

# Dominio de frecuencias

Series Cronológicas

Silvia Rodríguez Collazo

Facultad de Ciencias Económicas y de Administración

Hasta ahora hemos trabajado con la idea de que una serie de tiempo presenta ciertas regularidades en su comportamiento a través del tiempo, las que permiten caracterizar su trayectoria y las que pueden detectarse mediante herramientas adecuadas, en el dominio del tiempo, como son la función de autocovarianzas, autocorrelación y la función de autocorrelación parcial.

A partir de la forma de estas funciones se intenta representar mediante modelos lineales dichas regularidades.

Vamos a retomar el concepto de regularidad en la serie pero expresada de otra forma, **vamos a concebir a los procesos estacionarios como una combinación de senos y cosenos de distinta amplitud**. En este contexto, presentaremos nuevas herramientas para caracterizar al proceso.

**La representación espectral de una serie estacionaria, descompone a  $Y_t$  en una suma de componentes sinusoidales con coeficientes aleatorios incorrelacionados.**

Además de esta descomposición hay una correspondiente descomposición en sinusoides de la función de autocovarianzas de  $Y_t$ .

La descomposición espectral es para los procesos estocásticos estacionarios lo que es la representación de Fourier para las funciones determinísticas.

El análisis de los procesos estacionarios mediante su representación espectral se conoce usualmente como *análisis del dominio de las frecuencias o análisis espectral* de la serie de tiempo.

Es equivalente al análisis en el dominio del tiempo basado en la función de autocorrelación y autocorrelación parcial y da una visión alternativa al estudio de las series de tiempo (Brockwell y Davis (1991)).

El análisis en el dominio del tiempo y en el dominio de la frecuencia no son excluyentes, son dos enfoques complementarios que van a permitir caracterizar al proceso. Son distintas formas de mirar un mismo fenómeno.

Hemos visto los procesos ARMA(p,q) que se pueden representar en términos generales como:

$$y_t = \sum_{j=1}^p \phi_j y_{t-j} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t$$

Hemos prestado especial atención a la forma de la función de autocovarianzas  $\gamma_k$  para diferentes rezagos, como la función que contiene la información sobre el tipo de covariaciones, de regularidades que caracterizan al proceso.

Si la secuencia de autocovarianzas es absolutamente sumable, una forma de resumir las autocovarianzas es a través de una función escalar, llamada **función generatriz de autocovarianzas**, que se define como:

$$g_y(z) = \sum_{j=-\infty}^{\infty} \gamma_j z^j = \sum_{j=-\infty}^{\infty} \gamma_j (\cos(wj) - i \sin(wj)) = \sum_{j=-\infty}^{\infty} \gamma_j e^{-i w j} \quad (1)$$

donde  $z$  representa un escalar complejo. A partir del teorema de De Moivre podemos escribir:

$$z = \cos(w) - i \sin(w) = e^{-i w} \text{ con } w \in [0, 2\pi]$$

donde  $i = \sqrt{-1}$  y

$w$  es el ángulo (medido en radianes) que  $z$  forma con el eje de los reales.

## MA(1)

Sea el proceso  $y_t = \theta \varepsilon_{t-1} + \varepsilon_t = (1 + \theta L)\varepsilon_t = \theta(L)\varepsilon_t$

Donde las autocovarianzas relevantes para el MA(1) son:

$$\begin{aligned}\gamma_0 &= (1 + \theta^2)\sigma_\varepsilon^2 \\ \gamma_1 &= \gamma_{-1} = \theta\sigma^2 \\ \gamma_k &= 0 \quad \forall k \geq 2\end{aligned}$$

Dadas las autocovarianzas de este proceso, la función generatriz de autocovarianzas es:

$$\begin{aligned}g_y(z) &= \theta\sigma_\varepsilon^2 z^{-1} + (1 + \theta^2)\sigma_\varepsilon^2 z^0 + \theta\sigma_\varepsilon^2 z^1 \\ g_y(z) &= \sigma_\varepsilon^2 \left[ \theta z^{-1} + (1 + \theta^2)z^0 + \theta z^1 \right] \\ g_y(z) &= \sigma_\varepsilon^2 \left[ \theta z^{-1} + (1 + \theta^2) + \theta z^1 \right] \\ g_y(z) &= \sigma_\varepsilon^2 (1 + \theta z^{-1})(1 + \theta z) = \boxed{\sigma_\varepsilon^2 \theta(z)\theta(z^{-1})}\end{aligned}\tag{2}$$

Donde  $\theta(L) = 1 + \theta L$  (el polinomio introduce correlación a partir de un proceso ruido blanco para obtener  $Y_t$ ).

## MA(q)

Sea el proceso

$$Y_t - \mu = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} \dots + \theta_q \varepsilon_{t-q} = (1 + \theta_1 L + \theta_2 L^2 \dots + \theta_q L^q) \varepsilon_t$$

El proceso tiene una función de autocovarianzas:

$$\begin{aligned}\gamma_0 &= (1 + \theta_1^2 + \theta_2^2 \dots + \theta_q^2) \sigma^2 \\ \gamma_k &= (\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q) \sigma^2 & \forall k \leq q \\ \gamma_k &= 0 & \forall k > q\end{aligned}$$

Operando resulta que la función generatriz de autocovarianzas es:

$$g_Y(z) = \sigma_\varepsilon^2 (1 + \theta_1 z + \theta_2 z^2 \dots + \theta_q z^q) (1 + \theta_1 z^{-1} + \theta_2 z^{-2} \dots + \theta_q z^{-q})$$

