

Estimación y Predicción en Series Temporales

Procesos Estocásticos

Departamento de Procesamiento de Señales

Instituto de Ingeniería Eléctrica
Facultad de Ingeniería

2022

Biblio: Haykin, *Adaptive Filter Theory*, 4ta./5ta. edición. Capítulo 1 y Apéndice E.

Veremos dos aspectos básicos:

- 1 Propiedades de la **matriz de correlación** y de la **densidad espectral de potencia**.
 - 2 Modelización de un proceso de este tipo como la salida de un filtro lineal discreto excitado por ruido blanco.
- Un *proceso estocástico* no es tan solo una única función del tiempo, sino que representa un número infinito de realizaciones diferentes del proceso.
 - A una realización particular del proceso le llamamos *serie temporal*.
 - **Def:** Un *proceso estocástico es estrictamente estacionario (SSS)* si sus propiedades estadísticas son invariantes en el tiempo. Específicamente, para que un proceso estocástico representado por la serie $\mathbf{u}(n) = (u(n), u(n-1), \dots, u(n-M+1))$ sea SS, la

Caracterización parcial de un proceso

En la práctica no es posible determinar la función de probabilidad conjunta con muchas variables y debemos recurrir a una caracterización parcial del proceso con el primer y segundo momento.

Consideremos un proceso

$\mathbf{u}(n) = [u(n), u(n-1), \dots, u(n-M+1)]$.

- **Función valor medio:** $\mu(n) = \mathbb{E}[u(n)]$

- **Función de autocorrelación:**

$$r(n, n-k) = \mathbb{E}[u(n)u^*(n-k)], \quad k \in \mathbb{Z}$$

- **Función de autocovarianza:**

$$c(n, n-k) = \mathbb{E}[(u(n) - \mu(n))(u(n-k) - \mu(n-k))^*], \quad k \in \mathbb{Z}$$

- **Obs:** $c(n, n-k) = r(n, n-k) - \mu(n)\mu^*(n-k)$.

Le llamamos a este conjunto de funciones *caracterización parcial* del proceso.

Ventajas de la caracterización parcial:

- Es posible medir o estimar estas cantidades (e.g. mediante procedimientos de Monte Carlo).

Caracterización parcial de un proceso

Si el proceso es SSS, las ecuaciones se reducen a, $\forall n$:

$$\begin{aligned}\mu(n) &= \mu \text{ constante} \\ r(n, n-k) &= r(k) \\ c(n, n-k) &= c(k)\end{aligned}$$

Obs.: en el caso $k = 0$, tenemos $r(0) = \mathbb{E}[|u(n)|^2]$ y $c(0) = \sigma_u^2$.

Def: Un proceso se dice *estacionario en sentido amplio (WSS)* si $\mu(n) = \mu$ constante, y $r(n, n-k) = r(k) \forall n$.

Thm (Doob): Un proceso $\{u(n)\}$ SS es WSS $\Leftrightarrow \mathbb{E}[|u(n)|^2] < +\infty \forall n$. **Obs.:** Si un proceso Gaussiano es WSS \Rightarrow también es SS

Ergodicidad: Teorema ergódico en la media

- Las esperanzas son medias sobre los ensambles o realizaciones.
- También podemos definir medias en el tiempo.
- Si el proceso es **ergódico** en la media, podemos intercambiar ambos promedios. Para esto, debemos probar que **los promedios temporales convergen a los promedios sobre realizaciones**, en algún sentido estadístico. Un criterio de convergencia muy utilizado es el error cuadrático medio.

Sea $u(n)$ WSS, con $\mu = \mathbb{E}[u(n)]$ y

$c(k) = E[(u(n) - \mu)(u(n - k) - \mu)^*]$. Consideremos el promedio temporal $\hat{\mu}(N) = \frac{1}{N} \sum_{n=0}^{N-1} u(n)$.

