**MACHINE LEARNING PROJECT# 1 – GROUP 70**

**Group Members:**

Kaustuv Majumder ([kaustuvm@buffalo.edu](mailto:kaustuvm@buffalo.edu))

Venkat Ashish Nagabhirava (nagabhir@buffalo.edu)

Ieshaan Sharma

**Problem 1:**

LDA Accuracy = 97%

Observations:

• We observe discriminant boundaries for LDA which differentiate 5 classes.

• In LDA, the covariance for each class is assumed to be the same.

• Boundaries plotted are linear.

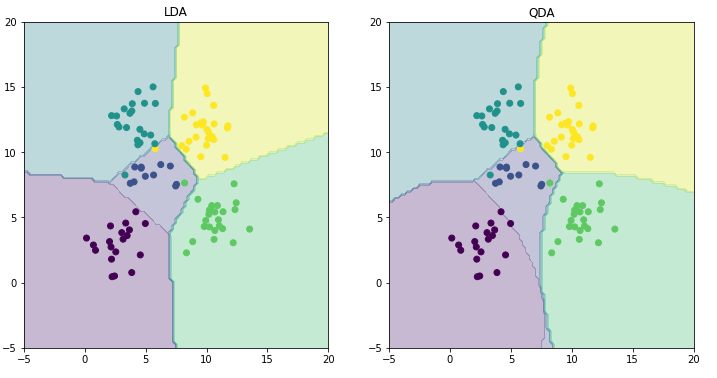
QDA Accuracy = 96%

Observations:

• We observe discriminant boundaries for QDA which differentiate 5 classes.

• In QDA, the covariance for each class is calculated separately.

• Boundaries plotted are parabolic (quadratic decision boundaries).



**Problem 2:**

MSE without intercept: 106775.36152087666

MSE with intercept: 3707.8401811278804

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

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

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:

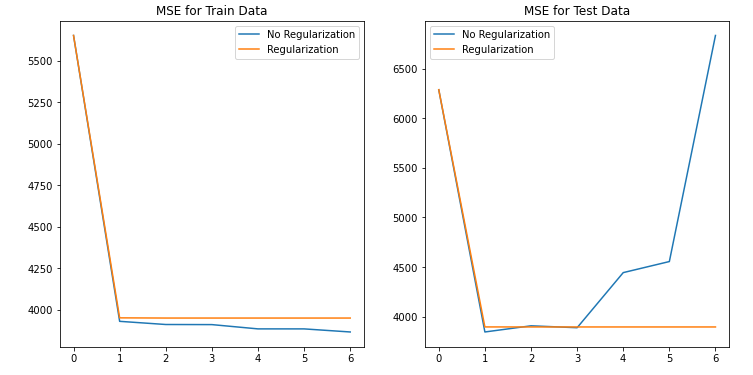
* The Mean Squared Error has increased by 200 after applying Gradient Descent, when compared with the MSE’s obtained in problem 3.
* Plot obtained is similar to the one obtained in problem 3 but there are many troughs and peaks with gradient descent.

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**



Inference from the above graphs:

i) MSE for Train Data:

From the above graph we can see that, when we have performed regularization, with increase in p values, there is not much difference in the MSE values. Whereas, when we do not have regularization i.e lambda = 0, the MSE values for train data decreases with increase in polynomial degree. This is because as the curve becomes non-linear, It tries to reduce the MSE, by adjusting according to the train data.

ii) MSE for Test Data:

From the above graph for Test data, we can see that, In case when there is no regularization, as the p value increases, there is a drastic increase in the MSE for test data. This is because, when regularization was not performed, the model gets adjusted according to the train data and with change in data (i.e with test data), the MSE increases drastically since It was made according to the test data. When regularization is performed, the Test MSE is also consistent. Thus, we should perform regularization as shown above.

Problem 6:

After comparing all the approaches,

Linear Regression,

Without Intercept:106775.36152087666

With Intercept:3707.8401811278804

|  |  |  |
| --- | --- | --- |
| Model Approach | Train MSE(least values) | Test MSE(least values) |
| With Ridge Regression | Optimal Lambda 0  2200(approx.) | Optimal Lambda 0.05(approx.)  2500(approx.) |
| Ridge Regression with Gradient Descent | Optimal Lambda 0  2200(approx.) | Optimal Lambda 0.05(approx.)  2500(approx.) |
| Non-Linear Regression (No Regularization) | P=6  3800(approx.) | P=1  3800(approx.) |
| Non-Linear Regression (Regularization) | P=1-6  3900(approx.) | P=1-6  3900(approx.) |

From the above obtained values,

* Our main objective is to reduce the Mean Squared Error on the test data, there has been improvement seen by adding the intercept and ridge regression to the test data.
* MSE has shown acceptable error estimation.
* RMSE( Root Mean Squared Error) has also been used as an error estimating metric but there is no difference seen the plots and values obtained.
* Regularization has shown downfall in the MSE for test data.
* No impact is made by Gradient Descent on the MSE of test data.
* Nonlinear Regression have increased the MSE on the whole when compared to linear regression.