## MA ( $\infty$ )

Sea  $Y_t = \mu + \psi(L)\varepsilon_t$

Donde  $\psi(L) = \psi_0 + \psi_1 L + \psi_2 L^2 + \dots$

con  $\sum_{k=0}^{\infty} |\psi_k| < \infty$

$Y_t$ , es un proceso estacionario cuya Función generatriz de autocovarianzas es

$$g_y(z) = \sigma_\varepsilon^2 \psi(z) \psi(z^{-1})$$

## AR(1)

$$Y_t = \phi Y_{t-1} + \varepsilon_t \implies Y_t - \phi Y_{t-1} = \varepsilon_t \implies (1 - \phi L) Y_t = \varepsilon_t \implies Y_t = \frac{\varepsilon_t}{(1 - \phi L)}$$

$\implies Y_t = \psi(L)\varepsilon_t$ , el proceso queda expresado como un MA( $\infty$ ).

donde  $\psi(L) = \frac{1}{(1 - \phi L)}$

y la función generatriz de autocovarianzas es:

$$g_y(z) = \sigma_\varepsilon^2 \frac{1}{(1 - \phi z)(1 - \phi z^{-1})}$$

## AR(p)

$$\begin{aligned}Y_t &= \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t \\ \varepsilon_t &= (1 - \phi_1 L + \phi_2 L^2 + \dots + \phi_p L^p) Y_t\end{aligned}\tag{3}$$

$$Y_t = \frac{\varepsilon_t}{(1 - \phi_1 L + \phi_2 L^2 + \dots + \phi_p L^p)} = \psi(L) \varepsilon_t$$

La Función generatriz de autocovarianzas tiene la siguiente expresión:

$$g_Y(z) = \frac{\sigma_\varepsilon^2}{(1 - \phi_1 z - \phi_2 z^2 - \dots + \phi_p z^p)(1 - \phi_1 z^{-1} - \phi_2 z^{-2} - \dots + \phi_p z^{-p})}$$

## ARMA(p,q)

Un proceso ARMA(p,q) puede escribirse como:

$$\Phi(L)Y_t = \theta(L)\varepsilon_t \implies Y_t = \frac{\theta(L)}{\Phi(L)}\varepsilon_t = \Psi(L)\varepsilon_t$$

La función generatriz de un proceso ARMA(p,q) se escribe de la siguiente forma:

$$g_Y(z) = \sigma^2 \left( \frac{\theta(z)\theta(z^{-1})}{\Phi(z)\Phi(z^{-1})} \right)$$



Es posible aproximar una función definida en un intervalo finito en los reales, con el nivel de precisión deseado, mediante una suma ponderada de funciones seno y coseno en frecuencias armónicas crecientes.

Resultados similares se obtienen en el caso de las secuencias.

Se puede entonces representar la trayectoria del proceso  $Y_t$  ( mediante el conjunto de observaciones o realizaciones del proceso,  $y_t$  ) con el grado de precisión deseado, mediante una suma ponderada de funciones periódicas de la forma  $\cos(wt)$  y  $\sin(wt)$  , donde  $w$  denota una frecuencia angular particular.

La esencia del análisis de Fourier es la representación de una secuencia de datos en términos de funciones sinusoidales.

En este marco, la herramienta clave para captar los ciclos relevantes, las periodicidades relevantes, es el **espectro poblacional** de la serie, el que se obtiene mediante una transformación de la función generatriz de autocovarianzas.

Si la función generatriz de autocovarianzas es evaluada en un escalar complejo  $z = \cos(w) - i \sin(w) = e^{-iw}$  y se divide por el factor  $2\pi$ , el resultado es el **espectro poblacional** del proceso.

El espectro es una función de  $w$ . Dado cualquier valor particular de  $w$  y una secuencia de autocovarianzas  $\{\gamma_j\}$ , se puede calcular el espectro  $S_y(w)$ .

$$S_y(w) = \frac{1}{2\pi} g_y(e^{-iwj}) = \frac{1}{2\pi} \left\{ \gamma_0 + 2 \sum_{j=1}^{\infty} \gamma_j \cos(wj) \right\} \quad (4)$$

**El espectro es la transformada de Fourier de la Función generatriz de autocovarianzas.**

En ambos enfoques es fundamental el **supuesto sobre la estacionariedad** de la serie, y la sumabilidad absoluta de las autocovarianzas, ya que esta condición permite que el espectro poblacional converja.

La función de autocovarianzas expresa la información en términos de rezagos, mientras que la densidad espectral expresa la misma información en términos de ciclos. Para describir este enfoque van a ser útil explicitar algunos puntos previos referidos a :

- Periodicidades en una serie de tiempo
- Series de Fourier

Nos interesa en este momento estudiar de qué manera se pueden representar las distintas periodicidades que intervienen en este análisis.

Para ello vamos a analizar la relación entre período y frecuencias, dicha relación será de utilidad para la interpretación de los resultados en el dominio de la frecuencia.

Mediante el análisis de Fourier es posible representar una secuencia de observaciones en términos de funciones sinusoidales.

¿Porqué este tipo de funciones?. Las funciones sinusoidales presentan ciertas propiedades que las convierte en herramientas útiles:

- El comportamiento de estas funciones ante cambios de escala. La amplitud de la senoide no depende ni del origen ni de la escala de la variable tiempo.
- La suma de sinusoides con frecuencia común es otra senoide con la misma frecuencia.

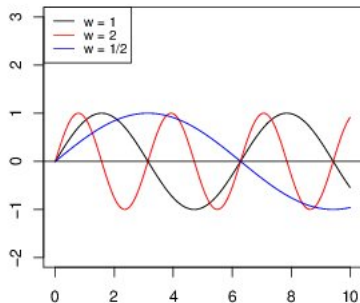
## Periodicidades en una serie de tiempo

En la figura se pueden ver los gráficos de la función  $\text{sen}(wt)$  para tres valores distintos de  $w$ , y donde  $t$  representa el tiempo.

Esta es una función "periódica", es decir que se repite cada cierto período, por esta razón es útil como ejemplo para ayudar a definir los conceptos de período, frecuencia y la relación entre ambos.

