CSE474/574

Programming Assignment 2

Regression and Classification

Compiled By:

Team 51

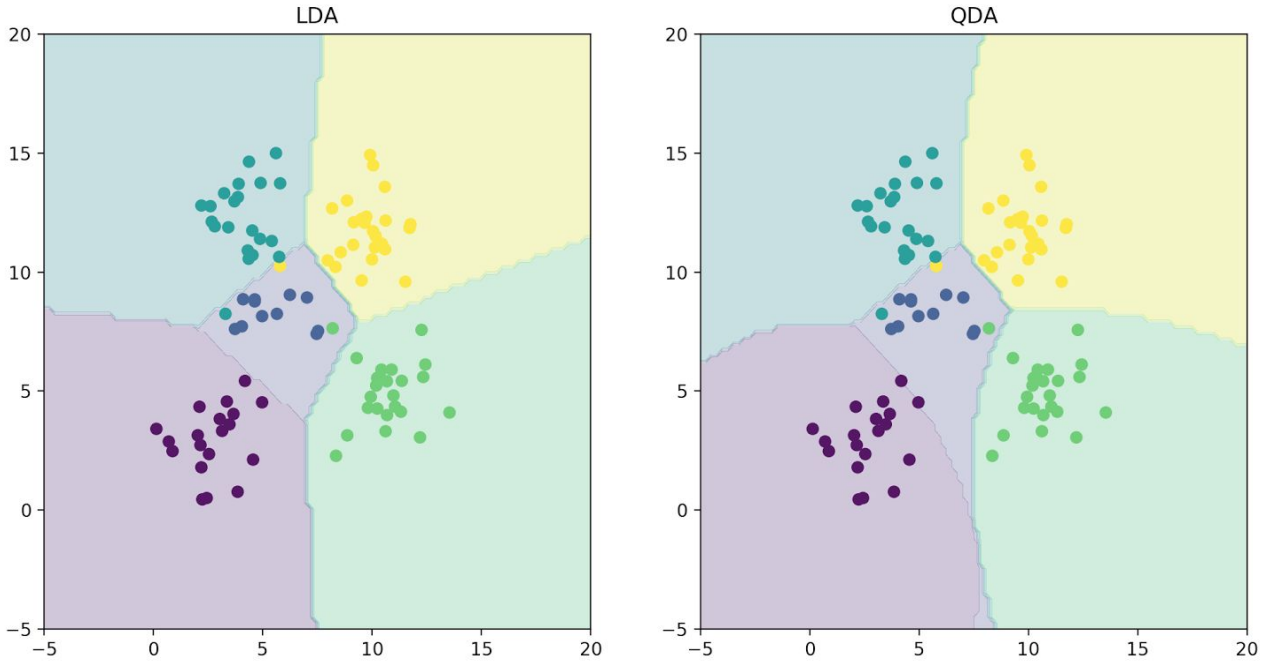
Naina Nigam (50208030)

Surabhi Singh (50208675)

Vanshika Nigam (50208031)

PROBLEM 1 : EXPERIMENT WITH GAUSSIAN DISCRIMINANTS

- **LDA Accuracy : 97%**
- **QDA Accuracy : 96%**



The above plots show decision boundaries for Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) . They depict LDA can learn only linear boundaries while QDA can learn quadratic boundaries. Both QDA and LDA are based on same probability function but the difference is that LDA assumes the same covariance for all classes i.e. it computes covariance over the whole training data while computing different means for each class, on the other hand QDA computes a separate covariance and mean for each possible output class.

PROBLEM 2 : EXPERIMENT WITH LINEAR REGRESSION

- **MSE** without intercept on Test Data **: 106775.361558**
- **MSE** with intercept on Test Data**: 3707.84018132**
- **MSE** without intercept on Training Data**: 19099.4468446**
- **MSE** with intercept on Training Data: **2187.16029493**

Inference : Linear Regression with intercept performs much better as compared to linear regression without intercept for both training and test data.

