Name: ..... ID Number: .....

1 Least-squares

The dataset can be downloaded at https://bit.ly/2RZ15t2

1. Load a dataset from `ex1_x.csv` as x and `ex1_z.csv` as z

2. Concatenate a bias vector with x as a new matrix X

3. Find the Least-squares solution by using the following equation

$$\hat{\theta} = (X^T X)^{-1} X^T z$$

4. Plot the Least-squares estimation of the function.

5. Plot the true function with $\theta = [3.75 \quad -1.15]^T$

6. Plot the measured data.

7. Observe and discuss the results in Question 4–6

8. Find the Least-squares solution by using pseudo-inverse

9. Plot the Least-squares estimation of the function calculated by pseudo-inverse.

10. Discuss on what you observe

2 Least-Squares - Multivariate Linear Regression

After the first lab session, I believe that you all get to know the Boston Housing Dataset. Thus, I do not need to give you any explaination about the data. The dataset can be downloaded at https://bit.ly/2NBaO7e.

1. Import dataset from `regression-datasets-housing.csv`

2. Find theta by using Gradient Descent Algorithm

3. Find theta by using Least-Squares

4. Compare an error between Gradient Descent Algorithm and Least-Squares. What do you observe?

5. Can these two algorithms obtain the optimal solutions?

6. Plot the importance of all the features. Which features is the most importance features in this dataset?

3 Least-Squares - Non-linear Data

1. Import dataset from `ex2_x.csv` as x and `ex2_z.csv` as Z. It is noted that z is generated by sine function with Gaussian random noise (S.D.=0.2).

2. Plot the data distribution

3. Obtain the least-squares solution of x and z and overlay the obtain function on the data distribution you plot in the previous question. Discuss on the results you get.

4. Find a series of models by increasing the polynomial degree.

5. Plot the result of each polynomial degree

6. Discuss the experimental results

7. Plot the importance of all the polynomial degree. Which degree(s) is/are the most importance feature? What is the appropriate number of degree?

8. What happens if you drop the terms corresponding to the weights of smallest magnitude?

9. What happens if you drop those terms and then re-compute the model?