

(86.09) PROCESOS ESTOCÁSTICOS

TRABAJO PRACTICO DE SIMULACIÓN N° 2: ESTIMACIÓN ESPECTRAL

Curso 02

2 de Agosto de 2020

Integrante		Padrón		Correo electrónico
Mendoza, Leonel	—	101153	—	lmendoza@fi.uba.ar
Rossi, Francisco	—	99540	—	frrossi@fi.uba.ar
Demarco, Facundo Agustin	—	100620	—	fdemarco@fi.uba.ar

Índice

1. Técnicas no paramétricas	1
1.1. Ejercicio 1 - Estimador sesgado e insesgado de la autocorrelación	1
1.1.a. Autocovarianza teórica	1
1.1.b. Estimador insesgado de la autocovarianza:	1
1.1.c. Estimador sesgado de la autocovarianza:	1
1.2. Ejercicio 2 - (In)consistencia del periodograma	3
1.2.a. 1. Grafiqué la verdadera densidad espectral de potencia de Y , S_Y	3
1.2.b. 2. Periodograma promedio y varianza del periodograma	4
1.2.c. 3. Periodograma con ventana de Bartlett	5
2. Técnicas paramétricas	5
2.1. Ejercicio teórico: Varianza del ruido estimada por máxima verosimilitud	5
2.2. Ejercicio 3 - Comparación entre el periodograma puro y el estimador de Welch	6
2.3. Ejercicio 4 - Comparación entre el estimador MV y el estimador de Welch	7
2.3.a. Coeficientes del modelo	7
2.3.b. Métricas de Akaike para cada orden (m)	8
2.3.c. Estimador de Welch, PSD verdadera y estimadores MV	8
2.3.d. Histograma: probabilidad de que el estimador de Akaike sea de orden m	9
2.4. Ejercicio 5 - Aplicación a un proceso con modelo desconocido	10
3. Conclusiones	11

1. Técnicas no paramétricas

1.1. Ejercicio 1 - Estimador sesgado e insesgado de la autocorrelación

1.1.a. Autocovarianza teórica

Siendo W un ruido blanco gaussiano de media nula y varianza unitaria, queremos obtener la autocovarianza del mismo, como la media es nula vale que la autocovarianza es igual a la autocorrelación, teniendo en cuenta que el ruido blanco es un proceso ESA:

$$C_W(\tau) = \mathbb{E}[W(t)W(t+\tau)] - \overbrace{\mathbb{E}[W(t)]}^{=0 \forall t \in \mathbb{R}} \overbrace{\mathbb{E}[W(t+\tau)]}^{=0 \forall t, \tau \in \mathbb{R}} = R_W(\tau)$$

Luego, como para cada intervalo de tiempo $\tau \in \mathbb{R} - \{0\}$ se cumple que $W(t)$ y $W(t+\tau)$ son V.A independientes :

$$C_{W_{teorica}}(k) = R_{W_{teorica}}(k) = \delta_{k,0} \quad (1)$$

1.1.b. Estimador insesgado de la autocovarianza:

Por otro lado, el estimador insesgado de la autocorrelación es:

Estimador insesgado de la autocorrelación

$$\hat{R}_{W_{insesgado}}(k) = \frac{1}{N-k} \sum_{n=k+1}^N W(n)W^*(n-k), 0 \leq k \leq N-1 \quad (2)$$

Como en este caso el proceso tiene media nula y el proceso es real, el estimador insesgado de la autocovarianza es:

Estimador insesgado de la autocovarianza

$$\hat{C}_{W_{insesgado}}(k) = \frac{1}{N-k} \sum_{n=k+1}^N W(n)W(n-k), 0 \leq k \leq N-1 \quad (3)$$

1.1.c. Estimador sesgado de la autocovarianza:

Por otro lado, el estimador sesgado de la autocorrelación es:

Estimador sesgado de la autocorrelación

$$\hat{R}_{W_{sesgado}}(k) = \frac{1}{N} \sum_{n=k+1}^N W(n)W^*(n-k), 0 \leq k \leq N-1 \quad (4)$$

Como en este caso el proceso tiene media nula y el proceso es real, el estimador sesgado de la autocovarianza es:

Estimador sesgado de la autocovarianza

$$\hat{C}_{W_{sesgado}}(k) = \frac{1}{N} \sum_{n=k+1}^N W(n)W(n-k), 0 \leq k \leq N-1 \quad (5)$$

En todos los casos usando que la autocovarianza es simétrica con respecto a $k = 0$ se reflejó el resultado obtenido para los valores de $0 \leq k \leq N - 1$ a los valores de $-(N - 1) \leq k \leq -1$.

