

SMAI ASSIGNMENT- 7 REPORT

MARCH 11

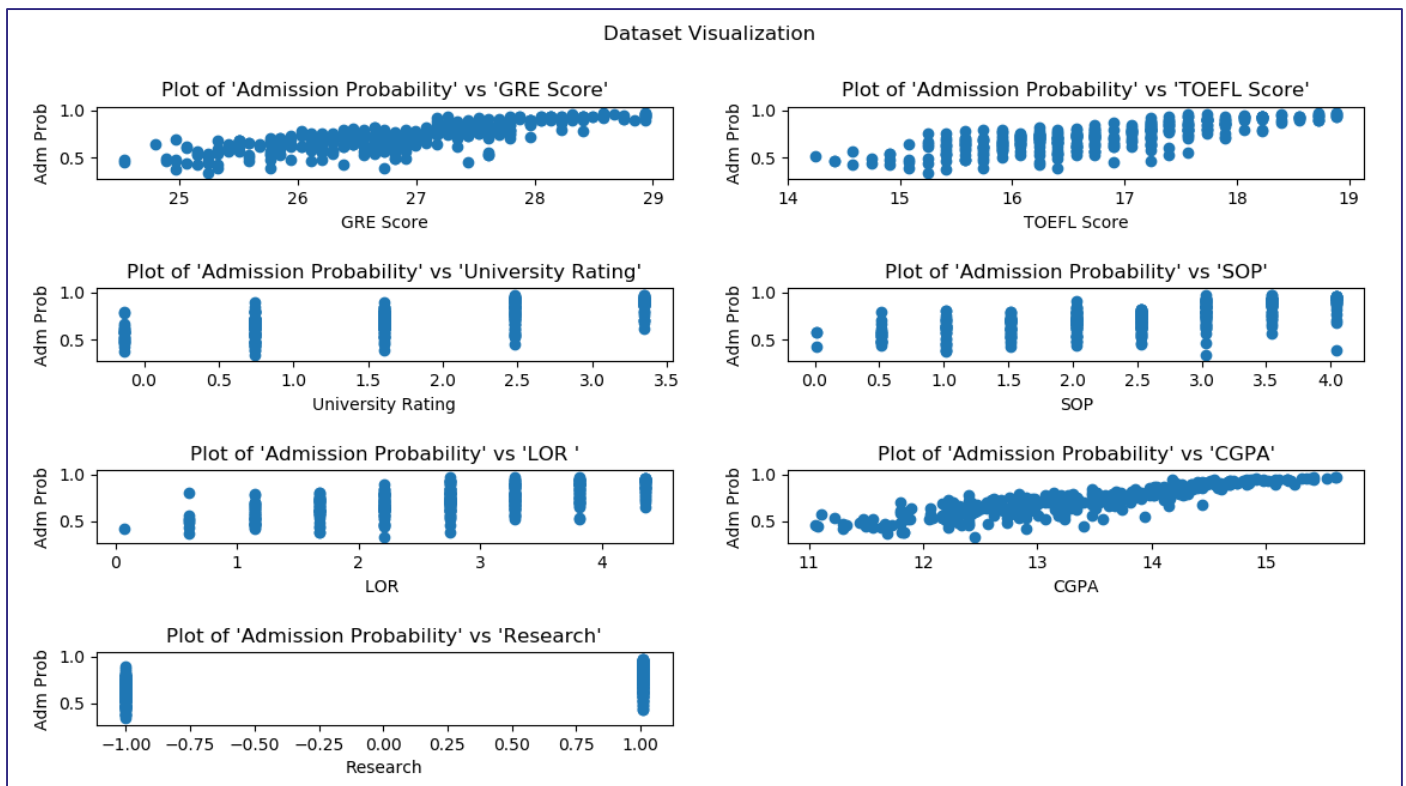
AYUSH KUMAR DWIVEDI
2018802002



Regularization and Cross Validation

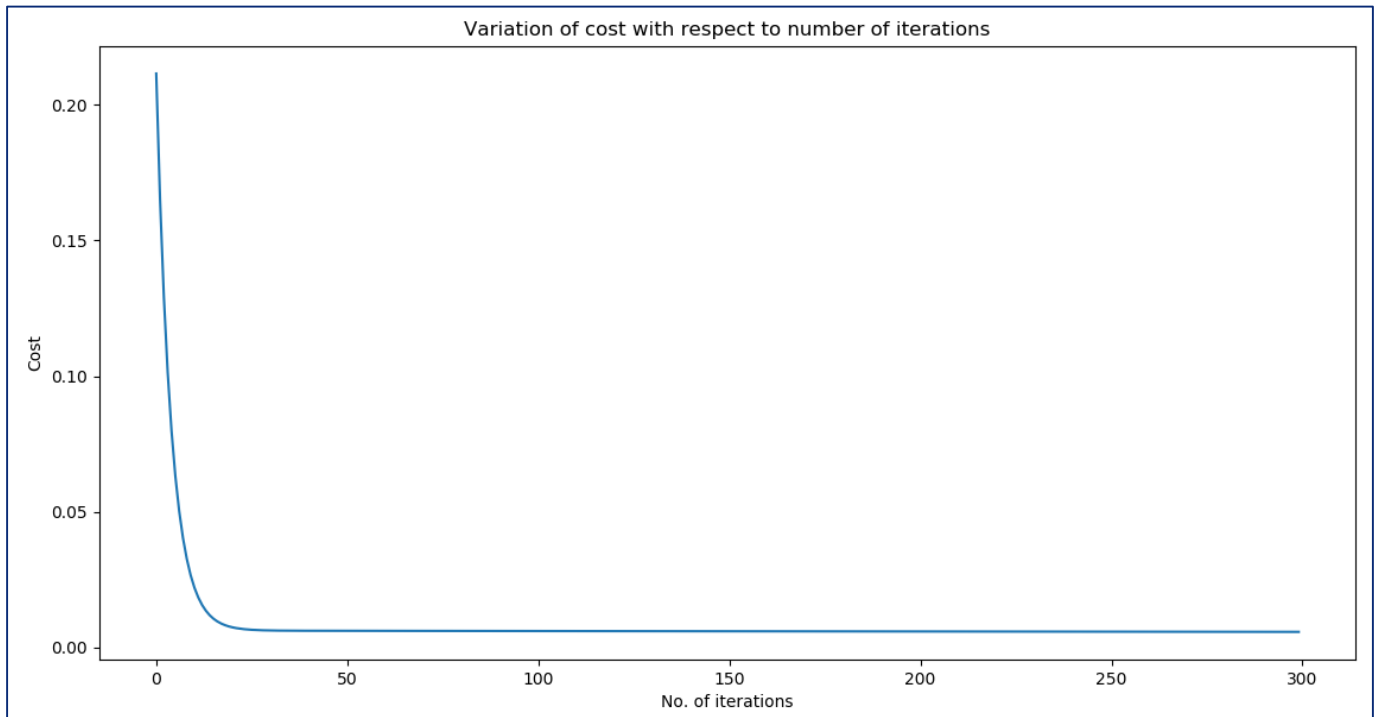
Data Visualization:

The following is the plot made for visualize the data before performing any other task:



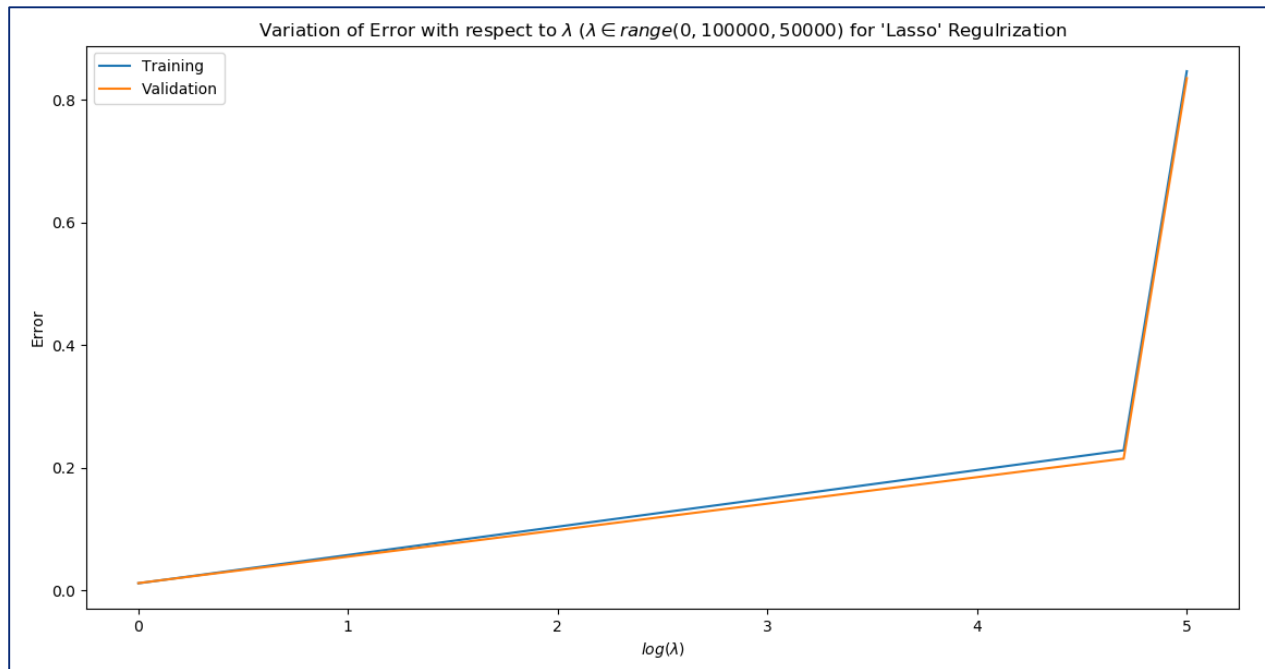
Checking if the Gradient Descent is Working properly or not ?

The following is the plot made for check the proper functioning of Gradient Descent:

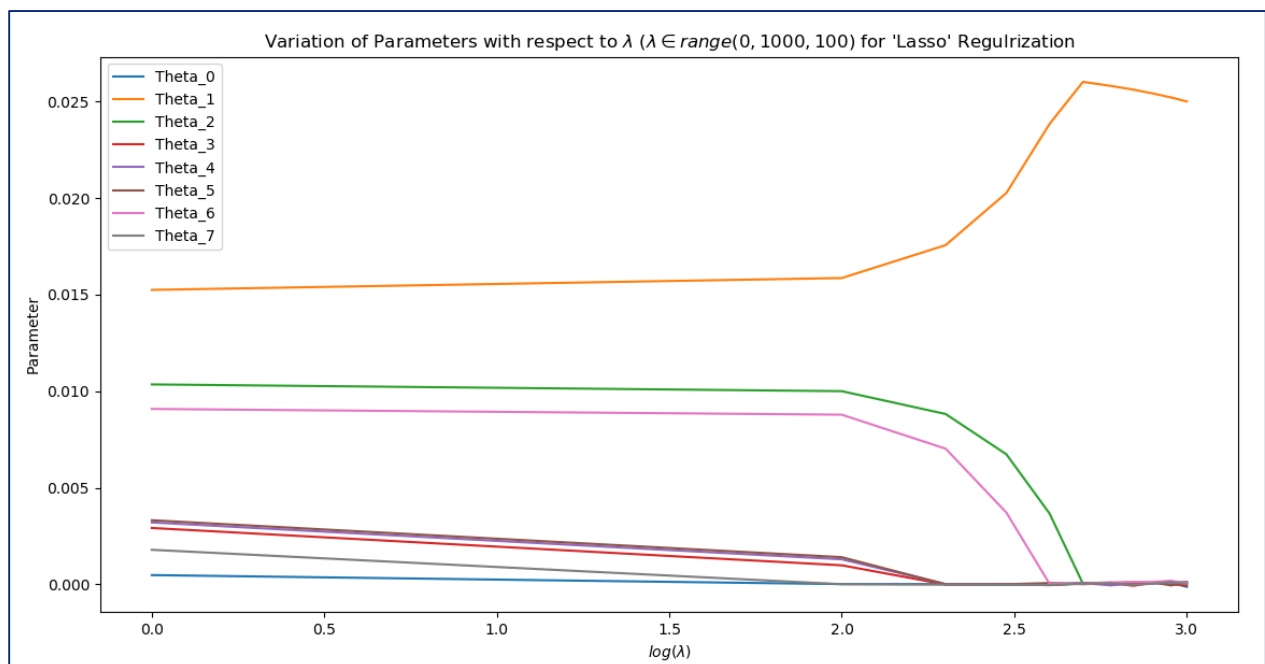


Answer 1: LASSO

I implemented Lasso regularization and got the following graph showing the variation of error w.r.t Lambda:

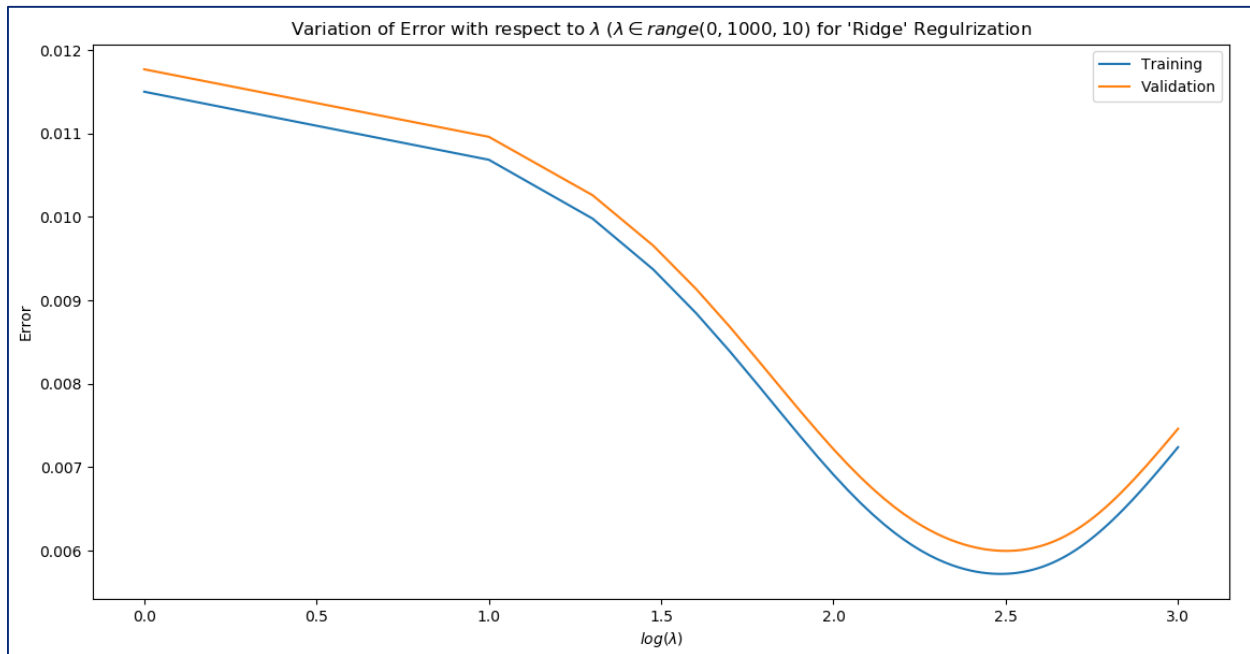


I also plotted the variation of various Parameters w.r.t Lambda. As expected, the weighted are penalized more and more as Lambda value increases and moves towards converging to zero.

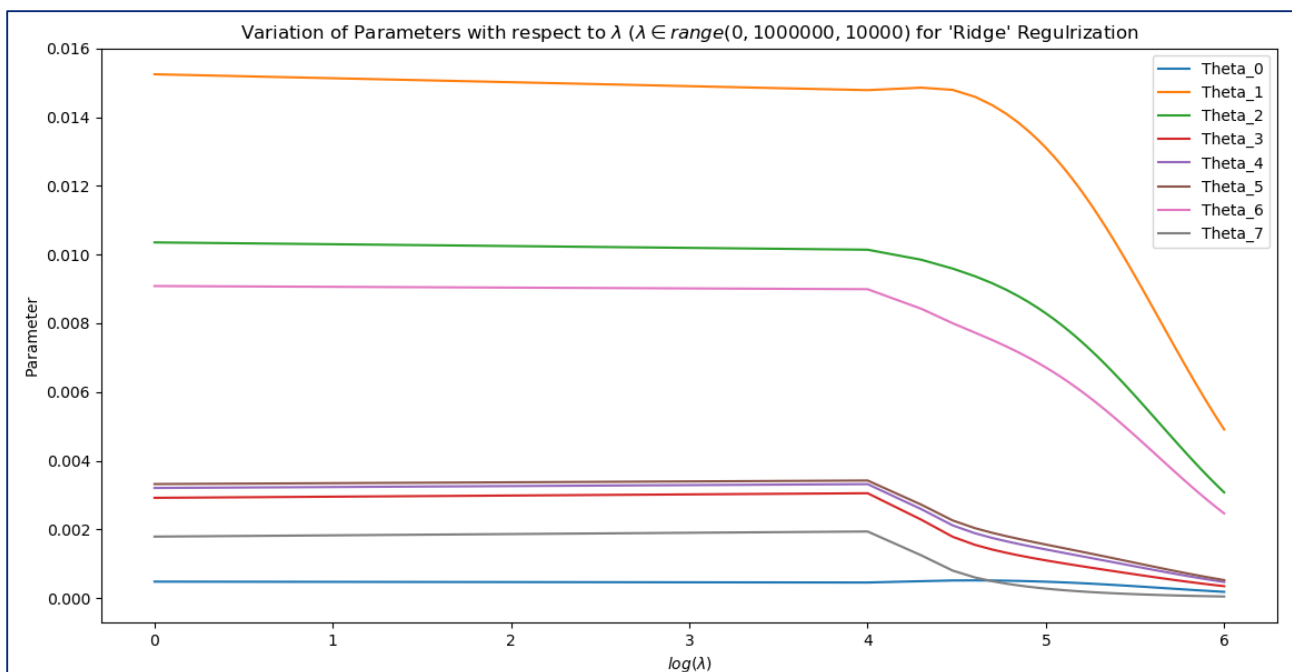


Answer 2: Ridge

I implemented Ridge regularization and got the following graph showing the variation of error w.r.t Lambda:



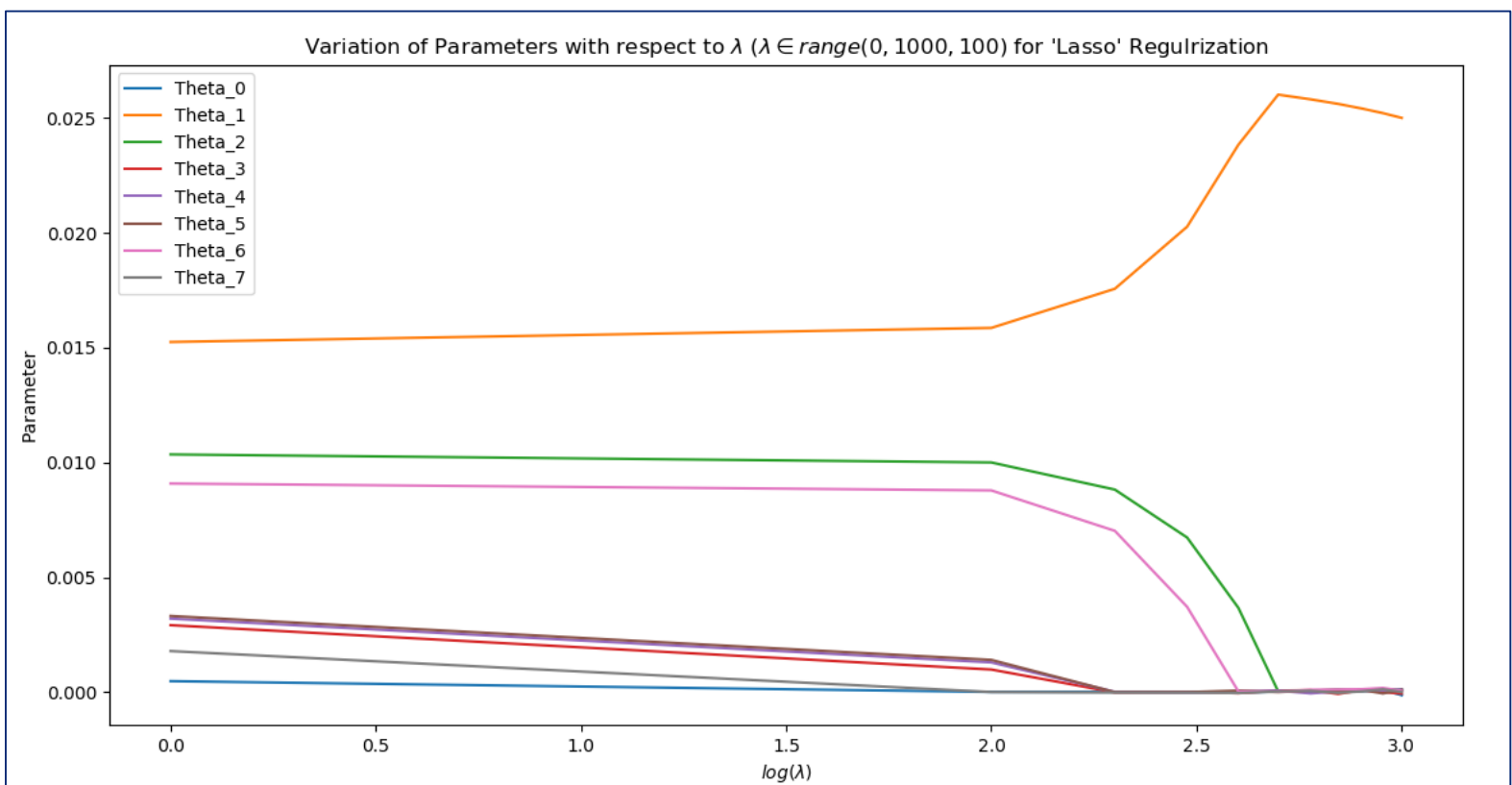
I also plotted the variation of various Parameters w.r.t Lambda. As expected, the weighted are penalized more and