

Homework 5: Construction of a Regression algorithm

Natasha Kubiak | ECE 517 | DUE: 10/23/2021

Intro:

The .mat and the .csv files provided is a dataset that contains a training and a test sets. Each set consists of the predictor input X_{train} and X_{test} for regressors y_{train} and y_{test} , corresponding to the response y of a physical model to a vector signal x of 19 dimensions. The train and test outputs consist of 81 samples and they are depicted in the figure below.

PART 1

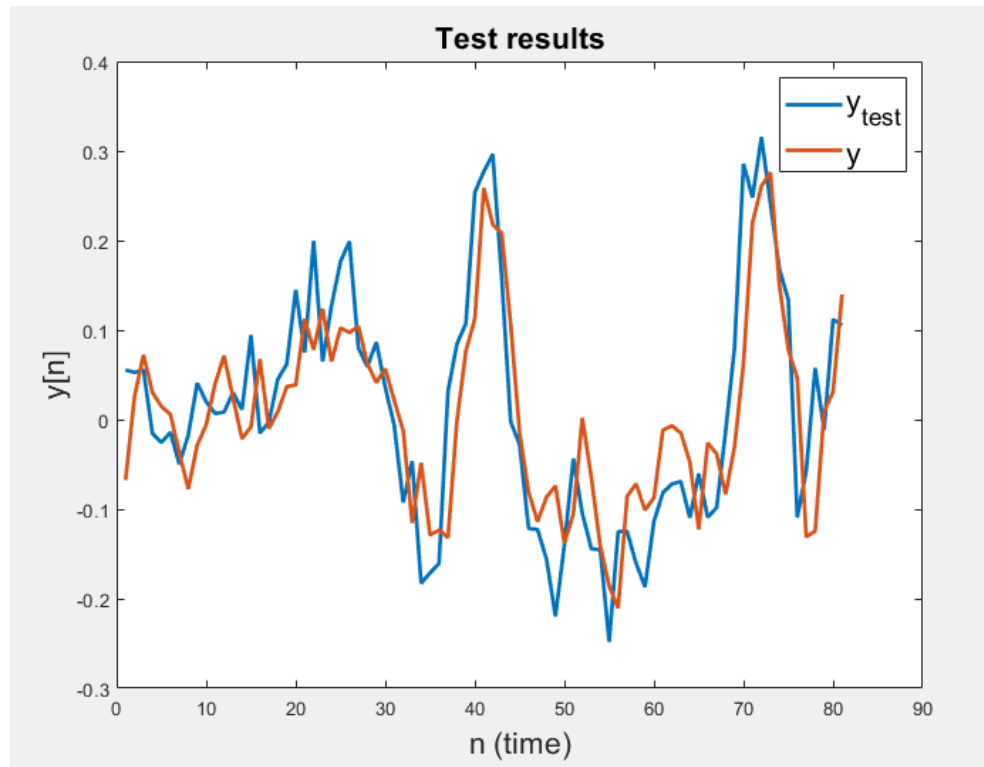
Construct a linear ridge regression that uses inputs X to predict **gamma**. Use 20% of the training data to validate **gamma**.

PART 1.1

The validation procedure:

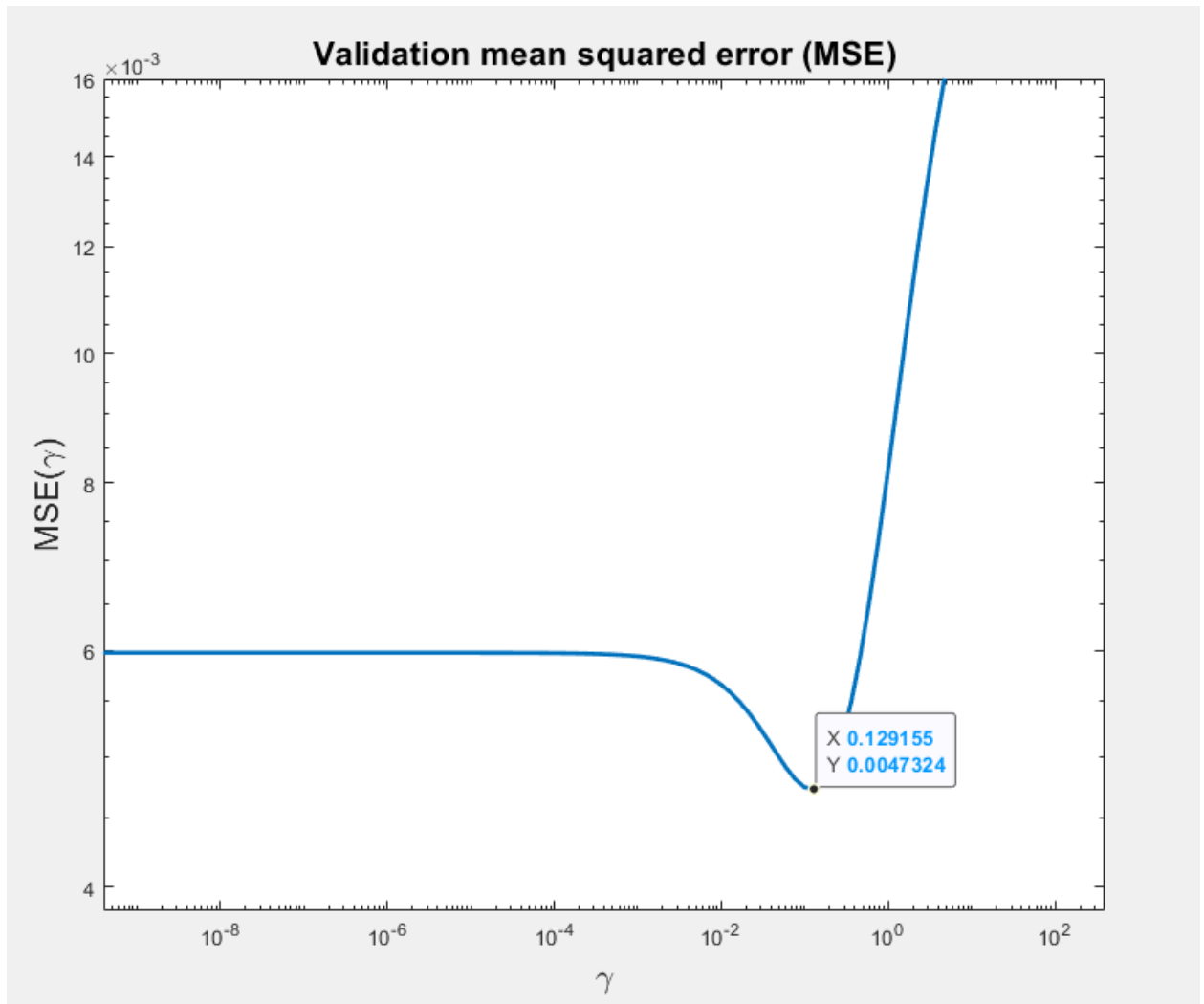
Training the predictor with 80% of the training data and a given value for Y

Running a test for the rest of the training data.



PART 1.2

- Computing the mean square error of the prediction.
- Repeat for a reasonable range of **gamma**

