

Procesos Estocásticos: Introducción

Cecilia Galarza

Procesos Estocásticos
Facultad de Ingeniería, Universidad de Buenos Aires

1er cuatrimestre 2025

Procesos vectoriales

Función de correlación cruzada entre procesos

En muchos análisis es interesante entender cómo se comportan dos o más procesos en forma conjunta.

Definición (CCF, *Cross-Correlation Function*)

Dados dos procesos escalares $X(t)$ y $Y(t)$, la función de correlación cruzada es

$$R_{X,Y}(t_1, t_2) = \mathbb{E}[X(t_1) Y^T(t_2)].$$

puede no estar

$$R_{X,Y}(t_1, t_2) = E[X(t_1) \cdot Y(t_2)] \stackrel{?}{=} E[X(t_2) \cdot Y(t_1)] = R_{Y,X}(t_2, t_1)$$

Sí es correcto que:

$$R_{X,Y}(t_1, t_2) = R_{Y,X}(t_2, t_1)$$

→ No necesariamente van a ser iguales

Procesos conjuntamente ESA (CESA)

Al considerar dos procesos, podemos analizar su estacionariedad en forma conjunta también. Para ello, extendemos el concepto de ESA del siguiente modo:

Definición (Procesos Conjuntamente ESA)

Sean $X(t)$, $Y(t)$ dos procesos ESA. Decimos que son Conjuntamente Estacionarios en Sentido Amplio (CESA) si

$$R_{X,Y}(t, t + \tau) = f(\tau).$$

Procesos CESA

Sean X, Y dos procesos CESA, $\mathbb{E}[X(t)Y^T(t + \tau)] = R_{X,Y}(\tau)$.

- Si $R_{X,Y}(\tau) = 0$ para todo τ , X e Y son procesos ortogonales .
- Si X e Y son independientes entonces $R_{X,Y}(\tau) = \mu_X\mu_Y$ para todo τ .
- $R_{X,Y}(\tau) = R_{Y,X}^*(-\tau)$.

Por definición, $R_{X,Y}(\tau) = \mathbb{E}[X(t)Y^T(t + \tau)] = \mathbb{E}[Y^T(t + \tau)X(t)] = \mathbb{E}[Y^T(t')X(t' - \tau)] = R_{Y,X}^*(-\tau)$.

$R_{X,Y}(\tau)$ puede ser distinto a $R_{Y,X}^*(+\tau)$. Por ende, $R_{X,Y}(\tau)$ puede NO ser par.

Nota con estas diapositivas, tienen pequeños errores

- $|R_{X,Y}(\tau)| \leq \sqrt{R_X(0)R_Y(0)}$.

Es una consecuencia directa de la desigualdad de Cauchy-Schwarz.

Ejemplo: dos procesos senoidales

$X_1(t) = A_1 \cos(\omega_1 t + \Theta_1), \quad X_2(t) = A_2 \cos(\omega_2 t + \Theta_2), \quad t \in \mathbb{R}$,
 donde $A_1, A_2 \sim \mathcal{N}(0, 1)$, $\Theta_1, \Theta_2 \sim \mathcal{U}(0, 2\pi)$ independientes.

- Vimos que $\mu_{X_i} = 0$ y $R_{X_i}(\tau) = \frac{1}{2} \cos(\omega_i \tau)$, $i = 1, 2$.
- Con argumentos similares, tenemos que

$$\begin{aligned} R_{X_1, X_2}(t, t + \tau) &= \mathbb{E}[X_1(t)X_2(t + \tau)] \\ &= \mathbb{E}[A_1]\mathbb{E}[\cos(\omega_1 t + \Theta_1)]\mathbb{E}[A_2]\mathbb{E}[\cos(\omega_2(t + \tau) + \Theta_2)] = 0, \end{aligned}$$

que no depende de t . Por lo tanto, X_1 y X_2 son PE CESA.

Ejemplo: dos procesos de tiempo discreto CESA

- Consideremos los procesos

$$X_1(n) = W(n), \quad X_2(n) = W(n) + 2W(n-1), \quad n \in \mathbb{Z},$$

donde W es ruido blanco de media nula y varianza unitaria.