PROBLEM 3 : EXPERIMENT WITH RIDGE REGRESSION

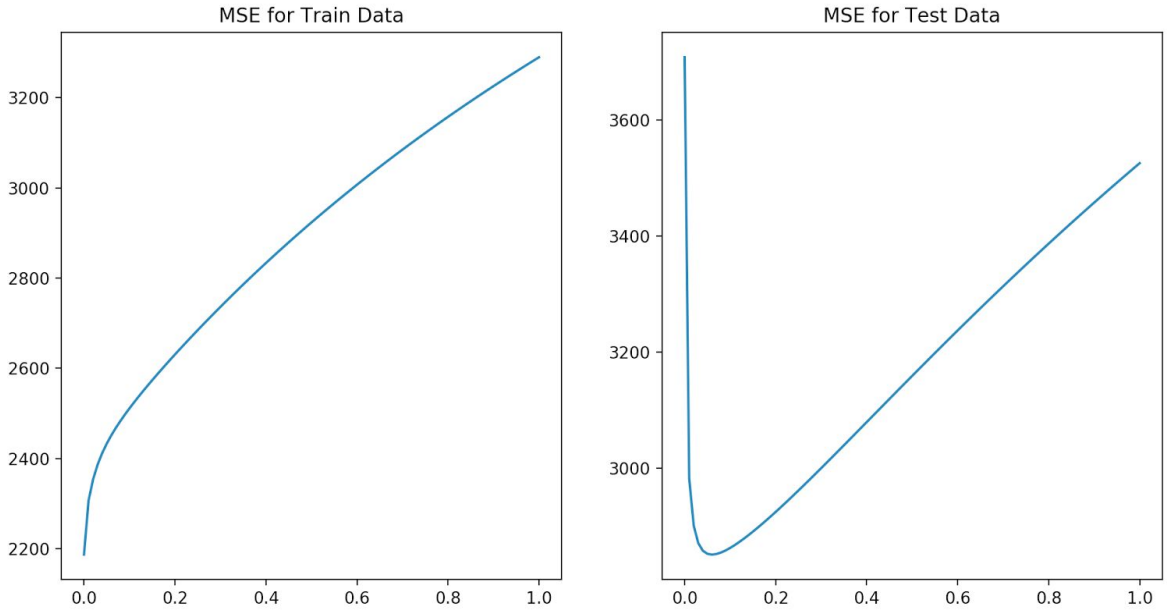
MSE with Ridge Regression with intercept for Lambda=0.06 which qualifies to be the optimal solution on two data sets are as follows:

- **Training Data: 2451.52849064**
- **Test Data: 2851.33021344**

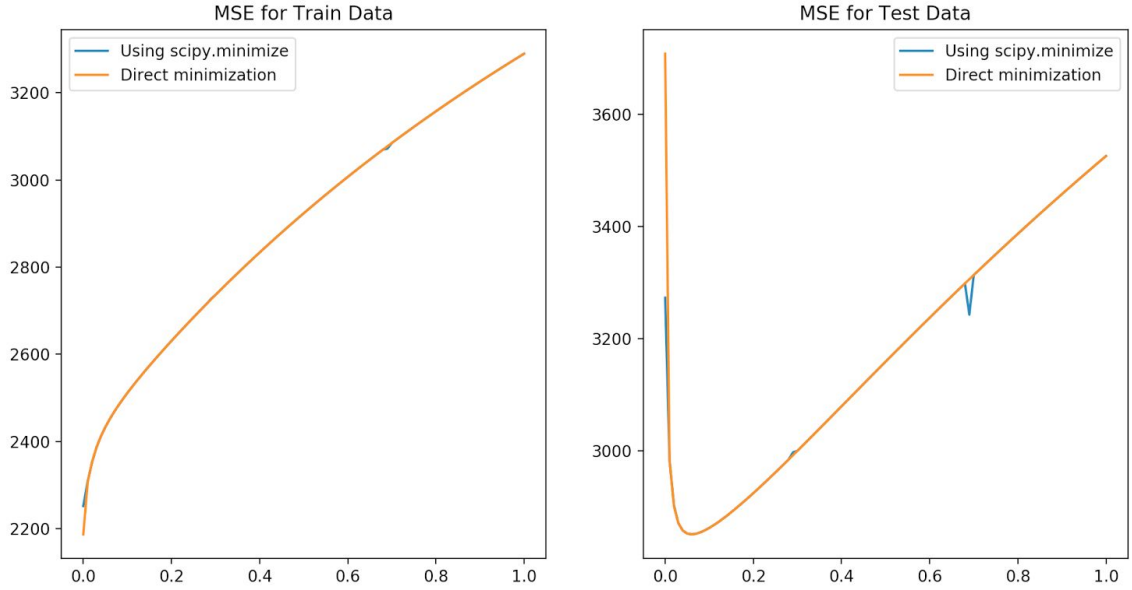
Below is the table for various values of lambda (0-0.2)

Lambda	MSE on Test Data	MSE for Train Data
0	3707.840181	2187.160295
0.01	2982.44612	2306.832218
0.02	2900.973587	2354.071344
0.03	2870.941589	2386.780163
0.04	2858.00041	2412.119043
0.05	2852.665735	2433.174437
0.06	**2851.330213**	**2451.528491**
0.07	2852.349994	2468.077553
0.08	2854.879739	2483.365647
0.09	2858.444421	2497.740259
0.1	2862.757941	2511.432282
0.11	2867.637909	2524.600039
0.12	2872.962283	2537.3549
0.13	2878.645869	2549.776887
0.14	2884.626914	2561.924528
0.15	2890.85911	2573.841288
0.16	2897.306659	2585.559875
0.17	2903.941126	2597.105192
0.18	2910.739372	2608.4964
0.19	2917.682164	2619.748386
0.2	2924.753222	2630.872823

Below is the graphical interpretation for the same.



