

# CSL603 - Machine Learning

## Lab 2

Aditya Gupta  
2015CSB1003

September 9, 2017

### Linear Ridge Regression

Given  $X$  and  $Y$  we will find  $W$  that minimizes  $J(W)$ , the error function and are defined as:

$$f(X) = \underbrace{\begin{pmatrix} 1 & x_{11} & \cdots & x_{1D} \\ 1 & x_{21} & \cdots & x_{2D} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{N1} & \cdots & x_{ND} \end{pmatrix}}_X \underbrace{\begin{pmatrix} w_0 \\ w_1 \\ \vdots \\ w_D \end{pmatrix}}_W = \underbrace{\begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_N \end{pmatrix}}_Y$$

$$\min_W J(W) \equiv \min_W \frac{1}{2} (XW - Y)^T (XW - Y) + \lambda \|W\|^2$$

Which when solved gives us:

$$W = (X^T X + \lambda I)^{-1} X^T Y$$

### Observations

The following observations were obtained:

- A particular value of  $\lambda$  (say 0) was chosen and then the magnitude of entries in the weights  $W$  was compared and one by one the least significant ones were discarded and the mean squared error changed in the following way:
- The effect of  $\lambda$  on error was observed for different partitions of the data into training and testing sets. The average mean squared error for 100 repetitions for splitting-fractions varying from 1% to 99% and lambda values from 0 to 100 was observed. The surface corresponding to the average mean absolute error observed was:

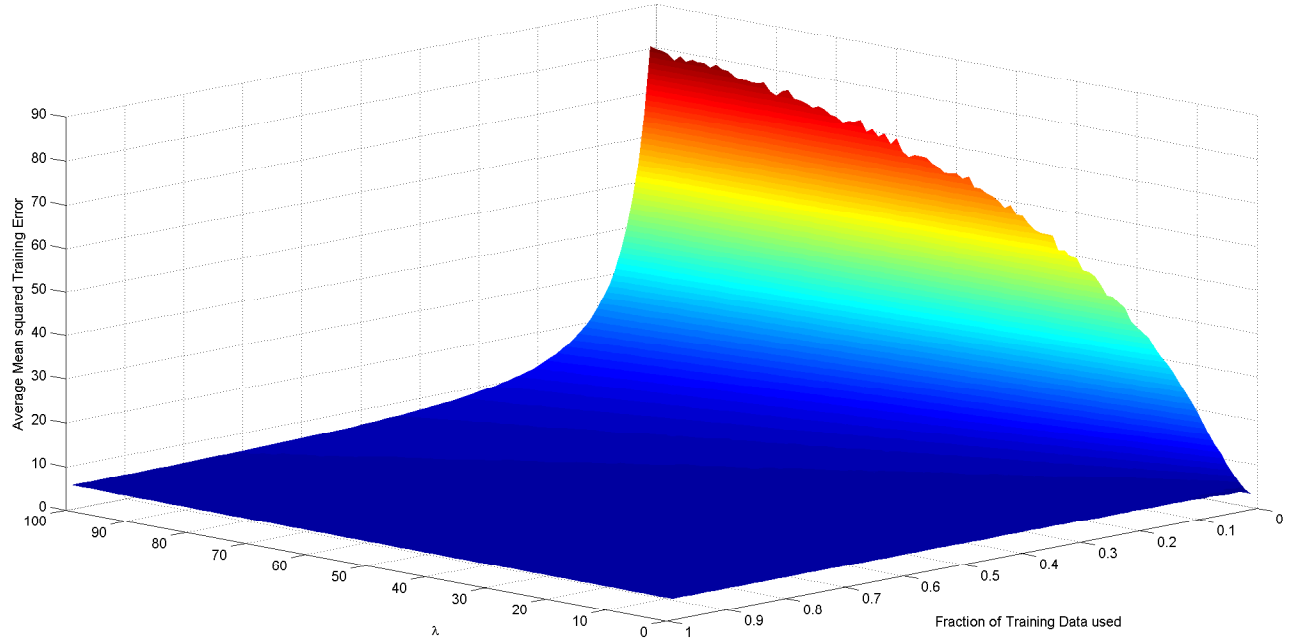


Figure 1: Average Mean Squared Error for various values of training set fraction and  $\lambda$  values used in Ridge Regression for Training Data.

