

Question 1:

What is the optimal value of alpha for ridge and lasso regression? What will be the changes in the model if you choose to double the value of alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?

Answer to Q1:

Alpha for ridge is ~ 10 and for Lasso is ~ 50 .

If we increase the Alpha, the constraint will be a lot more, we impose more constraint on the model. If the Alpha is more the Model become underfit model.

After increasing the Alpha, the coefficient variables are reduced. Even though the coefficients are reduced Model will be strong.

Question 2:

You have determined the optimal value of lambda for ridge and lasso regression during the assignment. Now, which one will you choose to apply and why?

Answer to Q2:

Which will give the highest and best accuracy. Lasso regression will try to reduce the coefficients, Lasso regression will do more better when number of significant predictors are very less.

Question 3:

After building the model, you realised that the five most important predictor variables in the lasso model are not available in the incoming data. You will now have to create another model excluding the five most important predictor variables. Which are the five most important predictor variables now?

Answer to Q3:

Which are variables having the highest coefficients are the most important predictors. We remove the top five variables having the highest coefficients from lr coefficients results and run the lr coefficients again to get the next top five variables.

Question 4:

How can you make sure that a model is robust and generalisable? What are the implications of the same for the accuracy of the model and why?

Answer to Q4:

If alpha is zero, then the model is overfitting.

If the alpha is higher, means the more regularization, than model becomes underfitting and if alpha is lower means lesser the regularization than the model will be overfitting.