Seung Jun Baek

1 Outline

In this assignment, you are asked to perform linear regression to estimate the linear model with noise.

2 Specification

In this assignment, you are asked to write a Python code that estimates the linear model with noise.

Your data is generated from the model

$$b = Ax_{\text{true}} + \epsilon$$

where output data is given by

$$b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

and input data

$$A = \begin{bmatrix} 1 & a_1 & a_2 & \dots & a_{n-1} \\ 1 & a_1 & a_2 & \dots & a_n \\ 1 & a_1 & a_2 & \dots & a_n \end{bmatrix}$$

Also

$$\epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_m \end{bmatrix}$$

is a noise vector, ϵ_i, $i = 1, \dots, m$ is normally distributed with mean 0 and standard deviation 0.1. $x_{\text{true}} \in \mathbb{R}^n$ is some unknown vector which we want to estimate.

To estimate $x_{\text{true}} \in \mathbb{R}^n$, we would like to use linear regression. In other words, want to find parameter vector $x \in \mathbb{R}^n$ which minimizes the mean square error (MSE) such that

$$\text{minimize } \frac{1}{m}\|Ax - b\|_2^2$$

(Or you can simply minimize $\|Ax - b\|_2^2$. In other words, factor $1/m$ does not affect the solution) The goal is to write a Python code to find x that minimizes MSE.

In the blackboard you will find the attachment named `hw1.py`. It contains the "skeleton" code for implementing linear regression for the above problem. But the code is not complete. **Your goal is to fill in the make the code complete for linear regression.**

You will use `minimize` function provided by Python to minimize the MSE. Your code will use `minimize` function from `scipy` library which implements a generic minimization. Documents on python `minimize` can be found at `https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.minimize.html`. In summary, you need to do the following two things:

1. Properly define the MSE function.
2. Use `minimize` function to perform minimization of MSE.

3 What to submit

- Submit a modified Python file `hw1.py`. The requirements are
  1. complete the function `MSE(x,A,b)`. In other words, fill in the following

```
# Mean Square Error function
def MSE(x, A, b):
    # parameter:
    # x: vector with shape (n,)
    # A: matrix with shape (m,n)
    # b: vector with shape (m,)
    ##############################
    # PART 1: completing MSE function
    ##############################
    # You need to fill in here.
    # Your function must return the MSE between Ax and b

    return ...
```

  2. complete the line obtaining `estim`, which contains the solution to the minimum the MSE loss. In other words, fill in the following

```
##############################
# PART 2: Finding estimate using minimize function
##############################
# use minimize function to find the best parameter estimate
# you need to properly use minimize function below
estim = minimize(...)
```

- Upload your `hw1.py` file at Blackboard before deadline. (Please submit the file in time, no late submission will be accepted).

4 How to test your module

Run your completed code `hw1.py`. At the output console you will see something like

```
solution from minimize: [ 0.1499277  -0.21636316 -0.27117587
 0.91924725 -1.11037848  0.31362976
 -1.50370403 -0.30722973  0.09011209  0.74199764]
true x [ 0.15013674 -0.21648811 -0.26893758  0.91425614 -1.10981388
 0.3132908
 -1.50202904 -0.30718146  0.08650764  0.74085398]
error percentage: 0.3018923964660468 %
```

As seen in the above, if you have implemented the function correctly, your error should be below 1%.

5 Grading

- 5 points if your module works correctly. Specifically, if your error is less than 1%.
- 2 points if your module does not work correctly. Specifically, if your error is above 1%.
- 0 point if you do not submit the file by deadline.

In the blackboard, you can upload your file as many times as you like, before the deadline. The last uploaded file will be used for grading. After deadline, the submission menu will be closed and you will not be able to make submission.