Escola Tècnica Superior d'Enginyeria de Telecomunicació de Barcelona UNIVERSITAT POLITÈCNICA DE CATALUNYA DEPARTAMENT DE TEORIA DEL SENYAL I COMUNICACIONS	**Procesado de Señal** Fecha examen: 25 de Junio de 2009 Fecha de entrega de notas: 1 de Julio de 2009
Profesores: Ferran Marqués, Albert Oliveras, Josep Vidal	
Información adicional: • Duración del control: 3 h • No pueden utilizarse libros, ni apuntes, ni calculadoras, ni otros dispositivos electrónicos • Utilizad hojas separadas para resolver cada problema Justificad todos los resultados	

Problema 1 (35%): Disponemos de un conjunto de observaciones d(M-1),…, d(N-1) (M<N) de un proceso que sabemos que están generadas por la suma de dos componentes: un proceso e(n) blanco Gausiano de media nula y potencia σ_e^2 y una señal determinista z(n) procedente del filtrado lineal de una señal x(0),…, x(N-1) conocida con un filtro **h** de M coeficientes. El filtrado de las muestras x(n) mediante **h** puede escribirse en forma matricial como **z** = **X h**, donde

$$\mathbf{z}=\begin{bmatrix} z(M-1)\\ z(M)\\ \vdots\\ z(N-1)\end{bmatrix}\qquad \mathbf{X}=\begin{bmatrix} x(M-1) & x(M-2) & \cdots & x(0)\\ x(M) & x(M-1) & \cdots & x(1)\\ \vdots & \vdots & \ddots & \vdots\\ x(N-2) & x(N-3) & \cdots & x(N-M-1)\\ x(N-1) & x(N-2) & \cdots & x(N-M)\end{bmatrix}\in\mathbb{C}^{(N-M+1)\times M}\qquad \text{h}=\begin{bmatrix} h(0)\\ h(1)\\ \vdots\\ h(M-1)\end{bmatrix}$$

y el vector que contiene la señal **d**:

$$\mathbf{d}=\mathbf{e}+\mathbf{Xh}\qquad \mathbf{e}=\begin{bmatrix} e(M-1)\\ e(M)\\ \vdots\\ e(N-1)\end{bmatrix}\qquad \mathbf{d}=\begin{bmatrix} d(M-1)\\ d(M)\\ \vdots\\ d(N-1)\end{bmatrix}$$

El objetivo del ejercicio es doble. En primer lugar, se pretende obtener el estimador de **h** siguiendo el criterio de máxima verosimilitud. En segundo lugar, se quiere determinar un criterio para diseñar la señal x(n) óptima. Para todo ello se pide:

1. Determine la media de **d** y su matriz de covarianza. Escriba la función de verosimilitud de **d**.

$$E\{\mathbf{d}\}=E\{\mathbf{e}\}+\mathbf{Xh}=\mathbf{Xh}\qquad \mathbf{C_d}=E\left\{[\mathbf{d}-\mathbf{Xh}][\mathbf{d}-\mathbf{Xh}]^H\right\}=\sigma_e^2\mathbf{I}$$

$$f(\mathbf{d})=\frac{1}{\pi^P\sigma_e^{2P}}\exp\left(-\frac{1}{\sigma_e^2}(\mathbf{x}-\mathbf{Hh})^H(\mathbf{x}-\mathbf{Hh})\right)$$

2. ¿Cuál es la estimación de máxima verosimilitud de **h**?

$$\hat{\mathbf{h}}_{ML}=\arg\max_{\mathbf{h}}\ln f(\mathbf{d})\qquad \nabla_{\mathbf{h}^*}\ln f(\mathbf{d})=-\frac{1}{\sigma_e^2}\mathbf{X}^H(\mathbf{d}-\mathbf{Xh})=\mathbf{0}\qquad \hat{\mathbf{h}}_{ML}=\left(\mathbf{X}^H\mathbf{X}\right)^{-1}\mathbf{X}^H\mathbf{d}$$

