Examen Final **Procesado de Señal** **19 Enero 2006**

Profesores: Miguel A. Lagunas, Montserrat Nájar, Ana Pérez, Jaume Riba

Problema 1 (30%)

Se pretende estimar la ganancia del canal h en un sistema de comunicaciones a partir de una secuencia piloto determinista, conocida en el receptor $\underline{u}$ de K símbolos complejos de módulo unidad. Para ello, minimizaremos la potencia del error entre el canal estimado linealmente y el canal real (estimador de error cuadrático medio mínimo). La señal observada en estas condiciones es la siguiente:

$$\underline{y} = h\sqrt{P}\underline{u} + \underline{w} \in \mathbb{C}^{K\times 1}$$

donde P es la potencia transmitida durante los símbolos piloto, y $\underline{w}$ es ruido blanco Gaussiano de media nula y potencia s_w^2. La ganancia del canal h es constante durante la observación de los K símbolos, y puede considerarse una variable aleatoria de potencia s_h^2 conocida, incorrelada con el ruido. Se pide:

1. Determine el estimador $\hat{h}$ de error cuadrático medio mínimo del canal, a partir de $\min_{\underline{s}^H} E\left\{\left|h - \underline{s}^H\underline{y}\right|^2\right\}$ con $\hat{h} = \underline{s}^H\underline{y}$. Utilice para ello el lema de inversión[1].

$$\min_{\underline{s}^H} E\left\{\left|h - \hat{h}\right|^2\right\} \quad\Rightarrow\quad \underline{\underline{R}}_y\underline{s} = s_h^2\sqrt{P}\underline{u} \quad\Rightarrow\quad \underline{s} = \frac{\sqrt{P}s_h^2}{PKs_h^2 + s_w^2}\underline{u}$$

2. ¿Cuál es el sesgo del estimador? ¿Cómo depende del número de símbolos en la secuencia de entrenamiento?

$$E\{\hat{h}\} = E\left\{\frac{\sqrt{P}s_h^2}{PKs_h^2 + s_w^2}\underline{u}^H\underline{y}\right\} = \frac{PKs_h^2}{PKs_h^2 + s_w^2}h$$

El sesgo se anula para un número suficientemente grande de simbolos-piloto K.

3. Determine el error cuadrático medio (MSE) de la estimación.

$$MSE = E\left\{\left|h - \frac{\sqrt{P}s_h^2}{PKs_h^2 + s_w^2}\underline{u}^H\underline{y}\right|^2\right\} = E\left\{\left(h^* - \frac{\sqrt{P}s_h^2}{PKs_h^2 + s_w^2}\underline{y}^H\underline{u}\right)h\right\} = \frac{s_h^2 s_w^2}{PKs_h^2 + s_w^2}$$

---

[1] Si $\underline{\underline{R}} = \underline{\underline{B}} + \underline{g}\underline{g}^H$, entonces su inversa está definida como $\underline{\underline{R}}^{-1} = \underline{\underline{B}}^{-1} - \dfrac{\underline{\underline{B}}^{-1}\underline{g}\underline{g}^H\underline{\underline{B}}^{-1}}{1 + \underline{g}^H\underline{\underline{B}}^{-1}\underline{g}}$

4. A partir de la SNR de $\hat{h}$, definida como $SNR_{\hat{h}} = \dfrac{E\left\{\left|h\sqrt{P}\underline{s}^{H}\underline{u}\right|^{2}\right\}}{E\left\{\left|\underline{s}^{H}\underline{w}\right|^{2}\right\}}$, determine que ésta aumenta linealmente con K. ¿Qué condición debe cumplir $SNR_{\hat{h}}$ para que el MSE sea muy inferior a s_{h}^{2}?

$$SNR_{est} = \frac{E\left\{P\left|h\underline{s}^{H}\underline{u}\right|^{2}\right\}}{E\left\{\left|\underline{s}^{H}\underline{w}\right|^{2}\right\}} = \frac{PKs_{h}^{2}}{s_{w}^{2}} = K \cdot SNR \qquad \Rightarrow \qquad MSE = \frac{s_{h}^{2}}{SNR_{est}+1}$$

Debe ocurrir que SNR_{est} sea grande, o bien porque la SNR en recepción es grande o bien porque el número de símbolos-piloto K es el suficiente.

