Escola Tècnica Superior d'Enginyeria de Telecomunicació de Barcelona UNIVERSITAT POLITÈCNICA DE CATALUNYA **DEPARTAMENT DE TEORIA DEL SENYAL I COMUNICACIONS**	**Processament del Senyal** Data d'examen: 9-Juny-2010 Data notes provisionals: 22-Juny-2010 Període d'al.legacions: fins el 28-Juny-2010 Data notes definitives: 30-Juny-2010

Professors: Ferran Marques, Albert Oliveras, Gregori Vazquez.

Informacions addicionals:

- Durada de la prova: 3h.
- Justificar tots els resultats.
- No es poden utilitzar apunts, llibres, calculadores, telèfons mòbils, PDAs etc…
- RESOLDRE ELS PROBLEMES EN FULLS SEPARATS.

Problema 1 (1/3)

Disponemos de una observación de la forma: $y=\sum_{n=1}^{N}x_n$ donde $\{x_n\}_{1\le n\le N}$ son variables aleatorias Gaussianas reales, estadísticamente independientes, de media nula y varianza σ_x^2. *Para resolver el ejercicio, considere que N es una magnitud continua.*

1. Obtenga la estimación ML de N.

Vemos que y es Gaussiana al ser la suma de variables aleatorias Gaussianas y por tanto, caracterizada por su media y varianza. Tenemos:

$$E[y]=\sum_{n=1}^{N}E[x_n]=0$$

$$E[y^2]=\sum_{n=1}^{N}\sum_{m=1}^{N}E[x_nx_m]=\sum_{n=1}^{N}E[x_n^2]=N\sigma_x^2$$

Por tanto:

$$f(y/N)=\frac{1}{\sqrt{2\pi N\sigma_x^2}}e^{-\frac{y^2}{2N\sigma_x^2}}\Rightarrow \ln f(y/N)=-\frac{1}{2}\ln(2\pi N\sigma_x^2)-\frac{y^2}{2N\sigma_x^2}$$

El enunciado nos sugiere considerar N como una magnitud continua, de modo que:

$$\frac{\partial}{\partial N}\ln f(y/N)=-\frac{1}{2}\frac{2\pi\sigma_x^2}{2\pi N\sigma_x^2}+\frac{y^2}{2N^2\sigma_x^2}=-\frac{1}{2N}+\frac{y^2}{2N^2\sigma_x^2}$$

La solución ML se obtiene buscando el máximo:

$$\frac{\partial}{\partial N}\ln f(y/N)=-\frac{1}{2N}+\frac{y^2}{2N^2\sigma_x^2}=0\Rightarrow \hat{N}_{ML}=\frac{y^2}{\sigma_x^2}$$

2. Analice si la estimación ML de N es sesgada o no.

Tomamos el valor esperado del estimador ML:

$$E[\hat{N}_{ML}]=\frac{E[y^2]}{\sigma_x^2}=\frac{N\sigma_x^2}{\sigma_x^2}=N$$

y vemos que no presenta sesgo.

3. Obtenga la varianza de la estimación ML de N.

Tenemos:

$$\mathrm{var}\left(\hat{N}_{ML}\right) = E\left[\left(\hat{N}_{ML} - E\left[\hat{N}_{ML}\right]\right)^2\right] = E\left[\left(\hat{N}_{ML} - N\right)^2\right] = E\left[\hat{N}_{ML}^2\right] - N^2 =$$

$$= E\left[\frac{y^4}{\sigma_x^4}\right] - N^2 = \frac{3N^2\sigma_x^4}{\sigma_x^4} - N^2 = 2N^2$$

4. Obtenga el error cuadrático medio asociado a la estimación ML de N .

En el apartado (2.) hemos visto que el estimador es no sesgado y por tanto la varianza y el error cuadrático medio del estimador coinciden.

