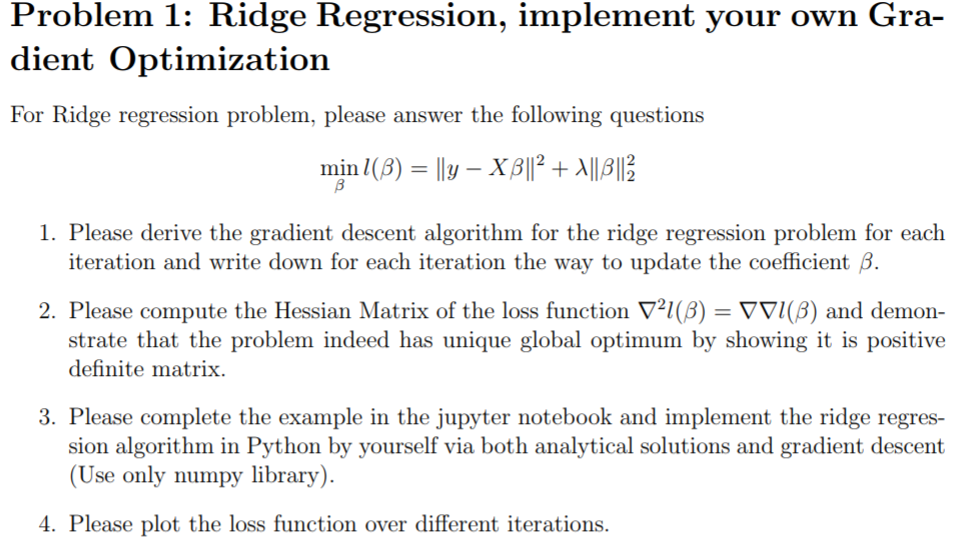
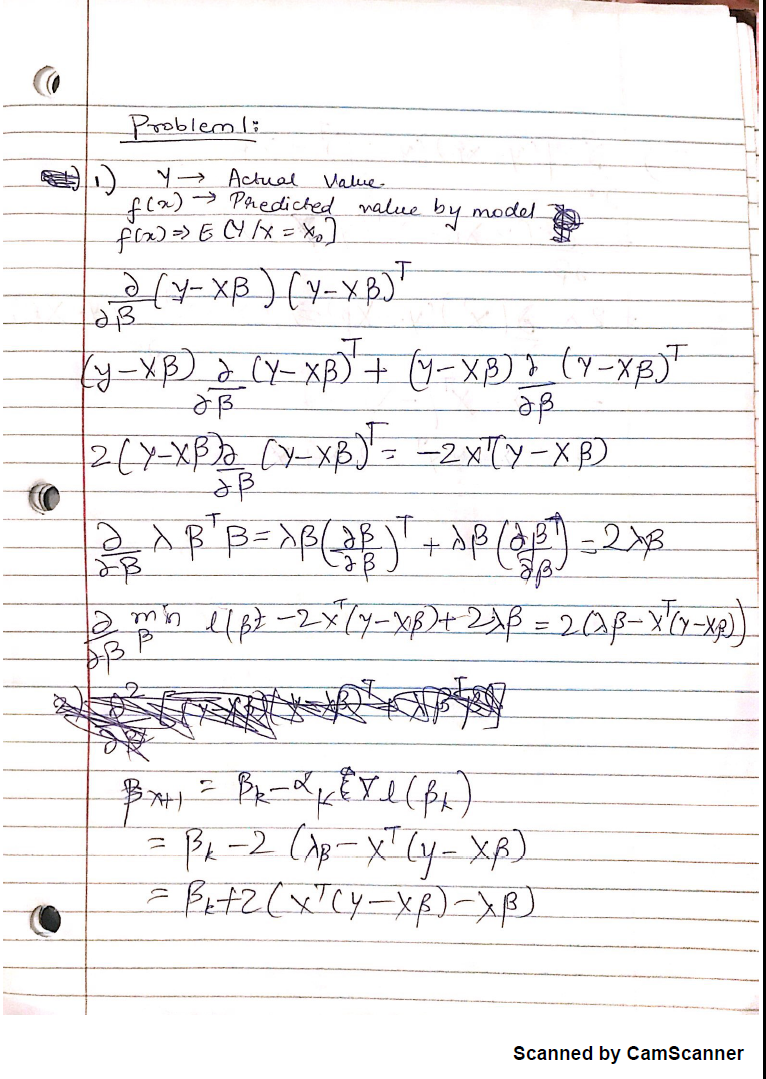
IEE 598: DATA SCIENCE FOR SYSTEM INFORMATICS:

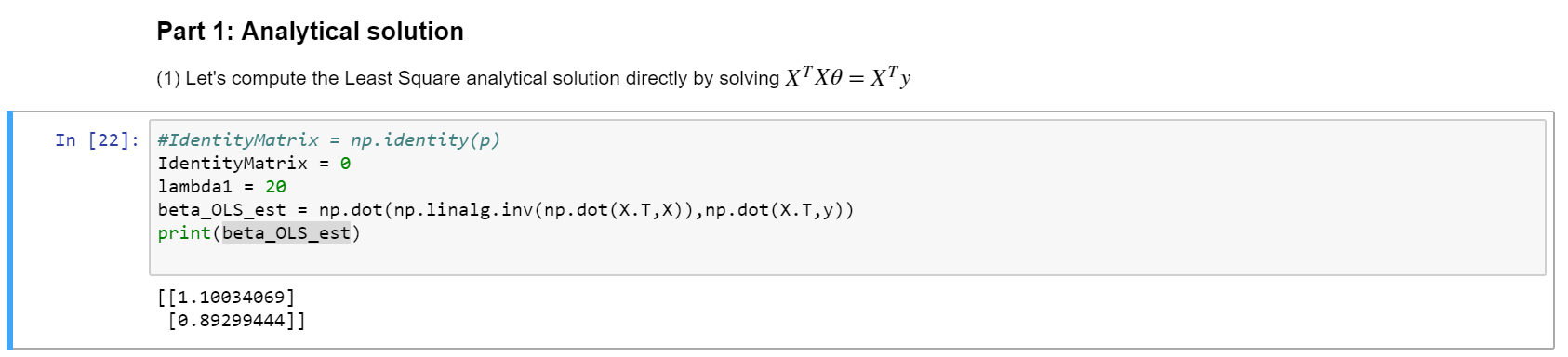
# ASSIGNMENT 2

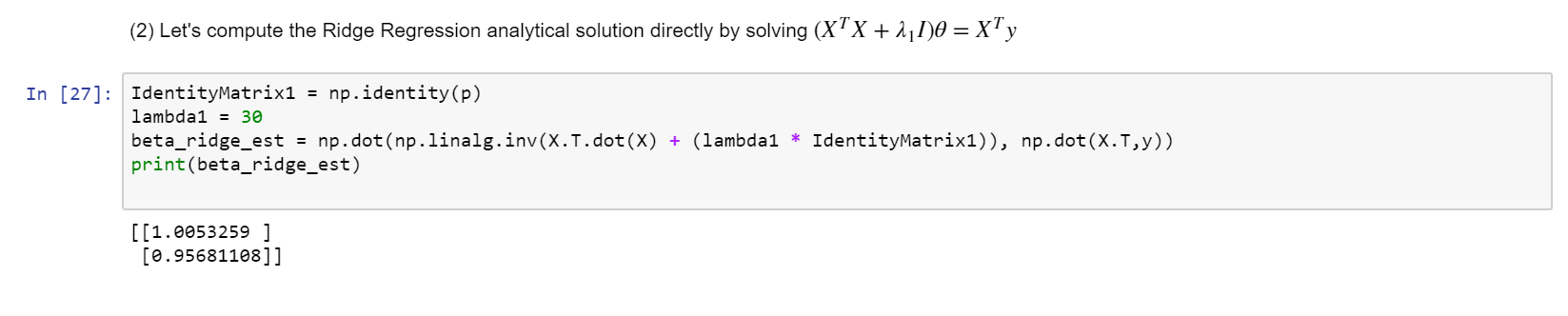
HETUL VARAIYA

1211306106









(b) Let's use Gradient Descent to update parameter until convergence, please try different stepsize to see the Convergence performance

Gradient descent with 1000 iterations. step\_size=0.00000 and initial beta=[-1.00000,-1.50000]

