Machine Learning Project Report

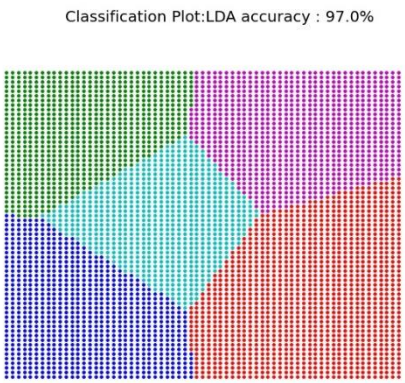
Assignment 2

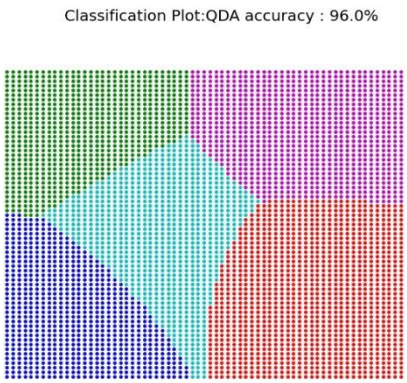
Ayush Jitendra Sanghvi (50132376)

Priyankar Nandi (50134010)

Avani Chandrashekhar Atre (50134720)

Problem 1





LDA accuracy -97.0%
QDA accuracy-96.0%

QDA and LDA are same except the fact that the covariance matrix in LDA is common to classes. The main difference between QDA and LDA classification is the bias variance trade off. QDA has more flexible decision boundaries compared to LDA as it is quadratic compared to linear. If there is a significant difference in the covariance's of the classes then using QDA for classification becomes an obvious choice. The difference in boundaries of QDA and LDA are mainly because of the above reasons.

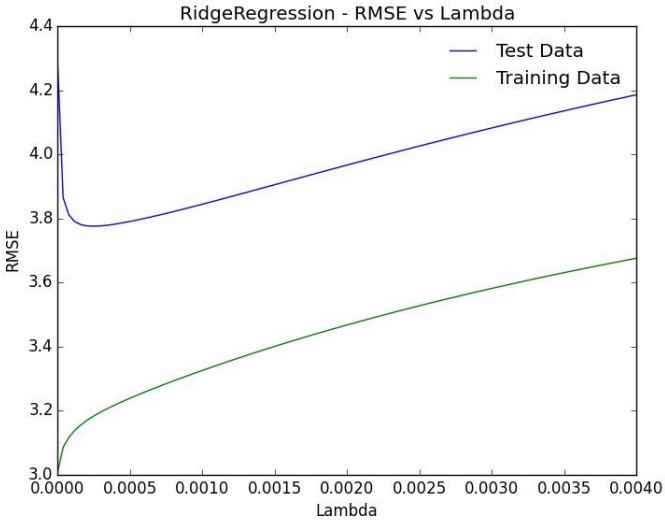
Problem 2

RMSE without intercept train:	8.88388057487
RMSE with intercept train:	3.0063021236
RMSE without intercept test:	23.1057743384
RMSE with intercept test:	4.3057172351

RMSE in data with intercept is lesser that the RMSE obtained in data without intercept. Hence the RMSE value with intercept is better compared to RMSE without intercept. This behavior is because with intercept we get a better fit for linear regression. Hence error is less for data with intercept.