- The above surface can be plotted into different graphs for few particular values of splitting-fractions and varying lambda and observing the change in average mean absolute error. These figures were observed:
- Now we noted the minimum average mean squared testing error for each training set fraction values. Also the corresponding  $\lambda$  value was observed. These figures were observed:
- The correspondence between the actual and predicted values was also observed. For perfect prediction, this should correspond to a straight line through origin at  $45^\circ$  degrees. The following figures were obtained:

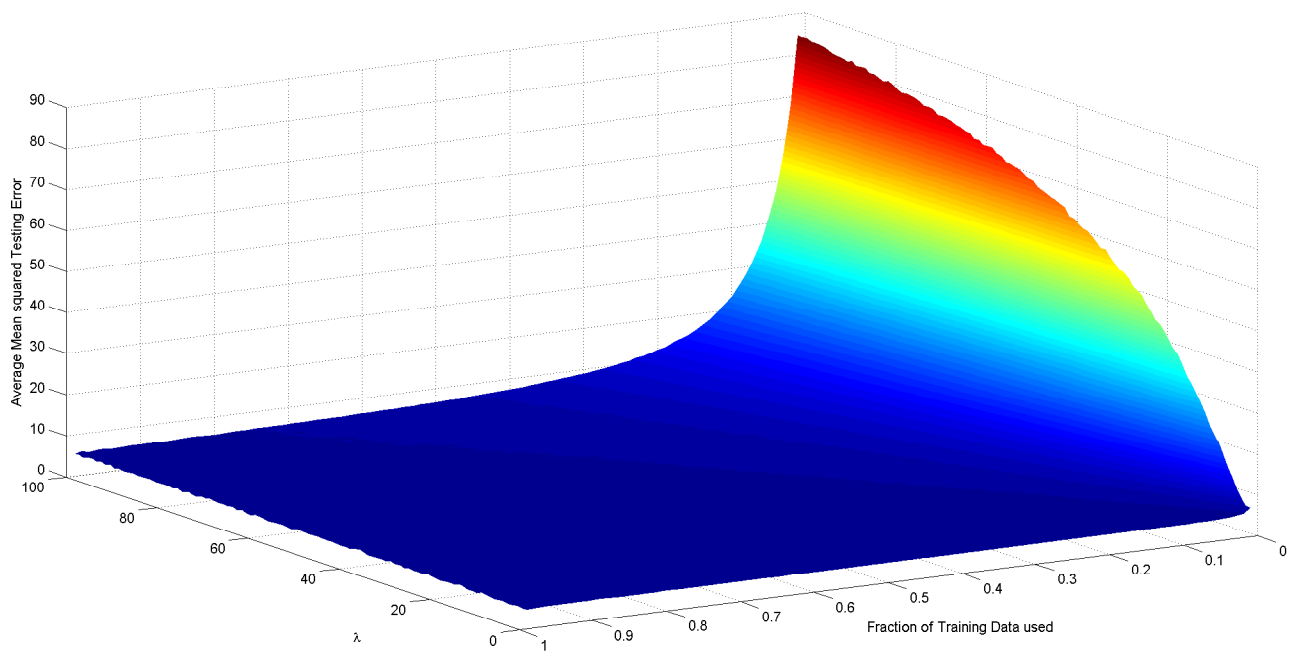


Figure 2: Average Mean Squared Error for various values of training set fraction and  $\lambda$  values used in Ridge Regression for Testing Data.