Iteration 0 --- beta:[ -1.00000, -1.50000] --- Cost: 19409.36672

Iteration 100 --- beta:[ -0.28992, -0.78622] --- Cost: 9167.58771

Iteration 200 --- beta:[ 0.19303, -0.29965] --- Cost: 4419.14312

Iteration 300 --- beta:[ 0.52133, 0.03220] --- Cost: 2217.58240

Iteration 400 --- beta:[ 0.74433, 0.25869] --- Cost: 1196.83799

Iteration 500 --- beta:[ 0.89564, 0.41345] --- Cost: 723.55781

Iteration 600 --- beta:[ 0.99813, 0.51936] --- Cost: 504.09982

Iteration 700 --- beta:[ 1.06739, 0.59200] --- Cost: 402.32230

Iteration 800 --- beta:[ 1.11403, 0.64198] --- Cost: 355.10561

Iteration 900 --- beta:[ 1.14527, 0.67654] --- Cost: 333.18552

final beta: [ 1.16602, 0.70058]

Gradient descent with 1000 iterations. step\_size=0.00000 and initial beta=[-1.00000,-1.50000]

Iteration 0 --- beta:[ -1.00000, -1.50000] --- Cost: 19409.36672

Iteration 100 --- beta:[ 0.52446, 0.03533] --- Cost: 2200.68375

Iteration 200 --- beta:[ 1.00011, 0.52133] --- Cost: 500.74653

Iteration 300 --- beta:[ 1.14620, 0.67747] --- Cost: 332.68642

Iteration 400 --- beta:[ 1.18880, 0.72985] --- Cost: 315.94372

Iteration 500 --- beta:[ 1.19895, 0.74955] --- Cost: 314.15388

Iteration 600 --- beta:[ 1.19898, 0.75890] --- Cost: 313.84726

Iteration 700 --- beta:[ 1.19590, 0.76493] --- Cost: 313.69311

Iteration 800 --- beta:[ 1.19192, 0.76984] --- Cost: 313.55972

Iteration 900 --- beta:[ 1.18772, 0.77433] --- Cost: 313.43380

final beta: [ 1.18352, 0.77862]

Gradient descent with 1000 iterations. step\_size=0.00001 and initial beta=[-1.00000,-1.50000]

Iteration 0 --- beta:[ -1.00000, -1.50000] --- Cost: 19409.36672

Iteration 100 --- beta:[ 1.16916, 0.70373] --- Cost: 321.87769

Iteration 200 --- beta:[ 1.19718, 0.76322] --- Cost: 313.73950

Iteration 300 --- beta:[ 1.18351, 0.77863] --- Cost: 313.31349

Iteration 400 --- beta:[ 1.17009, 0.79209] --- Cost: 312.95195

Iteration 500 --- beta:[ 1.15765, 0.80452] --- Cost: 312.64292

Iteration 600 --- beta:[ 1.14616, 0.81601] --- Cost: 312.37875

Iteration 700 --- beta:[ 1.13553, 0.82663] --- Cost: 312.15295

Iteration 800 --- beta:[ 1.12571, 0.83646] --- Cost: 311.95994

Iteration 900 --- beta:[ 1.11663, 0.84554] --- Cost: 311.79495

final beta: [ 1.10823, 0.85393]

Gradient descent with 1000 iterations. step\_size=0.00003 and initial beta=[-1.00000,-1.50000]

Iteration 0 --- beta:[ -1.00000, -1.50000] --- Cost: 19409.36672

Iteration 100 --- beta:[ 1.18349, 0.77867] --- Cost: 313.31257

Iteration 200 --- beta:[ 1.14611, 0.81606] --- Cost: 312.37760

Iteration 300 --- beta:[ 1.11657, 0.84560] --- Cost: 311.79387

Iteration 400 --- beta:[ 1.09322, 0.86894] --- Cost: 311.42943

Iteration 500 --- beta:[ 1.07478, 0.88738] --- Cost: 311.20189

Iteration 600 --- beta:[ 1.06020, 0.90195] --- Cost: 311.05984

Iteration 700 --- beta:[ 1.04869, 0.91346] --- Cost: 310.97114

Iteration 800 --- beta:[ 1.03959, 0.92256] --- Cost: 310.91577

Iteration 900 --- beta:[ 1.03240, 0.92975] --- Cost: 310.88120

final beta: [ 1.02672, 0.93543]

