CSE 574 – Introduction of Machine Learning
PROGRAMMING ASSIGNMENT - 2

GROUP 12

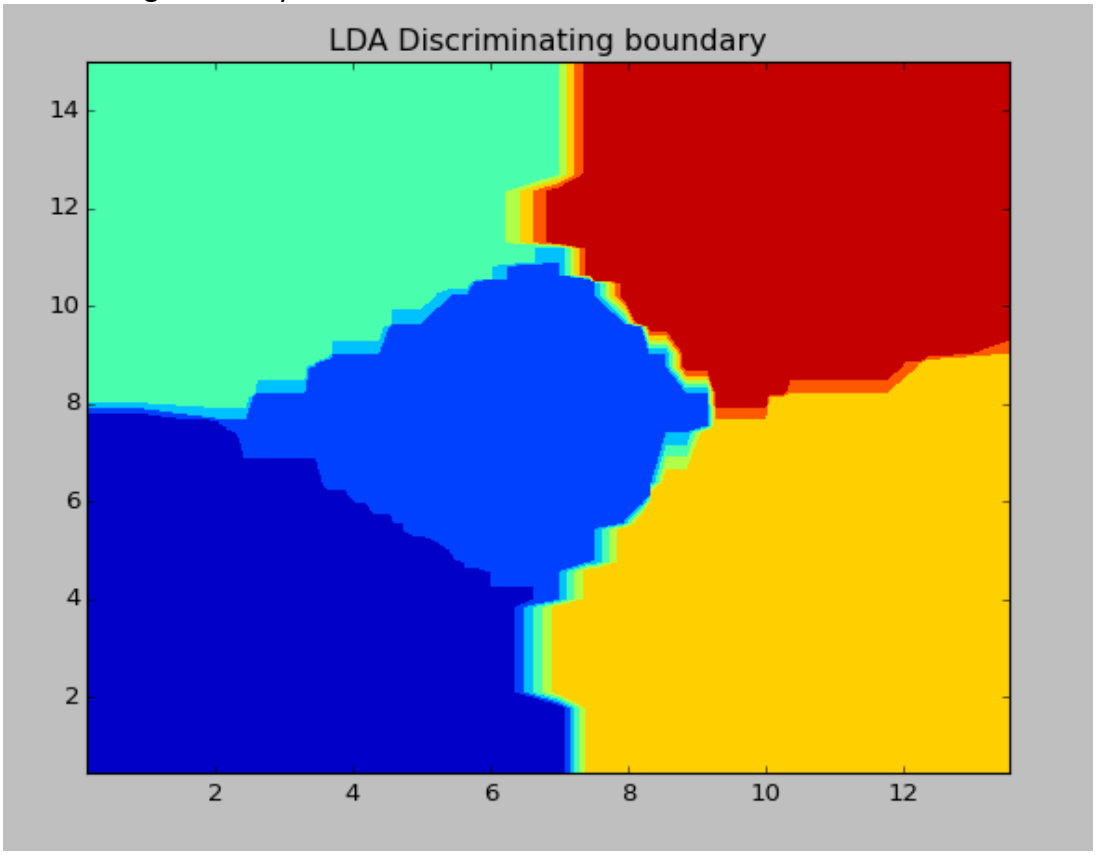
Prasanna Pai – 50132731
Gaurav Kshirsagar – 50134944
Vivin Rane - 50134206

