**Regularization**

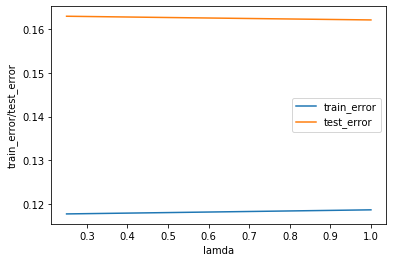
Perform the following regularizations on the curves for which you obtain the minimum and maximum training error from above.

1. Perform Lasso regression on the cost function as follows, vary λ = 0.25, 0.5, 0.75, 1

N=2

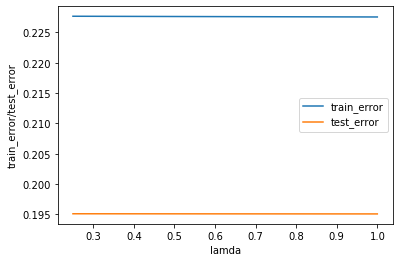


N=9

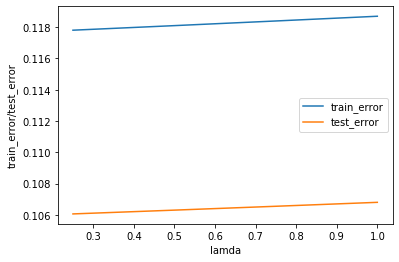


1. Perform Ridge regression on the cost function as follows, vary λ = 0.25, 0.5, 0.75, 1 :

N=2



N=9



1. What differences do you notice between the two kinds of regression? Which one would you prefer for this problem and why?

From the above graphs,

1. We see that both the regularizations perform almost the same on small number of features but as the number of features increases.
2. In our case, Lasso performs better because our aim through regularization was to reduce overfitting (if any) through penalizing the parameter values of the features and to ensure that training and test error approach towards each other.
3. Since the dataset is also very small, we may also infer that Lasso regularization performs better on small datasets because Ridge Regularization is also very powerful but for larger datasets.

For this problem, I will prefer Lasso Regularization because of small number of training examples and the graph for n=9 also shows that the training and test error approaches towards each other as the regularization parameter increases from 0.25 to 1.00.