Math 6020-1, Homework #5

Due date (hard): April 24, 2017 in class

1. Suppose π_1 and π_2 are two populations that are distributed respectively as univariate PDFs f_1 and f_2, where

$$f_1(x) = (1 - |x|)\boldsymbol{I}\{|x| \le 1\}, \quad \text{and} \quad f_2(x) = \tfrac{1}{2}\boldsymbol{I}\{|x| \le 1\}.$$

   Compute and plot the classification regions R_1 and R_2.

2. Let π_1 and π_2 respectively denote a $\mathrm{N}_p(\boldsymbol{\mu}_1\,, \boldsymbol{\Sigma}_1)$ and a $\mathrm{N}_p(\boldsymbol{\mu}_2\,, \boldsymbol{\Sigma}_2)$ population. In the case that $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$ we studied the problem of classifying every point $\boldsymbol{x} \in \mathbb{R}^p$ as π_1 or π_2. The end result was a linear discriminant analysis.

   (a) Prove that in the present more general case, R_1 is the collection of all points $\boldsymbol{x} \in \mathbb{R}^p$ such that

$$-\tfrac{1}{2}\boldsymbol{x}'\left(\boldsymbol{\Sigma}_1^{-1} - \boldsymbol{\Sigma}_2^{-1}\right)\boldsymbol{x} + \left(\boldsymbol{\mu}_1'\boldsymbol{\Sigma}_1^{-1} - \boldsymbol{\mu}_2'\boldsymbol{\Sigma}_2^{-1}\right)\boldsymbol{x} \ge L,$$

   where L is a constant that you should compute explicitly in terms of the costs $c(1 \mid 2)$ and $c(2 \mid 1)$ and priors p_1 and p_2, as well as the other parameters $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}_i$. Notice that the preceding yields a "quadratic discriminant" and not a linear discriminant.

   (b) Discuss, in good detail, how one could implement the preceding in a realistic setting where $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_1$, and $\boldsymbol{\Sigma}_2$ are unknowns.

3. Consider the following summary statistics from two (independent) data sets:

$$\overline{\boldsymbol{X}}_1 = \begin{bmatrix} 1.80 \\ 3.89 \end{bmatrix}, \quad \overline{\boldsymbol{X}}_2 = \begin{bmatrix} 5.50 \\ 8.01 \end{bmatrix}, \quad \boldsymbol{S}_1 = \begin{bmatrix} 1.21 & 1.10 \\ 1.10 & 2.00 \end{bmatrix}, \quad \boldsymbol{S}_2 = \begin{bmatrix} 1.06 & 1.62 \\ 1.62 & 2.50 \end{bmatrix}.$$

   (a) Compute the linear discriminant function. Use that linear discriminant function to classify the point $\boldsymbol{x} = [1.27\,, 4.45]'$ as either π_1 or π_2.

   (b) Under which theoretical assumptions do you expect your method to be reliable? For the present data, is it reasonable to assume that those assumptions are (or are at least partly) valid?

   (c) Suppose instead that $\boldsymbol{S}_2$ were replaced by the following (hypothetical) choice:

$$\boldsymbol{S}_2 = \begin{bmatrix} 41 & 40 \\ 40 & 41 \end{bmatrix}.$$

   Would you use the linear disciminant to classify $\boldsymbol{x}$? If not, then what would you do (and please do it!)?