} es un filtro causal. Cuando decimos que un proceso x_t satisface un modelo MA, eso es equivalente a decir que x_t es el resultado de aplicarle un filtro lineal causal a un proceso de ruido blanco ε_t . Un MA(q) tiene la forma:

$$y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}.$$

Ejemplo 7. Un alisado exponencial tiene la forma

$$y_t = \alpha x_{t-1} + (1 - \alpha)y_t,$$

donde $\alpha \in (0, 1)$. Podemos reescribir esa ecuación como $y_t = (1 - \alpha)^t y_0 + \alpha \sum_{k=0}^{t-1} (1 - \alpha)^k x_{t-k-1}$. Por tanto, cuando t sea grande, se parecerá mucho a un filtro de la forma

$$y_t = \sum_{k=0}^{\infty} \alpha (1 - \alpha)^k B^{k+1} x_t.$$

Cada filtro lineal tiene una función asociada llamada **función de transferencia**. Si lo representamos como en (7), su función de transferencia tiene la forma $\psi(z) = \sum_{k \in \mathbb{Z}} a_k z^k$. Esta función nos va a permitir obtener el resultado clave de esta sección.

Teorema 3. Si x_t e y_t tienen densidades espectrales $f_x(\omega)$ y $f_y(\omega)$ respectivamente y se cumple $y_t = \mathcal{F}(B)x_t$, entonces

$$f_y(\omega) = |\psi(e^{-i\omega})|^2 f_x(\omega). \quad (8)$$

La función $\omega \mapsto |\psi(e^{-i\omega})|^2$ se llama función de ganancia (aunque también se la llama en ocasiones, función de transferencia; nosotros reservamos ese nombre para $\psi(z)$). La función de ganancia nos dice como actúa el filtro sobre x_t en el dominio de la frecuencia. Si $|\psi(e^{-i\omega})|^2 < 1$, el filtro atenúa la frecuencia ω , mientras que si $|\psi(e^{-i\omega})|^2 > 1$ la amplifica.

Ejemplo 8. Este resultado también nos permite calcular el espectro de un proceso ARMA. Supongamos que x_t satisface

$$\phi(B)x_t = \theta(B)\varepsilon_t$$

donde ϕ tiene sus raíces fuera del círculo unidad. Entonces

$$x_t = \frac{\theta(B)}{\phi(B)} \varepsilon_t$$

Como la densidad espectral del ruido blanco es constante $\sigma^2/(2\pi)$, entonces la de x_t es

$$f(\omega) = \left| \frac{\theta(e^{-i\omega})}{\phi(e^{-i\omega})} \right|^2 \frac{\sigma^2}{2\pi}. \quad (9)$$

Si x_t es un ARMA(1,1), $(1-0,5B)x_t = (1-0,25B)\varepsilon_t$ entonces su espectro tiene la forma de la figura 7.

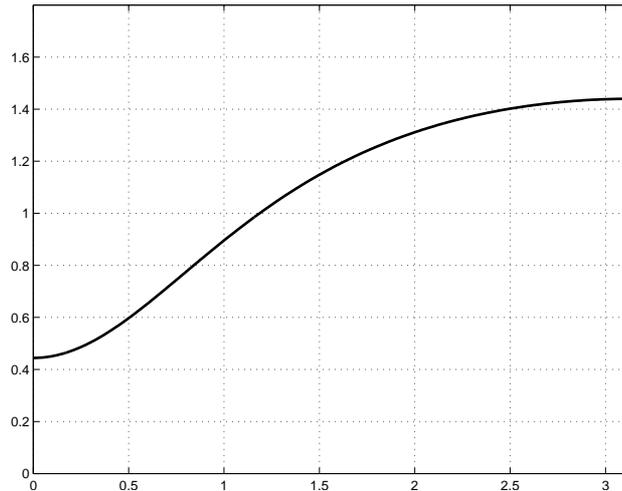


Figura 7: Densidad espectral de un ARMA $(1-0,5B)x_t = (1-0,25B)\varepsilon_t$.

Ejemplo 9. Supongamos que x_t es el del ejemplo anterior. Vamos a ver cómo afecta a su espectro el filtro que tiene como función de transferencia $\psi(z) = (1/2) + (1/2)z^2$. La función de ganancia del filtro es $(1/4)|1 + e^{-2i\omega}|^2$. En la figura 8 está representada la función de ganancia y se observa cómo atenúa las frecuencias en torno a $\pi/2$ o un cuarto de ciclo por unidad de tiempo. Debajo tenemos la densidad espectral de $y_t = \mathcal{F}(B)x_t$. Se ve cómo se parece a la de x_t en los extremos, pero están atenuadas las frecuencias cercanas a $\pi/2$.

3.1. Espectro de un proceso integrado

La representación espectral introducida en la sección 2.3 se demuestra solo para procesos estacionarios. Sin embargo en el tratamiento de series económicas a menudo hay que tratar con procesos integrados y por tanto, no estacionarios. La extensión del análisis espectral a procesos no estacionarios es problemática.

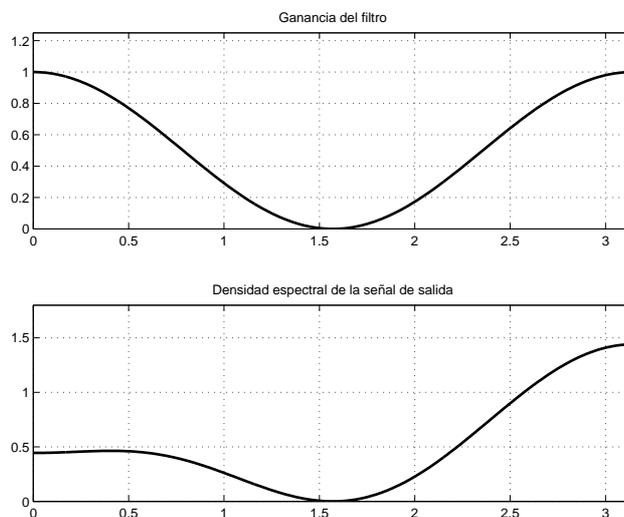


Figura 8: Ganancia del filtro y densidad espectral de la señal $y_t = \mathcal{F}(B)x_t$.

