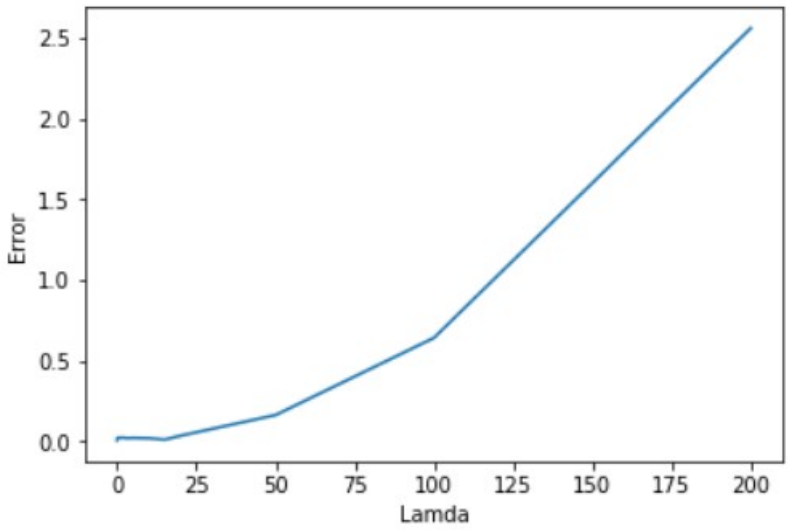
Regularization Techniques are used for:

- To reduce the computational complexity (scale down weights)
- Avoids Overfitting

1) Lasso Regression / L1 regularization:

Cost function : (Original Least Squares) + lamda*(sum of absolute value of coefficients)

Graph between lamda and error:

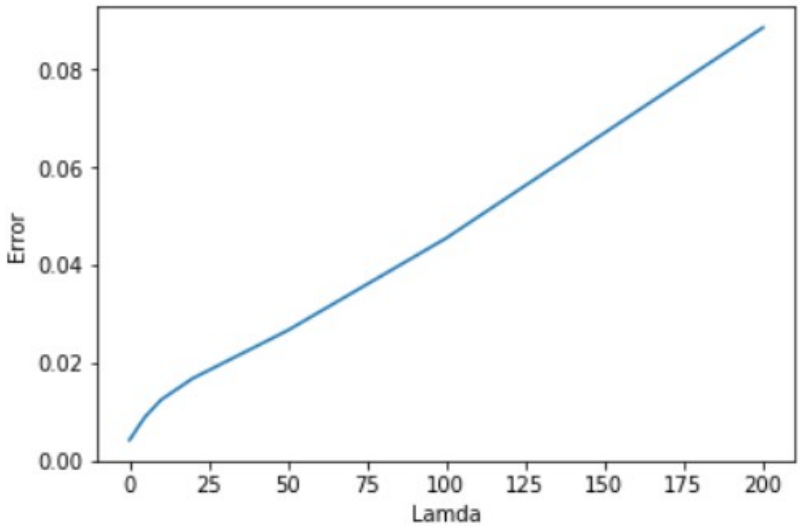


As lamda value increases,the parameters value decreases and the error increases

2)Ridge Regression / L2 Regularization:

Cost function : (Original Least Squares) + lamda*(sum of squares of coefficients)

Graph between lamda and error:



As lamda value increases,the parameters value decreases and the error increases

Question 3:

Original Least Squares measure has the property that it reduces bias.But It can have a huge variance.
Regularization helps in reducing variance but at the cost of introducing bias.

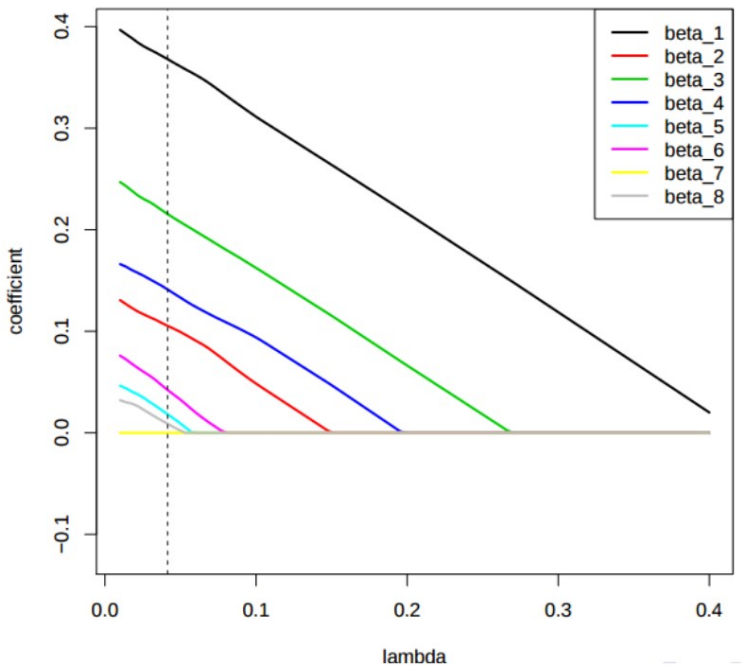
λ is the regularization parameter.
As λ becomes larger , both variance and bias increases.
So we need to choose optimal λ to get optimal Variance-Bias Tradeoff