Se considera la estimación iterativa de la ganancia del canal *h*, siendo $\underline{s}_{n}$ el vector de coeficientes en la iteración *n*.

5. Encuentre el error cuadrático medio (MSE) de la estimación en función del error en los coeficientes $\underline{v}_{n} = \underline{s}_{n} - \underline{s}_{opt}$

$$MSE_{n} = \frac{s_{h}^{2}s_{w}^{2}}{PKs_{h}^{2}+s_{w}^{2}} + \underline{v}_{n}^{H}\underline{\underline{R}}_{y}\underline{v}_{n}$$

6. Defina la ecuación de adaptación del algoritmo de gradiente.

$$\underline{s}_{n+1} = \underline{s}_{n} - m\left(\underline{\underline{R}}_{y}\underline{s}_{n} - s_{h}^{2}\sqrt{P}\underline{u}\right)$$

Para el análisis de convergencia del algoritmo de gradiente, la ecuación de adaptación puede escribirse como $\hat{\underline{v}}_{n}(i) = r_{i}^{n}\hat{\underline{v}}_{0}(i)$ donde $\hat{v}_{n}(i)$ es la componente i del vector $\hat{\underline{v}}_{n} = \underline{\underline{Q}}^{H}\underline{v}_{n}$ (las columnas de la matriz $\underline{\underline{Q}}$ son los autovectores de la matriz $\underline{\underline{R}}_{y}$)

7. Demuestre que $\underline{u}$ es un autovector de $\underline{\underline{R}}_{y}$
8. Obtenga los valores de r_{i} y defina el paso de adaptación para asegurar la convergencia del algoritmo en función de s_{w}^{2}, s_{h}^{2}, P y K

$$r_{i} = \left(1 - ml_{i}\right) \quad m < \frac{2}{l_{\max}} = \frac{2}{\left(PKs_{h}^{2}+s_{w}^{2}\right)}$$

9. Establezca la relación entre la velocidad de convergencia del algoritmo de gradiente, la SNR en recepción $SNR = \dfrac{Ps_{h}^{2}}{s_{w}^{2}}$ y el número de pilotos *K*.

$$velocidad \propto \frac{l_{\min}}{l_{\max}} = \frac{s_{w}^{2}}{PKs_{h}^{2}+s_{w}^{2}} = \frac{1}{KSNR+1}$$

El estimador va a utilizarse en un esquema de detección multiusuario en el que se reciben dos usuarios simultáneamente, con símbolos de información aleatorios y desconocidos $\underline{x}_{1}$ y $\underline{x}_{2}$:

$$\underline{y} = h_{1}\sqrt{P_{1}}\underline{x}_{1} + h_{2}\sqrt{P_{2}}\underline{x}_{2} + \underline{w} \qquad \text{donde} \qquad |h_{1}|P_{1} >> |h_{2}|P_{2}\ , \quad E\left\{\underline{x}_{i}\underline{x}_{j}^{H}\right\} = d_{i-j}\underline{\underline{I}}$$

y se decodifica con cancelación sucesiva de la siguiente forma: se detectan los símbolos del usuario 1 (el más potente) y se resta la contribución de éste de la señal recibida:

$$\underline{y}_2 = \underline{y} - \hat{h}_1\sqrt{P_1}\,\underline{x}_1$$

antes de detectar el usuario 2. Debido al error en la estimación de h_1, la SNR del usuario 2 en $\underline{y}_2$ se verá degradada.

10. Determine la SNR del usuario 2 en $\underline{y}_2$ (SNR_2'), y justifique (utilizando el resultado del apartado 4) que si K es grande ésta no se degrada ni siquiera para valores muy grandes de $SNR_1 = \frac{P_1\sigma_{h_1}^2}{\sigma_w^2}$ (justificando de esta forma el uso de un estimador de canal de error cuadrático medio mínimo).

