Statistical Pattern Recognition
Assignment 2

Ali Gholami

Department of Computer Engineering & Information Technology
Amirkabir University of Technology

http://ceit.aut.ac.ir/~aligholamee
aligholamee@aut.ac.ir

Abstract

In this assignment, we'll be focusing on the *Bayes Classifier*. We'll work with *Bayesian Discriminators* and *Bayes Error*. The *Bhattacharyya* error bound is also analyzed as an upper bound for the *Bayes Classifier* error. The detailed computations of *Bayesian Discriminators* are also given in an exact definition. Finally, we'll be going through a more practical example of a linear discriminator by classifying the flowers in the *Iris* dataset.

Keywords. *Linear Discriminator, Quadratic Discriminator, Bayes Classification, Bayes Error, Optimal Classification, Bhattacharyya Distance, Bhattacharyya Upper Bound, Iris Dataset, Iris Classification.*

1 Quadratic & Linear Discriminant Analysis

We consider a classification problem in dimension $d = 2$, with $k = 3$ classes where:

$$p(x \mid w_i) \sim N(\mu_i, \Sigma_i), \quad i = 1, 2, 3$$

and

$$\mu_1 = \begin{bmatrix} 0 \\ 2 \end{bmatrix},\ \mu_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix},\ \mu_3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \Sigma_i = \Sigma \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{3} \end{bmatrix},$$

(a) Calculate the discriminant function $g_i(x)$ for each class.

(b) Express your discriminant functions in the form of linear discriminant functions.

(c) Determine and plot the decision boundaries.

Solution

(a) The general form of a Bayesian discriminator is given below.

$$g_i(\underline{x}) = -\frac{1}{2}(\underline{x} - \underline{\mu}_i)^T \Sigma_i^{-1} (\underline{x} - \underline{\mu}_i) - \frac{1}{2} \log |\Sigma_i| + \log P(\omega_i) \tag{1.1}$$

In the problem case, the classes have the same covariance matrix, but the features have different variances. Since the Σ_i is diagonal, we'll have

$$g_i(\underline{x}) = -\frac{1}{2}(\underline{x}-\underline{\mu}_i)^T \begin{bmatrix} \sigma_1^{-2} & 0 & 0 & \dots & 0 \\ 0 & \sigma_2^{-2} & 0 & \dots & 0 \\ 0 & 0 & \sigma_3^{-2} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \sigma_N^{-2} \end{bmatrix} (\underline{x}-\underline{\mu}_i) - \frac{1}{2}\log \begin{vmatrix} \sigma_1^{-2} & 0 & 0 & \dots & 0 \\ 0 & \sigma_2^{-2} & 0 & \dots & 0 \\ 0 & 0 & \sigma_3^{-2} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \sigma_N^{-2} \end{vmatrix} + \log(P(\omega_i))$$

Since we have the following criteria:

$$(\underline{x} - \underline{\mu}_i)^T = \begin{bmatrix} x[1] - \mu_i[1] \\ x[2] - \mu_i[2] \\ x[3] - \mu_i[3] \\ x[4] - \mu_i[4] \\ \vdots \\ x[N] - \mu_i[N] \end{bmatrix}$$

where μ_{iN} denotes the N'th feature of class i. Removing the constant term for different classes, which is $x[k]^2$, we'll have the following results after the matrix multiplication and determinant computation:

$$g_i(\underline{x}) = -\frac{1}{2}\sum_{k=1}^{N} \frac{2x[k]\mu_i[k] + \mu_i[k]^2}{\sigma_k^2} - \frac{1}{2}\log\prod_{k=1}^{N}\sigma_k^2 + \log(P(\omega_i)) \tag{1.2}$$

One can simply find each discriminator, $g_i(\underline{x})$, by replacing the given information in the problem description in the formula given above. Thus we'll have the following results for the section (a).

$$g_1(\underline{x}) = -\frac{1}{2}(\frac{2x[1]*0+2}{1} + \frac{2x[2]*2+4}{\frac{1}{9}}) - \frac{1}{2}\log(1*\frac{1}{9}) + ?$$

$$g_2(\underline{x}) = -\frac{1}{2}(\frac{2x[1]*3+3}{1} + \frac{2x[2]*1+1}{\frac{1}{9}}) - \frac{1}{2}\log(1*\frac{1}{9}) + ?$$

$$g_3(\underline{x}) = -\frac{1}{2}(\frac{2x[1]*1+1}{1} + \frac{2x[2]*0+0}{\frac{1}{9}}) - \frac{1}{2}\log(1*\frac{1}{9}) + ?$$

The simplified results are

$$g_1(\underline{x}) = -18x[2] - \frac{1}{2}\log\frac{1}{9} - 19$$

$$g_2(\underline{x}) = -3x[1] + 9x[2] - \frac{1}{2}\log\frac{1}{9} - 6$$

$$g_3(\underline{x}) = -x[1] - \frac{1}{2}\log\frac{1}{9} - \frac{1}{2}$$

(b) The final results given above where in the format of a linear discriminant already. In order to lighten everything up, just assume the linear discriminant function as:

$$g_i(\underline{x}) = W_2 x[2] + W_1 x[1] + W_0$$

where the value of W_i is different for each of the discriminators.

$$g_1(\underline{x}) \quad W_2 = -18 \;\; W_1 = 0 \;\; W_0 = -\frac{1}{2}\log\frac{1}{9} - 19$$

$$g_2(\underline{x}) \quad W_2 = 9 \;\; W_1 = -3 \;\; W_0 = -\frac{1}{2}\log\frac{1}{9} - 6$$

$$g_1(\underline{x}) \quad W_2 = 0 \;\; W_1 = -1 \;\; W_0 = -\frac{1}{2}\log\frac{1}{9} - \frac{1}{2}$$

Each of the $g_i(\underline{x})$ represent a discriminator plane in the *3D* space.

(c)