Question 4:

Difference between Lasso and Ridge:

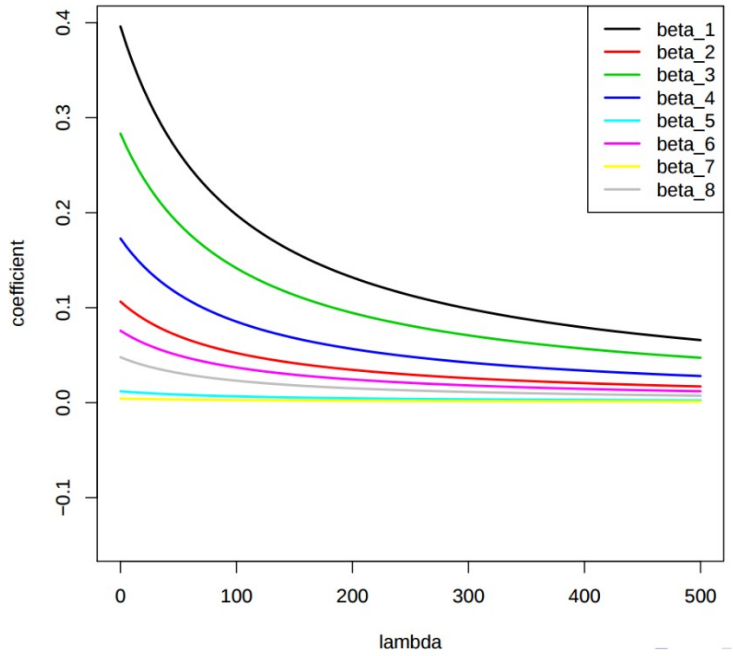
- Lasso does both parameter shrinkage and variable selection automatically because it zero out the co-efficients of collinear variables
- Ridge regression also does parameter shrinkage but can't zero coefficients
- Lasso tends to do well if there are a small number of significant parameters and the others are close to zero
- Ridge works well if there are many large parameters of about the same value

Effect of Lasso on Regression Weights



We can observe that coeffiecients are lowered as lamda increases.
The coefficients are scaled down and also set to zero if they are not relevant.

Effect of Ridge on Regression Weights

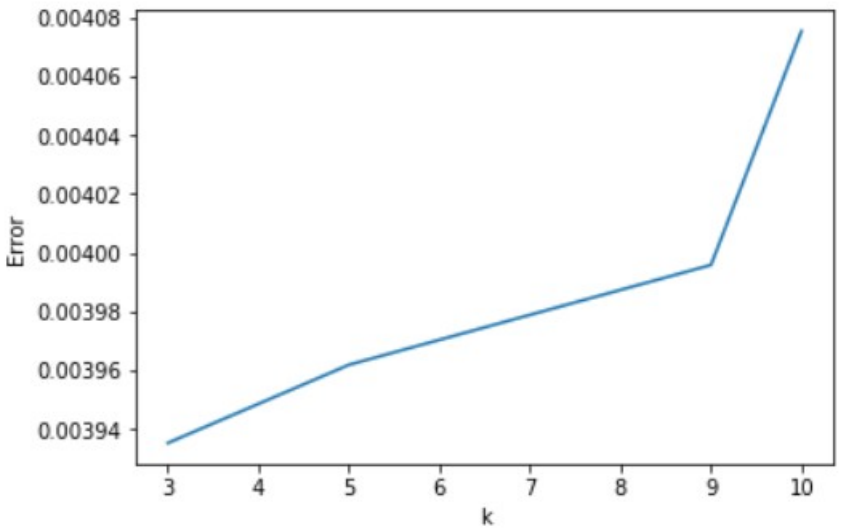


We can observe that the coefficients become lower as lamda increases but it does not enforce them to be zero .It will not get rid of irrelevant features but rather minimize their impact on the trained model.

Note:These graphs are taken from net for easy understanding of relation.

Question 5:

Graph with different values of k and error



Error increases with value of k.
By applying Cross validation with k=3,5 the error is less than normal L1 or L2 regularization.
Leave-One-Out Cross Validation takes huge amount of time to run.