5. Indique si el estimador ML de N es eficiente o no.

La cota de Cramer-Rao aplica a estimadores no sesgados, como el que nos ocupa. En el apartado (1.) hemos obtenido que:

$$\frac{\partial}{\partial N}\ln f(y/N) = -\frac{1}{2N} + \frac{y^2}{2N^2\sigma_x^2} = -\frac{1}{2N^2}\left(N - \frac{y^2}{\sigma_x^2}\right)$$

Lo que nos muestra que el estimador ML obtenido es eficiente, con una varianza mínima dada por $2N^2$ *que coincide con la obtenida en (3.).*

6. Indique como se modifica la estimación ML de N si en lugar de ser considerada como una magnitud continua es tratada de manera correcta como entera positiva.

Si la magnitud N *es entera la función ML* $f(y/N)$ *no es derivable en* N *, dado que sólo queda definida para valores enteros positivos. Tendremos que buscar el máximo por cálculo intensivo sobre todos los posibles valores de* N *.*

<u>*Opcionalmente*</u>*, si queremos ser más ambiciosos en la respuesta, tenemos que:*

$$\hat{N}_{ML} = \arg máx_N f(y/N) \Rightarrow \hat{N}_{ML} = N \; si \; f(y/N) > f(y/M) \text{ para todo } M \neq N \text{ entero positivo}$$

de modo equivalente:

$$\ln f(y/N) > \ln f(y/M)$$

Luego:

$$\ln f(y/N) = -\frac{1}{2}\ln\left(2\pi N\sigma_x^2\right) - \frac{y^2}{2N\sigma_x^2} > \ln f(y/M) = -\frac{1}{2}\ln\left(2\pi M\sigma_x^2\right) - \frac{y^2}{2M\sigma_x^2}$$

bien:

$$-\ln(N) - \frac{y^2}{N\sigma_x^2} > -\ln(M) - \frac{y^2}{M\sigma_x^2}$$

Problema 2 (1/3)

Frente al posible desvanecimiento de un canal, una técnica para paliar la pérdida del mensaje transmitido es realizar una transmisión adicional mediante un canal secundario, siendo esta segunda transmisión típicamente de peor calidad que la transmisión principal (*Multiple Description Coding*).

En este ejercicio se va a analizar un caso simple de este tipo de técnica. Para ello, se va a suponer que se transmite por el canal principal el proceso x(n), que se puede modelar como un proceso AR(1), donde la potencia del ruido blanco generador es s_w^2 y el coeficiente de correlación entre muestras consecutivas de señal es

$$\rho = \frac{r_x(1)}{r_x(0)}$$

1) Partiendo de las ecuaciones de Yule-Walker, hallad una expresión de las muestras de la autocorrelación de x(n) en función de la potencia de x(n) y el coeficiente r. Determinad el valor de la potencia del proceso x(n).

Por ser un proceso AR(1) se tiene la siguiente ecuación

$$r_x[l] = \rho r_x[l-1] + \sigma_w^2 \delta[l]$$

Evaluando la expresión en l = 0 y l = 1 se tiene

$$r_x[0] = \rho r_x[-1] + \sigma_w^2 \qquad\qquad r_x[1] = \rho r_x[0]$$

Como la señal x[n] es real, $r_x[l] = r_x[-l]$ y despejando

$$r_x[0] = \frac{\sigma_w^2}{1-\rho^2}$$

Dado que para valores del índice l > 0 se tiene la siguiente expresión

$$r_x[l] = \rho r_x[l-1] = \rho^2 r_x[l-2] = \cdots$$

se puede escribir la expresión siguiente:

$$r_x[l] = \rho^{|l|} r_x[0]$$

2) Determinad el orden del filtro lineal de predicción que permite, a partir de las muestras previas de *x*(*n*), minimizar la potencia del error de predicción y dad el valor de sus coeficientes. Calculad la potencia mínima del error de predicción obtenida, así como la ganancia de predicción G_1, definida como el cociente entre la potencia de la señal y la potencia del error.