Problem 3

Plot for Ridge Regression



Weights comparison for Ridge Regression and OLE Regression

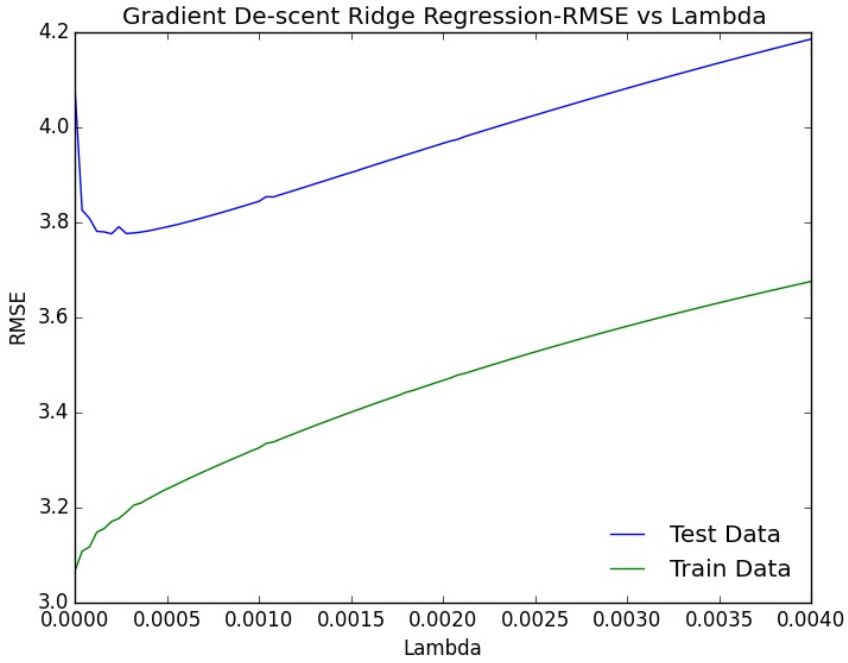
Observations:

- In rigde regression the weights in the higher range of suffer and hence the fall in same range. This is not the case with OLE regression and we can observe that the weights in the higher range do not suffer. Hence we observe such high deviation.
- The mean weight for OLE Regression is 882.81 which is greater than that of Ridge Regression which is 32.19 for corresponding optimal value of lambda.
- RMSE observed for test data in OLE Regression is: 4.30
  RMSE observed for training data in OLE Regression is: 3.00
  RMSE observed for test data in Ridge Regression is for optimal lambda: 3.77
  RMSE observed for training data in Ridge Regression is for optimal lambda: 3.18

- The root mean square error of training dataset in more in Rigde Regression Compared to that of OLE Regression with intercept but the root mean square error of test dataset is less that Ridge Regression Compared to OLE Regression.
- We found optimal value of lambda to be equal to 0.00024. The optimal value of lambda corresponds to minimum value of root mean square error. This result is mainly because of bias and variance trade off.
- For lambda less than 0.00024 the error observed is high because of under fitting problem. On the other hand where lambda is greater than 0.00024 because of rote learning the error increases because over fitting problem.
- Hence Ridge regression proves to be a better choice over OLE regression as the RMSES are lower.

Problem 4

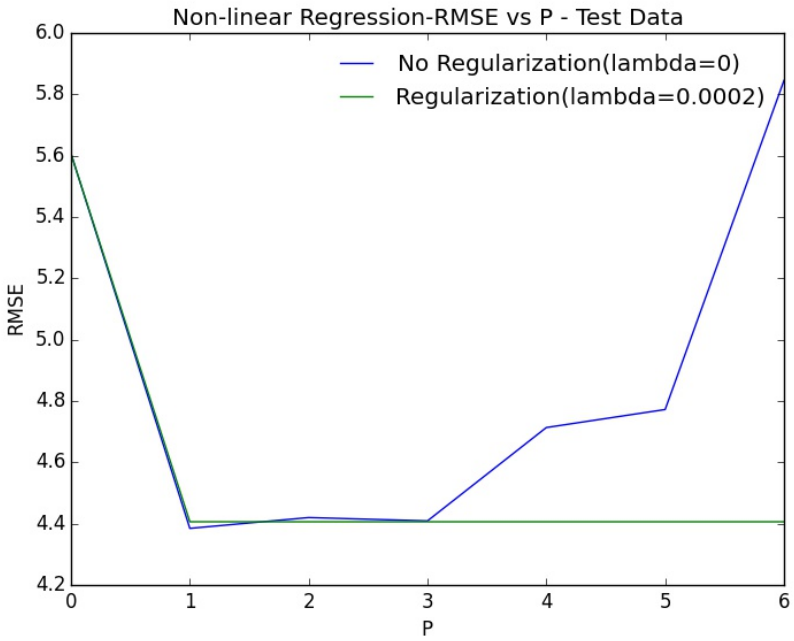
Plot for Ridge Regression with Gradient Descent



