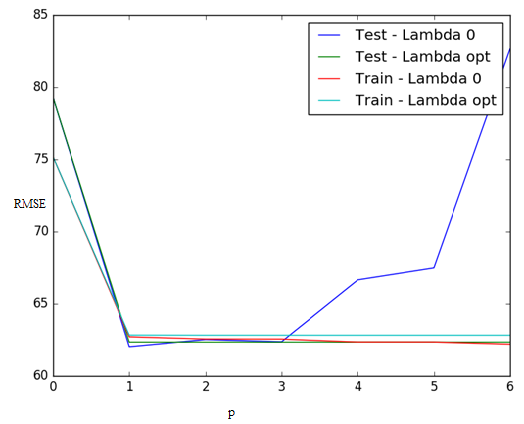
**Experiment 5: Non-linear Regression**

Non linear mapping, in this case polynomial expansion, is done to get non linear curves.

In this problem, using non linear mapping we have trained the weights for ridge regression in the *non-linear regression* function. We have used optimal λ value (0.0599) and λ as zero. The function was implemented so that we can convert a single attribute vector (N\*1) into a p attributed (N\*p+1) matrix.

The learning phase is similar to that used in problem 3, except that here we will use the (N\*p+1) matrix. Then the prediction for each data point is calculated with the weights obtained from learning function. The graph below shows the RMSE values for training and testing data with and without regularization.



*Figure 5*

**Observations: -**

When p=0, RMSE value is highest as no learning will be done on the training data. Further, the RMSE value begins decreasing after p=0, as is expected because of being trained by ridge regression.

When p=1 and λ as 0, RMSE increases steeply because there is no regularization factor, causing overfitting of data. This overfitting can be proven because, when p=1 and λ is optimal, RMSE value doesn’t increase. This is because regularization parameter (λ) takes care of the overfitting problem. RMSE for training data continues to be low, as the training was done on the data.

The RMSE values for both training and testing data with λ as zero and optimal value with different p values are tabulated in the table given below:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **p** | **Train when λ = 0** | **Test when λ = 0** | **p** | **Train when λ is opt** | **Test when λ is opt** |
| 0 | 75.17120818 | 79.28685132 | 0 | 75.1712145 | 79.28935782 |
| 1 | 62.69701275 | 62.00834404 | 1 | 62.81715398 | 62.32732645 |
| 2 | 62.54470138 | 62.5070244 | 2 | 62.80666908 | 62.32612357 |
| **3** | **62.53949684** | **62.35363292** | 3 | 62.80663223 | 62.32613129 |
| 4 | 62.33356293 | 66.658292 | **4** | **62.80663049** | **62.32613124** |
| 5 | 62.33303424 | 67.48948346 | 5 | 62.80663047 | 62.32613125 |
| 6 | 62.18427011 | 82.66473945 | 6 | 62.80663047 | 62.32613125 |

Optimum values of p for testing are observed as:

**p=3** for RMSE = 62.35363292 with no regularization, (i.e.) λ value zero

**p=4** for RMSE = 62.32613124 with regularization, (i.e.) λ is optimal value