La SNR para el usuario 2 en $\underline{y}_2$ vendrá dada por: $$\underline{y}_2 = \underline{y} - \hat{h}_1\sqrt{P_1}\,\underline{x}_1 = h_2\sqrt{P_2}\,\underline{x}_2 + \underline{w} + \left(h_1 - \hat{h}_1\right)\sqrt{P_1}\,\underline{x}_1$$ $$SNR_2' = \frac{P_2\sigma_{h_2}^2}{\sigma_w^2 + P_1\cdot MSE_1} = \frac{P_2\sigma_{h_2}^2\left(P_1\sigma_{h_1}^2K + \sigma_w^2\right)}{\sigma_w^2\left(P_1\sigma_{h_1}^2K + \sigma_w^2\right) + P_1\sigma_{h_1}^2\sigma_w^2} = SNR_2\frac{SNR_1K + 1}{\left(K+1\right)SNR_1 + 1}$$ $$\lim_{SNR_1\to\infty} SNR_2' = SNR_2\frac{K}{K+1}$$

Problema 2 (30%)

Dada una imagen $\underline{\underline{X}}_n$ de tamaño (Q, Q), se procede a representar esta mediante Q imágenes de rango uno formadas según sigue:

$$\underline{\underline{X}}_n = \sum_{r=1}^{Q} j_n(r).\left[\underline{u}_{r,n}.\underline{v}_{r,n}^H\right]$$

Donde los vectores $\underline{u}$ son ortonormales entre si, al igual que los $\underline{v}$ entre si.

a.- Comente las ventajas y dificultades de emplear esta representación para la transmisión de la imagen original.

Una alternativa diferente, sería usar la siguiente representación, también formada con vectores ortonormales:

$$\underline{\underline{X}}_n = \sum_{s=1}^{Q} \sum_{r=1}^{Q} j_n(s,r).\left[\underline{u}_s.\underline{u}_r^H\right]$$

b.- Indique las ventajas y/o los inconvenientes de emplear esta segunda transformación en vez de la primera.

Considerando una columna de la imagen original $\underline{X}_n$, de media nula y covarianza $\underline{\underline{R}} = E\left\{\underline{X}_n \underline{X}_n^H\right\}$ (nótese que la potencia es igual a $Traza\left(\underline{\underline{R}}\right)$), en un proceso de cuantificación la política óptima es asignar menos bits a los vectores de mayor probabilidad, es decir:

$$Bits\ para\ \underline{X}_n \propto -Ln\left(\Pr\left(\underline{X}_n\right)\right)$$

Asumiendo esta regla, el número de bits promedio necesarios para la transmisión de los vectores (asumiendo que éstos son Gausianos) vendría dado por:

$$B \propto -\int \Pr\left(\underline{X}_n\right).Ln\left(\Pr\left(\underline{X}_n\right)\right).d\underline{X}_n \propto \det\left(\underline{\underline{R}}_x\right) + Q$$

c.- Considerando fijada la potencia de cada componente y sabiendo que el determinante de una matriz de correlación es siempre menor o igual que el producto de los elementos de su diagonal, muestre que B es máximo cuando las componentes del vector están incorreladas.

Fijada la matriz de correlación y una vez obtenido B, la codificación conjunta de los vectores es complicada y sería deseable poderla realizar independientemente para cada componente. Con el fin de poder codificar cada componente independientemente, se aplicará una transformación a los vectores originales $\underline{Y}_n = \underline{\underline{U}}.\underline{X}_n$

d.- Demuestre qué condición ha de cumplir la transformación para no modificar la potencia de los vectores transformados, es decir, se ha de cumplir que $Traza\left(\underline{\underline{R}}_y\right) = Traza\left(\underline{\underline{R}}_x\right)$

e.- Indique ahora como se encuentra la matriz de transformación que permite codificar cada componente por separado. Compruebe que el determinante de la matriz de correlación de los datos transformados es el mismo que los originales, es decir, no se incrementa el número de bits necesarios para su codificación por componentes, con respecto a la codificación conjunta de los vectores originales.

f.- Cuáles son las similitudes y diferencias de la DFT con la transformación anterior y porqué se prefiere el uso de la DCT al de la DFT en procesado de imagen.

