CSE 574 Machine Learning – Assignment 1

Submitted by
1. Sajid khan (50248743)
2. Vinay Vardhaman (50248852)

Problem 1: Experiment with Gaussian Discriminators:

LDA Accuracy = 97.0
QDA Accuracy = 96.0

LDA uses single covariance matrix for all class calculations whereas QDA uses separate covariance matrix for every class. This behavior results in the following plots where LDA has straight lines for discriminating boundaries and QDA has curved boundaries for different classes.

The boundary plots for LDA and QDA are shown below.

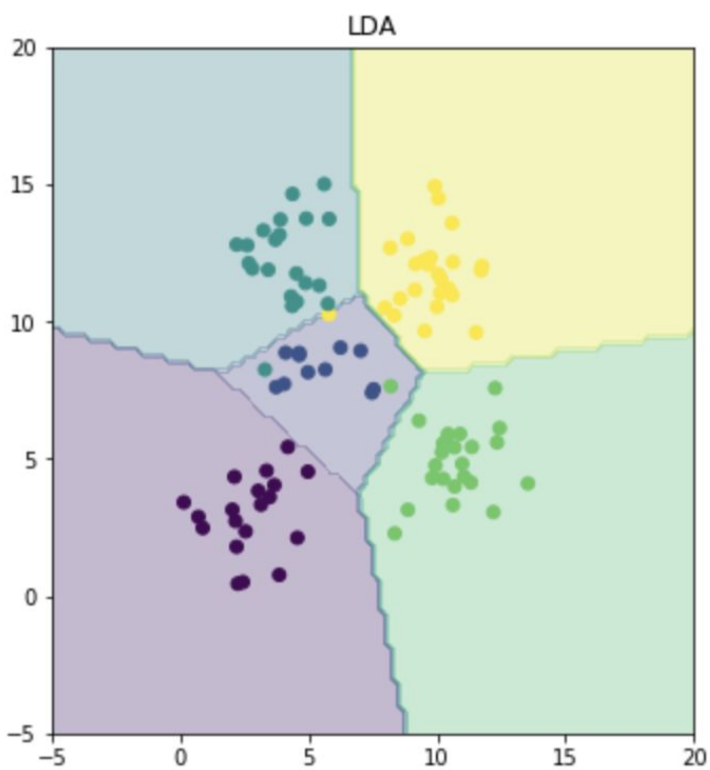

Fig 1

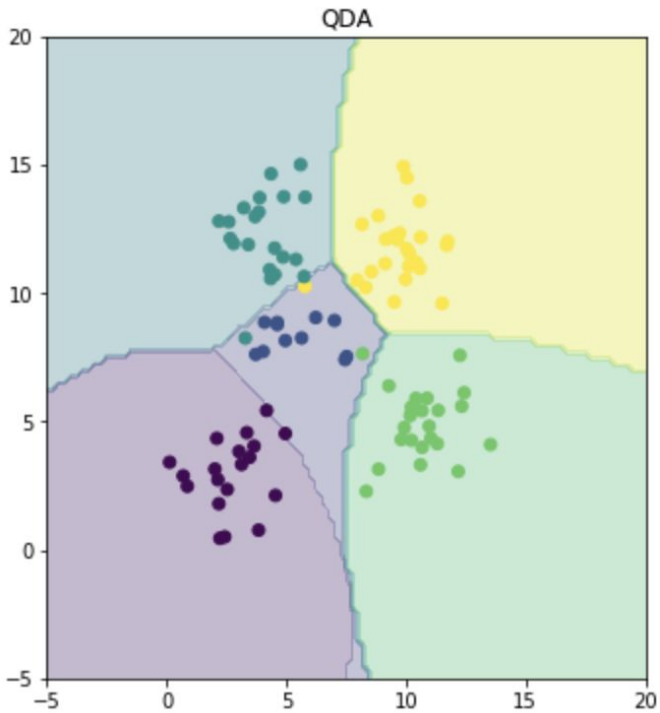

Fig 2

Problem 2: Experiment with Linear Regression

Data	MSE without Intercept	MSE with Intercept
Training Data	19099.44684457	2187.16029493
Testing Data	106775.3615526	3707.8401812

From the observations, it can be observed that MSE is the lowest when using a bias-intercept for both Training and Testing Data.