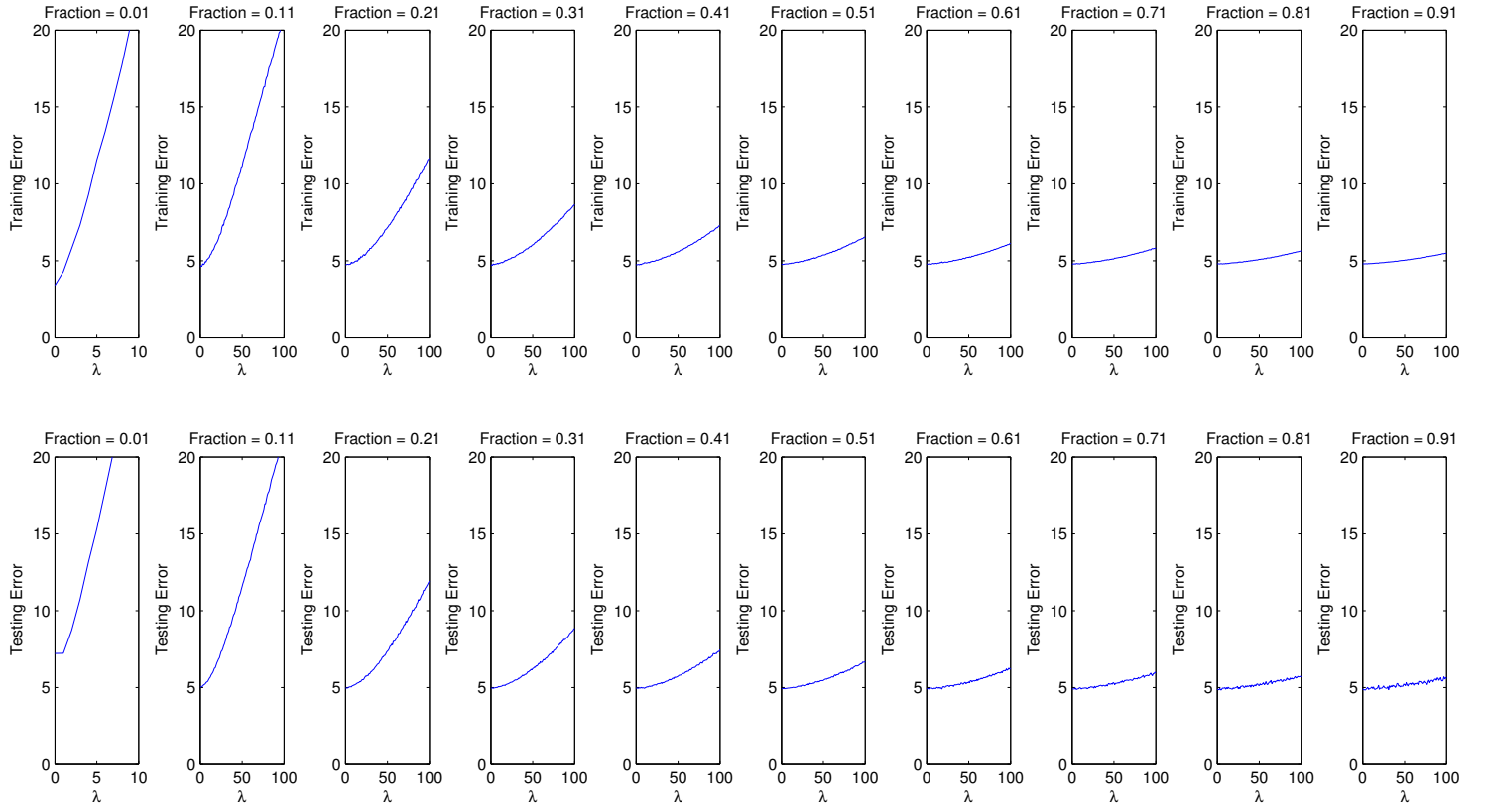


Figure 3: Average Mean Squared Error for various training set fractions varying against the  $\lambda$  values

