CSE 4/574
Gaussian Discriminant Analysis and Linear Regression

Rutuja Badve (rutujara@buffalo.edu)
Ramachandran Kulothangan (rkulothu@buffalo.edu)
Rudraksh Agarwal 50604938 (rudraksh@buffalo.edu)

Problem 1

Problem 2

$w = (X^T X)^{-1} X^T y$

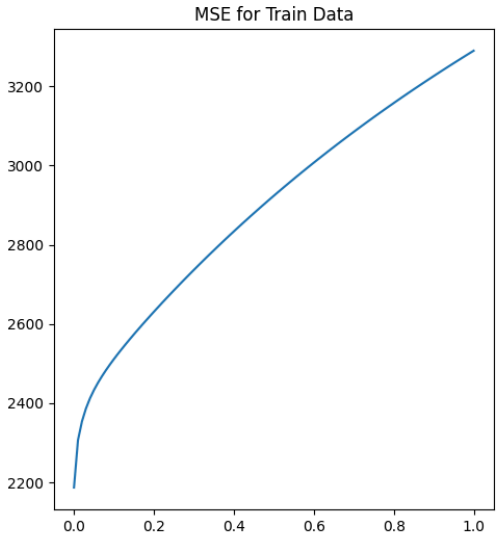
Train data
MSE without intercept: 19099.446844570746
MSE with intercept: 2187.1602949303892

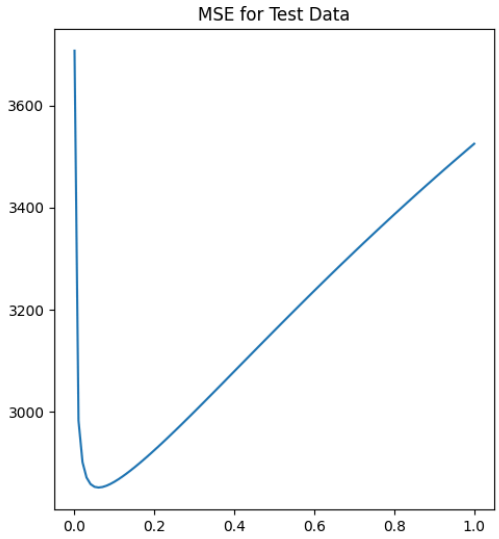
Test data
MSE without intercept: 106775.36153972965
MSE with intercept: 3707.8401811277313

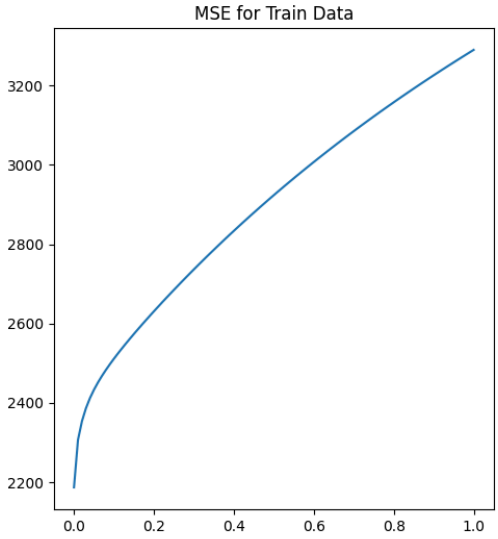
In both the datasets (train and test), MSE with intercept is significantly less than MSE without intercept, and thus MSE with intercept is better.

