COEN 140
Lab 4 Report
Stephen Tambussi - 00001469512

Problem 1

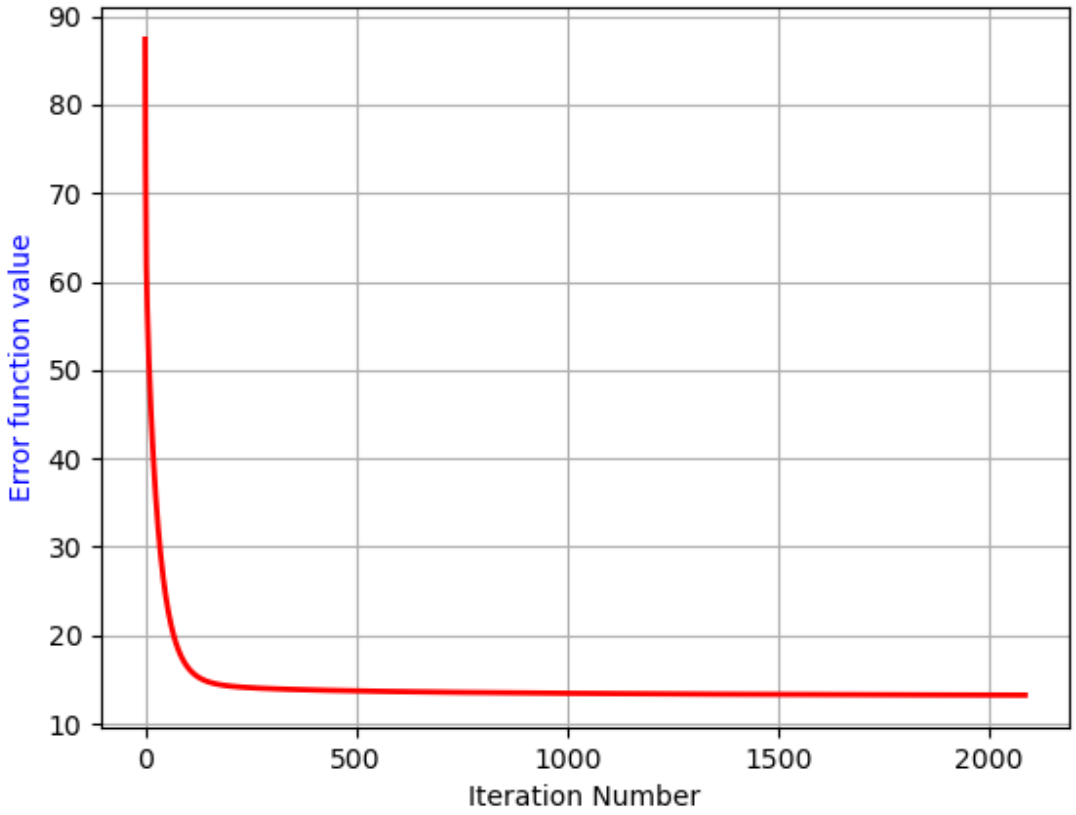
a. Algorithm exits at iteration: 2086
b.
    i. MSE of train: [[0.0166115]]
    ii. MSE of test: [[0.02126524]]
c. First 10 elements of weight vector:

```
[[ 2.36253626e-01]
 [-1.03715193e-02]
 [ 1.41353672e-02]
 [ 2.01145273e-03]
 [-2.83195947e-02]
 [-5.98489905e-03]
 [ 1.95060115e-02]
 [-9.72886571e-03]
 [ 1.93797089e-02]
 [ 8.20198684e-03]
```

d. Predicted crime rate of first 10 test samples:

```
[[ 0.06066813]
 [ 0.28074684]
 [ 0.1133951 ]
 [ 0.13719314]
 [ 0.14779631]
 [ 0.1643951 ]
 [ 0.74594688]
 [ 0.26401785]
 [ 0.24181028]
 [ 0.47839781]
```

e. The MSE values compared to lab 3 problem 1 are very similar. The MSE of the training data for lab 4 is actually greater than the MSE of lab 3 by approximately 0.0003. However, the MSE value of the testing data is smaller for lab 4 indicating greater performance of the model with gradient descent. The weight vector values are somewhat similar, but the predicted crime rate of the test samples appears to be more accurate in lab 4 than in lab 3.
f. Error function vs. iteration number



As shown in this graph, with increasing iterations the value of the error function decreases until iteration 2086, when the condition for the stopping criterion is met. The error function indicates how “close” the predicted crime rate for the training samples is to the actual target (actual crime rate) values of the training samples. The stopping criterion used was if the maximum absolute difference between any weight elements of the current iteration and previous iteration is less than the epsilon value. This stopping criterion allowed the model training to finish early. It also helped prevent the overfitting problem where the model weights become too adjusted to the training data and therefore the model will not generalize well.