Gradient descent with 1000 iterations. step\_size=0.00010 and initial beta=[-1.00000,-1.50000]

Iteration 0 --- beta:[ -1.00000, -1.50000] --- Cost: 19409.36672

Iteration 100 --- beta:[ 1.10794, 0.85422] --- Cost: 311.64931

Iteration 200 --- beta:[ 1.05203, 0.91012] --- Cost: 310.99472

Iteration 300 --- beta:[ 1.02658, 0.93556] --- Cost: 310.85915

Iteration 400 --- beta:[ 1.01500, 0.94714] --- Cost: 310.83107

Iteration 500 --- beta:[ 1.00973, 0.95241] --- Cost: 310.82526

Iteration 600 --- beta:[ 1.00733, 0.95481] --- Cost: 310.82406

Iteration 700 --- beta:[ 1.00624, 0.95590] --- Cost: 310.82381

Iteration 800 --- beta:[ 1.00574, 0.95640] --- Cost: 310.82375

Iteration 900 --- beta:[ 1.00551, 0.95662] --- Cost: 310.82374

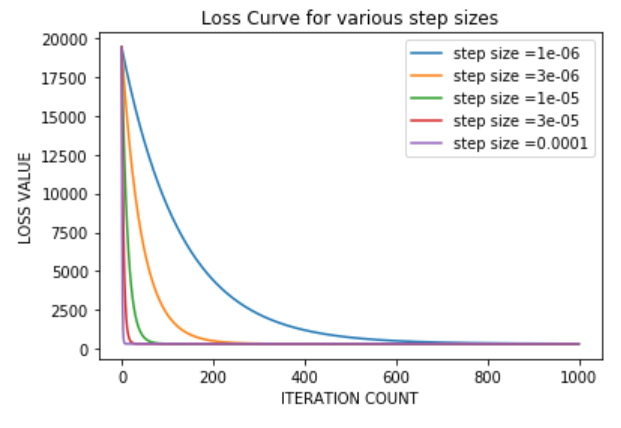
final beta: [ 1.00541, 0.95673]

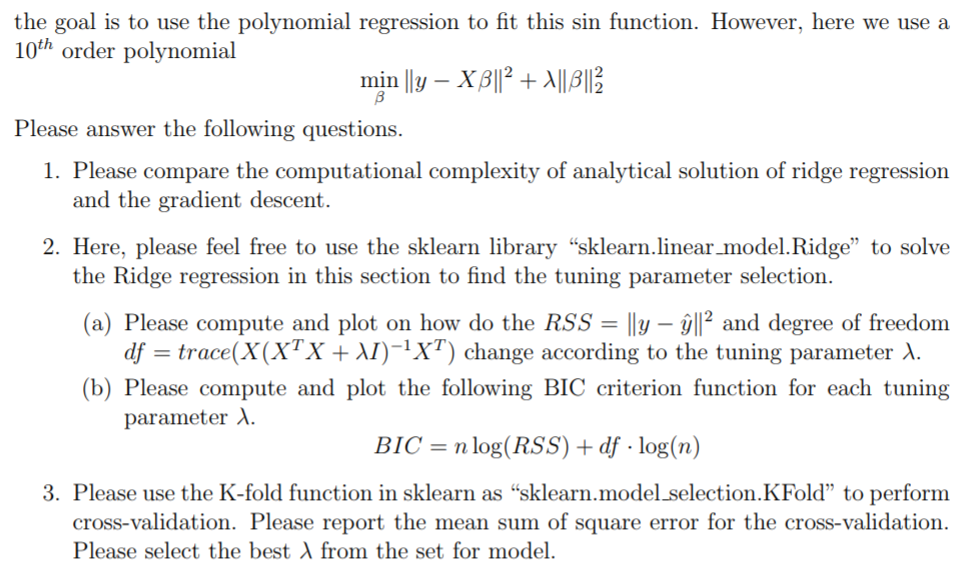
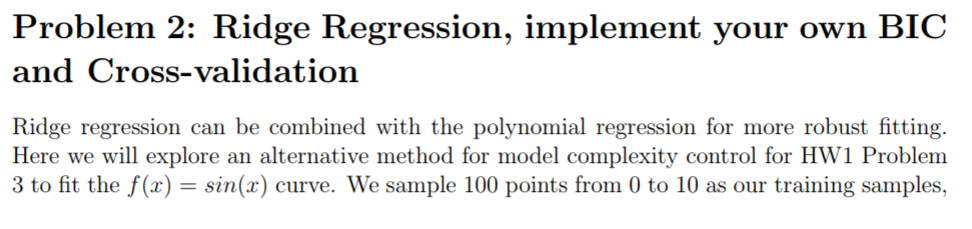
**Conclusion:**