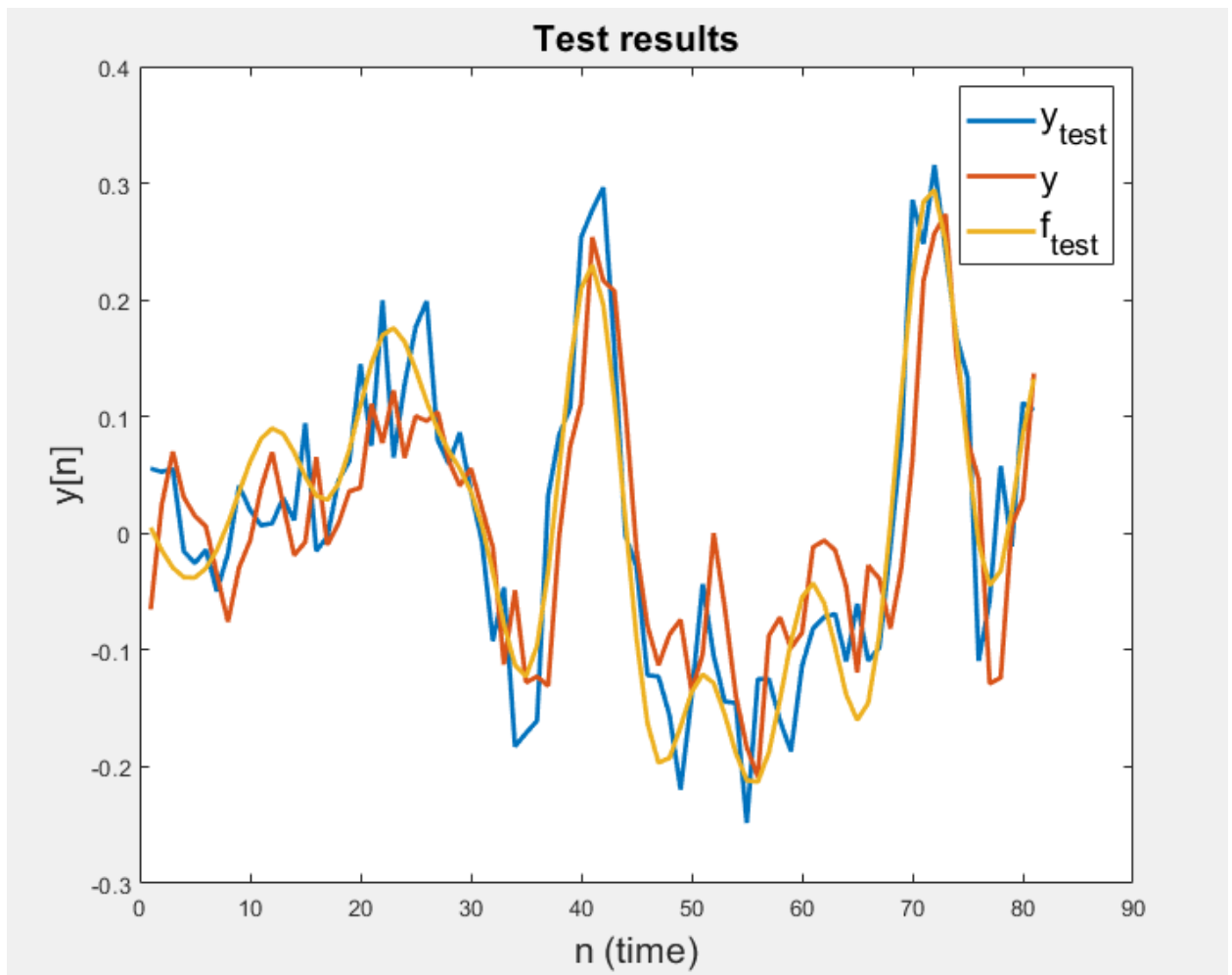


Ideal value for **gamma** was found to be 0.13, and the mean square error **MSE** as a result was 0.0047

PART 2

Choose the value of **Gamma** that produced the best result.

- A graph comparing the real and predicted data.
- A graph of the result of the validation square error.
- The value of the optimal validation and test square errors.