Report 1 – LDA and QDA

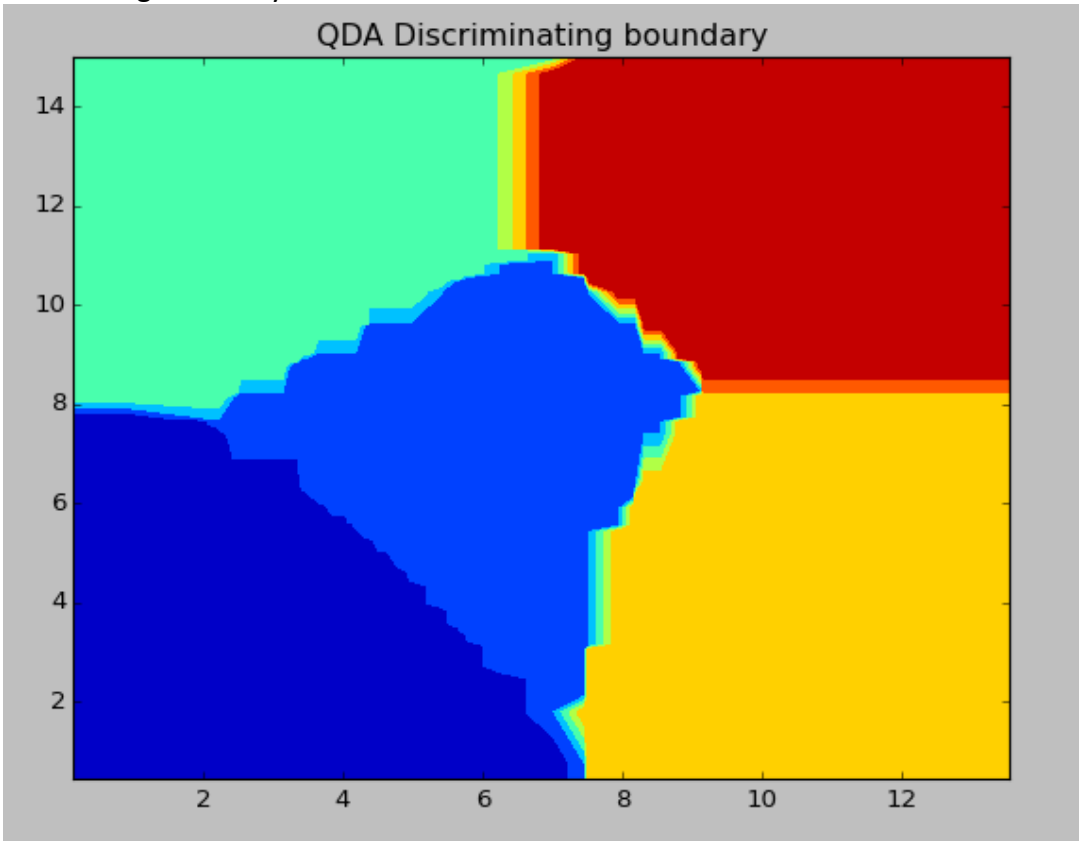
LDA Accuracy = 97.0

QDA Accuracy = 96.0

Plot of discriminating boundary for Linear Discriminators:



Plot of discriminating boundary for Quadratic Discriminators:



Explanation:
Accuracy for the LDA and QDA are comparable and nearing 100% for the given data set. Hence there is minimal difference between the boundaries of the two Gaussian discriminators. Still we can see that the boundaries for LDA are linear in fashion and the boundaries for the QDA are curved.

