

Análisis espectral

Fernando Borraz

Extracción de señales - dECON 2017

Análisis espectral

- Hasta ahora analizamos las propiedades de y_t en el dominio del tiempo
- $y_t = \alpha + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t$
- Ahora veremos el valor de y_t como una suma ponderada de funciones periódicas de la forma $\cos(\omega t)$ y $\sin(\omega t)$ donde ω denota una frecuencia particular
- $y_t = \alpha + \int_0^\Pi \alpha(\omega) \cos(\omega t) d\omega + \int_0^\Pi \delta(\omega) \sin(\omega t) d\omega$
- El objetivo es extraer el ciclo y determinar cuan importantes son los ciclos de diferentes frecuencias para explicar la evolución de Y

Análisis espectral

- Veremos que la representación de Cramer de los procesos estacionarios en cov brinda una forma matemáticamente rigurosa de expresar la noción de que las series macro contienen componentes asociados con fluctuaciones de diferentes frecuencias:

⇒ los movimientos lentos o de baja frecuencia asociados a la tendencia

⇒ los de frecuencia media que se vinculan al ciclo

⇒ y finalmente los movimientos rápidos o de alta frecuencia relacionados a factores estacionales e irregulares

Análisis espectral

- Esta teoría de extracción de ciclos se deriva de los números complejos y se conoce como dominio de frecuencia o análisis espectral
- Toda serie estacionaria tiene una representación en el dominio del tiempo y otra en el dominio de la frecuencia
- Veremos el teorema de la representación espectral que es la versión en el frecuencia del dominio del teorema de Wold
- Caracterizaremos las propiedades del espectro poblacional y su estimación

Análisis espectral

- La dependencia temporal en series económicas puede estudiarse en el dominio del tiempo o en el dominio de la frecuencia
- En el primer caso, es necesario partir explicitando el proceso estocástico que genera los datos como la suma de un componente sistemático y una perturbación aleatoria impredecible. La parte sistemática suele formularse como una simple combinación lineal de observaciones y perturbaciones pasadas, mientras que la parte impredecible se asimila a un proceso estocástico ruido blanco
- El análisis en el dominio de la frecuencia, o análisis espectral, interpreta el proceso estocástico de manera que el comportamiento en el tiempo de la variable es el resultado de la combinación (adición) de ciclos de distinta amplitud y duración

Análisis espectral

- El análisis espectral permite estudiar de qué forma las diferentes periodicidades o frecuencias contribuyen a la explicación de la variabilidad total de la serie
- Si bien los dos tipos de análisis aportan la misma información sobre la serie observada desde diferentes perspectivas, algunos aspectos del comportamiento de la serie pueden visualizarse más fácilmente en el dominio del tiempo y otros en el de la frecuencia
- El enfoque tradicional para el tratamiento de las series económicas ha sido en el dominio del tiempo. No obstante, la utilización del análisis espectral se revela como más conveniente cuando el analista enfrenta un problema que requiere la estimación de uno o varios de los componentes inobservables

Análisis espectral

- Por un lado, cada procedimiento de descomposición implica la aplicación de diferentes filtros o transformaciones de los datos originales, los cuáles tienen fuertes consecuencias en cuanto a la atenuación de algunos ciclos o al realzamiento de otros, siendo relevante analizar sus efectos, de modo de no arribar a conclusiones erróneas respecto a la señal de interés
- El análisis espectral proporciona las herramientas para conocer cómo son afectadas las series en este sentido, aportando información relevante acerca de los efectos de los filtros que se aplican a series de tiempo macroeconómicas

Análisis espectral

- Por otro lado, el análisis del comportamiento cíclico de una serie macroeconómica en el dominio del tiempo oculta un conjunto de influencias económicas que provienen de ciclos de diferentes amplitudes y longitudes, que pueden evidenciarse realizando la descomposición de las fluctuaciones en diferentes frecuencias cíclicas
- Antes de desarrollar el análisis espectral repasaremos números complejos

Números complejos: breve repaso

- Sean a, b dos números reales. Un número complejo z es:

- $z = a + bi$

- $i = \sqrt{-1}$

- a es la parte real y b es la parte imaginaria

Números complejos

- Para cada número complejo podemos definir su complejo conjugado (z^*) y el modulo (ρ_z):
- $z^* = a - bi$
- $\rho_z = \sqrt{a^2 + b^2}$
- El modulo es la distancia del número desde 0, su largo y $\rho_z = \rho_z^*$ para cada número complejo

Números complejos

- Sea ω el grado del ángulo formado por la representación de z y el eje real. Esto se denomina frecuencia
- Recordando de la definición de seno y coseno
- $\sin(\omega) = \frac{b}{\rho_z} \Rightarrow b = \rho_z \sin(\omega)$
- $\cos(\omega) = \frac{a}{\rho_z} \Rightarrow a = \rho_z \cos(\omega)$
- $z = \rho_z \cos(\omega) + \rho_z \sin(\omega)i$
- $z = \rho_z [\cos(\omega) + i \sin(\omega)]$ Esto es la representación de z en las coordenadas polares