- Por construcción, $\mu_{X_1} = 0$ y $R_{X_1}(k) = \delta(k)$. Para el proceso X_2 tenemos que

$$\mathbb{E}[X_2(n)] = \mathbb{E}[W(n) + 2W(n-1)] = 0$$

$$\begin{aligned}\mathbb{E}[X_2(n)X_2(n+k)] &= \mathbb{E}\{[W(n) + 2W(n-1)][W(n+k) + 2W(n+k-1)]\} \\ &= 5\delta(k) + 2\delta(k-1) + 2\delta(k+1).\end{aligned}$$

Entonces X_1 y X_2 son ESA.

- La función de correlación cruzada entre X_1 y X_2 es

$$R_{X_1, X_2}(n, n+k) = \mathbb{E}[X_1(n)X_2(n+k)] = \delta(k) + 2\delta(k-1),$$

que no depende de n . Del mismo modo, $R_{X_2, X_1}(n, n+k) = \delta(k) + 2\delta(k+1)$. Por lo tanto, X_1 y X_2 son PEs CESA.

Densidad espectral de potencia cruzada

- La *densidad espectral de potencia cruzada* caracteriza la relación entre procesos CESA en el dominio de la frecuencia

$$S_{X,Y}(\omega) = \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\left(\int_{-T/2}^{T/2} X(t) e^{-j\omega t} dt \right) \left(\int_{-T/2}^{T/2} Y(t) e^{+j\omega t} dt \right)^T \right].$$

ESTO es como una proyección (paralela)

- Se extiende el teorema de Wiener-Khinchin a este caso resultando

$$S_{X,Y}(\omega) = \mathcal{F}\{R_{X,Y}\}(\omega) = \int_{\mathbb{R}} R_{X,Y}(\tau) e^{-j\omega\tau} d\tau.$$

De forma similar, en tiempo discreto tenemos que

$$S_{X,Y}(\omega) = \mathcal{F}\{R_{X,Y}\}(\omega) = \sum_{k \in \mathbb{Z}} R_{X,Y}(k) e^{-j\omega k}.$$

Propiedades de la PSD cruzada

- En general, la PSD cruzada no es real porque $R_{X,Y}$ no es una función par. Sin embargo, satisface

$$S_{X,Y}^*(\omega) = S_{Y,X}(\omega).$$

Por otro lado, si $R_{X,Y}$ es una función real, la PSD cruzada es hermítica:

→ HERMÍTICA

$$S_{X,Y}^*(-\omega) = S_{X,Y}(\omega).$$

Compadre

La comcha de tu madre

Ejemplos

- Si X e Y son ortogonales, $R_{X,Y}(\tau) = 0$ y entonces $S_{X,Y}(\omega) = 0$.
- Sea X un proceso ESA. Armamos el proceso Y como

$$Y(t) = X(t - t_0), \quad t \in \mathbb{R},$$

donde $t_0 \in \mathbb{R}$ es un parámetro. Vemos que

$$R_X(\tau) = R_Y(\tau), \quad R_{X,Y}(\tau) = R_X(\tau - t_0).$$

Por lo tanto,

$$S_X(\omega) = S_Y(\omega), \quad S_{X,Y}(\omega) = e^{-j\omega t_0} S_X(\omega).$$

Parece q' ésta es LA propiedad

Procesos estocásticos vectoriales

Al considerar varios procesos en forma conjunta, es interesante plantear procesos de tipo vectorial.

Decimos que $\mathbf{X}(\xi, t)$ es un PE vectorial (PEV) si

$$\mathbf{X}(\xi, t) = \begin{bmatrix} X_1(\xi, t) \\ \vdots \\ X_n(\xi, t) \end{bmatrix}, \quad \begin{array}{l} \text{misma muestra} \\ \text{distintos o los} \end{array}$$

donde cada componente es un PE escalar.

Observen que el experimento aleatorio asociado, ξ , es el mismo para todos los componentes del vector.