Por tanto, aquí haremos una descripción informal, sin entrar en los detalles técnicos.

Podemos emplear la expresión (9) para generalizar el espectro a los procesos integrados. Supongamos que x_t satisface un modelo ARIMA que podemos escribir también como $\phi(B)x_t = \theta(B)\varepsilon_t$ pero permitiendo ahora que $\phi(z)$ tenga raíces unitarias. Entonces su densidad espectral $f(\omega) = |\theta(e^{-i\omega})/\phi(e^{-i\omega})|^2$ tendrá asíntotas verticales en algunas frecuencias.

Ejemplo 10. *Supongamos que $(1 - B)x_t = (1 - \theta B)\varepsilon_t$. Entonces, su densidad espectral es*

$$f(\omega) = \left| \frac{1 - \theta e^{-i\omega}}{1 - e^{-i\omega}} \right|^2.$$

Como el denominador se anula en $\omega = 0$, la densidad tiende a infinito cuando $\omega \rightarrow 0$.

Si un proceso cumple un modelo ARIMA estacional, podemos analizar su espectro empleando la ecuación (9). Supongamos que x_t cumple

$$\phi(B)\Phi(B^s)(1 - B)^d(1 - B^s)^D x_t = \theta(B)\Theta(B^s)\varepsilon_t.$$

Entonces la densidad espectral de x_t es

$$f(\omega) = \left| \frac{\theta(e^{-i\omega})\Theta(e^{-si\omega})}{(1 - e^{-i\omega})^d(1 - e^{-si\omega})^D\phi(e^{-i\omega})\Phi(e^{-si\omega})} \right|^2 \frac{\sigma^2}{2\pi}.$$

La densidad espectral tendrá asíntotas verticales en los puntos donde se anula el denominador, es decir, donde

- $1 - e^{-i\omega} = 0 \Rightarrow \omega = 0.$
- $1 - e^{-si\omega} = 0 \Rightarrow s\omega = 2\pi k$ con $k = 0, 1, \dots$. Por tanto, en ciclos por año: $0, 1, \dots, s/2.$

Por otra parte, los polinomios ϕ y Φ pueden tener raíces que sin ser unitarias estén cerca de la unidad. En esos puntos, el denominador se hará pequeño y se pueden encontrar máximos del espectro incluso aunque el modelo no tenga diferencias estacionales. Por ejemplo, si x_t cumple el modelo $(1 - \Phi \cdot B^s)x_t = \varepsilon_t$, y $\Phi \sim 1$, entonces $1 - \Phi \cdot e^{-si\omega} \sim 0$ cuando $s\omega = 2\pi k$. Por tanto aparecerán máximos para las mismas frecuencias, aunque no tan pronunciadas como cuando hay diferencias.

4. Interpretación del espectro de series estacionales

En las secciones anteriores hemos presentado algunas herramientas para el análisis espectral. A continuación vamos a emplearlas para describir posibles apariencias del espectro de una serie y a qué tipo de procesos corresponden. También veremos algunos ejemplos de cómo afectan los filtros lineales al espectro de una serie estacional.

4.1. Series periódicas

Como hemos explicado en 2.1.1, una serie periódica de periodo s tiene su espectro concentrado en unas pocas frecuencias. En el análisis de series temporales estacionales, es habitual normalizar la frecuencia de manera que estas frecuencias sean $0, 1, \dots, s/2$. Por tanto:

- Para series trimestrales, las frecuencias estacionales son 1 y 2 ciclos por año.

- Para series mensuales, las frecuencias estacionales son 1, 2, 3, 4, 5, 6 ciclos por año.

Si una serie x_t tiene una componente estacional determinista, se reflejará en su espectro en esas frecuencias. En la figura 9 vemos: (1) una serie que se obtiene como un seno de periodo 12, (2) un seno de periodo 6 y (3) una serie periódica de periodo 12, pero no sinusoidal. Todas ellas se representan junto con su periodograma. Las líneas verticales en el periodograma indican las frecuencias estacionales.

4.2. Series con estacionalidad, pero no periódicas

En la figura 10 representamos una serie que satisface un modelo AR estacional de la forma $x_t = 0,9x_{t-12} + \varepsilon_t$. También presenta máximos en las frecuencias estacionales, pero en el gráfico de la serie no se observa la estacionalidad tan claramente como en la serie periódica, ya que el perfil anual no se mantiene intacto de un año a otro, sino que va cambiando a lo largo del tiempo. Sin embargo, se aprecia con claridad en el periodograma.