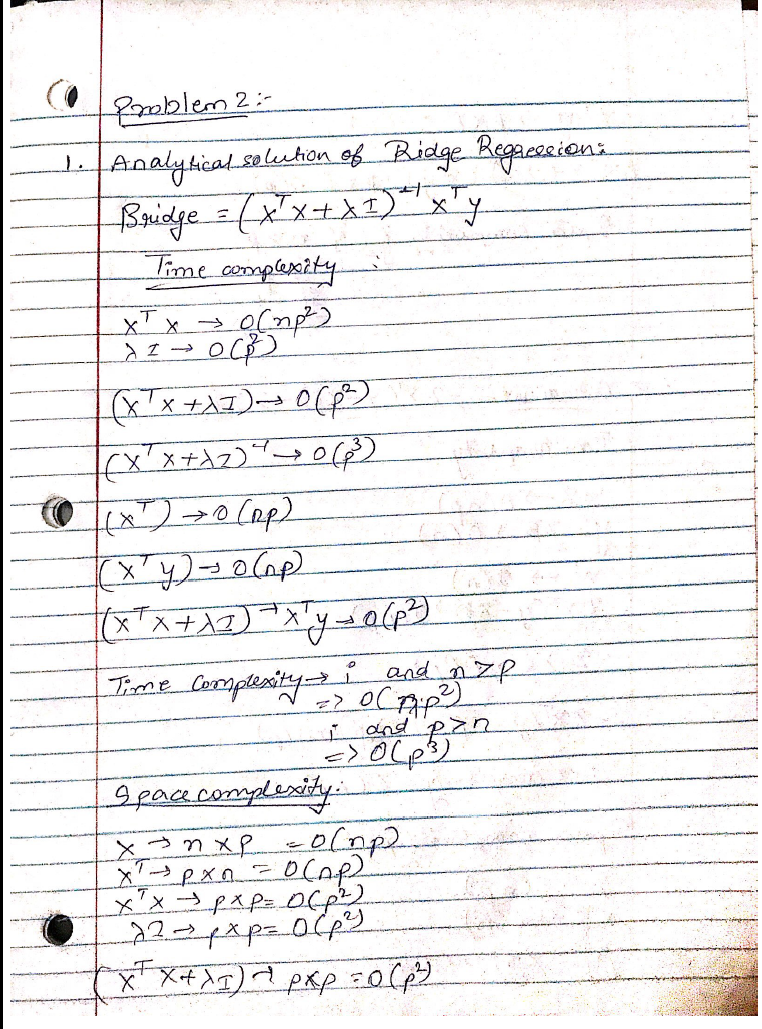
From the above result we can comment about the performance of the convergence that as the step size is increased the number of iterations required to converge to the constant cost decreases.