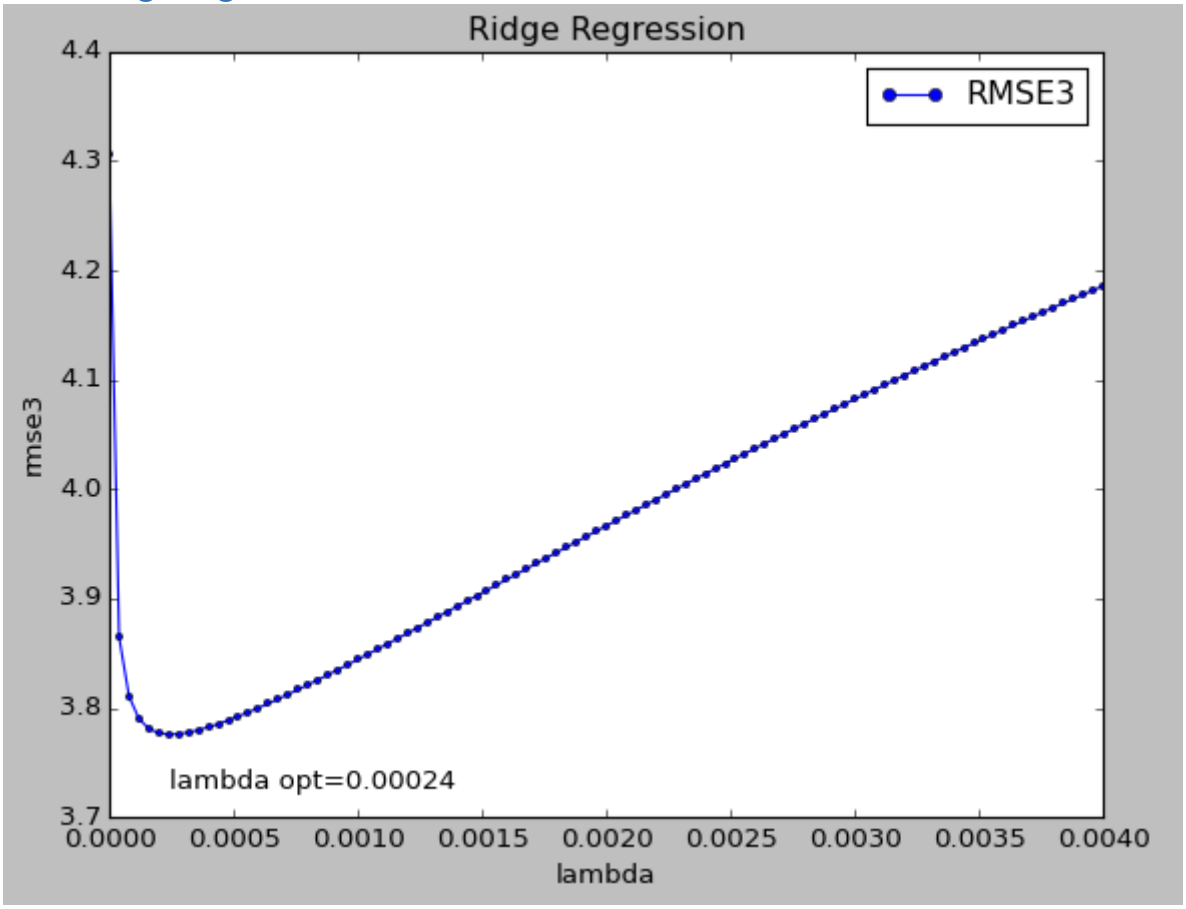
Report 2 – Linear Regression

RMSE without Intercept = 23.1057743384
RMSE with Intercept = 4.3057172351

RMSE without intercept on train data = 8.88388057487
RMSE with intercept on train data = 3.0063021236

RMSE with intercept is less as compared to without intercept. So, the data with intercept is better fit or more fit to this problem to predict the class than the one without intercept. RMSE can be interpreted as a deviation from the variance of this dataset. This RMSE tells us that the predictor is good.

