

COMP 421 – HOMEWORK 04

REPORT

Özgür Taylan Özcan

First, I read X and Y data from the given file and stored the data into separate vectors for training and test sets (x_train, x_test, y_train, y_test).

Regressogram

I implemented the regressogram function as instructed in section 8.8 in the textbook.

$$\hat{g}(x) = \frac{\sum_{t=1}^N b(x, x^t) r^t}{\sum_{t=1}^N b(x, x^t)}$$

where

$$b(x, x^t) = \begin{cases} 1 & \text{if } x^t \text{ is the same bin with } x \\ 0 & \text{otherwise} \end{cases}$$

After this, I formed the regressogram (with h=3) by using the function that I implemented. Then I plotted it along with the training and test data as stated in the homework description.

Finally, I calculated RMSE for my regressogram and printed it.

Running Mean Smoother

I implemented the running mean smoother function as instructed in section 8.8.1 in the textbook.

$$\hat{g}(x) = \frac{\sum_{t=1}^N w\left(\frac{x-x^t}{h}\right) r^t}{\sum_{t=1}^N w\left(\frac{x-x^t}{h}\right)}$$

where

$$w(u) = \begin{cases} 1 & \text{if } |u| < 1 \\ 0 & \text{otherwise} \end{cases}$$

After this, I formed the running mean smooth (with h=3) by using the function that I implemented. Then I plotted it along with the training and test data as stated in the homework description.

Finally, I calculated RMSE for my running mean smooth and printed it.

Kernel Smoother

I implemented the kernel smoother function as instructed in section 8.8.2 in the textbook.

$$\hat{g}(x) = \frac{\sum_i K\left(\frac{x-x^i}{h}\right) r^i}{\sum_i K\left(\frac{x-x^i}{h}\right)}$$

After this, I formed the kernel smooth (with $h=1$) by using the function that I implemented. Then I plotted it along with the training and test data as stated in the homework description.

Finally, I calculated RMSE for my kernel smooth and printed it.

Results

My regressogram is working as expected, without problems. I get the same results with the ones on the homework description.

However, I could not get good results for running mean smooth and kernel smooth. The plots and RMSE values are not 100% correct.