Ya que se trata de un proceso AR(1) y se propone utilizar un filtro causal, el filtro óptimo tendrá un único coeficiente. De esta manera, la solución de Wiener genérica $R_x \hat{\vec{h}} = \vec{r}_{xd}$ *queda* $r_x(0)\hat{h} = r_x(1)$ *y el filtro óptimo es*

$$\hat{h} = \frac{r_x(1)}{r_x(0)} = \rho$$

La expresión de la potencia del error es

$$E\left\{|e[n]|^2\right\} = r_d[0] - \vec{h}^T \vec{r}_{xd} = r_x[0] - \rho r_x[1] = r_x[0](1-\rho^2) = \sigma_w^2$$

Y por tanto se observa que la potencia del error de predicción coincide con la potencia del ruido generador del proceso AR(1). En este caso, la ganancia de predicción es

$$G_1 = \frac{r_x(0)}{E\left\{|e(n)|^2\right\}} = \frac{\sigma_w^2}{1-\rho^2}\frac{1}{\sigma_w^2} = \frac{1}{1-\rho^2}$$

Dado que el módulo del coeficiente de correlación ha de ser menor o igual a 1, la ganancia es mayor que la unidad. Además, se cumple que si la señal está muy correlada (r → 1), la ganancia de predicción tiende a infinito mientras que si la señal está muy decorrelada (r → 0), se obtiene poca mejora por el hecho de realizar la predicción.

Por el canal adicional se transmite una señal de peor calidad que, en este caso, vamos a representar como una versión submuestreada por un factor M del proceso $x(n)$: $z(n) = x(Mn)$.

3) Calculad la autocorrelación de la señal $z(n)$. Dad el valor del coeficiente de correlación entre muestras consecutivas para el proceso $z(n)$. Aprovechando el resultado del apartado (2), dad la expresión del filtro lineal de orden 1 que permite predecir las muestras de $z(n)$ a partir de sus muestras previas minimizando la potencia del error de predicción. Calculad la potencia mínima del error de predicción obtenida, así como la ganancia de predicción G_2.

$$r_z(l) = E\left\{z(n)z^*(n-l)\right\} = E\left\{x(Mn)\,x^*(M(n-l))\right\} = r_x(Ml)$$

Dada esta expresión de la autocorrelación, el nuevo coeficiente de correlación r' se puede hallar como

$$\rho' = \frac{r_z(1)}{r_z(0)} = \frac{r_x(M)}{r_x(0)} = \frac{\rho^M r_x(0)}{r_x(0)} = \rho^M$$

Por tanto, extrapolando los resultados del apartado (2) se puede dar las siguientes expresiones:

Filtro Lineal de orden 1: $\hat{h} = \dfrac{r_z(1)}{r_z(0)} = \rho^M$

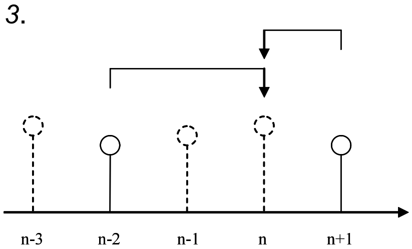
Potencia de error mínima:

$$E\left\{|e[n]|^2\right\} = r_z[0] - \rho^M r_z[1] = r_z[0](1-\rho^{2M}) = r_x[0](1-\rho^{2M}) = \sigma_w^2\,\frac{1-\rho^{2M}}{1-\rho^2}$$

Se puede observar que el error de predicción es de potencia mayor que en el caso anterior ya que la correlación entre la muestra que se utiliza para hacer la predicción y la que se desea predecir es menor.

Ganancia de predicción: $G_2 = \dfrac{r_z(0)}{E\left\{|e(n)|^2\right\}} = \dfrac{r_z(0)}{r_z(0)(1-\rho^{2M})} = \dfrac{1}{1-\rho^{2M}}$

Cuando el canal principal sufre un desvanecimiento, se recupera la señal recibida por el canal adicional. Para ello, se recuperan las *M-1* muestras de la señal $x(n)$ que se habían eliminado previamente de la señal $z(n)$. Estas muestras se recuperan mediante interpolación de las muestras de $z(n)$, lo cual introduce un retardo que se considera despreciable. Para estudiar este caso se va a tomar $M = 3$.



