CSE474/574 Introduction to Machine Learning
Programming Assignment 2
Classification and Regression

CSE 574: Assignment Group 33

Mayur Tale - 50169256

Vishwaksen Mane - 50168863

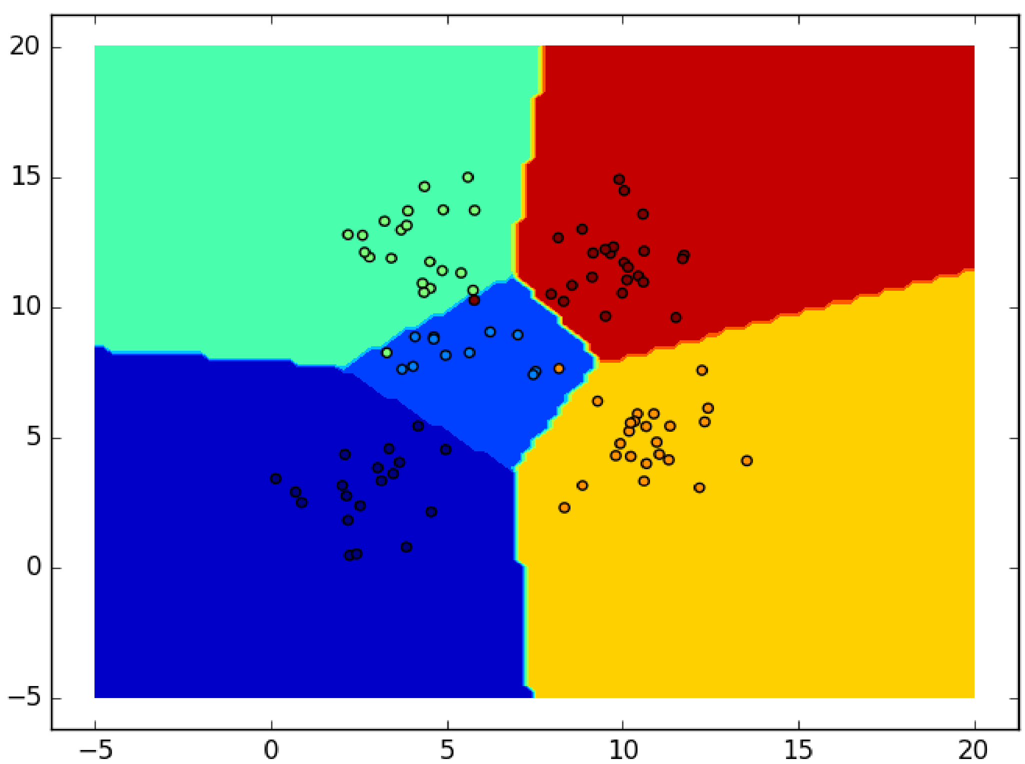
Abhijeet Deshpande - 50168465

Problem 1: Experiment with Gaussian Discriminators

```
[py27] C:\Users\Vishwaksen>python test_script.py
LDA Accuracy = 97.0
QDA Accuracy = 95.0
RMSE without intercept on testing data: 326.764994391
RMSE with intercept on testing data: 60.892037097
RMSE without intercept on training data: 138.20074835
RMSE with intercept on training data: 46.7670855937
Lambda: 0.0   Test Error: [ 60.8920371]   Train Error: [ 46.76708559]
Lambda: 0.00230769230769   Test Error: [ 56.70239382]   Train Error: [ 47.38333728]
Lambda: 0.00461538461538   Test Error: [ 55.70687022]   Train Error: [ 47.62757042]
Lambda: 0.00692307692308   Test Error: [ 55.11433679]   Train Error: [ 47.81941381]
Lambda: 0.00923076923077   Test Error: [ 54.7166248]   Train Error: [ 47.98075376]
Lambda: 0.0115384615385   Test Error: [ 54.4322698]   Train Error: [ 48.1207875]
Lambda: 0.0138461538462   Test Error: [ 54.22045209]   Train Error: [ 48.24467255]
```