Números complejos: resultados útiles

- 1. La formula de Euler nos permite escribir cualquier número complejo en terminos de la función exponencial
 $\Rightarrow \cos(\omega) + i \sin(\omega) = e^{i\omega} \Rightarrow z = \rho_z e^{i\omega}$
- 2. Teorema de De Moivre's: $z^k = \rho_z^k e^{ki\omega}$
 $\Rightarrow z^k = \rho_z^k [\cos(\omega) + i \sin(\omega)]$
- 3. Teorema de Riesz-Fischer: Sea $\{a_j\}_{j=-\infty}^{\infty}$ una secuencia de números complejos que satisfacen la condición de ser square-summability:
 $\Rightarrow \sum_{j=-\infty}^{\infty} (a_j)^2 < \infty$. El teorema establece que existe una función compleja conocida como la transformación de Fourier de la secuencia tal que: $f: [-\pi, \pi] \longrightarrow R$

$$f(\omega) = \sum_{j=-\infty}^{\infty} a_j e^{-i\omega j}$$

Dos importantes propiedades de la transformación de Fourier

- 1. La transformación de Fourier de una combinación lineal de secuencias es la combinación lineal de la transformación de Fourier de las secuencias
- 2. Dado f , la secuencia puede ser recuperada a través de la siguiente fórmula de inversión: $a_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\omega) e^{i\omega j} d\omega$. Se expresa cada elemento de la secuencia en términos de frecuencias

Función generatriz de autocovarianzas (FGA)

- El nexo entre el dominio del tiempo y el análisis espectral es la FGA, cuyo análisis es la base para el estudio de las propiedades cíclicas de series de tiempo
- La función de autocovarianzas y/o función de autorrelaciones contienen toda la información sobre la dependencia temporal de una serie
- En el dominio del tiempo la autocov de orden s , que se denota γ_j es
$$\gamma_j = COV (y_t, y_{t-j}) = E \{ [y_t - E (y_t)] [y_{t-j} - E (y_{t-j})] \}$$
- Para series estacionarias: *i*) $E (y_t) = E (y_{t-j})$, *ii*) $\gamma_j = \gamma_{-j}$

Función generatriz de autocovarianzas

- De manera general, una función generatriz registra la información sobre cierta secuencia
- Dada la secuencia a_0, a_1, a_2, \dots , probablemente compuesta por infinitos términos, la función generatriz de dicha secuencia se define como:

$$a(z) = \sum_{j=0}^{\infty} a_j z^j$$

- La variable z no necesariamente tiene interpretación. Puede decirse que es portadora de información sobre la forma en que se genera la secuencia
- Por lo que una forma de resumir información contenida en una secuencia de autocov es mediante el uso de una función de valor escalar denominada función generatriz de autocov

Función generatriz de autocovarianzas

- Sea el proceso estocástico real estacionario $\{y_t\}_{t=-\infty}^{\infty}$ con varianza γ_0 y autocov γ_j
- Se define $\{\gamma_j\}_{j=-\infty}^{\infty}$ como la secuencia de autocov. Si dicha secuencia es absolutamente summable, la función generatriz de autocov es: $g_y(z) = \sum_{j=-\infty}^{\infty} \gamma_j z^j$, donde el argumento de la función, z , es un número complejo

Función generatriz de autocovarianzas

- Ejemplo MA(1)
- $y_t = \alpha + \theta \varepsilon_{t-1} + \varepsilon_t$
- $E(y_t) = \alpha \quad \text{VAR}(y_t) = \gamma_0 = \sigma^2 (1 + \theta^2)$
- $\gamma_1 = \text{COV}(y_t, y_{t-1}) = E\{[y_t - E(y_t)][y_{t-1} - E(y_{t-1})]\} = E[(y_t - \alpha)(y_{t-1} - \alpha)] =$
- $= E[(\alpha + \theta \varepsilon_{t-1} + \varepsilon_t - \alpha)(\alpha + \theta \varepsilon_{t-2} + \varepsilon_{t-1} - \alpha)] = E[(\theta \varepsilon_{t-1} + \varepsilon_t)(\theta \varepsilon_{t-2} + \varepsilon_{t-1})] = \theta E(\varepsilon_{t-1}^2) = \theta \sigma^2$
- $\gamma_j = 0$ para $j \geq 2$

Función generatriz de autocovarianzas

- $g_y(z) = \sum_{j=-\infty}^{\infty} \gamma_j z^j = \gamma_{-1} z^{-1} + \gamma_0 z^0 + \gamma_1 z^1 =$
- $= \theta \sigma^2 z^{-1} + \sigma^2 (1 + \theta^2) z^0 + \theta \sigma^2 z^1 = \sigma^2 (1 + \theta z) (1 + \theta z^{-1})$
- Para el AR(1): $g_y(z) = \frac{\sigma^2}{(1-\phi z)(1-\phi z^{-1})}$

