**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 :

The value of alpha is the factor which decides upon the constraint for the regression polynomial .it represents the penalty terms of the regression . if alpha is zero then the polynomial is a simple linear regression . If alpha is doubled then the value cost function also increases and the coefficients decreases and extends towards zero .

**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 :

The lambda value for ridge and lasso are 26 and 20 respectively for the housing data set . After running the regression I see that the Ridge model produces better results with lambda = 26 and the R-squared value is optimum and it doesn’t signify overfitting or underfitting .

**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 :

The five most important predictor variables are :

|  |
| --- |
| * TotalBsmtSF |
| * BsmtFinSF2 |
| * BsmtUnfSF |
| * LotArea |
| * GarageArea |

**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 :

The model can be robust by the following :

* Reducing the error term
* No overfitting
* No underfitting