SOLUCION:

a)

La expresión dada $\underline{\underline{X}}_n = \sum_{r=1}^{Q} j_n(r).\underline{u}_{r,n}.\underline{v}_{r,n}^H$, también puede escribirse como $\underline{\underline{X}}_n = \underline{\underline{U}}_n.diag\,\underline{\underline{j}}_n.\underline{\underline{V}}_n^H$ indica que las matrices $\underline{\underline{U}}_n$ y $\underline{\underline{V}}_n$ diagonalizan la matriz original directamente. Por lo tanto por cada sub-imagen han de transmitirse ambas con lo que la información lateral ya superaría la de la imagen original por lo que es inviable.

b)

En este caso $\underline{\underline{X}}_n = \sum_{r=1}^{Q}\sum_{s=1}^{Q} f_n(r,s).\underline{u}_r.\underline{u}_s^H$ o bien $\underline{\underline{X}}_n = \underline{\underline{U}}.\underline{\underline{f}}_n.\underline{\underline{U}}^H$ que no diagonaliza y se trata de una mera transformación. Al mismo tiempo, esta transformación no tiene por que depender de cada sub-imagen y puede calcularse con un criterio sobre promedios de estas.

Con todo, si no se imponen criterios adicionales sobre esta transformación para que la transmisión de $\underline{\underline{f}}_n$ presente mas ventajas que la transmisión de la imagen original, la transformación carecería de utilidad alguna.

c)

Como el determinante de $\underline{\underline{R}}_x$ es siempre menor que el producto de los elementos de su diagonal, y estos valores están fijados por ser la potencia de cada componente, entonces B se hace máximo cuando los valores fuera de la diagonal son cero, es decir, cuando las componentes están incorreladas.

d)

Como la potencia es $P_y = Traza\left(\underline{\underline{R}}_y = E\left[\underline{Y}_n.\underline{Y}_n^H\right]\right) = Traza\left(\underline{\underline{U}}.\underline{\underline{R}}_x.\underline{\underline{U}}^H\right) =$ La traza tiene la propiedad de circular las matrices en su interior sin alterar su valor $= Traza\left(\underline{\underline{R}}_x.\underline{\underline{U}}^H.\underline{\underline{U}}\right)$.

Para que esta última expresión sea igual a $P_x = Traza\left(\underline{\underline{R}}\right)$ es claro que el requisito es que la matriz de transformación sea orthonormal.

Es decir, $\underline{\underline{U}}^H.\underline{\underline{U}} = \underline{\underline{I}}$

e)

La condición de que sean ortonormales unida a que deben hacer independientes las componentes del vector transformado obliga a que la matriz de transformación este formada por los autovectores de la matriz de correlación.

$\underline{\underline{R}}.\underline{\underline{U}} = \underline{\underline{U}}.diag\left[l\right]$ y $\underline{\underline{R}}_y = diag\left[l\right]$

Claramente B se mantiene (no se incrementa) pues el determinante de la nueva matriz de correlación (la de los datos transformados) es el producto de autovalores que es idéntico al determinante de la matriz original.

f)

La DFT es también una transformación ortonormal. La diferencia es que la matriz de correlación de los datos transformados no es diagonal. Dicho de otro modo, interpretando la DFT como la salida de un banco de filtros, tan solo cuando estos son ideales se consigue una correlación cruzada cero entre sus salidas. Tan solo cuando el tamaño de la muestra es grande (la DFT es de muchos puntos se daría esta situación). No obstante, su decorrelacion es muy buena y puede llevarse a cabo la codificación de sus componentes una a una con una escasa pérdida con respecto a la codificación conjunta.

Un proceso no-lineal tiende a extender la energía de unas frecuencias a otras. Dado que la DCT conlleva una extrapolación que hace par la función original esta resulta mas ventajosa, menor distorsión, que al extrapolar con ceros (como ocurre con la DFT para manejar en ambas la misma resolución frecuencial). Además, la DCT conlleva tan solo operaciones con números reales lo que simplifica la complejidad de la transformación.

Problema 3 (40%)

Sea $x(n)$ un proceso aleatorio real autoregresivo de primer orden (AR(1)). Suponga que dispone de las estimaciones $r(0)$ y $r(1)$ $(=r(-1))$, siendo $r(m)$ la función de autocorrelación del proceso $x(n)$. Se pretende estimar la densidad espectral de potencia del proceso a la frecuencia f=0 ($S_x(0)$) como función de $r(0)$ y $r(1)$ usando diferentes estimadores espectrales.

a.- Halle $S_x(0)_{co}$ en función de $r(0)$ y $r(1)$ usando el estimador correlograma (transformada de Fourier de la autocorrelación conocida con ventana rectangular). ¿Para qué valores de $r(1)/r(0)$ la estimación obtenida no tiene sentido?

