Solución Ejercicio 1

El presente ejercicio obtiene un algoritmo que permite obtener con bajo cálculo el autovector asociado al autovalor máximo de la matriz de correlación $\mathbf{R}_x = E\left[\mathbf{x}_n \mathbf{x}_n^H\right]$. Para ello se pide resolver los siguientes apartados.

a.- Considere la expresión (1), en donde $\mathbf{x}_n$ y $\mathbf{w}_n$ son vectores de Q componentes cada uno, y demuestre que la minimización que plantea da como resultado un autovector de $\mathbf{R}_x$. En dicha demostración halle el autovalor, λ, asociado e indique si se corresponde con λ_{max} o λ_{min}, justifíquelo.

$$\max_{\mathbf{w}} \mathbf{w}^H \mathbf{R}_x \mathbf{w} \qquad \qquad (1)$$
$$\textit{sujeto a } \mathbf{w}^H \mathbf{w} = 1$$

Solución:

La optimización se obtiene a través de $\dfrac{\partial L}{\partial \mathbf{w}^H} = \mathbf{0}$. Como $L = \mathbf{w}^H \mathbf{R}_x \mathbf{w} - l\,\mathbf{w}^H \mathbf{w}$ entonces $\mathbf{R}_x \mathbf{w} = l\,\mathbf{w}$

Para que cumpla $\mathbf{w}^H \mathbf{w} = 1 \Rightarrow l = \mathbf{w}^H \mathbf{R}_x \mathbf{w} = l_{\max}$

b.- Demuestre que se obtiene el mismo resultado si se plantea el problema como

$$\min_{\mathbf{w}} E\left[\left|\mathbf{x}_n - \mathbf{w}\mathbf{w}^H \mathbf{x}_n\right|^2\right] \qquad \qquad (2)$$
$$\textit{sujeto a} \quad \mathbf{w}^H \mathbf{w} = 1$$

NOTA: para $f = \left(\mathbf{a}^H \mathbf{b}\right)\left(\mathbf{a}^H \mathbf{c}\right) \rightarrow \dfrac{\partial f}{\partial \mathbf{a}^H} = \mathbf{b}\left(\mathbf{a}^H \mathbf{c}\right) + \left(\mathbf{a}^H \mathbf{b}\right)\mathbf{c}$

Solución:

En este caso $L = E\left[\left|\mathbf{x}_n\right|^2\right] - 2\mathbf{w}^H \mathbf{R}_x \mathbf{w} + \mathbf{w}^H \mathbf{R}_x \mathbf{w}\mathbf{w}^H \mathbf{w} - l\,(\mathbf{w}^H \mathbf{w} - 1)$

$$\frac{\partial L}{\partial \mathbf{w}^H} = \mathbf{0} \Rightarrow -2\mathbf{R}_x \mathbf{w} + \mathbf{R}_x \mathbf{w}\mathbf{w}^H \mathbf{w} + (\mathbf{w}^H \mathbf{R}_x \mathbf{w})\mathbf{w} = l\,\mathbf{w} \quad (*)$$

Para que cumpla $\mathbf{w}^H \mathbf{w} = 1 \Rightarrow l = -2\mathbf{w}^H \mathbf{R}_x \mathbf{w} + 2\mathbf{w}^H \mathbf{R}_x \mathbf{w} = 0$

Sustituyendo λ=0 en (*) se obtiene que $\mathbf{R}_x \mathbf{w} = (\mathbf{w}^H \mathbf{R}_x \mathbf{w})\mathbf{w}$ y coincide con el resultado del apartado anterior.