Se obtuvieron los gráficos de la **Fig. 1**

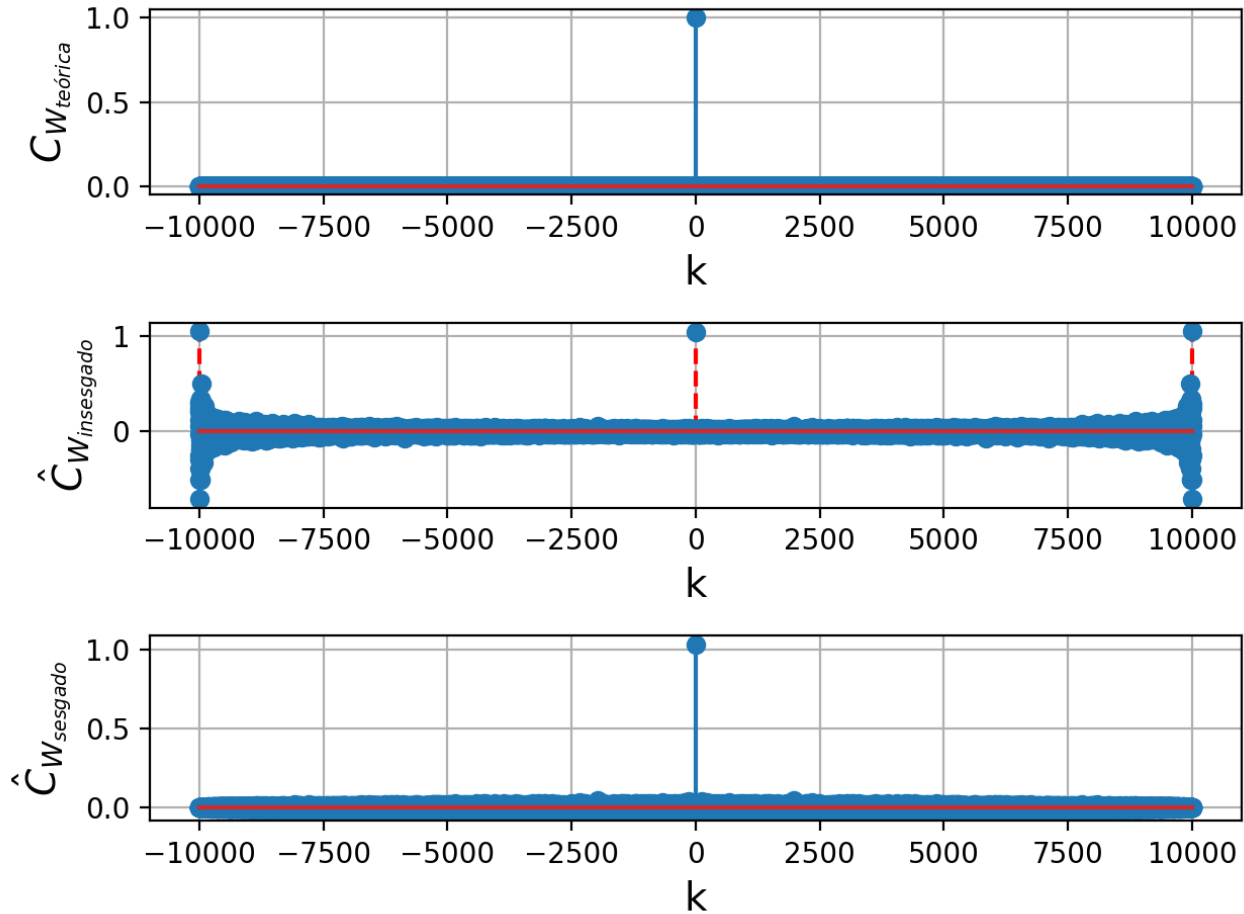


Figura 1: Autocovarianza teórica y estimadores.

Puede observarse como el estimador sesgado es un mejor estimador que el insesgado (al tener muchas muestras) dado que no se observa un comportamiento errático en las muestras de los extremos como sí así en el insesgado. Este comportamiento se debe a que cuando $N - k$ empieza a decrecer los valores del estimador insesgado se disparan a valores grandes.

De todas formas, cuando se quiera un estimador insesgado y solo se tenga interés en una región específica de la autocovarianza (sin contar con tantas muestras), se pueden descartar los extremos y contar con un estimador que no esté sesgado, ya que no es tan errático.

1.2. Ejercicio 2 - (In)consistencia del periodograma

Se considera un sistema LTI con transferencia:

$$H(z) = \frac{1 - 1,3817z^{-1} + 1,5632z^{-2} - 0,8843z^{-3} + 0,4096z^{-4}}{1 + 0,3544z^{-1} + 0,3508z^{-2} + 0,1736z^{-3} + 0,2401z^{-4}} \quad (6)$$

Asumiendo que el sistema es excitado por ruido blanco gaussiano de media nula y varianza unitaria denotado como $\{X_n\}_n$, tal que

$$Y(n) = (h * X)(n)$$

Siendo N el número de muestras disponibles del proceso X y consideramos el estimador de Blackman-Tukey:

$$\hat{S}_{BT}(e^{j\omega}) = \sum_k w(k) \hat{R}(k) e^{-j\omega k} \quad (7)$$

Tomando $w(k) = 1 \forall k$ y $M = N$ se tiene el periodograma clásico.