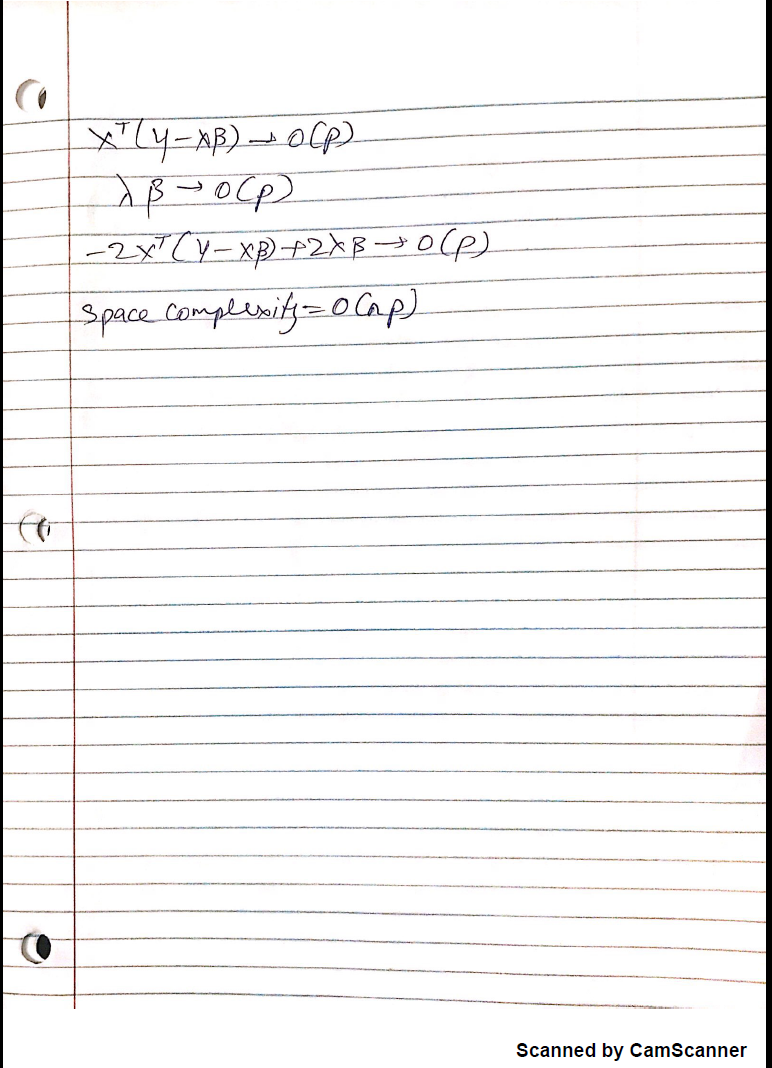
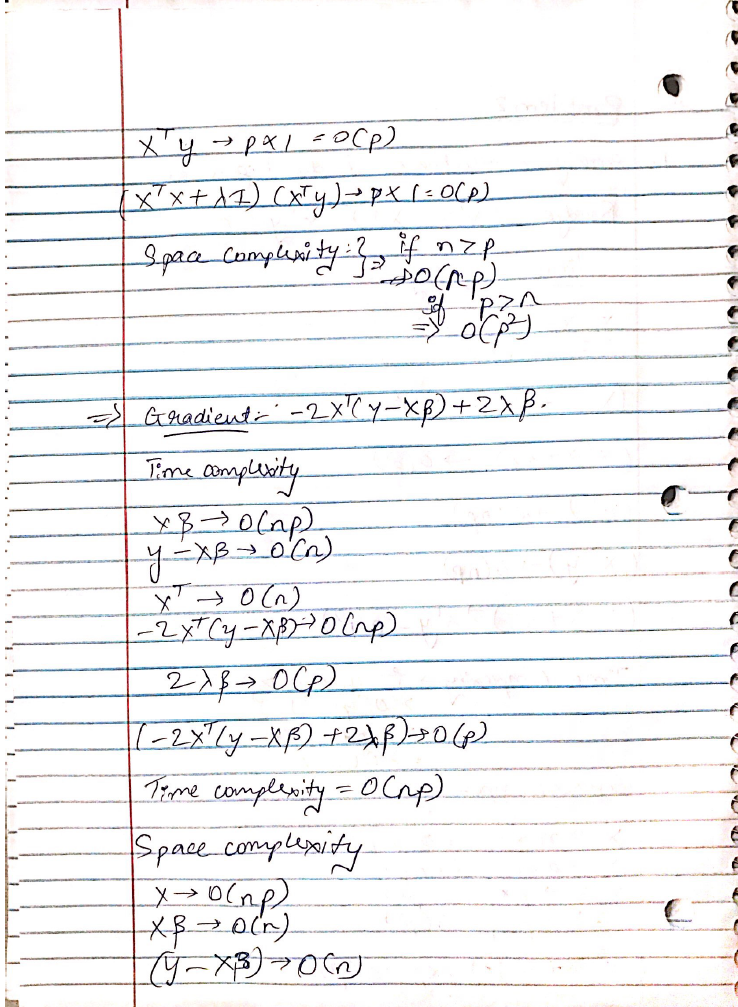
Hence mathematically, we can say step size is directly proportional to the Performance of the convergence.

