

Machine Learning

Assignment 2

Yousef Jaber, Ali Hassan, Ahmad Temoor

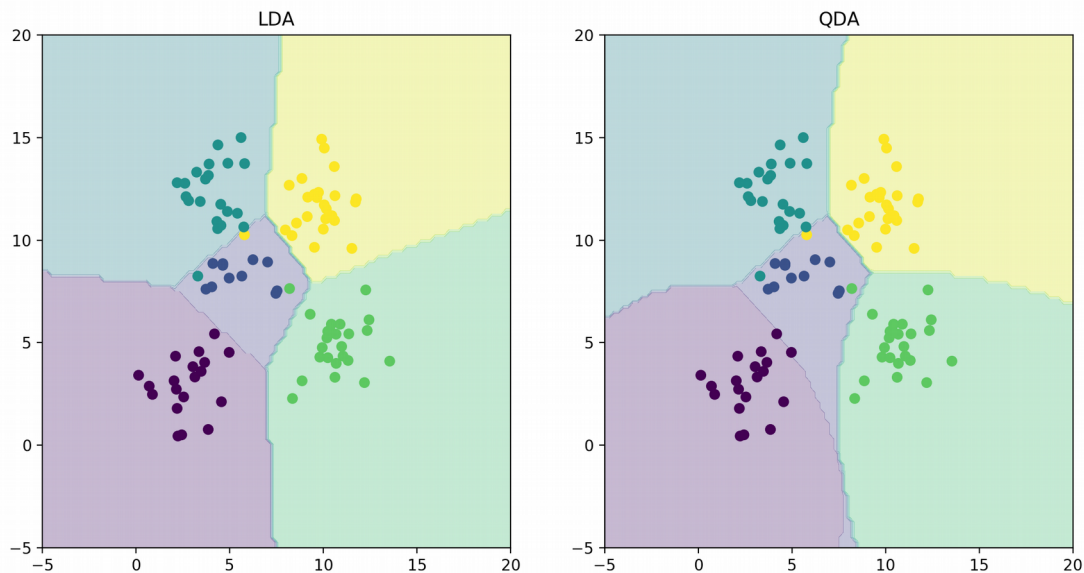
## Problem 1

**Report the accuracy of LDA and QDA on the provided test data set**

Accuracy for LDA: 97

Accuracy for QDA: 96

**Also, plot the discriminating boundary for linear and quadratic discriminators.**



**Explain why there is a difference in the two boundaries**

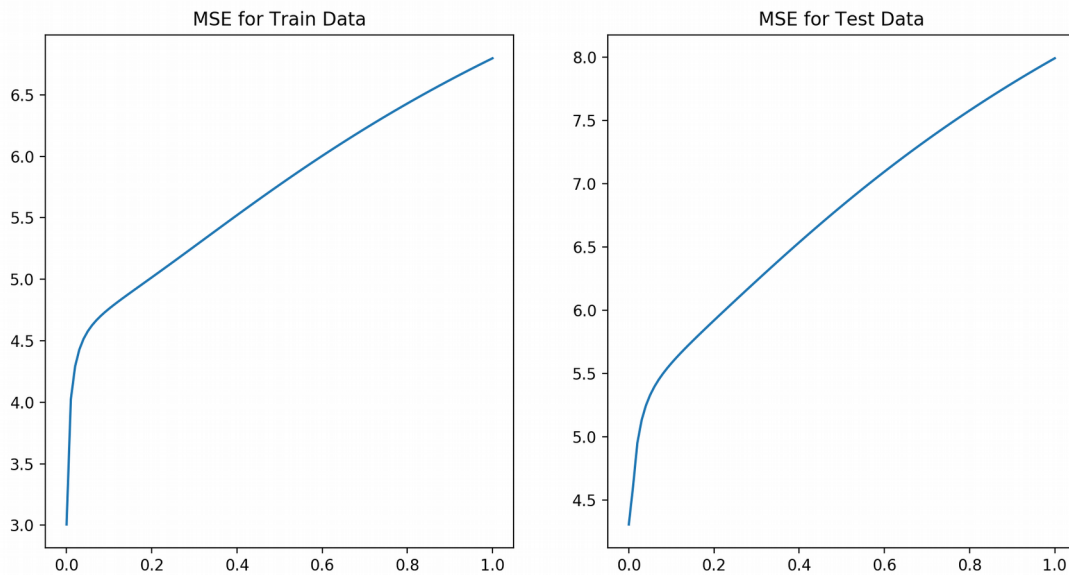
LDA boundary vs QDA boundary: LDA can only learn linear boundaries, generated by the linear equation  $D = a_1x_1 + a_2x_2 + \dots + a_ix_i + b$ . The QDA is generated by a quadratic equation, which offers more flexibility, but may prove to be more complicated as well as it deals with more complex boundaries spaces, making it slightly easier to misclassify.

In order to generate the prediction of the label, QDA utilizes covariance matrices specific to every individual class, while LDA simply analyzes the covariance of the data set as a whole.

## Problem 2

Calculate and report the MSE for training and test data for two cases: first, without using an intercept (or bias) term, and second with using an intercept

	Test Data	Training Data(Retreived by zooming in on plot)
MSE with bias term:	4.30571724	3.00630 (base of line)
MSE without bias term:	23.10577434	



**Which one is better?**

A smaller MSE value indicates a more accurate line of best fit (boundary). So according to the values above, including the intercept is better as it produces a lower MSE.

### Problem 3

Calculate and report the MSE for training and test data using ridge regression parameters using the `testOLERegression` function that you implemented in Problem 2. Use data with intercept.

Plot not generated by script

### Problem 4

Plot the errors on train and test data obtained by using the gradient descent based learning by varying the regularization parameter  $\lambda$ .

Plot not generated by script

### Problem 5

Training data error: lowest at  $p = 6$

Test data error: lowest at  $p = 1$