Report 3 – Ridge Regression

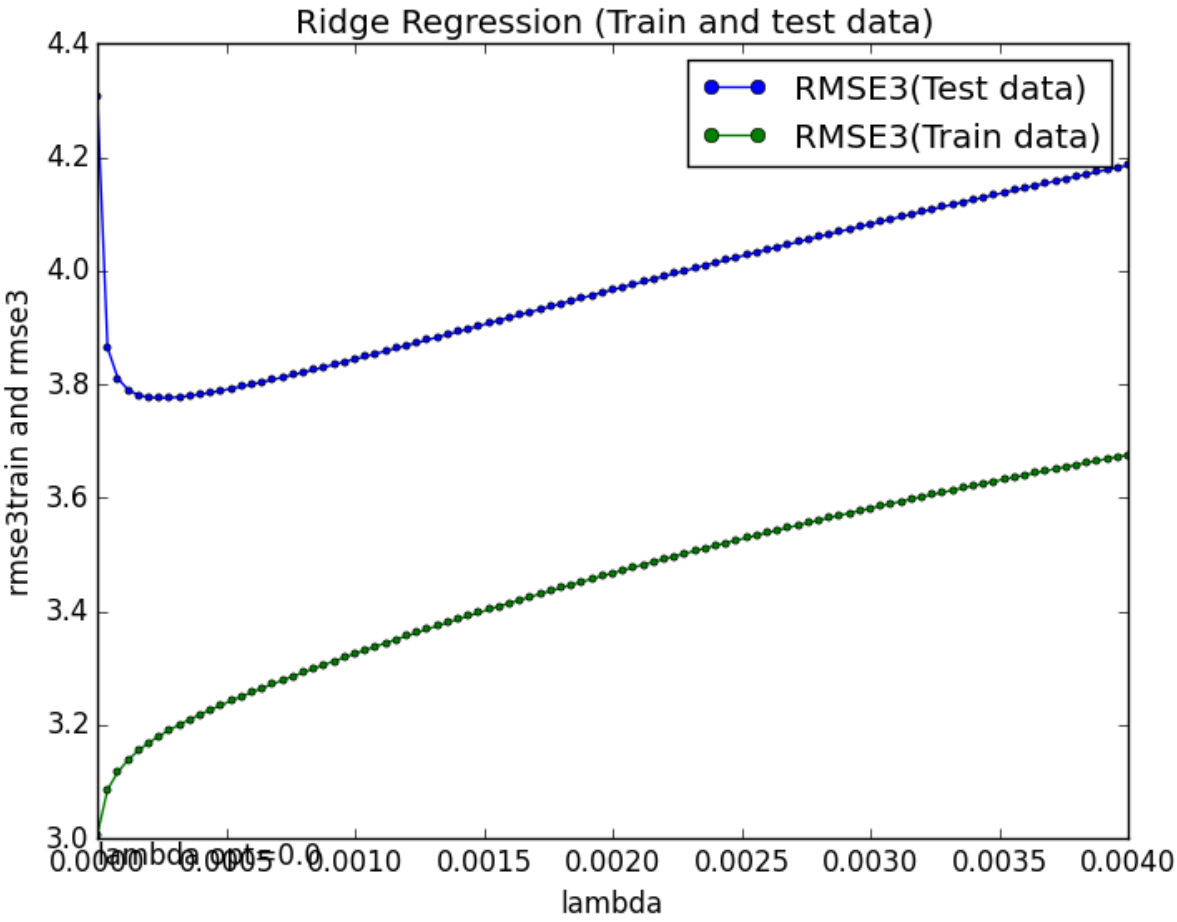


RMSE using Ridge Regression parameters

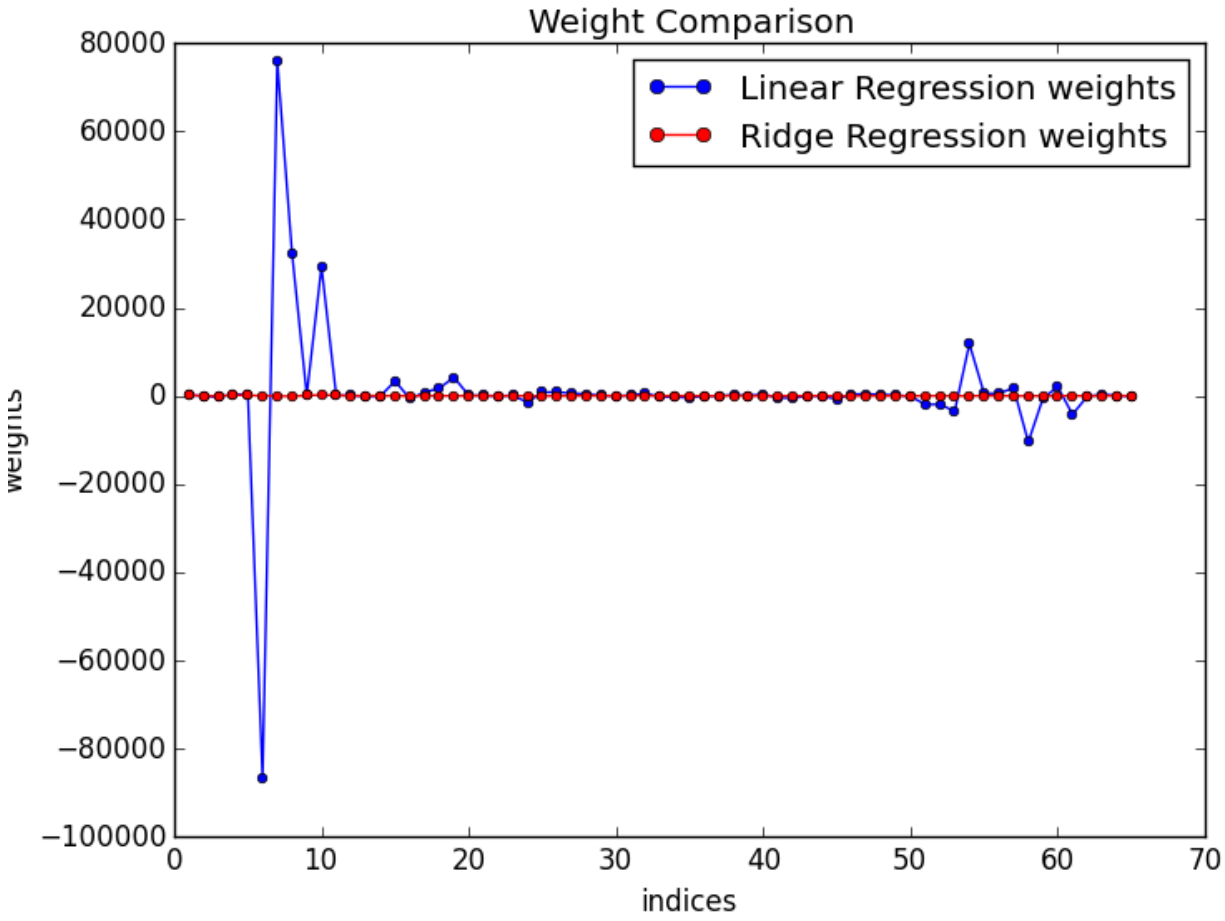
Optimal value of lambda = 0.00024

Least value of RMSE = 3.77582332 at this above value

As seen from the plot, Ridge Regression parameters are better suited to find least RMSE value as compared to Linear Regression.