4.3. Series con tendencia

En la figura 11 representamos primero una serie con tendencia determinista (lineal). La estacionalidad está oculta en el espectro porque la asíntota de la densidad en la frecuencia cero, que corresponde a la tendencia, es muy dominante. Sin embargo, si aplicamos una diferencia regular a la serie, anulamos este efecto y permitimos que la estacionalidad se haga visible. Esto nos sugiere que **cuando las series tengan tendencia, puede ser conveniente comprobar la presencia de estacionalidad en la serie diferenciada regularmente.**

4.4. Efecto calendario en el espectro

Como se dijo en 2.1.3, los fenómenos asociados a frecuencias mayores que dos ciclos por año en datos trimestrales o seis en datos mensuales no se pueden observar directamente, pero sí indirectamente a través de sus frecuencias alias. Si una variable que se observa, por ejemplo, mensualmente se puede ver como acumulación de datos diarios, entonces es posible que se vea afectada por fluctuaciones semanales. La frecuencia asociada a la semana no se puede definir

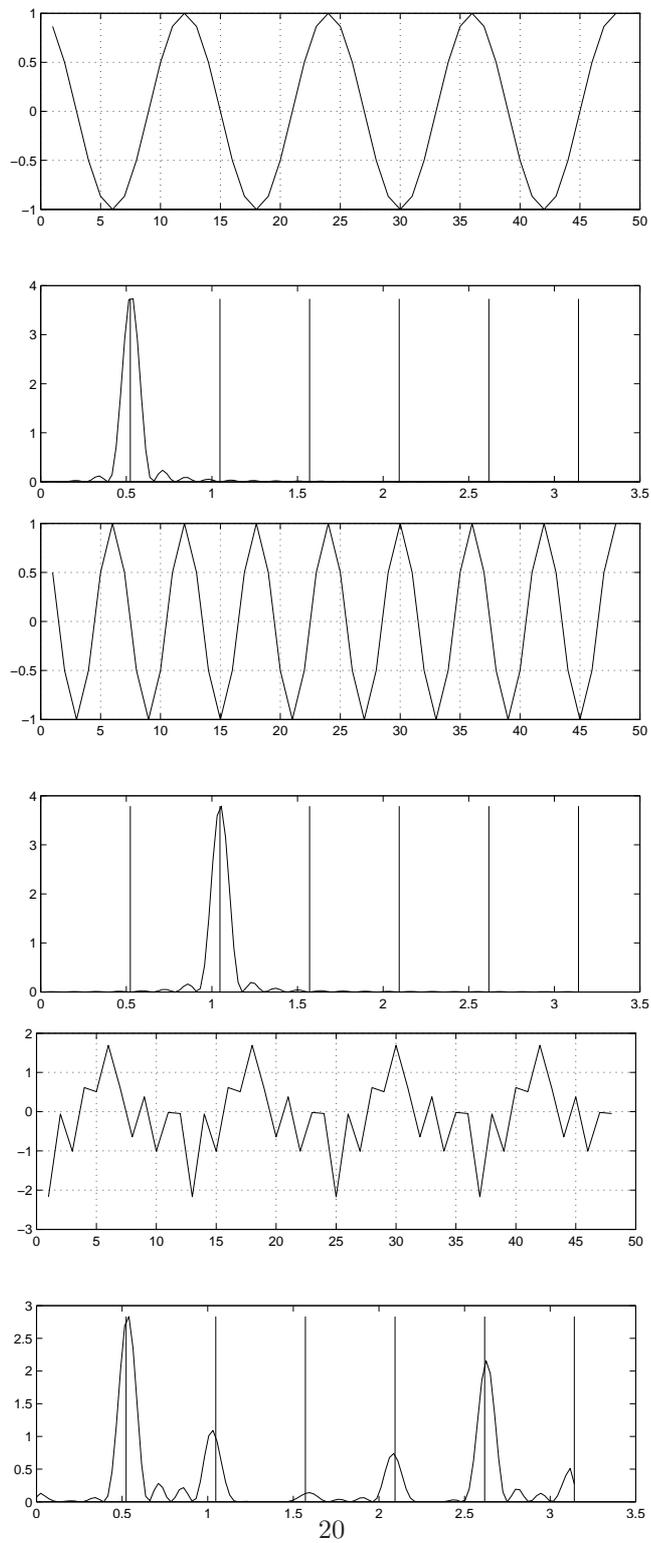


Figura 9: Dos series sinusoidales de frecuencias $2\pi/12$ y $2\pi(2/12)$ (periodo 12 y 6) y una serie periódica de periodo 12 no sinusoidal.

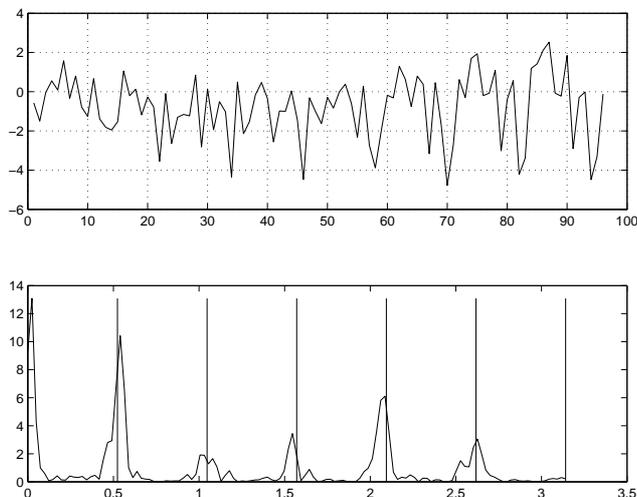


Figura 10: Serie AR estacional.

de forma totalmente precisa en meses, porque los meses tienen distinta duración. Sin embargo, si hacemos la aproximación de suponer que todos tienen un número de días promedio, digamos, $(4 \cdot 365 + 1)/48 = 30,4375$, entonces, el periodo es aproximadamente $p = 7/30,4375 = 0,23$. Como $2/p = 8,6964$, entonces $\text{mod}([2/p], 2) = 0$ y por tanto el periodo alias es

$$p_a = \frac{2}{1 - 2/p + [2/p]} = 2,8718.$$

En consecuencia, la frecuencia alias es de 4.1786 ciclos por año o 2.19 rad/mes. Para datos trimestrales, tendríamos una frecuencia de 0.5357 ciclos por año o 0.28 rad/trimestre.

