PROCESADO DE LA SEÑAL (Semestre P2010) **Grupo 20**
Gregori Vázquez

Ejercicio 1: Un resultado fundamental en Procesado de la Señal es el siguiente. Consideramos dos variables aleatorias reales conjuntamente Gaussianas S y X, con medias, varianzas y covarianzas, todas *<u>conocidas</u>:*

$$E[S] = \mu_S,\ E[X] = \mu_X$$
$$\mathrm{var}(S) = \sigma_S^2,\ \mathrm{var}(X) = \sigma_X^2$$
$$\mathrm{cov}(S, X) = \rho\sigma_S\sigma_X$$

Consideramos que S es el término de señal y X la observación. A partir de la distribución Gaussiana conjunta $f_{S,X}(s,x)$ puede deducirse fácilmente cualquiera de las dos distribuciones marginales $f_S(s)$, $f_X(x)$ o condicionadas $f_{X/S}(x/s)$, $f_{S/X}(s/x)$ siendo esta última de la forma:

$$f_{S/X}(s/x) = \frac{f_{X/S}(x/s) f_S(s)}{f_X(x)} = \frac{1}{\sqrt{2\pi\sigma_{S/X}^2}} \exp\left(-\frac{(s - \mu_{S/X}(x))^2}{2\sigma_{S/X}^2} \right)$$

con:

$$\mu_{S/X}(x) = \mu_S + \rho\frac{\sigma_S}{\sigma_X}(x - \mu_X)$$
$$\sigma_{S/X}^2 = (1 - \rho^2)\sigma_S^2$$

En el caso más general, el conocimiento inicial que tenemos del término de señal S es μ_S y nos planteamos como se modifica dicho conocimiento cuando disponemos de una nueva observación $X = x$:

a. Obtenga la estimación MAP de S cuando se dispone de una nueva observación $X = x$.

Res.: Vemos que:

$$\hat{S}_{MAP} = \arg\, máx_S \quad \ln f_{S/X}(s/x) \Rightarrow \frac{d}{ds}\ln f_{S/X}(s/x) = 0 \Rightarrow$$
$$\Rightarrow \hat{S}_{MAP} = \mu_{S/X}(x) = \mu_S + \rho\frac{\sigma_S}{\sigma_X}(x - \mu_X)$$

Podemos comentar que la estimación MAP en este caso no exige del conocimiento del modelo de señal concreto del problema, ni de ningún prior del parámetro, dado que la estadística a maximizar es conocida.

b. Indique en que casos la nueva observación $X = x$ es más o menos informativa en función de todos los estadísticos conocidos de primer y segundo orden del problema.

Res.: Observamos en la solución anterior que:

- *El conocimiento nuevo $X = x$ influye en la solución cuanto mayor sea el coeficiente (normalizado) de covarianza cruzada ρ. El coeficiente establece el grado de correlación (dependencia en el caso Gaussiano) entre S y X.*
- *Del mismo modo, si la dispersión en S es elevada, el término σ_S será grande. Ello supone que el conocimiento inicial sobre S a través de μ_S es poco representativo del valor real, prestando mayor atención a la observación, en la medida en que $\left(x-\mu_X\right)$ sea elevado en relación a la dispersión observada en X, es decir, σ_X. Si la observación se ajusta a su media $x \approx \mu_X$, el modelo está ajustado y no se actualiza.*

c. Obtenga la varianza asociada a la estimación MAP anterior.

Res.: Tenemos que:

$$E\left[\hat{S}_{MAP}\right] = \mu_S \Rightarrow E\left[\left(\hat{S}_{MAP} - E\left(\hat{S}_{MAP}\right)\right)^2\right] = E\left[\left(\frac{\rho\sigma_S}{\sigma_X}\left(x-\mu_X\right)\right)^2\right] = \rho^2\sigma_S^2 \leq \sigma_S^2$$

dado que vemos que $E\left[\hat{S}_{MAP}\right] = \mu_S$.

d. Indique si la nueva observación aumenta o reduce la varianza asociada a la estimación MAP respecto a la varianza inicial σ_S^2.

Res.: Como $|\rho| \leq 1$, vemos que la varianza siempre se reduce o como máximo se mantiene.

En las mismas condiciones, consideramos ahora un caso más concreto, con un modelo de señal dado por:

$$X = S + W$$

con W un término de ruido de media cero y potencia σ_W^2.

e. Obtenga el estimador MAP para este caso.