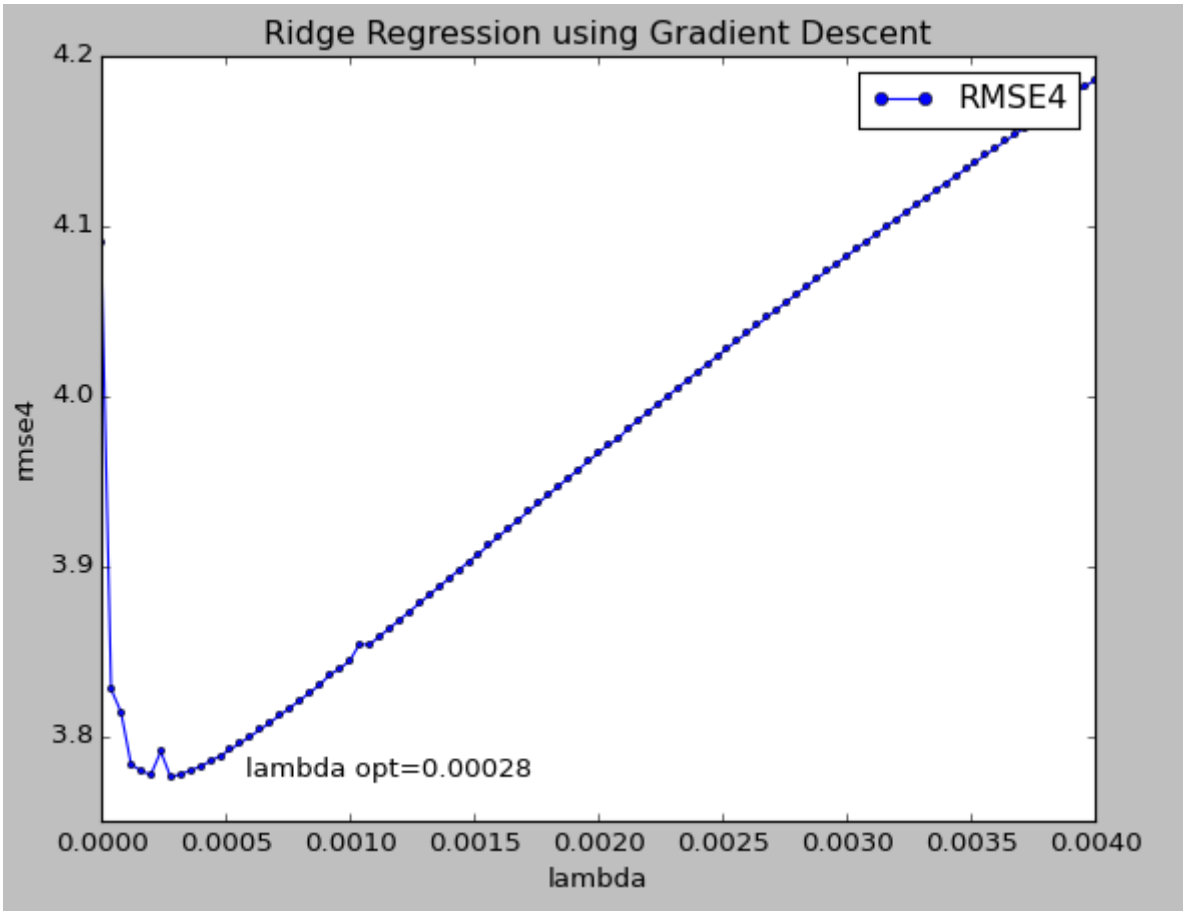
Comparison between test and train data



Comparison between weights learnt in OLE and Ridge Regression:

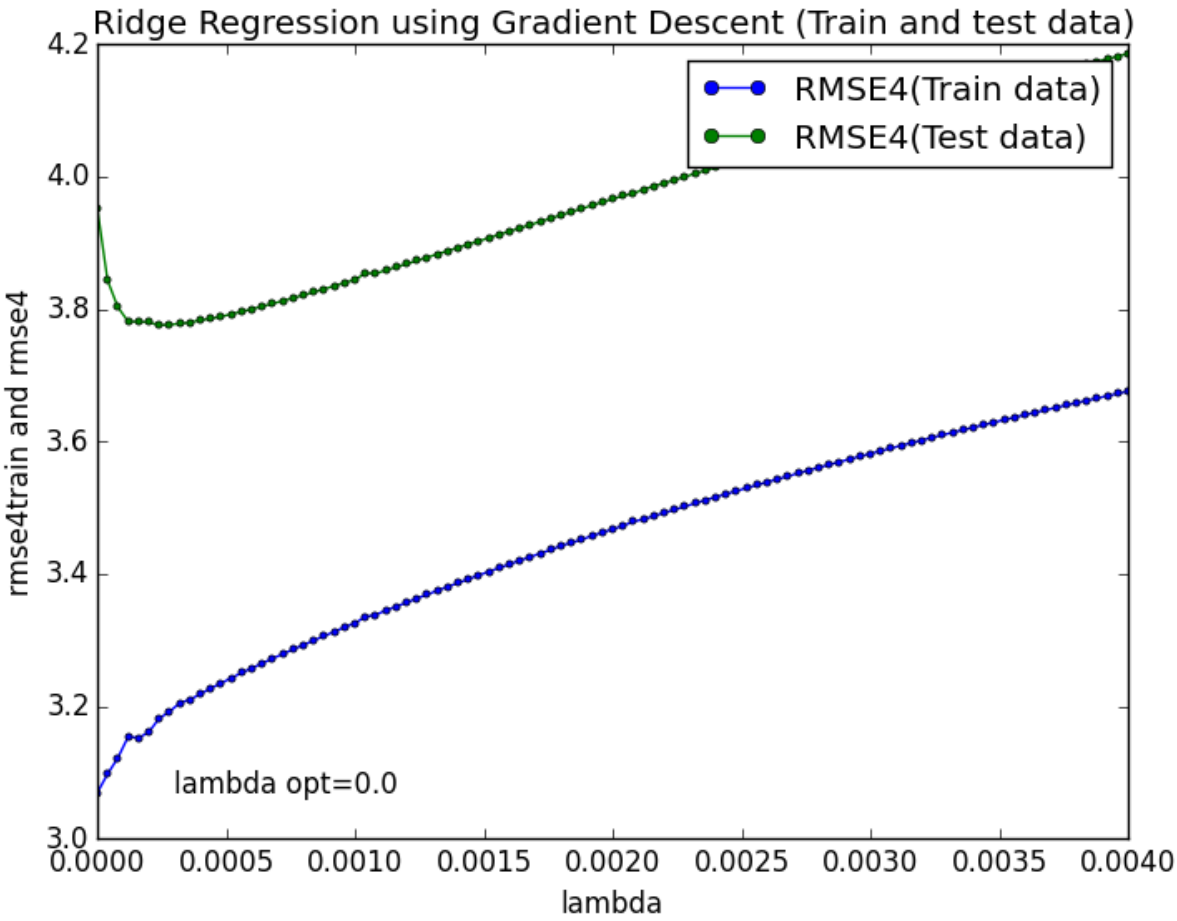
For RMSE matrices, please refer appendix at the end of report.

