SML

January 2020

Assignment 1

Q1. Consider the following decision rule for a two-category one-dimensional problem:
Decide ω_1 if $x > \theta$; otherwise decide ω_2.
(a) Show that the probability of error for this rule is given by

$$P(error) = P(\omega_1)\int_{-\inf}^{\theta} p(x|\omega_1)dx + P(\omega_2)\int_{\theta}^{\inf} p(x|\omega_2)dx \tag{1}$$

(b) By differentiating, show that a necessary condition to minimize P(error) is that θ satisfy $p(\theta|\omega_1)P(\omega_1) = p(\theta|\omega_2)P(\omega_2)$

Q2. Let the conditional densities for a two-category one-dimensional problem be given by the Cauchy distribution

$$p(x|\omega_i) = \frac{1}{\pi b}\frac{1}{1+(\frac{x-a_i}{b})^2}, \quad i = 1, 2 \tag{2}$$

Assuming $P(\omega_1) = P(\omega_2)$, show that $P(\omega_1|x) = P(\omega_2|x)$ if $x = (a_1 + a_2)/2$, i.e., the minimum error decision boundary is a point midway between the peaks of the two distributions, regardless of b.

Q3. Suppose we have three equi-probable categories in two dimensions with the following underlying distributions:

$$p(x|\omega_1) \sim N(0, I)$$
$$p(x|\omega_2) \sim N(\begin{bmatrix}1\\1\end{bmatrix}, I)$$
$$p(x|\omega_3) \sim 0.5N(\begin{bmatrix}0.5\\0.5\end{bmatrix}, I) + 0.5N(\begin{bmatrix}-0.5\\0.5\end{bmatrix}, I)$$

By explicit calculation of posterior probabilities, classify the point x = $\begin{bmatrix}0.3\\0.3\end{bmatrix}$ for minimum probability of error.

Q4. a. Write a procedure to generate random samples according to a normal distribution $N(\mu, \Sigma)$ in d dimensions.

b. Write a procedure to calculate the discriminant function for a given normal distribution with $\Sigma = \sigma^2 I$ and prior probability $P(\omega_i)$.

c. Compare the discriminant function's values for two different distributions $N(\mu_1 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, \Sigma_1 = \sigma^2 I)$ and $N(\mu_2 = \begin{bmatrix} -0.5 \\ 0.5 \end{bmatrix}, \Sigma_2 = \sigma^2 I)$ in $d = 2$ dimensions. Assume the test sample to be $X = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $P(\omega_1) = 1/3$ and $P(\omega_2) = 2/3$.

In a general process, you would be given several samples from two (or more) classes. Counting each class' frequency will give the priors. With these samples as d dimensional vectors, you can estimate mean and covariance using MLE or other techniques, which is a part of later lecture. This computed info is sufficient for computing discriminants and thereby classifying the sample into one of the classes.