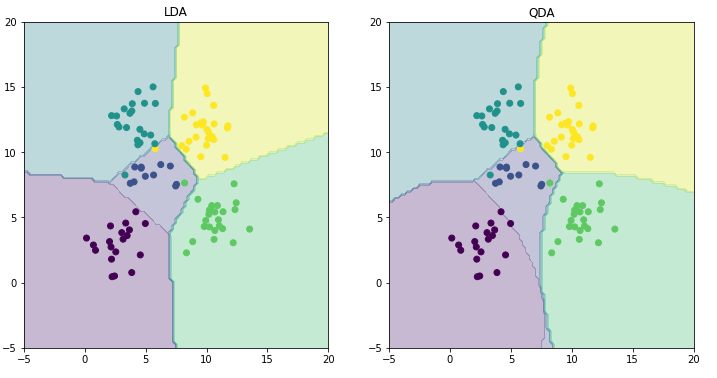
Problem 1:

LDA Accuracy = 97%

QDA Accuracy = 96%



TODO – Explain the difference in the boundaries

Problem 2

MSE without intercept mean of magnitude of weights: 0.20690517560805177

MSE with intercept mean of magnitude of weights: 0.28581612984613397

MSE without intercept 106775.36152087666

MSE with intercept 3707.8401811278804

TODO – Explain which one is better

From the above obtained results MSE with intercept is better than without intercept

Mean Squared Error obtained without intercept is 25 times more than with intercept

As the need for intercept is more, from this it can be inferred that there are many zeros in the features given.

It can also be inferred that by assigning mean of y as the output for data points which have empty elements there is downfall of Mean Squared Error(MSE).

Problem 3

Chart

Description automatically generated

TODO - Compare the two approaches in terms of errors on train and test data. What is the optimal value

for lambda and why?

From the above plotted graphs,

* MSE for Train Data:

Optimal value of Lambda is equal to “ZERO”, as the MSE is least when lambda is at 0.

It can be inferred that the increase MSE is quite proportional MSE.

MSE is at peak when lambda is 1.

* MSE for Test Data:

Optimal value of Lambda is equal to “0.05”, as the MSE is least when lambda is at 0.05.

Unlike train data regularization is required for test data

At 0 the MSE is at its peak.

The peak MSE values are,

* Train Data: >3200
* Test Data: > 3600

Least MSE values are,

* Train Data: 2200
* Test Data: <3000

From both the values it can be stated that Mean Squared Error is high in test data.

Train data don’t require regularization whereas test data requires regularization.

Problem 4

Chart, line chart, histogram

Description automatically generated

TODO - Compare with the results obtained in Problem 3.

From the above graphs it can be inferred that,

Direct Minimization

* Gradient descent didn’t make much of difference in the MSE’s obtained.
* The values obtained before and after gradient descent completely match.

Using scipy.minimize:

MSE for Train Data:

* The Mean Squared Error has increased by 200 after applying Gradient Descent, when compared with the MSE’s obtained in problem 3.
* Plot obtained is similar to the one obtained in problem 3 but there are many troughs and peaks with gradient descent.

MSE for Test Data:

* Gradient descent caused a small variation in MSE’s associated to lambda’s similar to the variation caused in train data.
* The structure of plot is infrequent with many troughs and peaks.

The plots and values are comparable with the ones obtained in problem 3.

Problem 5

Chart, line chart

Description automatically generated