La minimización de (2) puede interpretarse como el diseño de un vector $\mathbf{w}$ que minimiza el error
$\mathbf{e} = \mathbf{x} - \mathbf{w}.\,y$, siendo $y = \mathbf{w}^H \mathbf{x}$
Con el objetivo de obtener un método de baja complejidad para calcular $\mathbf{w}$ (que no presente problemas numéricos cuando $\mathbf{R}_x$ se estima con bajo número de muestras), se propone obtener $\mathbf{w}_n$ iterativamente a través de (3)

$$\min_{\mathbf{w}_n} \left|\mathbf{w}_n - \mathbf{w}_{n-1}\right|^2 \qquad \qquad (3)$$
$$\textit{sujeto a } \mathbf{w}_n y_n = \mathbf{x}_n$$

con $y_n = \mathbf{w}_{n-1}^H \mathbf{x}_n$, ya que se considera que para señales de variación lenta o estacionarias, la diferencia entre $\mathbf{w}_n$ y $\mathbf{w}_{n-1}$ es pequeña. Al minimizar la diferencia entre vectores **w** de iteraciones sucesivas en (3), se busca que dicho diseño garantice la convergencia del algoritmo.

c.- Halle el vector $\mathbf{w}_n$ que verifica (3) y demuestre que se puede calcular como

$$\mathbf{w}_n = \mathbf{w}_{n-1} - m\left(\mathbf{w}_{n-1} y_n - \mathbf{x}_n\right) y_n^* \qquad (4)$$

dé el valor de μ

Solución:

$$L = \mathbf{w}_n^H \mathbf{w}_n + \mathbf{w}_{n-1}^H \mathbf{w}_{n-1} - \mathbf{w}_n^H \mathbf{w}_{n-1} - \mathbf{w}_{n-1}^H \mathbf{w}_n + (\mathbf{w}_n^H y_n^* - \mathbf{x}_n^H)?$$

$$\frac{\partial L}{\partial \mathbf{w}_n^H} = \mathbf{0} \Rightarrow \mathbf{w}_n - \mathbf{w}_{n-1} + ?y_n^* = \mathbf{0} \ \ (**)$$

Para que se cumpla $\mathbf{w}_n y_n = \mathbf{x}_n \Rightarrow l = \frac{1}{|y_n|^2}\left(\mathbf{w}_{n-1} y_n - \mathbf{x}_n\right)$

Sustituyendo en (**) se obtiene $\quad \mathbf{w}_n = \mathbf{w}_{n-1} - m\left(\mathbf{w}_{n-1} y_n - \mathbf{x}_n\right) y_n^*$ con $m = \frac{1}{|y_n|^2}$

d.- Compruebe que la solución óptima de (4) en régimen permanente es el autovector asociado al autovalor máximo de $\mathbf{R}_x$.

Solución:

$$E\left[\mathbf{w}_n - \mathbf{w}_{n-1}\right] \Rightarrow \mathbf{w}\left(\mathbf{w}^H \mathbf{R}_x \mathbf{w}\right) = \mathbf{R}_x \mathbf{w}$$

e.- Teniendo en cuenta que (4) pretende resolver (2) ¿es (4) un algoritmo de gradiente instantáneo? Justifique la respuesta.

Solución:

El gradiente de la potencia del error hallada en (2) es $\nabla x = -2\mathbf{R}_x \mathbf{w} + \mathbf{R}_x \mathbf{w}\mathbf{w}^H \mathbf{w} + (\mathbf{w}^H \mathbf{R}_x \mathbf{w})\mathbf{w}$

Su valor instantáneo es

$$\nabla x_{inst} = -\mathbf{x}_n \mathbf{x}_n^H \mathbf{w} + \left(\mathbf{w}^H \mathbf{x}_n \mathbf{x}_n^H \mathbf{w}\right)\mathbf{w} =$$
$$= \left(\mathbf{w}_{n-1} y_n - \mathbf{x}_n\right) y_n^*$$

Por lo tanto, (4) es un algoritmo de gradiente instantáneo

f.- Indique las similitudes y diferencias de (4) con un algoritmo NMLS. Para ello:

- Dé la expresión general del NMLS
- Compare dicha expresión general con (4) y dé la expresión del error **e**(n) y de los datos

