Analysis Report

1 Normal Equation and Gradient Descent Algorithm without regularizer :-

1.1 Hypothesis parameters using Normal Equation Algorithm are :-

We got idea that if our matrix is singular then we could make it invertible by eleminating redundant features or we could make a new feature using the exsisting feature to eleminate the problem of singularity of matrix. Following are the results obtained after using mean normalization.
$\theta = [[1.37658083e+03][2.15081925e-01][1.11110312e+02][8.69446160e+02][3.97676261e+02]$
$[4.05611236e+02][2.73607644e+02][3.30685106e+02][7.78219890e+02][7.66179987e+02]$
$[2.57447262e+02][5.68257407e+02]]$

1.2 Hypothesis parameters using Gradient Descent Algorithm are:-

We used different $\alpha = 0.01,0.001,0.0001$ and epoch = 100,1000,10000 repectively.Finally our algorithm converged best at $\alpha = 0.0001$ and epoch = 10000 with minimal possible error.
$\theta = [[-0.14265545][0.13801749][0.0412469][0.04138947][0.04141092][0.04133055][0.04139053][0.0413724]$
$[0.0413582][0.04144458][0.04146955][0.04140794]]$

1.3 Performance of predictor/hypothesis in both the cases :-

Ans. Perfomance of the predictor in normal equation was found to better than gradient descent since the calculated squared error in gradient was 0.42740064225945834 whereas in normal equation was 0.42128128593783093. In general we got idea that if number of features are more and data points are less then gradient descent will perform much better as time complexity of normal equation is $O(n^3)$ but if number of features are less and data points are more then normal equation will perform better.

MIT2018072