- $\hat{\mu}(N)$ es un estimador insesgado de μ para todo N .
- **Decimos que $u(n)$ es ergódico en la media en el sentido del MSE si**

$$\lim_{N \rightarrow +\infty} \mathbb{E}[|\hat{\mu}(N) - \mu|^2] = 0$$

$$\mathbb{E}[|\hat{\mu}(N) - \mu|^2] \stackrel{\text{operando}}{=} \frac{1}{N^2} \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} c(n-k)$$
$$\stackrel{l=n-k}{=} \frac{1}{N} \sum_{n=-N+1}^{N-1} (1 - |l|/N) c(l).$$

Tenemos entonces una **CNS para la ergodicidad en media**:

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{n=-N+1}^{N-1} (1 - |l|/N) c(l) = 0$$

(Teorema de ergodicidad en la media)

El teorema de ergodicidad en la media puede ser utilizado para otros promedios temporales del proceso. Por ejemplo, aplicándolo a

$$\hat{r}(k, N) = \frac{1}{N} \sum_{n=0}^{N-1} u(n)u^*(n-k), \quad k = 0, 1, \dots, N-1,$$

diremos que $u(n)$ es **ergódico en correlación en el sentido del MSE** si

$$\lim_{N \rightarrow +\infty} \mathbb{E}[|\hat{r}(k, N) - r(k)|^2] = 0, \quad \forall k$$

Matriz de correlación

Consideramos el proceso estocástico

$$\mathbf{u}^T(n) = [u(n), u(n-1), \dots, u(n-M+1)].$$

Su matriz de (auto)correlación es $\mathbf{R} = \mathbb{E}[\mathbf{u}(n)\mathbf{u}^H(n)]$.

Matriz de correlación de un proceso WSS

- Si el proceso es WSS desaparece la dependencia en n :

$$\mathbf{R} = \begin{bmatrix} r(0) & r(1) & \dots & r(M-1) \\ r(-1) & r(0) & \dots & r(M-2) \\ \vdots & \ddots & \ddots & \vdots \\ r(-M+1) & \dots & r(-1) & r(0) \end{bmatrix}$$

- **Obs:** $r(0) \in \mathbb{R}$, $r(k) \in \mathbb{C} \forall k \neq 0$.

Propiedades de la matriz de correlación de un proceso WSS

Prop. 1: \mathbf{R} es Hermítica

Esto es $\mathbf{R} = \mathbf{R}^H$. Esto surge inmediatamente de la definición de $r(k)$, que verifica $r(-k) = r^*(k)$. Entonces:

- Para definir \mathbf{R} , sólo necesitamos $r(k)$ para $k = 0, 1, \dots, M - 1$.
- Si el proceso es real, entonces \mathbf{R} es real y simétrica.

Prop. 2: \mathbf{R} es Toeplitz

Esto quiere decir que dentro de cada diagonal, sus elementos son iguales.

- Es consecuencia del carácter WSS del proceso.
- Hay algoritmos específicos para inversión y descomposiciones de matrices que aprovechan la estructura Toeplitz para bajar el costo computacional

Propiedades de la matriz de correlación de un proceso WSS

Prop. 3: \mathbf{R} es semi-definida positiva

- Si bien puede ser singular, rara vez ocurre (casos patológicos)
- *Dem:* Sea \mathbf{x} un vector complejo, consideremos $y = \mathbf{x}^H \mathbf{u}(n)$. Tenemos:
 $0 \leq \mathbb{E}[|y|^2] = \mathbb{E}[yy^*] = \mathbb{E}[\mathbf{x}^H \mathbf{u}(n) \mathbf{u}^H(n) \mathbf{x}] = \mathbf{x}^H \mathbf{R} \mathbf{x}$. Para que valga la igualdad, debe haber una dependencia lineal entre las variables aleatorias que constituyen $\mathbf{u}(n)$ (cosa que ocurre esencialmente sólo cuando $\mathbf{u}(n)$ es el muestreo de $K \leq M$ sinusoides).
- Podemos considerar en la práctica que \mathbf{R} es definida positiva.

Prop. 4: Relación entre las matrices \mathbf{R}_M y \mathbf{R}_{M+1}

Valores y vectores propios de la matriz de correlación

- Queremos encontrar $\mathbf{q} \neq \mathbf{0}$ tal que $\mathbf{R}\mathbf{q} = \lambda\mathbf{q}$ para cierto λ .
- De forma equivalente, $(\mathbf{R} - \lambda\mathbf{I})\mathbf{q} = \mathbf{0} \Leftrightarrow \det(\mathbf{R} - \lambda\mathbf{I}) = 0$.
- Veremos propiedades de los valores y vectores propios de \mathbf{R} , muchas de ellas son consecuencia de que \mathbf{R} es hermítica y semi-definida positiva.