(c) Normally, it is good to visualize the loss function over time, please plot the loss (cost) function for theta in each iteration.







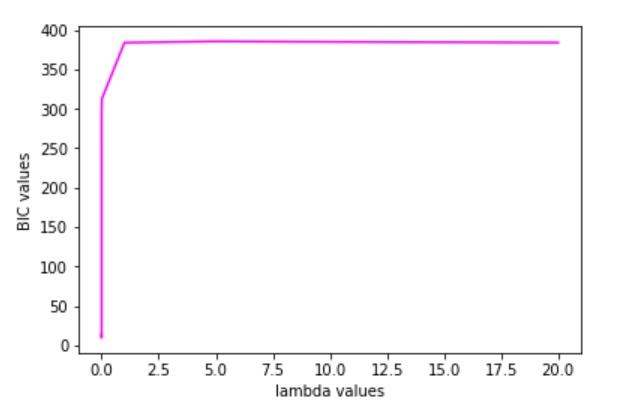


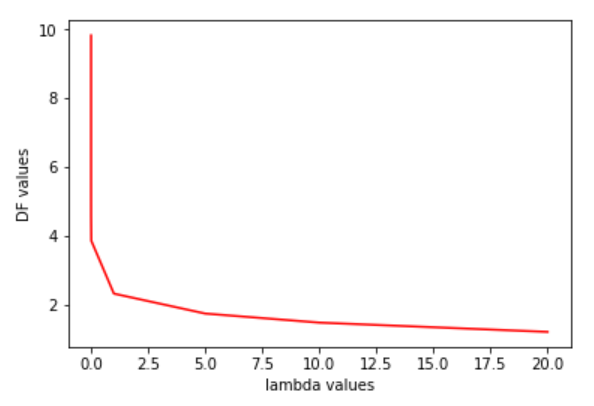


The values displayed above are the corresponding values that increase or decrease as the value of Lambda value increases.

When Lambda values are increased the BIC used as a tuning parameter for selection increases and shoots up drastically when the value of lambda is changed from 5 to 6. The BIC indicates how the predictors will affect the response. Therefore, the increase in BIC values results in the decrease of degrees of freedom as the number of regressors are reduced due to their less significant effect on the response variable. The BIC values stabilize and is highest when the lambda value is 10.

The above explanation is explained by the plot shown on the next page.