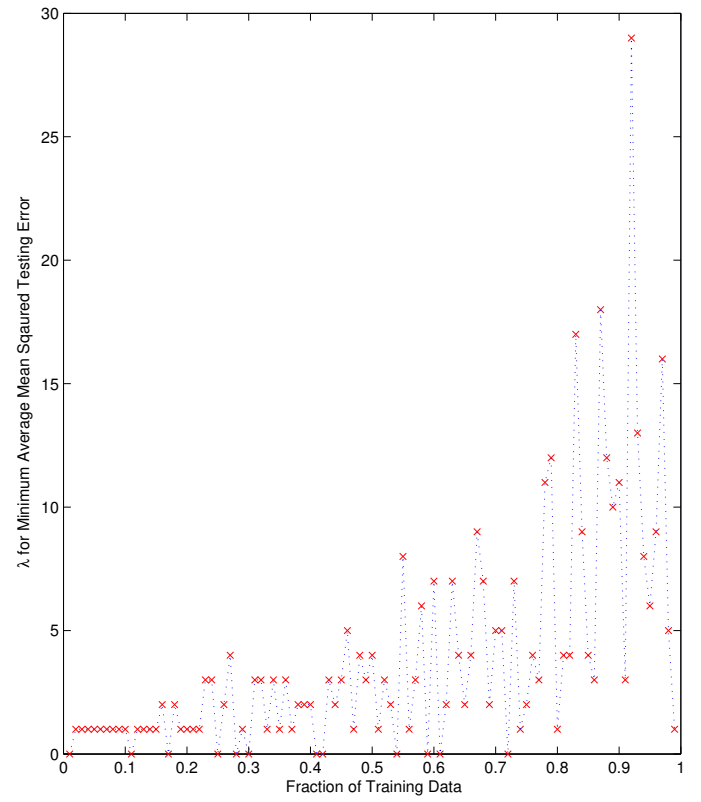
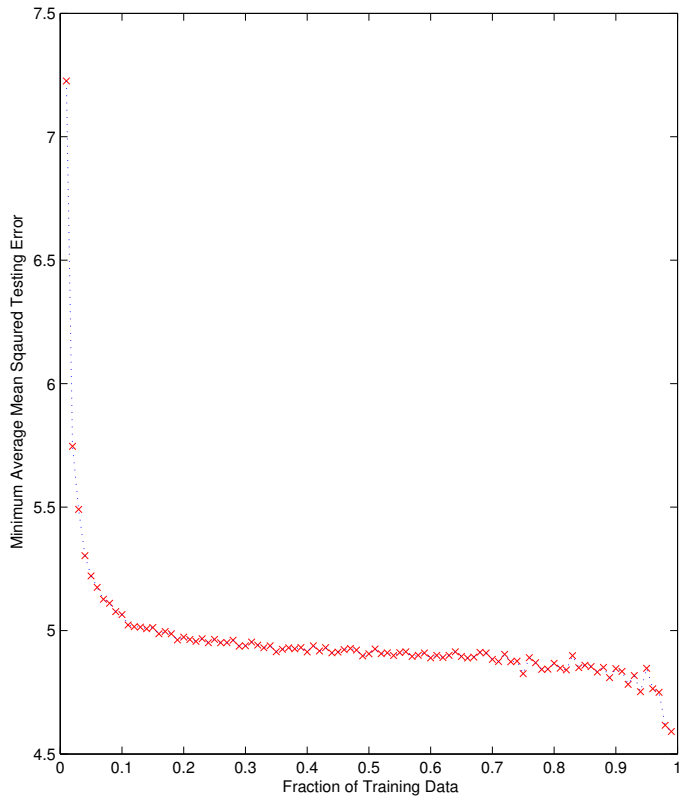


Figure 4: Minimum Average Mean Squared Error for various values of training set fraction and the corresponding  $\lambda$  values.

