HW1

Name

Problem 1

(a)

The calculation steps of finding the minimum.

Show that it is actually a minimum.

(b)

Code for gradient descent algorithm with constant step size and with back-tracking line search. Print out the minimum found in both cases.

1.

Discuss how you selected the constant step size used in your code.

2.

Show the code for plotting and two plots generated:

Plot 1: the value of the objective function vs. number of iterations for constant step size.

Plot 2: the value of the objective function vs. number of iterations for backtracking.

3.

Code and the generated plot of eta_k vs. iteration k.

Comment on the result.

Problem 2

(a)

Show your code of gradient descent with constant step size and with back-tracking line search.

Read in the two datasets. Apply the gradient descent algorithm. Record the results you found using gradient descent with constant step size and with back-tracking line search for each dataset. (4 results in total)

1.

Discuss how you selected the constant step size.

Discuss the convergence criterion and tolerance parameter used.

2.

Show your code of estimating beta_hat using lm command in R or sklearn.linear_model in Python.

For each estimate obtained using gradient descent, calculate the norm of the difference between it and the corresponding estimate obtained using inbuilt functions.

3.

Show the code and the generated plots (4 in total).

Plot 1: objective function value vs number of iteration for GD with constant step size on dataset1.

Plot 2: objective function value vs number of iteration for GD with back-tracking on dataset1.

Plot 3: objective function value vs number of iteration for GD with constant step size on dataset2.

Plot 4: objective function value vs number of iteration for GD with back-tracking on dataset2.

(b)

Show your code of Polyak and Nesterov momentum methods with constant step size and with back-tracking line search. (2 functions in total, one for Polyak, one for Nesterov)

Apply the methods on the two datasets. Record the results you found using the two momentum methods with constant step size and with back-tracking line search for each dataset. (8 results in total)

1.

Similar to Problem 2 Part (a) 2.

2.

Similar to Problem 2 Part (a) 3. 8 plots in total. Make the comment as required.

3

Based on your results in Part (a) 2&3 and Part (b) 1& 2, make the comment as required.