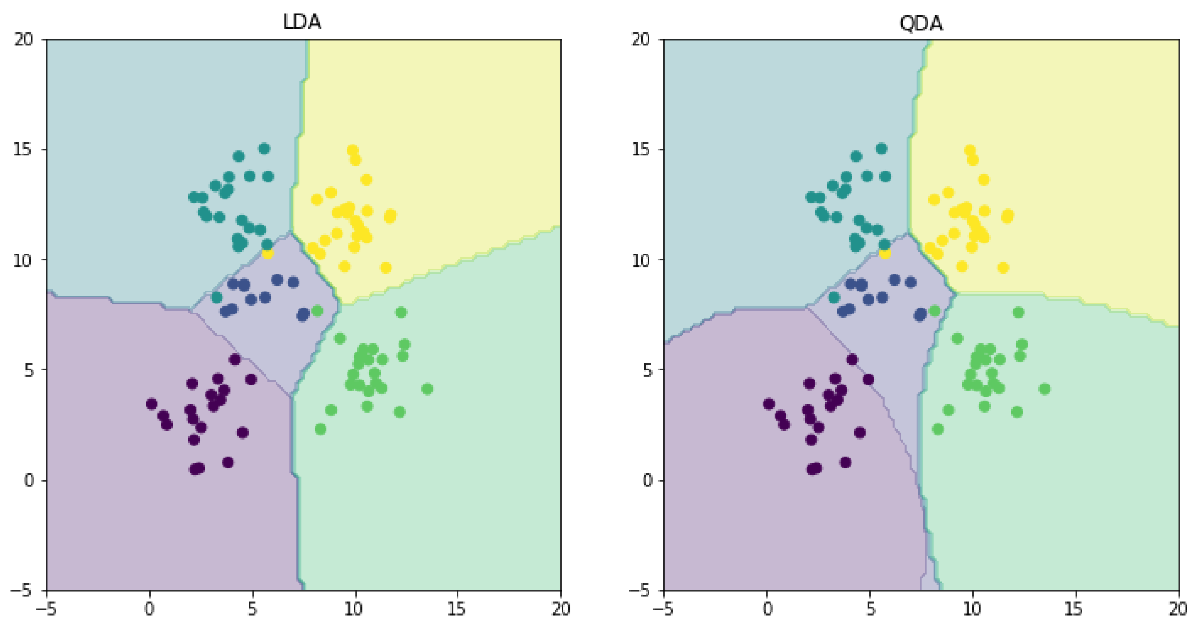
CSE574 Introduction to Machine Learning
Programming Assignment 1
Classification and Regression

Name: Sai krishna kanneti
UBIT Name: skanneti

Name: Yashas JS Shamraju
UBIT Name: yashasjs

Report 1:

The decision boundary for LDA is Linear and decision boundary for QDA is Nonlinear because of the difference in mathematical representation of Discriminant Analysis techniques.



The QDA fits the data well because the misclassification rate is less for QDA as compared to LDA.

Report 2:

The mse for training and test data for two cases : one without intercept and one with intercept are given below.

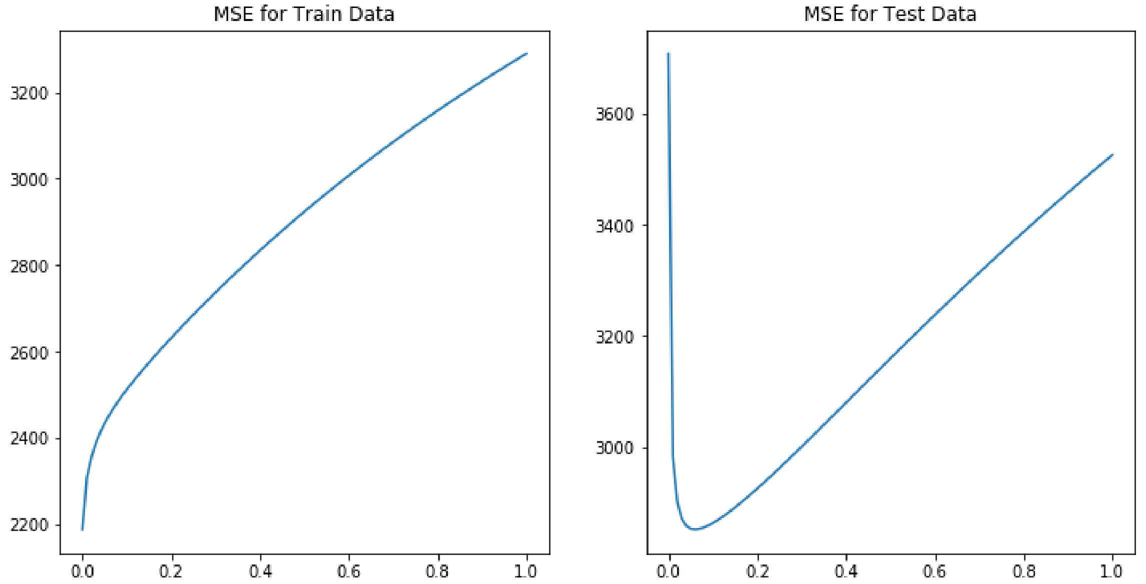
MSE without intercept: 106775.361429
MSE with intercept: 3707.84018164

The mean squared error with intercept is better

Report: 3

The mean squared error for training data increases gradually as lambda value increases.

