STAT 102B: Homework 4

George Michailidis

Due electronically through
BruinLearn/Gradescope on Wednesday June 4 at 11:00 pm

Problem 1:

Consider the function

$$f(x) = \frac{1}{4}x^4 - x^2 + 2x$$

Using the pure version of Newton's algorithm **report** x_k **for** $k = 20$ (after running the algorithm for 20 iterations) based on the following 5 initial points:

1. $x_0 = -1$

2. $x_0 = 0$

3. $x_0 = 0.1$

4. $x_0 = 1$

5. $x_0 = 2$

(i) What do you observe?

(ii) How can you fix the issue reported in (i)?

Problem 2:

Consider the data in the train_data.csv file.

The first 600 columns correspond to the predictors and the last column to the response y.

Implement that proximal gradient algorithm for Lasso regression, by modifying appropriately your code from Homework 1.

To select a good value for the regularization parameter λ use the data in the validation_data.csv to calculate the sum-of-squares error validation loss.

Show a plot of the training and validation loss as a function of iterations.

Report the number of regression coefficients estimated as zero based on the best value of λ you selected.