Densidad espectral y representación espectral

- Un concepto relacionado es el espectro de un proceso $\{y_t\}_{-\infty}^{\infty}$
- Sea el proceso square-summable, por el teorema de Riesz-Fischer podemos obtener la transformación de Fourier de la secuencia de autocov (que es igual a la FGM evaluada en $z=e^{-i\omega}$) :

$$\implies f_y(\omega) = \sum_{k=-\infty}^{\infty} \gamma_k e^{-i\omega k} = g_y(e^{-i\omega})$$

- El espectro es: $s_y(\omega) = \frac{1}{2\pi} f_y(\omega) = g_y(e^{-i\omega}) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_k e^{-i\omega k}$
- El espectro también puede ser escrito como:

$$s_y(\omega) = \frac{1}{2\pi} \left[\gamma_0 + \sum_{k=1}^{\infty} \gamma_k (e^{-i\omega k} + e^{i\omega k}) \right]$$

Densidad espectral y representación espectral

- A pesar del término i , la función es real. Para verlo, usar la representación de los números complejos en las coordenadas polares y el teorema de De Moivre's para escribir $e^{i\omega} = \cos(\omega) + i \sin(\omega)$ y $e^{-i\omega} = (e^{i\omega})^{-1} = \cos(-\omega) + i \sin(-\omega)$
- Usando: $\cos(-\omega) = \cos(\omega)$ y $\sin(-\omega) = -\sin(\omega) \implies$
- $e^{-i\omega} = \cos(\omega) - i \sin(\omega) \implies e^{i\omega} + e^{-i\omega} = 2 \cos(\omega)$
- Espectro = $s_y(\omega) = \frac{1}{2\pi} \left[\gamma_0 + 2 \sum_{k=1}^{\infty} \gamma_k \cos(\omega k) \right]$

Espectro

- Simétrico en 0 (por lo que se lo suele representar en $[0, \pi]$ más que en $[-\pi, \pi]$). Esto es pues $\gamma_k = \gamma_{-k}$ y $\cos(-\omega k) = \cos(\omega k)$
- El espectro se conoce también como densidad espectral y se puede mostrar que es no negativo e integra 1 sobre su dominio
- $\cos(\omega\tau + 2\pi k\tau) = \cos(\omega\tau)$ para cualquier k, τ entero \implies el espectro es una función periódica de $\omega : S(\omega) = S(\omega + 2\pi k)$. Conocer el valor de $S(\omega)$ para cualquier valor de ω en $[0, \pi]$ implica conocer $S(\omega)$ para cualquier valor de ω
- Como $\cos(\omega)$ va de 1 a -1 cuando ω va de 0 a $\pi \implies$ el espectro decrece en ω monotonamente si $\theta < 0$ y monotonamente crece si $\theta > 0$ en $(0, \pi)$

Espectro

- Ejemplo 1: ruido blanco $s_y(\omega) = \frac{1}{2\pi}\sigma^2$
- Ejemplo 2: MA(1) $s_y(\omega) = \frac{1}{2\pi}\sigma^2 (1 + \theta e^{-i\omega})(1 + \theta e^{i\omega}) = \frac{1}{2\pi}\sigma^2 (1 + \theta e^{i\omega} + \theta e^{-i\omega} + \theta^2 e^{-i\omega} e^{i\omega}) = \frac{1}{2\pi}\sigma^2 (1 + 2\theta \cos(\omega) + \theta^2)$

Espectro

- Ejemplo 3: AR(1)

- $s_Y(\omega) = \frac{1}{2\pi} \sigma^2 \frac{1}{(1-\phi e^{-i\omega})(1-\theta e^{i\omega})} =$

- $= \frac{\sigma^2}{2\pi} \frac{1}{(1+\phi^2-2\phi \sin(\omega))}$

- Si $\phi > 0$ el espectro es monótono decreciente mientras que si $\phi < 0$ es creciente en ω en $[0, \pi]$

Espectro poblacional

- Por la transformación de Fourier podemos escribir las autocov de la serie usando la información del espectro mediante:
- $\gamma_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} s_y(\omega) e^{i\omega k} d\omega$
- $Var(y_t) = \gamma_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} s_y(\omega) d\omega$
- Podemos calcular la varianza asociada con los componentes periódicos aleatorios con frecuencia menor o igual que ω_1 mediante $\frac{1}{2\pi} \int_{-\omega_1}^{\omega_1} s_y(\omega) d\omega$
- $\frac{1}{2\pi} \int_{-\omega_1}^{\omega_1} s_y(\omega) d\omega$ es la proporción de la varianza de y_t que puede ser atribuida a un componente inobservable con frecuencia igual o menor que ω_1

Teorema representación espectral

- Cramer y Leadbetter (1967)
- Todo proceso estacionario en cov $\{Y_t\}_{-\infty}^{\infty}$ con autocov absolutely summable:
$$\sum_{k=-\infty}^{\infty} |\gamma_k| = \gamma_0 + 2 \sum_{k=1}^{\infty} |\gamma_k| < \infty$$
- para cualquier frecuencia fija $\omega \in [0, \pi]$, se pueden definir dos V.A. $\alpha(\omega), \delta(\omega)$ incorrelacionadas con media 0 e igual varianza tales que se puede representar el proceso como:

$$y_t = \alpha + \int_0^{\pi} \alpha(\omega) \cos(\omega t) d\omega + \int_0^{\pi} \delta(\omega) \sin(\omega t) d\omega$$

