**CSE474 (19035)** : Introduction to Machine Learning

**Members:** Jaimeen Ahn, Tyler Craven, & Soo Kyeong Kang (Group 6)

Programming Assignment 2

Classification and Regression

**CONTENT**

1. Report 1: Experiment with Gaussian Discriminators
2. Report 2: Experiment with Linear Regression
3. Report 3: Experiment with Ridge Regression
4. Report 4: Using Gradient Descent for Ridge Regression Learning
5. Report 5: Non-linear Regression
6. Report 6: Interpreting Results

1. Experiment with Gaussian Discriminators

Train both methods using the sample training data (**sample\_train**). Report the accuracy of LDA and QDA on the provided test data set (**sample\_test**). Also, plot the discriminating boundary for linear and quadratic discriminators. The code to plot the boundaries is already provided in the base code. Explain why there is a difference in the two boundaries.

Accuracy of LDA: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Accuracy of QDA: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. Experiment with Linear Regression

Calculate and report the MSE for training and test data for two cases: first, without using an intercept (or bias) term, and second with using an intercept. Which is better?

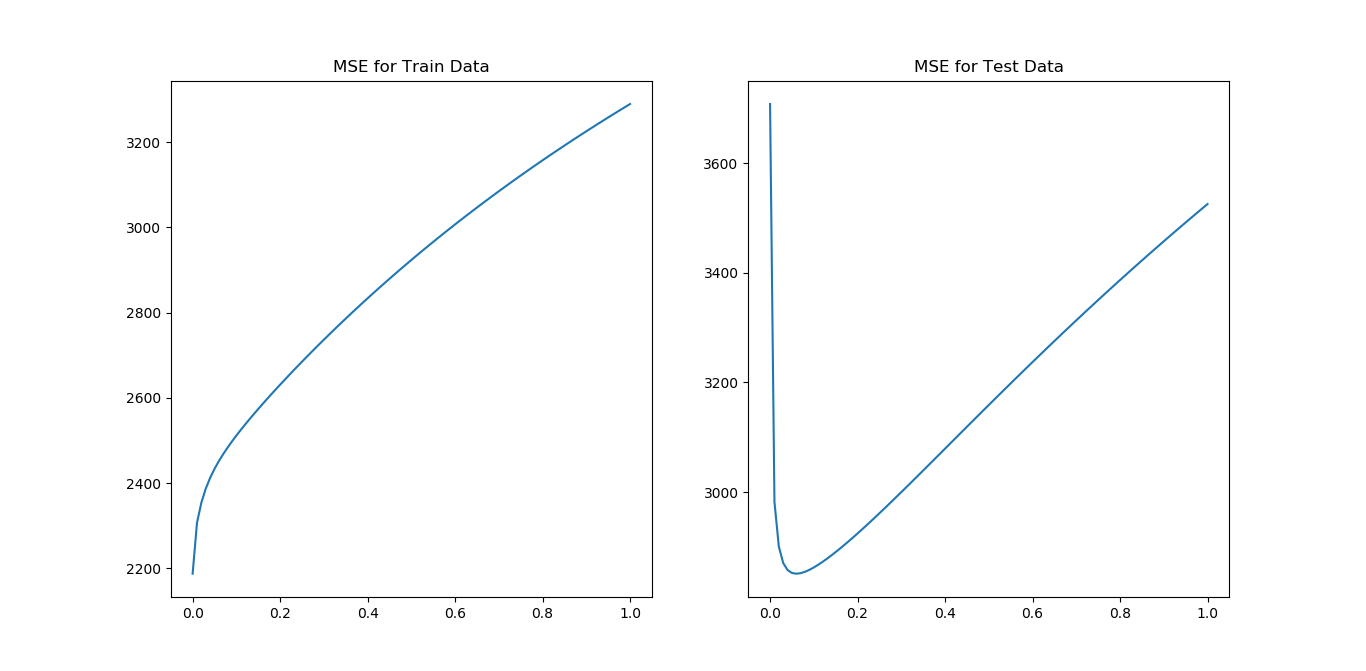
MSE WITHOUT INTERCEPT = 106775.3615112

MSE WITH INTERCEPT = 3707.84018105

Since the MSE WITH INTERCEPT had a lower error, it is the superior regression.