1.2.a. 1. Grafiqué la verdadera densidad espectral de potencia de Y , S_Y

La verdadera densidad espectral de potencia de Y que la denotaremos de ahora en adelante $S_Y(\omega)$ se puede obtener a partir del teorema de Wiener-Khintchine y propiedades de los sistemas LTI:

Dada la transferencia H que define un sistema LTI se tiene:

$$S_Y(e^{j\Omega}) = \mathcal{F}\{R_Y[n]\} = |H(e^{j\Omega})|^2 S_X(e^{j\Omega}) \quad (8)$$

En este caso como X es un proceso de ruido blanco de varianza unitaria $S_X(e^{j\Omega}) = 1$, luego:

$$S_Y(e^{j\Omega}) = |H(e^{j\Omega})|^2 \quad (9)$$

Buscamos los polos y ceros de la transferencia y obtenemos:

$$polos = \begin{cases} p_{1,2} = 0,31778647 \pm 0,62368046 j \\ p_{3,4} = -0,49498647 \pm 0,49499783 j \end{cases} \quad (10)$$

$$ceros = \begin{cases} c_{1,2} = 0,125156 \pm 0,79015796 j \\ c_{3,4} = 0,565694 \pm 0,56566482 j \end{cases} \quad (11)$$

Luego podemos reescribir la transferencia anterior como:

$$H(z) = \frac{(z - c_1)(z - \bar{c}_1)(z - c_3)(z - \bar{c}_3)}{(z - p_1)(z - \bar{p}_1)(z - p_3)(z - \bar{p}_3)} \quad (12)$$

Luego,

$$|H(z)|^2 = \frac{|(z - c_1)(z - \bar{c}_1)(z - c_3)(z - \bar{c}_3)|^2}{|(z - p_1)(z - \bar{p}_1)(z - p_3)(z - \bar{p}_3)|^2} \quad (13)$$

Finalmente:

$$S_Y(e^{j\Omega}) = \frac{|(e^{j\Omega} - c_1)(e^{j\Omega} - \bar{c}_1)(e^{j\Omega} - c_3)(e^{j\Omega} - \bar{c}_3)|^2}{|(e^{j\Omega} - p_1)(e^{j\Omega} - \bar{p}_1)(e^{j\Omega} - p_3)(e^{j\Omega} - \bar{p}_3)|^2} \quad (14)$$

Se obtuvo el gráfico de la misma en Python que se muestra en el **Fig. 2**:

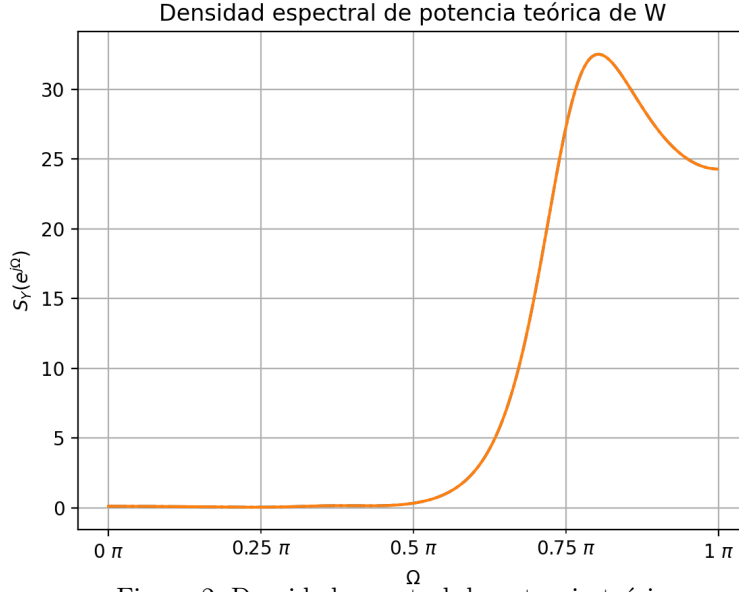


Figura 2: Densidad espectral de potencia teórica.

1.2.b. 2. Periodograma promedio y varianza del periodograma

Se generaron $J = 100$ realizaciones de largo $N = 64$ y $N = 512$ de X . Se obtuvo el periodograma correspondiente para cada realización de X con la función `scipy.signal.periodogram` y se obtuvieron el periodograma promedio y su varianza como:

Periodograma promedio:
$$\bar{\hat{S}}_Y(e^{j\Omega}) = \frac{1}{J} \sum_{i=1}^J \hat{S}_{Y,i}(e^{j\Omega}) \quad (15)$$

Varianza del periodograma:
$$\sigma_{\hat{S}_Y}^2(e^{j\Omega}) = \frac{1}{J-1} \sum_{i=1}^J \left(\hat{S}_{Y,i}(e^{j\Omega}) - \bar{\hat{S}}_Y(e^{j\Omega}) \right)^2 \quad (16)$$

Luego se realizó un gráfico de S_y superpuesto con las curvas de $\bar{\hat{S}}_Y$ y $\sigma_{\hat{S}_Y} \pm \bar{\hat{S}}_Y$ para cada N . Este gráfico está impreso en la **Fig. 3**