Prop. 1: Los v. p. de \mathbf{R} , $\lambda_1, \lambda_2, \dots, \lambda_M$ son reales positivos.

Dem.:

$$\mathbf{R}\mathbf{q}_i = \lambda_i\mathbf{q}_i \Rightarrow \mathbf{q}_i^H \mathbf{R}\mathbf{q}_i = \lambda_i \mathbf{q}_i^H \mathbf{q}_i \Leftrightarrow \lambda_i = \underbrace{\frac{\mathbf{q}_i^H \mathbf{R}\mathbf{q}_i}{\mathbf{q}_i^H \mathbf{q}_i}}_{\text{coef. de Rayleigh de } \mathbf{q}_i} \geq 0.$$

Prop. 2: Sea $k > 0$ entero. Si λ es v.p de \mathbf{R} , λ^k es v.p de \mathbf{R}^k . Si además $\lambda > 0$, λ^{-k} es v.p de \mathbf{R}^{-k} .

Dem.: basta con multiplicar recursivamente por \mathbf{R} la igualdad $\mathbf{R}\mathbf{q} = \lambda\mathbf{q}$.

Valores y vectores propios de la matriz de correlación

Prop. 3: Los \mathbf{q}_i asociados a los v. p. λ_i , $i = 1, \dots, M$ son ortogonales entre sí.

$$\begin{aligned}\mathbf{R}\mathbf{q}_i &= \lambda_i \mathbf{q}_i \Rightarrow \mathbf{q}_j^H \mathbf{R}\mathbf{q}_i = \lambda_i \mathbf{q}_j^H \mathbf{q}_i, \\ \mathbf{R}\mathbf{q}_j &= \lambda_j \mathbf{q}_j \Rightarrow \mathbf{q}_j^H \underbrace{\mathbf{R}^H}_{\mathbf{R}} = \lambda_j \mathbf{q}_j^H \Rightarrow \mathbf{q}_j^H \mathbf{R}\mathbf{q}_i = \lambda_j \mathbf{q}_j^H \mathbf{q}_i.\end{aligned}$$

Restando las ecuaciones en azul, $0 = (\lambda_i - \lambda_j) \mathbf{q}_j^H \mathbf{q}_i$, por lo que $\forall i \neq j$, $\mathbf{q}_j^H \mathbf{q}_i = 0$.

Valores y vectores propios de la matriz de correlación

Prop. 4: Teorema Espectral

Sean $\mathbf{q}_1, \dots, \mathbf{q}_M$ correspondientes a $\lambda_1, \dots, \lambda_M$ distintos.

Sea $\mathbf{Q} = [\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_M]$, con $\mathbf{q}_i^H \mathbf{q}_j = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j. \end{cases}$

Sea $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_M)$.

Entonces:

- \mathbf{Q} es unitaria i.e. $\mathbf{Q}^H \mathbf{Q} = \mathbf{I}$, o de forma equivalente $\mathbf{Q}^{-1} = \mathbf{Q}^H$.
- $\mathbf{\Lambda} = \mathbf{Q}^H \mathbf{R} \mathbf{Q}$.
- $\mathbf{R} = \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^H = \sum_{i=1}^M \lambda_i \mathbf{q}_i \mathbf{q}_i^H$ (descomposición espectral).

Prop. 5

$$\text{Tr}(\mathbf{R}) = \sum_{i=1}^M \lambda_i.$$

Valores y vectores propios de la matriz de correlación

Prop. 6 (surge del Teorema de los círculos de Gershgorin)

$$\lambda_{max} \leq \sum_{k=0}^{M-1} |r(k)|, \quad \lambda_{min} \geq r(0) - \sum_{k=1}^{M-1} |r(k)|.$$

Prop. 7: sobre el número de condición

La matriz de correlación \mathbf{R} es mal condicionada si el cociente $\lambda_{max}/\lambda_{min}$ es grande. Definiciones:

- Número de condición: $\chi(\mathbf{A}) = \|\mathbf{A}\| \|\mathbf{A}^{-1}\|$.
- De forma general la norma de una matriz \mathbf{A} inducida por una norma vectorial dada se define como se define como $\|\mathbf{A}\| = \max \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|}$.
- La norma espectral se define como la norma inducida por la norma 2,

$$\|\mathbf{A}\|_S^2 = \max \frac{\|\mathbf{A}\mathbf{x}\|_2^2}{\|\mathbf{x}\|_2^2} = \max \frac{\mathbf{x}^H \mathbf{A}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{x}} = \lambda_{max}(\mathbf{A}^H \mathbf{A}).$$

Como $\mathbf{A}^H \mathbf{A}$ es hermítica y definida positiva, sus v.p. son positivos, $\Rightarrow \|\mathbf{A}\|_S = \sqrt{\lambda_{max}(\mathbf{A}^H \mathbf{A})}$.

Prop. 7 (número de condición, continuación)

Aplicación de la norma espectral a \mathbf{R} :

$\mathbf{R}^H = \mathbf{R} \Rightarrow$ Si λ_{max} es el mayor v.p. de \mathbf{R} , λ_{max}^2 mayor v.p. de

$\mathbf{R}^H \mathbf{R} \Rightarrow \|\mathbf{R}\|_S = \lambda_{max}$.

De la misma forma, $\|\mathbf{R}^{-1}\|_S = 1/\lambda_{min}$.

Finalmente,

$$\chi(\mathbf{R}) = \frac{\lambda_{max}}{\lambda_{min}}.$$

Si $\chi(\mathbf{R})$ es grande, se dice que \mathbf{R} está mal condicionada.

Resolver problemas inversos con matrices de número de condición grande conduce a fuertes inestabilidades, amplificación del ruido, etc. ¿Porqué? Justificar y discutir posibles soluciones al problema.

Descomposición de Karhunen-Loève

Sea $\mathbf{u}(n) = [u(n), u(n-1), \dots, u(n-M+1)]$ una realización de un proceso WSS, de media nula y matriz de correlación \mathbf{R} . Sean $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_M$ los vectores propios asociados a los valores propios de \mathbf{R} . Entonces vale la expansión

$$\mathbf{u}(n) = \sum_{i=1}^M c_i(n) \mathbf{q}_i, \quad \text{con } c_i(n) = \mathbf{q}_i^H \mathbf{u}(n).$$

¿Cuánto valen media y correlación de los $c_i(n)$?

$$\mathbb{E}[c_i(n)] = \mathbf{q}_i^H \mathbb{E}[\mathbf{u}(n)] = 0, \quad \mathbb{E}[c_i(n) c_j^*(n)] = \begin{cases} \lambda_i & i = j, \\ 0 & i \neq j \end{cases}$$

Además, $\sum_{i=1}^M |c_i(n)|^2 = \|\mathbf{u}(n)\|^2$.

Modelado de Rango Bajo

Supongamos $\lambda_1 > \lambda_2 > \dots > \lambda_M$, y que tenemos información a priori que para $p < M$, los $\lambda_{p+1}, \dots, \lambda_M$ son muy pequeños. Consideramos la aproximación:

$$\hat{\mathbf{u}} = \sum_{i=1}^p c_i(n) \mathbf{q}_i$$

$\mathbf{c}(n) = [c_1(n), c_2(n), \dots, c_p(n)]$ representación de rango bajo de $\mathbf{u}(n)$.

¿Cuánto vale el error cuadrático medio de aproximación?

$$\mathbf{e}(n) = \mathbf{u}(n) - \hat{\mathbf{u}}(n) = \sum_{i=p+1}^M c_i(n) \mathbf{q}_i.$$

$$\begin{aligned} \epsilon &= \mathbb{E}[\|\mathbf{e}(n)\|^2] = \mathbb{E}[\mathbf{e}(n)^H \mathbf{e}(n)] \\ &= \mathbb{E} \left[\sum_{i=p+1}^M \sum_{j=p+1}^M c_i^*(n) c_j(n) \mathbf{q}_i^H \mathbf{q}_j \right] \\ &= \sum_{i=p+1}^M \sum_{j=p+1}^M \mathbb{E}[c_i^*(n) c_j(n)] \mathbf{q}_i^H \mathbf{q}_j = \sum_{i=p+1}^M \lambda_i. \end{aligned}$$

