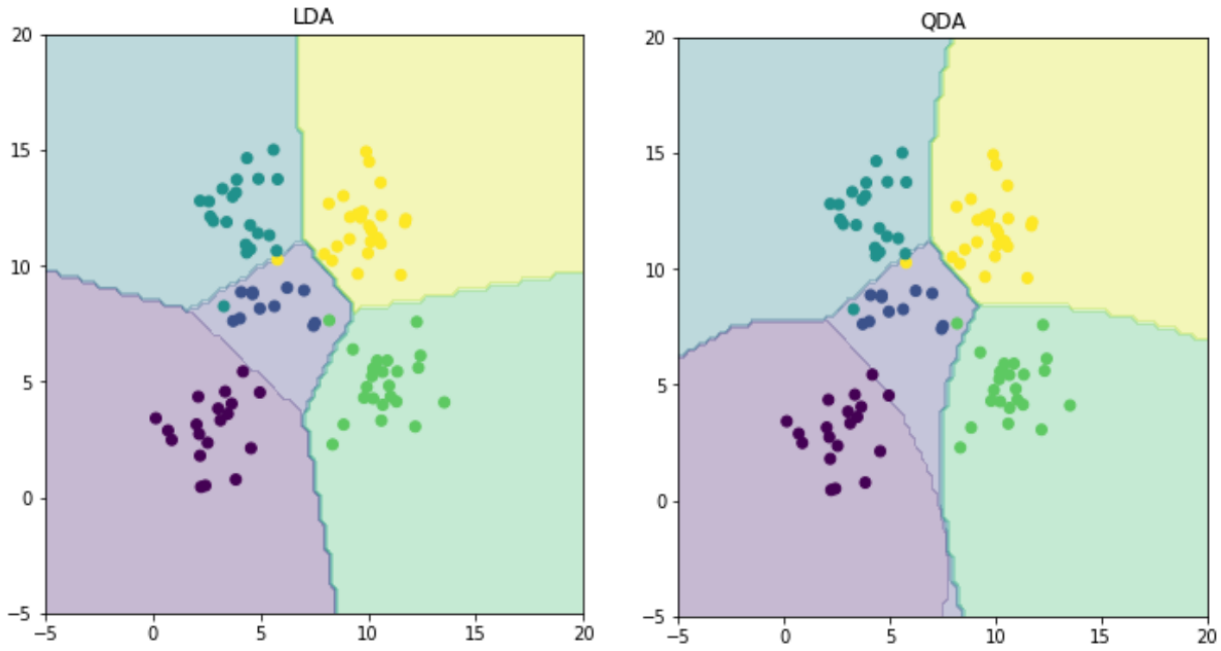
CSE 474 – Programming Assignment 2
Classification and Regression

Group 28
Team Members: Ankit Nerlakanti, Vikram Garu, and Veronica Ng

Problem 1: Experiment with Gaussian Discriminators

Accuracy of LDA: 97.0
Accuracy of QDA: 96.0



Our reported accuracies of LDA and QDA shows that the accuracy of LDA is greater than that of QDA. The reason for this is because when calculating the LDA, the same covariance is used throughout all the classes. On the other hand, when calculating QDA, there are different covariance for each class. This can be seen on the plots where the LDA plot has lines which are more linear while the QDA plot has lines which are more curved. Thus showing that the accuracy of LDA would be greater than that of the accuracy of QDA.

Problem 2: Experiment with Linear Regression

Case 1 – Without using an intercept (or bias term)
Training Data: 19099.446844570863
Test Data: 106775.36155227982