more as Lambda value increases and moves towards converging to zero.

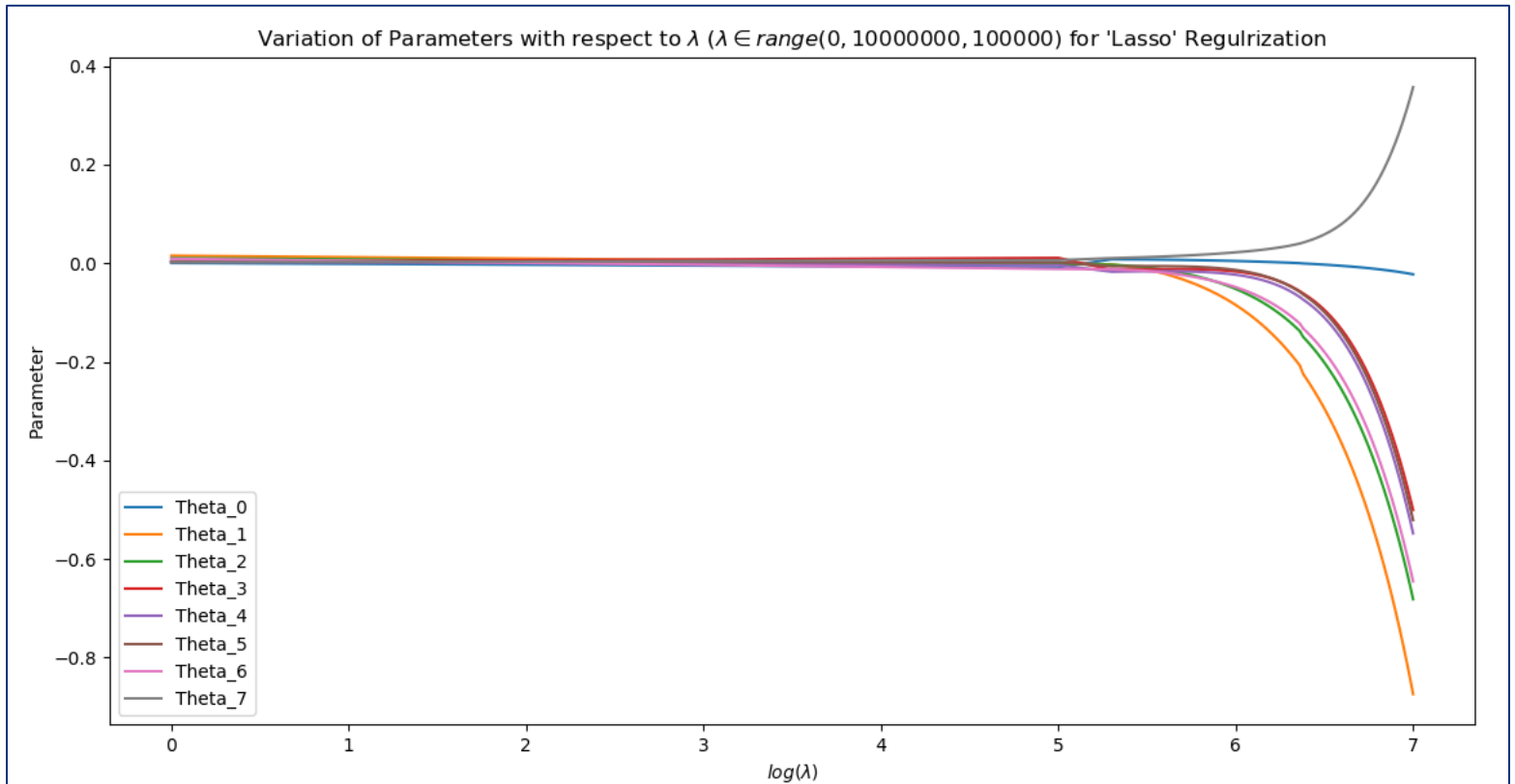
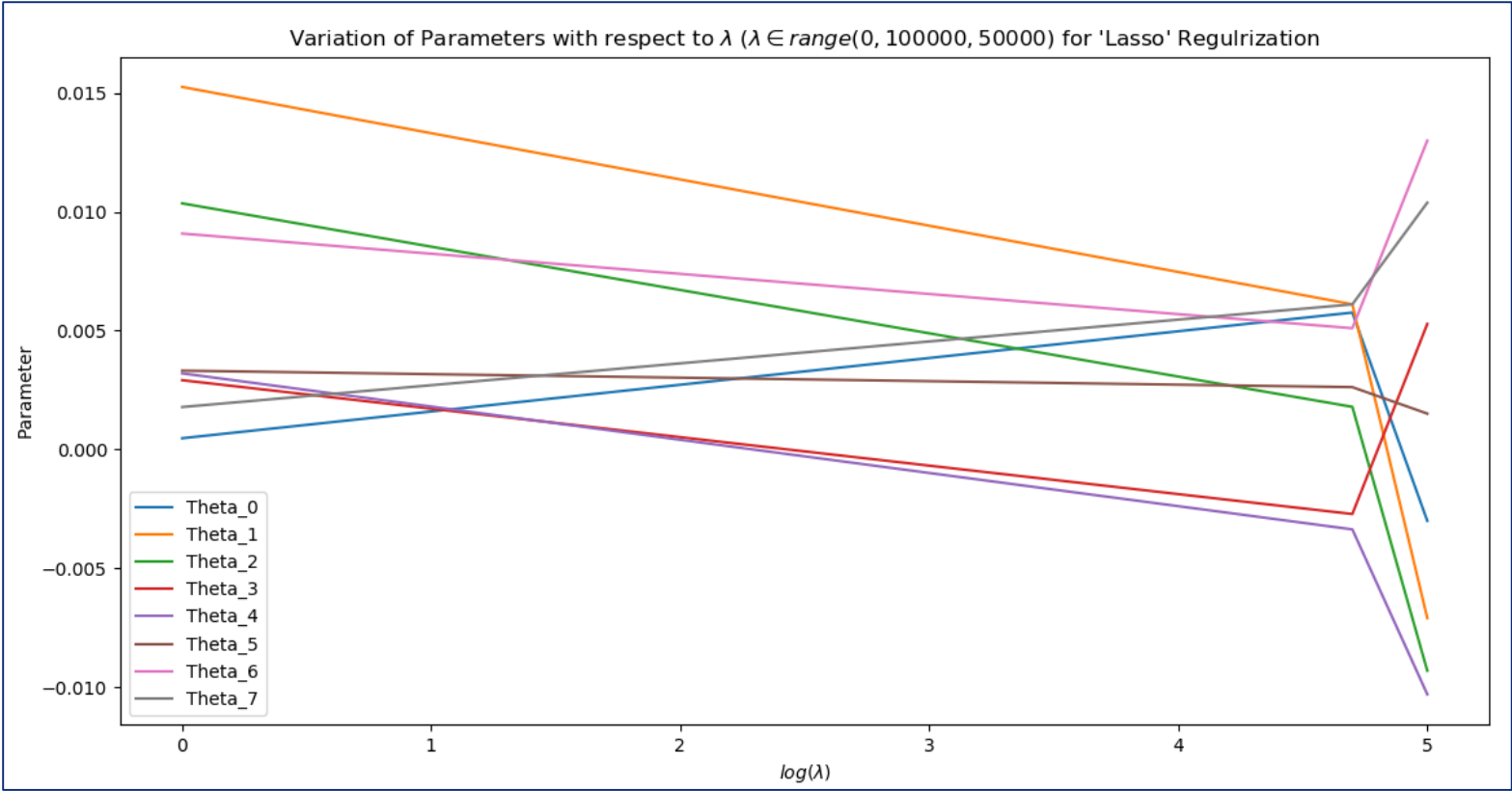


