Franklin Pinnock

James Hsu

Yuzhe Wang

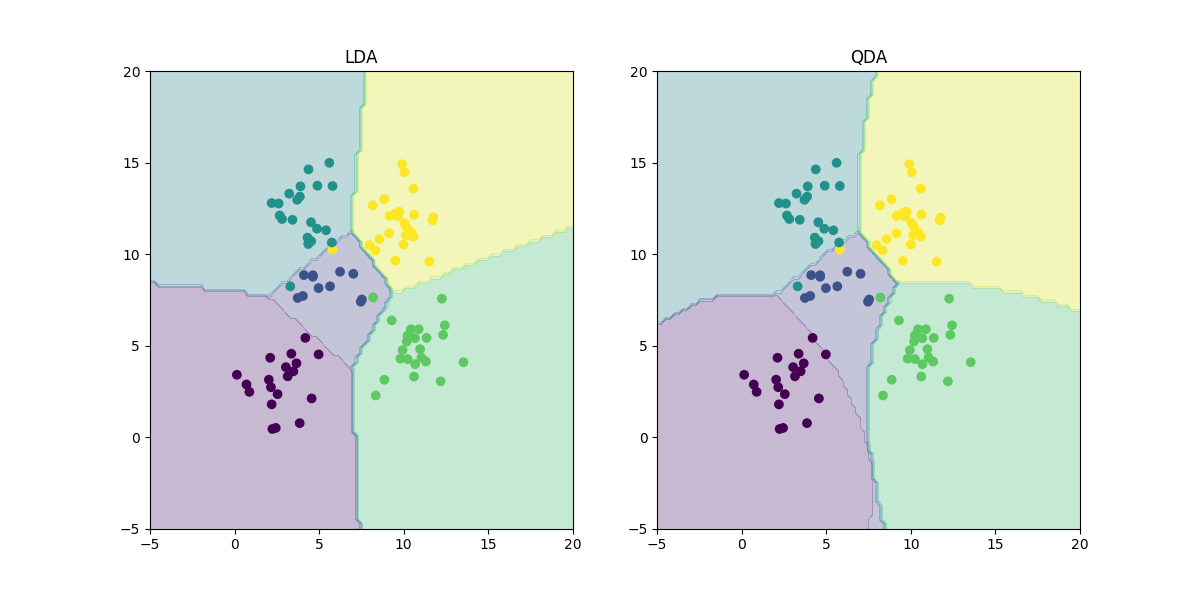
Group 30

CSE474 Introduction to Machine Learning

**Programming Assignment 2**

*Classification and Regression*

**Problem 1: Gaussian Discriminators**



LDA Accuracy = 0.97

QDA Accuracy = 0.96

The reason why there are difference between LDA and QDA is because LDA is linear and is only learning and testing a single d\*d covariance matrix while QDA is quadratic and is learning a list od d\*d covariance matrixes.