4.5. Interpretación de la escala

Cuando se examina el espectro de una serie ajustada para detectar estacionalidad o efecto calendario residual, al interpretar los picos que se observan debe tenerse en cuenta lo siguiente:

- Cuando comparamos el resultado de ajustar la misma serie con distintos métodos o modelos, la magnitud de los picos en frecuencias estacionales o asociadas al efecto calendario nos orienta sobre la importancia de la componente estacional que queda.

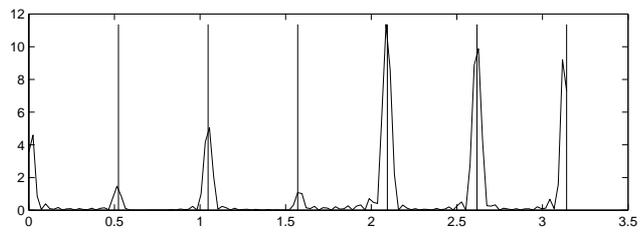
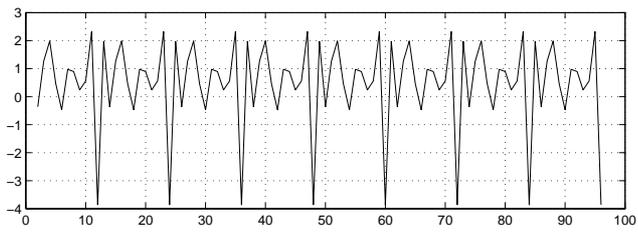
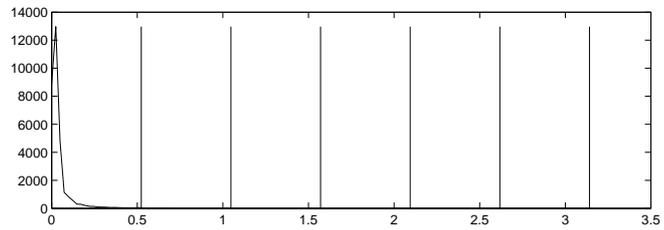
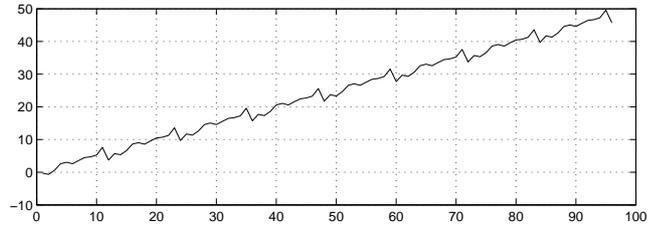


Figura 11: Serie con tendencia determinista y la misma serie diferenciada regularmente.

- Por el contrario, al comparar el resultado del ajuste de distintas series, el tamaño de los picos no nos informa sobre lo bien ajustadas que estén las series en términos relativos.

Vamos a ilustrar esta idea con dos ejemplos. En la figura 12, tenemos el periodograma de las series ajustadas que se obtienen a partir de una misma serie bruta mediante dos modelos distintos. En el panel superior vemos que queda un pico en la frecuencia asociada al ciclo semanal de aproximadamente cinco, mientras que en el inferior su magnitud es de poco más de dos. En el primer caso, la variabilidad debida a la componente de calendario residual es de un 13% de la de la serie original, mientras que en el segundo es de aproximadamente 5%. El mayor tamaño del pico espectral indica realmente un peor ajuste.

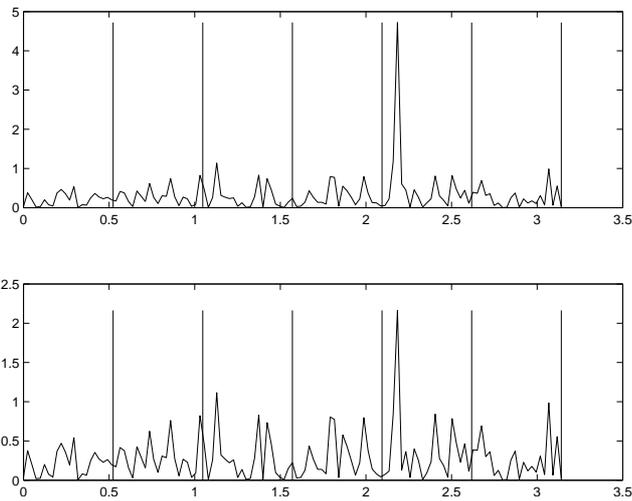


Figura 12: Escala: periodograma de una serie ajustada mediante dos modelos distintos.

Por el contrario, en la figura 13 vemos el resultado del ajuste de dos series distintas. El pico de frecuencia 2.19 en la segunda es aproximadamente la mitad de alto que el de la primera. Sin embargo, ambos representan una fracción similar de la variabilidad total de las dos series (alrededor del 8%). Esto significa que el efecto calendario residual es de una importancia semejante en ambos casos.

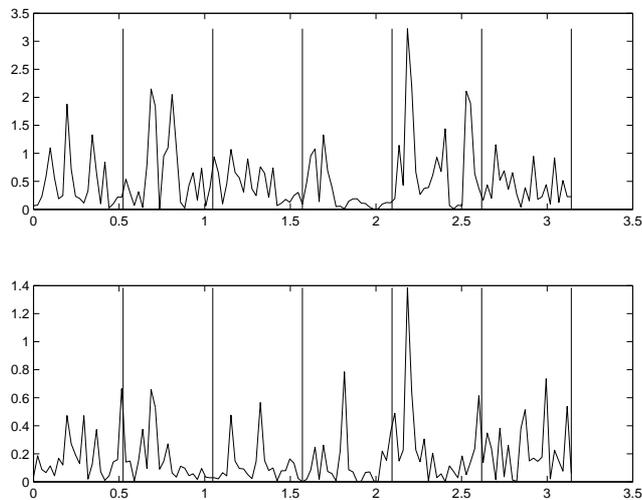


Figura 13: Escala: periodogramas de dos series ajustadas.

