# Problem Statement - Part II

## 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 value of alpha for Ridge is 0.1 and for Lasso is 0.0001

When we double the value of alpha for Ridge, R square is 0.954782461208471 on train and 0.8831228509329014 on test

Similarly, on doubling alpha for lasso, R square is 0.9519235144389346 on train set and 0.881828103113388 On test set

No major changes in the predictors. The predictors for Ridge and lasso after change in alpha are shown below

**Ridge**

| | Feaure | Coef |
|---|---|---|
| 0 | MSSubClass | 11.291348 |
| 125 | RoofMatl_Membran | 0.317977 |
| 99 | Condition2_PosA | 0.301731 |
| 130 | RoofMatl_WdShngl | 0.292224 |
| 127 | RoofMatl_Roll | 0.272377 |
| 124 | RoofMatl_CompShg | 0.239699 |
| 103 | Condition2_RRNn | 0.226943 |
| 97 | Condition2_Feedr | 0.221041 |
| 126 | RoofMatl_Metal | 0.213785 |
| 122 | RoofStyle_Shed | 0.209401 |

**Lasso**

| | Feaure | Coef |
|---|---|---|
| 0 | MSSubClass | 11.924826 |
| 268 | SaleType_ConLD | 0.152409 |
| 68 | Neighborhood_Crawfor | 0.134708 |
| 84 | Neighborhood_StoneBr | 0.090545 |
| 238 | Functional_Typ | 0.082569 |
| 14 | LowQualFinSF | 0.078173 |
| 66 | Neighborhood_ClearCr | 0.075431 |
| 13 | 2ndFlrSF | 0.074151 |
| 134 | Exterior1st_BrkFace | 0.065264 |
| 83 | Neighborhood_Somerst | 0.064766 |

## 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?

- when lambda is 0 -r square is 1
- lambda is increased slightly 0.001, we see wigglyness is subdued to some extent and r2 values decreases by a certain amount so does the beta values
- lambda is 0.1, seems to be a good fit, and beta values are pushed down further therefore lambda of 0.1 gave us a good fit..
- for a high value of lambda= 1, model starts underfitting, r2 value would be really really low and also the beta coef pushed down further
- as the value of lambda increases, you will notice that beta coef pushed closer and closer to 0.. this is how regularization affects the model fit.
- Comparatively r2 score of ridge is slightly higher than lasso.

## Question 3

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?

- LotFrontage 0.006413
- OverallCond 0.033467
- YearRemodAdd 0.023564
- MasVnrArea 0.004776
- BsmtFinSF2 0.001995
- BsmtUnfSF 0.002140

## Question 4

How can you make sure that a model is robust and generalizable? What are the implications of the same for the accuracy of the model and why?

- The model should be generalized so that the accuracy on test is not lower than the training set. The model should be able to generalize well on unseen data. Overfitting is a phenomenon wherein a model becomes highly specific to the data on which it is trained and fails to generalise to other unseen data points in a larger domain. A model that has become highly specific to a training data set has 'learnt' not only the hidden patterns in the data but also

the noise and the inconsistencies in it. In a typical case of overfitting, a model performs quite well on the training data but fails miserably on the test data.

- Simpler models are more robust and this is a summary of Bias-Variance Tradeoff. If a model is complex, it will swing wildly with small changes in training data. It is always advisable to use simple yet robust model. This means the moment we change the input data, complex model will not perform well on unseen data unlike the simple model.

- Use a model that's resistant to outliers.

Thus, simple models are generalizable and robust