$$S_x(0)_{co} = \sum_{m=-1}^{1} r(m) = r(0) + 2r(1)$$

Para $r(1) < -\frac{r(0)}{2}$, la estimación no tiene sentido ya que se obtiene una valor negativo de la densidad espectral de potencia, $S_x(0)_{co} < 0$.

b.- Halle $S_x(0)_{pe}$ en función de $r(0)$ y $r(1)$ usando el estimador Blackman-Tukey con ventana triangular (transformada de Fourier de la autocorrelación conocida con ventana triangular). Justifique que este estimador no presenta el problema del correlograma. Explique cómo obtendría esta estimación de un modo natural a partir de K bloques de tamaño 2 de los datos $x(n)$.

$$S_x(0)_{pe} = \sum_{m=-1}^{1} \frac{2-|m|}{2} r(m) = r(0) + r(1)$$

En este caso, como $|r(1)| \le |r(0)|$ se obtiene siempre una estimación positiva de la densidad espectral de potencia, $S_x(0)_{pe} \ge 0$.

Como este estimador es el periodograma promedio, podemos obtenerla directamente a partir de K bloques de longitud 2 del siguiente modo:

$$\hat{S}_x(0)_{pe} = \frac{1}{K}\sum_{k=0}^{K-1} |x(2k) + x(2k+1)|^2$$

c.- Halle $S_x(0)_{ca}$ en función de $r(0)$ y $r(1)$ usando el estimador de Capon. Demuestre que $S_x(0)_{ca} = S_x(0)_{pe}$. Compruebe que para $r(1)/r(0) = 0.5$, $S_x(0)_{ca} = 1.5r(0)$.

$$S_x(0)_{ca} = \frac{\underline{1}^T \underline{\underline{R}}^{-1} \underline{1}}{\underline{1}^T \underline{\underline{R}}^{-2} \underline{1}} \qquad \begin{cases} \underline{\underline{R}} = \begin{bmatrix} r_0 & r_1 \\ r_1 & r_0 \end{bmatrix} \\ \underline{\underline{R}}^{-1} = \dfrac{1}{r_0^2 - r_1^2} \begin{bmatrix} r_0 & -r_1 \\ -r_1 & r_0 \end{bmatrix} \\ \underline{\underline{R}}^{-2} = \dfrac{1}{\left(r_0^2 - r_1^2\right)^2} \begin{bmatrix} r_0 & -r_1 \\ -r_1 & r_0 \end{bmatrix} \begin{bmatrix} r_0 & -r_1 \\ -r_1 & r_0 \end{bmatrix} = \dfrac{1}{\left(r_0^2 - r_1^2\right)^2} \begin{bmatrix} r_0^2 + r_1^2 & -2r_0r_1 \\ -2r_0r_1 & r_0^2 + r_1^2 \end{bmatrix} \end{cases}$$

$$S_x(0)_{ca} = \left(r_0^2 - r_1^2\right) \frac{[1,1]\begin{bmatrix} r_0 & -r_1 \\ -r_1 & r_0 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}}{[1,1]\begin{bmatrix} r_0^2 + r_1^2 & -2r_0r_1 \\ -2r_0r_1 & r_0^2 + r_1^2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}} = \frac{\left(r_0^2 - r_1^2\right)\left(r_0 - r_1\right)}{\left(r_0^2 + r_1^2 - 2r_0r_1\right)} = \frac{\left(\left(r_0 - r_1\right)\left(r_0 + r_1\right)\right)\left(r_0 - r_1\right)}{\left(r_0 - r_1\right)^2} = r_0 + r_1$$

Se obtiene el mismo resultado que en el apartado b, con lo que

$$S_x(0)_{ca} = S_x(0)_{pe} = r_0 + r_1 = r_0\left(1 + \frac{r_1}{r_0}\right).$$

Se comprueba facilmente que para $r(1)/r(0) = 0.5$, $S_x(0)_{ca} = 1.5r(0)$

d.- Dé la expresión general de $r(m)$ en función de $r(0)$ y $r(1)$ sabiendo que se trata de un proceso AR(1).

Un proceso AR(1) tiene una función de autocorrelación $r(m) = r(0)a^{|m|}$.

