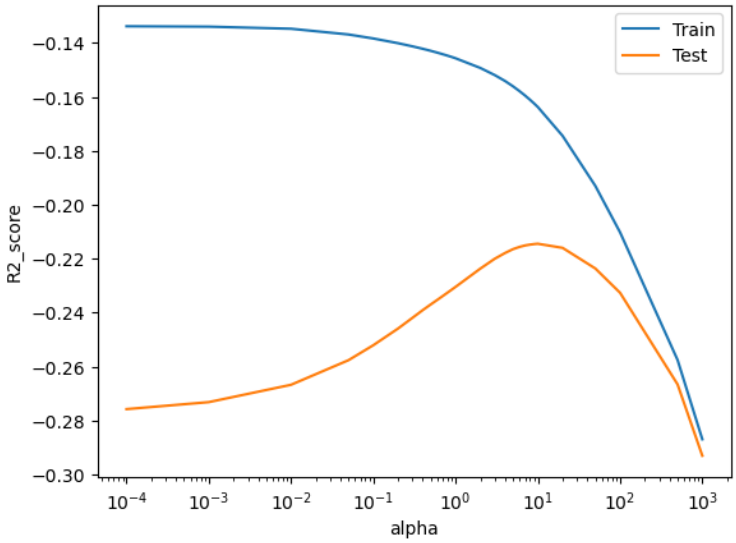
Question 1: What is the optimal value of alpha for ridge and lasso regression? What will be the changes in the model if you choose to double the value of alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?

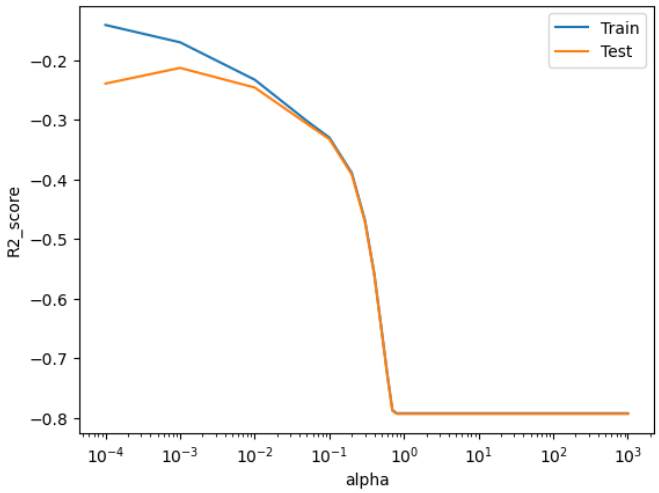
Ans. Optimal values of Alpha for:

a. Lasso Alpha: 0.01
b. Ridge Alpha: 10.0

Ridge graph:



Lasso graph:



There is no significant change in the values of r2 after doubling the alpha value for Lasso

Below is with Alpha = 0.001

- r2 train : 0.94
- r2 test : 0.92
- rss1 : 74.52
- rss2 : 24.85
- mse train : 0.06
- mse train : 0.09

Below is with Alpha = 0.002

- r2 train : 0.93
- r2 test : 0.92
- rss1 : 85.9
- rss2 : 24.19
- mse train : 0.07
- mse train : 0.08

There is no significant change in the values of r2 after doubling the alpha value for Ridge

Below is with Alpha = 10.0

- r2 train : 0.94
- r2 test : 0.92
- rss1 : 67.78
- rss2 : 25.19
- mse train : 0.06
- mse train : 0.09

Below is with Alpha = 20.0

- r2 train : 0.93
- r2 test : 0.92
- rss1 : 75.17
- rss2 : 24.6
- mse train : 0.06
- mse train : 0.08

Below are the important predictors for Ridge

Ridge_Coffe_Double_Alpha	Betas
OverallQual_9	0.2533
GrLivArea	0.2476
OverallQual_8	0.2288
Functional_Typ	0.1812
Exterior1st_BrkFace	0.1801
Neighborhood_Crawfor	0.1619
TotalBsmtSF	0.1472
OverallCond_9	0.1445
KitchenAbvGr_1	0.1244

Fireplaces_2	0.1242

Below are the important predictors for Lasso

Lasso_Coffe_Double_Alpha	Betas
GrLivArea	0.3596
OverallQual_8	0.1616
TotalBsmtSF	0.1331
YearBuilt	0.1233
BsmtQual_Ex	0.1118
YearRemodAdd	0.1118
BsmtFinSF1	0.1074
GarageArea	0.0947
GarageCars_3	0.0902
LotArea	0.0798

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:

1. Both the models have got decent R2 Squared value for both the train and test cases
2. The use of the model depends on the following:
    - If there are lots of variables and feature selection is required irrespective of the significance of each variables, then we can use lasso. Using Lasso, will consider only the features with very significant betas, but might lose some of the significant featues
    - If we want to consider all the features and we want our coefficients to be minimal then, we can chose ridge. The ridge will try to fit all the significant variables even if the beta is too low.

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:

Top 5 Significant variables which was in the initial Lasso model which won't be available in the incoming data would be :

1. **OverallQual_9 : 0.567886**
2. **OverallQual_8 : 0.409132**
3. **SaleCondition_Alloca : 0.340809**
4. **GrLivArea : 0.312970**
5. **Exterior1st_BrkFace : 0.299889**

Now, After that we dropped the above values and re-created the model. The new alpha for the lasso is till same that is 0.001. There's isn't significant difference in the r2 value.

After re-creating the new lasso model, following are the top 5 predictors :

1. **Exterior2nd_BrkFace 0.322839**
2. **2ndFlrSF 0.296624**
3. **KitchenAbvGr_1 0.291586**
4. **Neighborhood_StoneBr 0.263139**
5. **Neighborhood_Crawfor 0.262614**

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?

Answers :

1. A model is set to be robust when it performs well in both the training as well as test data and performance doesn't get affected on unseen data. Such models are generalised and can fit to unseen data without any issues with the performance.
2. The Robustness can also been seen from the fact that the model is not overfitting, that is it is not just learning the training data and performing great on training set but performing poorly on unseen data. As the complexity of the model increases the tendency of overfitting increases, thus the model becomes less generalized and will tend to have high variance.
3. Regularization helps in penalizing the model when it becomes too complex. This in turns prevents overfitting of the model.
4. Hyperparameter tuning helps in finding the optimal values of alpha, such that the model performs the best with significant betas.

Implication of Accuracy:

1. The accuracy of the model for the training data as well as the testing data without relying on too much pattern, should be significant.
2. A generalized model might have a lower accuracy due to the bias and variance trade off to find the optimal fitting params.
3. The regularization might penalize the model accuracy but it helps in identifying the optimal alpha values, and the model might perform well on any unseen training set
4. Overfitting will have a very higher accuracy for the training data while will perform poorly on the unseen data.