4) Dad el valor de los coeficientes del filtro lineal que permite obtener la muestra $x(n)$ a partir de las dos muestras más próximas en el tiempo de la señal $z(n)$ con menor potencia del error de predicción. Calculad la potencia mínima del error de predicción obtenida, así como la ganancia de predicción G_3.

La solución de Wiener genérica $R_x\hat{\vec{h}} = \vec{r}_{xd}$ *queda*

$$R_x = \begin{bmatrix} r_x[0] & r_x[-3] \\ r_x[3] & r_x[0] \end{bmatrix} \qquad \vec{r}_{xd} = \begin{bmatrix} r_x[-2] \\ r_x[1] \end{bmatrix} \qquad \begin{bmatrix} r_x[0] & \rho^3 r_x[0] \\ \rho^3 r_x[0] & r_x[0] \end{bmatrix}\begin{bmatrix} h_0 \\ h_1 \end{bmatrix} = \begin{bmatrix} \rho^2 r_x[0] \\ \rho r_x[0] \end{bmatrix}$$

$$\begin{bmatrix} 1 & \rho^3 \\ \rho^3 & 1 \end{bmatrix}\begin{bmatrix} h_0 \\ h_1 \end{bmatrix} = \begin{bmatrix} \rho^2 \\ \rho \end{bmatrix}$$ *y solucionando el sistema* $$\begin{bmatrix} h_0 \\ h_1 \end{bmatrix} = \frac{1}{1-\rho^6}\begin{bmatrix} \rho^2 - \rho^4 \\ \rho - \rho^5 \end{bmatrix}$$

La potencia del error viene dada por

$$E\left\{|e[n]|^2\right\} = r_d[0] - \vec{h}^T \vec{r}_{xd} = r_x[0] - \frac{\rho-\rho^5}{1-\rho^6} r_x[2] - \frac{\rho^2-\rho^4}{1-\rho^6} r_x[1] =_x [0] - \frac{\rho-\rho^5}{1-\rho^6}\rho^2 r_x[0] - \frac{\rho^2-\rho^4}{1-\rho^6}\rho r_x[0]$$

$$E\left\{|e[n]|^2\right\} = r_x[0]\frac{1-\rho^2-\rho^4+\rho^6}{1-\rho^6}$$

Y la ganancia de predicción: $$G_3 = \frac{r_x(0)}{E\left\{|e(n)|^2\right\}} = \frac{1-\rho^6}{1-\rho^2-\rho^4+\rho^6}$$

Problema 3 (1/3)

La inestabilitat dels algorismes adaptatius es pot observar si els coeficients del vector $\underline{h}[n]$ prenen valors elevats (valors molt grans de coeficients del vector $\underline{h}[n]$ pot portar a una degeneració de les prestacions del l'algorisme, especialment si s'utilitzen processadors amb artimètica finita). Una idea per evitar aquests efectes seria introduïr un factor de penalització en l'equació d'adaptació del vector $\underline{h}[n]$. Per això, es defineix una nova funció de cost a minimitzar:

$$\varepsilon^2[n] = E\left\{\left|e[n]\right|^2\right\} + \alpha\left\|\underline{h}[n]\right\|^2$$

on $\alpha \geq 0$, és el factor que pondera la penalització dels valors grans de $\underline{h}[n]$ i $\left\|\underline{h}[n]\right\|^2 = \underline{h}^T[n]\,\underline{h}[n]$.

Es demana:

1. La funció gradient de $\varepsilon^2[n]$ respecte de $\underline{h}[n]$.
2. Demostrar que la introducció del terme $\alpha\left\|\underline{h}[n]\right\|^2$ en la funció de cost és equivalent a sumar al senyal $x[n]$ un soroll blanc de mitjana nul·la i variància α (suposant que el senyal i el soroll són independents).
3. Determinar l'equació d'actualització de $\underline{h}[n]$ segons el mètode del gradient.
4. Determinar el marge de valors de μ que assegura la convergència de l'algorisme adaptatiu.