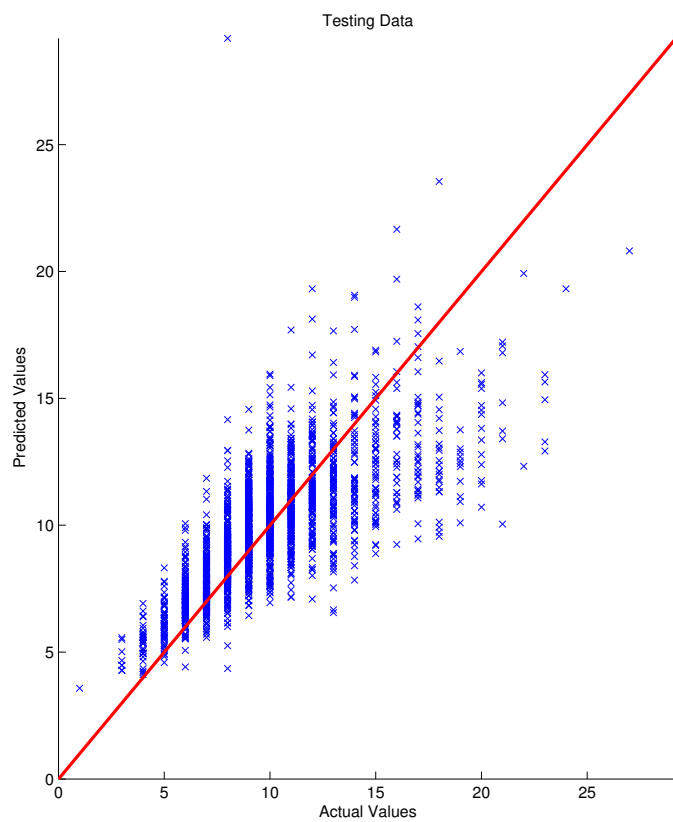
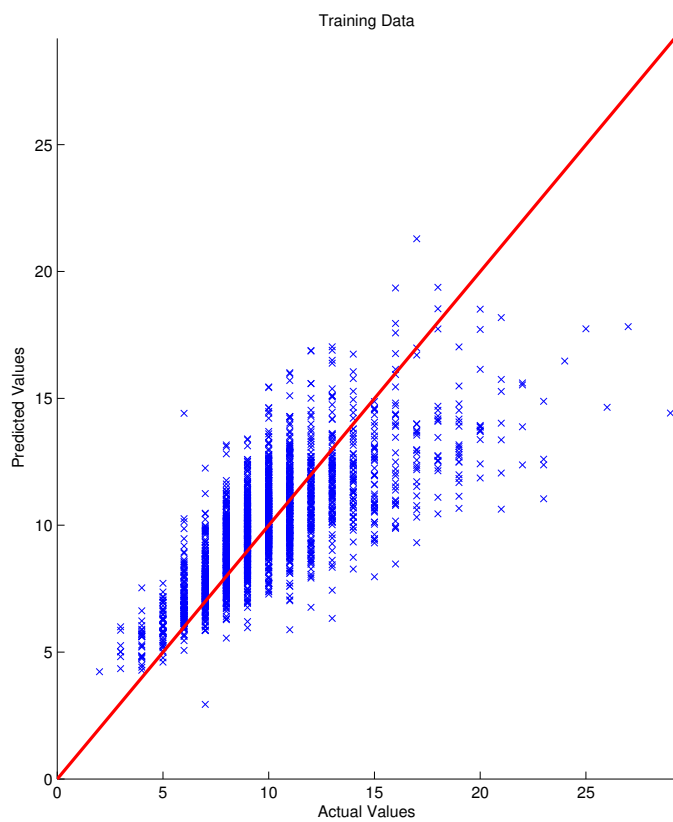


Figure 5: Relation between the actual data set values and predicted values for training and testing data. The reference line  $y = x$  is shown in red.