¿Utiliza o importante?

Caracterización de PEs vectoriales

La media del PEV es el vector de medias. No preguntes por los calzoncillos. cuack.

Definimos la media y la función de autocorrelación del PEV como

$$\boldsymbol{\mu}_{\mathbf{X}}(t) = \mathbb{E}[\mathbf{X}(t)] = \begin{bmatrix} \mathbb{E}[X_1(t)] \\ \vdots \\ \mathbb{E}[X_n(t)] \end{bmatrix} = \begin{bmatrix} \mu_{X_1}(t) \\ \vdots \\ \mu_{X_n}(t) \end{bmatrix},$$

$$\begin{aligned} \mathbf{R}_{\mathbf{X}}(t, t + \tau) &= \mathbb{E}[\mathbf{X}(t)\mathbf{X}^T(t + \tau)] \\ &= \begin{bmatrix} R_{X_1, X_1}(t, t + \tau) & \cdots & R_{X_1, X_n}(t, t + \tau) \\ \vdots & \ddots & \vdots \\ R_{X_n, X_1}(t, t + \tau) & \cdots & R_{X_n, X_n}(t, t + \tau) \end{bmatrix}. \end{aligned}$$

Proceso vectorial ESA

Definición (Proceso vectorial ESA)

Un proceso aleatorio vectorial $\mathbf{X}(t) \in \mathbb{C}^n$ es ESA si:

$$\mathbb{E}[\mathbf{X}(t)] = \mu_{\mathbf{X}}(t) = \mu_{\mathbf{X}}$$

$$\mathbf{R}_{\mathbf{X}}(t, t + \tau) = \mathbb{E}[\mathbf{X}(t)\mathbf{X}^T(t + \tau)] = \mathbf{R}_{\mathbf{X}}(\tau).$$

Las condiciones de estacionariedad requieren que:

- Cada componente del PEV sea ESA, de modo que

$$\mathbb{E}[X_i] = \mu_i \quad \text{y} \quad R_{X_i, X_i}(t, t + \tau) = R_{X_i}(\tau) \quad \forall i = 1, \dots, n.$$

- Las componentes sean CESAs entre sí, de modo que

$$R_{X_i, X_j}(t, t + \tau) = R_{X_i, X_j}(\tau), \quad \forall i \neq j.$$

$\mathbf{X}(t)$ es ESA $\iff X_1, \dots, X_n$ son CESAs.

Ejemplo: Ruido blanco vectorial

Sean X_i , $i = 1, \dots, n$, procesos de ruido blanco con $\mathbb{E}[X_i] = 0$ y $\mathbf{R}_{X_i}(\tau) = \alpha_i \delta(k)$ independientes entre sí. Definimos el ruido blanco vectorial como

→ Autocorr.

$$\mathbf{X}(t) = \begin{bmatrix} X_1(t) \\ \vdots \\ X_n(t) \end{bmatrix}, \quad t \in \mathbb{R}.$$

- $\mathbb{E}[\mathbf{X}(t)] = \mu_{\mathbf{X}}(t) = \mathbf{0}$, $\forall t \in \mathbb{R}$.

- $\mathbf{R}_{\mathbf{X}}(t, t + \tau) = \mathbf{R}_{\mathbf{X}}(\tau) = \begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \alpha_n \end{bmatrix} \delta(\tau)$.

Estimadores de la correlación

Procesos ergódicos: un tema para pensar

En muchas aplicaciones, sólo tenemos acceso a realizaciones del proceso y no a una descripción analítica de su estadística. Es utilizando sólo estas observaciones que hay que estimar μ_X o R_X .

$$\mu_X(t) = \mathbb{E}[X(\xi, t)] \simeq \frac{1}{N} \sum_{n=1}^N X(\xi_n, t) \quad \xi_n \in \Xi \quad , \quad N \text{ grande}$$

Repetiendo el experimento un número grande de veces y promediando las realizaciones, nos estaremos acercando a la media como dice la Ley de Grandes Números.