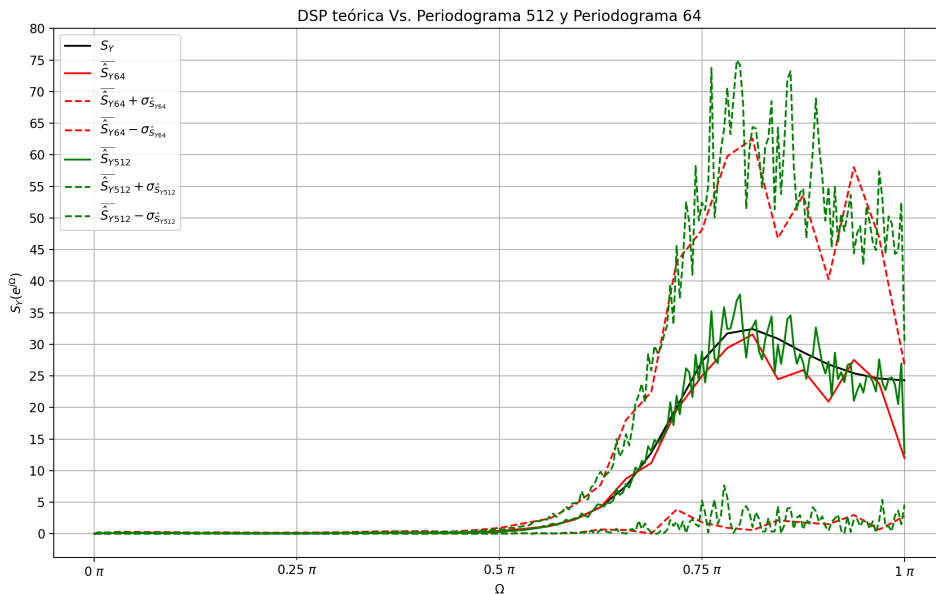


Figura 3: PSD teórica Vs. Periodograma 512 y Periodograma 64.

A medida que aumenta el N podemos ver que la varianza permanece constante, si bien como tiene más puntos parecería ser mas errática el valor medio de la varianza permanece constante.

1.2.c. 3. Periodograma con ventana de Bartlett

Ahora se realizó el mismo gráfico que en el punto anterior pero con $N=256$ comparando este con las mismas curvas, pero utilizando una ventana de Bartlett (triangular) con $M = \frac{N}{4}$ y $M = \frac{N}{16}$ en vez de la ventana rectangular.

En la siguiente figura se muestran los resultados obtenidos:

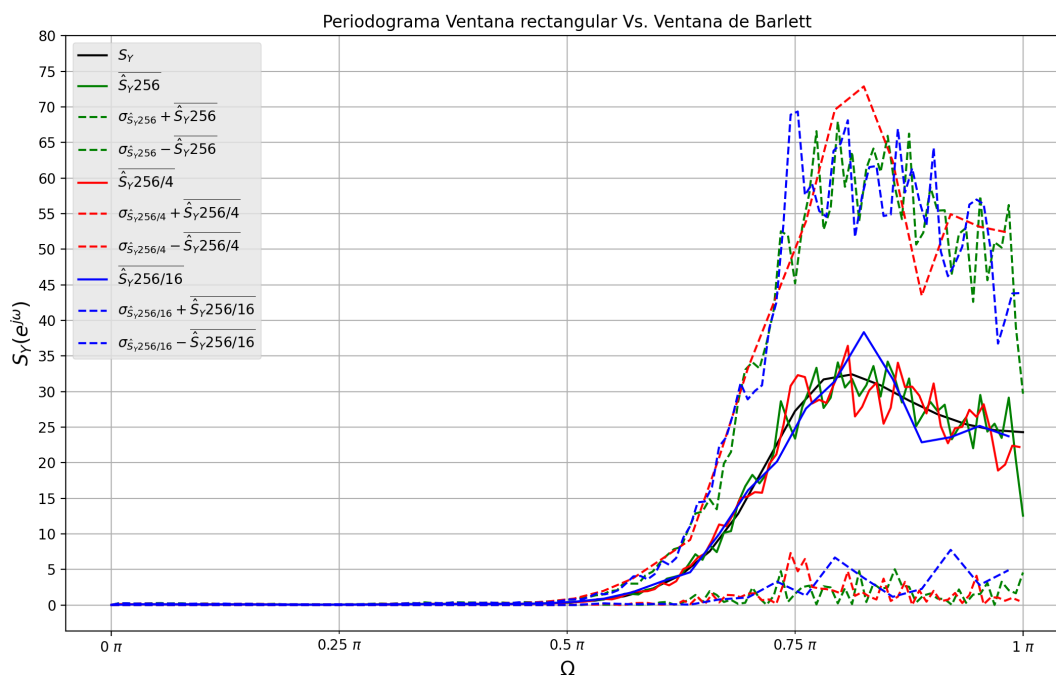


Figura 4: Periodogramas Ventana Rectangular Vs. Ventana Bartlett

En la **Fig. 4** se observa que el promedio es más preciso con la ventana de Bartlett ya que al ventanear con una ventana que atenúa el efecto de los extremos y por ende, la varianza del valor promedio es menor.

Si disminuimos el M , la ventana triangular se va a ensanchar, por lo tanto produce que el ancho del lóbulo principal de su transformada se haga angosta, ergo, el periodograma tiene una mayor resolución. Por otro lado, si mejoramos la resolución, la varianza aumentará como se muestra en el gráfico. Inversamente ocurre si aumentamos el M , la ventana se va a volver más angosta, lo que hace que los lóbulos secundarios bajen su altura entonces no va a haber potencia donde no debería luego, disminuye su varianza sin embargo, el lóbulo principal se hará más ancho y esto produce que haya una resolución mala.

2. Técnicas paramétricas

2.1. Ejercicio teórico: Varianza del ruido estimada por máxima verosimilitud

Para obtener un modelo autoregresivo de un determinado proceso se propone utilizar el criterio de información de Akaike (AIC), para el mismo es necesario estimar los parámetros del modelo (coeficientes y varianza del ruido), para distintos ordenes (m).