4.6. Efecto de los filtros sobre la estacionalidad

En la sección 3 ya vimos de forma general cómo los filtros lineales transforman el espectro de una serie. Veamos ahora algunos ejemplos con series estacionales.

4.6.1. Efecto de la diferencia estacional

En la figura 14 podemos ver una serie generada según el modelo $x_t = x_{t-12} + \varepsilon_t$. A esta serie le aplicamos el filtro $1 - B^{12}$. Debajo de la serie original, representamos la serie filtrada. A la derecha de cada una vemos el periodograma. Se puede ver cómo el filtro ha mitigado las frecuencias estacionales, pero también la frecuencia cero. Esto es consecuencia de que la función de ganancia se anula en las frecuencias estacionales y en el cero.

4.6.2. Efecto de la media móvil

Supongamos ahora que no queremos eliminar las frecuencias cercanas a cero para conservar la tendencia y para empleamos que $1 - B^{12} = (1 - B) \cdot (1 + B + \dots + B^{11})$. Eso sugiere filtrar con una media móvil $(1 + B + \dots + B^{11})/11$. En la figura 15 podemos ver el efecto de este filtro sobre una serie generada

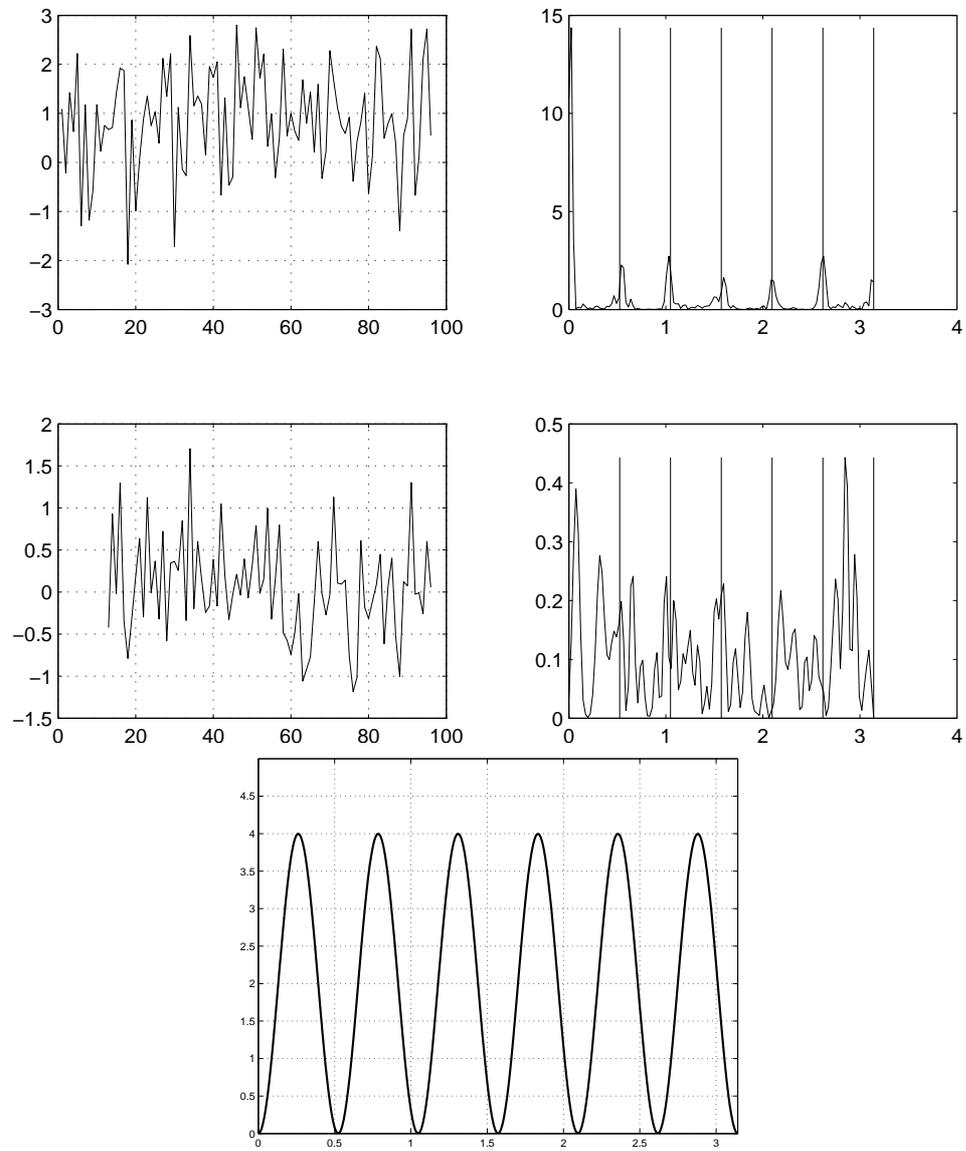


Figura 14: Series original y filtrada y función de ganancia del filtro $1 - B^{12}$.

con el mismo modelo que el ejemplo anterior. Se puede observar que además de mitigar las frecuencias estacionales, también se han mitigado las frecuencias que hay entre ellas y en particular las frecuencias altas. Para confirmar esto solo hay que observar la función de ganancia. Esto se puede interpretar de la siguiente manera: con la media móvil además de eliminar la componente estacional, se reduce la componente irregular.