3. Calcule el sesgo del estimador $\mathbf{h}_{ML}$ calculado en el apartado 2.

$$E\{\hat{\mathbf{h}}_{ML}\}=\left(\mathbf{X}^H\mathbf{X}\right)^{-1}\mathbf{X}^HE\{\mathbf{d}\}=\left(\mathbf{X}^H\mathbf{X}\right)^{-1}\mathbf{X}^HE\{\mathbf{e}+\mathbf{Xh}\}=\left(\mathbf{X}^H\mathbf{X}\right)^{-1}\mathbf{X}^H\mathbf{Xh}=\mathbf{h}$$

4. Calcule $\mathbf{C_h}=E\left\{[\mathbf{h}_{ML}-E\{\mathbf{h}_{ML}\}][\mathbf{h}_{ML}-E\{\mathbf{h}_{ML}\}]^H\right\}$ en función únicamente de σ_e^2 y de **X**.

$$\mathbf{C_h}=E\left\{[\mathbf{h}_{ML}-E\{\mathbf{h}_{ML}\}][\mathbf{h}_{ML}-E\{\mathbf{h}_{ML}\}]^H\right\}=E\left\{\left[\left(\mathbf{X}^H\mathbf{X}\right)^{-1}\mathbf{X}^H\mathbf{d}-\mathbf{h}\right]\left[\left(\mathbf{X}^H\mathbf{X}\right)^{-1}\mathbf{X}^H\mathbf{d}-\mathbf{h}\right]^H\right\}=$$

$$=E\left\{\left[\left(\mathbf{X}^H\mathbf{X}\right)^{-1}\mathbf{X}^H(\mathbf{e}+\mathbf{Xh})-\mathbf{h}\right]\left[\left(\mathbf{X}^H\mathbf{X}\right)^{-1}\mathbf{X}^H(\mathbf{e}+\mathbf{Xh})-\mathbf{h}\right]^H\right\}=$$

$$=E\left\{\left[\left(\mathbf{X}^H\mathbf{X}\right)^{-1}\mathbf{X}^H\mathbf{e}\right]\left[\left(\mathbf{X}^H\mathbf{X}\right)^{-1}\mathbf{X}^H\mathbf{e}\right]^H\right\}=E\left\{\left(\mathbf{X}^H\mathbf{X}\right)^{-1}\mathbf{X}^H\mathbf{e}\mathbf{e}^H\mathbf{X}\left(\mathbf{X}^H\mathbf{X}\right)^{-1}\right\}=$$

$$=\left(\mathbf{X}^H\mathbf{X}\right)^{-1}\mathbf{X}^HE\{\mathbf{e}\mathbf{e}^H\}\mathbf{X}\left(\mathbf{X}^H\mathbf{X}\right)^{-1}=\sigma_e^2\left(\mathbf{X}^H\mathbf{X}\right)^{-1}\mathbf{X}^H\mathbf{X}\left(\mathbf{X}^H\mathbf{X}\right)^{-1}=\sigma_e^2\left(\mathbf{X}^H\mathbf{X}\right)^{-1}$$

Diseñaremos la señal x(n) de forma que se minimice la suma de las varianzas de los coeficientes del filtro estimados: $\mathrm{var}\{\mathbf{h}_{ML}\} = traza(\mathbf{C}_\mathbf{h})$.

5. ¿Como deben ser los autovalores de la matriz $\mathbf{X}^H\mathbf{X}$ para que $traza(\mathbf{C}_\mathrm{h})$ sea mínima? Tenga en cuenta la desigualdad entre la media aritmética y la media armónica: para cualquier conjunto de valores a_1, a_2, …, a_n positivos entonces $H \leq A$ con

$$\frac{1}{H} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{a_i} \qquad\qquad A = \frac{1}{n}\sum_{i=1}^{n} a_i$$

cumpliéndose la igualdad si $a_1 = a_2 = \ldots = a_n$.