- Es una formulación alternativa a la del dominio tiempo y expresa el proceso como una suma ponderada de un continuo de componentes cíclicos que difieren por la frecuencia con la que completan sus ciclos

Análisis espectral

- La representación de Cramer de los procesos estacionarios en cov brinda una forma matemáticamente rigurosa de expresar la noción de que las series macro contienen componentes asociados con fluctuaciones de diferentes frecuencias:
 - ⇒ los movimientos lentos o de baja frecuencia asociados a la tendencia
 - ⇒ los de frecuencia media que se vinculan al ciclo
 - ⇒ y finalmente los movimientos rápidos o de alta frecuencia relacionados a factores estacionales e irregulares
- La pregunta clave es ¿qué frecuencias son dominantes y cuáles de menor importancia para explicar los movimientos de y_t ?
- Una respuesta sencilla a dicha cuestión la proporciona la función de densidad espectral o espectro poblacional

Análisis espectral

- De esta forma, el espectro constituye un recurso para realizar la descomposición de la varianza de una serie por sus frecuencias. El área debajo del espectro en el intervalo $[-\pi, \pi]$, es igual a la varianza de y

- $$\int_{-\pi}^{\pi} S_y(\omega) d\omega = \int_0^{\pi} 2S_y(\omega) d\omega = \gamma_0$$

- Esta propiedad brinda alguna idea sobre la interpretación de la varianza en el dominio del tiempo, ya que la misma es la suma del espectro a lo largo de todas las frecuencias entre $-\pi$ y π . La varianza de una serie económica se distribuye desigualmente entre las frecuencias -excepto en el caso de un proceso ruido blanco-, de forma que el componente de crecimiento, el componente de ciclo económico (business cycle) y el componente estacional contribuyen de manera diferente a la variabilidad de la serie analizada

Análisis espectral

- Una formulación más general es

- $$\int_{-\pi}^{\pi} S_y(\omega) e^{i\omega\tau} d\omega = \int_0^{\pi} 2S_y(\omega) \cos(\omega\tau) d\omega = \gamma_{\tau}$$

- Así, el espectro de un proceso estocástico contiene la misma información que la función generatriz de autocovarianzas, ya que el mismo es simplemente una combinación lineal de las autocovarianzas. Sin embargo, la diferencia fundamental es que la función de densidad espectral arroja luz sobre la importancia de los componentes cíclicos para diferentes frecuencias

Análisis espectral

- Puede definirse una función acotada no decreciente $F(\omega)$ llamada función de distribución espectral que cumple:

- $$g_y(\tau) = \int_{-\infty}^{\infty} e^{i\omega\tau} F(d\omega)$$

- donde $g_y(\tau)$ es la función generatriz de momentos. La función de distribución espectral determina una medida $F(A)$, llamada distribución espectral de la serie de tiempo, donde $S_y(\omega) = F'(\omega)$ y sobre la banda de frecuencias A se cumple que

- $$F(A) = \int_A S_y(\omega) d\omega$$

Análisis espectral

- El cociente $\frac{F(A)}{\gamma_0}$ proporciona información sobre la contribución de los componentes asociados a las frecuencias incluidas en A a la varianza del proceso
- Así, al visualizar el gráfico de la función de densidad espectral de una serie, un pico puede interpretarse como un indicador de un componente cíclico con un período o longitud de onda constante e igual a $\frac{2\pi}{\omega}$ donde ω es la frecuencia correspondiente (esto es pues el valor coseno se repite cada $\frac{2\pi}{\omega}$ períodos)
- $p = \frac{2\pi}{\omega}$ (p periodicidad del ciclo y ω frecuencia)

Análisis espectral

- Recordar que el valor del $\cos(\omega t)$ se repite cada $\frac{2\pi}{\omega}$ períodos
- Por lo que una frecuencia de ω se corresponde con un período de $\frac{2\pi}{\omega}$
- Si por ej. $\omega = \frac{3\pi}{2}$ el ciclo se repite a si mismo cada $p = \frac{2\pi}{\omega} = \frac{2\pi}{\frac{3\pi}{2}} = \frac{4}{3}$ períodos
- El período más corto de un ciclo es uno que se repite cada $2 = \frac{2\pi}{\pi}$ períodos
- La frecuencia más baja es $\omega = \frac{2\pi}{T}$ que se corresponde a un período de T
- Si un ciclo se repite cada más de T períodos no se puede inferir nada con T observaciones
- Frecuencia $\omega_j = \frac{2\pi j}{T}$ se corresponde con un período de $\frac{2\pi}{\omega_j} = \frac{2\pi}{\frac{2\pi j}{T}} = \frac{T}{j}$

