Bayesian Classification (Klasyfikacja Bayesowska)

Exercises

1. Implementation of LDA and QDA methods for two classes

   - For each method write two functions.
     - A function that takes as an input a matrix of coefficients and a vector of classes and it trains the model (for the training data).
     - A function that takes as an input a matrix of coefficients and it returns posterior distributions (for the test data).
   - Compare the performance of the implemented functions with `lda` and `qda` (package: `MASS`) for the dataset `Default` (package: `ISLR`)

2. Comparison of LDA and QDA methods on simulated data.

   - Consider two simulation schemes.
     - **Scheme 1:** Training dataset contains $n = 1000$ observations, $p = 2$ features and a binary variable that is generated from the Bernoulli distribution with probability of success 0.5. Features of the observations from the class 0 are generated independently from a normal standard distribution. Features of the observations from the class 1 are generated independently from a normal distribution (mean 1, variance 1). **LDA assumptions are satisfied**
     - **Scheme 2:** Training dataset contains $n = 1000$ observations, $p = 2$ features and a binary variable that is generated from the Bernoulli distribution with probability of success 0.5. Features of the observations from the class 0 are generated from a two-dimensional normal distribution (mean 0, varianve 1, correlation 0.8). Features of the observations from the class 1 are generated from a two-dimensional normal distribution (mean 1, variance 1, correlation -0.8). **LDA assumptions are not satisfied**
   - Compare LDA and QDA for both schemes (compute accuracy on the training set).