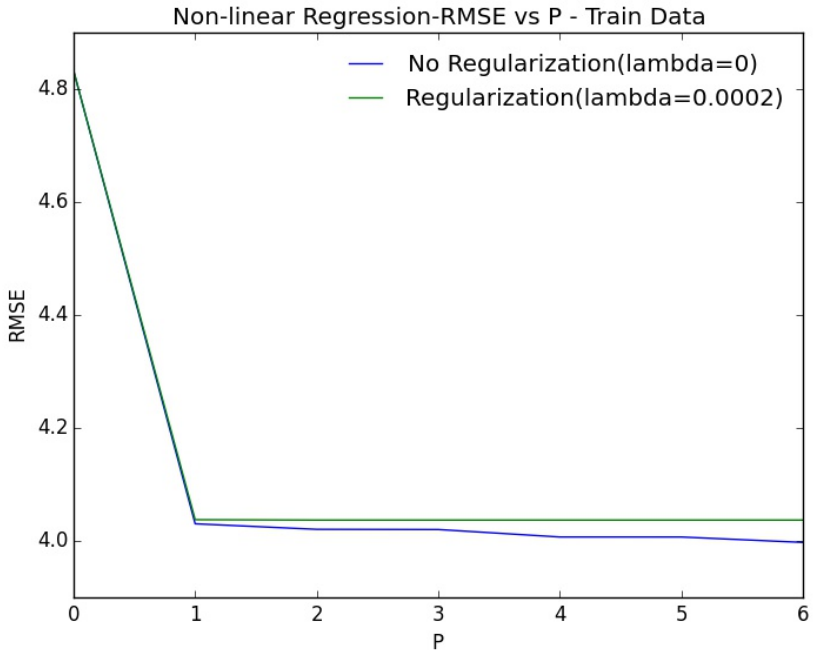
- When we observe graphs for RMSE vs Lambda are very similar for Ridge Regression and Ridge Regression with gradient descent.
- Hence for the given data set the gradient descent gives the same result as that of ridge regression.

Problem 5

Plot for Nonlinear Regression for Test Data



Plot for Nonlinear Regression for Train Data



lambda=0	No Regularization	
p	Test Data	Train Data
0	5.60642702	4.83218828
1	4.38465206	4.03031662
2	4.41991408	4.02052568
3	4.40906767	4.02019112
4	4.71345303	4.00695319
5	4.77222714	4.00691920
6	5.84527978	3.99735628

lambda=0.0002	Regularization	
p	Test Data	Train Data
0	4.83218883	5.60663297
1	4.04042614	4.41228423
2	4.03982307	4.41214157
3	4.03982064	4.41214105
4	4.03982053	4.41214103
5	4.03982053	4.41214103
6	4.03982053	4.41214103

Lambda = 0

The most optimal results for Training Data is 3.99 obtained for p=6.
The most optimal results for the Test Data is 4.38 obtained for p=1.

Lambda = 0.0002

The most optimal results for Training Data is 4.41 obtained for p=6.
The most optimal results for the Test Data is 4.03 obtained for p=4.

Problem 6

Comparisons:

RMSE can be used as a metric to compare the different approaches.

OLE regression:

RMSE without intercept train: 8.88388057487
RMSE with intercept train: 3.0063021236
RMSE without intercept test: 23.1057743384
RMSE with intercept test: 4.3057172351

Ridge regression:

RMSE observed for test data for optimal lambda: 3.77
RMSE observed for training for optimal lambda: 3.18

Gradient Descent for Ridge Regression Learning:

The RMSE values are almost similar to that of Ridge Regression

Non-linear Regression:

Lambda = 0
The most optimal results for Training Data is 3.99 obtained for p=6.
The most optimal results for the Test Data is 4.38 obtained for p=1.
Lambda = 0.0002
The most optimal results for Training Data is 4.41 obtained for p=6.
The most optimal results for the Test Data is 4.03 obtained for p=4.

Conclusion:

Ridge regression and Ridge regression with gradient descent have low RMSE. Hence, for producing high accuracy results ridge regression and ridge regression with gradient descent is preferred.