Report 4 – Ridge Regression with Gradient Descent

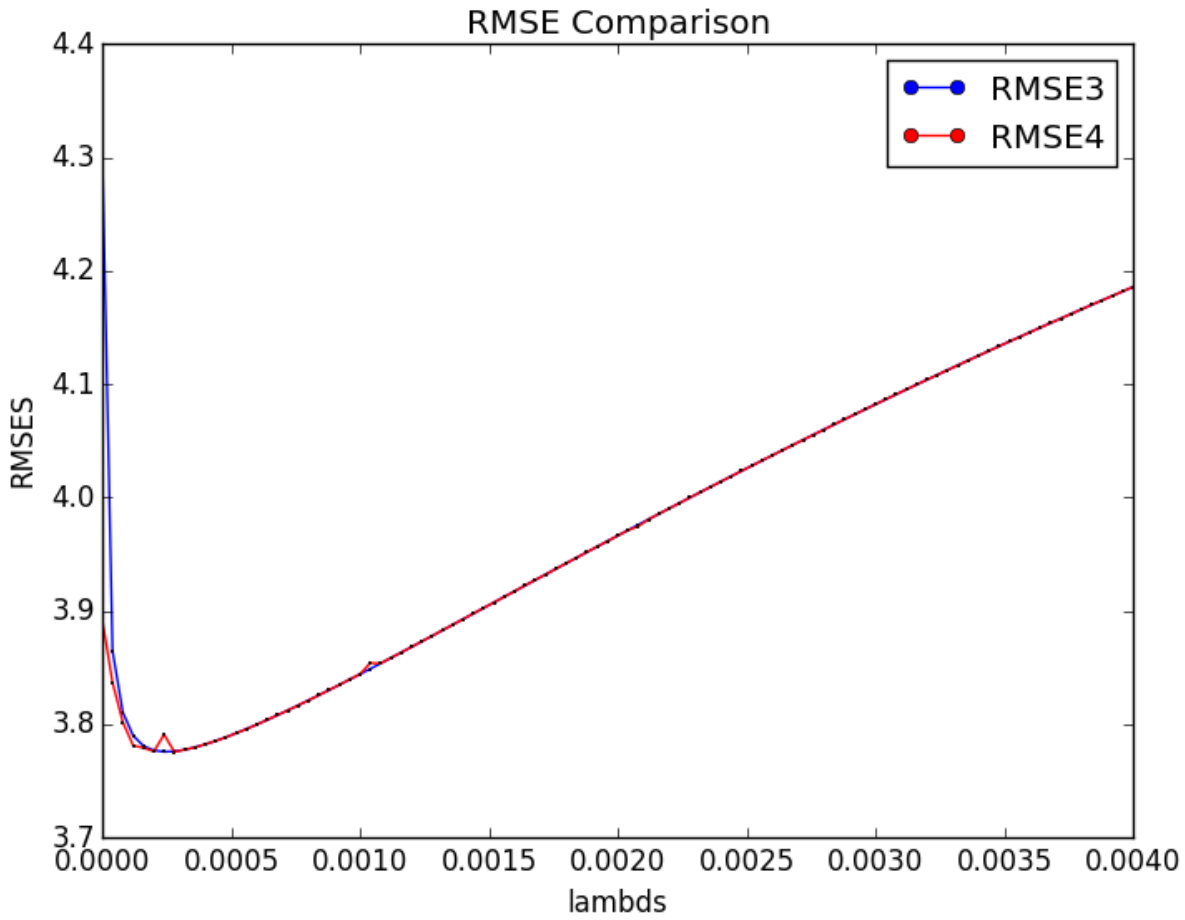


RMSE using Ridge Regression with Gradient Descent

As we can see, the RMSE values are pretty similar to normal Ridge Regression with just a few changes. On observing closely, gradient descent method provides a slightly lower value of RMSE which is 3.77576878 against the value reported by the normal method 3.77582332