Para simplificar el problema asumimos que las primeras m muestras de Y son fijas y actúan como condiciones iniciales.

Partimos del logaritmo de la función de máxima verosimilitud, queremos hallar $\hat{\Theta}_{MV} = [\hat{\sigma}_{W,MV}^2, \hat{\mathbf{a}}_{MV}]$.

$$\log(L(\mathbf{a}, \sigma_W^2)) = -\frac{N-m}{2} \log(2\pi\sigma_W^2) - \frac{1}{2\sigma_W^2} \|\tilde{\mathbf{y}} - \mathbf{Y}^T \mathbf{a}\|^2 \quad (17)$$

Donde, $\tilde{\mathbf{y}} = [y(N), \dots, y(m+1)]^T$, $\mathbf{Y} = [\mathbf{y}(N-1), \dots, \mathbf{y}(m)] \in \mathbb{R}^{m \times (N-m)}$, $\mathbf{a} = [a_1, \dots, a_m]^T$

Es evidente en la ecuación 17 que el estimador MV de \mathbf{a} , $\hat{\mathbf{a}}_{MV}$, es la solución por cuadrados mínimos del sistema, ya que maximiza el logaritmo de la función de verosimilitud para un valor de a el valor mínimo de la norma en cuestión:

$$\tilde{\mathbf{y}} = \mathbf{Y}^T \mathbf{a}$$

Y debe satisfacer la ecuación normal:

$$\mathbf{Y}\mathbf{Y}^T \hat{\mathbf{a}}_{MV} = \mathbf{Y}\tilde{\mathbf{y}}$$

Si $\mathbf{Y}\mathbf{Y}^T$ es inversible,

$$\hat{\mathbf{a}}_{MV} = (\mathbf{Y}\mathbf{Y}^T)^{-1} \mathbf{Y}\tilde{\mathbf{y}}$$

Para hallar la varianza del ruido estimada por máxima verosimilitud se encuentra el valor de σ_W^2 que maximice el logaritmo de la función, esto lo encontramos derivando respecto de la varianza e igualando a cero.

$$\begin{aligned} \frac{\partial(\log(L(\hat{\mathbf{a}}_{MV}, \sigma_W^2)))}{\partial \sigma_W^2} &= 0 \\ -\frac{N-m}{2} \frac{1}{2\pi\sigma_W^2} 2\pi - \left[\frac{-2\|\tilde{\mathbf{y}} - \mathbf{Y}^T \hat{\mathbf{a}}_{MV}\|^2}{4\sigma_W^4} \right] &= 0 \\ \frac{\|\tilde{\mathbf{y}} - \mathbf{Y}^T \hat{\mathbf{a}}_{MV}\|^2}{2\sigma_{W,MV}^4} &= \frac{N-m}{2\sigma_{W,MV}^2} \end{aligned}$$

Finalmente obtenemos la forma del estimador de máxima verosimilitud para la varianza del ruido:

$$\sigma_{W,MV}^2 = \frac{\|\tilde{\mathbf{y}} - \mathbf{Y}^T \hat{\mathbf{a}}_{MV}\|^2}{N-m}$$

2.2. Ejercicio 3 - Comparación entre el periodograma puro y el estimador de Welch

Se generaron dos gráficos, uno para $N = 1000$ y otro con $N = 5000$. Se superponen el periodograma, el estimador de Welch y la PSD verdadera. En el caso de $N = 1000$ el estimador de Welch se utilizaron segmentos de 50 muestras con solapamiento del 50 % y para $N = 5000$ se utilizaron ventanas de 250 muestras con solapamiento de 50 %, en ambos casos se utilizó una FFT de 5000 puntos y la misma realización de Y .

En la **Fig. 5** se muestra el caso de $N = 1000$ y en la **Fig. 6** el caso de $N = 5000$

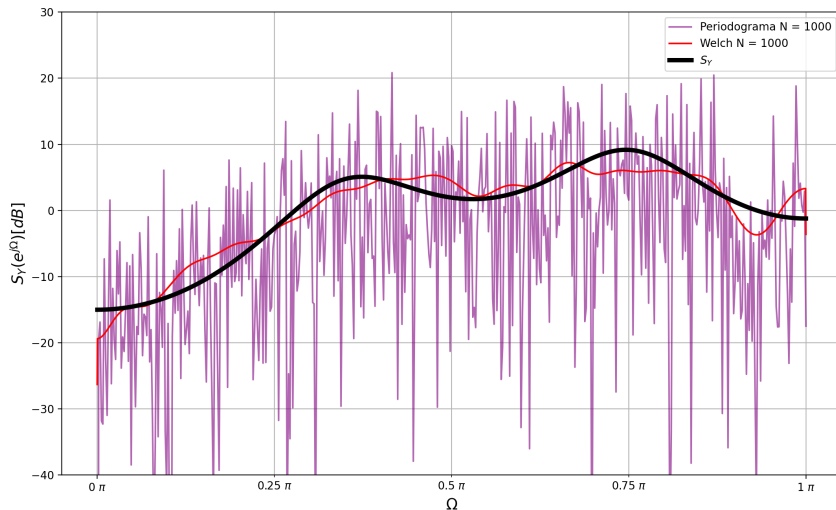


Figura 5: Periodograma, Welch y PSD teórica N= 1000.

