# Subjective Questions and Answers

## 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 1

The Optimal Values of Alpha obtained from Lasso and Ridge are given below:

- Optimal value of Alpha for Ridge 0.00079901
- Optimal value of Alpha for Lasso 0.00058701

If the value chosen for the alpha is doubled from the current one, the following changes will happen to both Lasso and Ridge given below:

- Common changes
  - Test Accuracy is reduced.
  - Test Error increases.
  - Bias is increased slightly.
  - Variance is reduced.
- Different Changes
  - Ridge
    - Coefficients value are brought closer to zero further.
  - Lasso
    - Coefficients of many columns becomes zero compared to before.

The Most Important Predictor after the change is implemented are given below as screenshot from notebook both positive and negative coefficients:

| | Columns | Coefficients |
|---|---|---|
| 73 | KitchenQual_Gd | -0.243036 |
| 74 | KitchenQual_TA | -0.219580 |
| 57 | BsmtQual_Gd | -0.201435 |
| 72 | KitchenQual_Fa | -0.184110 |
| 78 | GarageType_Basment | -0.167135 |

| | Columns | Coefficients |
|---|---|---|
| 86 | SaleType_New | 0.253796 |
| 77 | Functional_Typ | 0.258413 |
| 27 | Neighborhood_Crawfor | 0.331792 |
| 5 | GrLivArea | 0.405415 |
| 37 | Neighborhood_StoneBr | 0.451443 |

## 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 2

We selected **Lasso** as final model as it had higher accuracy, lower MSE, Lower RSS and implicit feature selection which automatically does the feature selection. It also has lower gap between the train and test accuracy. The feature selection performed allow creation of simpler model which is easier to deploy, understand and is more robust and generalized.

Comparison screenshot from notebook:

| | Ridge | Lasso |
|---|---|---|
| MSE Test | 0.090746 | 0.072316 |
| MSE Train | 0.064382 | 0.059215 |
| R2 Test | 0.867015 | 0.894023 |
| R2 Train | 0.918182 | 0.924749 |
| RSS Test | 34.574249 | 27.552460 |
| RSS Train | 57.107245 | 52.523712 |

## 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 3

The most important features given by Lasso model which have highest absolute value of coefficients which include both negative and positive coefficients are given below:

| | Columns | Coefficients |
|---|---|---|
| 112 | SaleType_New | 0.250234 |
| 100 | Functional_Typ | 0.333132 |
| 34 | Neighborhood_Crawfor | 0.336607 |
| 5 | GrLivArea | 0.438137 |
| 48 | Neighborhood_StoneBr | 0.505559 |

After removing these from dataset and rebuilding the model with new optimal alpha we get following columns as most important features:

| | Columns | Coefficients |
|---|---|---|
| 55 | HouseStyle_2.5Fin | 0.934974 |
| 122 | Functional_Sev | 0.646186 |
| 142 | SaleCondition_Alloca | 0.495348 |
| 18 | MSZoning_RH | 0.467169 |
| 131 | GarageQual_TA | 0.449777 |

## 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 4

The Model is robust and generalisable when it can handle changes in training data showing low variance and gives high accuracy on unseen data meaning it hasn't overfit.

The Model is said to be robust when it has low variance and low bias as per Occam's Razor Principle, which stats that model should be simple as it can understand the data and learn general patterns but shouldn't be too simple that it fails to learn the patterns of data.

The training accuracy is reduced as we make model simpler but test accuracy is dependent upon both variance and bias which is explained by both the Variance Bias Trade off.

In practice, however, we often cannot have a model with a low bias and a low variance. As the model complexity increases, the bias reduces, whereas the variance increases and, hence, the trade-off.