Assignment 4

1. Generate 50 real numbers for the variable X from the uniform distribution U [0,1]
2. Construct the training set T = { $(x_1,y_1),(x_2,y_2),\ldots\ldots,(x_{50},y_{50})$} using the relation

   $Yi = \sin(2\pi(x_i)) + \epsilon_i$ where $\epsilon_i \sim N(0,0.25)$ and x_i is from **R**.

3. In the similar way construct a testing set of size 500
   i,e. Test = { $(x'_1,y'_1),(x'_2,y'_2),\ldots\ldots,(x'_{50},y'_{50})$}.

4. Estimate the Regularized Least Square kernel regression model by using the training set T. Plot the training estimate along with the original target function. Also, compute the Test RMSE.

5. Randomly select 5 training points to say $x'_1, x'_2, x'_3, x'_4, x'_5$ and scale(multiply) their corresponding $y'_1, y'_2, y'_3, y'_4, y'_5$ by 4. So these 5 points are to be considered as outliers. Estimate the Regularized Least Square kernel regression model by using the modified training set T. Plot the training estimate along with the original target function.
   Also, compute the RMSE.

6. Write your observation regarding the performance of the Least Square Regression model in presence of outliers.

7. Repeat the experiment (1-5) for the L1-norm loss kernel regression model. For obtaining the solution of the L1-norm loss kernel regression model, you need to use the gradient descent or stochastic gradient descent algorithm. Write your observation regarding the performance of L1- norm kernel Regression model in presence of outliers.

Note:- Start working on this assignment and ask any doubts and difficulties regarding this in the lab class.

.