

Gokulnath Advanced Regression - Part 2

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 optimal values for alpha for ridge and lasso regression thought to us in the module are -

```
params = {'alpha': [0.0001, 0.001, 0.01, 0.05, 0.1,
                    0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 2.0, 3.0,
                    4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0, 20, 50, 100, 500, 1000 ]}
```

When alpha decreases, model becomes simpler and vice versa

Most important predictor variable turns to be `RoofMatl`

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:

Optimal value for alpha ranges would be 5.0

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:

5 most important variables would be -

	Linear	Ridge	Lasso
MSSubClass	-2.809523e-03	-0.012368	-0.0
LotFrontage	-1.868274e-03	-0.027193	-0.0

	Linear	Ridge	Lasso
LotArea	1.764853e-02	-0.002910	0.0
OverallQual	2.159026e-01	0.028435	0.0
OverallCond	8.321607e-02	0.074484	0.0

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:

We can apply the regularization technique lasso and ridge regression and compare the R2 score, RSS and RMSE to confirm if they are optimal for our model.