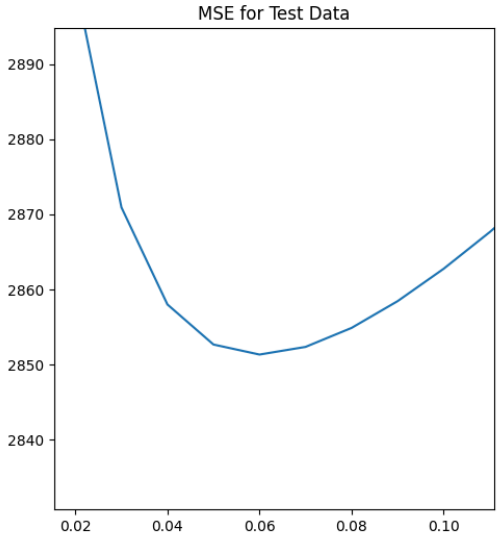
Problem 3

$w = (\lambda I + X^T X)^{-1} X^T y$



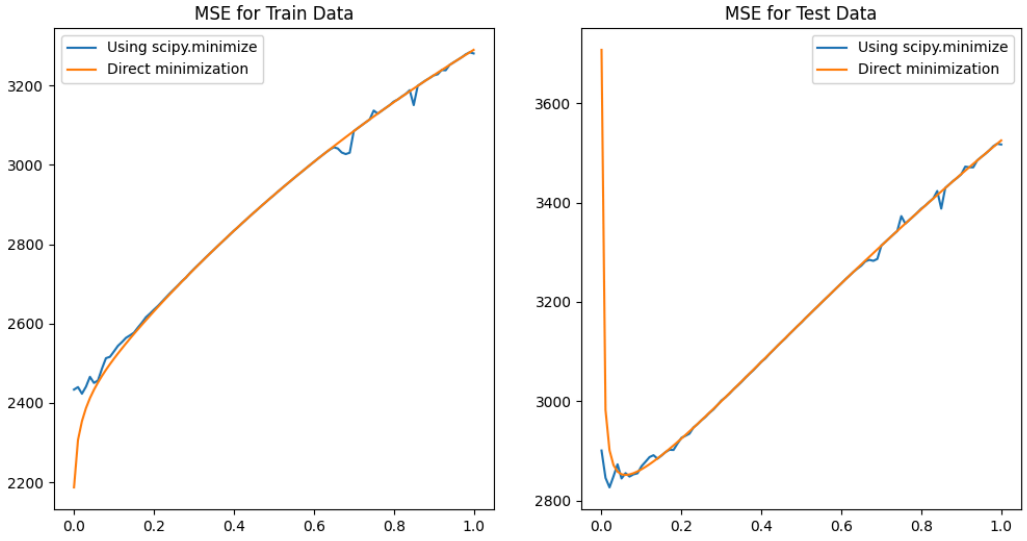




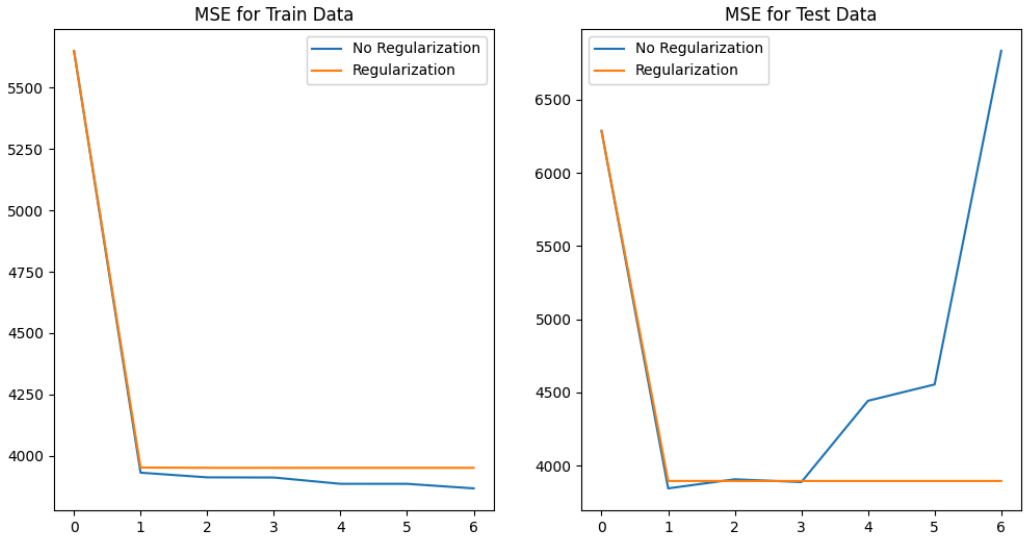


lambda_opt = 0.06 (at this lambda, MSE for test data is lowest (~2851)

Problem 4



Problem 5



For $\lambda = 0$, the optimum value of p = 1
For $\lambda = 0.06$, the optimum value of p = 1

Problem 6