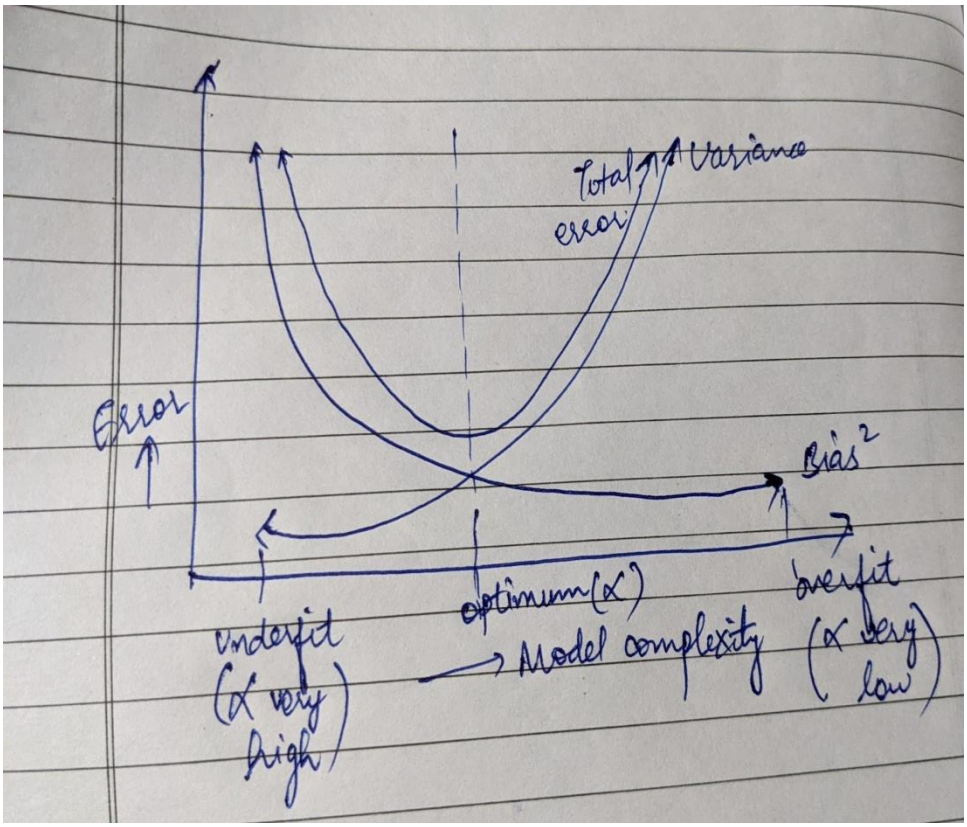
Question1:
What is the optimal value of alpha for ridge and lasso regression? What will be the changes in the model if you choose double the value of alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?

Answer:
The optimum value of alpha is the point where optimum value of bias is trade off for a significant reduction in variance of the model.

This value gives us the best model fit (not too complex as otherwise it will over fit and not too simple as otherwise it will under fit).

For this optimum value, Total Error would be least as described in below figure.



If the value of lambda is doubled, the model will tend to underfit as it would become simpler. Below are the most important Features for Ridge and Lasso when lambda is doubled:

Ridge:

1. GrLivArea
2. TotalBsmtSF
3. OverallQual_Very Excellent

4. OverallQual_Excellent
5. 2ndFlrSF

Lasso:

1. GrLivArea
2. OverallQual_Very Excellent
3. OverallQual_Excellent
4. TotalBsmtSF
5. OverallQual_Very Good

Question2: 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 of lambda for Ridge comes as 100 and for Lasso comes out as 1000.

We'll use Ridge as this is giving better results than Lasso in terms of R square value, RSS and MSS. Also, it is noticed that using Lasso Model, coefficients of nearly 239 features out of 284 turned out to be zero. This implies large value of lambda (as most beta coefficients are 0) and hence the Lasso model may be underfitting and very simple.

Hence, Ridge is better in this case.

Question3: After building the model, you realized 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:

The most important predictor variables dropped are:

	Linear	Ridge	Lasso
GrLivArea	2.757902e+04	1.649518e+04	31502.432848
OverallQual_Very Excellent	9.884924e+03	1.083716e+04	13883.237995
OverallQual_Excellent	9.779989e+03	9.094474e+03	12050.746298
OverallQual_Very Good	8.444018e+03	6.972468e+03	8948.955803
TotalBsmtSF	-1.857570e+04	1.128082e+04	8242.783688

Now, the important variables for Lasso are:

	Lasso
BsmtFinSF1	31800.823767
BsmtUnfSF	24413.062272
2ndFlrSF	23428.785081
BsmtFinSF2	8641.916722
LotArea	8126.472387

Question4: 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:
Model is robust and generalisable when small changes in predictor variables don't drastically reduce the performance of the model. If the performance of the model fluctuates widely with changes in the predictor variables, the model is not robust.

To make sure, the model is robust, we can:

1. Treat the outliers effectively.
2. The training dataset should not be biased, i.e. it should contain the wholesum data which represents the entire dataset.
3. Optimizing value of lambda in case of Lasso and Ridge, so that model is optimally complex. It's not underfitting, nor overfitting.
4. Not using the test data for training the model. Test Data should be unseen by the model and should be used only for making predictions.

   If the model is robust and generalisable, the difference between accuracy of test data and train data should be low. For making model general, train data accuracy can decrease slightly but that is trade off by increase in test data accuracy.