Procesos ergódicos: un tema para pensar

Sin embargo, no siempre es posible repetir el experimento y es importante buscar condiciones bajo las cuales el promedio temporal de una única realización permita obtener la media del proceso. O sea, si ξ_o identifica a la realización particular registrada, buscamos condiciones bajo las cuales

$$\mu_X \simeq \langle X(t) \rangle_T = \frac{1}{T} \int_0^T X(\xi_o, t) dt, \quad \text{en tiempo continuo}$$

$$\mu_X \simeq \langle X(n) \rangle_M = \frac{1}{M} \sum_{n=0}^M X(\xi_o, n), \quad \text{en tiempo discreto}$$

En ese caso se dice que el proceso es *ergódico en la media*.

Procesos ergódicos

Teorema (Ergodicidad en la media)

Un proceso ESA $X(t)$ es ergódico en la media, es decir

$$\lim_{T \rightarrow \infty} \mathbb{E} \left\{ [\langle X(t) \rangle_T - \mu_X]^2 \right\} = 0$$

si y sólo si

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T}^T \left(1 - \frac{|u|}{T} \right) C_X(u) du = 0 \quad \text{en tiempo continuo}$$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=-(N-1)}^{N-1} \left(1 - \frac{|k|}{N} \right) C_X(k) = 0 \quad \text{en tiempo discreto}$$

Esta condición pide que C_X decaiga a 0 rápidamente, de modo que las muestras resulten descorrelacionadas rápidamente. Así, al observar una realización suficientemente larga, estamos circulando por muchas realizaciones del ensamble.

Estimación de los indicadores de procesos ESA

Vamos a utilizar los conceptos de ergodicidad para estimar la R_X de un proceso ESA. Supongamos que sólo tenemos acceso a realizaciones del proceso y no una descripción analítica de su estadística. Sin embargo, podemos concluir, a partir de observaciones adicionales, que el proceso es ESA y ergódico. En ese caso, queremos estimar las funciones R_X y S_X .

Estimador de R_X

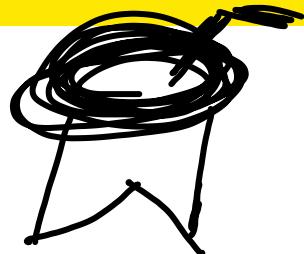
Sea $X(n)$ un proceso ESA en tiempo discreto. Recordemos que

$$R_X(k) = \mathbb{E}[X(n)X(n+k)].$$

Si no tenemos una descripción analítica de $F_{X(n)X(n+k)}$ cómo hacemos para estimar R_X ?

Una idea es utilizar las N muestras de una realización $x(n), n = 0, \dots, N - 1$.

Estimadores de la correlación: estimador insesgado



Definición (Estimador insesgado de la correlación)

El estimador insesgado de $R_X(k)$ es

En los bordes, la estimación es peor, me faltan datos

$$\hat{R}_X(k) = \frac{1}{N-k} \sum_{n=0}^{N-k-1} x(n)x^T(n+k), \quad 0 \leq k \leq N-1$$

Si hacés el dibujito tiene sentido $\hat{R}_X(-k) = \hat{R}_X(k)$
 Es claro que \hat{R}_X depende de la realización observada.

Si el proceso fuera vectorial:

$$\hat{R}_{\vec{x}}(k) = \frac{1}{N-k} \sum_{m=0}^{N-k-1} \vec{x}(m) \cdot \vec{x}^T(m+k)$$

Señales era más directido

Estimador insesgado: Análisis del sesgo

Considerando la aleatoriedad de $\hat{R}_X(k)$, calculamos su esperanza.

$$\mathbb{E}[\hat{R}_X(k)] = \frac{1}{N-k} \sum_{n=0}^{N-k-1} \underbrace{\mathbb{E}[X(n)X^T(n+k)]}_{R_X(k) \text{ constante}}$$

Por definición
de multiplicación
(literal. No entiendo x_3')
hizo énfasis acá, pero se ve q'alguna se confundió con esto)
 $\forall k, \mathbb{E}[\hat{R}_X(k)] = R_X(k)$ entonces el estimador es *insesgado*.

