Alex Ivensky

ECE 1395

Homework 2

1.

A black background with white numbers

Description automatically generated

2.

A screen shot of a computer code

Description automatically generated

3.

A screen shot of a computer

Description automatically generated

There is a significant difference. The gradient descent conducted on this data set had too low of a learning rate due to the minimal decrease in the cost over the iterations, and too few iterations. Increasing both values makes the gradient descent response closer to the normal equation response.

4b.

A graph of blue dots

Description automatically generated

4c.

A black background with white text

Description automatically generated

4e.

A black background with white text

Description automatically generated

A graph with blue lines

Description automatically generated

4f.

A graph of a horse power

Description automatically generated

4g & 4h:

A black background with white numbers

Description automatically generated

The errors in one run of the program will appear different, but running the program multiple times shows different results. This is due to the random initialization of theta. This shows that both algorithms on average produce the same error.

4i.

|  |  |
| --- | --- |
|  |  |
|  |  |

In figure 1, the learning rate is too low, so the gradient descent algorithm is slow to converge, and does not converge in the given 300 iterations. In figures 2 and 3, they both converge, but figure 3 converges much faster due to its higher learning rate. In figure 4, the algorithm diverges in the last two iterations because the learning rate is too high.

5a.

A black background with white numbers

Description automatically generated

A black background with white letters

Description automatically generated

5b.

A black background with white numbers

Description automatically generated

5c.

