Lectures 5, 6

Reading: Sections 4.1-4.3,

Notes: We will start with classification using linear discriminant analysis. Under some assumptions this method divides the feature space into regions, each of which corresponds to a unique predicted class.

Problems:

1. Hastie, problem 4.2

2. In the discussion of section 4.2, consider a dataset \( \{(x_i, y_i)\}_{i=1}^{24} \) with 4 categories \((k = 1, \ldots, 4)\), in which the category \( y_i \) of data point \( x_i \) is \( i \text{ mod } 4 + 1 \) (i.e., the category alternates successively to the next one in the group of 4).
   (a) What does the data matrix \( Y \) look like?
   (b) Let \( y \) be the first column of \( Y \). If we replace the data matrix \( Y \) by \( y \), show that the vector \( \beta \) for this standard regression is just the first column of the matrix \( B \).
   (c) Generalize this statement to all columns of \( B \).
   (d) What does this tell you that the regression using \( Y \) is doing? Can it be interpreted precisely in terms of the individual regressions arising from replacing \( Y \) by one of its columns?

3. Derive equation (4.9)