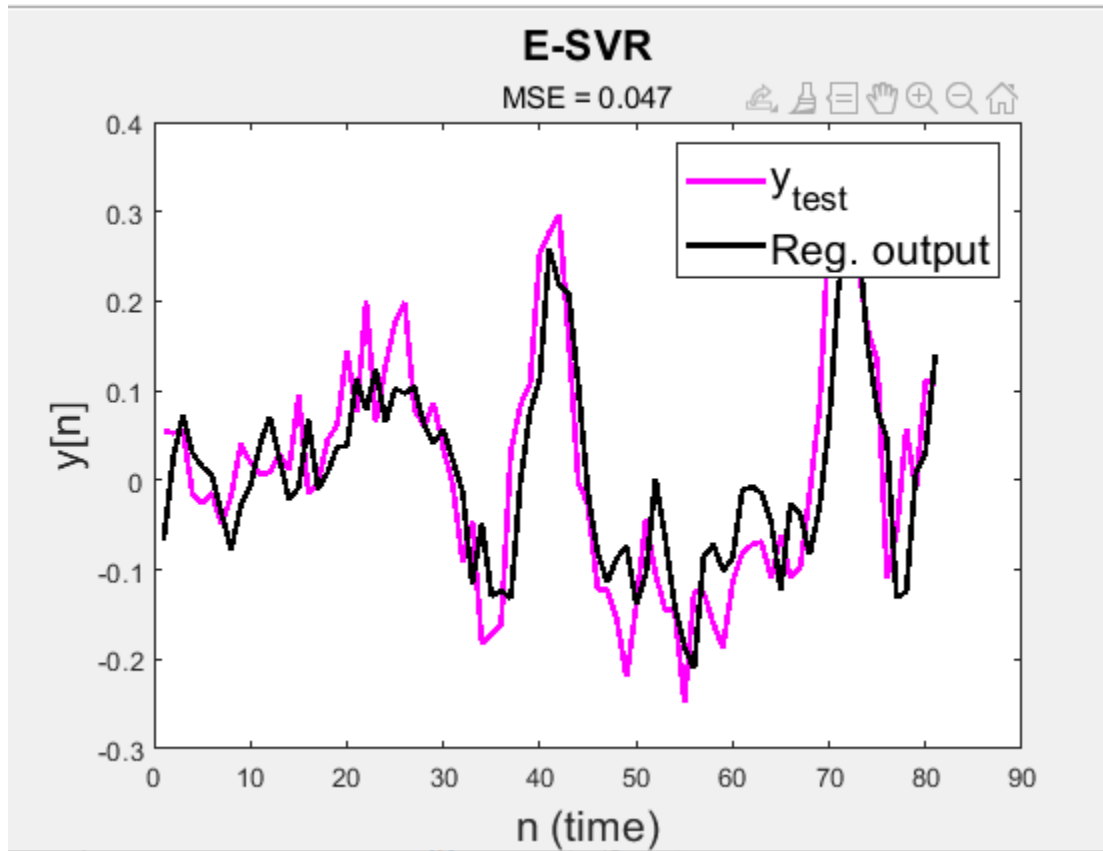
PART 3

- Provide the following results of the experiment:
- A graph comparing the real and predicted data.
- A graph of the result of the validation square error.
- The value of the optimal validation and test square errors.
- A written comparison of the results of this experiment and the previous ones.

PART 3.1

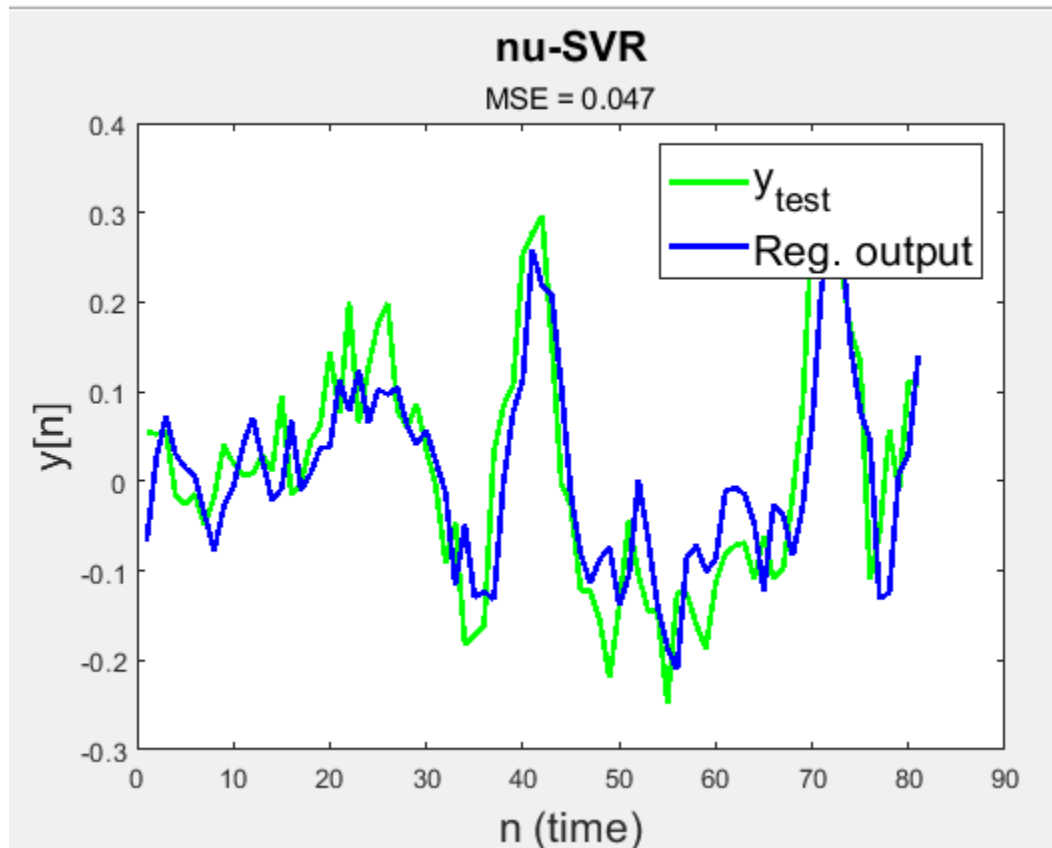
ϵ — VECTOR REGRESSION

The criterion is the same: Find a set of constraints that define a set of loss or slack variables to define an empirical risk. Minimize the empirical risk and the complexity in the same function.



Here I had the MSE = 0.047, which was the same as the MSE from the linear ridge regression. The optimal test error in this case turned out to be $\epsilon = .01$ with a C value of 5.

NU-SUPPORT VECTOR REGRESSION



I observed this graph looked the same as the previous two, for ϵ -SVR, and the ridge regression SVR. The MSE was also the same as the previous two. I am not sure if I graphed this correctly or wrote the functions for the MSE right. However I do observe the MSE is the same for all three methods. I understand the purpose of the addition of ν , and the constraints we use for the dual. I observed all three methods look very similar however, through implementation they are very different.