Las funciones graficadas pueden asociarse con periodicidades distintas, cuando  $w = 1/2$  el gráfico muestra movimientos más "suaves" y corta una sola vez el eje horizontal. En el otro extremo si  $w = 2$  el  $\text{sen}(2t)$  los movimientos son "menos suaves" y corta más veces el eje horizontal.

**Sen( $wt$ ) en distintas frecuencias**



**El período de la función se define como la distancia entre dos máximos relativos consecutivos.**

Por ejemplo  $\sin(w)$  tiene un período  $p = 2\pi$ , lo que implica que  $\sin(t) = \sin(2\pi t)$  para todo  $t$ . El período está vinculado al valor de  $w$ .

**El término  $w$  denota a la frecuencia angular (ciclos por unidad de tiempo); su relación con el período es  $p = \frac{2\pi}{w}$ .**

En la figura podía verse la relación inversa entre  $p$  y  $w$ , por ejemplo cuando  $w = 2$  el período es menor o de otro modo los máximos relativos de  $\sin(2t)$  son más próximos entre sí. El período de  $\sin(2t)$  es  $p = \frac{2\pi}{2} = \pi$  y por eso hay tres máximos relativos en el intervalo.

Uno de los principales resultados del análisis de Fourier es que cualquier función definida en un intervalo finito de reales, puede ser aproximada mediante una conveniente combinación ponderada, de funciones seno y coseno en las frecuencias armónicas crecientes.

Sea la series de tiempo  $Y_t$  para una muestra de  $T$  observaciones, es posible derivar la expresión:

$$Y_t = \sum_{j=1}^M \alpha_j \cos(w_j t) + \beta_j \sin(w_j t)$$

con  $0 < w_1 < w_2 \dots < w_k < \pi$  y

$w_j = 2\pi j / T$  múltiplo de la **frecuencia fundamental**  $w_1 = 2\pi / T$

Donde los elementos de la secuencia finita pueden ser expresados en términos de senos y cosenos.

Esta expresión se conoce como **descomposición de Fourier de  $Y_t$** .

Donde los coeficientes  $\alpha_1, \beta_1 \dots \alpha_k, \beta_k$  son **variables aleatorias** y se les conoce como coeficientes de Fourier

cuya  $E(\alpha_j) = E(\beta_j) = 0$  y  $Var(\alpha_j) = Var(\beta_j) = \sigma_j^2$  para  $j = 1, \dots, k$ .

**Las secuencias  $\alpha_j$  y  $\beta_j$  son serialmente incorrelacionadas y mutuamente incorrelacionados.**

$Y_t$  como todo proceso estacionario con media cero puede ser expresado como una superposición de sinusoides incorrelacionadas con frecuencias  $w \in [0, \pi]$ .

En general, un proceso estacionario puede escribirse como una superposición de infinitas sinusoides, más que un número finito de ellas, como en el ejemplo.

- $w_j = w_1 j$  con  $w_1 = \frac{2\pi}{T}$  es la frecuencia fundamental.
- $n = \frac{T}{2}$  si  $T$  es par, y  $n = \frac{T-1}{2}$  cuando es impar.
- Los coeficientes  $\alpha_j$  y  $\beta_j$  son los llamados coeficientes de Fourier.

La representación de Fourier es el fundamento del **teorema de representación espectral**, que es el análogo del teorema de Wold pero en el dominio de las frecuencias y constituye la pieza teórica fundamental del análisis.

La analogía con el teorema de Wold es que ambos brindan una manera de representar cualquier proceso que sea estacionario.

El **teorema de representación espectral** establece que si  $Y_t$  es un proceso estocástico estacionario en covarianza , con autocovarianzas absolutamente sumables.

Para cualquier frecuencia fija  $w \in [0, \pi]$  , se definen las variables aleatorias  $\alpha(w)$  y  $\delta(w)$  . El proceso  $Y_t$  se puede representar como :

$$Y_t = \mu + \int_0^{\pi} [\alpha(w)\cos(wt) + \delta(w)\sin(wt)]dw \quad (5)$$

Donde  $\alpha(\cdot)$  y  $\delta(\cdot)$  tienen media cero y para  $0 < w_1 < w_2 < w_3 < w_4 < \pi$ ,

$\int_{w_1}^{w_2} \alpha(w)dw$  está incorrelacionada con  $\int_{w_3}^{w_4} \alpha(w)dw$

y  $\int_{w_1}^{w_2} \delta(w)dw$  está incorrelacionada con  $\int_{w_3}^{w_4} \delta(w)dw$

y para cualquier  $0 < w_1 < w_2 < \pi$  y  $0 < w_3 < w_4 < \pi$ ,  $\int_{w_1}^{w_2} \alpha(w)dw$  está incorrelacionada con  $\int_{w_3}^{w_4} \delta(w)dw$ .

**Este teorema de representación espectral, muestra que un proceso estacionario, con función de autocovarianzas absolutamente sumable, puede ser escrito como la yuxtaposición de procesos elementales de naturaleza periódica (en el tiempo continuo mediante integrales y en tiempo finito, mediante sumatorias).**



De acuerdo a Piccolo (1990) desde un **punto de vista intuitivo**, el espectro es una función de un proceso estacionario que permite ver de modo claro eventuales componentes cíclicos.

Desde un **punto de vista matemático**, el espectro es la transformada de Fourier de una función o sucesión numérica.

Desde un **punto de vista estadístico**, el espectro es una descomposición de la varianza mediante la contribución de las distintas frecuencias angulares a la variabilidad de la serie.

Sea  $Y_t$  un proceso estacionario con secuencia de autocovarianzas  $\gamma_j$  absolutamente sumable.