Problem 3: Experiment with Ridge Regression

Lambda	Testing MSE	Training MSE
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

The optimal value of lambda for which Testing data has lowest MSE is 0.06

MSE with intercept using OLE Approach

Test Data – 3707.840181
Training Data – 2187.160295

MSE with intercept using Ridge Regression Approach for optimal lambda (=0.06)
Test Data – 2851.330213
Training Data – 2451.528491

From the above results, it can be observed that testing data error can be reduced by using ridge regression approach. Thus, ridge regression is a better approach than the OLE approach for the current problem.

The plot below shows MSE plotted with Training and Test data against lambda

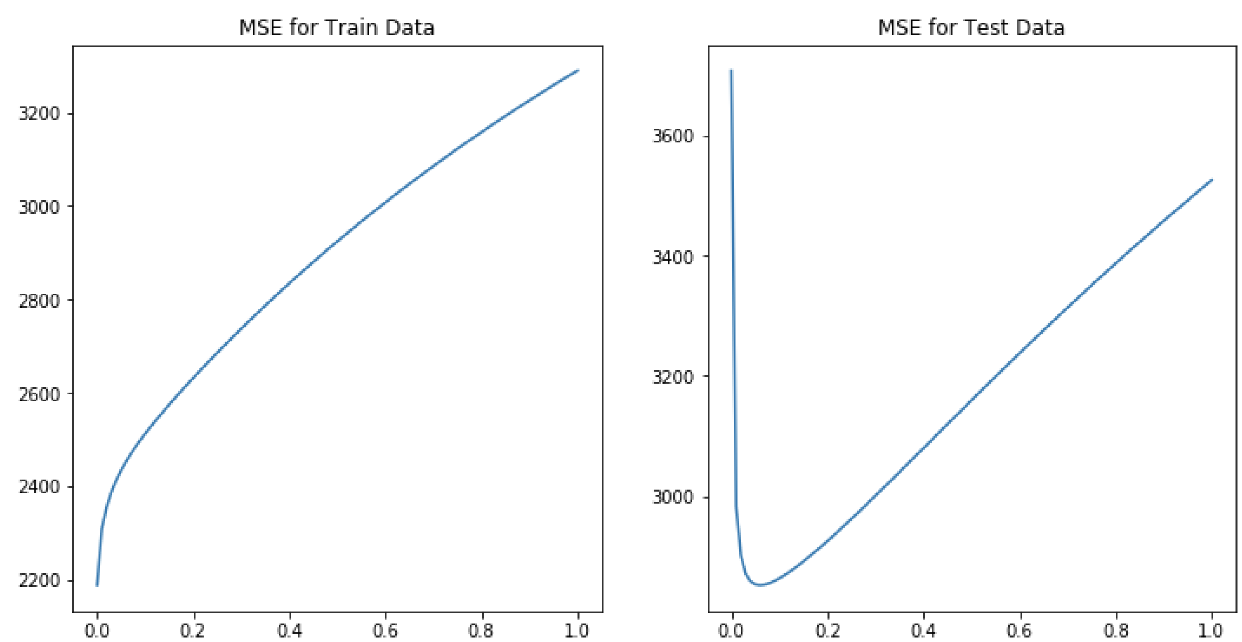

Fig 3

Problem 4: Using Gradient Descent for Ridge Regression Learning

Lambda	Testing MSE	Training MSE
0	2900.5515	2900.549034
0.01	2935.6472	2825.234646
0.02	2839.2449	2839.244879
0.03	2825.6953	2841.285801
