# Assignment: Advanced Regression

**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:**

We found out that optimised Alpha Value for both ridge and lasso is around 0.0001, since the alpha we choose is too small it is almost equivalent to just an Ordinary Least Squares Regression model. Our objective to maintain a good fit without overfitting or under fitting. Our model should be with low bias and low variance.

Choosing alpha to be 0.0001 had the following R-squared values

***Train Data:***

- Ridge R-Squared regression: 0.84
- Lasso R-Squared regression: 0.83

***Test Data:***

- Ridge R-Squared regression: 0.83
- Lasso R-Squared regression: 0.83

Since my alpha is too small doubling it didn't make a huge difference but on further increasing our R-square was dropping a bit.

So, the larger is the alpha, the higher is the smoothness constraint. So, the smaller the value of alpha, the higher would be the magnitude of the coefficients.

**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:**

Although we have an optimal value of lambda and ridge to be similar we don't see a much difference in their r-squared value but I will go with Lasso Regression.

Our objective is to reduce the number of features or independent variables for prediction which is additional benefit of the Lasso regression technique.

**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:**

Five most important predictor variables and with their coefficients are as follows

| Column | Coefficient |
|---:|---:|
| MSZoning_FV | 0.523643 |
| MSZoning_RL | 0.463164 |
| MSZoning_RH | 0.449951 |
| MSZoning_RM | 0.360489 |
| Exterior1st_BrkComm | -0.320382 |

Now, on rebuilding the model here are top important features

| Column | Coefficient |
|---:|---:|
| Neighborhood_ClearCr | 0.219974 |
| HouseStyle_2.5Fin | 0.204474 |
| CentralAir_Y | 0.198654 |
| Functional_Sev | -0.26531 |
| BsmtCond_NoBasement | -0.18656 |

- Neighbourhood : Clear Creek
- House Style : Two and one-half story: 2nd level finished
- Houses having Central Air
- Houses which are Severely Damaged negatively impact Sales Price
- Houses with No Basement negatively impact Sales Price

**Question 4**

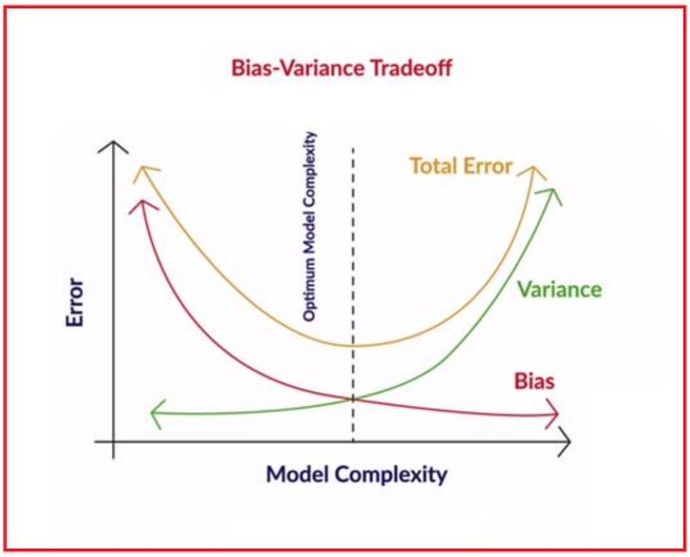
How can you make sure that a model is robust and generalisable? What are the implications of the same for the accuracy of the model and why?

**Answer:**

We can make a model robust and more generalizable by following Occam's razor rules, which are

- Making the model simple as necessary, but not too simpler.
- Out of two model simpler one needs to be chosen.
- Advantages of simplicity are generalisability, robustness, making few assumptions and less data required for learning.

Here is the model complexity trade-off.



**Bias-Variance Trade-off**

- Bias measures how accurately a model can describe the actual task at hand
- Variance measures how flexible the model is with respect to changes in the training data
- As complexity increases, bias reduces and variance increases, and we aim to find the optimal point where the total model error is the least

We should be choosing a model with less variance and less bias. More Bias leads to very simple model causing under fitting and more variance leads to overfitting the model. So we need keep a decent complexity to achieve a good fit.