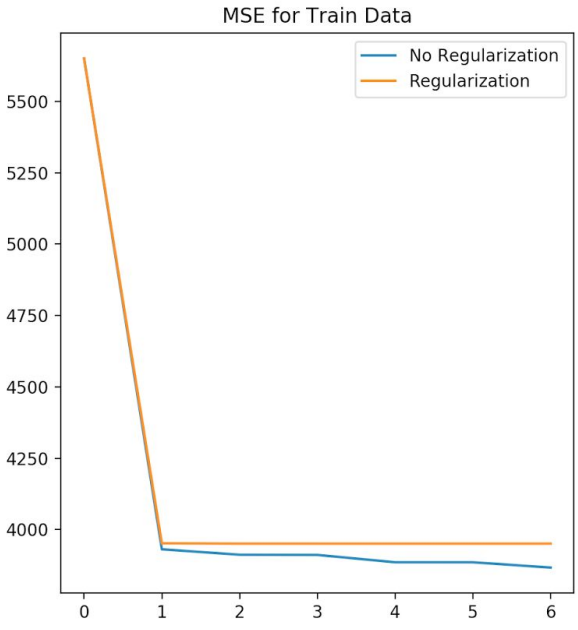
PROBLEM 4: USING GRADIENT DESCENT FOR RIDGE REGRESSION LEARNING

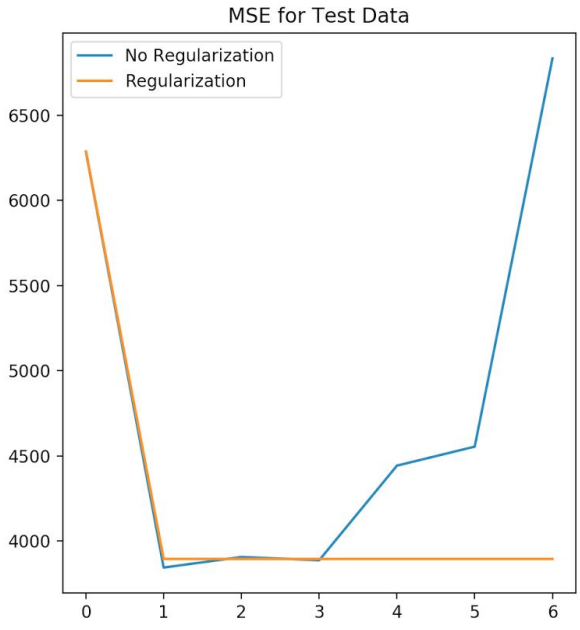


We can observe that the graphs plotted for training and testing set for Ridge Regression and Gradient Descent based Ridge Regression does not differ much. Moreover, the optimal lambda values also does not vary much for both cases.

PROBLEM 5 : NON-LINEAR REGRESSION

	MSE on Test Data		MSE on Train Data	
P	Lambda=0.0	Lambda=0.06	Lambda=0.0	Lambda=0.06
0	6286.404792	6286.881967	5650.710539	5650.711907
1	3845.03473	3895.856464	3930.915407	3951.839124
2	3907.128099	3895.584056	3911.839671	3950.687312
3	3887.975538	3895.582716	3911.188665	3950.682532
4	4443.327892	3895.582668	3885.473068	3950.682337
5	4554.830377	3895.582669	3885.407157	3950.682335
6	6833.459149	3895.582669	3866.883449	3950.682335





It can be observed from the given graph and table that when lambda is 0, the optimal value of p is 1 and when lambda is 0.06(the optimal value of lambda), p is 4.

PROBLEM 6 : INTERPRETING RESULTS

- **LINEAR REGRESSION :**

  MSE without intercept on Test Data : 106775.361558

  MSE with intercept on Test Data: 3707.84018132

  MSE without intercept on Training Data: 19099.4468446

  MSE with intercept on Training Data: 2187.16029493

- **RIDGE REGRESSION**

  For Optimal lambda value : 0.06 following are the observations :

  MSE on Training Data: 2451.52849064

  MSE on Test Data: 2851.33021344

- **NON LINEAR REGRESSION**

  For lambda =0, p=1:

  - MSE on Training Data =3930.915407
  - MSE on Testing Data = 3895.856464

  For lambda =0.06,p=4 :

  - MSE on Training Data=3950.682335,
  - MSE on Testing Data =3895.582669

CONCLUSION

Thus after the assignment following observations can be made :

1. Lower values of MSE(Mean Squared Error) indicate better fit as this indicates how close the observed data points are to the model's predicted values.
2. The Linear Regression with intercept is better for training data.
3. The Ridge Regression with intercept is better for testing data getting us a better fitting.