0.04	2848.3401	2872.409744
0.05	2836.1898	2844.861158
0.06	2877.1543	2855.990849
0.07	2846.093	2878.484354
0.08	2901.0847	2862.58984
0.09	2861.4713	2861.47173
0.1	2892.2993	2892.298491
0.11	2876.874	2871.12665
0.12	2876.3309	2871.35036
0.13	2881.2384	2881.727064
0.14	2881.9192	2897.091439
0.15	2893.4715	2889.524927
0.16	2897.0191	2897.019364
0.17	2903.129	2903.476315
0.18	2912.1917	2907.728725
0.19	2920.4867	2920.486653

0.2	2924.9315	2925.385806

The optimal value of lambda using Gradient descent has lowest MSE for Testing data at 0.03

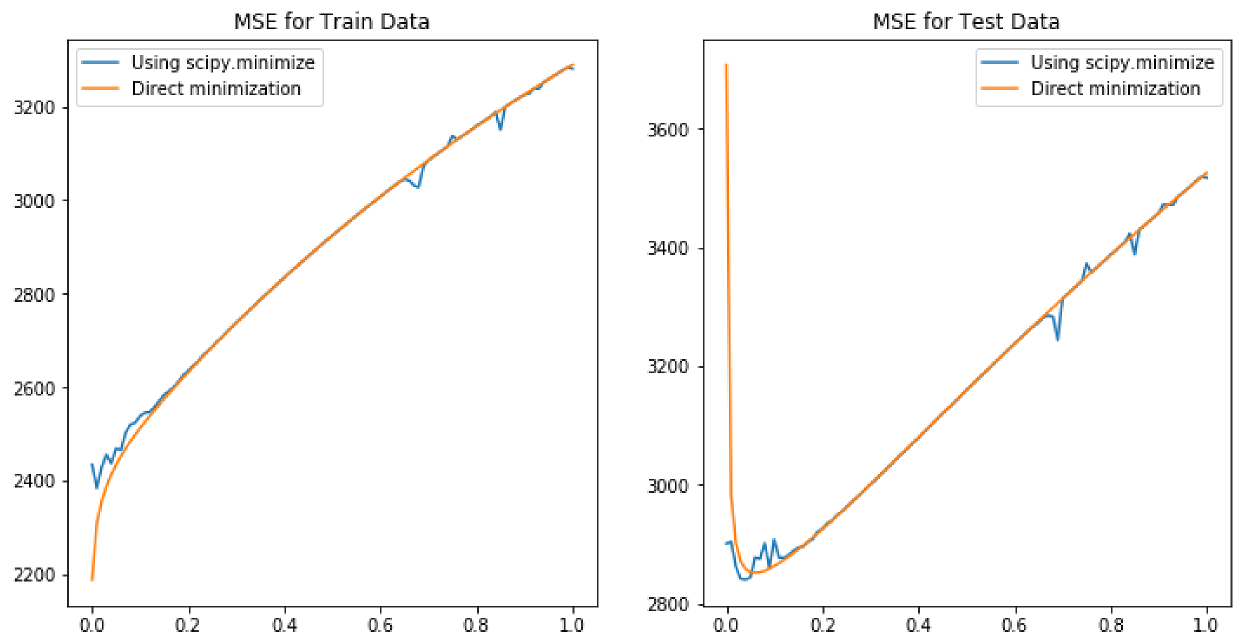

Fig 4

The graph above plots lambda against Training and Test data for Ridge regression and Gradient descent learning. It can be observed that plots for Gradient descent are almost similar to ridge regression learning in problem 3. However, the lowest MSE for Training data can be observed at 0.06 and 0.03 for Ridge regression with optimal lambda and gradient descent respectively.