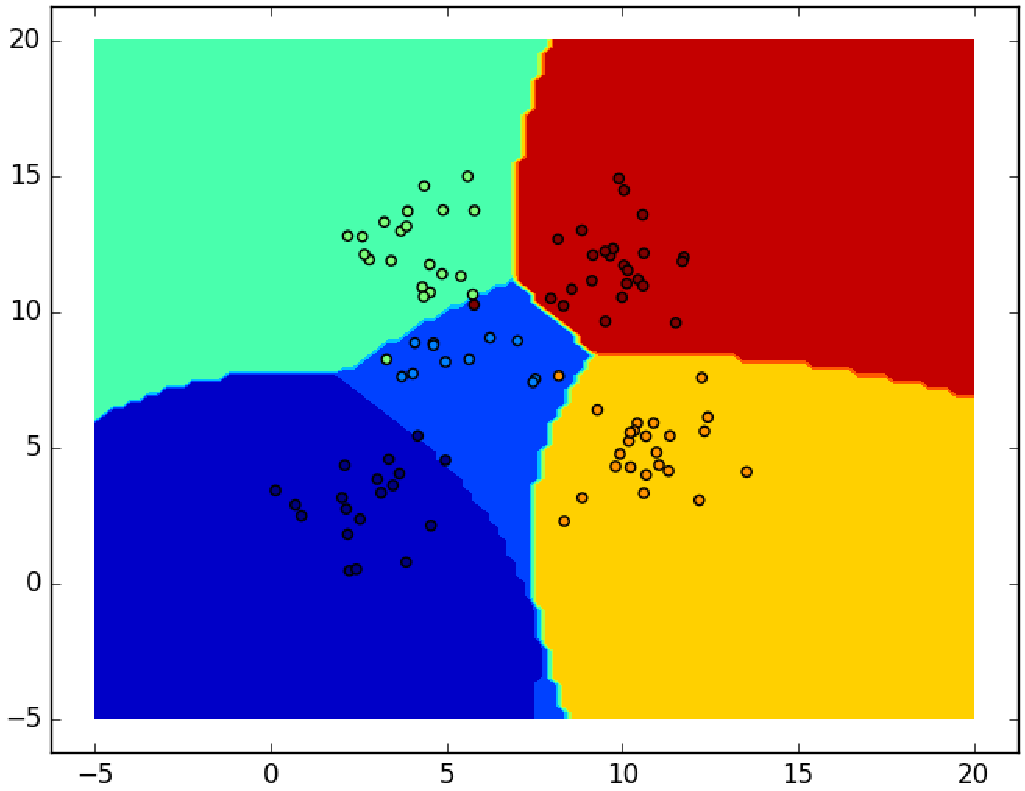
The above highlighted text shows LDA accuracy as 97% and QDA accuracy as 95%. The Discriminating Boundaries for LDA and QDA are as shown in the plots below. We observed that the difference in boundaries is due to the fact that QDA uses Covariance of each class whereas LDA uses the Covariance of whole data to generate the prediction.

LDA Plot:



This plot describes the boundaries generated by LDA

QDA Plot:



This plot describes the boundaries generated by QDA

Problem 2: Experiment with Linear Regression

The below snapshot describes the Root Mean Square Error for Testing and Training Data with and without Intercept obtained for Linear Regression.

```
[py27] C:\Users\Vishwaksen>python test_script.py
LDA Accuracy = 97.0
QDA Accuracy = 95.0
RMSE without intercept on testing data: 326.764994391
RMSE with intercept on testing data: 60.892037097
RMSE without intercept on training data: 138.20074835
RMSE with intercept on training data: 46.7670855937
Lambda: 0.0   Test Error: [ 60.8920371]   Train Error: [ 46.76708559]
Lambda: 0.00230769230769   Test Error: [ 56.70239382]   Train Error: [ 47.38333728]
Lambda: 0.00461538461538   Test Error: [ 55.70687022]   Train Error: [ 47.62757042]
Lambda: 0.00692307692308   Test Error: [ 55.11433679]   Train Error: [ 47.81941381]
Lambda: 0.00923076923077   Test Error: [ 54.7166248]   Train Error: [ 47.98075376]
Lambda: 0.0115384615385   Test Error: [ 54.4322698]   Train Error: [ 48.1207875]
Lambda: 0.0138461538462   Test Error: [ 54.22045209]   Train Error: [ 48.24467255]
Lambda: 0.0161538461538   Test Error: [ 54.05801403]   Train Error: [ 48.35577845]
Lambda: 0.0184615384615   Test Error: [ 53.93072625]   Train Error: [ 48.45649757]
Lambda: 0.0207692307692   Test Error: [ 53.82933013]   Train Error: [ 48.54861502]
Lambda: 0.0230769230769   Test Error: [ 53.74753245]   Train Error: [ 48.63350719]
Lambda: 0.0253846153846   Test Error: [ 53.68090489]   Train Error: [ 48.71226068]
Lambda: 0.0276923076923   Test Error: [ 53.62624203]   Train Error: [ 48.78574889]
Lambda: 0.03   Test Error: [ 53.58116823]   Train Error: [ 48.85468415]
```

