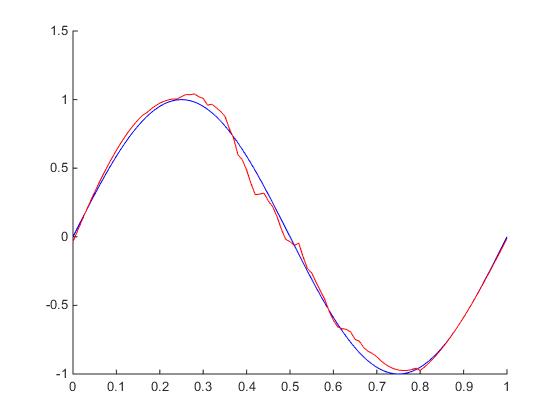
## The Report of 6th Programming Homework

1. One leaf node in a regression tree may indicate a regression function rather than a constant. Implement a regression tree generation algorithm when the leaf nodes indicate 3rd order polynomials. Test your program with the dataset that you had generated in the Programming Exercise 1 in Section 3.2, and compare the results with those of simple linear regression.

In this experiment, the key point is to find the threshold that can split x into two dataset, then use build the regression tree (poly fit ) using these two dataset, and change the threshold to make minimum error. The step are as follow

1. Generate x randomly and calculate corresponding y using
2. Initialize the threshold with 0
3. Split the dataset into two part based on whether x is larger than threshold
4. Use these two dataset to polyfit y using 3rd polynomials
5. Calculate the error
6. Change threshold and repeat step 3-5
7. Find the minimum error and use corresponding regression tree as model

I generate 100 dataset to get 100 regression tree and then generate a test dataset with equal interval and using these 100 regression tree to get predicted value and make an average. And the plot the average value to compared with , the result are as follow:



Summarize: From the picture above we can see the regression tree whose leaf nodes indicate 3rd order polynomials can fit the function very good and its effect is almost the same as ridge regression’s.