4.6.3. Efecto de la diferencia regular

En 4.3, ya vimos cómo la diferencia regular, aplicada a una serie con tendencia determinista podía sacar a la luz la parte estacional al eliminar del espectro las frecuencias bajas asociadas a la tendencia. Veamos ahora qué pasa con una serie que es integrada pero no tiene una tendencia determinista. En este caso, el modelo es

$$(1 - B)(1 - B^{12})x_t = \varepsilon_t - 0,9\varepsilon_{t-12}.$$

Podemos ver cómo también en este caso la estacionalidad se aprecia mejor en el espectro de la serie diferenciada³

4.6.4. Ejemplos con datos reales

En la figura 17, representamos el Índice de Producción Industrial en el periodo 2000-2010. Debajo están las series obtenidas tras aplicar ∇ , ∇_{12} y $\nabla \cdot \nabla_{12}$. Vemos como la serie original tiene gran parte de la energía espectral concentrada cerca de la frecuencia cero. En la serie diferenciada regularmente se puede como el espectro se concentra en las frecuencias estacionales. En la serie diferenciada regular y estacionalmente se ve el pico de frecuencia 2.19 que corresponde al efecto calendario.

En la figura 18 se puede ver una representación análoga para la serie de Cifras de Negocios de la Industria en el periodo 2002-2010.

³Por otra parte, hay que tener en cuenta que el filtro distorsiona la estructura estacional de la serie. Observando la función de ganancia se puede ver que la primera frecuencia estacional, que corresponde a un ciclo por año está mitigada, mientras que las otras están amplificadas según factores crecientes.

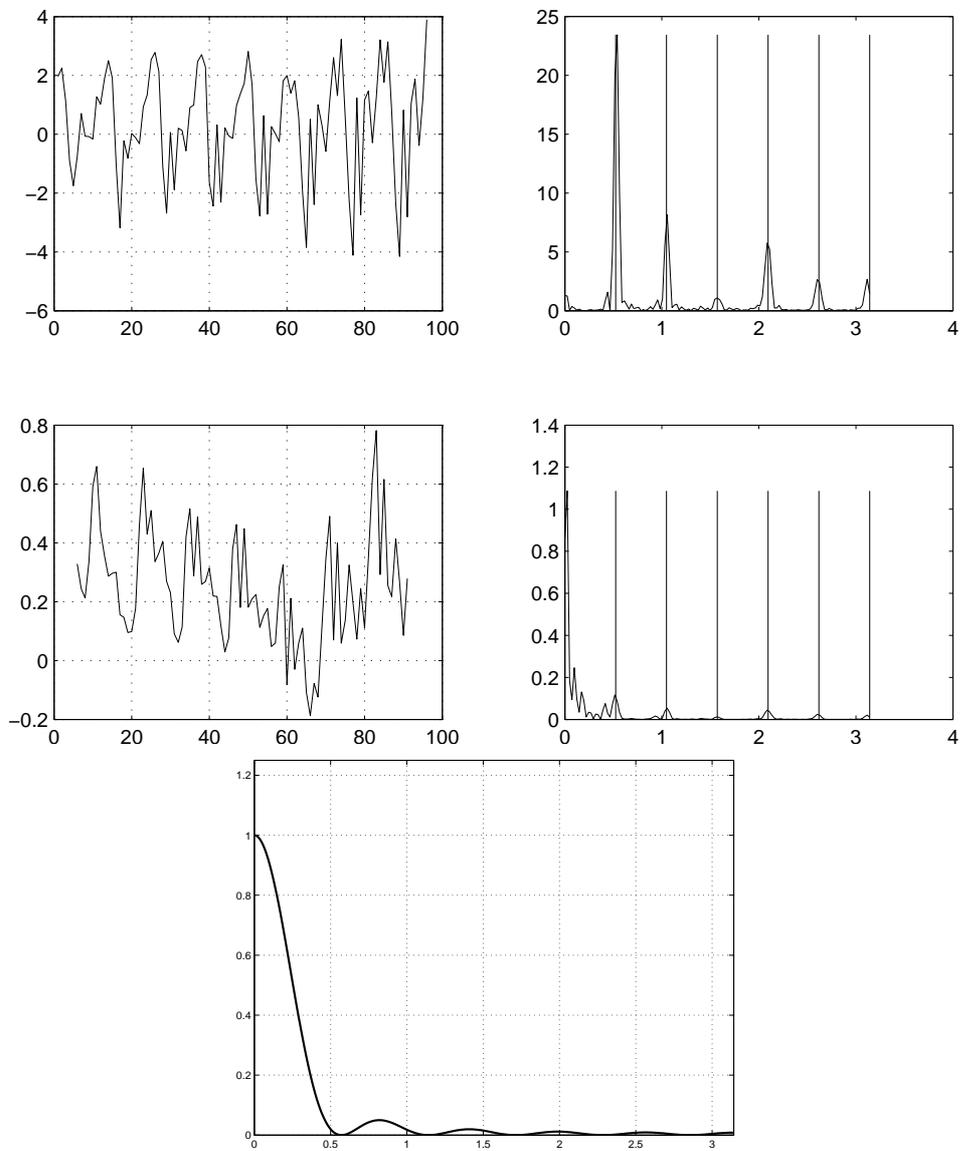


Figura 15: Series original y filtrada y función de ganancia del filtro $(1 + B + \dots + B^{11})$.