$traza\left[\left(\mathbf{X}^H\mathbf{X}\right)^{-1}\right] = \sum_{i=1}^{M}\frac{1}{\lambda_i} = \frac{M}{H}$ donde λ_i son los autovalores de $\mathbf{X}^H\mathbf{X}$, y son positivos por ser $\mathbf{X}^H\mathbf{X}$ una matriz hermítica. Para que la traza sea mímina H=A y por tanto todos los autovalores han de ser iguales.

6. Relacione la matriz $\mathbf{X}^H\mathbf{X}$ con la matriz de autocorrelación de la señal x(n). A la vista del resultado obtenido en el apartado (5) ¿cómo debe ser la función de correlación de x(n)?

$E\left\{\mathbf{X}^H\mathbf{X}\right\} = (N - M + 1)\mathbf{R}_x$ Para que todos los autovalores de la matriz de correlación sean iguales la función de correlación de x(n) debe ser una delta.

Nota: La función de densidad de probabilidad Gausiana para un vector aleatorio $\mathbf{x} \in \mathbb{C}^{P\times 1}$ es

$$f(\mathbf{x}) = \frac{1}{\pi^P \det(\mathbf{C}_x)}\exp\left(-(\mathbf{x}-\mathbf{m}_x)^H \mathbf{C}_x^{-1}(\mathbf{x}-\mathbf{m}_x)\right)$$

Problema 2 (35%): Sea una señal real x(n) que se puede modelar como un proceso AR(1) donde la potencia del ruido blanco generador es σ_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 ρ. 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

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

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

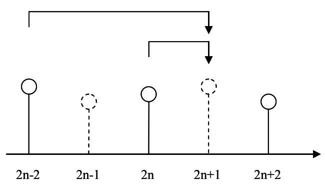
$$r_x[l] = \rho r_x[l-1]$$

se puede escribir la expresión siguiente:

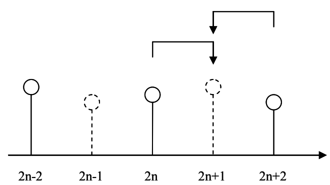
$$\boxed{r_x[l] = \rho^{|l|} r_x[0]}$$

De esta señal se tiene una versión z(n) que se ha obtenido eliminando las muestras impares de x(n).

Para reconstruir la señal x(n) a partir de las muestras de la señal z(n) se propone trabajar con dos predictores distintos, ambos de 2 coeficientes. Se utilizará en ambos casos la solución con error cuadrático medio mínimo, y se compararan tanto en su versión adaptativa como no adaptativa.



Predictor A



Predictor B

2) Hallad los coeficientes del predictor A que permite predecir las muestras impares de x(n) a partir de las muestras de z(n). Justificad el resultado. Calculad la potencia del error de predicción que se obtiene.

Si sólo disponemos de las muestras pares previas a la muestra a predecir, el filtro de 2 coeficientes procesará las muestras z[n] (= x([2n]) y z[n-2] (= x[2n-2]) para obtener una estimación de x[2n+1]. Así, el vector de entrada al filtro (vector de datos x[n]) y la señal de referencia (d[n]) serán

$$\underline{x}^T[n] = \left[z[n], z[n-1]\right] = \left[x[2n], x[2n-2]\right] \qquad\qquad d[n] = x[2n+1] \qquad .$$

La solución de Wiener genérica $R_x \hat{\vec{h}} = \vec{r}_{xd}$ queda (como la señal x[n] es real, $r_x[l] = r_x[-l]$)

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

$$\begin{bmatrix} 1 & \rho^2 \\ \rho^2 & 1 \end{bmatrix} \begin{bmatrix} h_0 \\ h_1 \end{bmatrix} = \begin{bmatrix} \rho \\ \rho^3 \end{bmatrix}$$ y solucionando el sistema $$\boxed{\begin{bmatrix} h_0 \\ h_1 \end{bmatrix} = \begin{bmatrix} \rho \\ 0 \end{bmatrix}}$$

