Assignment 1

1. Generate 20 real number for the variable X from the uniform distribution U [0,1]
2. Construct the training set T = { $(x_1,y_1),(x_2,y_2),\ldots\ldots,(x_{20},y_{20})$} using the relation

$$Yi = \sin(2\pi x_i) + \epsilon_i \text{ where } \epsilon_i \sim N(0,0.25)$$

3. In the similar way construct a testing set of size 50
   I,e. Test = { $(x'_1,y'_1),(x'_2,y'_2),\ldots\ldots,(x'_{50},y'_{50})$}
4. Estimate the Least Square polynomial regression model of order M= 1,2, 3, 9, using the training set T.

   For example for M=1 , we need to estimate

   $F(x) = \beta_1 x + \beta_0$

   For M = 2

   $F(x) = \beta_2 x^2 + \beta_1 x + \beta_0$.

5. List the value of coefficients of estimated polynomial regression models for each case.
6. Obtain the prediction on testing set and compute the RMSE for polynomial regression models for order M =1,2,3 and 9 .
7. Plot the estimate obtained by polynomial regression models for order M =1,2,3 and 9 for training set along with $y_1, y_2, \ldots, y_{50}$. Also plot our actual mean estiamte $E(Y/X) = \sin(2\pi x_i)$.
8. Plot the estimate obtained by polynomial regression models for order M =1,2,3 and 9 for testing set along with $y'_1, y'_2, \ldots, y'_{50}$. Also plot the $\sin(2\pi x'_i)$.
9. What happens when we increase the value of M. Note down your observations.
10. Also ,try to find the statistical reasons behind your observation.
11. Enjoy now.

Note:- Do not use any inbuilt functions of MATLAB.

Pritam.