En (4) tenemos un algoritmo del tipo LMS normalizado o NLMS:

$$w_n = w_{n-1} - \frac{a}{|\mathbf{y}_n|^2} \mathbf{y}_n e^*(n)$$

con la diferencia de que la señal de datos en (4) es un escalar y el error es un vector

Ejercicio 2

En este ejercicio se considera el diseño de un codificador de imagen óptimo. La imagen se define como un vector aleatorio **u** de dimensión Nx1 con media nula y matriz de covarianza **R**. El vector **u** se transforma con la transformación lineal definida por la matriz compleja **A** de dimensión NxN. Esta transformación da lugar al vector complejo **v** cuyas componentes están incorreladas entre sí. Cada componente del vector **v** se cuantifica independientemente. El vector cuantificado $\mathbf{v}'$ se transforma linealmente con la matriz **B** dando lugar al vector $\mathbf{u}'$.

$$\xrightarrow{\mathbf{u}} \boxed{\text{Tranformación lineal } \mathbf{A}} \xrightarrow{\mathbf{v}} \boxed{\text{Cuantificador}} \xrightarrow{\mathbf{v}'} \boxed{\text{Tranformación lineal } \mathbf{B}} \xrightarrow{\mathbf{u}'}$$

El objetivo es encontrar las matrices de transformación **A** y **B** que minimizan la distorsión, definida como el error cuadrático medio entre los vectores **u** y $\mathbf{u}'$.

$$D = \frac{1}{N} E\left\{ Tr\left[\left(\mathbf{u} - \mathbf{u}'\right)\left(\mathbf{u} - \mathbf{u}'\right)^T \right] \right\}$$

a) Defina la distorsión D en función de los vectores transformados **v** y $\mathbf{v}'$ y de las matrices de transformación **A** y **B**.

Solución: $$D = \frac{1}{N} E\left\{ Tr\left[\left(\mathbf{A}^{-1}\mathbf{v} - \mathbf{B}\mathbf{v}'\right)\left(\mathbf{A}^{-1}\mathbf{v} - \mathbf{B}\mathbf{v}'\right)^H \right] \right\}$$

b) Obtenga la matriz **B** que minimiza la distorsión D en función de la matriz **A** y de las matrices de correlación: $\mathbf{R}_{\mathbf{v}'\mathbf{v}'} = E\left[\mathbf{v}'\mathbf{v}'^H\right]$ y $\mathbf{R}_{\mathbf{v}\mathbf{v}'} = E\left[\mathbf{v}\mathbf{v}'^H\right]$.

Solución: $$\mathbf{B} = \mathbf{A}^{-1}\mathbf{R}_{\mathbf{v}\mathbf{v}'}\mathbf{R}_{\mathbf{v}'\mathbf{v}'}^{-1}$$

El cuantificador óptimo que minimiza la potencia del ruido de cuantificación genera una señal cuantificada ortogonal al error de cuantificación: $E\left[\left(\mathbf{v}' - \mathbf{v}\right)\mathbf{v}'^H\right] = \mathbf{0}$.

c) Considerando el cuantificador óptimo, demuestre que la distorsión mínima puede definirse como:

$$D = \frac{1}{N} Tr\left[\mathbf{A}^{-1}\mathbf{F}\mathbf{A}\mathbf{R}\right] \quad siendo \quad \mathbf{F} = diag\left\{ f\left(b_k\right) \right\}$$

donde $f\left(b_k\right)$ es la distorsión de un cuantificador de b_k bits cuando la varianza de la entrada es la unidad, siendo el error de cuantificación:

$$\sigma_q^2\left(k\right) = E\left[\left|v\left(k\right) - v'\left(k\right)\right|^2\right] = E\left[\left|v\left(k\right)\right|^2\right] f\left(b_k\right)$$

Solución: $$\mathbf{R}_{\mathbf{v}\mathbf{v}'}\mathbf{R}_{\mathbf{v}'\mathbf{v}'}^{-1} = \mathbf{I} \Rightarrow \mathbf{B} = \mathbf{A}^{-1}$$

$$D = \frac{1}{N} Tr\left[\mathbf{A}^{-1}\mathbf{F}\mathbf{R}_{\mathbf{v}\mathbf{v}}\left(\mathbf{A}^{-1}\right)^H\right]$$

$$\mathbf{R}_{\mathbf{v}\mathbf{v}} = \mathbf{A}\mathbf{R}\mathbf{A}^H$$

d) Minimice la distorsión con respecto a la transformación **A** y demuestre que la transformación óptima corresponde a la transformada KL.