**El espectro poblacional es también una transformación de la función generatriz de autocovarianzas.**

Si la función generatriz de autocovarianzas es evaluada en un escalar complejo  $z = \cos(w) - i \sin(w) = e^{-iw}$  y se divide por el factor  $2\pi$ , el resultado es el espectro poblacional del proceso.

El espectro es una función continua valuada en los reales, función de  $w$ .

Dado cualquier valor particular de  $w$  y una secuencia de autocovarianzas  $\{\gamma_j\}$ , se puede calcular el espectro  $S_y(w)$ .

$$S_y(w) = \frac{1}{2\pi} g_y(e^{-iw}) = \frac{1}{2\pi} \left\{ \gamma_0 + 2 \sum_{j=1}^{\infty} \gamma_j \cos(wj) \right\} \quad (6)$$

**En  $S_y(w)$  se resume la variabilidad del proceso que generó la secuencia  $Y_t$ , es por tanto, una herramienta fundamental para el estudio de series temporales en el dominio de la frecuencia.**

Por su forma de construcción, el espectro verifica las siguientes propiedades:

- Si  $\gamma_j$  constituyen las autocovarianzas de un proceso débilmente estacionario con la secuencia  $\{\gamma_j\}_{-\infty}^{+\infty}$  absolutamente sumable, entonces  $S_y(w)$  **es una función de valores reales, continua y no negativa**.
- El espectro es **simétrico** respecto de  $w = 0$ , debido a que  $\cos(wt) = \cos(-wt) \forall w$ . Esto implica que  $S_y(w) = S_y(-w)$ .
- Como  $\cos(wj) = \cos[(w + 2\pi\tau)j]$ , para cualquier  $j$  y  $\tau$  enteros, el espectro es una función periódica de  $w$ :  $S_y(w) = S_y(w + 2\pi\tau)$ .  
Así, el conocimiento del valor de  $S_y(w)$  para cualquier valor de  $w$  en el intervalo  $[0, \pi]$ , implica el conocimiento de  $S_y(w)$  para cualquier valor de  $w$ .

**El espectro puede verse como una descomposición de la varianza de la serie por frecuencias.**

El área bajo el espectro en el intervalo  $[-\pi, \pi]$ , es la varianza de  $Y_t$ :

$$\int_{-\pi}^{\pi} S_y(w)dw = \gamma_0 \quad (7)$$

De forma más general, dado que  $S_y(w)$  es no negativo, se puede calcular:

$\int_{-w_1}^{w_1} S_y(w)dw$  para cualquier  $w_1$  entre 0 y  $\pi$ .

Este resultado, puede interpretarse como la porción de la varianza de  $Y_t$ , asociada con las frecuencias  $w$  que son  $\leq |w_1|$ .

Para interpretar la información que brinda el espectro consideremos el vínculo entre el dominio de la frecuencia y el dominio del tiempo, el vínculo entre el espectro poblacional  $S_y(w)$  y la función de autocovarianzas  $\gamma_j$ . Dicha relación, se conoce como **teorema Wiener-Kinchine** para procesos estacionarios.

$$\int_{-\pi}^{\pi} S_y(w) e^{-i w j} dw = \gamma_j \quad (8)$$

A partir de la ecuación se puede ver que dado  $S_y(w)$  es posible obtener las autocovarianzas de la serie y si,  $j=0$ , a partir del área bajo el espectro se obtiene el valor de la varianza de la serie.

Por otro lado ecuación define el espectro poblacional de la serie como una función de  $\gamma_j$ .

En conjunto, ambas ecuaciones determinan la existencia de una relación entre el espectro poblacional y la estructura de autocovarianzas, y por lo tanto determinan la equivalencia entre el dominio del tiempo y el dominio de la frecuencia para el estudio de series de tiempo.

- ❶ Si se considera el caso particular cuando  $j = 0$ , se puede ver que:

$$\int_{-\pi}^{\pi} S_Y(w) = \gamma_0$$

Es decir que la variabilidad total de la serie es igual al área bajo el espectro poblacional. Además, como el espectro es siempre no negativo, la integral de  $S_Y(w)$  entre  $w_a$  y  $w_b$  puede ser interpretada como la porción de varianza explicada por el conjunto de periodicidades que se corresponden con las frecuencias entre  $w_a$  y  $w_b$ , las que pertenecen al intervalo  $[0, \pi]$ . **Utilizando el espectro poblacional es posible descomponer la serie en componentes vinculados a periodicidades con distinto rango de frecuencias.**

- ❷ Por otro lado, como  $S_Y(w)$  es simétrico en torno a cero, se cumple que:

$$\int_{-w_b}^{w_b} S_Y(w) dw = 2 \int_0^{w_b} S_Y(w) dw$$

es la porción de la varianza de  $Y_t$  que puede ser atribuida a componentes aleatorios periódicos con frecuencias menores o iguales a  $w_b$  en valor absoluto.

Si se considera el caso de un proceso ruido blanco, es decir una secuencia  $\varepsilon_t$  de variables aleatorias iid, con media cero y varianza constante  $V(\varepsilon_t) = \sigma^2$ , la secuencia de autocovarianzas  $\gamma_j = 0$  para todo  $j > 0$  y  $\gamma_0 = \sigma_\varepsilon^2$ .

Utilizando la ecuación 6 el resultado es que

$$S_\varepsilon(w) = \frac{\sigma_\varepsilon^2}{2\pi}$$

El espectro poblacional es constante a lo largo de todas las frecuencias. **Todas las frecuencias tienen el mismo peso** para explicar la variabilidad de un proceso ruido blanco. De aquí viene el término ruido blanco: así como una luz blanca, consiste en un número infinito de frecuencias cuyo peso es el mismo.