Based on the above results, we observed that Linear Regression performs better with Intercept for both Training and Testing Data. The RMSE without intercept is observed to be much higher as compared to RMSE with intercept on training and testing data.

We also observed that the lower value of RMSE with intercept for training data shows better accuracy possibly due to smaller data.

Problem 3: Experiment with Ridge Regression

RMSE with Intercept for OLE Regression

Training Data: 46.7670855937

Testing Data: 60.892037097

RMSE with Intercept for Ridge Regression

Training Data: 49.5929123628

Testing Data: 53.3978483971

We observed that Ridge Regression is better than Linear Regression with intercept since we observed that the RMSE value of testing data obtained using Ridge Regression is **smaller** than the RMSE value of testing data obtained using Linear Regression.

The RMSE values with/without intercept on training/testing data are highlighted in below snapshot.

```
Lambda: 0.06Test Error: [ 53.3978484]Train Error: [ 49.51291236]
Lambda: 0.0623076923077Test Error: [ 53.39846238]Train Error: [ 49.55283755]
Lambda: 0.0646153846154Test Error: [ 53.40010546]Train Error: [ 49.5918539]
Lambda: 0.0669230769231Test Error: [ 53.40267628]Train Error: [ 49.6300337]
Lambda: 0.0692307692308Test Error: [ 53.40608667]Train Error: [ 49.66744179]
Lambda: 0.0715384615385Test Error: [ 53.41025959]Train Error: [ 49.70413654]
Lambda: 0.0738461538462Test Error: [ 53.41512745]Train Error: [ 49.74017056]
Lambda: 0.0761538461538Test Error: [ 53.42063071]Train Error: [ 49.77559143]
Lambda: 0.0784615384615Test Error: [ 53.42671677]Train Error: [ 49.81044222]
Lambda: 0.0807692307692Test Error: [ 53.43333897]Train Error: [ 49.84476205]
Lambda: 0.0830769230769Test Error: [ 53.44045577]Train Error: [ 49.87858647]
Lambda: 0.0853846153846Test Error: [ 53.44803011]Train Error: [ 49.91194789]
Lambda: 0.0876923076923Test Error: [ 53.45602879]Train Error: [ 49.94487588]
Lambda: 0.09Test Error: [ 53.46442201]Train Error: [ 49.97739748]
Ridge Regression RMSE with intercept on testing data: 53.3978483971
Ridge Regression RMSE with intercept on training data: 49.5129123628
Ridge Regression RMSE without intercept on testing data: 197.289999356
Ridge Regression RMSE without intercept on training data: 147.325318113
```