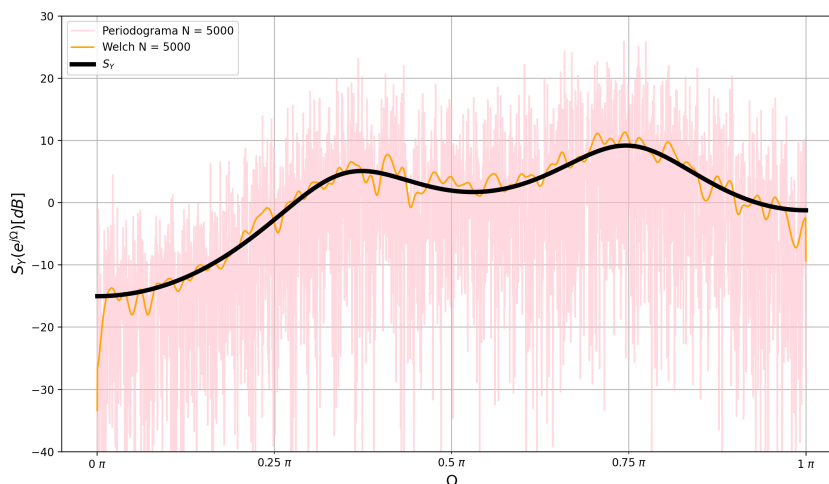


Figura 6: Periodograma, Welch y DSP teórica N= 5000.

En el caso del periodograma, no se observa ninguna mejora al cambiar la cantidad de muestras ya que este método no es consistente y su varianza no disminuye con las muestras, por lo que se puede ver únicamente un cambio en la cantidad de muestras graficadas.

Comparando el estimador de Welch con el periodograma, se puede ver que Welch presenta una varianza mucho menor, ya que la curva de densidad es mas cercana punto a punto a la PSD teorica que el periodograma.

Cuando se reduce la varianza del estimador de Welch, se necesita usar una ventana mas ancha (con un overlap cte), esto hace que la cantidad de puntos a graficar son menos que en el periodograma normal, perdiendo resolucion a cambio de disminuir la varianza.

2.3. Ejercicio 4 - Comparación entre el estimador MV y el estimador de Welch

Se calculó la métrica de akaike para cada orden $m \in 1, \dots, 10$ y se eligió el modelo candidato con la métrica mínima. (En el ejemplo se muestra que el orden coincidió, pero como se puede observar en el histograma de la **Fig. 8** no siempre coincide.

2.3.a. Coeficientes del modelo

Tabla 1: Coeficientes reales Vs. Estimados

	Coeficientes verdaderos	Coeficientes estimados orden 4	Coeficientes estimados (min akaike)
a_0	1	1	1
a_1	0,3544	0,3613	0,3613
a_2	0,3508	0,3195	0,3195
a_3	0,1736	0,1812	0,1812
a_4	0,2401	0,2449	0,2449

Al saber que originalmente las muestras provienen de un sistema AR-4, tiene sentido que el orden estimado por el criterio de Akaike sea $m = 4$ con mayor probabilidad (ver **Fig. 8**, luego de varias ejecuciones del algoritmo hecho en `python`, se observo por lo general $m = 4$ (como se verá a continuación).

Con respecto a los coeficientes, el criterio de **máxima verosimilitud**, se estimaron coeficientes cercanos a los del sistema original. Estos datos se muestran en la tabla 1 (Coeficientes reales Vs. Estimados).

El estimador de la varianza se obtuvo a partir de los coeficientes y es: $\hat{\sigma}_{W,MV}^2 = 0,9974$.

A partir de esto podemos definir el proceso definido por el criterio de Akaike como:

$$Y(n) + 0,3613 Y(n-1) + 0,3195 Y(n-2) + 0,1812 Y(n-3) + 0,2449 Y(n-4) = 0,9974 W(n) \quad (18)$$

2.3.b. Métricas de Akaike para cada orden (m)

Tabla 2: Diferencias entre las métricas de akaike para diferentes ordenes.

Orden	AIC(m)	$\Delta(m)$
1	5952.46	262
2	5818.51	128
3	5800.41	110
4	5690.44	0
5	5693.97	3.53
6	5696.46	6.02
7	5696.51	6.07
8	5698.40	7.96
9	5698.08	7.64
10	5697.63	7.19

Teniendo en cuenta:

El valor de la métrica $\Delta(m)$ indica el grado de soporte que le dan las muestras al modelo de orden m en relación al modelo elegido. Considere la siguiente tabla:

$\Delta(m)$	Soporte empírico para el modelo de orden m
0 – 2	Sustancial
4 – 7	Considerablemente menos que el elegido
> 10	Esencialmente ningún soporte empírico

Al ver que los órdenes próximos al de mínimo índice ($m = 5$, $m = 6$, y $m = 7$ de la tabla 2) tienen una diferencia métrica dentro de la categoría de *Considerablemente menos que el elegido*, indica que las muestras le dan un grado de soporte inferior en comparación al orden que se obtuvo ($m = 4$). El resto de ordenes no tienen soporte empírico alguno.

2.3.c. Estimador de Welch, PSD verdadera y estimadores MV

En la **Fig. 7** se muestran superpuestos el estimador de Welch la PSD verdadera, el estimador MV de orden 4 y el estimador MV del orden que indica la métrica de Akaike, en este caso estos últimos dos, coinciden.