1. Experiment with Ridge Regression

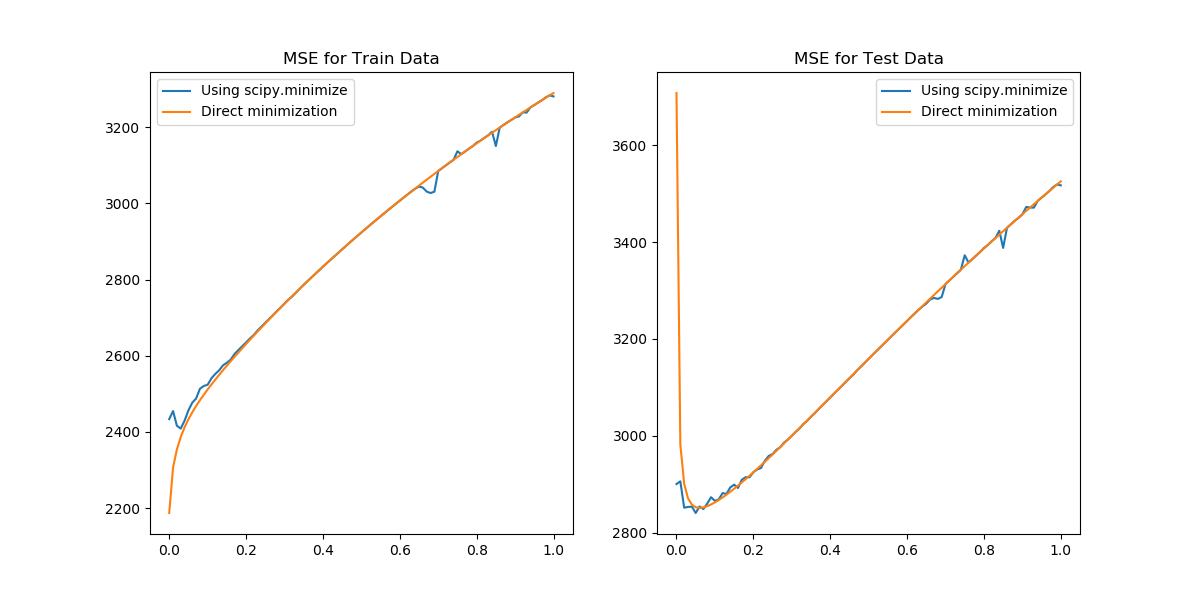
Calculate and report the MSE for training and test data using ridge regression parameters using the testOLERegression function that you implemented in Problem 2. Use data with intercept. Plot the errors on train and test data for different values of λ. Vary λ from 0 (no regularization) to 1 in steps of 0.01. Compare the relative magnitudes of weights learnt using OLE (Problem 2) and weights learnt using ridge regression. Compare the two approaches in terms of errors on train and test data. What is the optimal λ and why?



The optimal λ is approximately 0.05 because it has the lowest MSE on the Test Data.

1. Using Gradient Descent for Ridge Regression Learning

**Plot the errors on train and test data obtained by using the gradient descent based learning by varying the regularization parameter λ. Compare with the results obtained in Problem 3.**



With using the gradient descent based learning, the best regularization parameter λ is still close to 0.05 as seen in problem 3, but may be closer to 0.025.

1. Non-linear Regression

Using the λ = 0 and the optimal value of λ found in Problem 3, train ridge regression weights using the non-linear mapping of the data. Vary *p* from 0 to 6. Note that *p* = 0 means that using a horizontal line as the regression line, *p* = 1 is the same as linear ridge regression. Compute the errors on train and test data. Compare the results for both values of λ. What is the optimal value of *p* in terms of test error in each setting? Plot the curve for the optimal value of *p* for both values of λ and compare.

1. Interpreting Results

Compare the various approaches in terms of training and testing error. What metric should be used to choose the best setting?