E1 213 Pattern Recognition and Neural Networks

Practice Problems: Set 1

1. Consider a 2-class problem with one dimensional feature space. Let the class conditional densities be: $f_0(x) = e^{-x}, \; x > 0$, and $f_1(x) = 1/2a, \; x \in [-a, \; a], \; a > 0$. The prior probabilities are equal. Assume we are using 0–1 loss. Find the Bayes classifier. For the case when $a = 0.25$, find Bayes error.

Answer: Since we are using 0–1 loss and since prior probabilities are equal, the Bayes classifier is: $h_B(x) = 0$ if $f_0(x) > f_1(x)$ and $h_B(x) = 1$ otherwise.

Since $f_0(x) = 0$ for $x < 0$, we have $h_B(x) = 1$ for $x < 0$. (Actaully, we need not worry about the region $x < -a$ because a pattern coming from that region has probability zero. But since we want to think of h_B as a function on $\Re$, we can assign class-1 in that region).

Now consider the region $x \geq 0$. If $a \leq 0.5$ then $(1/2a) \geq 1$ and hence $f_1(x) \geq f_0(x)$ for $0 \leq x \leq a$ and there after $f_1(x) = 0$. Hence till a we classify into class-1 and beyond that we classify into class-0.

If $a > 0.5$ then $f_0(x) > f_1(x)$ if $e^{-x} > (1/2a)$. This is true if $x < \ln(2a)$.

Putting all this together, the Bayes classifier for this problem is the following:

- If $a \leq 0.5$ then

$$h_B(x) = \begin{cases} 1 & \text{if } x \leq a \\ 0 & \text{if } x > a \end{cases}$$

- If $a > 0.5$ then

$$h_B(x) = \begin{cases} 1 & \text{if } x \leq 0 \\ 0 & \text{if } 0 < x < \ln(2a) \\ 1 & \text{if } \ln(2a) \leq x \leq a \\ 0 & \text{if } x > a \end{cases}$$

For $a = 0.25$, we classify into class-1 till 0.25 and into class-0 after that. Hence, we make an error if we get a class-0 pattern with $x \leq 0.25$. Hence, bayes error is

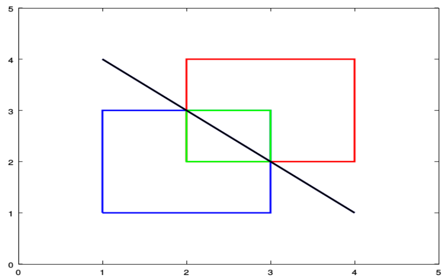
$$0.5 \int_0^{0.25} e^{-x} \, dx = 0.5(1 - e^{-0.25}) \approx 0.11$$

2. Consider a 2-class PR problem with feature vectors in $\Re^2$. The class conditional density for class-I is uniform over $[1,3] \times [1,3]$ and that for class-II is uniform over $[2,4] \times [2,4]$. Suppose the prior probabilities are equal and we are using 0–1 loss. Consider line given by $x + y = 5$ in $\Re^2$. Is this a Bayes Classifier for this problem? Is Bayes Classifier unique for this problem? If not, can you specify two different Bayes classifiers? Suppose the class conditional densities are changed so that the density for class-I is still uniform over $[1,3] \times [1,3]$ but that for class-II is uniform over $[2,5] \times [2,5]$. Is the line $x + y = 5$ a Bayes classifier now? If not, specify a Bayes classifier now. Is the Bayes classifier unique now? For this case of class conditional densities, suppose that wrongly classifying a pattern into class-I is 10 times more expensive than wrongly classifying a pattern into class-II. Now, what would be a Bayes classifier?

Answer: Consider the first case. With equal priors and 0–1 loss, the decision of Bayes classifier is simply based on which class conditional density has a higher value.

It is easy to see that in $([1,3] \times [1,3]) - ([2,3] \times [2,3])$ Bayes classifier would assign Class-I because the other class conditional density is zero. Similarly in $([2,4] \times [2,4]) - ([2,3] \times [2,3])$ the decison is Class-II.

The situation is as shown in the figure below.



The only thing remaining to be decided is what to do in the overlapping region shown as a green rectangle in the figure. In this regions, both class conditional densities have the same value and hence it does not matter which class you assign. (This is like having an arbitrary rule to 'break ties' while deriving Bayes rule).

Thus, the Bayes classifier here is not unique. For example, we can assign all points in the green rectangle in the figure to Class-I or Class-II thus giving two different Bayes classifiers.

The line $x + y = 5$ shown in the figure is also a Bayes classifier here. It assigns half the points in the green rectangle to one class and the other half to the other class.

Now consider the case where class-II is uniform over $[2, 5] \times [2, 5]$. Now at all these points the value of class conditional density is $1/9$. Hence, in the common region now, the density of class-I has higher value (namely, $(1/4)$). Hence, all points in the common region have to be assigned to class-I. Thus, the line is no longer a Bayes classifier. Also, the Bayes classifier is unique here (assuming the domain, $\mathcal{X}$, to be $([1, 3] \times [1, 3]) \cup ([2, 5] \times [2, 5])$).

Now consider the case where we are not using 0–1 loss and wrogly classifying into class-I is 10 times costlier than wrongly classifying into class-II. Since priors are equal, this means we will put something into class-I only if the value (at that point) of $f_I(x)$ is 10 times that of $f_{II}(x)$. In the common region, the ratio is (9/4). Hence, now the Bayes classifier would put all points in the common region in Class-II.

3. Consider a 2-class problem with one dimensional feature vector and class conditional densities being normal. Specify a simple special case where Bayes classifier, Min-max classifier and Neyman-Pearson classifier would all be the same.

Answer: To specify a special case, we can assume what we need about class conditioal densities, loss function, and, for NP classifier, on the bound on type-I error.

Since we are given that the class conditional densities are normal, if we assume equal variances and 0–1 loss, then we know Bayes and minmax classifiers are same. This is a single threshold based classifier. So, if we take the type-I error of Bayes classifier as the value for α in NP classifier, all the three classifiers are same. Thus the special case is the following.

We assume the two class conditional densities have the same variance. We use 0–1 loss. We take $\alpha = \int_{(\mu_0+\mu_1)/2}^{\infty} f_0(x)\, dx$.