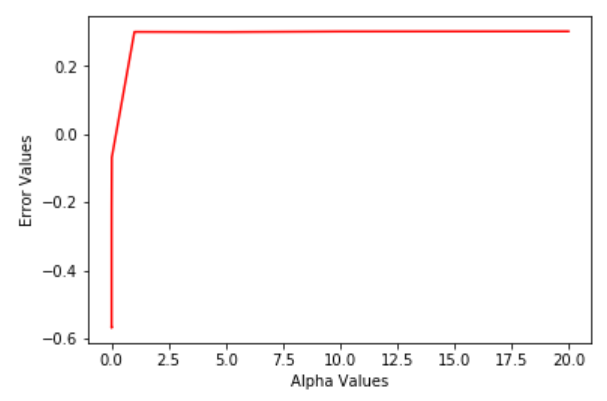


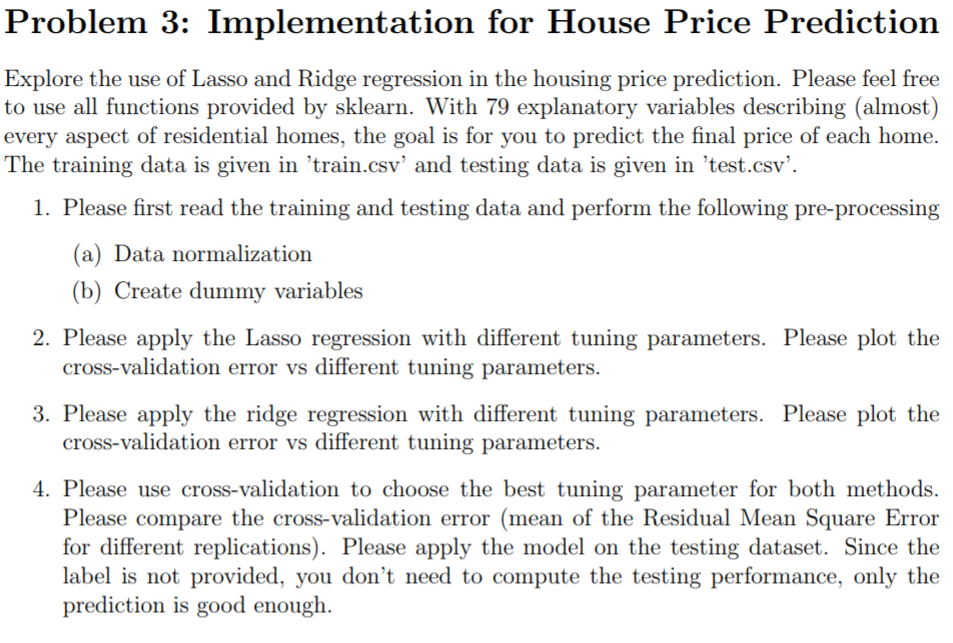




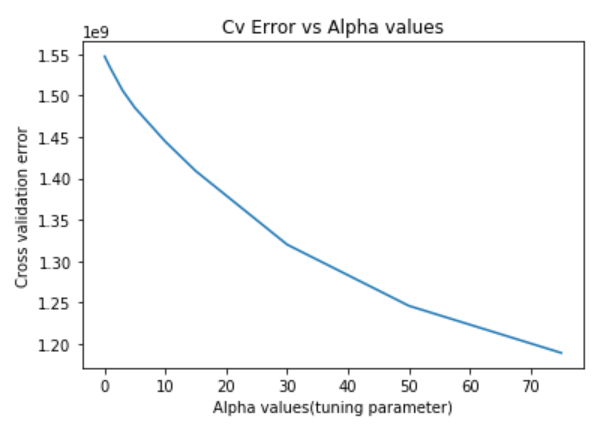
The minimum MSE value is: -0.5681743544671313

The below plot shows the effect alpha values on the error.

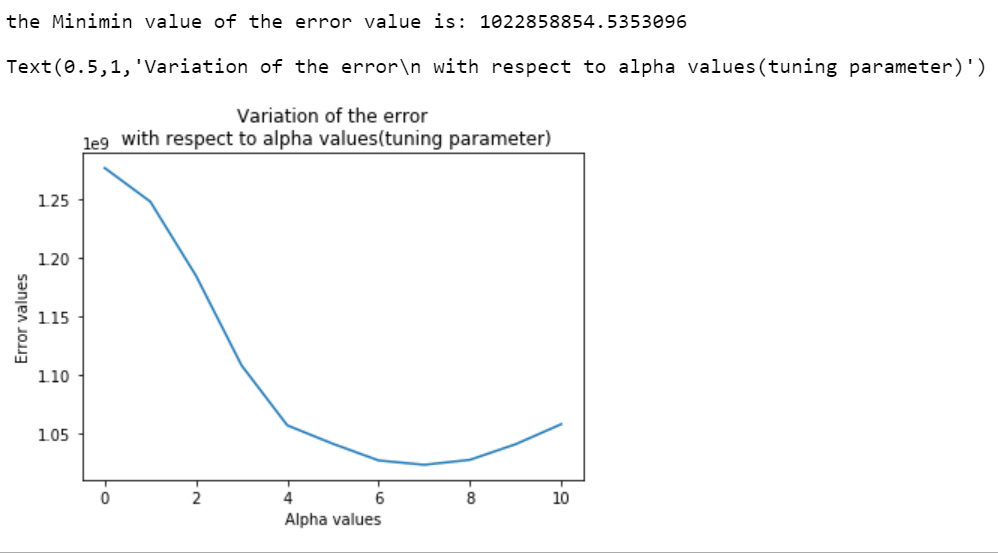




2. Model 2. Cross validation error vs. Alpha values (tuning parameters) for the model using LASSO regression

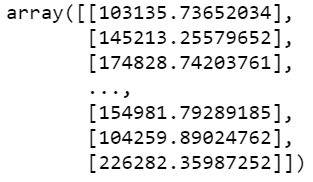


3. Model 1. Cross validation error vs tuning parameters for the model with Ridge regression:



4. Applying the fitted models to predict the response variable values:

Predicted values for Ridge regression:



Predicted Values for LASSO regression:

