Mehmet Mert Bezirgan

Engr 421 – homework 4

**Problem**: Nonparametric Regression

**Solution:**

For nonparametric regression we have a dataset contains 133 pairs of data where we are going to assign first 100 for training and other 33 for testing our predictions. After importing the data, we implement regressogram, running mean smoother and kernel smoother. Formulas of these smoothers are as given.

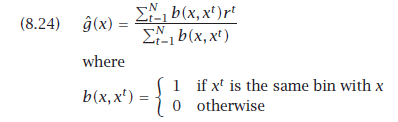


Figure 1 regressogram

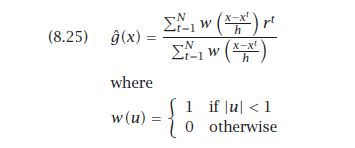


Figure 2 running mean smoother

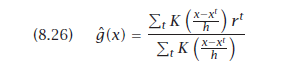


Figure 3 kernel smoother

We are going to calculate root mean squared error for all 3 smoothers and compare results of 3 algorithm.

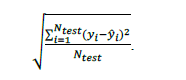


Figure 4 root mean squared error

1-Regressogram

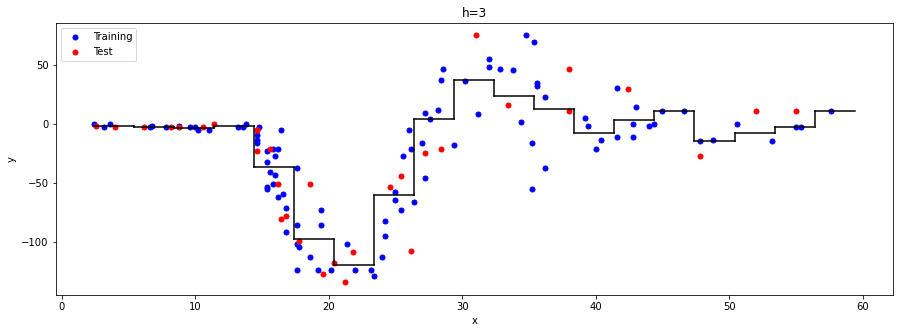
We developed a regressogram smoother for given data and after training with bin width 3 and origin 0 we got the following graph. 

Figure 5 regressogram graph

The black line is the value that we are going to predict if an x value from that interval comes to our algorithm. After drawing graph we calculate RMSE error as 23.3291



Figure 6 rmse output 1

2-Running Mean Smoother

Using equation in figure 2 we developed a running mean smoother and drew graph of it too.

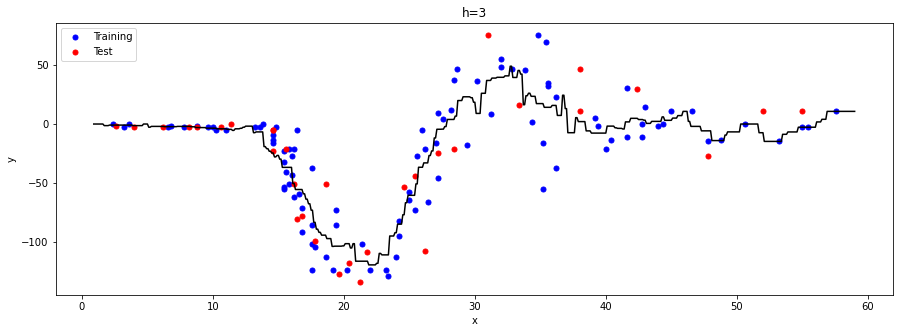


Figure 7 running mean smoother graph

The bin width is still 3. We also calculated RMSE for this smoother too and got the value 23.8403.



Figure 8 rmse output 2

3-Kernel Smoother

In this step we are going to use Kernel function while calculating result as in figure 3. We draw the graph of kernel smoother too.

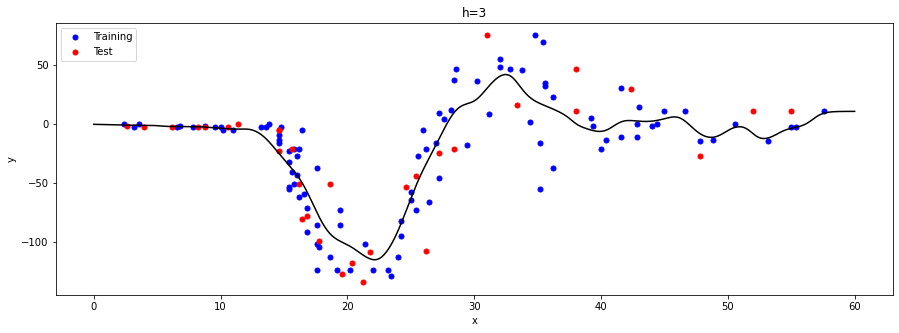


Figure 9 graph of kernel smoother

For error calculation we calculate RMSE for kernel smoother too and got the result 24.1672. For this smoother we changed bin width to 1.



Figure 10 rmse output 3