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

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