Exercise no 1

FirstName LastName StudentNumber

Statistical Machine Learning

December 31, 2017

1 Optimization

For each part of the following exercise argue if derivative is applicable for minimization, and why derivative equal zero gives the minimum.

Exercise 1.1 Find the least squares estimator $\hat{\beta}_0 = \text{argmin } S(\beta_0)$, in which

$$S(\beta_0) = \sum_{i=1}^{n}(y_i - \beta_0)^2.$$

Solution 1.1

Exercise 1.2 Find the least squares estimator $\hat{\beta}_1 = \text{argmin } S(\beta_1)$, in which

$$S(\beta_1) = \sum_{i=1}^{n}(y_i - \beta_1 x_{1i})^2.$$

Solution 1.2

Exercise 1.3 Find the least squares estimator $(\hat{\beta}_0, \hat{\beta}_1) = \text{argmin } S(\beta_0, \beta_1)$, in which

$$S(\beta_0, \beta_1) = \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_{1i})^2.$$

Solution 1.3

Exercise 1.4 Find the least squares estimator $(\hat{\beta}_1, \hat{\beta}_2) = \text{argmin } S(\beta_1, \beta_2)$, in which

$$S(\beta_1, \beta_2) = \sum_{i=1}^{n}(y_i - \beta_1 x_{1i} - \beta_2 x_{2i})^2.$$

Solution 1.4

Exercise 1.5 How do you minimize if you cannot differentiate the function S, but you know S is differentiable?

Solution 1.5

Exercise 1.6 How do you minimize if you know S is not differentiable (such as the absolute loss function $S(\beta_0) = \sum_{i=1}^{n} |y_i - \beta_0|$.

Solution 1.6

2 Linear Algebra

Exercise 2.1 Show that any matrix in the form $\mathbf{A}^\top \mathbf{A}$ is positive semi-definite where $\mathbf{A}^\top$ is the transpose of $\mathbf{A}$.

Solution 2.1

Exercise 2.2 How can you use this result in optimization?

Solution 2.2

Suppose you have a code that solves the systems of linear equations $\mathbf{A}\mathbf{x} = \mathbf{b}$ where $\mathbf{A}_{p\times p}$ and $\mathbf{b}_{p\times 1}$ both are known and $\mathbf{x}$ is unknown.

Exercise 2.3 When does this system of linear equation have at least one solution? Why?

Solution 2.3

Exercise 2.4 When does this system of linear equation has infinite solutions? Why?

Solution 2.4

Exercise 2.5 When does this system of linear equation does not have any solution?

Solution 2.5

Exercise 2.6 When does this system of linear equation have exactly one solution? Write a pseudo code that finds this solution.

Solution 2.6

Exercise 2.7 How can you find the inverse of a matrix using a code that solves this system of linear equations?

Solution 2.7

3 Mathematical Statistics

Exercise 3.1 What is Fisher information, observed information, and Hessian. How they are related and why they are useful?

Solution 3.1

Exercise 3.2 Find the maximum likelihood estimator for univariate θ, if $y_i \mid x_i \sim N(x_i\theta, 1)$, and then find the asymptotic variance of $\hat{\theta}_{\mathrm{MLE}}$ using the closed form of $\hat{\theta}_{\mathrm{MLE}}$, using Fisher information, using observed information, using Hessian.

Solution 3.2