Analysis Report

1 Normal Equation and Gradient Descent Algorithm with regularizer :-

1.1 Hypothesis parameters using Normal Equation Algorithm with regularization

Following are the results obtained after using mean normalization. Without mean normalization there was lot of chattering and our algorithm was failing to converge.
$\theta = [[-4.63024605e-01][1.56785583e-01][6.18328233e-05][6.24359518e-05][9.37648359e-05]$ $[2.00184105e-05][2.46306334e-05][1.53985864e-05][5.40463783e-06][4.71981248e-05]$ $[6.00614100e-05][2.74224103e-05]]$

1.2 Hypothesis parameters using Gradient Descent Algorithm with regularization are:-

We used different $\alpha = 0.01, 0.03, 0.001, 0.003, 0.0001, 0.0003$ and epoch $= 100, 100, 1000, 1000, 10000, 10000$ repectively.Finally our algorithm converged best at $\alpha = 0.01$ and epoch $= 10000$ with minimal possible error. $\theta = [[-0.5848375][0.37900657][0.17450523][0.17501759][0.17818519][0.17071102][0.17145451]$ $[0.17158816][0.17008154][0.17391457][0.17429542][0.17197108]]$

1.3 Performance of predictor/hypothesis in terms of accuracy with regualrizer and without regualrizer

Ans. We can see from the squared error functions that perfomance of the predictor with regualriser is better to predict values of new points as it eliminates the problem of overfitting/high variance which was earlier in the predictor. For our problem there was not much gain but in general we got idea that overfitting could be avoided using regularizaiton and moreover it can help in making our $(X^T.X)$ matrix invertible in normal equation if it was earlier not possible.

MIT2018072