Polynomial of Degree : 1
Starting Cost theta initiated as one: 1.6982862077051877
Polynomial of Degree : 2
Starting Cost theta initiated as one: 2.6320538340598687
Polynomial of Degree : 3
Starting Cost theta initiated as one: 3.5625176902420317
Polynomial of Degree : 4
Starting Cost theta initiated as one: 4.480386783555601
Polynomial of Degree : 5
Starting Cost theta initiated as one: 5.384800114493669
Polynomial of Degree : 6
Starting Cost theta initiated as one: 6.276989139522175
Polynomial of Degree : 7
Starting Cost theta initiated as one: 7.158478749063768
Polynomial of Degree : 8
Starting Cost theta initiated as one: 8.030589888578536
Polynomial of Degree : 9
Starting Cost theta initiated as one: 8.894348859937836

Performing regularization on minimum train cost polynomial

Minimum error: 0.0003689405325816304

Lambda: 0.25
Final Lasso Train cost:: 0.00033742345810168706
Final Ridge Train cost:: 0.040007954745503296
Lambda: 0.5
Final Lasso Train cost:: 0.00033006930764983817
Final Ridge Train cost:: 0.03997715908226119
Lambda: 0.75
Final Lasso Train cost:: 0.0003245743580898326
Final Ridge Train cost:: 0.03995064404323816
Lambda: 1
Final Lasso Train cost:: 0.0003209260065673399
Final Ridge Train cost:: 0.03992843578446011

Performing regularization on maximum train cost polynomial

Maximum error: 0.09967924136211806

Lambda: 0.25
Final Lasso Train cost:: 0.09920820024943446
Final Ridge Train cost:: 0.101802342539069
Lambda: 0.5
Final Lasso Train cost:: 0.0992080178571436
Final Ridge Train cost:: 0.10180293682692926
Lambda: 0.75
Final Lasso Train cost:: 0.09920799121844102
Final Ridge Train cost:: 0.1018038899531708
Lambda: 1
Final Lasso Train cost:: 0.09920809219642739
Final Ridge Train cost:: 0.10180520191587451

Plots for minimum train cost polynomial
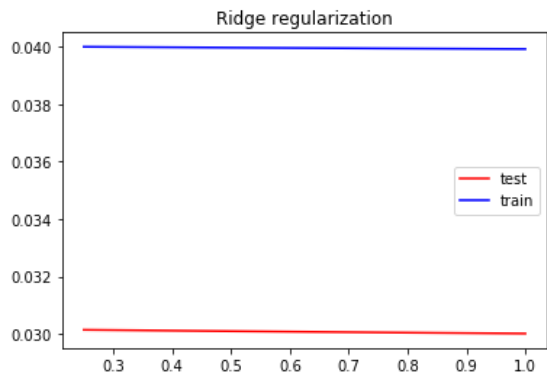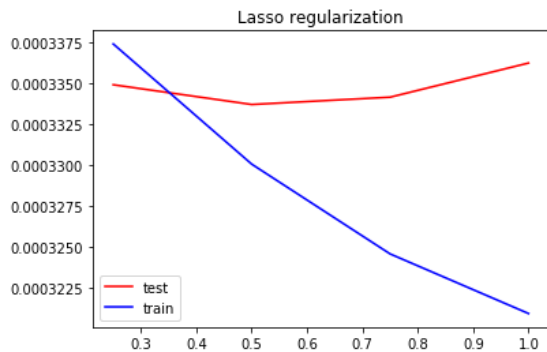
Plots for maximum train cost polynomial
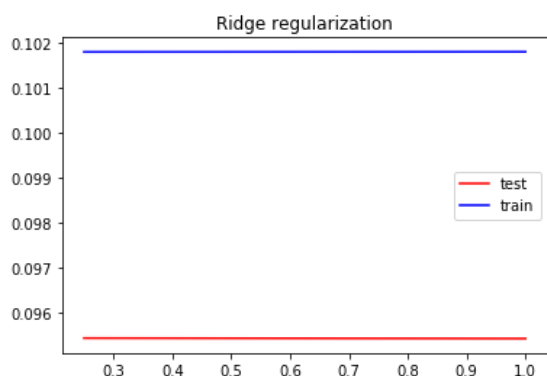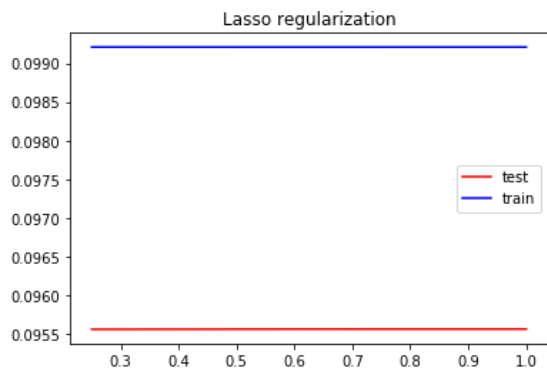
From the plots we see that only minimum train cost is affected by regularization(as expected) and since our motive was to increase bias, we hope to get high train error, from the plots hence Ridge regression seems to do a better job

We also see that as lambda increases the Lasso regression train cost keeps on decreasing while the ridge regression train cost although constant, is much higher than the original minimum error, indicating that Lasso might still overfit the data

I will use Ridge regression for this problem