En el cas en que $\underline{h}[n]$ sigui de dos coeficients $\underline{h}[n] = \begin{bmatrix} h_0[n] \\ h_1[n] \end{bmatrix}$ i $\alpha = 0{,}5$, $\underline{p} = \underline{r}_{xd} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ i $\underline{\underline{R}}_x = \begin{bmatrix} 1 & 0{,}5 \\ 0{,}5 & 1 \end{bmatrix} =$
$= \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1{,}5 & 0 \\ 0 & 0{,}5 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$.

5. Dibuixar la superfície d'error (en corves de nivell) en funció de $h_0[n]$ i $h_1[n]$. Indicar clarament quin és el punt mínim i el seu valor, els eixos principals i la direcció de les corves de nivell.
6. Si $\underline{h}[0] = \underline{0}$, dibuixar aproximadament l'evolució de $h_0[n]$ i $h_1[n]$ en funció de n, pels valors de μ=0,1, μ=0,25 i μ=1.

(Solució)

1.

$$\begin{aligned}\varepsilon^2[n] &= E\left\{\left|\left(d[n]-y[n]\right)\right|^2\right\} + \alpha\,\underline{h}^T[n]\,\underline{h}[n] = \\ &= E\left\{\left|d[n]\right|^2\right\} - 2E\left\{\underline{h}^T[n]\underline{x}[n]d[n]\right\} + E\left\{\underline{h}^T[n]\underline{x}[n]\underline{x}^T[n]\underline{h}[n]\right\} + \alpha\,\underline{h}^T[n]\,\underline{h}[n] = \\ &= r_d[0] - 2\underline{h}^T[n]\underline{p} + \underline{h}^T[n]\underline{\underline{R}}_x\,\underline{h}[n] + \alpha\,\underline{h}^T[n]\,\underline{h}[n]\end{aligned}$$

on $\nabla\varepsilon^2[n] = -2\underline{p} + 2\underline{\underline{R}}_x\underline{h}[n] + 2\alpha\,\underline{h}[n]$

2. Suposant que el senyal i el soroll són independents, el resultat obtingut en l'apartat anterior es pot interpretar que la matriu $\left(\underline{\underline{R}}_x + \alpha\underline{\underline{I}}\right)$ és el nou senyal d'entrada composat per un senyal amb matriu d'autocorrelació $\underline{\underline{R}}_x$ i per un soroll de variància α:

$$\begin{aligned}\varepsilon^2[n] &= r_d[0] - 2\underline{h}^T[n]\underline{p} + \underline{h}^T[n]\underline{\underline{R}}_x\,\underline{h}[n] + \alpha\,\underline{h}^T[n]\,\underline{h}[n] \\ &= r_d[0] - 2\underline{h}^T[n]\underline{p} + \underline{h}^T[n]\left(\underline{\underline{R}}_x + \alpha\underline{\underline{I}}\right)\underline{h}[n]\end{aligned} \quad \text{i } \nabla\varepsilon^2[n] = -2\underline{p} + 2\left(\underline{\underline{R}}_x + \alpha\underline{\underline{I}}\right)\underline{h}[n]$$

També es pot observar en aquest resultat coincideix amb el resultat conegut (sense el terme $\alpha\,\underline{h}^T[n]\,\underline{h}[n]$ si fem que α=0).

3.

$$\left.\begin{aligned}\underline{h}[n+1] &= \underline{h}[n] + \mu\left(-\nabla\varepsilon^2[n]\right) \\ \nabla\varepsilon^2[n] &= -2\underline{p} + 2\left(\underline{\underline{R}}_x + \alpha\underline{\underline{I}}\right)\underline{h}[n]\end{aligned}\right\} \quad \underline{h}[n+1] = \underline{h}[n] - 2\mu\left(\underline{p} - \left(\underline{\underline{R}}_x + \alpha\underline{\underline{I}}\right)\underline{h}[n]\right)$$