## MA(1)

$$\begin{aligned}Y_t &= \varepsilon_t + \theta \varepsilon_{t-1} \\ \gamma_0 &= (\theta^2 + 1)\sigma_\varepsilon^2 \\ \gamma_1 &= \theta \sigma_\varepsilon^2 \\ \gamma_i &= 0 \quad \forall i > 1\end{aligned}$$

Con la anterior estructura de autocovarianzas el espectro poblacional queda definido de la siguiente manera:

$$S_y(w) = \frac{1}{2\pi} \left\{ \gamma_0 + 2 \sum_{j=1}^{\infty} \gamma_j \cos(wj) \right\} \quad (9)$$

$$S_y(w) = \frac{1}{2\pi} \left[ (1 + \theta^2)\sigma_\varepsilon^2 + 2 \sum_{j=1}^1 \theta \sigma_\varepsilon^2 \cos(wj) \right]$$

$$S_y(w) = \frac{\sigma_\varepsilon^2}{2\pi} (\theta^2 + 2\cos(w)\theta + 1)$$

La pendiente del espectro poblacional depende del signo del parámetro (si se supone que  $\sigma_\varepsilon^2 = 1$ ), y en el caso que  $\theta$  es positivo las frecuencias bajas son las que explican en mayor medida la variabilidad del proceso.

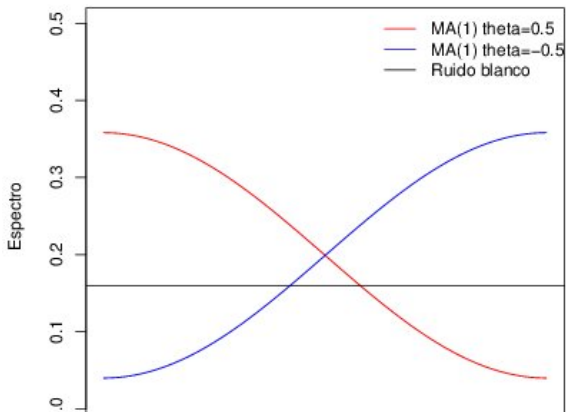
En cambio cuando el valor de  $\theta$  es negativo los resultados se invierten.



## El espectro de un MA(1)

Para ver el sentido que tiene la forma del espectro se debe recordar que las frecuencias bajas están asociadas a períodos más largos, en el primer caso los períodos largos explican la mayor parte de la variabilidad de la serie, mientras que en el segundo caso la mayor variabilidad está explicada con los períodos de más corto plazo, por lo que el proceso va a presentar un comportamiento más "errático".

**Algunos Espectros**



## AR(1)

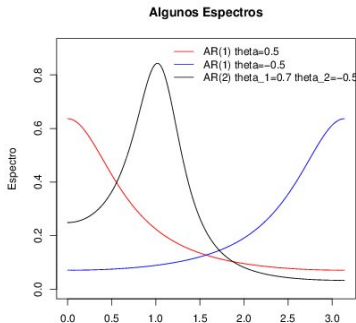
$$Y_t = \phi Y_{t-1} + \varepsilon_t \implies Y_t = \frac{1}{1-\phi L} \varepsilon_t$$

El espectro poblacional se expresa como sigue:

$$S_y(w) = \frac{1}{2\pi} g_y(e^{-iw}) \implies S_y(w) = \frac{1}{2\pi} \left[ \sigma^2 \psi(e^{iw}) \psi(e^{-iw}) \right] \quad (10)$$

$$S_y(w) = \frac{1}{2\pi} \frac{\sigma^2}{(1 - \phi e^{iw})(1 - \phi e^{-iw})}$$

$$S_y(w) = \frac{1}{2\pi} \frac{\sigma^2}{(1 + \phi^2 - 2\phi \cos(w))}$$



Vamos a tratar otro aspecto del análisis desde el dominio de las frecuencias, vamos a vernos el análisis espectral para observar el efecto de la aplicación de un determinado **filtro lineal** a un proceso estocástico.

En términos generales un filtro **lineal** tiene la siguiente estructura

$$h(L) = \sum_j h_j L^j$$

donde  $L$  es el operador de retardos.

Recordemos la representación de Wold de un proceso ARMA(p,q) estacionario,

$$Y_t = \Psi(L)\varepsilon_t$$

por lo que se puede observar cómo todos los ARMA(p,q) son el resultado de aplicar un filtro  $\Psi(L)$  a un proceso ruido blanco,  $\varepsilon_t$ .

Se parte de un proceso estocástico sin correlaciones, sin regularidades,  $\varepsilon_t$  y mediante la aplicación del filtro,  $\Psi(L)$ , se introducen las correlaciones.

En el marco del análisis espectral vamos a observar la relación del espectro de un ruido blanco con el de un ARMA(p,q) y observaremos las periodicidades que se transfieren, cuales se anulan y cuáles se incrementan mediante la utilización de una función llamada **función de transferencia**.

Sea  $X_t$  el resultado de aplicar el filtro lineal  $h(L)$  a la proceso  $Y_t$ :

$$X_t = h(L)Y_t$$

Podemos escribir la relación entre los espectros poblacionales de  $X_t$  y  $Y_t$ :

$$S_x(w) = h(e^{-iw})h(e^{iw})S_y(w) = f(w)S_y(w) \quad (11)$$

La función  $f(w)$  es la llamada **función de transferencia**.

Esta función permite analizar el efecto del filtro  $h(L)$  sobre la serie  $Y_t$ .

Podemos escribir entonces a  $f(w)$  como

$$f(w) = \frac{S_x}{S_y}$$

Se puede decir que cuando  $f(w_a) > 1$  el filtro aumenta el peso de la frecuencia  $w_a$  en la varianza de la serie, y reducirá el peso en aquellas frecuencias donde la función de transferencia sea menor que la unidad. En el caso que  $f(w_a) = 0$  el filtro anula el peso de esta frecuencia, ( $w_a$ ).