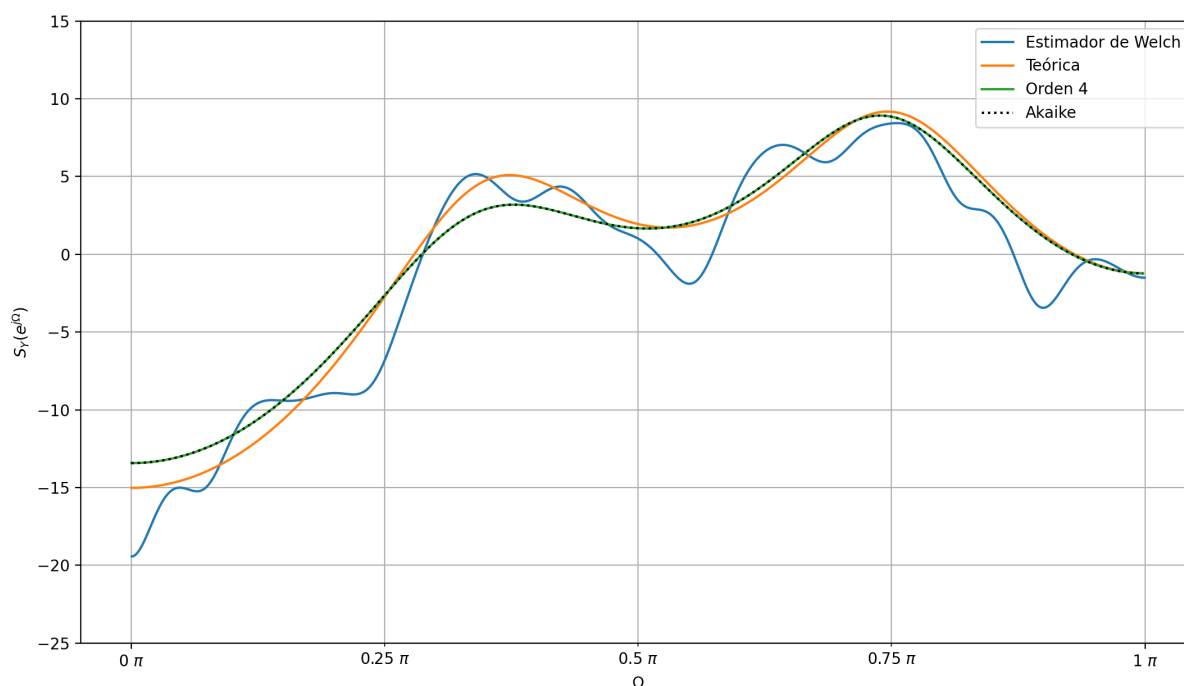


Figura 7: Estimador de Welch, PSD verdadera, Estimador de MV de orden 4 y akaike.

La **Fig. 7** muestra la PSD verdadera del AR-4, el estimador MV de orden 4 (y el de Akaike que es el mismo) y el estimador de Welch. Notar como el hecho de modelar la señal a priori como un proceso AR-m hace que la estimación sea mejor (siempre que sea correcto asumir que el sistema es AR-m); que en comparación con la estimación de Welch se ve mucho mas cercana a la teórica y con menor varianza.

2.3.d. Histograma: probabilidad de que el estimador de Akaike sea de orden m

Para obtener información sobre la probabilidad de cada valor de m sea el que determina el criterio de Akaike, se realizó 2000 veces en forma independiente el experimento de generar 1000 muestras de Y y se estimó en cada caso el orden del modelo según el criterio de Akaike. De esta manera se obtuvo el histograma de la **Fig. 8**.

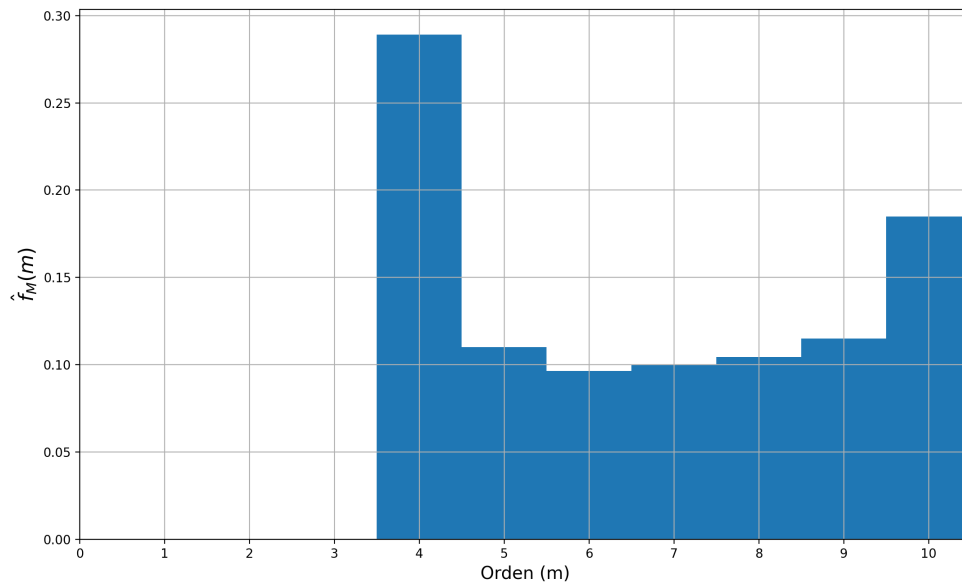


Figura 8: Histograma (probabilidad de que el estimador de Akaike elija cada orden posible).