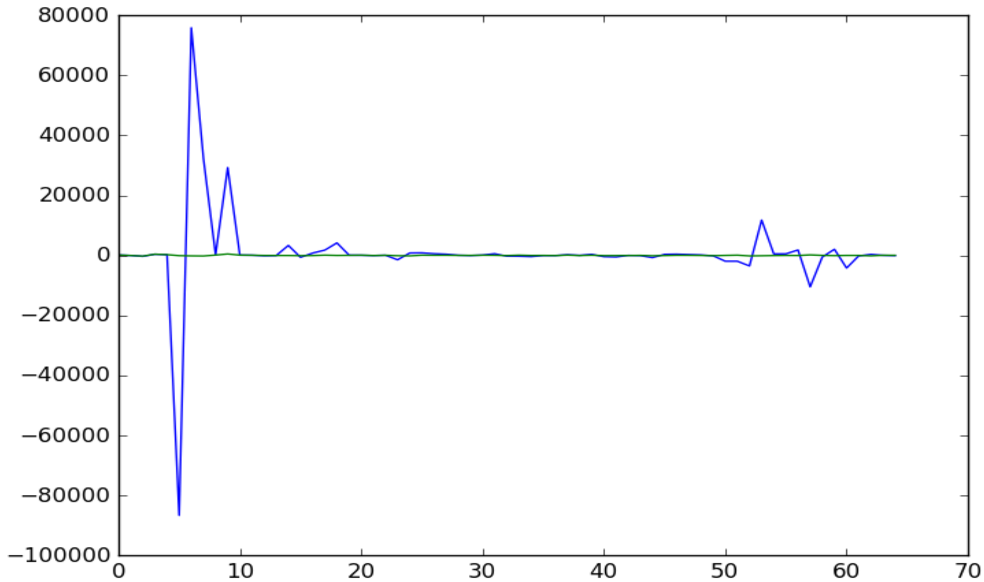
Optimal Lambda → 0.06

```
Lambda: 0.0   Test Error: [ 58.13340827]   Train Error: [ 47.36817616]
Lambda: 0.00230769230769   Test Error: [ 56.78086862]   Train Error: [ 47.48483205]
Lambda: 0.00461538461538   Test Error: [ 55.38935668]   Train Error: [ 47.7699075]
Lambda: 0.00692307692308   Test Error: [ 55.09168751]   Train Error: [ 47.85135038]
Lambda: 0.00923076923077   Test Error: [ 54.72168571]   Train Error: [ 48.00434072]
Lambda: 0.0115384615385   Test Error: [ 54.47548752]   Train Error: [ 48.11419505]
Lambda: 0.0138461538462   Test Error: [ 54.20903558]   Train Error: [ 48.2657637]
Lambda: 0.0161538461538   Test Error: [ 54.05623643]   Train Error: [ 48.360502]
Lambda: 0.0184615384615   Test Error: [ 53.91110997]   Train Error: [ 48.46603388]
Lambda: 0.0207692307692   Test Error: [ 53.8642351]   Train Error: [ 48.52221508]
Lambda: 0.0230769230769   Test Error: [ 53.72114126]   Train Error: [ 48.671324]
Lambda: 0.0253846153846   Test Error: [ 53.68307389]   Train Error: [ 48.71283674]
Lambda: 0.0276923076923   Test Error: [ 53.62789845]   Train Error: [ 48.78638908]
Lambda: 0.03   Test Error: [ 53.58016979]   Train Error: [ 48.85631004]
Lambda: 0.0323076923077   Test Error: [ 53.54386089]   Train Error: [ 48.91997407]
Lambda: 0.0346153846154   Test Error: [ 53.51299546]   Train Error: [ 48.98104515]
Lambda: 0.0369230769231   Test Error: [ 53.49007284]   Train Error: [ 49.03902853]
Lambda: 0.0392307692308   Test Error: [ 53.4661761]   Train Error: [ 49.0965874]
Lambda: 0.0415384615385   Test Error: [ 53.44848437]   Train Error: [ 49.1490852]
Lambda: 0.0438461538462   Test Error: [ 53.43499206]   Train Error: [ 49.19983031]
Lambda: 0.0461538461538   Test Error: [ 53.42375395]   Train Error: [ 49.24901124]
Lambda: 0.0484615384615   Test Error: [ 53.412062]   Train Error: [ 49.30191906]
Lambda: 0.0507692307692   Test Error: [ 53.40834619]   Train Error: [ 49.34235593]
Lambda: 0.0530769230769   Test Error: [ 53.40359543]   Train Error: [ 49.38676741]
Lambda: 0.0553846153846   Test Error: [ 53.40016568]   Train Error: [ 49.43001301]
Lambda: 0.0576923076923   Test Error: [ 53.39847537]   Train Error: [ 49.47193522]
Lambda: 0.06   Test Error: [ 53.3978335]   Train Error: [ 49.5129164]
Lambda: 0.0623076923077   Test Error: [ 53.39850344]   Train Error: [ 49.55277729]
Lambda: 0.0646153846154   Test Error: [ 53.40011764]   Train Error: [ 49.59186034]
Lambda: 0.0669230769231   Test Error: [ 53.40270081]   Train Error: [ 49.63004307]
Lambda: 0.0692307692308   Test Error: [ 53.40606653]   Train Error: [ 49.66744903]
Lambda: 0.0715384615385   Test Error: [ 53.41038235]   Train Error: [ 49.70435078]
Lambda: 0.0738461538462   Test Error: [ 53.41513707]   Train Error: [ 49.740171]
Lambda: 0.0761538461538   Test Error: [ 53.42063542]   Train Error: [ 49.77560376]
Lambda: 0.0784615384615   Test Error: [ 53.4267154]   Train Error: [ 49.81044044]
Lambda: 0.0807692307692   Test Error: [ 53.4321852]   Train Error: [ 49.84553787]
Lambda: 0.0830769230769   Test Error: [ 53.44045347]   Train Error: [ 49.87859562]
Lambda: 0.0853846153846   Test Error: [ 53.44809295]   Train Error: [ 49.91196564]
Lambda: 0.0876923076923   Test Error: [ 53.45601802]   Train Error: [ 49.94489957]
Lambda: 0.09   Test Error: [ 53.46441977]   Train Error: [ 49.97739692]
```