Solución: $$\frac{\partial D}{\partial \mathbf{A}} = -\frac{1}{N}\left(\mathbf{A}^{-1}\mathbf{F}\mathbf{A}\mathbf{R}\mathbf{A}^{-1}\right)^T + \frac{1}{N}\left(\mathbf{R}\mathbf{A}^{-1}\mathbf{F}\right)^T \Rightarrow \mathbf{F}\left(\mathbf{A}\mathbf{R}\mathbf{A}^{-1}\right) = \left(\mathbf{A}\mathbf{R}\mathbf{A}^{-1}\right)\mathbf{F}$$

$\mathbf{F}$ diagonal $\Rightarrow \mathbf{A}\mathbf{R}\mathbf{A}^{-1}$ diagonal

$\mathbf{A}\mathbf{R}\mathbf{A}^{H}$ diagonal $\Rightarrow \mathbf{A}\mathbf{A}^{H} = \mathbf{I}$

A autovectores de **R**

Utilizando una transformación arbitraria **A** y definiendo $\mathbf{B} = \mathbf{A}^{-1}$, la distorsión puede definirse como:

$$D = \frac{1}{N}\sum_{k=0}^{N-1} E\left[\left|v(k) - v'(k)\right|^2\right] = \frac{1}{N}\sum_{k=0}^{N-1}\sigma_k^2 f\left(b_k\right)$$

donde σ_k^2 es la varianza del coeficiente transformado v(k).

Considere la función de distorsión del cuantificador en función del número de bits $f\left(b\right) = 2^{-2b}$.

e) Obtenga el número de bits necesario para cuantificar los coeficientes transformados asumiendo un valor dado de distorsión D.

Solución: $$M = \sum_{k=0}^{N-1}\frac{1}{2}\log_2\frac{\sigma_k^2}{D} = \frac{1}{2}\log_2\left(\prod_{k=0}^{N-1}\sigma_k^2\right) - \frac{N}{2}\log_2 D$$

f) Demuestre que el número de bits necesario es mínimo cuando **A** es la transformación KL.

Solución: $$\prod_{k=0}^{N-1}\sigma_k^2 = \det\left(\mathbf{R}\right)$$

$$M_{KL} = \frac{1}{2}\log_2\left(\det\left(\mathbf{R}\right)\right) - \frac{N}{2}\log_2 D$$

$$M - M_{KL} = \frac{1}{2}\log_2\left(\frac{\prod_{k=0}^{N-1}\sigma_k^2}{\det\left(\mathbf{R}\right)}\right) \geq 0$$

NOTAS: $Tr\left[\mathbf{XY}\right] = Tr\left[\mathbf{YX}\right]$ $\qquad \dfrac{\partial\left[Tr\left[\mathbf{XY}\right]\right]}{\partial \mathbf{X}} = \mathbf{Y}^T$

$$\frac{\partial Tr\left[\mathbf{X}^{-1}\mathbf{YXZ}\right]}{\partial \mathbf{X}} = -\left(\mathbf{X}^{-1}\mathbf{YXZX}^{-1}\right)^T + \left(\mathbf{ZX}^{-1}\mathbf{Y}\right)^T$$

Si $\mathbf{XY} = \mathbf{YX}$ y **X** es diagonal, **Y** es también diagonal

$\det\left(\mathbf{X}\right) \leq \prod_{k=1}^{N} x\left(k,k\right)$, donde $x\left(k,k\right)$ es el elemento de la fila k y columna k de la matriz **X.**