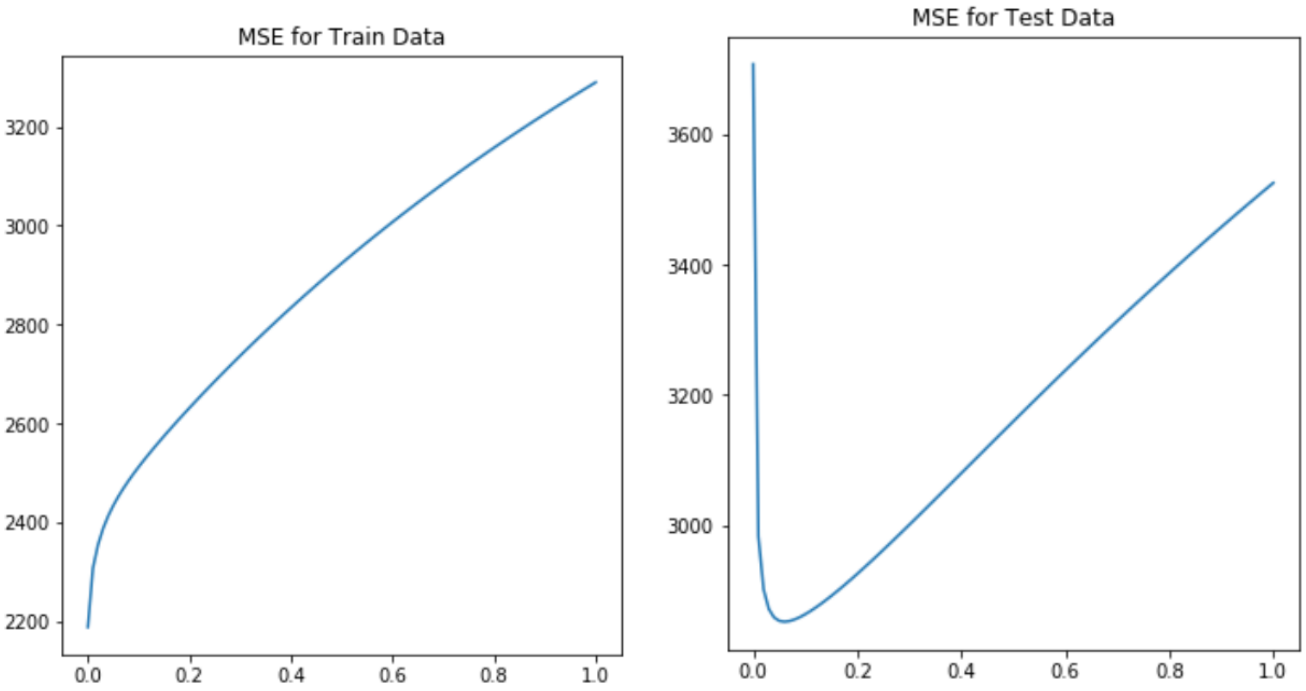
Case 2 – With using an intercept
Training Data: 2187.1602949303892

Test Data: 3707.8401817889676

After the calculations of the MSE for each case, we can see that the case with using an intercept is better than the case without using an intercept. It clearly shows that when using an intercept there is more error as compared to when using an intercept. Thus, the MSE when using an intercept is better.

Problem 3: Experiment with Ridge Regression

OLE Regression using an intercept:
Training Data: 2187.1602949303892
Test Data: 3707.8401817889676