Comparison between training data and test data

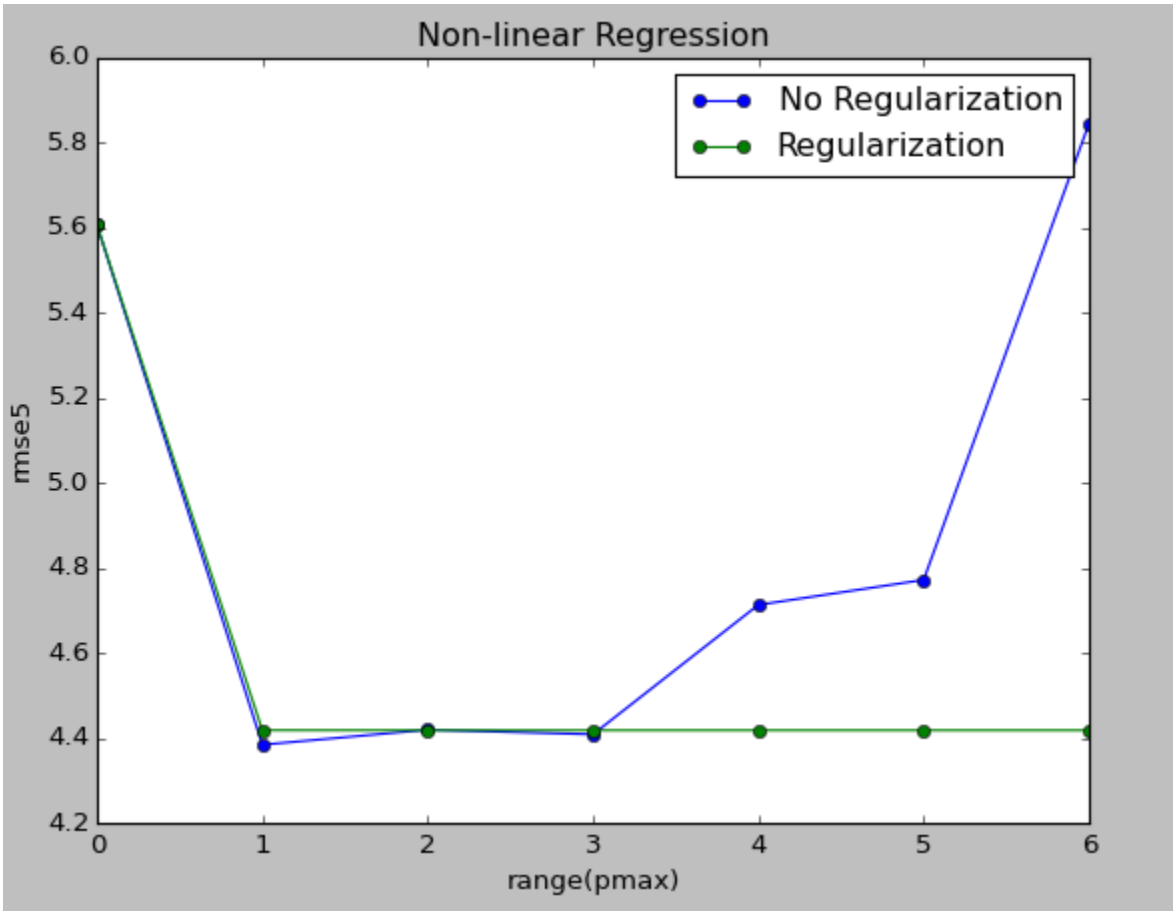


Comparison of the RMSEs from Problem 3 and Problem 4

We can see that starting value for Gradient Descent is lower than in normal Ridge Regression. Normal Ridge Regression has a smoother curve while the other one has a few outliers. Both have almost similar values towards for higher lambda values.

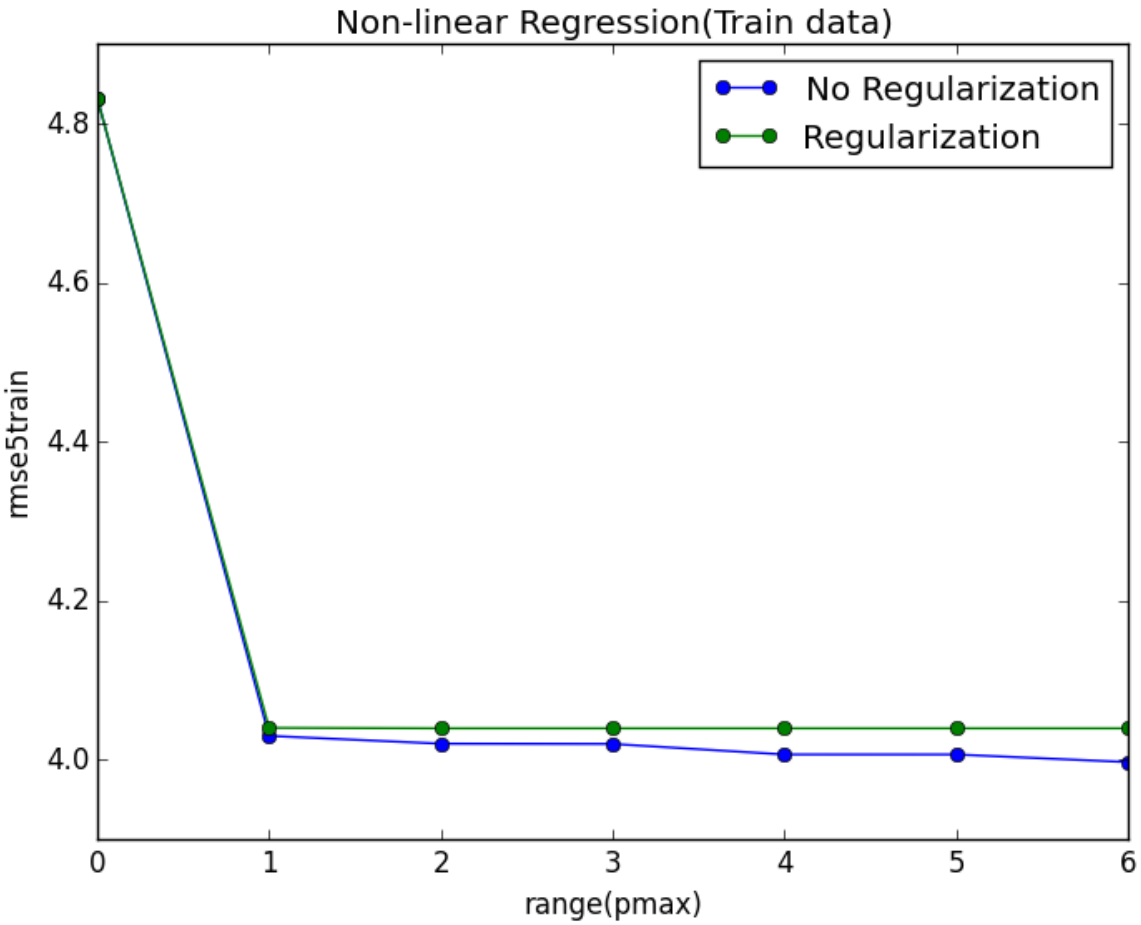
Report 5 – Non-linear Regression

Plot for optimal value of p for both values of lambda



The optimal value of p is 2 where RMSE value both regularization and no regularization have minimum error. The RMSE with regularization remains almost constant for the remaining values of p