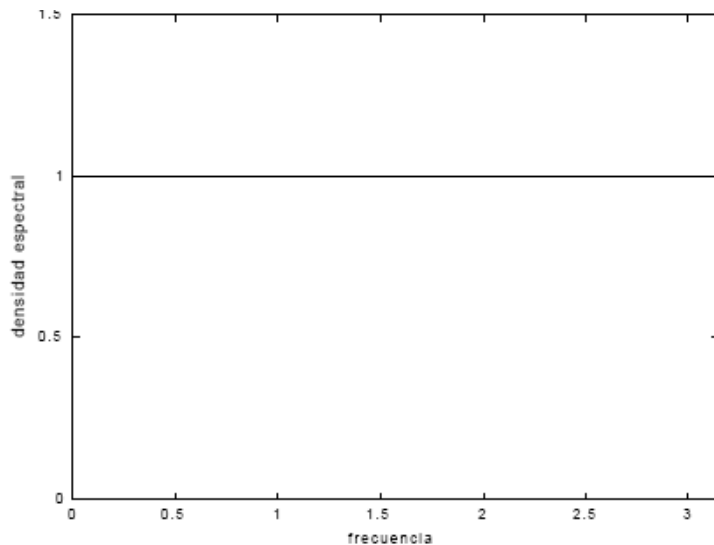
Frecuencias y períodos en series trimestrales y anuales

Periodicidad	Frecuencia			
	Datos trimestrales		Datos anuales	
infinita	0	0	0	0
16 años	$\pi/32$	0.0982	$\pi/8$	0.3927
12 años	$\pi/24$	0.1309	$\pi/6$	0.5236
10 años	$\pi/20$	0.1571	$\pi/5$	0.6283
8 años	$\pi/16$	0.1964	$\pi/4$	0.7854
6 años	$\pi/12$	0.2618	$\pi/3$	1.0472
5 años	$\pi/10$	0.3142	$\pi/2.5$	1.2566
4 años	$\pi/8$	0.3927	$\pi/2$	1.5708
3 años	$\pi/6$	0.5236	$\pi/1.5$	2.0944
2 años	$\pi/4$	0.7854	π	3.1416
estacional	$\pi/2$	1.5708	n.c.	n.c.
6 meses	π	3.1416	n.c.	n.c.

Casos particulares del espectro

- 1) Si el proceso generador de los datos es un ruido blanco el espectro es una recta horizontal. Un ruido $Y_t = \varepsilon_t$ tiene varianzas $\gamma_0 = \sigma^2$ y autocovarianzas $\gamma_\tau = 0$ para $\tau \neq 0$
- Así, la función generatriz de momentos es $g_y(z) = \sigma^2$ y el espectro es constante e igual a $s_y(e^{-i\omega}) = \frac{\sigma^2}{2\pi}$
- El gráfico del espectro de un ruido blanco con varianzas 2π se presenta en la próxima figura. Se caracteriza por ser plano en el nivel $\frac{\sigma^2}{2\pi}$ para todas las frecuencias. De aquí viene el término ruido blanco: así como una luz blanca, consiste en un número infinito de frecuencias cuyo peso es el mismo
- ¿Cuánto es el área debajo de la densidad espectral entre las frecuencias $-\pi$ y π ?

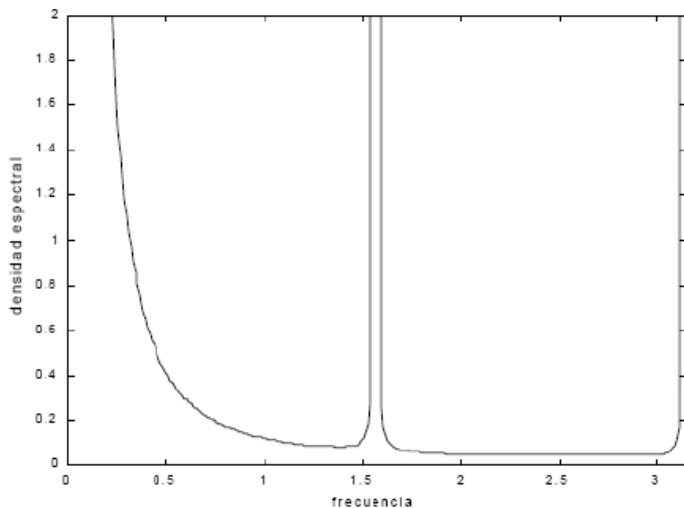
Función de densidad espectral de un ruido blanco



Análisis espectral

- 2) El espectro de una serie trimestral con estacionalidad y tendencia presenta tres picos marcados en las frecuencias 0 , $\frac{\pi}{2}$ y π
- Esto significa en términos de períodos, que los componentes que se asocian a un período infinito, anual y dos trimestres son importantes para explicar la variabilidad total de la serie en cuestión, respectivamente. Justamente, el período infinito se asocia a una señal suave o con fluctuaciones lentas (frecuencia cero), que es característica esencial del componente de tendencia
- Por otra parte, las series trimestrales estacionales tienen la particularidad de presentar regularidades o ciclos que se repiten casi sistemáticamente cada año y al interior del año cada dos trimestres. Esto es representado por los picos en las frecuencias $\frac{\pi}{2}$ y π