La potencia del error viene dada por

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

Dado que la señal x[n] es un proceso AR(1), la parte predecible de la muestra x[2n+1] puede obtenerse a partir de la muestra previa x[2n] = z[n] y cualquier otra muestra que se encuentre a mayor distancia de la muestra a predecir no aporta información adicional. De esta manera, el coeficiente asociado a la muestra más alejada a la muestra a predecir z[n-1] = x[2n-2] se anula mientras que el coeficiente asociado a la muestra más cercana a la muestra a predecir z[n] = x[2n] toma el valor del coeficiente de correlación.

La potencia del error de predicción es la potencia del ruido blanco generador ya que con un único coeficiente se consigue predecir perfectamente la señal AR(1).

3) Hallad ahora los coeficientes del predictor B que permite predecir las muestras impares de x(n) a partir de las muestras de z(n). Calculad la potencia del error de predicción que se obtiene con este predictor. Justificad los dos resultados, comparándolos con los obtenidos en el apartado (2).

Si ahora disponemos de las muestras pares previa y posterior a la muestra a predecir, el filtro de 2 coeficientes procesará las muestras z[n] (= x([2n]) y z[n+1] (= x[2n+2]) para obtener una estimación de x[2n+1]. Así, el vector de entrada al filtro (vector de datos x[n]) y la señal de referencia (d[n]) serán

$$\underline{x}^T[n] = \left[z[n+1], z[n]\right] = \left[x[2n+2], x[2n]\right] \qquad d[n] = x[2n+1] \quad .$$

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

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

$$\begin{bmatrix} 1 & \rho^2 \\ \rho^2 & 1 \end{bmatrix} \begin{bmatrix} h_0 \\ h_1 \end{bmatrix} = \begin{bmatrix} \rho \\ \rho \end{bmatrix}$$ y solucionando el sistema $$\boxed{\begin{bmatrix} h_0 \\ h_1 \end{bmatrix} = \frac{\rho}{1+\rho^2} \begin{bmatrix} 1 \\ 1 \end{bmatrix}}$$

La potencia del error viene dada por

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

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

En este caso, las dos muestras de que se dispone están a igual distancia de la muestra a predecir. Dado que la señal x[n] es un proceso AR(1), la parte predecible de la muestra x[2n+1] puede obtenerse a partir de la muestra previa x[2n] = z[n] y de la muestra posterior x[2n+2] = z[n+1] con igual precisión. De esta manera, los coeficientes de ambas muestras toman el mismo valor.

La potencia del error de predicción es ahora menor que la potencia del ruido blanco generador ya que al promediar dos muestras de ruido independientes se consigue reducir la potencia del error.

Si la solución de ambos predictores se ha obtenido mediante un sistema adaptativo que utiliza el algoritmo LMS:

4) Hallad para ambos casos el rango de valores del parámetro μ del sistema que asegura la convergencia a la solución de Wiener.

Para poder establecer el rango de valores del parámetro μ que asegura la convergencia del sistema, se debe hallar los autovalores de la matriz de correlación de los datos de entrada x[n] ya que

$$0 < \mu < \frac{2}{\lambda_{MAX}}$$

Debe destacarse que el ejercicio no pide una cota más restrictiva del rango de convergencia sino el rango propiamente y, por tanto, es necesario hallar estos autovalores.