4. La condició de convergència ara és que: $\left|1 - 2\mu\left(\lambda_i + \alpha\right)\right| < 1 \quad \forall i$ on $\lambda_i + \alpha$ son els autovalors de la matriu $\left(\underline{\underline{R}}_x + \alpha\underline{\underline{I}}\right)$ i λ_i els de la matriu $\underline{\underline{R}}_x$, per tant els valors de μ en que l'algorisme adaptatiu convergira son:

$$0 < \mu < \frac{1}{(\lambda_{\max} + \alpha)}$$

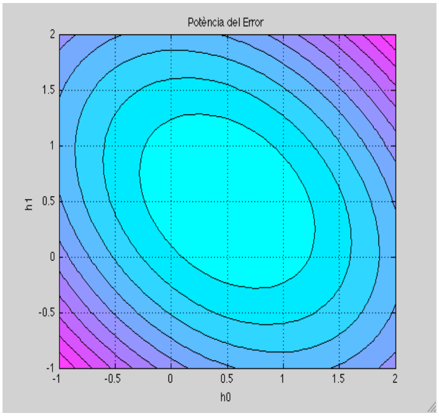
5.

El punt d'error mínim vindra donat pel filtre òptim $\underline{h}_0 = \left(\underline{\underline{R}}_x + \alpha \underline{\underline{I}}\right)^{-1} \underline{p} = \begin{bmatrix} 1{,}5 & 0{,}5 \\ 0{,}5 & 1{,}5 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{3}{4} & \frac{-1}{4} \\ \frac{-1}{4} & \frac{3}{4} \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$

la potència del error en el mínim vindra donada per:

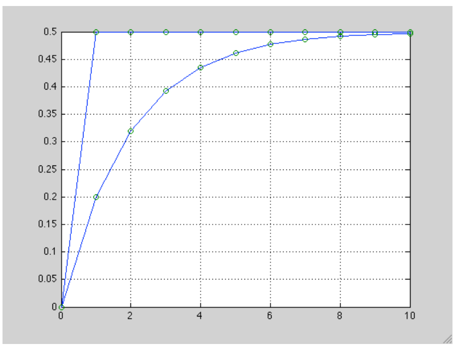
$$\varepsilon_{\min}{}^2 = r_d[0] + \alpha - \underline{h}_o{}^T \underline{p} = r_d[0] + 0{,}5 - [0{,}5 \quad 0{,}5] \begin{bmatrix} 1 \\ 1 \end{bmatrix} = r_d[0] - 0{,}5$$

Els autovalors de la matriu $\left(\underline{\underline{R}}_x + \alpha \underline{\underline{I}}\right)$ son 2 i 1 i els seus autovectors associats $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ i $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix}$, per tant el semieix mes gran sera 1 en la direcció $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix}$ i el mes petit $\frac{1}{2}$ (associat a l'invers de 2) en la direcció $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ tal com es mostra en la figura següent:



on el valor mínim de la potència del error es dona en el punt $\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$ i el seu valor és $r_d[0] - 0{,}5$ (o 0,5 si s'ha considerat que d[n] és una seqüència de potència unitat).

6. Tal com s'ha defnit l'algorime adaptatiu, la convergència en aquest cas estarà assegurada pels valors de $0 < \mu < \frac{1}{(\lambda_{\max} + \alpha)} = \frac{1}{(1{,}5 + 0{,}5)} = \frac{1}{2}$ per tant podem afirmar que l'algorisme adaptatiu per $\mu = 1$ no convergirà per $\mu = 0{,}25$ i $\mu = 0{,}1$ l'algorisme tant $h_0[n]$ com $h_1[n]$ covergiran al mateix valor 0,5. Evidenment si $\mu = 0{,}1$ ho farà més lentament que amb $\mu = 0{,}25$. Com que $\mu = 0{,}25$ és la meitat del valor maxim que pot tenir μ per aquest valor convergirà amb la màxima velocitat: 1 sola iteració.



Tampoc hi haura error de desajust ja que s'utilitza el valor del gradient i no una estimació instantània com en el cas del LMS.