Un filtro ampliamente conocido y utilizado es el **operador de primera diferencia regular**

$$\Delta Y_t = Y_t - Y_{t-1}$$

. Es el resultado de aplicar a  $Y_t$  el filtro lineal  $(1 - L)$ . Para analizar los efectos de la aplicación de este filtro, escribimos a  $X_t$  como

$$X_t = h(L)Y_t = (1 - L)Y_t$$

luego calculamos la función de transferencia y visualizamos el resultado gráficamente.

Como  $Y_t$  es un proceso estacionario, podemos expresarlo como:  $Y_t = \Psi(L)\varepsilon_t$ , por tanto a  $X_t$  lo escribimos como:

$$X_t = h(L)\Psi(L)\varepsilon_t \equiv \Psi^*(L)\varepsilon_t,$$

con

$$\Psi^*(L) = h(L)\Psi(L)$$

. La secuencia de coeficientes asociados con el operador compuesto  $\{\Psi_j^*\}_{j=-\infty}^{+\infty}$  resulta ser absolutamente sumable por lo que la función generatriz de autocovarianzas de  $X_t$  en consecuencia puede ser calculada como:

$$g_X(z) = \sigma^2 \Psi^*(z) \Psi^*(z^{-1}) = \sigma^2 h(z) \Psi(z) \Psi(z^{-1}) h(z^{-1}) = h(z) h(z^{-1}) g_Y(z)$$

Reescribimos la función generatriz de autocovarianzas de  $X_t$ :

$$g_x(z) = h(z)h(z^{-1})g_y(z)$$

O sea que aplicar el filtro  $h(L)$  a una serie implica multiplicar a la función generatriz de autocovarianzas por  $h(z)h(z^{-1})$ .

Escribamos el espectro poblacional de  $X_t$  a partir de la FGAutocov.

$$\begin{aligned}g_x(z) &= h(z)h(z^{-1})g_y(z) \\S_x(w) &= \frac{1}{2\pi}g_x(e^{-iw}) = \frac{1}{2\pi}h(e^{-iw})h(e^{iw})g_y(e^{iw}) \\S_x(w) &= h(e^{-iw})h(e^{iw})S_y(w)\end{aligned}\tag{12}$$

Recordemos que  $e^{-iw} = \cos(w) - i\sin(w)$  por tanto la función de transferencia del filtro primera diferencia  $\Delta$  tiene esta expresión:

$$\begin{aligned}f_{\Delta}(w) &= (1 - e^{-iw})(1 - e^{iw}) = 2 - e^{iw} - e^{-iw} \\f_{\Delta}(w) &= 2 - (\cos(w) - i\sin(w)) - (\cos(w) + i\sin(w))\end{aligned}$$

y operando se llega a la expresión  $f_{\Delta}(w) = 2(1 - \cos(w))$ .



Si se considera el **operador de primera diferencia estacional** para series trimestrales,

$$\Delta^4(Y_t) = Y_t - Y_{t-4}$$

. Y escribimos a  $X_t$  como el resultado de aplicar el filtro a  $Y_t$

$$X_t = (1 - L^4)Y_t$$

La función de transferencia es:

$$f_{\Delta^4}(w) = h(e^{4iw})h(e^{-4iw}) = (1 - e^{4iw})(1 - e^{-4iw}) = 2 - e^{4iw} - e^{-4iw}$$

$$f_{\Delta^4}(w) = 2 - e^{4iw} - e^{-4iw}$$

Recordando que  $e^{4iw} = \cos(4w) - i\sin(4w)$   
por tanto la función de transferencia es:

$$f_{\Delta^4}(w) = 2(1 - \cos(4w))$$







A pesar que estos filtros son diseñados para eliminar el peso de las frecuencias mencionadas, **incorporan distorsiones** adicionales a la serie original.

El análisis espectral permite ver que otros efectos (aparte de los buscados) son incorporados a la serie al aplicar estos filtros.

En particular cuando se utiliza *la primera diferencia regular*, se aumenta el peso de las frecuencias más altas y por lo tanto provoca que el proceso resultante del filtrado sea más ruidoso" que el original.

Mientras que cuando se aplica *la primer diferencia estacional*, se realiza la frecuencia  $w = \frac{\pi}{4}$  que se asocia a los períodos de 8 trimestres y de 2,66 trimestres que corresponde a la frecuencia  $\frac{3\pi}{4}$ .

Es importante considerar estos efectos ya que cuando se analiza la serie filtrada, se pueden confundir las características de los datos con los efectos introducidos por el filtro que no son propios del proceso.

Describiremos las distintas formas de estimar el espectro poblacional dada una muestra  $y_1, y_2, \dots, y_T$ .

La estimación del espectro se puede realizar de distintas formas : el periodograma muestral, estimación de tipo paramétrico del espectro y la estimación no paramétrica.

## El Periodograma Muestral

La forma más directa consiste en estimar el espectro poblacional a partir de las autocovarianzas. Primero se obtienen la autocovarianzas muestrales y luego se incluyen en la ecuación 13; el resultado se conoce como periodograma muestral,  $\hat{S}_y(w)$ .

$$\hat{S}_y(w) = \frac{1}{2\pi} \left\{ \hat{\gamma}_0 + 2 \sum_{j=1}^{T-1} \hat{\gamma}_j \cos(wj) \right\} \quad 0 \leq w \leq \pi \quad (13)$$

El área bajo el periodograma muestral, es la varianza muestral.

$$\int_{-\pi}^{\pi} \hat{S}_y(w) = \hat{\gamma}_0$$

El periodograma muestral es también simétrico entorno a  $w = 0$  y se puede escribir:

$$2 * \int_0^{\pi} \hat{S}_y(w) = \hat{\gamma}_0$$

Consideremos que el número de observaciones es  $T$ , impar.  $Y_t$  puede ser expresado en términos de funciones periódicas con  $M = ((T-1)/2)$  distintas frecuencias tal que:  $w_1 = 2\pi/T, w_2 = 4\pi/T, \dots, w_M = 2M\pi/T$ . Si se considera la regresión por mínimos cuadrados ordinarios de  $y_t$  con una constante y varios términos senos y cosenos:

$$Y_t = \mu + \sum_{j=1}^M \alpha_j \cos[w_j(t-1)] + \delta_j \sin[w_j(t-1)] + \varepsilon_t \quad (14)$$