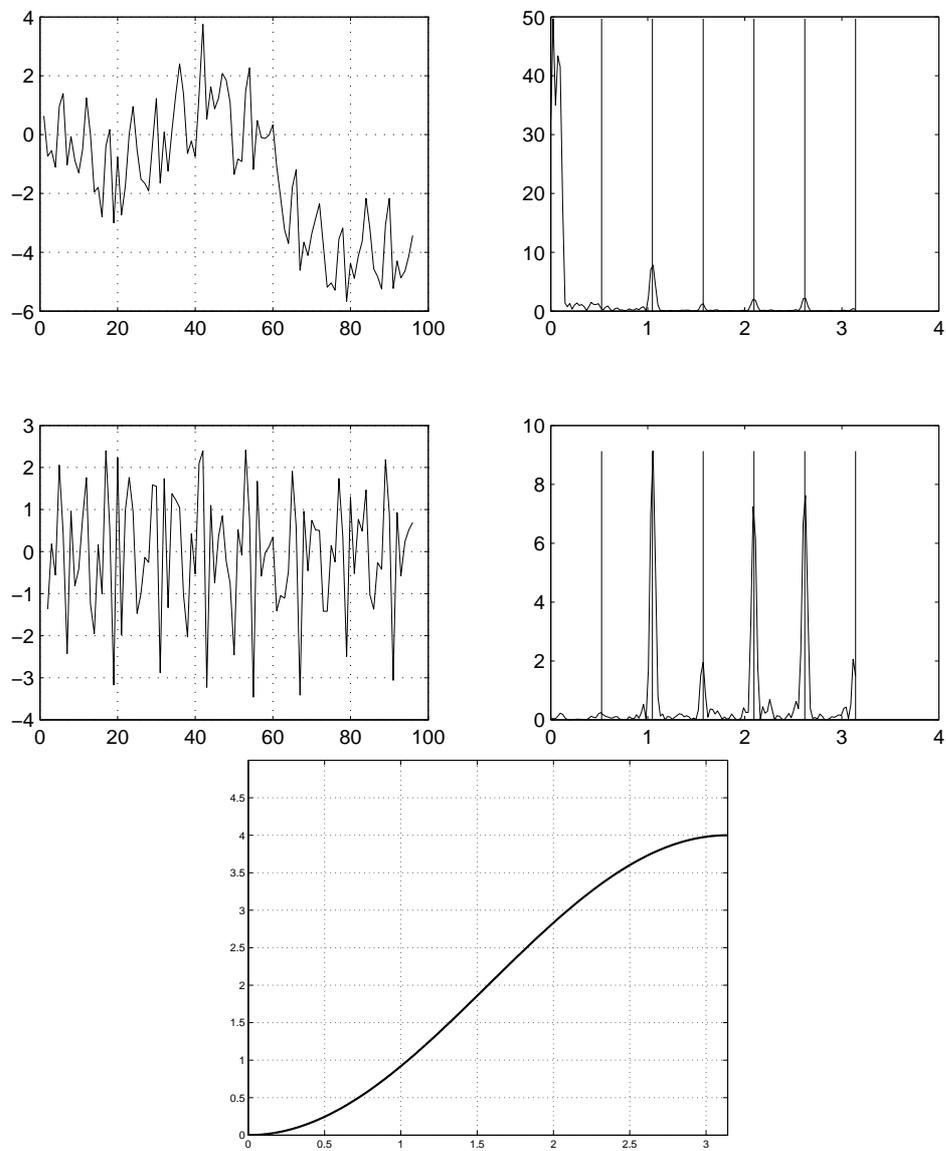


Figura 16: Series original y filtrada y función de ganancia del filtro $1 - B$.

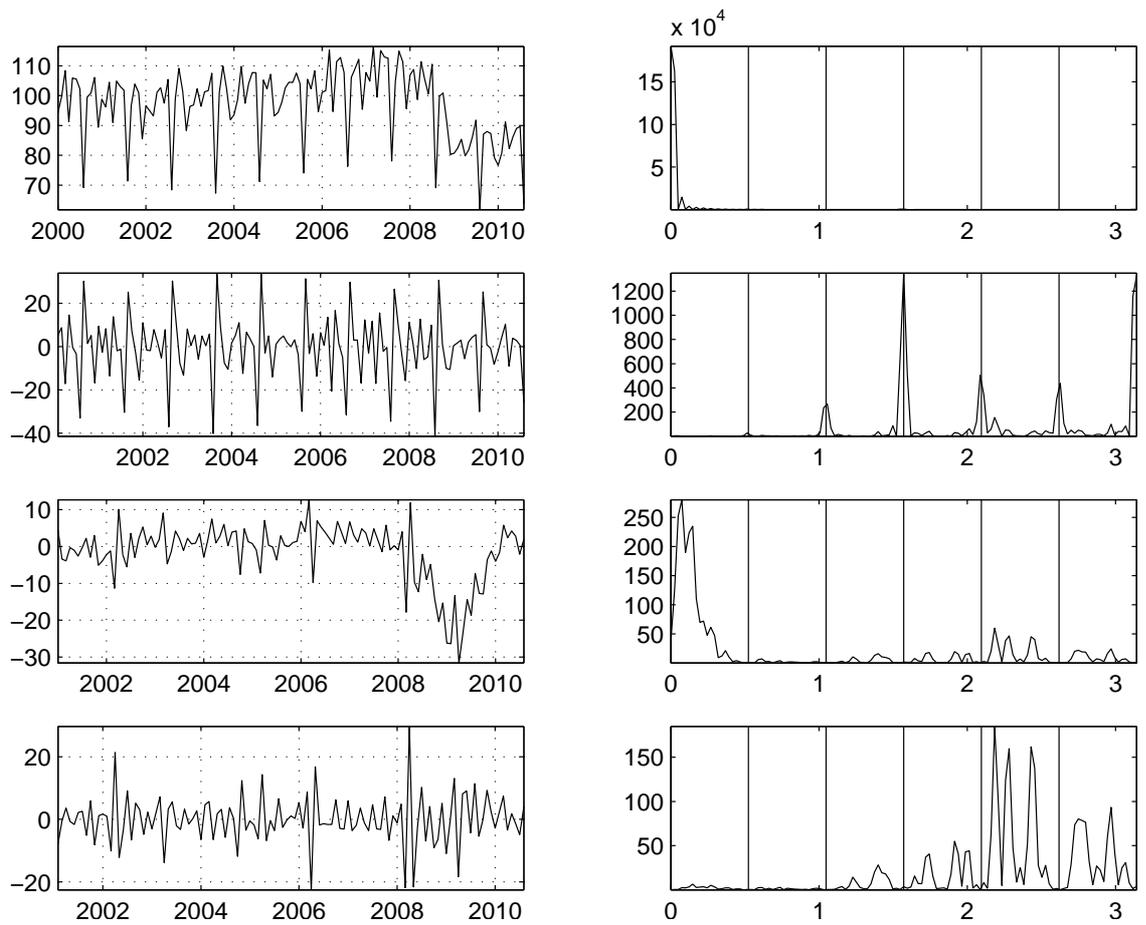


Figura 17: IPI: serie original y series diferenciadas con ∇ , ∇_{12} y $\nabla \cdot \nabla_{12}$, todas ellas junto con su periodograma.

