

COEN 140

Lab 3 Report

Stephen Tambussi - 00001469512

Problem 1

- a. MSE values
 - i. Training data: $[[0.01630465]]$
 - ii. Testing data: $[[0.02126774]]$
- b. First 10 values of optimal weight vector:
 - $[[2.36115262e-01]$
 - $[-1.54906138e-02]$
 - $[3.35459838e-03]$
 - $[1.27547102e-02]$
 - $[-2.99835603e-02]$
 - $[-2.49633935e-02]$
 - $[2.42430772e-02]$
 - $[-1.26473288e-02]$
 - $[2.03432879e-02]$
 - $[-3.20098178e-02]$
- c. Predicted crime rate of first 10 test samples:
 - $[[5.42331530e-02]$
 - $[2.83258436e-01]$
 - $[1.16474255e-01]$
 - $[1.15555411e-01]$
 - $[1.55865217e-01]$
 - $[1.33145752e-01]$
 - $[7.60943848e-01]$
 - $[2.52983564e-01]$
 - $[2.31453712e-01]$
 - $[5.11709643e-01]$

Problem 2

- a. MSE values and optimal weight vector
 - i. Training data: $[[0.001034]]$
 - ii. Testing data: $[[0.53967314]]$
 - iii. First 10 values of optimal weight vector:
 - $[[2.48993770e-01]$
 - $[5.53500158e-02]$
 - $[-3.97965119e-02]$
 - $[-2.23470491e-02]$
 - $[2.48021165e-02]$

[9.82015016e-02]
[5.23273576e-01]
[4.49260891e-01]
[-2.33720994e-02]
[3.41409119e-01]

b. MSE values and optimal weight vector for ridge regression

- i. Training data: [[0.03235784]]
- ii. Testing data: [[0.04836619]]
- iii. First 10 values of optimal weight vector:

[[0.11747502]
[0.00977899]
[-0.00618524]
[-0.00708916]
[-0.00766872]
[-0.00645346]
[0.00017213]
[0.01161043]
[0.00608684]
[-0.00065077]

- c. In part 2a, with only a 100 training samples used the model had a very small MSE value for the training data. However, the MSE value for the testing data was much higher, indicating lower performance due to an inadequate amount of training data for the linear regression model. For part 2b, the same amount of training data was used (100 samples) with a ridge regression model. This model's MSE value for the training data was higher than the model for part 2a, but its MSE for the testing data was much lower indicating improved performance over a data-constrained linear regression model.