Question 1

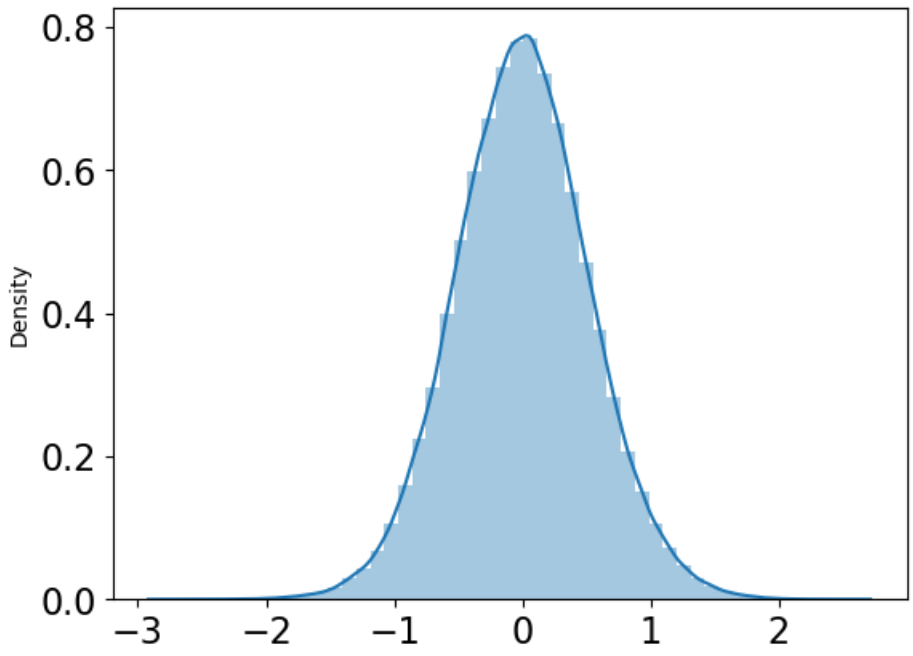
What is the optimal value of alpha for ridge and lasso regression? What will be the changes in the model if you choose double the value of alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?

Answer:

The change will be in R2 Square. Ridge looks better in this case.

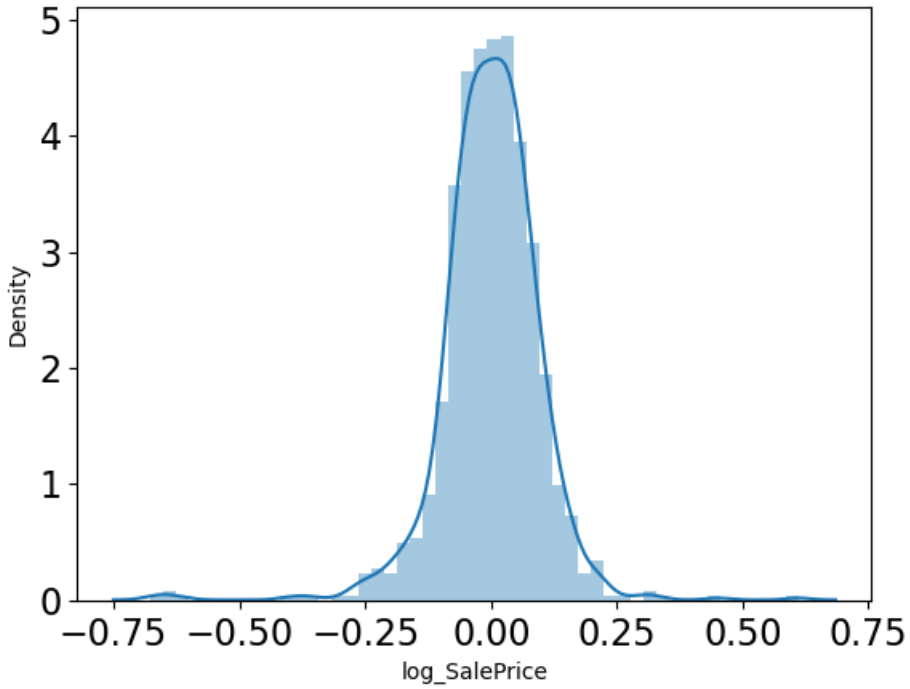
Ridge = 500

Train R2 Square : 0.91
Test R2 Square : 0.89



Lasso = 0.001

Train R2 Square : 0.94
Test R2 Square : 0.86



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:

I will go for Ridge because the train and test data R2 square is not having difference much.

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?

GrLivArea

RoofMatl_Tar&Grv

OverallCond

GarageQual_Fa

GarageCond_Fa