DIGITAL SIGNAL PROCESSING

THEORY
(Time: 70 minutes. Mark: 3/8)

T1.- A one dimensional binary decision problem with equally probable hypotheses is characterized by the following likelihoods:

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

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

a) Provide expressions for the decider with a minimum error probability, and the value of the following probabilities: P_{FA}, P_M and P_e.

b) A designer does not know the true likelihoods, but he can take three values of x under each of the hypotheses, obtaining the following observations:

- given H_0: $x_0^{(1)}$ =1'56, $x_0^{(2)}$ =2'64, $x_0^{(3)}$ =3'30
- given H_1: $x_1^{(1)}$ =3'16, $x_1^{(2)}$ =4'40, $x_1^{(3)}$ =4'73

The designer applies a 1-NN decider to classify subsequent samples.

Obtain the intervals of x for which the 1-NN scheme will lead to decisions D_0 and D_1.

c) From the previous results, compute the new probabilities P'_{FA}, P'_M, and P'_e for the 1-NN decider, and compare these results with those achieved in Section a).

d) Explain what would happen if the designer could take a very large number of labeled samples to work with the 1-NN scheme, providing approximations for P''_{FA}, P''_M, and P''_e.

(30 min; 1p)

T2.- Consider the problem of designing a regression model using a set of K labeled training samples $\{\mathbf{x}^{(k)}, s^{(k)}\}$, with dim(s)=1 and dim($\mathbf{x}$)=100. It is known that, for the problem under study, some of the input variables do not contain any relevant information, and their use would lead to a regression model with poor generalization capabilities.

Explain how this problem could be solved with a variable selection approach based on genetic algorithms; in particular, explain how the binary encoding of the solutions could be carried out, and discuss a plausible implementation of the fitness function that would allow to determine the generalization capability of each individual of the population.

(20 min; 1 p)

T3.- Given the following cost functions:

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

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

Please, provide, if possible, the values for $\{a_i, b_i\}$, with i=1,2, that make $\hat{s}_i = \hat{s}_{ms}$, $\hat{s}_i$ being the optimum estimator associated to cost C_i, i=1,2, and $\hat{s}_{ms}$ being the minimum mean square error estimator. Consider for your answers the two following cases:

a) $p(s\,|\,\mathbf{x})$ is symmetric with respect to its mean.

b) $p(s\,|\,\mathbf{x})$ is any pdf.

(20 min; 1 p)

DIGITAL SIGNAL PROCESSING

PROBLEMS

(Time: 180 minutes. Mark: 5/8)

P1.- A system provides two observations, x_1 and x_2, that under both hypotheses, H_0 and H_1, are independent and identically distributed, following

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

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

for i=1,2. Assume that both hypotheses are equally probable.

a) Obtain the MAP decider based on x_1 and give the error probability for a given observation (i.e., Pr{error | x_1}), and the unconditional probability of error (P_e = Pr{error}).

b) Let DMAP1 be the classifier obtained in section a), and $g(x_1)$ = Pr{error | x_1} its conditional probability of error. Assume that, when $g(x_1)>a$ (with $0<a<0.5$), variable x_2 is observed and, to keep on using decisions based on thresholded variables only, observation x_1 and the decision of DMAP1 are discarded. In this case, a new decider based on x_2 is employed. The new classifier, that uses also the MAP criterion, will be denoted as DMAP2.
Sketch on the x_1-x_2 plane the decision regions for the combined scheme DMAP1-DMAP2 for arbitrary a.

c) Calculate the overall probability of error of the joint scheme DMAP1-DMAP2.

d) Which is the maximum reduction of the error probability that can be achieved by using the joint scheme, with respect to the case in which decider DMAP1 is used alone?

e) Compare the performance of the combined decider DMAP1-DMAP2 with respect to the MAP classifier that uses simultaneously observations x_1 and x_2

(75 min; 2.5 p)

P2.- Consider two random variables, s and x, following the joint pdf

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

for c a constant.

a) After representing the support of the pdf, compute the value of c.

b) Obtain the expressions of the marginal pdfs of s and x, p(s) and p(x).

c) Derive analytically the minimum mean square error estimator of s as a function of x, $\hat{s}_{ms}(x)$. Depict the function of this estimator over the representation of the support of p(s,x), and discuss if the estimator could have been determined directly.

d) Calculate the mean square error $E\{e^2_{ms}(x)\}$ provided by the previous estimator.

e) Obtain the linear minimum mean square error estimator of s as a function of x, $\hat{s}_{lms}(x)$. Depict this new function over the support of p(s,x), discussing the results.

f) Obtain the mean square error of the linear estimator from e), $E\{e^2_{lms}(x)\}$, comparing this result with $E\{e^2_{ms}(x)\}$.

g) What would happen if a designer perceives (for instance, by analyzing a number of samples) that there exists (statistically) different behaviors for intervals 0<x<1 and 1<x<2, and designs a different optimum linear estimator for each of these intervals ($\hat{s}_{Alms}(x)$ and $\hat{s}_{Blms}(x)$, respectively)? Verify analytically the proposed solution.

(105 min; 2.5 p)