Tsinghua-Berkeley Shenzhen Institute
Learning from Data
Fall 2018

Programming Assignment 2

Issued: Monday 29th October, 2018 **Due:** Monday 5th November, 2018

2.1. Suppose we are given a dataset $\{(x^{(i)}, y^{(i)}): i = 1, 2, \dots, m\}$ consisting of m independent examples, where $x^{(i)} \in \mathbb{R}^n$ is an n-dimension vector, and $y^{(i)} \in \{1, 2, \dots, K\}$. We will model the joint distribution of (x, y) based on:

$$y^{(i)} \sim \text{Multinomial}(\phi),$$
$$x^{(i)}|y^{(i)} = j \sim \mathcal{N}(\mu_j, \Sigma_j),$$

where the parameter ϕ_j gives $p(y^{(i)} = j)$ for each $j \in \{1, 2, \dots, K\}$.

(a) (3 points) Linear Discriminant Analysis (LDA) is a special case of Gaussian Discriminant Analysis (GDA) which assumes that the classes have a common covariance matrix $\Sigma_j \equiv \Sigma, \forall j$. Now suppose all the Σ_j's are not equal, and we will get the Quadratic Discriminant Analysis (QDA). The estimations for QDA are similar to those for LDA, except that separate covariance matrices must be estimated for each class. Give the maximum likelihood estimate of Σ_j's for the case $K = 2$.

(b) (7 points) Complete the codes for LDA and QDA. This link provides some information about LDA and QDA.

`Pa2_2018.py` will walk you through this exercise.

Notice:

- Submit your codes together with a document containing your proof for part (a).
- You may need some knowledge of matrix calculus to finish this proof.
- Do NOT modify pre-written codes in the file.
- If your implementation is right, for the given data, the accuracy of LDA is about 0.44, and the accuracy of QDA is about 0.47.