Departamento de Teoría de la Señal y Comunicaciones **18/06/05**
UCIIIM

DIGITAL SIGNAL PROCESSING

THEORY
(Time: 60 minutes. Marks: 3/8)

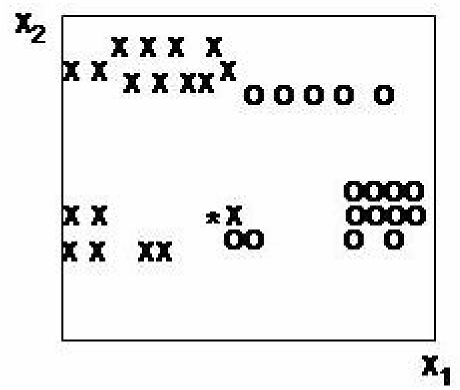
T1.- The k-Nearest Neighbors probability density estimation algorithm (k-NN) can be used for the estimation of a posteriori probabilities. Hence, it can be used for classification problems.

a) Starting from the non-parametric expression for probability density estimation, show that the expression for a posteriori probability estimation is:

$$\hat{\Pr}(C_c|x) = \frac{k_c}{k}$$

where k_c is the number of samples which belong to class c and k is the number of neighbors.

b) Determine the result obtained when classifying through k-NN the test sample shown in the figure, for k = 1, 3 and 5. Considering that the test sample actually belongs to class 1, which conclusions can you take based on the variation of the result with k ?



Class 0: **x**, Class 1: **o**, Test sample: *****

(20 min; 1p)

T2.- A r.v. s takes positive values. In order to perform bayesian estimation, the following cost function is considered:

$$C(s,\hat{s}) = a\,|s-\hat{s}|^k, \quad k > 0$$

Which conditions should a and k satisfy so that $\hat{s} = \hat{s}_{ms}$?

(20 min; 1 p)

T3.- Consider the following non-linear function:

$$f_\alpha(z) = \mathrm{sign}(z)\,|z|^\alpha; \quad \alpha > 0$$

a) Draw $f_\alpha(z)$ for 3 values of α which generate 3 qualitatively different curves.
b) Obtain the expression for a perceptron with two input variables, a hidden layer with 4 neurons and one output that uses $f_\alpha(z)$ as activation function.
c) Derive the update algorithm for each one of the weights of the network described in b), following the steepest descent rule for square cost. Find also an expression for the update of the α parameter of each neuron.

(20 min; 1 p)

DIGITAL SIGNAL PROCESSING

PROBLEMS

(Time: 135 minutes. Points: 5/8)

P1.- Our goal is to classify objects that belong to one out of two classes H_0 and H_1, through the observation of two of its dimensions, x_1 and x_2. x_1 and x_2 are statistically independent between them and their likelihoods are given by:

$$p(x_1 \mid H_0) = \begin{cases} 1 - x_1/2, & 0 < x_1 < 2 \\ 0, & \text{otherwise} \end{cases} \qquad p(x_1 \mid H_1) = \begin{cases} x_1/2, & 0 < x_1 < 2 \\ 0, & \text{otherwise} \end{cases}$$

$$p(x_2 \mid H_0) = \begin{cases} 1, & 0 < x_2 < 1 \\ 0, & \text{otherwise} \end{cases} \qquad p(x_2 \mid H_1) = \begin{cases} 1, & 1 < x_2 < 2 \\ 0, & \text{otherwise} \end{cases}$$

The classes are equiprobable and the costs related to the 2 kinds of classification errors are unitary. On the other hand, the cost of measuring x_1 is negligible while the cost of measuring x_2 is C_M.

a) Determine the average cost C_B obtained when applying the bayesian clasisfier on x_1, x_2.

b) With the aim of reducing the cost calculated in a), a sequential clasification procedure is going to be used: classify according to x_1 if the decision is clear enough, or measure x_2 and classify according to both x_1 and x_2. Given the symmetry shown by the likelihoods of x_1 around the decision threshold, the process is halted in the first step if the distance between x_1 and this threshold is larger than γ ($0< \gamma <1$).

- b1) Determine the decision threshold for the first step.
- b2) Determine the probability of taking a decision in just one step, Pr(1).
- b3) Considering the probability calculated in b2) and the average costs related to deciding in one and two steps, determine the average cost of the sequential classification, C_S.
- b4) Comment about the best selection of the γ parameter. Is there any limitation on the C_M value ?

(75 min; 2.5 p)

P2.- Consider the r.v. x given by:

$$x = s\,t$$

where s and t are independent variables with distributions given by:

$$p_s(s) = (m+1)s^m \qquad 0 \le s \le 1$$
$$p_t(t) = 1 \qquad 0 \le t \le 1$$

where $m \ge 0$.

1. Determine p(x).
2. Determine the MMSE estimator of s based on the observation of x, $\hat{s}_{ms}$.
3. Determine the MAP esimator of s based on the observation of x, $\hat{s}_{map}$.
4. Determine the bias of the MAP estimator.

(Notice that in some expressions the cases m=0 and m>0 must be treated separately.)

(60 min; 2.5 p)