Los valores ajustados de esa regresión, son de la forma:

$$Y_t = \hat{\mu} + \sum_{j=1}^M \hat{\alpha}_j \cos[w_j(t-1)] + \hat{\delta}_j \sin[w_j(t-1)] \quad (15)$$

- con  $\hat{\mu} = \bar{y}$
- $\hat{\alpha}_j = 2/T \sum_{t=1}^T y_t \cos[w_j(t-1)] ; j=1,2,\dots,M$
- $\hat{\delta}_j = 2/T \sum_{t=1}^T y_t \sin[w_j(t-1)] ; j=1,2,\dots,M$
- $\hat{\gamma}_0 = 1/2 \sum_{j=1}^M \hat{\alpha}_j^2 + \hat{\delta}_j^2$

Fuller (1976) determina la distribución asintótica del cociente entre el periodograma y el espectro poblacional,  $S_y(w)$ , cuando  $w \neq 0$  y un tamaño muestral  $T$  suficientemente grande. Formalmente establece que

$$\frac{2\hat{S}_y(w)}{S_y(w)} \xrightarrow{d} \chi^2(2)$$

Si se considera  $\lambda \neq w$ , la convergencia en distribución anterior se mantiene

$$\frac{2\hat{S}_y(\lambda)}{S_y(\lambda)} \xrightarrow{d} \chi^2(2)$$

Para ver las propiedades del periodograma como estimador del espectro se debe tener en cuenta que toda variable aleatoria que se distribuye  $\chi^2(2)$  tiene esperanza 2 entonces el periodograma es un estimador asintóticamente insesgado del espectro.

Sin embargo la varianza del estimador no depende de  $T$  y por lo tanto no proporciona un estimador consistente del espectro para una frecuencia dada.

El estimador no es más preciso a medida que se incrementa el tamaño de la muestra ya que aumenta el número de autocovarianzas a estimar.

En concreto la varianza de una  $\chi^2(2)$  es 4 y por lo tanto el desvío del periodograma es igual al valor del espectro.

$$E\left(\frac{2\hat{S}_y(w)}{S_y(w)}\right) \cong 2 \Rightarrow E(\hat{S}_y(w)) \cong S_y(w)$$

$$\frac{4V(\hat{S}_y(w))}{S_y(w)^2} \cong 4 \Rightarrow \sqrt{V(\hat{S}_y(w))} \cong S_y(w)$$

La última ecuación tiene al menos dos consecuencias importantes que descartan al periodograma como estimador.

- Los intervalos de confianza son muy amplios y carecen de sentido como tales
- Al tener mucha varianza en cada  $w$  el periodograma muestral es sumamente irregular y presenta muchos picos que no corresponden al proceso.

La estimación no paramétrica del espectro se construye a partir del periodograma muestral  $\hat{S}_y(w)$  promediando sus valores más cercanos a  $w$ .

Para cada  $w$  se promedian los valores del periodograma en frecuencias en el entorno de  $w$ , el estimador esta determinado por la amplitud del rango utilizado para promediar y los ponderadores de cada valor. La estructura de los ponderadores constituyen el núcleo de *kernel* sobre el que se basa la estimación.

Se comienza por definir un conjunto de frecuencias  $B = \{\lambda \in (w_{j-m}, w_{j+m})\}$  de largo  $L = 2m + 1$  y centrada en  $w$  que constituye el rango sobre el que se realiza el promedio.

Este tipo de estimadores esta basado en dos supuestos básicos:

- Se supone que  $S_y(w)$  es relativamente constante en las frecuencias que pertenecen a la banda  $B$ .
- A partir del resultado de Fuller ya mencionado, se plantea que  $\hat{S}_y(w)$  y  $\hat{S}_y(\lambda)$  son aproximadamente independientes en muestras grandes si  $w \neq \lambda$ .



A partir de estas condiciones  $S_y(w)$  es estimado mediante un promedio ponderado de los valores de  $\hat{S}_y(\lambda)$  para los  $\lambda \in B$ .

El promediar variables aproximadamente independientes reduce la variabilidad por lo que debería proporcionar un mejor estimador del espectro que el periodograma muestral.

Así el nuevo estimador de  $S(w_j)$  queda como sigue:

$$S_y^*(w_j) = \sum_{h=-m}^m k(w_{j+h}, w_j) \hat{S}_y(w_{j+h})$$

Los términos  $k(w_{j+h}, w_j)$ , denominados ponderadores de *kernel* asignan cuál es el peso relativo de cada frecuencia, y cumplen que  $\sum_{h=-m}^m k(w_{j+h}, w_j) = 1$ .

La formulación de estos ponderadores es muy general ya que permite que los ponderadores cambien para cada  $w_j$  aunque es más común que un conjunto de ponderadores no dependa de  $j$ , y por otro lado el término  $h$  indica la distancia de cada frecuencia respecto a la frecuencia de interés  $w_j$ , por esta razón en general los ponderadores son decrecientes en  $h$ .

Existen diferentes tipos de funciones *kernel*. Una aproximación que se suele utilizar es  $k(w_{j+h}, w_j)$  proporcional a  $m+1-|h|$  y como  $\sum_{h=-m}^m [m+1-|h|] = (m+1)^2$  los ponderadores propuestos son:

$$k(w_{j+h}, w_j) = \frac{m+1-|h|}{(m+1)^2}$$

El kernel utilizado por defecto en R para la estimación del espectro poblacional es el kernel de Daniell modificado.