En el histograma podemos ver que efectivamente el modelo de orden 4 es el que mayor probabilidad tiene de ser el elegido según el criterio de Akaike seguido por el modelo de orden 10. Esta información en conjunto con la **Tabla 2** nos confirma que el modelo de orden 4 es el mejor modelo a aplicar según este criterio.

2.4. Ejercicio 5 - Aplicación a un proceso con modelo desconocido

En este ejercicio se aplica el criterio de Akaike para, a partir de una señal determinada por un conjunto de $N = 5000$ muestras obtener un modelo autoregresivo utilizando el estimador de máxima verosimilitud y obteniendo el orden indicado por la métrica de Akaike. Luego se compara la densidad espectral de potencia obtenida con este modelo con el estimador de Welch. Se consideró que los órdenes posibles están en el rango $m \in \{1, 20\} \subseteq \mathbb{N}$.

El experimento dio por resultado un modelo de *orden 14*, *AR-14*:

$$\begin{aligned}
 Y(n) + 1,7568 Y(n-1) + 1,2056 Y(n-2) - 0,0472 Y(n-3) - 0,5886 Y(n-4) \\
 - 0,3766 Y(n-5) - 0,1305 Y(n-6) - 0,0949 Y(n-7) - 0,0106 Y(n-8) \\
 + 0,1779 Y(n-9) + 0,2578 Y(n-10) + 0,1195 Y(n-11) \\
 - 0,0511 Y(n-12) - 0,0910 Y(n-13) - 0,0323 Y(n-14) = 0,9790W(n)
 \end{aligned}$$

Se obtuvo el estimador de Welch y la PSD del modelo obtenido, se observan en el gráfico de la **Fig. 9**:

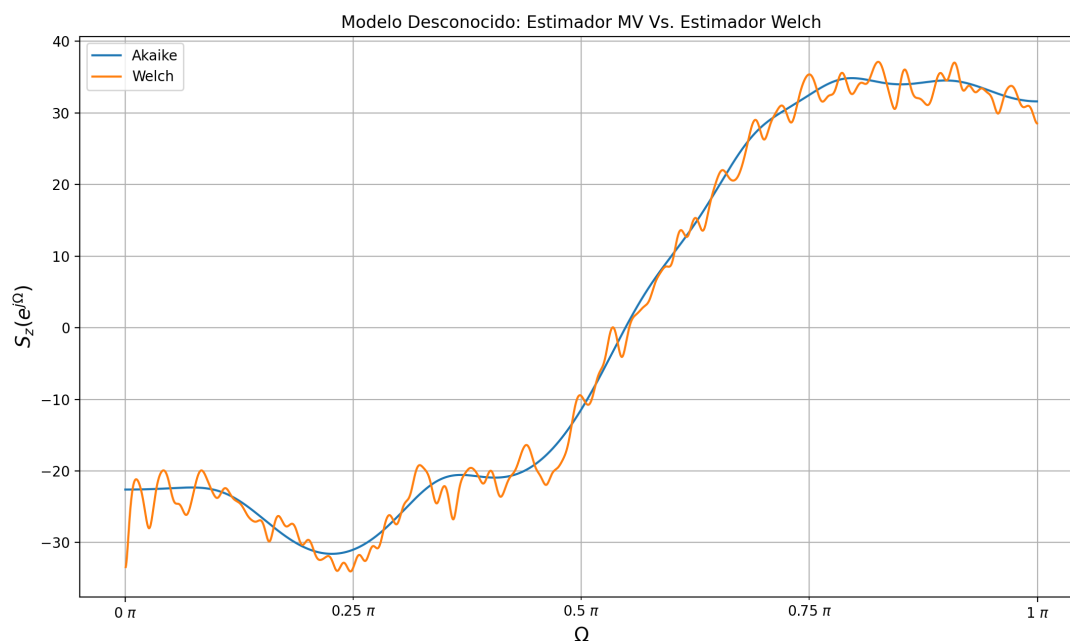


Figura 9: Estimación por MV vs Estimación de Welch

Se puede observar en la figura como el valor medio de ambos estimadores coincide, siendo la varianza del modelo obtenido mucho menor al del estimador de Welch, el criterio de Akaike y el estimador de MV sería implementado en caso de querer obtener un modelo de un sistema a partir del cual se puede generar la señal en cuestión, en caso de querer hacer un análisis de la señal el estimador de Welch cumple esta determinada función.

3. Conclusiones

En conclusión podemos decir que el estimador de Welch tiene menor varianza que el periodograma a coste de una pérdida en la resolución en frecuencia del mismo, este es uno de los más utilizados. Por otra parte, el estimador sesgado de la autocorrelación es menos errático en las muestras de los extremos y por eso se utiliza este y no el insesgado. También vimos que el criterio de akaike nos muestra que al momento de realizar un modelo podemos obtener el orden del mismo a partir del estimador de máxima verosimilitud y obtener un modelo de la señal como un proceso AR-4, estimando los parámetros de la misma. A diferencia de los otros estimadores este define un proceso del cual se puede generar esta señal.