Exercise no 1

FirstName LastName StudentNumber

Statistical Machine Learning

February 3, 2018

1 Optimization

Exercise 1.1 Show that $S(\boldsymbol{\beta}) = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^\top(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$ is a convex function.

Solution 1.1

2 Implementation

Install Anaconda 3 (along with Python 3.6). Use "jupiter notebook" to execute your python codes.

Exercise 2.1 Copy paste the following code in your jupyter notebook and upload the grphs for polynomial order=1, order=2, order=3. You may play with the order from 1 to 10 and let me know which order think is a better fit? Why?

```
# library for data handling and statistics

import pandas as pd

# library for numerical computation

import numpy as np

# library for plot

import matplotlib.pyplot as plt

# library for plot style

import seaborn as sns

# put plots in jupyter window

%matplotlib inline

# define the path of data

PATH = 'http://www-bcf.usc.edu/~gareth/ISL/Advertising.csv'

# read a CSV data file

advertising = pd.read_csv(PATH, usecols=[1,2,3,4])

# seaborn plot

sns.regplot(advertising.TV, advertising.sales, order=3, ci=None, scatter_kws={'color'



# control axes

plt.xlim(-10,310)

plt.ylim(ymin=0);
```

Solution 2.1

3 Mathematical Statistics

Exercise 3.1 In this exercise you see that the mean and the median both impose a constant model on data, but the constant is estimated differently.

- Show that the mean is the minimizer of of the constant model $y_i = \beta_0 + \varepsilon_i$, using $S(\beta_0) = \sum_{i=1}^{n}(y_i - \beta_0)^2$
- Show that the median is the minimizer of of the constant model $y_i = \beta_0 + \varepsilon_i$, using $S(\beta_0) = \sum_{i=1}^{n}|y_i - \beta_0|$
- Construct a loss function $S(\beta_0)$ that produces the trimmed-mean as $\hat{\beta}_0$.

Solution 3.1

Exercise 3.2 Show maximum likelihood estimator for linear model $y_i = \mathbf{x}_i^\top \beta + \varepsilon_i$ coincides with the least squares if $\varepsilon_i \sim N(0, \sigma^2)$, for a known σ^2.

Solution 3.2

Exercise 3.3 Recommend a density function for ε_i that leads to median regression. Median regression minimizes $S(\boldsymbol{\beta}) = \sum_{i=1}^{n}|\mathbf{y}_i - \mathbf{x}_i^\top \boldsymbol{\beta}|$

4 Computation

Exercise 4.1 How do you compute the least squares estimator $\hat{\boldsymbol{\beta}}$ using for a given response vector $\mathbf{y}$ and design matrix $\mathbf{X}$

- LU decomposition

- QR decomposition
- SVD decomposition

Solution 4.1