Los ponderadores que conforman este kernel tienen la particularidad de que los extremos de la media móvil en conjunto pesan como una observación, es decir si se recuerda que  $L = 2m+1$ , los ponderadores son

$$\left\{ \frac{1}{2(L-1)}, \dots, \frac{1}{(L-1)}, \dots, \frac{1}{2(L-1)} \right\}$$

por ejemplo con  $m=1$   $k = \{\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\}$ .

Se puede mejorar esta aproximación aplicando de forma recursiva este kernel. Retomando el ejemplo anterior se aplica dos veces seguidas este kernel con  $m=1$ , el resultado es una media móvil de largo 5 ( $L=5$ ) con ponderadores:

$$\left\{ \frac{1}{16}, \frac{4}{16}, \frac{6}{16}, \frac{4}{16}, \frac{1}{16} \right\}$$

La elección de  $m$ , el parámetro de ancho de la media móvil, en general se basa en criterios subjetivos.

Mientras que el periodograma es asintóticamente insesgado, aunque con varianza grande, el estimador del espectro poblacional que se basa en el promedio del periodograma para diferentes frecuencias tiene menor varianza, pero genera algún sesgo.

Para estudiar las propiedades asintóticas de  $S_y^*$  se debe exigir una condición adicional para el conjunto de ponderadores, cuando  $\frac{m}{n} \rightarrow 0$  y  $m, n \rightarrow \infty$ , entonces debe cumplirse:

$$\sum_{h=-m}^m k^2(w_{j+h}, w_j) \rightarrow 0$$

Bajo las condiciones anteriores

- $E(S_y^*(w_j)) \rightarrow S_y(w_j)$
- $(\sum_{h=-m}^m k^2(w_{j+h}, w_j))^{-1} V(S_y^*(w_j)) \rightarrow S_y^2(w_j)$
- $\frac{2L_h S_y^*(w_j)}{S_y(w_j)} \sim \chi_{2L_h}^2$

Donde  $L_h = (\sum_{h=-m}^m k^2(w_{j+h}, w_j))^{-1}$ . Las condiciones anteriores indican que  $S_y^*$  es un estimador asintóticamente insesgado, consistente y que reduce la varianza respecto al periodograma muestral. La última de las condiciones permitirá construir intervalos de confianza del estimador.

El sesgo de este estimador proviene del supuesto de que el espectro es constante en las frecuencias consideradas en la media móvil, por lo que la severidad del sesgo depende de la rapidez con que desciende el espectro poblacional y del tamaño del ancho de banda.

Una forma de elegir  $m$  podría basarse en los gráficos del espectro estimado usando diferentes valores de dicho parámetro y hacer un juicio subjetivo de cuáles son los valores más plausibles.

Un estimador del espectro bastante popular utiliza la función *kernel* de Barlett modificando la que está dada por

$$k^* = \begin{cases} 1 - \frac{j}{q+1} & j = 1, 2, \dots, q \\ 0 & \text{para } j > q \end{cases}$$

El estimador de Barlett es entonces

$$\hat{S}_Y(w) = (2\pi)^{-1} \left\{ \hat{\gamma}_0 + 2 \sum_{j=1}^q \left[ 1 - \frac{j}{(q+1)} \right] \hat{\gamma}_j \cos(w_j) \right\}$$

Las autocovarianzas de orden mayor a  $q$  son tratadas como si fueran cero o como si  $Y_t$  fuera un proceso MA( $q$ ).

Las opciones anteriores para estimar el espectro poblacional no suponen ninguna característica sobre el proceso generador de los datos.

Una manera alternativa es especificar un modelo para los datos, estimar los parámetros de dicho modelo y estimar el espectro suplantando los parámetros por sus estimaciones en la fórmula teórica del espectro.

La técnica de estimación espectral autorregresiva desarrollada por Parzen (1974) y Akaike (1969) se basa en identificar un modelo AR(p) y estimar sus parámetros,

$$\hat{S}_y(w) = \frac{\hat{\sigma}_\varepsilon^2}{2\pi} \frac{1}{|1 - \hat{\phi}_1(e^{-iw}) - \dots - \hat{\phi}_p(e^{-ipw})|^2} \quad (16)$$

Esto se conoce como el estimador autoregresivo del espectro, el valor de  $p$  puede ser obtenido utilizando el criterio de Akaike para la selección de modelos.

Se debe tener en cuenta que no es necesario que la serie siga un proceso tipo AR, ya que los procesos estocásticos estacionarios tienen representación  $AR(\infty)$  y este puede ser aproximado por un proceso autorregresivo de orden finito.

La distribución asintótica del estimador de la ecuación 16 ha sido estudiada por Berk (1974), que bajo ciertas condiciones muestra que el estimador espectral autorregresivo es insesgado, consistente y asintóticamente normal.

Permitiendo obtener intervalos de confianza para el valor del espectro, definiendo  $C = \sqrt{\frac{2p}{n}}$  y  $z_{\alpha/2}$  el valor que acumula  $\alpha/2$  de probabilidad en una normal típica se establece que:

$$\frac{\hat{S}_y(w)}{1 + Cz_{\alpha/2}} \leq S_y(w) \leq \frac{\hat{S}_y(w)}{1 - Cz_{\alpha/2}}$$

Box.G; Jenkins.G (1976) «*Time Series Analysis: Forecasting and Control*»

Hamilton. J. (1994) «*Time Series Analysis*». Princeton University Press.

Koopmans.L.H. (1983) «A Spectral Analysis Primer» *Handbook of Statistics, Vol.3*

Piccolo. D (1990) «*Introduzione all'analisi delle serie storiche*»

Shumway.R.; Stoffer,D.(2006) «*Time Series Analysis and its Applications with R Examples*»

Wei.W (2006) «*Time Series Analysis, Univariate and Multivariate Methods*»