

Subjective Questions:

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?

The optimal Alpha for Both Ridge and Lasso is:

For Ridge:10

For Lasso:0.001

The R2 Value and RMSE of both ridge and Lasso is

For Ridge:

R2 score (train) : 0.9136672341521263

R2 score (test) : 0.8917121671352861

RMSE (train) : 0.11517486054214349

RMSE (test) : 0.13676364734434512

For Lasso:

R2 score (train) : 0.9124710628019927

R2 score (test) : 0.8948549391422506

RMSE (train) : 0.11597001019042258

RMSE (test) : 0.13476443089321966

If we Increase the Alpha by Double

	Ridge (alpha=9.0)	Lasso (alpha=0.0001)	Ridge (alpha = 18.0)	Lasso (alpha = 0.0002)
MSSubClass	-0.011661	-0.008571	-0.010957	-0.011847
LotFrontage	0.012396	0.010040	0.012641	0.011596
LotArea	0.033427	0.031834	0.033401	0.033196
OverallQual	0.079905	0.083460	0.079999	0.080197
OverallCond	0.048175	0.046879	0.047136	0.049136
YearBuilt	-0.034869	-0.036322	-0.033172	-0.037313
YearRemodAdd	-0.014089	-0.014019	-0.014716	-0.013380
BsmtExposure	0.014686	0.013437	0.014500	0.014710
BsmtFinSF1	0.049503	0.048072	0.049470	0.049080
BsmtFinType2	0.007958	0.007319	0.007981	0.007742
BsmtUnfSF	0.028459	0.024747	0.028274	0.027910
CentralAir	0.010604	0.009992	0.010739	0.010305
1stFlrSF	0.104704	0.111633	0.100796	0.110756
2ndFlrSF	0.100888	0.104775	0.096151	0.107350
BsmtFullBath	0.025992	0.024223	0.025618	0.026194
HalfBath	0.014837	0.012690	0.015425	0.013641
KitchenQual	0.014325	0.014006	0.015043	0.013594
TotRmsAbvGrd	0.013216	0.007411	0.015969	0.008659
Functional	-0.024278	-0.024576	-0.024022	-0.024609
Fireplaces	0.019824	0.019201	0.020396	0.019070
GarageFinish	0.011346	0.009122	0.011481	0.010732

The Above Feature after doubled the Alpha value

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?

Ans: While I building the Model I got Eactly Same Answer for Both Ridge and Lasso, Eventhough I got Same answer for Both Model I go for Lasso because Lasso explains which are the Predictors are Not important for Model

i.e., It makes some variables as 0, according to the Occams razors Principle choose which one is simpler., For building model Lasso is simpler

After Building the Model I end up with:

For Ridge :

R2 score (train) : 0.9136672341521263

R2 score (test) : 0.8917121671352861

For Lasso:

R2 score (train) : 0.9124710628019927

R2 score (test) : 0.8948549391422506

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?

Ans:After Creating the Lasso model

Lasso

1stFlrSF	0.111633
2ndFlrSF	0.104775
OverallQual	0.083460
BsmtFinSF1	0.048072

OverallCond 0.046879

After dropped the above Variables I got:

Lasso

BsmtFinSF 0.091798

TotRmsAbvGrd 0.088590

BsmtUnfSF 0.068099

GarageQual 0.047703

KitchenQual 0.046318

TotRmsAbvGrd0.088590

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?

ANS:Robust: If our model has very minute difference between Test and Train then our is called the Robust,If our model has low Bias and Low Variance then our model will be robust Example,Our Model Train Score is 90 and our Test Score is 86-89 then our model is called Robust

Generalization: If our model Adopt the New change in the model, then Our model is called Generalization Model, Example: If our model Learn something in new data and Perform well on Unseen Data then our model is called Generalization Model.