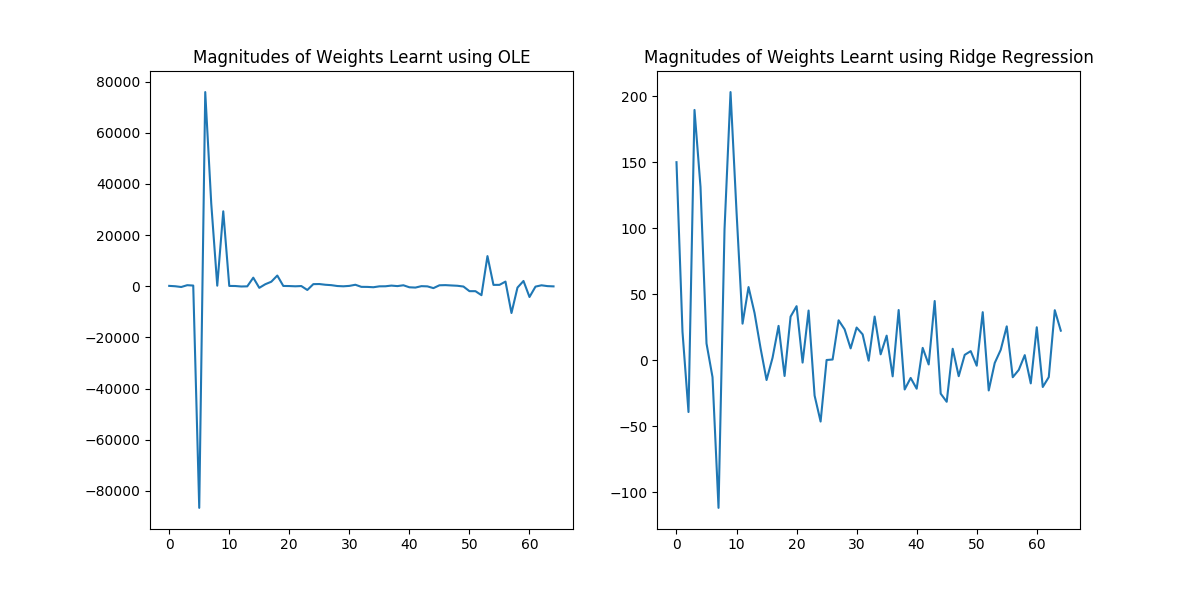
**Problem 2: Linear Regression**

**MSE using OLE**

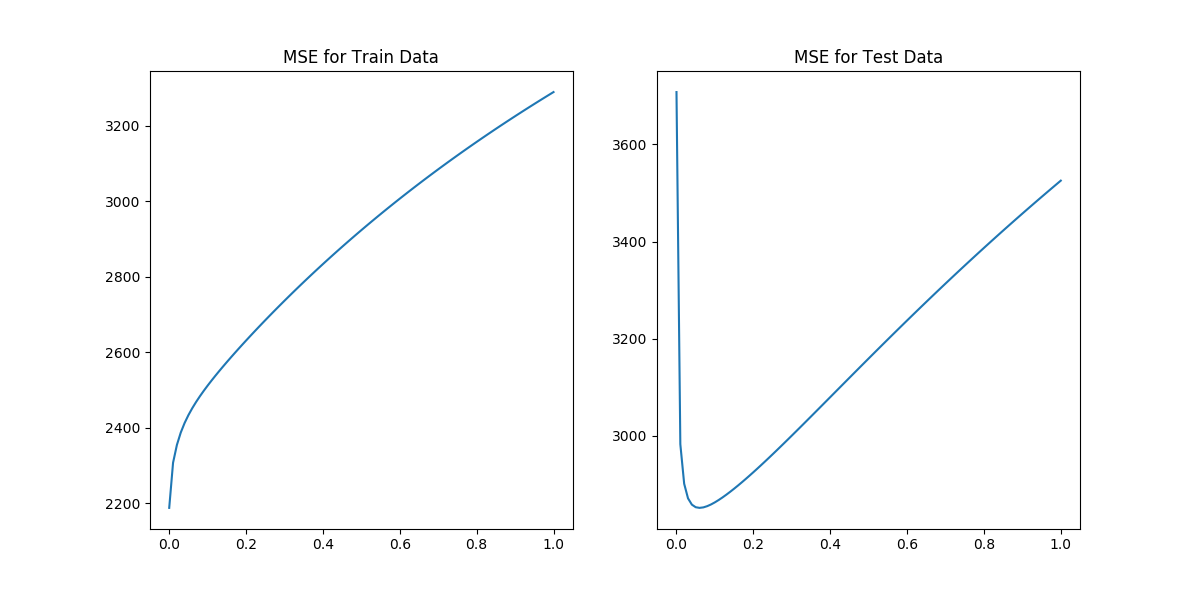
|  |  |  |
| --- | --- | --- |
|  | without intercept | with intercept |
| Training Data | 19099.44684457 | 2187.16029493 |
| Test Data | 106775.36155512 | 3707.84018132 |

MSE with intercept is better since lesser Mean Square Error is better. MSE with intercept in both test data and training data has a lower value than MSE without intercept. Therefore, we choose MSE with intercept is better.

**Problem 3: Ridge Regression Learning**



* The magnitudes of the weights for OLE are way larger than for Ridge Regression because the regularization term does not favor high weights.