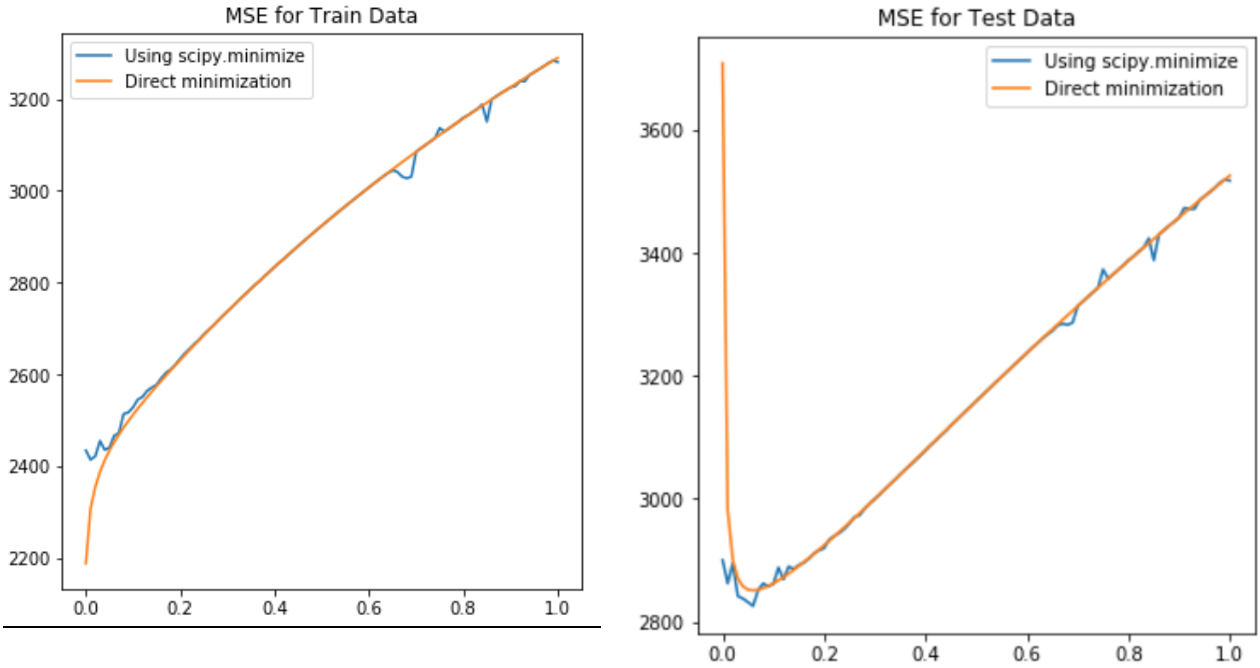


Based on observations, Ridge Regression using an intercept is better than OLE Regression using an intercept as there are less errors in Test Data.

The optimal lambda would be 0.06 as shown on the two graphs and as the Test Data was the least amongst all the other lambdas.

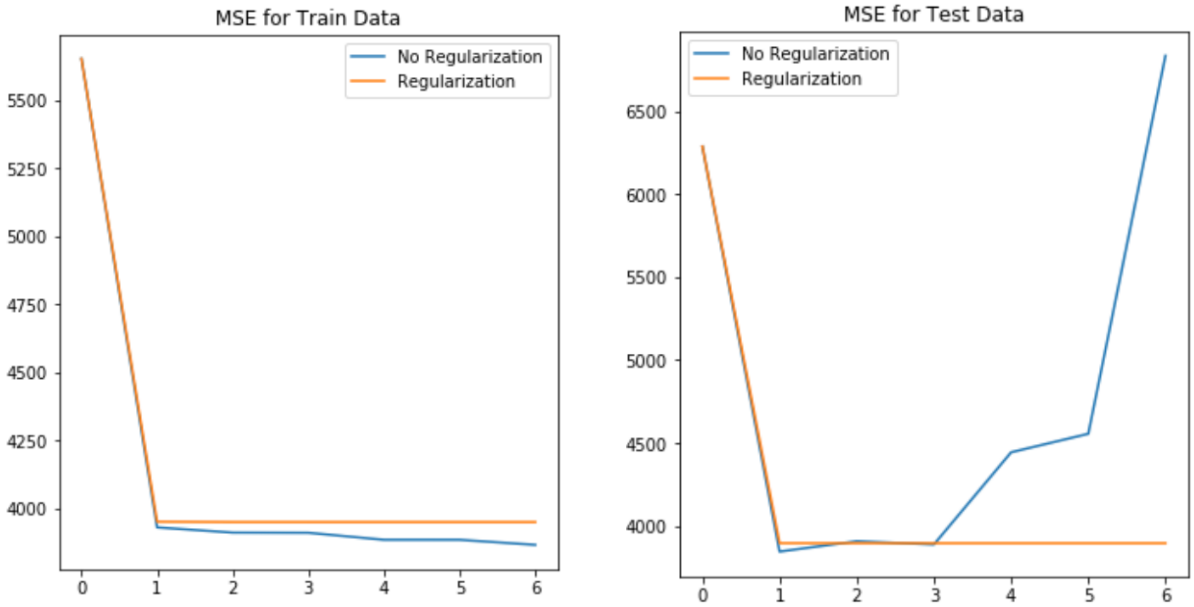
Problem 4: Using Gradient Descent for Ridge Regression Learning

Plots of the errors on the train and test data obtained by using the gradient descent:



As shown on the graphs, we can see that the trend lines are almost nearly the same. Also from the graphs we can tell that using Gradient Descent would be better.

Problem 5: Non-Linear Regression



As shown on the graph MSE for Train Data, having no regularization is better as it decreases more than with regularization. On the other hand, on the graph MSE for Test Data, having regularization is better as from p = 3 to p = 6, errors increase with no regularization. Therefore, it is better to have regularization even though for Train Data having no regularization would be better because the difference of having no regularization and regularization for Train Data is small and they are going at a steady pace with each other. As having regularization would be better, the optimal p would be p = 1 since starting from p = 1, constant values start to begin from there.

Problem 6: Interpreting Results

Based on the results obtained in Problem 4, it would be best to use Gradient Descent for Ridge Regression Learning.