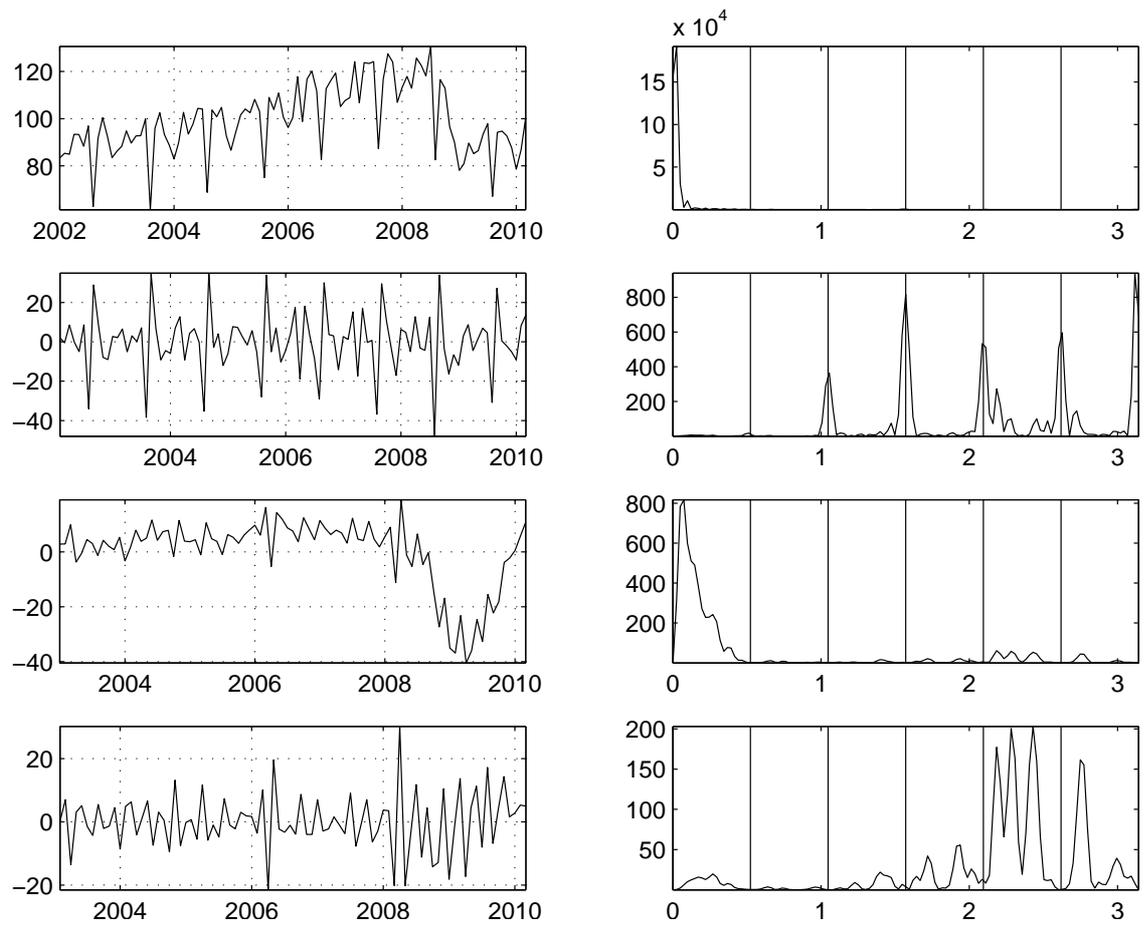


Figura 18: ICN: serie original y series diferenciadas con ∇ , ∇_{12} y $\nabla \cdot \nabla_{12}$, todas ellas junto con su periodograma.

Referencias

- [1] Brillinger, D. R. y Krishnaiah P. R. eds. (1983) *Handbook of Statistics 3. Time Series in the Frequency Domain*. Amsterdam: North Holland.
- [2] Findley, D. Frequency Domain Analyses of SEATS and X11/12-ARIMA Seasonal Adjustment Filters for Short and Moderate-Length Time Series. Working paper, Statistical Research Division. U.S. Census Bureau.
- [3] Hamilton, J. D. (1994) *Time Series Analysis*, Princeton, New Jersey: Princeton University Press.
- [4] Hillmer, S. C. y Tiao, G. C. (1982) An ARIMA-Model-Based Approach to Seasonal Adjustment. *Journal of the American Statistical Association* 77, 63–70.
- [5] Koopmans, L. K. (1995) *The spectral analysis of time series*, New York: Academic Press.
- [6] Melis Maynar, F. (1992) Agregación temporal y solapamiento o aliasing. *Revista Estadística Española* 34, 309-346.
- [7] Pollock, D. S. G., The spectral analysis of a stationary stochastic process. Notas de clase.
URL: <http://www.le.ac.uk/users/dsgp1/COURSES/LEIMETZ/SPECTRAL.pdf>.
- [8] Van der Vaart, A. W. *Time series: course notes*.