Densidad espectral de serie con tendencia y estacionalidad



Análisis espectral

- Si la serie contiene una raíz unitaria, el pico del espectro será infinito en la frecuencia 0
- Sea el AR(1): $y_t = \phi y_{t-1} + \varepsilon_t$
- El espectro es: $s_y(\omega) = \frac{\sigma^2}{2\pi(1-\phi e^{-i\omega})(1-\phi e^{i\omega})}$
- $\lim_{\phi \rightarrow 1} S_y(0) = \infty$

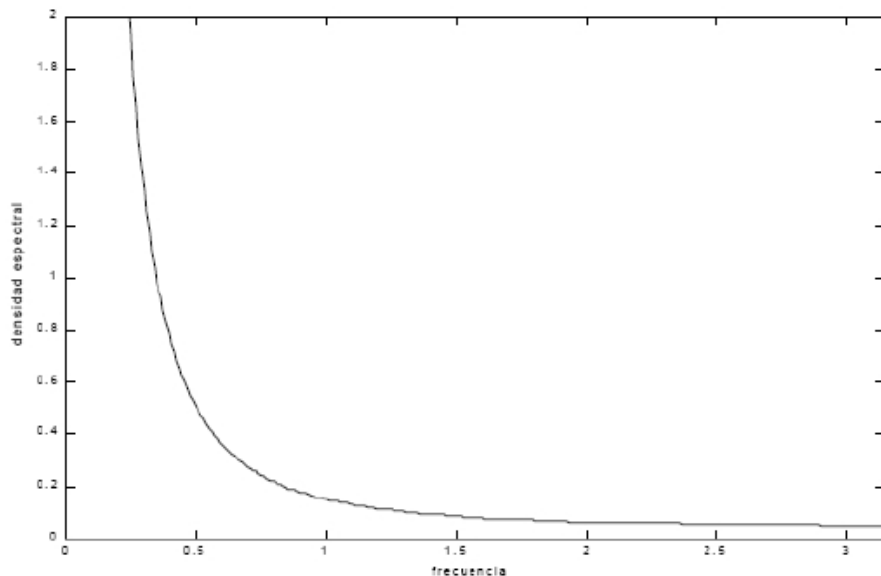
Análisis espectral

- Una observación importante es que la frecuencia más alta (período más bajo) sobre la que se tiene información directa es π , la que se conoce como la frecuencia de Nyquist
- Con datos trimestrales no se pueden detectar ciclos con una mayor frecuencia que dos trimestres, $\omega = \pi$, problema que se relaciona con el tema de agregación en el tiempo o muestro
- Este se refiere al proceso por el cual una señal continua se mide a través de una secuencia de números discretos, por ejemplo de periodicidad trimestral
- El inconveniente más serio relacionado a esto se denomina aliasing, e implica que si una serie contiene ciclos con una frecuencia mayor que la trimestral (por ej. mensual) estos serán imputados a ciclos con frecuencia entre 0 y π

Análisis espectral

- 3) Granger (1966) plantea que la mayoría de las series macro anuales presentan una "forma espectral típica"
- Su característica más evidente es la elevada importancia de los componentes asociados con las frecuencias bajas (aún cuando la tendencia del proceso hay sido previamente removida)

Forma típica de densidad espectral de variable macro anual



Estimación espectro poblacional: 1) Periodograma muestral

- El espectro poblacional fue expresado en función de la secuencia $\{\gamma_j\}_{j=0}^{\infty}$ que son los momentos poblacionales de segundo orden
- Sea ahora, un conjunto de T observaciones $\{y_1, y_2, \dots, y_T\}$
- Se pueden calcular hasta $T-1$ autocov muestrales a partir de

- $$\hat{\gamma}_j = \left\{ \begin{array}{l} \frac{1}{T} \sum_{t=j+1}^T (y_t - \bar{y}) (y_{t-j} - \bar{y}) \quad \text{para } j = 0, 1, 2, \dots, T-1 \\ \hat{\gamma}_{-j} \quad \text{para } j = 0, 1, 2, \dots, T-1 \end{array} \right\}$$

- donde \bar{y} es la media muestral

Periodograma muestral

- Análogamente al espectro poblacional se define el periodograma muestral como

- $$\hat{S}_y(\omega) = \frac{1}{2\pi} \sum_{j=-T+1}^{T-1} \hat{\gamma}_j e^{-i\omega j} \quad \text{ó}$$

$$\hat{S}_y(\omega) = \frac{1}{2\pi} \left[\hat{\gamma}_0 + 2 \sum_{j=1}^{T-1} \hat{\gamma}_j \cos(\omega j) \right]$$

- De la misma forma que para el espectro poblacional:

$$\int_{-\pi}^{\pi} \hat{S}_y(\omega) d\omega = 2 \int_0^{\pi} \hat{S}_y(\omega) d\omega = \hat{\gamma}_0$$

Periodograma muestral

- El periodograma muestral constituye una forma obvia de estimar el espectro poblacional. Sin embargo, esta aproximación presenta serias limitaciones. Sea el proceso