Answer 3: Role of Lambda

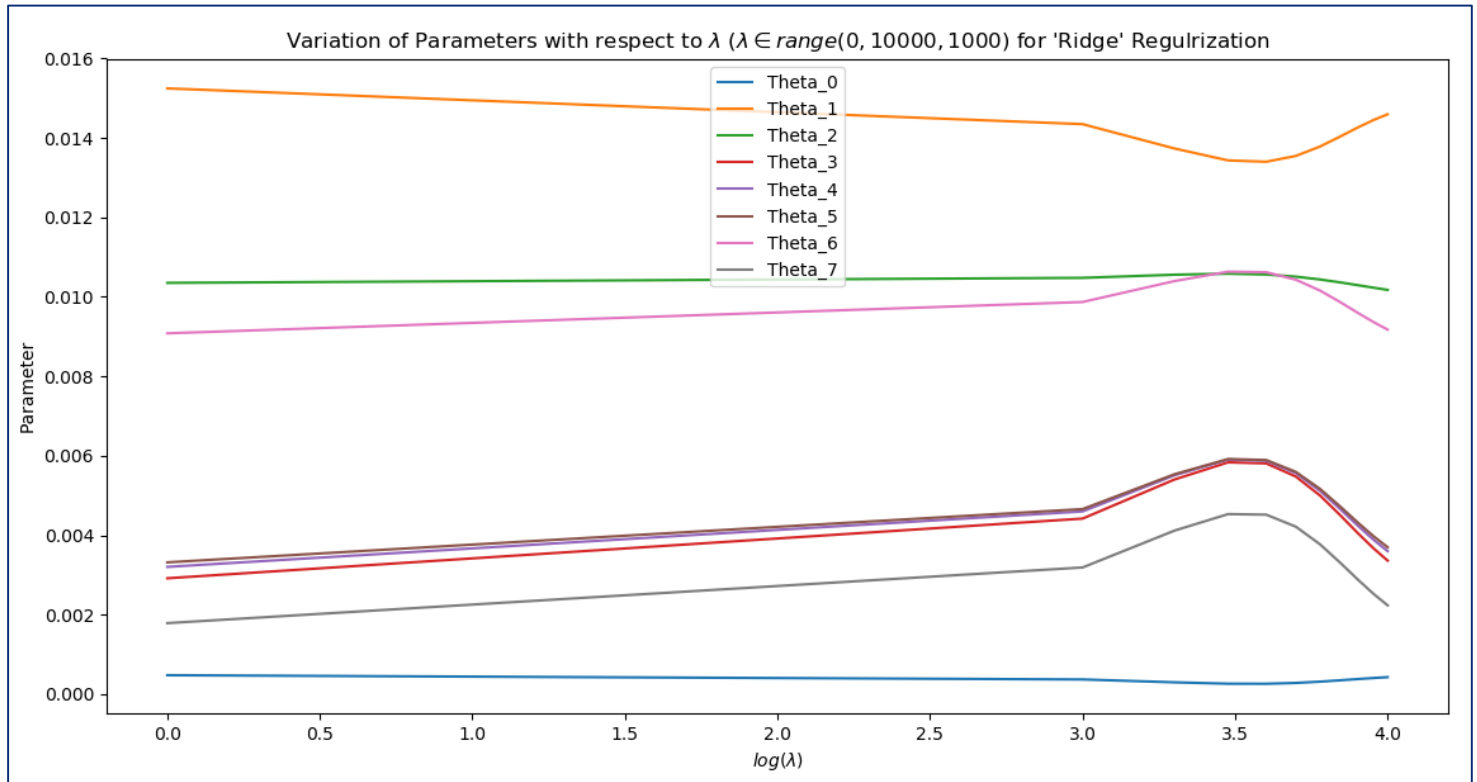
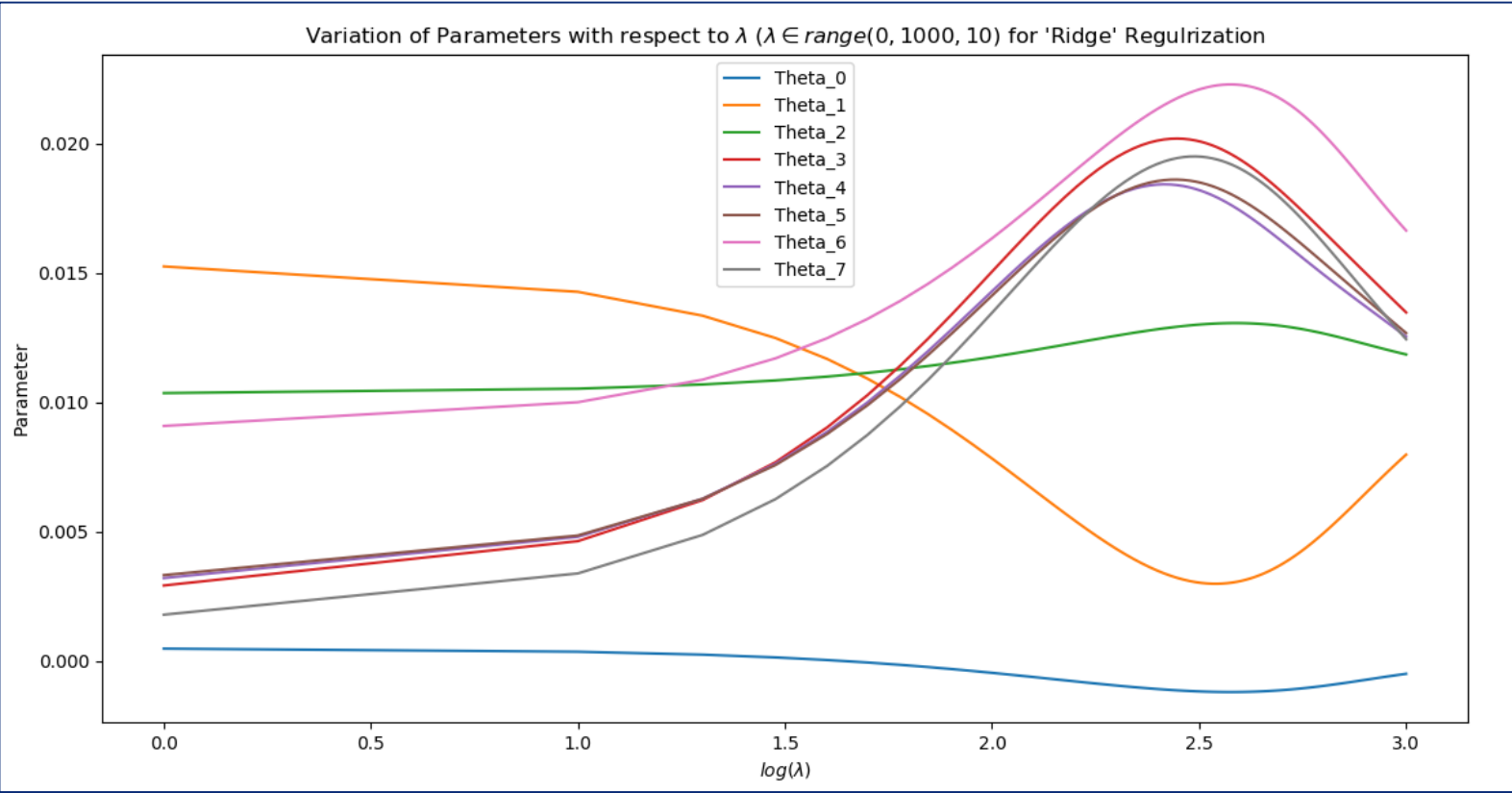
As lambda increases, the weights are penalized more and more, the models tend towards becoming more simple and the variation is decreased but with an increase in bias.

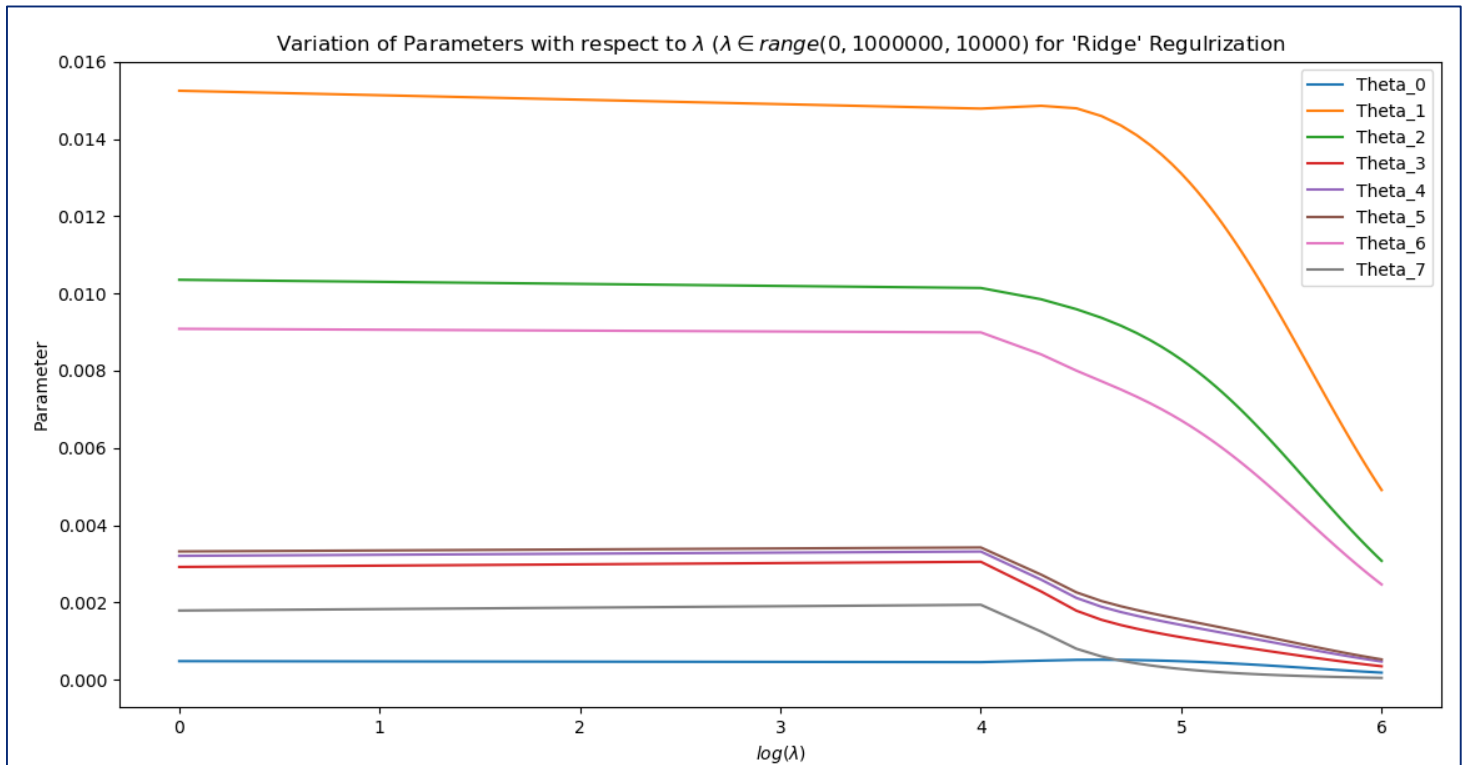
The decrease in parameters, tending towards zero with increase in Lambda values can be vied from following plots.





For Ridge:





Answer 4: Analysis of two different types of regression techniques

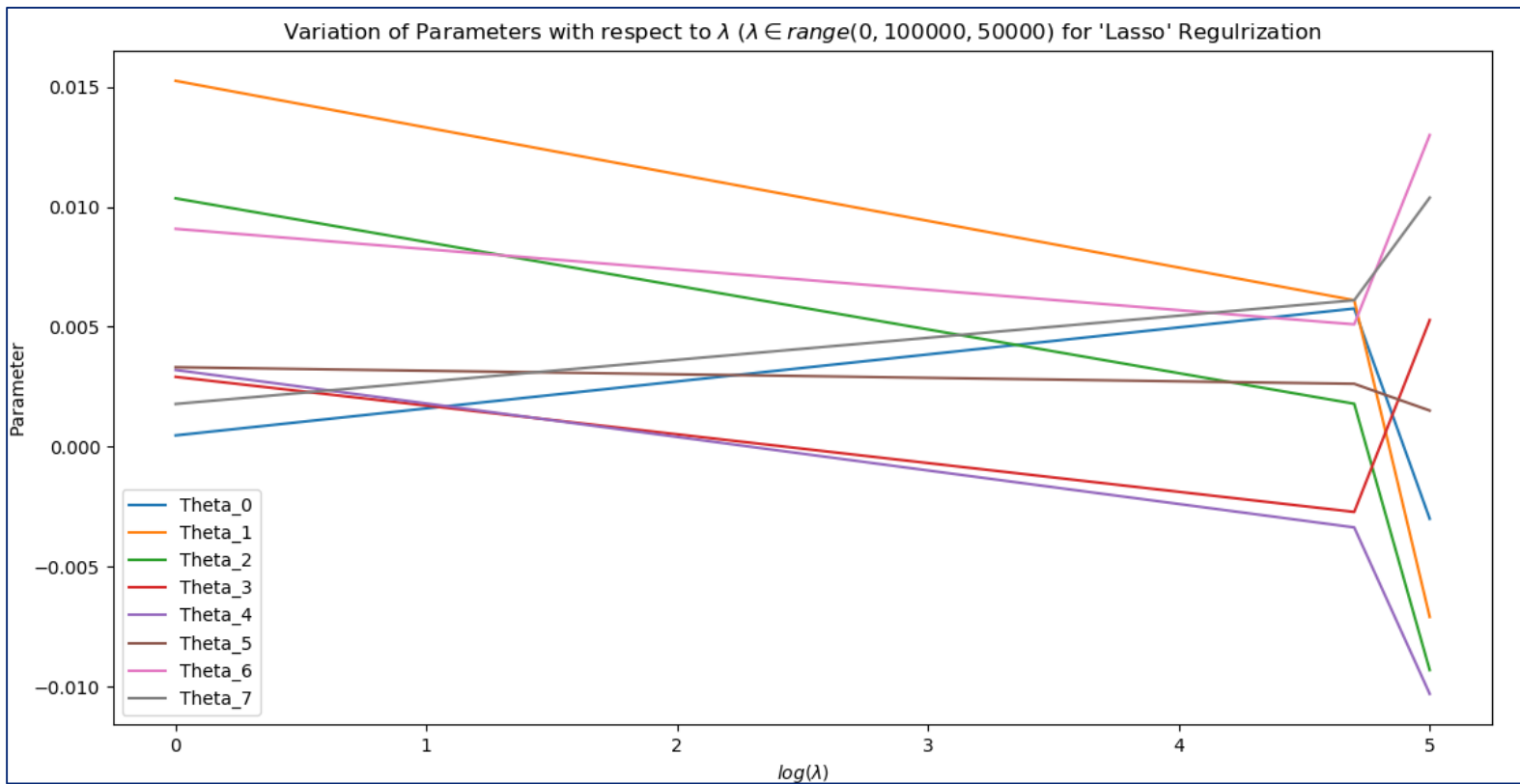
Ridge:

Instead of forcing the parameters to be exactly zero, it penalizes them if they are too far from zero, thus enforcing them to be small in a continuous way. This way, we decrease model complexity while keeping all variables in the model. This, basically, is what Ridge Regression does.

Lasso:

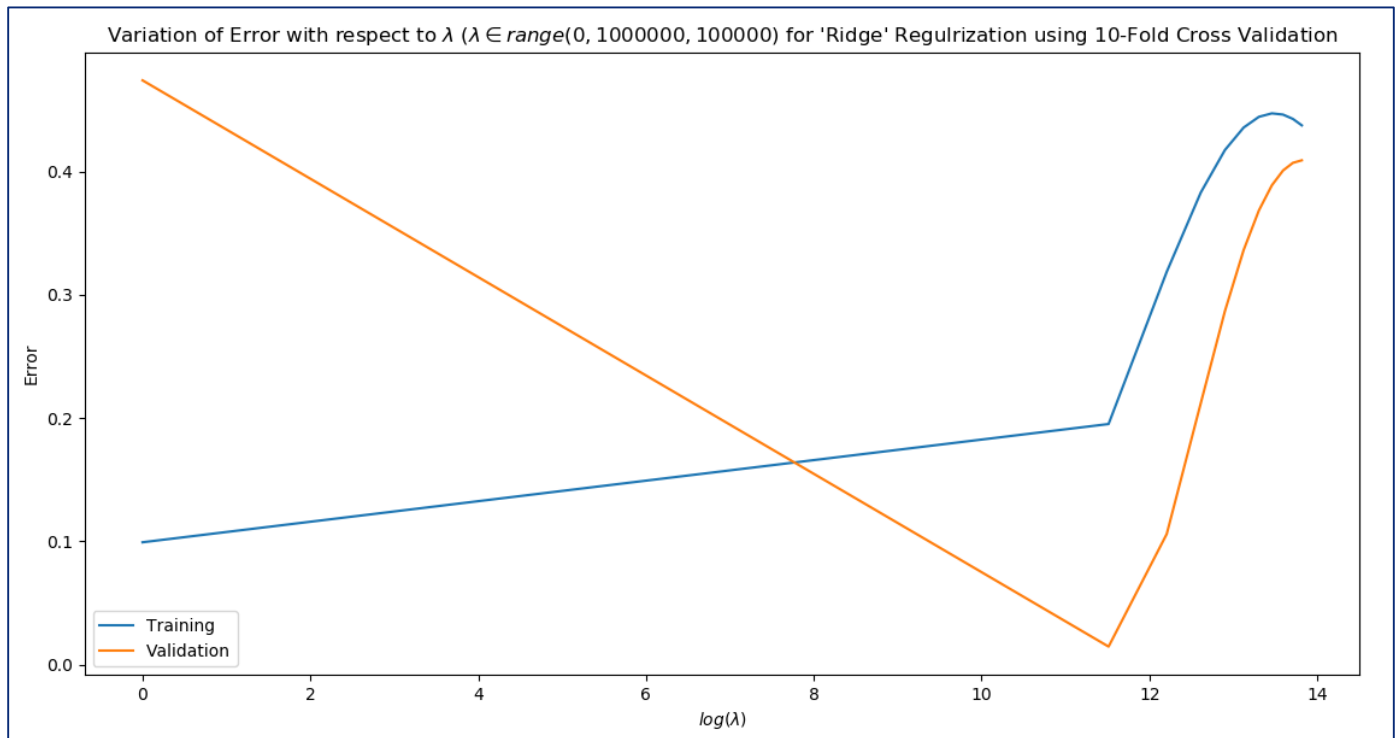
Lasso is quite similar conceptually to ridge regression. It also adds a penalty for non-zero coefficients, but unlike ridge regression which penalizes sum of squared coefficients (the so-called L2 penalty), lasso penalizes the sum of their absolute values (L1 penalty). As a result, for high values of λ , many coefficients are exactly zeroed under lasso, which is never the case in ridge regression.

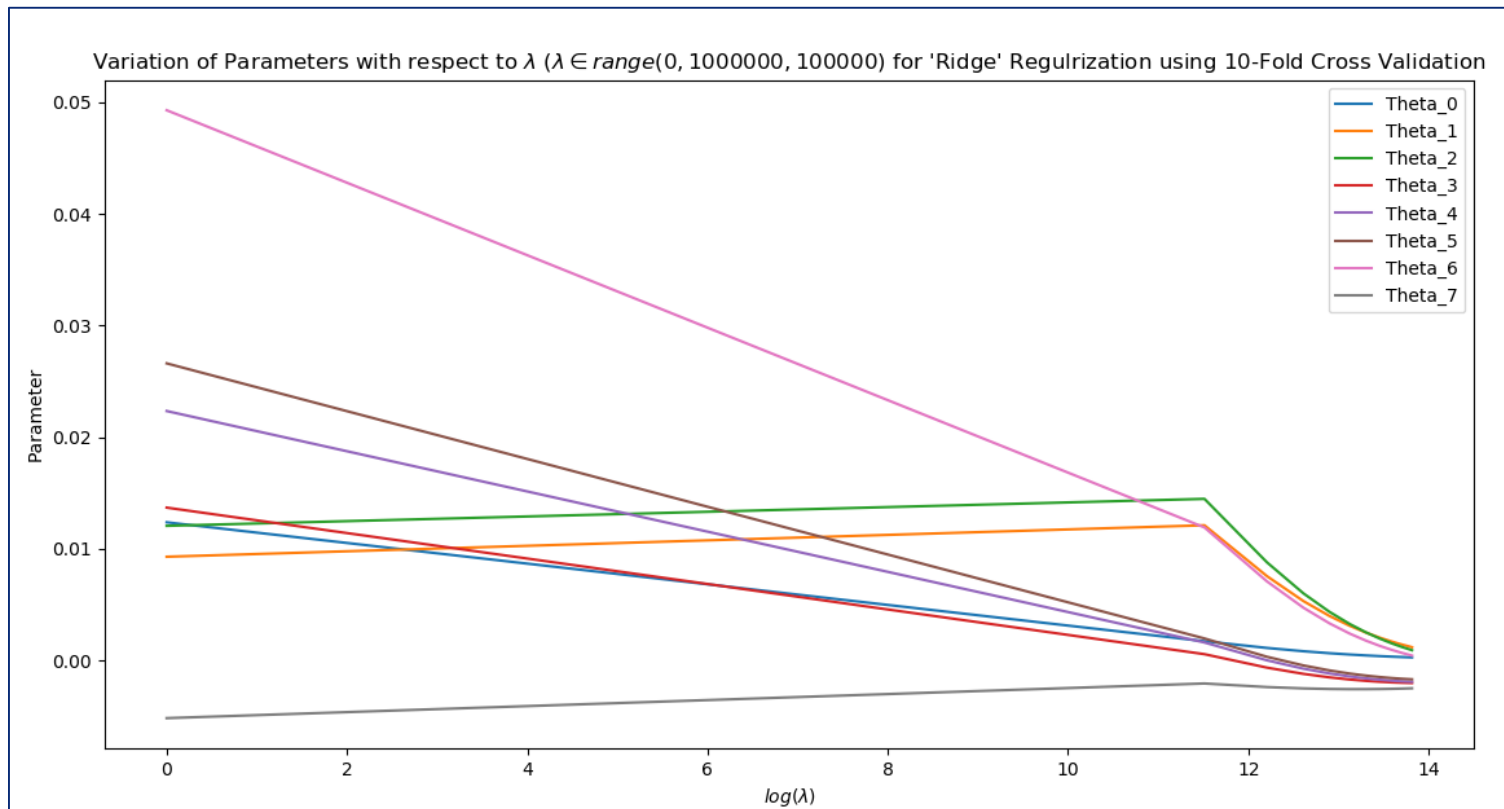
Lasso generates very sparse weights:



Answer 5: Cross Validation

I implemented both k-Fold Cross Validation with K=10 and also Leave One Out Cross validation





To check proper function of GD, I plotted the error vs iters curve for all the different values of lambda after performing cross validation:

