

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?

Optimal value of lamda is 5.0 in ridge and 0.0002 in lasso.
Slight increase in mean square error increased in both regression.
Predictor variable changed to

Features	Coefficient
LotArea	0.0375
OverallQual	0.1162
OverallCond	0.0369
BsmtQual	0.0335
BsmtFinSF1	0.0509
GrLivArea	0.1219
GarageArea	0.0460
MSZoning_FV	0.0550
MSZoning_RL	0.0514
MSZoning_RM	0.0039

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?

The Mean Squared error in case of Ridge and Lasso are:

Ridge - 0.01482

Lasso - 0.01457

The Mean Squared Error of Lasso is slightly lower than that of Ridge

Also, since Lasso helps in feature reduction (as the coefficient value of one of the feature became 0), Lasso has a better over Ridge.

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?

the factors that generally affect the price are the Zoning classification, Living area square feet, Overall quality and condition of the house, Foundation type of the house, Foundation stone and neighborhood.

If we drop these columns accuracy of the model decreases

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?

To make model robust and generalisable 3 features are required:

1. Model accuracy should be $> 70-75\%$
2. P-value of all the features is < 0.05
3. VIF of all the features are < 5

Thus we are sure that model is robust and generalisable.