Problema: Para k cercano a N , la estimación promedia pocos valores, lo que conduce en general a valores más erráticos.

Estimadores de la correlación: estimador sesgado

Definición (Estimador sesgado de la correlación)

El estimador sesgado de $R_X(k)$ es

$$\hat{R}_X(k) = \frac{1}{N} \sum_{n=0}^{N-k-1} x(n)x^T(n+k), \quad 0 \leq k \leq N-1$$

$$\hat{R}_X(-k) = \hat{R}_X(k)$$

Estimador sesgado: Análisis del sesgo

En este caso, la esperanza es

$$\begin{aligned}\mathbb{E}[\hat{R}_X(k)] &= \frac{1}{N} \sum_{n=0}^{N-k-1} \mathbb{E}[X(n)X^T(n+k)] \\ &= \frac{N-k}{N} R_X(k) = \left(1 - \frac{k}{N}\right) R_X(k), \quad 0 \leq k \leq N-1 \\ &= \underbrace{\left(1 - \frac{|k|}{N}\right)}_{v(k)} R_X(k), \quad -N+1 \leq k \leq N-1\end{aligned}$$

Otra vez
la ventaja esta → ESTO agrega un sesgo y aplasta a los
extremos.

Todo esto xq 'me faltan datos. NO SE BANCAN LOS
DATOS, ZURDITOS

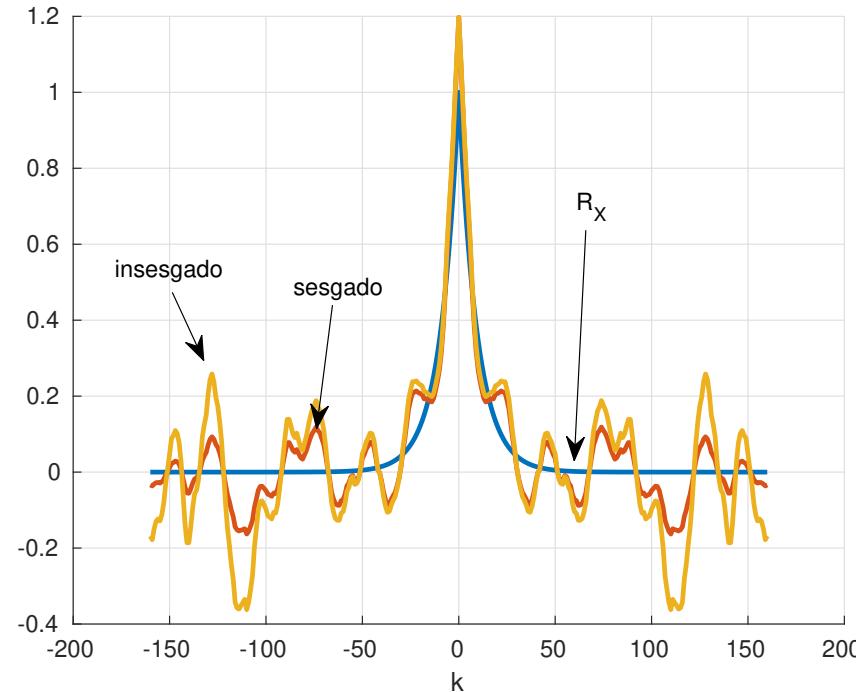
Estimador sesgado: Análisis del sesgo

- $\mathbb{E}[\hat{R}_Y(k)] \neq R_Y(k)$, entonces el estimador es *sesgado*.
- $v(k) = \left(1 - \frac{|k|}{N}\right)$ es una ventana triangular (ventana de Barlett) que se anula en $|k| = N$.
- De este modo, se mejora la performance para k grande.

Estimador sesgado e insesgado de $R_X(k)$

Autocorrelación $R_X(k)$ de proceso ESA.

EL ROJO ES MÁS
ACHATADO EN LOS
EXTREMOS Y COMO
CORRESPONDE



- Curva azul: $R_X(k)$ calculada utilizando la estadística de $X(n)$
- Curva roja: $\hat{R}_X(k)$ con el estimador sesgado
- Curva naranja: $\hat{R}_X(k)$ con el estimador insesgado

Estimador de la densidad espectral

Definición (Periodograma)

Sea $\hat{R}_X(k)$ el estimador sesgado de la autocorrelación. Luego, el periodograma es

$$\hat{S}_X(\omega) = \sum_{k=-N+1}^{N-1} \hat{R}_X(k) e^{-j\omega k}.$$

El periodograma es un estimador de la densidad espectral de potencia. Veremos su comportamiento estadístico.

Sesgo del periodograma

Considerando la esperanza del estimador sesgado, tenemos:

Esperanza del periodograma:

$$\mathbb{E} [\hat{S}_X(\omega)] = \sum_{k=-(N-1)}^{N-1} \mathbb{E} [\hat{R}_X(k)] e^{-j\omega k}$$

$$= \sum_{k=-(N-1)}^{N-1} v(k) R_X(k) e^{-j\omega k}.$$

¿QUÉ ES
ESTA FÓRMULA?

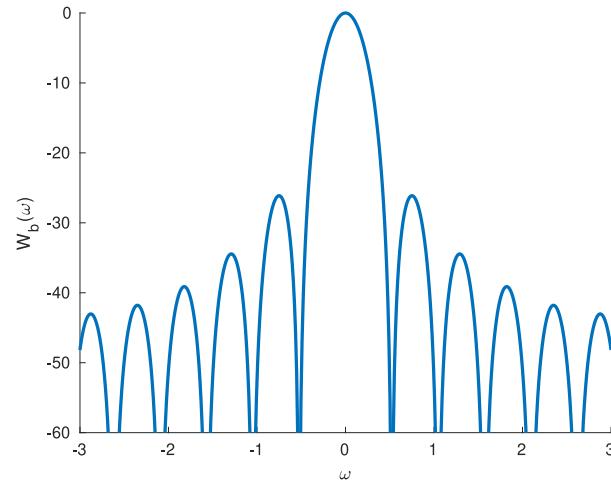
Sea $V(\omega) = \mathcal{F}(v)$. Utilizando las propiedades de Fourier,

$$\mathbb{E} [\hat{S}_X(\omega)] = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_X(\omega - \psi) V(\omega - \psi) d\psi , \quad V(\omega) = \frac{1}{N} \left[\frac{\sin(N\omega/2)}{\sin(\omega/2)} \right]^2$$

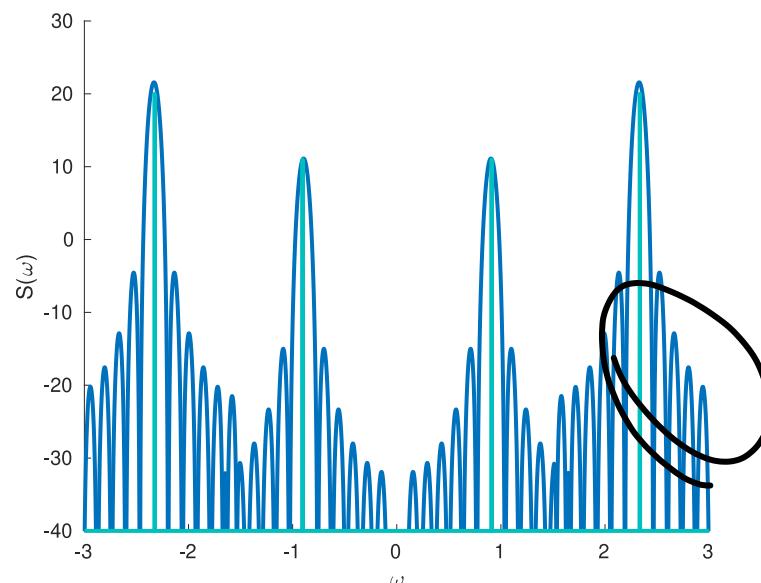
Cuando N crece, la ventana se angosta en frecuencia y el sesgo tiende a cero. Decimos que \hat{S}_X es *asintóticamente insesgado*.

Sesgo del periodograma

Ejemplo: Análisis de una suma de dos senoides



$V(\omega)$: DTFT de la ventana de Barlett
Ancho lóbulo principal: $8\pi/N$



$S(\omega)$: Densidad espectral de potencia
ESTO parece ruido, pero es la DTFT de mi ventana
¿Cuál es la mejor ventana?

Sesgo del periodograma

En líneas generales, el ventaneo introduce los siguientes fenómenos:

- Los lóbulos secundarios producen un escurrimiento o *leakage* de potencia.
- El lóbulo principal suaviza o produce *smearing* (borroneado) del espectro.
- El lóbulo principal de Barlett decrece a la mitad en $1/N$. Luego, f_1 , f_2 van a quedar confundidas bajo el mismo lóbulo principal si

$$|f_2 - f_1| \lesssim 1/N,$$

- La resolución del periodograma es del orden de $1/N$ (frecuencia discreta).

Ejemplo: resolución del periodograma (I)

- Consideremos el siguiente modelo

$$Y(n) = \sin [\omega_0 n] + \sin \left[2\pi \left(f_0 + \frac{\alpha}{N} \right) n \right], \quad n = 1, \dots, N$$

α es la diferencia de frecuencia entre las senoides.

- Consideramos $N = 256$, de modo que la resolución debería ser, aprox:

$$1/N = 0,0039.$$

Ejemplo: resolución del periodograma (II)

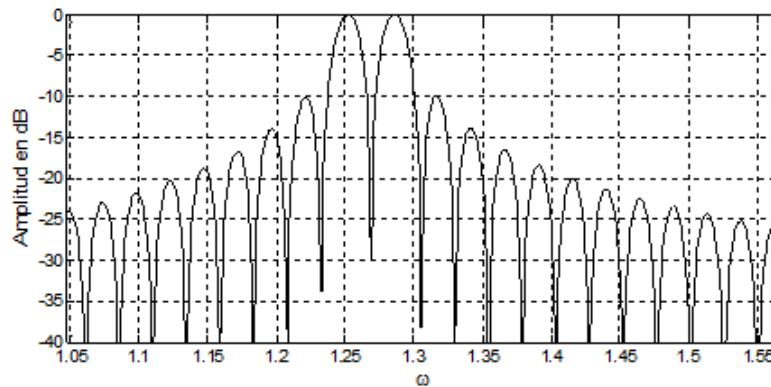


Figura: Espectro para $\alpha = 1$. $\frac{\alpha}{N} = 0,039$. FFT de 8192 puntos

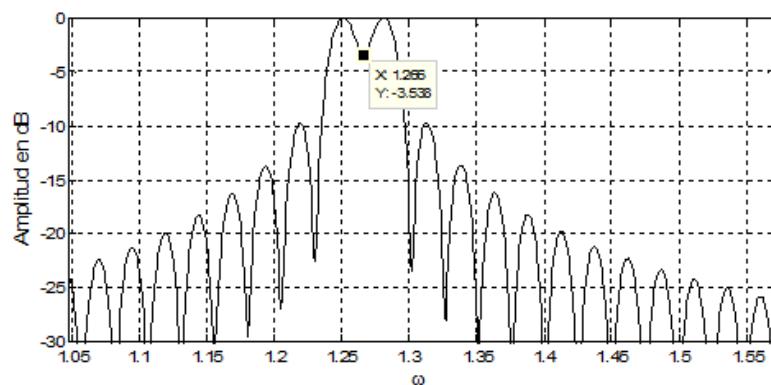


Figura: Espectro para $\alpha = 0,75$. $\frac{\alpha}{N} = 0,029 < 0,039$. FFT de 8192 puntos

Ejemplo: resolución del periodograma (III)

Observaciones:

- La resolución se ve afectada también por la fase relativa de las señales.
- Si las señales tienen diferente amplitud, una puede enmascarar a la otra.