The mean squared error for test data decreases initially for lambda changing from 0 to 0.06 and the error value increases linearly for changing lambda value from 0.06 to 1.0.



The optimal value of lambda for which the error is minimum: 0.06.

The optimal lambda is: 0.06 because the test error is at the lowest and the training error is also low as compared to other lambda values.

In order to compare the test and training errors using OLE and Ridge regression, the OLE wit hout intercept has a training error of:106775. The training error for OLE with intercept is only considered for comparison.
The training error using Ridge regression for all lambda [0:1] is smaller than the error generat ed using the OLE. The test error using Ridge regression is also less than the error generated u sing the OLE for any lambda ranging from [0:1].

The mean squared error for training data using ridge regression is less as compared to mse for training data using OLE with intercept.

```
OLS Weights,Ridge Weights
(148.15487599577682, 150.19472397176366)
(1.274852097936673, 21.70175431644263)
(-293.3835223638307, -39.06015175437952)
(414.7254484111618, 189.76879579631392)
(272.08913436140665, 131.4363176855414)
(-86639.45710110664, 12.871667628403861)
(75914.46800303459, -12.64706671183519)
(32341.62281191349, -111.67740784965426)
(221.1012148067457, 99.43808809617671)
(29299.551186800003, 203.3144709679367)
(125.23036027719354, 112.56860288788033)
(94.41108334060846, 27.873071479649692)
(-93.86286324125831, 55.60808949902476)
(-33.728279988163195, 35.94955734779689)
(3353.1977149397135, 9.320488702638784)
(-621.0963050425053, -14.822271894177039)
(791.7365331239998, 2.083102994634764)
(1767.7603889631573, 26.192734905566727)
(4191.674055218697, -11.762473145028858)
(119.4381209365747, 33.19690741592031)
(76.61034003713758, 41.121909680934344)
(-15.200129300381377, -1.6292317088961603)
(82.24245936752413, 37.81703093790463)
(-1456.6620842922712, -26.61591914225288)
(827.3867028136738, -46.276109237421146)
(869.290952321142, 0.3382075474677606)
(586.2344951646519, 0.7128583977422451)
(427.02672668453306, 30.42616757192365)
(90.24676900660415, 23.581207173806675)
(-17.88762240516371, 9.146373000958487)
(141.69677382925875, 24.92991385654428)
(582.8193843052723, 19.75759867147805)
(-234.0375106096617, -0.17672982987872388)
(-256.0714522051858, 33.25175428654336)
(-385.1774005240295, 4.638032286264842)
(-33.41767372499453, 18.795892478966778)
(-10.735006606297247, -12.039934460366016)
(257.10718885164533, 38.20746890409774)
(59.95545918145217, -21.98664620368227)
(383.72804236452794, -13.23383199985664)
(-404.1583897320379, -21.406483439723267)
(-514.2864343624933, 9.528354748283363)
(38.36366419686237, -3.041096104038836)
(-44.61028890890884, 45.032658564385244)
(-729.6435309264343, -25.145735497322057)
(377.4083367405692, -31.314978437268664)
(439.79429032979533, 8.845983390976446)
(308.51437332414207, -11.920017822675785)
(189.85967871081084, 4.234694788755782)
(-109.77379701776226, 7.0763703200074595)
(-1919.656979471445, -3.990913457251441)
(-1924.6337759569287, 36.57518819565828)
(-3489.7952769305557, -22.716173017082447)
(11796.9687333107, -1.9058973479557988)
```

```
(530.6744147985592, 8.069498614001194)
(543.3059032186866, 25.844578430507966)
(1821.0751798856072, -12.711259075677354)
(-10463.980688720942, -7.17700708703142)
(-516.6276109277969, 4.034333813087301)
(2064.359173813369, -17.361366882305298)
(-4199.413353249431, 25.14496419150153)
(-140.49570524846786, -20.103956284123683)
(374.1570901129162, -12.621459379374944)
(51.47574915125733, 38.086175375388706)
(-46.449273037549574, 22.553800909883247)
```

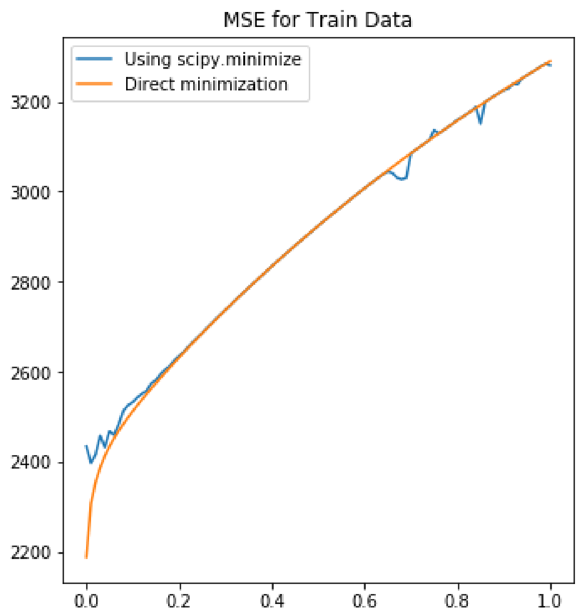
Report 4:

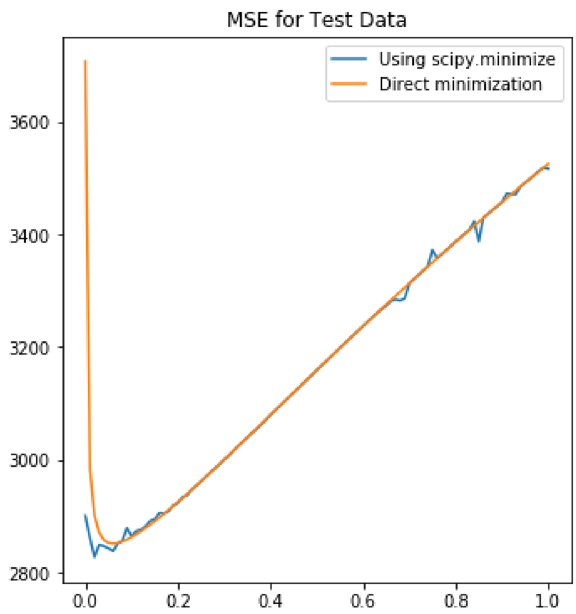
The errors obtained on training data and test data using the gradient descent is almost similar to the results from the direct minimization (the ridge regression). There is very little change in error for training data at lambda value changing from 0 to 0.06.

For all other lambda values, the training error plots obtained using the gradient descent and training error plots obtained using direct minimization is overlapping.

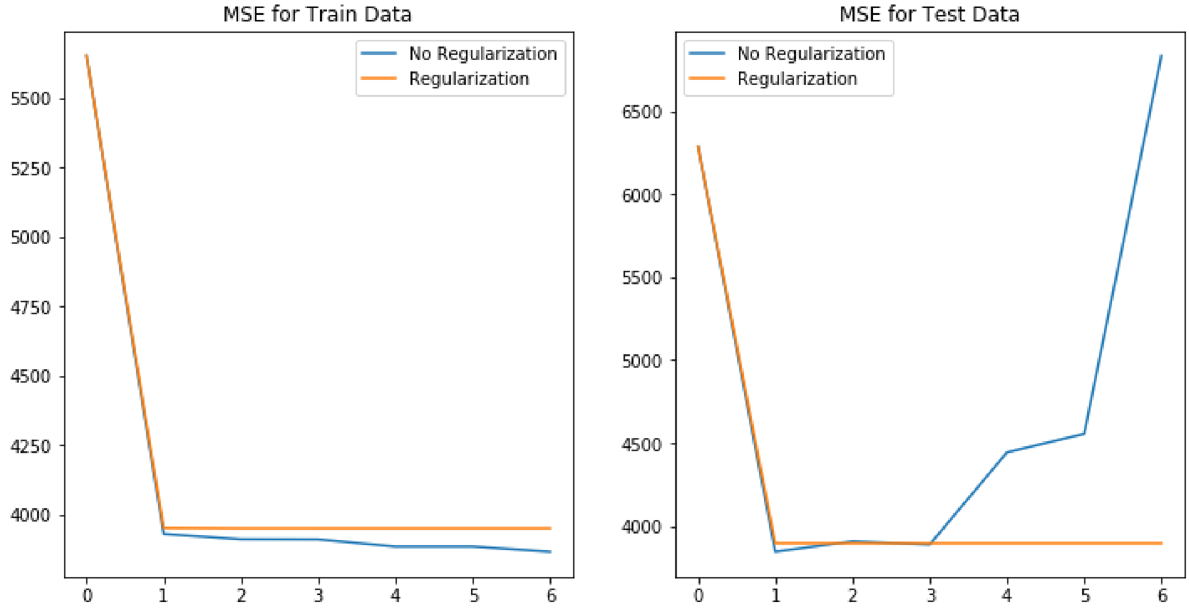
There is very little change in error for test data at lambda value changing from 0 to 0.06.

For all other lambda values, the test error plots obtained using the gradient descent and test error plots obtained using direct minimization is overlapping.





Report 5:



For training data, mse with lambda=0 is less as compared to mse with optimal lambda value for any p value.

We could see that the mse for training data decreases as the p value increases. There is a downside though as the model may face overfitting or won't generalize well for other un seen data.

For test data, the mse is minimum with optimal lambda value.

The optimal value of p in terms of test error is: p=1.

Report 6:

In general, Regularization helps us to provide better results.

Using a gradient descent with a regularization parameter lambda=0.06 gives the minimum error values as compared to other parameters. The root mean squared error or mean squared error is considered a metric to choose the best setting.