Exercise no 4

FirstName LastName StudentNumber

Statistical Machine Learning

February 25, 2018

1 Mathematical Statistics

Exercise 1.1 Read Section 2 of `http://www-stat.stanford.edu/~tibs/ftp/lars.pdf`, then solve Ex 3.25 of Elements of Statistical Learning (ESL) in page 97.

Solution 1.1

Exercise 1.2 Consider two interesting functions related to Bernoulli distribution, and defined over $x \in (0, 1)$.

$$f(x) = x\{\log(1-x) - \log x\} - \log(1-x) \tag{1}$$

$$g(x) = x(1-x) \tag{2}$$

- Plot these two functions.
- Argue how these two functions are related to the concept of information and the concept of entropy.
- Write the Tylor expansion of $f(x)$ up to a quadratic term and compare with $g(x)$.
- Find $\lim_{x\to 0} x \log(x)$. How does this limit help to define $f(x)$ over $x \in [0, 1]$.

2 Optimization

Exercise 2.1 Many of quadratic multivariate methods fall on eigenvalue problem.

- For given positive definite matrices $\mathbf{B}$ and $\mathbf{W}$, show

$$\lambda_{\max} = \max \frac{\mathbf{x}^\top \mathbf{B}\mathbf{x}}{\mathbf{x}^\top \mathbf{W}\mathbf{x}}, \qquad \mathbf{e}_{\max} = \operatorname{argmax} \frac{\mathbf{x}^\top \mathbf{B}\mathbf{x}}{\mathbf{x}^\top \mathbf{W}\mathbf{x}}$$

where $\lambda_{\max}$ is the maximum eigenvalue of $\mathbf{B}\mathbf{W}^{-1}$.

Hint: use the result of the principal components.

- How do you think this optimization problem is related to data classification?

Solution 2.1

3 Computation

Exercise 3.1 Suppose the data vector $\mathbf{x}_{p\times 1}$mean vector $\boldsymbol{\mu}_{p\times 1}$ and variance covariance matrix $\boldsymbol{\Sigma}_{p\times p}$ are given. How do you compute the log likelihood of multivariate normal distribution efficiently?

Solution 3.1