- $$y_t = \sum_{j=0}^{\infty} \Psi_j \varepsilon_{t-j}$$

- donde la secuencia $\{\Psi_j\}_{j=0}^{\infty}$ es absolutamente sumable y $\{\varepsilon_t\}_{t=-\infty}^{\infty}$ es una secuencia de ruidos blancos independientes

Periodograma muestral

- Sea $S_y(\omega)$ el espectro poblacional y supóngase $S_y(\omega) > 0$ para todo ω
- Fuller demuestra que para todo $\omega \neq 0$ y un tamaño de muestra lo suficientemente grande $\frac{\hat{2S}_y(\omega)}{S_y(\omega)} \sim \chi_{(2)}^2$
- Además para $\lambda \neq \omega$, $\frac{\hat{2S}_y(\lambda)}{S_y(\lambda)} \sim \chi_{(2)}^2$ y es aproximadamente independiente de la primera
- Dado que una variable que se distribuye $\chi_{(2)}^2$ tiene esperanza igual a 2, $E \left[\frac{\hat{2S}_y(\omega)}{S_y(\omega)} \right] \cong 2$
- Y como $S_y(\omega)$ es una magnitud poblacional: $E \left[\hat{S}_y(\omega) \right] \cong S_y(\omega)$

Periodograma muestral

- De esta forma, si el tamaño de muestra es lo suficientemente grande, el periodograma muestral proporciona un estimador aproximadamente insesgado del espectro poblacional
- Sin embargo, el intervalo de confianza a un 95% para una variable $\chi^2_{(2)}$, es (0,05; 7,4), de forma que resulta poco probable que $\hat{S}_y(\omega)$ sea menor que 0,025 y/o mayor que 3,7 veces el verdadero valor de $S_y(\omega)$
- La gran amplitud de dicho intervalo permite establecer que el periodograma muestral como estimador del espectro poblacional es insatisfactorio

Periodograma muestral

- Por otro lado, se espera que un estimador mejore a medida que se utiliza un tamaño de muestra mayor
- Sin embargo, $\hat{S}_y(\omega)$ no es consistente, es decir que no se vuelve “más preciso” a medida que aumenta el tamaño de la muestra, porque con éste también se incrementa el número de autocovarianzas a estimar
- Otra desventaja es que tiene una apariencia “irregular”, aún cuando se refiere a un proceso ruido blanco, pues el número de frecuencias de Fourier en un intervalo dado se incrementa de forma aproximadamente lineal con el tamaño de la muestra

Estimación espectro poblacional: 2) Estimación paramétrica

- Ya sea en el caso de que el verdadero proceso generador de los datos sea un proceso ARMA(p,q) o que pueda ser aproximado por el mismo, entonces los parámetros del proceso pueden estimarse por máxima verosimilitud y el espectro teórico puede calcularse a partir de ellos
- Este procedimiento consta de dos etapas básicas. i) se estiman los coeficientes de un proceso ARMA(p,q)
- $$y_t = \alpha + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$
- por máxima verosimilitud, donde los valores de p y q se establecen de forma que el modelo se ajuste lo más adecuadamente posible al verdadero proceso generador de los datos

Estimación espectro poblacional: 2) Estimación paramétrica

- ii) Se sustituyen los parámetros estimados $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \sigma^2$ en la expresión del espectro

- $$S_y(\omega) = \frac{1}{2\pi} \left[\frac{1 + \theta_1 e^{-i\omega} + \theta_2 e^{-2i\omega} + \dots + \theta_q e^{-iq\omega}}{1 - \phi_1 e^{-i\omega} - \phi_2 e^{-2i\omega} - \dots - \phi_p e^{-ip\omega}} \right] \left[\frac{1 + \theta_1 e^{i\omega} + \theta_2 e^{2i\omega} + \dots + \theta_q e^{iq\omega}}{1 - \phi_1 e^{i\omega} - \phi_2 e^{2i\omega} - \dots - \phi_p e^{ip\omega}} \right] \sigma^2 \quad \omega \in [0, \pi]$$

- Parzen (1969) demuestra que si el modelo ajustado es autorregresivo de orden p

- $$\text{Var} \left[\hat{S}_y(\omega) \right] \approx \frac{2p S_y^2(\omega)}{T}$$

- Es decir que la varianza del espectro se incrementa con p . Esto plantea un trade-off entre un mejor ajuste del modelo y la minimización de la varianza del espectro

Estimación espectro poblacional: 2) Estimación paramétrica

- Berk (1974) muestra que el estimador espectral basado en la modelización ARMA(p,q) es insesgado, consistente y asintóticamente normal
- Al basarse en un modelo teórico tiende a ser más suave que el estimado mediante las técnicas no paramétricas, que se verá a continuación, y parece mostrar una alta resolución siendo capaz de seleccionar picos angostos (Priestley,1981)
- Aún si el modelo está incorrectamente especificado, si las autocovarianzas del verdadero proceso se encuentran cercanas a las de una especificación ARMA(p,q) este procedimiento debería proporcionar una útil estimación del espectro poblacional (Hamilton, 1994)

Aplicación: producción industrial USA, Hamilton (1994)