Problem 5: Non-linear Regression

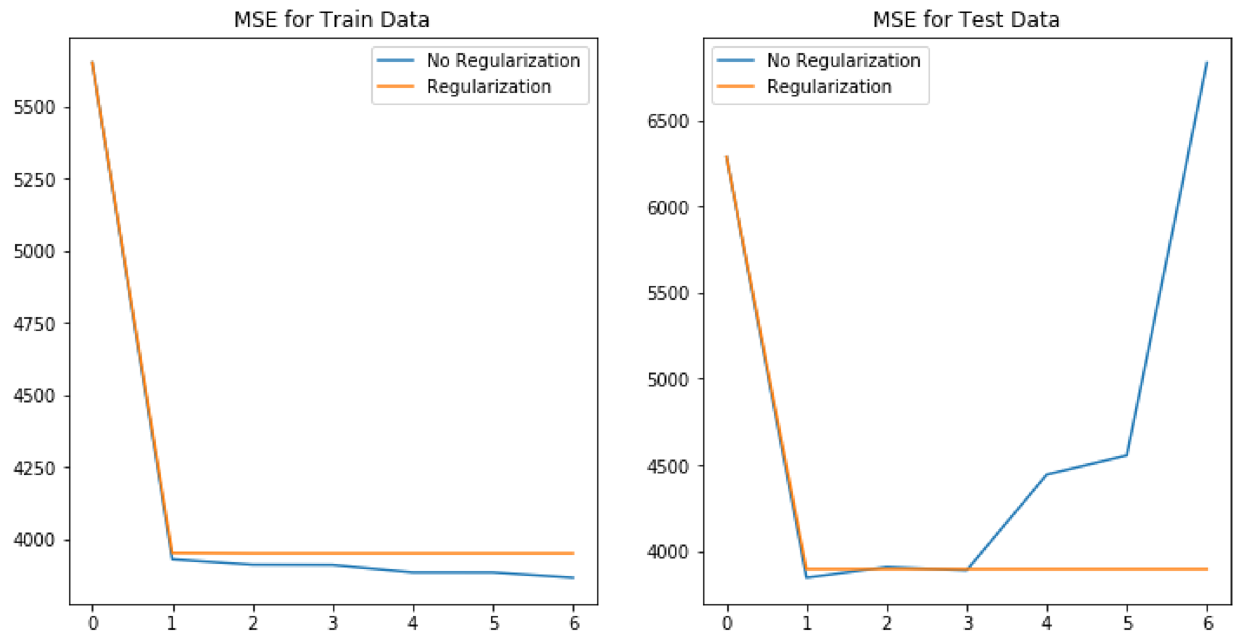

Fig 5: MSE for train data and test data when λ = 0 and λ = 0.06 (calculated from problem 3)

When λ =0 the minimum MSE can be observed from the graphs.

For Training Data MSE is minimum at P = 6.
For Testing Data MSE is minimum at P = 1.

For λ = 0.06(Optimal value calculated from problem 3)

P	MSE for Training Data	MSE for Testing Data
0	5650.711907	6286.881967
1	3951.839124	3895.856464
2	3950.687312	3895.584056
3	3950.682532	3895.582716
4	3950.682337	3895.582668
5	3950.682335	3895.582669

6	3950.682335	3895.582669

From the above table the optimal values of P in terms of test error are found to be
For Training Data: p = 6
For Testing Data: p = 4

Problem 6: Interpreting Results

Procedure	Testing MSE	Training MSE
OLE regression without Intercept	106775.3615526	19099.44684457
OLE regression with Intercept	3707.8401812	2187.16029493
Ridge regression with optimal lambda	2851.330213	2451.528491
Ridge regression with Gradient Descent	2825.6953	2841.285801
Non Linear Regression without regularization	3845.03473017	3866.88344945
Non Linear Regression with regularization	3895.58266828	3950.68233514

Conclusion:
From the table above, it can be concluded that **Ridge regression with Gradient descent** has the lowest value of MSE for Testing data. Contrarily, **OLE regression with Intercept** has the lowest MSR for Training data.