Modelado de Rango Bajo

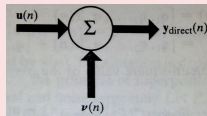
Aplicación: Transmisión de $\mathbf{u}(n)$ por un canal ruidoso

Suponemos ruido del canal $\mathbf{v}(n)$ aditivo, WGN, independiente de la señal. Tenemos

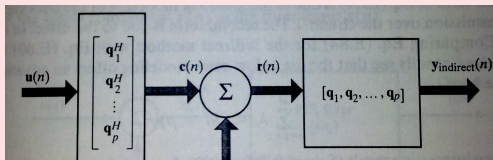
- $\mathbb{E}[\mathbf{u}(n)\mathbf{v}^H(n)] = \mathbf{0}$
- $\mathbb{E}[\mathbf{v}(n)\mathbf{v}^H(n)] = \sigma^2\mathbf{I}$

Transmisión:

- Directa:



- Indirecta:



Aplicación: Transmisión de $\mathbf{u}(n)$ por un canal ruidoso

Transmisión directa:

$$y_D(n) = u(n) + v(n)$$

$$\epsilon_D = \mathbb{E}[\|\mathbf{y}_D(n) - \mathbf{u}(n)\|^2] = \mathbb{E}[\|\mathbf{v}(n)\|^2]$$

$$= \sum_{i=1}^M \mathbb{E}[|v_i(n)|^2] = M\sigma^2.$$

Transmisión indirecta:

$$\begin{cases} \mathbf{c}(n) = \underbrace{[\mathbf{q}_1, \dots, \mathbf{q}_p, \mathbf{0}, \dots, \mathbf{0}]^H}_{\tilde{\mathbf{Q}}^H} \mathbf{u}(n) \\ \mathbf{r}(n) = \mathbf{c}(n) + \mathbf{v}(n) \end{cases}$$

$$\Rightarrow \mathbf{y}_I = \tilde{\mathbf{Q}}^H [c_1(n), \dots, c_p(n), 0, \dots, 0]^T + \tilde{\mathbf{Q}}^H [v_1(n), \dots, v_p(n), 0, \dots, 0]^T$$

Aplicación: Transmisión de $\mathbf{u}(n)$ por un canal ruidoso

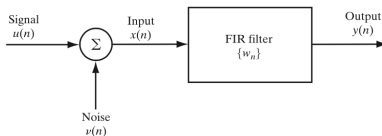
Transmisión indirecta (cont):

$$\begin{aligned}\Rightarrow \epsilon_I &= \mathbb{E}[\|\mathbf{y}_I(n) - \mathbf{u}(n)\|^2] = \mathbb{E} \left[\left\| \underbrace{\hat{\mathbf{u}}(n)}_{-\sum_{i=p+1}^M \sqrt{\lambda_i} \mathbf{q}_i} + \sum_{i=1}^p v_i(n) \mathbf{q}_i \right\|^2 \right] \\ &= \sum_{i=p+1}^M \lambda_i + p\sigma^2 \quad (= \text{error de aprox. señal} + \text{ruido filtrado}).\end{aligned}$$

Ejercicio: establecer en qué condiciones es mejor la transmisión indirecta a la directa.

$$\epsilon_I < \epsilon_D \Leftrightarrow \sum_{i=p+1}^M \lambda_i + p\sigma^2 < M\sigma^2 \Leftrightarrow \sum_{i=p+1}^M \lambda_i < (M-p)\sigma^2.$$

Filtros propios (Eigenfilters)



- $u(n)$ WSS, media cero, mat. de correlación \mathbf{R} , $v(n)$ WGN, σ^2
- $u(n)$ y $v(n)$ no correlacionados: $\mathbb{E}[u(n)v^*(m)] = 0 \quad \forall m, n$.

Objetivo: encontrar el filtro óptimo \mathbf{w} que maximice la SNR a la salida.

Ejercicio: Calcular la potencia media a la salida, y sus contribuciones debido a la señal y al ruido

$$y(n) = \mathbf{w} * (\mathbf{u} + \mathbf{v})(n) = \sum_{k=0}^{M-1} w(k)(u(n-k) + v(n-k))$$

$$\begin{aligned} \mathbb{E}[|y(n)|^2] &= \sum_{k,k'} w(k) \mathbb{E}[(u(n-k) + v(n-k))(u(n-k') + v(n-k'))^*] \\ &= \sum_k w(k) (\mathbb{E}[u(n-k)u(n-k')^*] + \mathbb{E}[|v(n-k)|^2] \delta(k-k')) \end{aligned}$$

Filtros propios (Eigenfilters)

Ejercicio: encontrar \mathbf{w} que maximice $SNR = \frac{\mathbf{w}^H \mathbf{R} \mathbf{w}}{\sigma^2 \mathbf{w}^H \mathbf{w}}$.