Res.: Para este caso, no sólo conocemos el problema estadísticamente, sino que también el modelo de señal concreto. Tenemos que:

$$\hat{S}_{MAP} = \mu_{S/X}\left(x\right) = \mu_S + \rho\frac{\sigma_S}{\sigma_X}\left(x-\mu_X\right) = \mu_S + \frac{\mathrm{cov}(S,X)}{\sigma_X^2}\left(x-\mu_X\right)$$

con:

$$\begin{aligned}
&\operatorname{cov}(S,X) = E[(S-\mu_S)(X-\mu_X)] = E[(S-\mu_S)(X-\mu_S)] = \\
&= E[(S-\mu_S)(S+W-\mu_S)] = E[(S-\mu_S)(S-\mu_S)] + E[(S-\mu_S)W] = \\
&= E\left[(S-\mu_S)^2\right] = \sigma_S^2
\end{aligned}$$

Donde hemos hecho uso de que: $E[W]=0$ *o alternativamente también* $\mu_X = E[X] = E[S+W] = \mu_S$ *para este caso. Por otro lado:*

$$\sigma_X^2 = E\left[(X-\mu_X)^2\right] = E\left[(S+W-\mu_S)^2\right] = E\left[(S-\mu_S)^2\right] + E\left[W^2\right] = \sigma_S^2 + \sigma_W^2$$

Finalmente:

$$\hat{S}_{MAP} = \mu_S + \frac{\operatorname{cov}(S,X)}{\sigma_X^2}(x-\mu_X) = \mu_S + \frac{\sigma_S^2}{\sigma_S^2+\sigma_w^2}(x-\mu_X) = \mu_S + \frac{SNR}{SNR+1}(x-\mu_X)$$

f. Particularice la solución para los casos límite $SNR \to 0$ y $SNR \to \infty$, siendo $SNR = \dfrac{\sigma_S^2}{\sigma_W^2}$. Justifique el comportamiento de la solución en cada caso. (Comentario: a pesar de la notación, la SNR no es estrictamente una relación señal a ruido en este caso, aunque si mantiene una cierta relación con esta).

Res.: Vemos que:

$$\lim_{SNR\to 0} \hat{S}_{MAP} = \lim_{SNR\to 0}\left\{\mu_S + \frac{SNR}{SNR+1}(x-\mu_X)\right\} \to \mu_S$$

Cuando tenemos que $SNR = \sigma_S^2/\sigma_W^2$ *es una magnitud muy pequeña, la dispersión del parámetro que queremos estimar es mucho menor que potencia de ruido en la observación, dicho de otro modo, o bien* μ_S *es una muy buena representación de* S *o bien, las observaciones* $X = x$ *son tan ruidosas que no aportan información (o ambas circunstancias a la vez).*

En el caso opuesto:

$$\lim_{SNR\to\infty} \hat{S}_{MAP} = \lim_{SNR\to\infty}\left\{\mu_S + \frac{SNR}{SNR+1}(x-\mu_X)\right\} \to \mu_S + (x-\mu_X) = \mu_S + (x-\mu_S) = x$$

Vemos que o bien la dispersión en S *es muy elevada y/o la observación está libre de ruido* $\sigma_W^2 \to 0$*, en otras palabras, el conocimiento inicial sobre* S *a través de* μ_S *no es representativo y/o la observación es muy informativa.*

Ejercicio 2: Una propiedad de la matriz de autocorrelación $\underline{\underline{R}}$ es la siguiente. Si tomamos una observación vectorial compleja $\underline{X}(n)$ compuesta por N componentes, sabemos que $\underline{\underline{R}} = E\left[\underline{X}(n)\underline{X}^H(n)\right] = \underline{\underline{Q}}\,\underline{\underline{\Lambda}}\,\underline{\underline{Q}}^H$, con $\underline{\underline{Q}}$ y $\underline{\underline{\Lambda}}$ las matrices de autovectores y autovalores, respectivamente, ambas dependientes de los estadísticos de segundo orden de los datos observados. Cuando N es suficientemente grande, la matriz de autovectores muestra el siguiente comportamiento:

$$\lim_{N\to\infty} \underline{\underline{Q}} \to \underline{\underline{F}} = \frac{1}{\sqrt{N}}\left[\underline{S}_0\ \underline{S}_1\ \underline{S}_2\ \ldots\ \underline{S}_{N-1}\right]$$

$$con: \quad \underline{S}_k = \left[1\ \exp\left(j2\pi k/N\right) \ldots \exp(j2\pi(N-1)k/N)\right]^H$$

es decir, la matriz de autovectores tiende a la matriz de Fourier $\underline{\underline{F}}$, independientemente de los datos analizados.

Para la matriz de autovalores el comportamiento es el siguiente:

$$\lim_{N\to\infty} \underline{\underline{\Lambda}} \to diag\left(S_{xx}(0)\,S_{xx}(1)\ldots S_{xx}(N-1)\right)$$

donde $S_{xx}(k)$ es la *densidad espectral de potencia* de la señal a la frecuencia $f = k/N$.

a. A partir de la definición general de espectro promedio de potencia, justifique la siguiente relación:

$$S_{xx}(k) = \lim_{N\to\infty} \frac{1}{N}\underline{S}_k^H \underline{\underline{R}}\,\underline{S}_K$$

