Departamento de Teoría de la Señal y Comunicaciones **27/06/07**
UCIIIM

TRATAMIENTO DIGITAL DE SEÑALES

TEORÍA

(Tiempo: 70 minutos. Puntos: 3/8)

T1.- La densidad de probabilidad tipo Erlang de orden N viene dada por la expresión

$$p(x) = \frac{a^N x^{N-1} \exp(-ax)}{(N-1)!} \qquad x > 0 \quad y \quad a > 0$$

Supóngase que N es conocida. Considerando que la media viene dada por m = N/a, determínense:

a) El estimador ML de su media, $\hat{m}_{ML}$, a partir de K observaciones independientes de la variable .

b) El sesgo de $\hat{m}_{ML}$.

c) ¿Es $\hat{m}_{ML}$ un estimador consistente en varianza?

(20 min; 1p)

T2.-

a) Los estimadores no paramétricos: histograma, k-vecinos más próximos y ventanas de Parzen con núcleo rectangular; pueden expresarse de la forma

$$\hat{p}(x) = \frac{k}{KV_R}$$

siendo K el número total de muestras y k el número de muestras contenido en el volumen V_R. Pese a admitir una expresión común, son estimadores diferentes. Explíquese por qué.

b) Compruebe si un estimador 1-NN cumple o no la propiedad

$$\int_{-\infty}^{\infty} \hat{p}(x)\,dx = 1$$

considerando el caso de una única muestra disponible $x^{(1)}$.

(20 min; 1 p)

T3.- Considérese el algoritmo regularizado de Newton caracterizado por la siguiente regla de actualización:

$$\mathbf{w}^{(k+1)} = \mathbf{w}^{(k)} - \left[\varepsilon\mathbf{I} + \mathbf{H}_w\left(\mathbf{w}^{(k)}\right)\right]^{-1}\mathbf{g}_w\left(\mathbf{w}^{(k)}\right) \tag{1}$$

Análogamente al caso del algoritmo secuencial de gradiente, es posible definir un algoritmo secuencial de Newton, cuya formulación se obtiene de sustituir en (1) aproximaciones del Hessiano y del gradiente, $\mathbf{H}_w(w^{(k)})$ y $\mathbf{g}_w(w^{(k)})$ respectivamente, calculadas únicamente a partir de la muestra k-ésima del conjunto de entrenamiento. Para una red que implementa su salida como $y^{(k)} = w_0 + \mathbf{w}^T\mathbf{x}^{(k)} = \mathbf{w}_e^T\mathbf{x}_e^{(k)}$, y suponiendo coste cuadrático:

a) Demuéstrese que la regla de actualización del algoritmo secuencial de Newton es:

$$\mathbf{w}_e^{(k+1)} = \mathbf{w}_e^{(k)} + (d^{(k)} - y^{(k)})\left(\varepsilon\mathbf{I} + \mathbf{x}_e^{(k)}\mathbf{x}_e^{(k)T}\right)^{-1}\mathbf{x}_e^{(k)} \tag{2}$$

b) Aplicando el siguiente lema de inversión (identidad de Woodbury):

$$\left(\mathbf{A} + \mathbf{c}\mathbf{c}^T\right)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{c}\mathbf{c}^T\mathbf{A}^{-1}}{1 + \mathbf{c}^T\mathbf{A}^{-1}\mathbf{c}}$$

siendo **c** un vector columna y **A** una matriz de dimensiones adecuadas, muéstrese que dicha regla de actualización es equivalente al algoritmo NLMS.

c) ¿Qué ventajas de índole práctica ofrece el algoritmo NLMS frente al empleo directo de la expresión (2)? ¿Y frente al LMS sin normalización?

(30 min; 1 p)

TRATAMIENTO DIGITAL DE SEÑALES

PROBLEMAS

(Tiempo: 150 minutos. Puntos: 5/8)

P1.- Un analista construye un estimador lineal de una v.a. s en función de otra observable x a partir de un conjunto de datos etiquetados $\{s^{(k)}, x^{(k)}\}$, recurriendo a sustituir en la formulación correspondiente al estimador de mínimo error cuadrático medio los estadísticos reales ($E\{s\}, E\{x\}, v_{xx}, v_{sx}, v_{ss}$) por sus estimaciones muestrales ($\overline{s}, \overline{x}, \overline{v}_{xx}, \overline{v}_{sx}, \overline{v}_{ss}$, respectivamente), que se consideran constantes. El analista no sabe que el par (s,x) sigue una distribución

$$G\left\{\mathbf{0}, \begin{bmatrix} v & \rho v \\ \rho v & v \end{bmatrix}\right\}.$$

a) Establézcase el estimador que calcula el analista, $\hat{s}_a(x)$, y compárese su error cuadrático medio teórico (que es el que resultará de su repetida aplicación a nuevas observaciones) con el que proporcionaría el estimador lineal de error cuadrático medio mínimo (estimador óptimo teórico), $\hat{s}_t(x)$.

b) En un momento posterior el analista percibe, por consideraciones físicas, que s y x tienen medias nulas; y diseña un nuevo estimador, $\hat{s}'_a(x)$, con los valores anteriores de las covarianzas muestrales y medias nulas. ¿En qué difiere $\hat{s}'_a(x)$ de $\hat{s}_a(x)$?

c) Calcúlese la ventaja (reducción de error cuadrático medio) que proporciona $\hat{s}'_a(x)$ sobre $\hat{s}_a(x)$.

d) ¿Podría el analista percibir la ventaja calculada en el apartado c) si sustituyese en las expresiones de los errores cuadráticos medios los estadísticos reales por sus valores estimados?

(75 min; 2.5 p)

P2.- Considérense las v.a. x_0 y x_1, independientes e idénticamente distribuidas con ddp p(x) continua de media m y mediana M. Tras observar el valor de x_0, y sin conocer x_1, debe decidirse entre las hipótesis:

$$H_0: x_0 > x_1$$
$$H_1: x_1 \geq x_0$$

a) Diséñese el decisor de mínima probabilidad de error (MAP).

b) Determínese la probabilidad de error del decisor obtenido en el apartado anterior.

c) Supóngase ahora que los costes asociados a las decisiones son

$$C_{00}=C_{01}=x_1$$
$$C_{10}=C_{11}=x_0$$

(obsérvese que los costes no dependen de la hipótesis, pero sí de las observaciones). Determínese el decisor de mínimo coste medio (a x_0 dado).

d) Determínese una expresión para el coste medio global del decisor del apartado anterior.

(75 min; 2.5 p)