Filtro óptimo:

$$\left\{ \begin{array}{l} \max_{\mathbf{w}} SNR \\ \text{sujeto a } \mathbf{w}^H \mathbf{w} = 1 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} \max_{\mathbf{w}} \mathbf{w}^H \mathbf{R} \mathbf{w} \\ \text{sujeto a } \mathbf{w}^H \mathbf{w} = 1 \end{array} \right.$$

\Leftrightarrow

\mathbf{w} vector propio asociado al mayor valor propio de \mathbf{R} ($\lambda_{max}, \mathbf{q}_{max}$)

$$\left\{ \begin{array}{l} \mathbf{w}_{opt} = \mathbf{q}_{max} \\ SNR_{opt} = \frac{\lambda_{max}}{\sigma^2} \end{array} \right.$$

Densidad Espectral de Potencia

La matriz de correlación es una descripción temporal de las propiedades estadísticas de segundo orden del proceso.

Aplicando la DTFT a la secuencia de autocorrelaciones $\{r(k)\}$ obtenemos la **densidad espectral de potencia (o espectro de potencia) del proceso**.

Def.:

$$S(\omega) = \sum_{k=-M+1}^{M+1} r(k) \exp(-i\omega k), \quad -\pi \leq \omega < \pi.$$

Tenemos además la fórmula de inversión

$$r(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S(\omega) \exp(i\omega k) d\omega, \quad k = 0, \pm 1, \pm 2, \dots, \pm M - 1.$$

Densidad Espectral de Potencia: Filtrado de un proceso

$$\mathbf{u}(n), S_u(\omega) \longrightarrow \boxed{\{h(l)\}, H(\omega)} \longrightarrow \mathbf{y}(n), S_y(\omega)$$

¿Cuánto vale $S_y(\omega)$?

Para cada realización de $\mathbf{u}(n)$, tenemos $y(n) = \sum_{l=0}^{M-1} h(l)u(n-l)$.

$$\begin{aligned} r_y(k) &= \mathbb{E}[y(n)y^*(n-k)] = \sum_{l=0}^{M-1} \sum_{l'=0}^{M-1} h(l)h^*(l')\mathbb{E}[u(n-l)u^*(n-k-l')] \\ &= \sum_{l=0}^{M-1} \sum_{l'=0}^{M-1} h(l)h^*(l')r_u(k+l'-l) = \sum_{l=0}^{M-1} h(l) \sum_{l'=0}^{M-1} h^*(l')r_u(k+l'-l) \end{aligned}$$

Entonces

$$\begin{aligned} S_y(\omega) &= \sum_k r_y(k) \exp(-i\omega k) \\ &= \sum_{l=0}^{M-1} h(l) \exp(-i\omega l) \sum_{l'=0}^{M-1} h^*(l') \exp(i\omega l') \sum_k r_u(k+l'-l) \exp(-i\omega(k+l'-l)) \\ &= H(\omega)H^*(\omega)S_u(\omega) = |H(\omega)|^2 S_u(\omega). \end{aligned}$$

Densidad Espectral de Potencia: Propiedades

Prop.1: La DEP de un proceso WSS discreto es 2π -periódica

$$\begin{aligned} S(\omega + 2n\pi) &= \sum_{k=-M+1}^{M-1} r(k) \exp(-i(\omega + 2n\pi)k) \\ &= \sum_{k=-M+1}^{M-1} r(k) \exp(-i2n\pi k) \exp(-i\omega k) = S(\omega) \end{aligned}$$

Prop.2: La DEP de un proceso WSS discreto es real

$$\begin{aligned} S(\omega) &= r(0) + \sum_{k=1}^{M-1} r(k) \exp(-i\omega k) + \sum_{k=-M+1}^{-1} r(k) \exp(-i\omega k) \\ &= r(0) + \sum_{k=1}^{M-1} r(k) \exp(-i\omega k) + \sum_{k=1}^{M-1} r(-k) \exp(i\omega k) \end{aligned}$$