Inference:

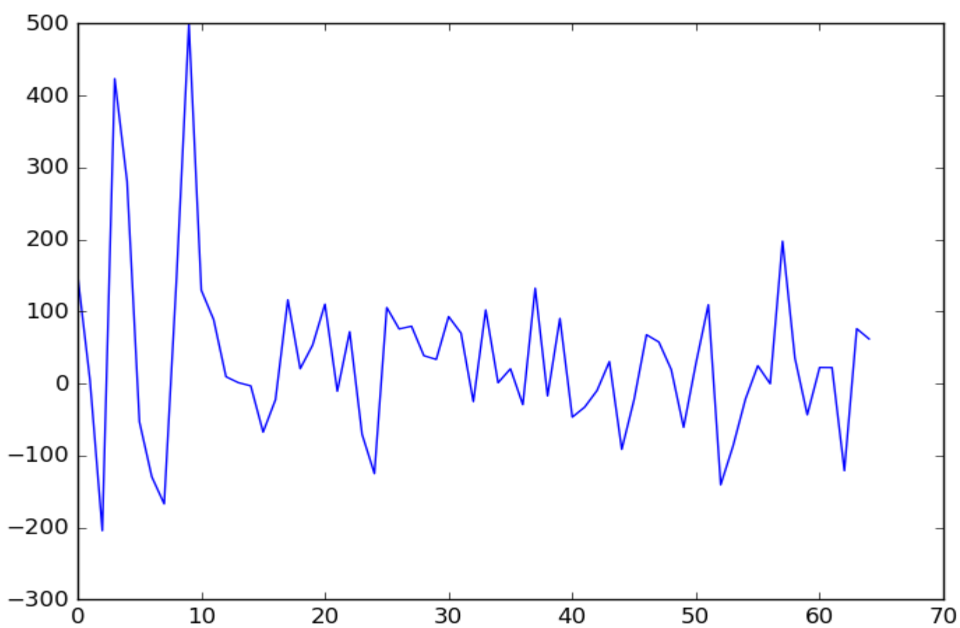
As seen in the above screenshot, the RMSE for testing and training data tends to decrease as the Lambda value varies from 0 to 0.06. Moreover, we observed that as the Lambda value varies from 0.06 to 0.09, the RMSE for testing and training data tends to increase considerably. Thus we observed that the RMSE tends to converge for **Lambda = 0.06**. The above highlighted text indicates an **<u>Optimal Value of Lambda = 0.06</u>** which leads to minimum RMSE.

Magnitude of weights learnt using OLE:



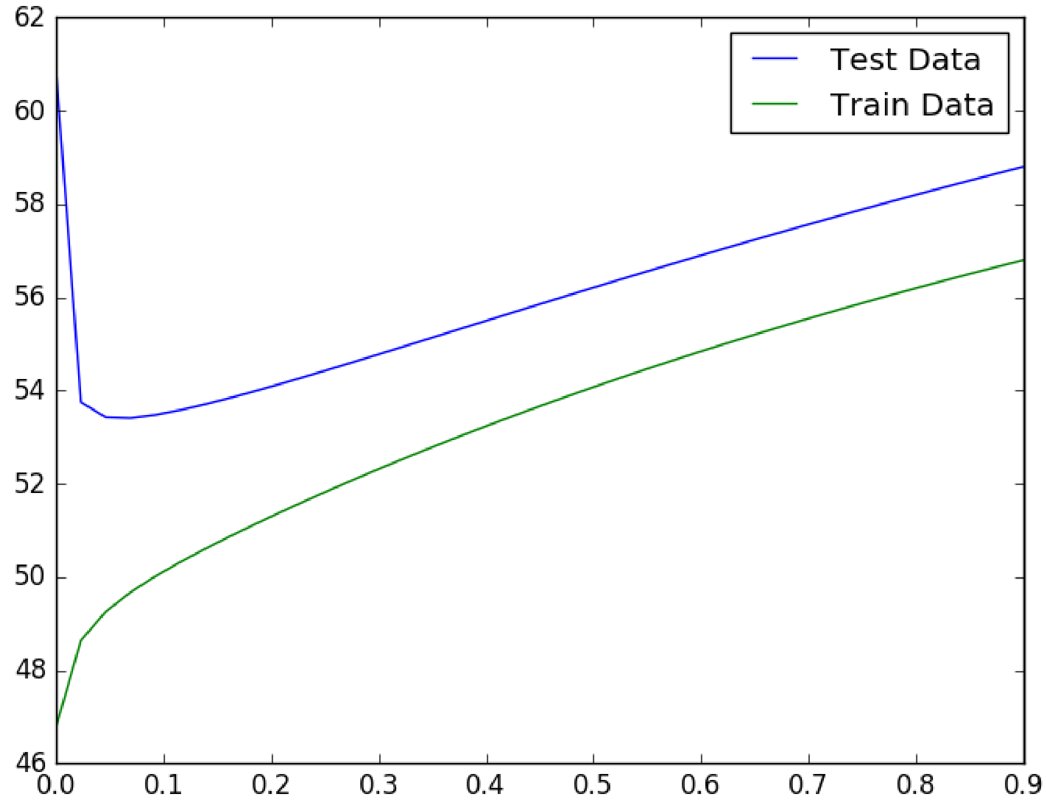
The above plot depicts the magnitude of weights learnt using Linear Regression with Intercept.

Magnitude of weights learnt using Ridge Regression:



The above plot depicts the magnitude of weights learnt using Ridge Regression with Intercept.

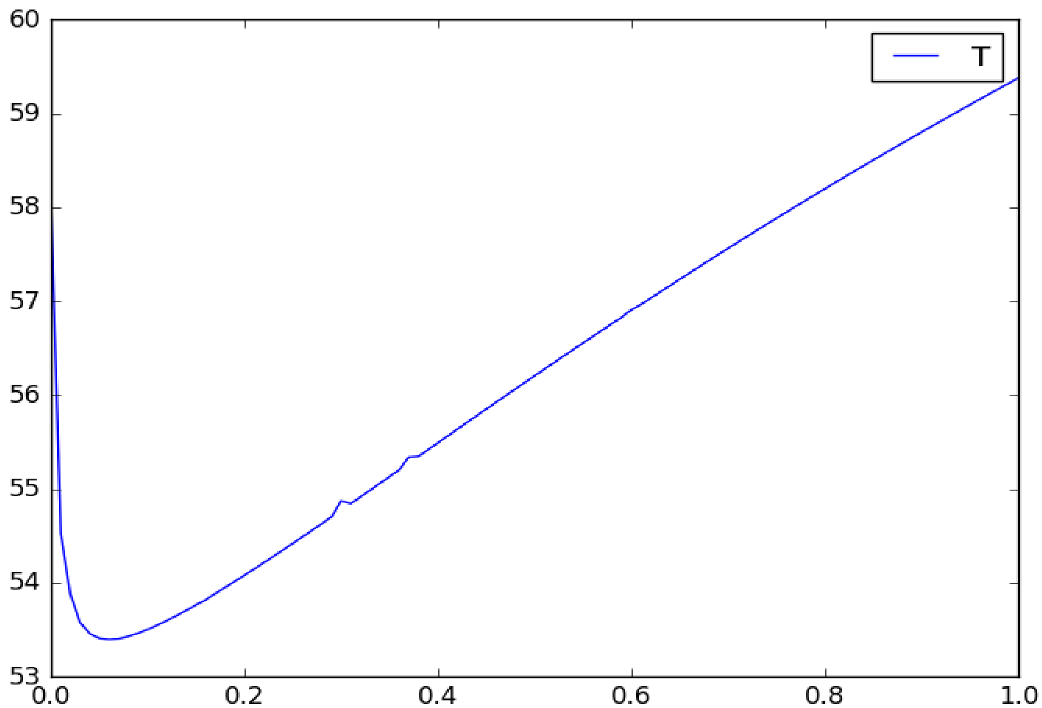
Comparison of RMSE for Training and Testing Data plotted against values of Lambda:



The above plot shows the variation of RMSE for training and testing data plotted against different values of Lambda.