En ambos casos se tiene la tiene la misma matriz de datos

$$R_x = \begin{bmatrix} r_x[0] & r_x[2] \\ r_x[2] & r_x[0] \end{bmatrix} = \begin{bmatrix} \sigma_x^2 & \sigma_x^2\rho^2 \\ \sigma_x^2\rho^2 & \sigma_x^2 \end{bmatrix}$$

Por tanto, los autovalores se pueden obtener resolviendo la ecuación:

$$|R_x - \lambda I| = \begin{vmatrix} \sigma_x^2 - \lambda & \sigma_x^2\rho^2 \\ \sigma_x^2\rho^2 & \sigma_x^2 - \lambda \end{vmatrix} = 0 \qquad \text{que da lugar a} \qquad \lambda = \sigma_x^2\left(1 \pm \rho^2\right)$$

Y, por tanto $\lambda_{MAX} = \sigma_x^2\left(1+\rho^2\right)$, con lo que se llega a $\boxed{0 < \mu < \frac{2}{\sigma_x^2(1+\rho^2)}}$

5) Dada la siguiente expresión para el desajuste

$$M \approx \mu \sum_i \lambda_i$$

determinad, suponiendo fijo el valor del parámetro μ, cómo varía el desajuste para ambos predictores si (i) varía el valor del coeficiente de correlación entre muestras consecutivas de x(n) manteniéndose igual su potencia y (ii) varía la potencia del proceso x(n) manteniéndose igual el valor de su coeficiente de correlación entre muestras consecutivas.

Dado que se tiene $M \approx \mu \sum_i \lambda_i$ se puede sustituir los autovalores hallados en la expresión y se llega a

$$M \approx \mu \sum_i \lambda_i = \mu\left[\sigma_x^2\left(1+\rho^2\right) + \sigma_x^2\left(1-\rho^2\right)\right] = 2\mu\sigma_x^2$$

i) Como se puede observar, si varía el coeficiente de correlación, manteniéndose constante la potencia de la señal, el desajuste no varía ya que éste no depende de la correlación entre muestras consecutivas.

ii) Por el contrario, el desajuste sí que es proporcional a la potencia de la señal y, aunque se mantenga constante el coeficiente de correlación, si la potencia varía, cambia la forma del paraboloide de la superficie de error y varía el desajuste.

Problema 3 (15%): Un inesperado accidente ha hecho desaparecer los datos originales de cómo se hizo una larga simulación sobre filtros adaptativos. La información que se ha podido salvar es la siguiente:

- La matriz de autocorrelación de la señal real x(n) de entrada era $\underline{\underline{R}}_x = \begin{bmatrix} 1.25 & 0.5 \\ 0.5 & 1.25 \end{bmatrix}$
- El filtro h era un filtro FIR de dos coeficientes.
- La salida del filtro h estaba perturbada por un ruido blanco gaussiano v(n).
- Se utilizó el algoritmo LMS con un filtro w de dos coeficientes con la siguiente ecuación de actualización del filtro:

  $$\underline{w}(n+1) = \underline{w}(n) + 2\mu e(n)\underline{x}(n)$$

- El valor del desajuste, estimado como $M \approx \mu \sum_i \lambda_i$, era 0,125.
- El diagrama de bloques del sistema simulado y la gráfica de error cuadrático promediado sobre 10,000 simulaciones eran los siguientes:

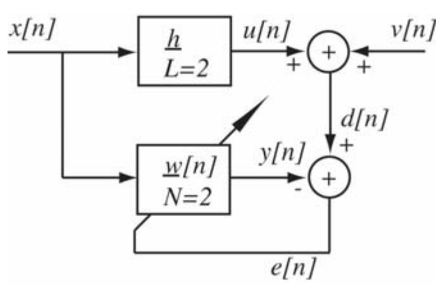
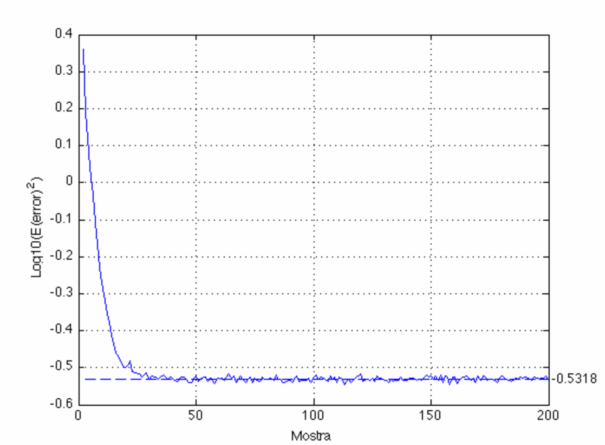