**MSE using OLE**

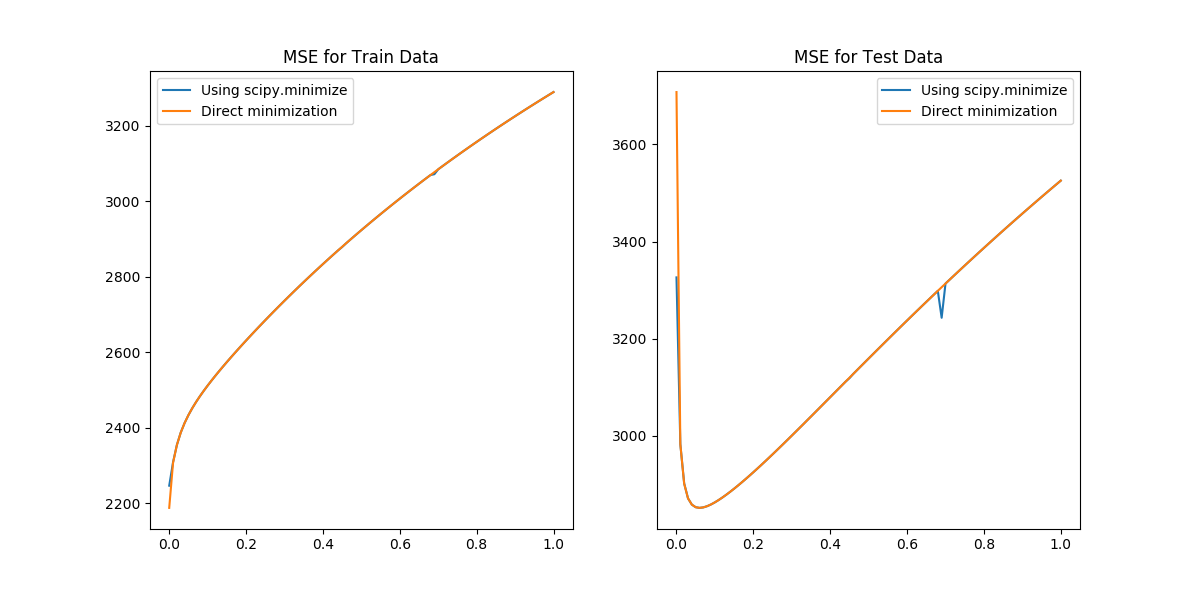
|  |  |  |
| --- | --- | --- |
|  | without intercept | with intercept |
| Training Data | 19099.44684457 | 2187.16029493 |
| Test Data | 106775.36155512 | 3707.84018132 |

**MSE using Ridge Regression (using optimal lambda = 0.06)**

|  |  |  |
| --- | --- | --- |
|  | Training Data | Test Data |
| with intercept | 2851.33021344 | 2451.52849064 |

* The optimal value for lambda is 0.06 since this value minimizes our MSE for the data.
* When comparing OLE vs Ridge Regression for Training and Test Data the MSE for Ridge Regression is significantly lower than OLE therefore Ridge Regression is the preffered way to learn weights for this data set.

**Problem 4: Using Gradient Descent**



* It looks very similar to problem 3, its almost identical; however, the result from problem 4 is not as smooth as the result from problem 3.

**Problem 5: Non-linear Regression**



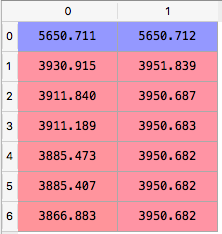
Train Data:

**p Lambda=0 Optimal Lambda**



Test Data:

**p Lambda=0 Optimal Lambda**



Train Data:

Lambda=0 is 3845.035 when p=1

Optimal Lambda is 3095.583 when p=6

Test Data

Lambda=0 is 3885.407 when p=5

Optimal Lambda is 3950.682 when p=6

Problem 6

We can use MSE as a metric to compare OLE and Ridge Regression to weigh out the better approach.

OLE Regression:

**MSE using OLE**

|  |  |  |
| --- | --- | --- |
|  | without intercept | with intercept |
| Training Data | 19099.44684457 | 2187.16029493 |
| Test Data | 106775.36155512 | 3707.84018132 |

Ridge Regression:

**MSE using Ridge Regression (using optimal lambda = 0.06)**

|  |  |  |
| --- | --- | --- |
|  | Training Data | Test Data |
| with intercept | 2851.33021344 | 2451.52849064 |

**Conclusion:**

The metric that should be used to choose the best setting is MSE. We can see from our data that Ridge Regression and Gradient Descent for Ridge Regression produce significantly lower error in the general case than OLE.