Res.: A partir de la definición original de espectro promedio de potencia:

$$S_{xx}(f) = \lim_{N\to\infty} \frac{1}{N} E\left[\left|\sum_{n=0}^{N-1} x(n)e^{-j2\pi fn}\right|^2\right] = \lim_{N\to\infty} \frac{1}{N} E\left[\left|\underline{S}_f^H\,\underline{X}(n)\right|^2\right] =$$

$$= \lim_{N\to\infty} \frac{1}{N}\underline{S}_f^H E\left[X(n)\underline{X}^H(n)\right]\underline{S}_f = \lim_{N\to\infty} \frac{1}{N}\underline{S}_f^H \underline{\underline{R}}\,\underline{S}_f$$

Para el enunciado del problema, tomamos los valores discretos $f = k/N$.

b. Demuestre que la expresión anterior es correcta asintóticamente ($N\to\infty$), utilizando para ello el comportamiento límite de las matrices $\underline{\underline{Q}}$ y $\underline{\underline{\Lambda}}$ descritos en el enunciado.

Res.: Tenemos,

$$\underline{\underline{F}}^H \underline{\underline{F}} = \underline{\underline{F}}\,\underline{\underline{F}}^H = \underline{\underline{I}} \Rightarrow \underline{\underline{F}}^H \underline{S}_k = \sqrt{N}\left[0\;0\ldots 1_{k-\acute{e}sima}\ldots 0\right]^T = \sqrt{N}\underline{1}_k$$

luego, entre otras opciones:

$$S_{xx}(k) = \lim_{N\to\infty} \frac{1}{N} \underline{S}_k^H \underline{\underline{R}}\underline{S}_k = \lim_{N\to\infty} \frac{1}{N} \underline{S}_k^H \underline{\underline{F}}\,\underline{\underline{\Lambda}}\,\underline{\underline{F}}^H \underline{S}_k =$$
$$= \lim_{N\to\infty} \frac{1}{N} tr\left[\underline{\underline{\Lambda}}\,\underline{\underline{F}}^H \underline{S}_k \underline{S}_k^H \underline{\underline{F}}\right] = \lim_{N\to\infty} \frac{N}{N} tr\left[\underline{\underline{\Lambda}}\underline{1}_k \underline{1}_k^H\right] = S_{xx}(k)$$

c. Deduzca analíticamente el estimador de Capon para un valor de N finito.

Res.: A partir de las notas/apuntes de la asignatura (capítulos 1 y 3), se obtiene la solución conocida:

$$P_{xx}(k) = \frac{1}{\underline{S}_k^H \underline{\underline{R}}^{-1} \underline{S}_k}$$

d. Analice la solución de Capon anterior para el caso límite $N \to \infty$ y comente la solución, comparándola con la obtenida en (b.).

Res.: Seguimos el mismo razonamiento que en (b.), de modo que ahora:

$$\lim_{N\to\infty} \ln\left(P_{xx}(k)\right) = \lim_{N\to\infty} -\ln\left(\underline{S}_k^H \underline{\underline{R}}^{-1} \underline{S}_k\right) = \lim_{N\to\infty} -\ln\left(\underline{S}_k^H \underline{\underline{F}}\,\underline{\underline{\Lambda}}^{-1} \underline{\underline{F}}^H \underline{S}_k\right) =$$
$$= \lim_{N\to\infty} -\ln\left(tr\left[\underline{\underline{\Lambda}}^{-1} \underline{\underline{F}}^H \underline{S}_k \underline{S}_k^H \underline{\underline{F}}\right]\right) = \lim_{N\to\infty} -\ln\left(N tr\left[\underline{\underline{\Lambda}}^{-1} \underline{1}_k \underline{1}_k^H\right]\right) =$$
$$= \lim_{N\to\infty} \ln\left(S_{xx}(k)/N\right) \to 0$$

Dado que el estimador de Capon estima potencia y no densidad espectral de potencia, vemos que asintóticamente tiende a cero, lo que es razonable al tener que distribuir la potencia de la señal en un número arbitrariamente grande de bins frecuenciales, o si se quiere, al tener que medir la potencia a la salida de un filtro de ancho de banda cada vez menor. Si analizamos el filtro:

$$\underline{W}_k = \frac{1}{\underline{S}_k^H \underline{\underline{R}}^{-1} \underline{S}_k} \underline{\underline{R}}^{-1} \underline{S}_k \Rightarrow BW \approx \left\|\underline{W}_k\right\|^2 = \frac{\underline{S}_k^H \underline{\underline{R}}^{-2} \underline{S}_k}{\left(\underline{S}_k^H \underline{\underline{R}}^{-1} \underline{S}_k\right)^2} \to \frac{1}{N}$$

Para órdenes elevados del filtro, el ancho de banda es N^{-1}. De hecho, la ecuación anterior muestra tal afirmación. Sería razonable formular para órdenes elevados:

$$S_{xx}(k) = NP_{xx}(k) = N \frac{1}{\underline{S}_k^H \underline{\underline{R}}^{-1} \underline{S}_k}$$