Diagrama de bloques



Gráfica de la evolución del error cuadrático promediado sobre 10,000 simulaciones

Para ayudar a resolver el problema de los datos perdidos, se pide:

1. ¿Cuál era el valor de μ utilizado en la simulación?

El valor de μ el podem obtenir de l'expressió: $M \approx \mu \sum_{i=1}^{2} \lambda_i = 0{,}125$ on $\sum_{i=1}^{2} \lambda_i = Tr\left(\underline{\underline{R}}_x\right) = 2{,}5$

per tant: $\mu = \dfrac{0{,}125}{2{,}5} = 0{,}05$

2. Dad el valor de la varianza σ_v^2 del ruido blanco v(n).

En l'algorisme LMS, malgrat tenir un error de desajust del 12,5%, tenim que el $\lim_{k\to\infty} E\{\underline{h}[k]\} = \underline{h}_{opt}$. El que podem observa en la gràfica és un a aproximació del valor esperat en el límit, pel promig de 10000 realitzacions d'adaptació (amb 200 mostres cadascuna). Per tant la zona plana de la gràfica correspon mes a la potència del soroll $v[n]$ que a l'error de desajust. Una fita de la potència màxima de la potència de $v[n]$ és: $\sigma_v^2 < 0{,}29$.

Nota: $\log_{10}(0.29) \approx -0.5318$

Problema 4 (15%): Se desea codificar vectores de la señal x de dimensión 2×1 de los que se sabe que su matriz de correlación es

$$\underline{\underline{R}}_x = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

a partir de la cuantificación de los coeficientes transformados obtenidos como $\underline{y} = [y_1, y_2]^T = \underline{A}^H \underline{x}$. Se pide:

a) Suponed que se utilizan cuantificadores uniformes i la matriz de transformación óptima. Cuál es la SNR sobre la señal decodificada, si se asignan 2 bits al elemento de y de más potencia y 1 bit al otro elemento?

L'error de quantificació en el nostre cas vindrà donat per:

$$J = \frac{1}{12}\sum_{k=1}^{N}\Delta_k^{\ 2} = C\left(\frac{pot(y_1)}{4^{B_1}} + \frac{pot(y_2)}{4^{B_2}}\right) = \frac{C}{16}(5-3\rho)$$ on N=2B1=2bits i B2=1bit

$pot(y_1) = 1+\rho$ i $pot(y_2) = 1-\rho$, que son els autovalors de la matriu $\underline{\underline{R}}_x$.

Es a dir: $\underline{\underline{R}}_y = \begin{vmatrix} 1+\rho & 0 \\ 0 & 1-\rho \end{vmatrix} = \underline{\underline{A}}^H \underline{\underline{R}}_x \underline{\underline{A}}$. Per tant: $\frac{S}{N} = \frac{2\, r_x[0]}{(5-3\rho)}\frac{16}{C} = \frac{32}{(5-3\rho)C}$

b) ¿Mejoraría la SNR de la señal recibida si antes de cuantificarla se hiciera predicción lineal sobre los coeficientes transformados, es decir, si se buscase un *a* óptimo que minimizase la potencia de $y_1 - a y_2$? Razonad la respuesta.

Si la matriu $\underline{\underline{R}}_y = \begin{vmatrix} 1+\rho & 0 \\ 0 & 1-\rho \end{vmatrix}$ és diagonal, vol dir que les mostres de y_1 i de y_2 estan incorrelades, per tant y_2 no conté cap informació amb la que predir y_1, pel que no es possible dissenyar cap predictor per aquest cas (*a*=0).