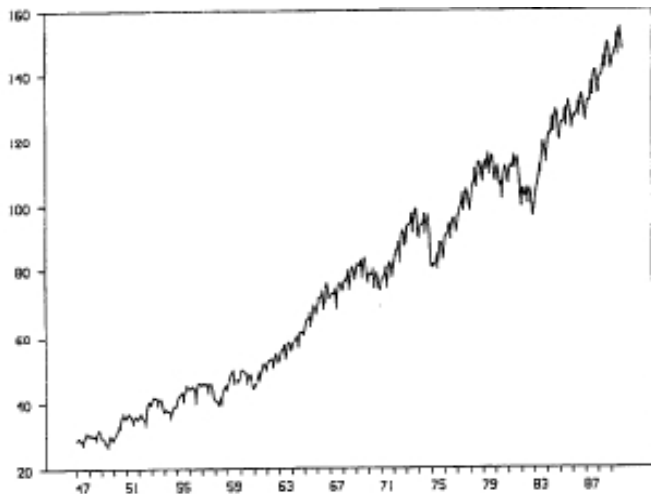


FIGURE 6.3 Federal Reserve Board's seasonally unadjusted index of industrial production for U.S. manufacturing, monthly 1947:1 to 1989:11.

Aplicación: producción industrial USA - periodograma

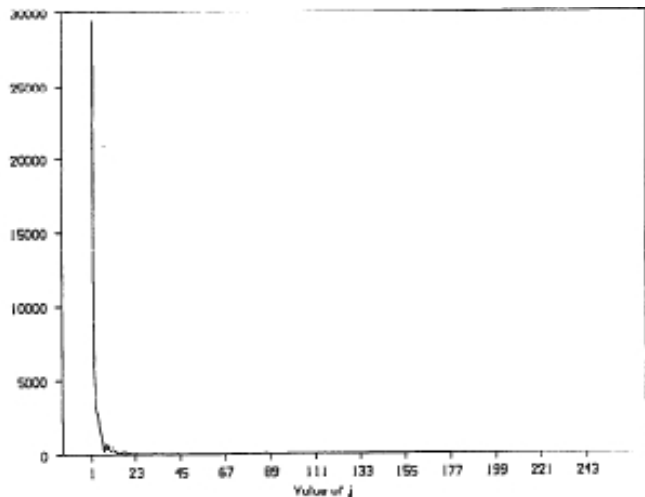


FIGURE 6.4 Sample periodogram for the data plotted in Figure 6.3. The figure plots $f.(\omega_j)$ as a function of j , where $\omega_j = 2\pi j/T$.

Aplicación: espectro primera diferencia

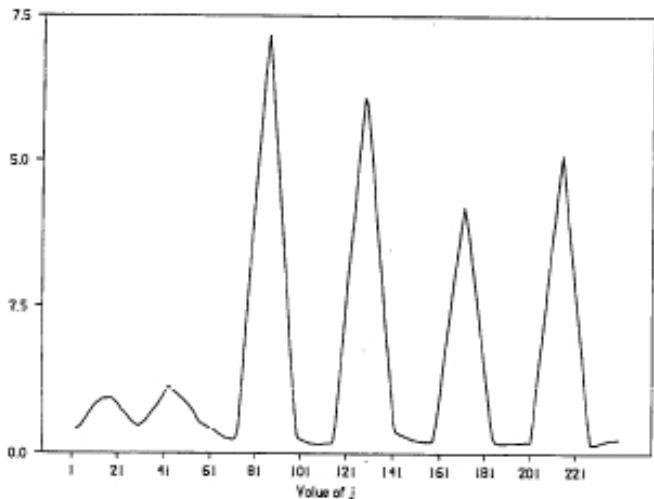


FIGURE 6.5 Estimate of the spectrum for monthly growth rate of industrial production, or spectrum of 100 times the first difference of the log of the series in Figure 6.3.

Aplicación: espectro tasa de crecimiento interanual

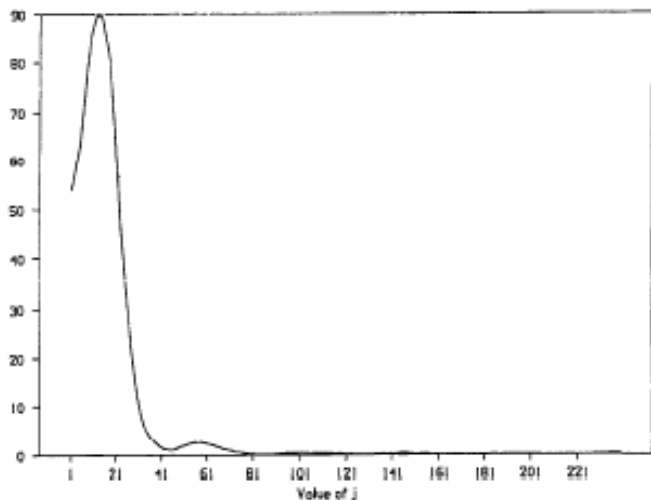


FIGURE 6.6 Estimate of the spectrum for year-to-year growth rate of monthly industrial production, or spectrum of 100 times the seasonal difference of the log of the series in Figure 6.3.