**Problem 3: Experiment with Ridge Regression**

Issues associated with Linear Regression are as follows:

* Linear Regression is susceptible to outliers.
* Linear Regression is too simplistic- which leads to the problem of under fitting.

On the other hand, Ridge Regression or Tikhonov Regression helps in reducing the impact of correlated inputs.

For this problem, the weight vector (w) is learnt from the given dataset (training data matrix – x & corresponding labels – y) by implementing the *learnRidgeRegression* function.

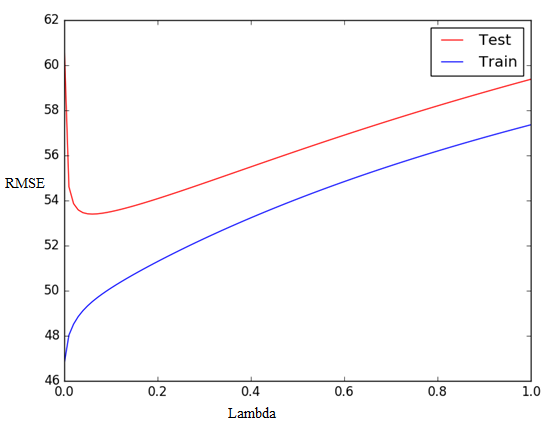
The impact of outliers on the weights is controlled by the regularization factor – λ. This ultimately helps in avoiding overfitting.

The regularized square loss function:



And the weight w is obtained by:



****

*Figure 3*

**Observations: -**

From the graph above, it’s clear that by varying the λ value, the RMSE value changes too. RMSE decreases steeply with increasing λ value. Initially, the λ value being zero, causes the second term containing λ become zero in the objective function above. But as the λ value increases from 0, the term plays a part in the minimization and emphasizes how complex the model can get. λ value controls the weight values leading to lower RMSE. But as λ value goes beyond 0.0599, the RMSE begins increasing which shows that the λ puts a tighter restriction on the weight vectors from growing leading to increased errors.

**Comparison of Relative Magnitudes of Weights:-**

Its observed that OLE tends to vary the weights more sharply than ridge regression. Relative to weights learnt using Ridge Regression, the weights learnt using OLE are higher.



Figure 4

Calculating the mean error from Linear Regression and Ridge Regression, we get the following table:

|  |  |  |
| --- | --- | --- |
| **Training/Testing Data** | **Mean RMSE from OLE** | **Mean RMSE from Ridge Regression** |
| Training Data | 46.76708559 | 56.2805719779 |
| Testing Data | 60.89203717 | 53.7038339189 |

As seen from the table, Linear Regression is seen to perform better in the learning phase. However, the mean error for the testing data. This is an example of the previously mentioned shortcomings of Linear Regression. Ridge Regression, on the other hand, performs better than Linear Regression with a mean error of 53.7038339189 on test data.

Finally, the Optimal value of λ calculated from this problem is as follows:-

**Optimum** λ value was observed at **0.0599** with the lowest **RMSE** error of **53.3979**.

As mentioned before, the first term in the equation for error is for error minimization and the second term controls the emphasize on weights. The reason why error is the lowest when λ is 0.0599 is because at that value both the terms in the equation for error are balanced.