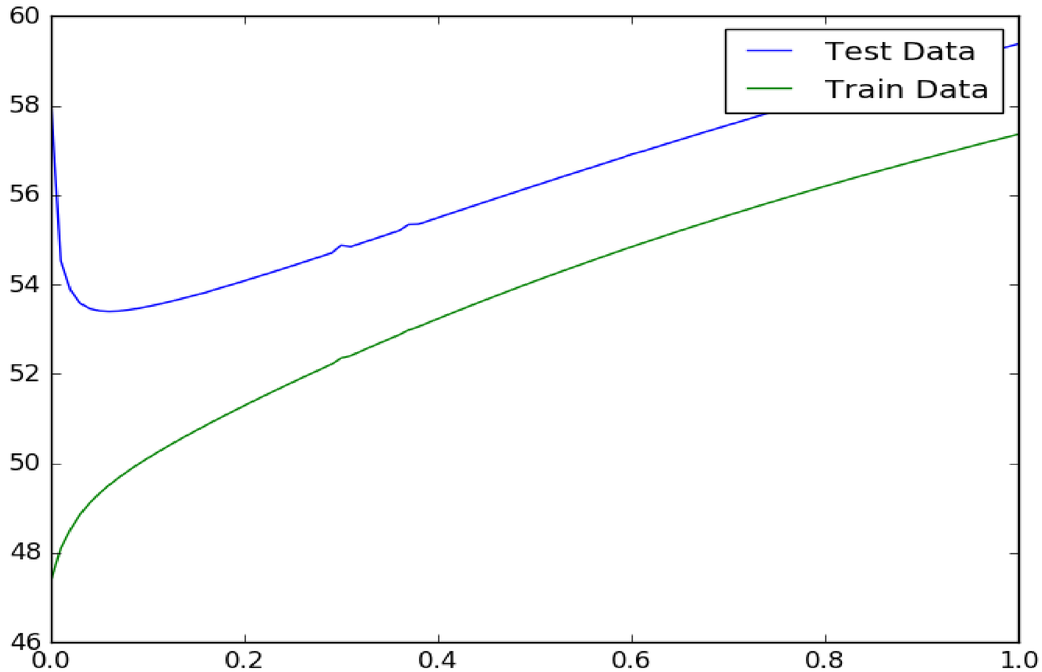
Problem 4: Gradient Descent for Ridge Regression Learning

Gradient Descent based RMSE Plot against Testing Data with Intercept:



The above plot describes the RMSE obtained using gradient descent on testing data with intercept.

Gradient Descent based RMSE Plot against Training and Testing Data with Intercept:



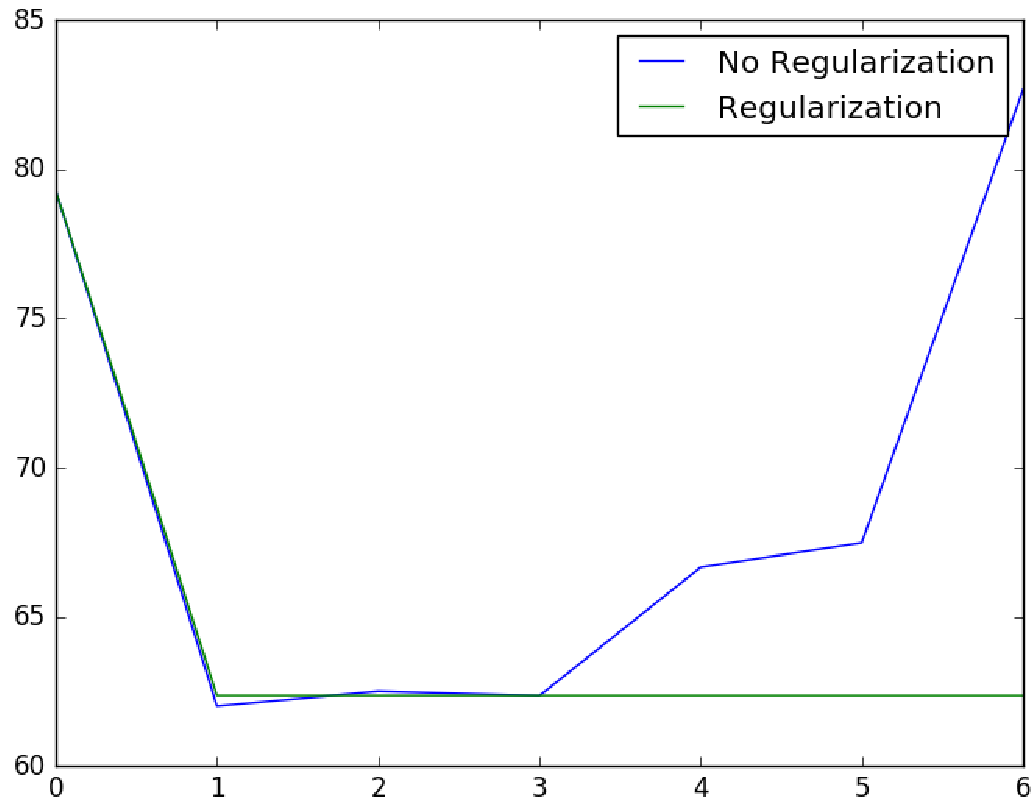
The above plot describes the variation of Gradient Descent based RMSE on training and testing data with intercept.

