TRATAMIENTO DIGITAL DE SEÑALES

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

T1.- Un problema unidimensional binario de hipótesis equiprobables presenta las siguientes verosimilitudes:

$$p(x|H_0)=\begin{cases} 1/4 & , 0<x<4 \\ 0 & , \text{en otro caso} \end{cases}$$

$$p(x|H_1)=\begin{cases} 1/2 & , 3<x<5 \\ 0 & , \text{en otro caso} \end{cases}$$

a) Establézcase el decisor de mínima probabilidad de error y calcúlense las características P_{FA}, P_M y P_e.

b) Un diseñador no conoce dichas verosimilitudes, pero le resulta posible tomar tres valores de x bajo cada una de las hipótesis, que resultan ser

- bajo H_0: $x_0^{(1)}$ =1’56, $x_0^{(2)}$ =2’64, $x_0^{(3)}$ =3’30

- bajo H_1: $x_1^{(1)}$ =3’16, $x_1^{(2)}$ =4’40, $x_1^{(3)}$ =4’73

El diseñador aplica un decisor tipo 1-NN para clasificar las muestras posteriores.

Indíquense los márgenes de valores de x para los que el esquema 1-NN llevará a tomar las decisiones D_0 y D_1.

c) A partir de lo anterior, calcúlense las características P'_{FA}, P'_M, P'_e para el decisor 1-NN, y compárense los resultados con los del decisor óptimo calculado en a).

d) Explíquese lo que ocurriría si el diseñador pudiese tomar un número muy elevado de muestras etiquetadas para utilizarlas con el esquema 1-NN, dando aproximaciones para P''_{FA}, P''_M, P''_e.

(30 min; 1p)

T2.- Se desea diseñar un modelo de regresión a partir de un conjunto de datos de entrenamiento $\{\mathbf{x}^{(k)}, s^{(k)}\}$ que contiene K pares, con dim(s)=1 y dim($\mathbf{x}$)=100. Sin embargo, se sabe que, para el problema bajo análisis, algunas de las variables del espacio de entrada no contienen información relevante, y que su uso daría lugar a un modelo con una pobre capacidad de generalización.

Indíquese cómo solucionaría el problema mediante un procedimiento de selección de variables, utilizando una búsqueda basada en algoritmos genéticos; en concreto, explique cómo realizaría la codificación binaria de las posibles soluciones, y discuta cómo implementaría la función de ajuste ("fitness") que permita determinar la capacidad de generalización de cada individuo de la población.

(20 min; 1 p)

T3.- Dadas las siguientes funciones de coste:

$$C_1(s,\hat{s}) = |s-\hat{s}+a_1| + (s-\hat{s}-b_1)^2$$

$$C_2(s,\hat{s}) = a_2(s-\hat{s})^2 + b_2 \log(s-\hat{s}), \text{ con } a_2, b_2 > 0$$

Indíquese si existen valores (indicando cuáles son en caso afirmativo) a_i, b_i, con i=1,2, que hacen que se verifique $\hat{s}_i = \hat{s}_{ms}$, siendo $\hat{s}_i$ el estimador óptimo asociado al coste C_i, i=1,2, y $\hat{s}_{ms}$ el estimador de error cuadrático medio mínimo, cuando:

a) $p(s\,|\,\mathbf{x})$ es simétrica respecto a su media.

b) Para cualquier $p(s\,|\,\mathbf{x})$.

(20 min; 1 p)

TRATAMIENTO DIGITAL DE SEÑALES

PROBLEMAS

(Tiempo: 180 minutos. Puntos: 5/8)

P1.- Un sistema genera dos observaciones, x_1 y x_2, que, tanto bajo hipótesis H_0 como H_1, son independientes e idénticamente distribuidas, siendo

$$p(x_i|H_1) = 2x_i, \qquad 0 < x_i < 1$$

$$p(x_i|H_0) = 2(1-x_i), \qquad 0 < x_i < 1$$

para i=1,2. Suponga hipótesis equiprobables.

a) Determine el decisor MAP basado en x_1 y calcule la probabilidad de error dada la observación (es decir Pr{error | x_1}), y la probabilidad de error global (es decir Pr{error}).

b) Sea DMAP1 el decisor del apartado a), y $g(x_1)$ = Pr{error | x_1} su probabilidad de error condicional. Suponga que, si $g(x_1)>a$ (siendo $0<a<0.5$), se observa x_2 y, con objeto de seguir aplicando decisión por umbral, se descartan x_1 y la decisión de DMAP1. En su lugar, se aplica un segundo decisor, basado en x_2 y también MAP, que llamaremos DMAP2.

   Represente gráficamente sobre el plano x_1-x_2, para un valor de a arbitrario, las regiones de decisión del esquema conjunto DMAP1-DMAP2.

c) Determine la probabilidad de error global del esquema conjunto DMAP1-DMAP2.

d) Determine la máxima reducción de la probabilidad de error global que puede conseguirse utilizando el esquema conjunto, respecto al decisor DMAP1.

e) Compare las prestaciones del decisor conjunto DMAP1-DMAP2 con las del decisor MAP que utiliza simultáneamente x_1 y x_2

(75 min; 2.5 p)

P2.- Las vv.aa. s y x obedecen a la ddp conjunta

$$p(s,x) = c, \quad \begin{cases} 0 < s < 1 \\ s < x < 2s \end{cases}$$

siendo c una constante.

a) Tras representar el soporte de la ddp, calcúlese el valor de c.

b) Establézcanse las expresiones de las ddp marginales p(s) y p(x).

c) Determínese analíticamente la forma del estimador de error cuadrático medio mínimo de s a la vista de x, $\hat{s}_{ms}(x)$. Trácese dicha forma sobre la representación del soporte de p(s,x), y discútase si es posible determinarla directamente.

d) Calcúlese el error cuadrático medio $E[e^2_{ms}(x)]$ que proporciona la aplicación del estimador anterior.

e) Determínese la forma del estimador lineal de error cuadrático medio de s a la vista de x, $\hat{s}_{lms}(x)$. Trácese dicha forma sobre la representación del soporte de p(s,x), y coméntese su aspecto.

f) Calcúlese el error cuadrático medio $E[e^2_{lms}(x)]$ que proporciona la aplicación del estimador anterior, y compárese con $E[e^2_{ms}(x)]$.

g) ¿Qué ocurre si se percibe (p.ej., visualizando muestras) que hay un comportamiento (estadístico) distinto para 0<x<1 y 1<x<2, y se diseña un estimador lineal óptimo diferente para cada uno de estos intervalos? ($\hat{s}_{Alms}(x)$ y $\hat{s}_{Blms}(x)$, respectivamente). Verifíquese analíticamente la solución que se proponga.

(105 min; 2.5 p)