Prop.3: La DEP de un proceso WSS discreto real es par

$$\begin{aligned} S(-\omega) &= \sum_{k=-M+1}^{M-1} r(k) \exp(i\omega k) \stackrel{l=-k}{=} \sum_{l=-M+1}^{M-1} r(-l) \exp(-i\omega l) \\ &= \sum_{l=-M+1}^{M-1} r^*(l) \exp(-i\omega l) \stackrel{r^*(l)=r(l)}{=} \sum_{l=-M+1}^{M-1} r(l) \exp(-i\omega l) = S(\omega) \end{aligned}$$

Prop.4: El área bajo la curva $\omega \mapsto S(\omega)/2\pi$ es el valor cuadrático medio del proceso, o su potencia

Si $k = 0$, de la fórmula de inversión

$$r(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S(\omega) d\omega.$$

Prop.5: La DEP es positiva

Definimos el filtro $H(\omega) := \mathbb{I}_{\{|\omega - \omega_c| \leq \Delta\omega\}}$. La salida y de filtrar $u(n)$ por H conduce a $S_y(\omega) = S_u(\omega) |H(\omega)|^2$. Luego,

$$\begin{aligned} 0 \leq \mathbb{E}[y^2(n)] &= r_y(0) = \frac{1}{2\pi} \int_{\omega_c - \Delta\omega}^{\omega_c + \Delta\omega} S_u(\omega) \\ &\sim_{\Delta\omega \rightarrow 0} \frac{1}{2\pi} \int_{\omega_c - \Delta\omega}^{\omega_c + \Delta\omega} \left(S_u(\omega_c) + \frac{dS}{d\omega}(\omega_c)(\omega - \omega_c) \right) d\omega = \frac{1}{2\pi} 2\Delta\omega S \end{aligned}$$

$$\Rightarrow S(\omega_c) \geq 0 \quad \forall \quad \omega_c \in [-\pi, \pi).$$

Prop.6: Si λ_m , $m = 1, \dots, M$ son los v.p de \mathbf{R} , entonces $\min_{\omega} S(\omega) \leq \lambda_m \leq \max_{\omega} S(\omega)$. Además, $\chi(\mathbf{R}) = \frac{\lambda_{max}}{\lambda_{min}} \leq \frac{S_{max}}{S_{min}}$.

Sabemos que $\lambda_m = \frac{\mathbf{q}_m^H \mathbf{R} \mathbf{q}_m}{\mathbf{q}_m^H \mathbf{q}_m}$, con $\mathbf{q}_m = [q_m(1), q_m(1), \dots, q_m(M)]^T$.

$$\begin{aligned} \mathbf{q}_m^H \mathbf{R} \mathbf{q}_m &= \sum_{k=1}^M \sum_{l=1}^M q_m(k)^* r(l-k) q_m(l) \\ &= \frac{1}{2\pi} \sum_{k=1}^M \sum_{l=1}^M q_m(k)^* q_m(l) \int_{-\pi}^{\pi} S(\omega) \exp(i\omega(l-k)) d\omega \end{aligned}$$

Prop. 6 (cont.)

Si llamamos $Q'(\omega)$ a la DTFT de \mathbf{q}_m^* , i.e.

$Q'(\omega) = \sum_{k=1}^M \mathbf{q}_m^*(k) \exp(-i\omega k)$, tenemos que

$$\mathbf{q}_m^H \mathbf{R} \mathbf{q}_m = \frac{1}{2\pi} \int_{-\pi}^{\pi} S(\omega) |Q'(\omega)|^2 d\omega.$$

De la misma forma,

$$\mathbf{q}_m^H \mathbf{q}_m = \frac{1}{2\pi} \int_{-\pi}^{\pi} |Q'(\omega)|^2 d\omega.$$

Luego,

$$\lambda_m = \frac{\int_{-\pi}^{\pi} S(\omega) |Q'(\omega)|^2 d\omega}{\int_{-\pi}^{\pi} |Q'(\omega)|^2 d\omega},$$

por lo que

$$\min S(\omega) \leq \lambda_m \leq \max S(\omega).$$