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

**Group Members:**

Kaustuv Majumder ([kaustuvm@buffalo.edu](mailto:kaustuvm@buffalo.edu)) (50360369)

Venkat Ashish Nagabhirava ([nagabhir@buffalo.edu](mailto:nagabhir@buffalo.edu)) (50388929)

Ieshaan Sharma ([ieshaans@buffalo.edu](mailto:ieshaans@buffalo.edu)) (50367364)

**Problem 1: LDA & QDA**

LDA Accuracy = 97%

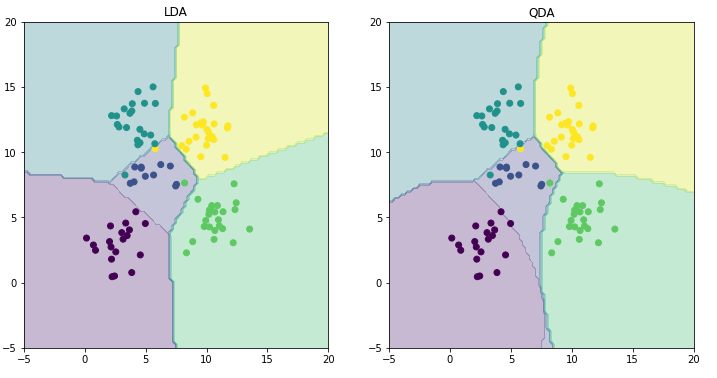
Observations:

* 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: Linear Regression**

Test Data

* MSE without intercept: 106775.36152087666
* MSE with intercept: 3707.8401811278804

Train Data

* MSE without intercept: 19099.446844570663
* MSE with intercept: 2187.160294930391

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: Ridge Regression**

Chart

Description automatically generated

From the above plotted graphs,

* MSE for Train Data:

The MSE is least when lambda is at 0.

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

MSE is at peak when lambda is 1.

* MSE for Test Data:

Optimal value of Lambda is equal to “0.06 ”, as the MSE is least.

Unlike train data, regularization is required for test data.

At 0 the MSE is at its peak.

The peak MSE values are,

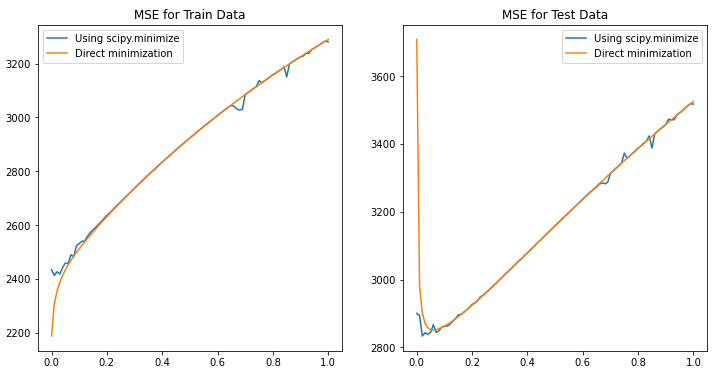
* Train Data: >3200
* Test Data: > 3600

Least MSE values are,

* Train Data: 2200
* Test Data: <3000

**The Optimal regularization parameter (lambda) will be considered is 0.06 because for Test Data, MSE is least at lambda =0.06.**

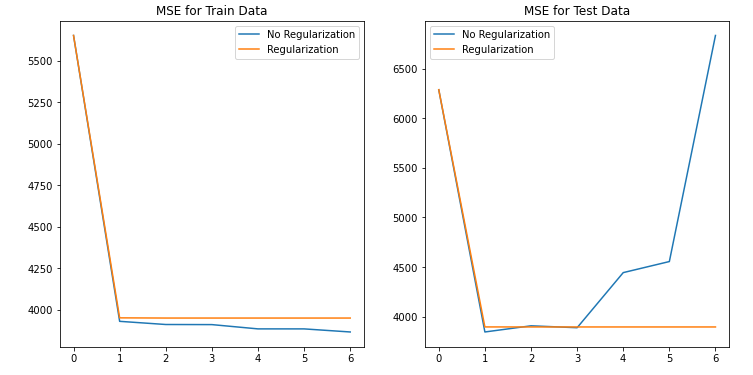
**Problem 4: Gradient Descent**



From the above graph, it can be said that training and testing error for ridge regression using gradient descent is almost same with the error obtained in problem 3 (without gradient descent).

However, distortions can be observed for small and large values of lambda. This may be attributed to the use of minimize function that can slow down the process considerably.

**Problem 5: Non-Linear Regression**



Inference from the above graphs:

Regularization: Lambda considered is 0.06.

No Regularization: Lambda is 0.

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, 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.

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: Overall Interpretation**

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 0.06(approx.)  2500(approx.) |
| Ridge Regression with Gradient Descent | Optimal Lambda 0  2200(approx.) | Optimal Lambda 0.06(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.