Para $m = 1$ tenemos que $a = \frac{r_1}{r_0}$, con lo que: $r(m) = r(0)\left(\frac{r_1}{r_0}\right)^{|m|}$

e.- Halle $S_x(0)_{ar}$ (estimación espectral parámetrica AR(1)) a partir de la extrapolación de $r(m)$ obtenida en el apartado anterior. Compruebe que para $r(1)/r(0) = 0.5$, $S_x(0)_{ca} = 3r(0)$.

Ayuda: $\sum_{m=0}^{\infty} a^m = \frac{1}{1-a}$, para $|a| < 1$. Nota: alternativamente puede resolver también este apartado a partir de la expresión general del estimador espectral AR en función de la matriz de autocorrelación.

$$S_x(0)_{ar} = \sum_{m=-\infty}^{\infty} r(m) = r(0) + 2\sum_{m=1}^{\infty} r(m) = r(0)\left(1 + 2\frac{r(1)}{r(0)}\sum_{m=0}^{\infty}\left(\frac{r(1)}{r(0)}\right)^m\right) =$$

$$S_x(0)_{ar} = r(0)\left(1 + 2\frac{r(1)}{r(0)}\frac{1}{1 - \frac{r(1)}{r(0)}}\right) = r(0)\left(1 + 2\,r(1)\frac{1}{r(0) - r(1)}\right) = r\,(0)\frac{r(0) + r(1)}{r(0) - r(1)} = r\,(0)\frac{1 + \frac{r(1)}{r(0)}}{1 - \frac{r(1)}{r(0)}}$$

Se comprueba facilmante que para $r(1)/r(0) = 0.5$, $S_x(0)_{ca} = 3r(0)$

Solución alternativa:

$$S_x(0)_{ar} = \frac{\underline{d}^T \underline{\underline{R}}^{-1} \underline{d}}{\left|\underline{1}^T \underline{\underline{R}}^{-1} \underline{d}\right|^2}$$

$$\underline{\underline{R}} = \begin{bmatrix} r_0 & r_1 \\ r_1 & r_0 \end{bmatrix}$$

$$\underline{\underline{R}}^{-1} = \frac{1}{r_0^2 - r_1^2}\begin{bmatrix} r_0 & -r_1 \\ -r_1 & r_0 \end{bmatrix}$$

$$S_x(0)_{ar} = \left(r_0^2 - r_1^2\right) \frac{[1,0]\begin{bmatrix} r_0 & -r_1 \\ -r_1 & r_0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}}{\left\| [1,1]\begin{bmatrix} r_0 & -r_1 \\ -r_1 & r_0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\|^2} =$$

$$S_x(0)_{ar} = \frac{\left(r_0^2 - r_1^2\right) r_0}{\left(r_0 - r_1\right)^2} = \frac{\left(r_0 - r_1\right)\left(r_0 + r_1\right) r_0}{\left(r_0 - r_1\right)\left(r_0 - r_1\right)} = r(0)\frac{r(0)+r(1)}{r(0)-r(1)} = r(0)\frac{1+\frac{r(1)}{r(0)}}{1-\frac{r(1)}{r(0)}}$$

f.- Para el caso de un proceso AR(1) paso-bajo, demuetre que se cumple siempre que $S_x(0)_{ar} > S_x(0)_{ca}$. Analice e interprete en particular la estimación AR el caso en que $r(1)/r(0) = 1$. En base a esto, compare la diferente capacidad resolutiva de AR y Capon.

Si el proceso es paso-bajo significa que $r(1) > 0$. Entonces, $\frac{r(1)}{r(0)} > 0$ y $\frac{1}{1-\frac{r(1)}{r(0)}} > 1$. Entonces:

$$\frac{S_x(0)_{ar}}{S_x(0)_{ca}} = \frac{1}{1-\frac{r(1)}{r(0)}} > 1$$

de lo que se deduce que:

$S_x(0)_{ar} > S_x(0)_{ca}$.

En el caso de que $r(1)/r(0) = 1$, tenemos un proceso que es un nivel de continua puro, con lo que su densidad espectral de potencia es una delta de dirac en f=0. Con el estimador de capon se obtiene

$S_x(0)_{ca} = 2r_0$

mientras que no el estimador AR se obtiene:

$S_x(0)_{ar} = \infty$

siendo este último el único que el proporciona el valor corrector de la densidad espectral. Esto no es más que la constatación del carácter general más resolutivo de los estimadores paramétricos en comparación con los no paramétricos.