Training data plot



Report 6

We compare in this project various regressions namely – Linear regression, Ridge regression, Ridge regression with gradient descent and Non- linear regression. Following are the results -

We use the RMSE value to choose the best setting as we want the error to be least. Although this may not be the most appropriate way to tell, we have no other metric we consider while calculating regression parameters. After comparing the various approaches, we find that Ridge Regression using Gradient Descent is the best setting.

Appendix

rmses3

[

[4.30571724]	[4.00929053]
[3.86462284]	[4.01398299]
[3.81049727]	[4.01865759]
[3.79001317]	[4.02331402]
[3.78095159]	[4.02795203]
[3.77702915]	[4.03257138]
[3.77582332]	[4.03717185]
[3.77622035]	[4.04175324]
[3.77764401]	[4.04631538]
[3.7797691]	[4.05085811]
[3.78239921]	[4.05538129]
[3.78540893]	[4.0598848]
[3.7887143]	[4.06436852]
[3.79225674]	[4.06883235]
[3.79599386]	[4.07327623]
[3.79989393]	[4.07770007]
[3.80393252]	[4.08210383]
[3.80809026]	[4.08648745]
[3.81235149]	[4.09085089]
[3.81670321]	[4.09519414]
[3.82113448]	[4.09951717]
[3.82563592]	[4.10381998]
[3.8301994]	[4.10810256]
[3.83481776]	[4.11236493]
[3.83948466]	[4.11660709]
[3.84419444]	[4.12082907]
[3.848942]	[4.12503089]
[3.85372272]	[4.12921258]
[3.8585324]	[4.1333742]
[3.86336721]	[4.13751576]
[3.86822364]	[4.14163734]
[3.87309846]	[4.14573897]
[3.8779887]	[4.14982071]
[3.88289163]	[4.15388263]
[3.88780472]	[4.15792479]
[3.89272562]	[4.16194726]
[3.89765218]	[4.1659501]

[3.90258239] [3.90751438] [3.91244643] [3.91737693] [3.9223044] [3.92722745] [3.93214481] [3.93705527] [3.94195773] [3.94685116] [3.95173461] [3.95660718] [3.96146806] [3.96631648] [3.97115174] [3.97597317] [3.98078017] [3.98557218] [3.99034868] [3.99510919] [3.99985327] [4.0045805] …contd in next column	[4.16993339] [4.17389721] [4.17784163] [4.18176674] [4.18567262]

]

rmse3train

[

[3.00630212] [3.08616671] [3.11726746] [3.13868758] [3.15518692] [3.16881598] [3.18062699] [3.19121702] [3.20094915] [3.21005578] [3.21869174] [3.2269635] [3.2349462] [3.24269403]	[3.46241379] [3.46749213] [3.47252652] [3.47751754] [3.48246578] [3.48737181] [3.4922362] [3.49705951] [3.50184229] [3.50658508] [3.5112884] [3.5159528] [3.52057878] [3.52516686]

[3.2502468]
[3.25763425]
[3.26487886]
[3.27199788]
[3.27900467]
[3.28590973]
[3.29272134]
[3.29944616]
[3.30608954]
[3.31265587]
[3.31914877]
[3.32557125]
[3.33192585]
[3.33821476]
[3.34443986]
[3.35060277]
[3.35670497]
[3.36274774]
[3.36873229]
[3.37465969]
[3.38053094]
[3.38634698]
[3.39210871]
[3.39781695]
[3.40347252]
[3.40907619]
[3.41462871]
[3.4201308]
[3.42558317]
[3.43098651]
[3.43634149]
[3.44164878]
[3.44690902]
[3.45212285]
[3.4572909]
…contd in next column

[3.52971754]
[3.53423131]
[3.53870867]
[3.5431501]
[3.54755606]
[3.55192702]
[3.55626344]
[3.56056577]
[3.56483446]
[3.56906994]
[3.57327264]
[3.57744299]
[3.58158139]
[3.58568827]
[3.58976402]
[3.59380903]
[3.5978237]
[3.60180841]
[3.60576353]
[3.60968944]
[3.6135865]
[3.61745508]
[3.62129551]
[3.62510815]
[3.62889333]
[3.6326514]
[3.63638269]
[3.64008751]
[3.64376618]
[3.64741903]
[3.65104635]
[3.65464846]
[3.65822564]
[3.66177819]
[3.66530641]
[3.66881057]
[3.67229095]
[3.67574783]

]