Inference:

We observed that the above plots for Gradient Descent based Ridge Regression can be compared with the Ridge Regression plots to infer that there is very small change in the values for RMSE showing minor peaks in the prior one. Also, we may infer that the Optimal Value of Lambda = 0.06 for both the Regressions with and without Gradient Descent remains same.

Problem 5: Non Linear Regression

RMSE Regularization Plot against Different Values of P:



Observations:

The above plot shows RMSE values against Value of P with and without Regularization. Thus based on the above plot, we observed that Non-Linear Regression with Regularization performs better than without Regularization.

As seen above, we observed that the RMSE without Regularization decreases as value of P varies from 0 to 1 and increases rapidly and unevenly as value of P varies from 3 to 6. and is optimal for P=1.

Moreover, we observed that RMSE with Regularization decreases as value of P changes from 0 to 1. However, we observed that RMSE value almost tends to decrease slightly or remain constant as value of P fluctuates between 1 to 6.

Based on these facts, we observed that optimal value for P is 1 for Lambda=0 and optimal P is 1 for Optimal Lambda i.e. 0.06

P Value	Testing Error with Lambda = 0	Testing Error with Lambda = Optimal
0	79.2868513165	79.2898604296
1	62.0083440367	62.416796333
2	62.5070243981	62.4146141215
3	62.3536329193	62.4146033867
4	66.6582919959	62.4146030051
5	67.4894834581	62.4146030085
6	82.6647394523	62.4146030086

As seen above, we executed our code so as to display and record the variations in testing error with Lambda as Zero and Lambda as Optimal (i.e. 0.06) against different values of P.

The highlighted text shown above depicts that we obtained minimum error values for testing error with Lambda=0 at P=1 and Lambda=Optimal (i.e. 0.06) at P=4.

Inference:

Optimal P for Lambda = 0	**1**
Optimal P for Lambda = Optimal (i.e. 0.06	**4**

Problem 6: Interpreting Results

OLE Regression:

RMSE with Intercept on Testing Data: 60.892037091
RMSE without Intercept on Testing Data: 326.764994391
RMSE with Intercept on Training Data: 46.7670855937
RMSE without Intercept on Training Data: 138.20074835

Ridge Regression:

RMSE with Intercept on Testing Data: 53.3976483971
RMSE without Intercept on Testing Data: 197.289999356
RMSE with Intercept on Training Data: 49.5929123628
RMSE without Intercept on Training Data: 147.325318113

Optimal Lambda for Ridge Regression: 0.06

Non-Linear Regression:

	Optimal Testing Error	**Optimal P Value**
Lambda = 0	62.0083440367	P=1
Lambda = 0.06	62.4146030051	P=4

Inference:

Based on the observations presented above, it can be inferred that Ridge Regression performs better for Testing data whereas OLE regression performs better for Training data giving minimum RMSE.

The Optimal value for Lambda is 0.06 for Ridge Regression. So we can conclude that, Ridge Regression is the best setting compared to others by taking the average RMSE values for Training and Testing data. Thus Gradient Descent based Ridge Regression approach performs better on training and testing data.