

EEL5840 Elements of Machine Intelligence - HW 2

Madhusudan Govindraju 39267182

The steps done to solve the assignment has been outlined in detail along with the assumptions and the data used. First during the training phase, we calculate the optimal weights W from the training data, using the formula

$$W = R^{-1} * P$$

where

$$R = XX^T$$

$$P = X^T Y$$

For this step we use the following set of hyperparameters(λ) and filter orders.

Table 1: Filter Orders and λ used

Filter Order	4	8	15	25	30
λ	0.001	0.020	0.04	0.08	0.1

The other assumption is that the input is a time series with 1feature per sample and there are N samples. where $N = 3000$ in training $N = 1000$ in validation & test.

The following experiments are done with the above mentioned assumptions

1. We use the above mentioned formula to calculate the optimal weight for the validation data set. The set of optimal weights W obtained for the following filter orders and λ is shown in table ?? , along wth the Mean Squared Error obtained.

With the hyper parameter, filter order and optimal weights obtained from the step above, to calculate the mean squared error. where Y is the desired output and \bar{Y} is the obtained output. This is implemented on the validation data set to obtain the MSE.

The table 1 shows us the MSE obtained for the different values of filter order and λ .

We plot a 3D surface plot - the performance pot as shown in figure1 From these values of MSE we choose the lowest point in the surface which will correspond to the best Weight, hyper parameter and the filter order. We proceed to use those selected values on the test data set to check the error obtained though time. The point with the lowest point in the surface plot in figure 1 will be the optimum point. That is called the local minimum(or even Global minimum in this case).

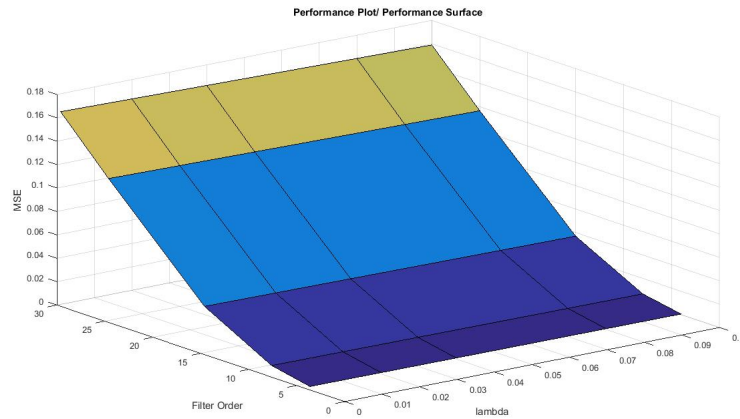
From figure 2 we choose the parameters that give the best performance to test on the test.mat data.

The best performance is obtained with the following parameters, filter order: 4, lambda: 1.000000e-01, MSE: 5.519148e-04

2. We now use these parameters and run it on the test.mat data and calculate the MSE obtained. The MSE obtained for different filter orders are as follows
filterorder : 4 – MSE : 3.811019e + 00
filterorder : 8 – MSE : 2.046562e + 01
filterorder : 30 – MSE : 3.390014e + 02

The figures 3,4,5 depict the errors obtained vs input over time.

Figure 1: Performance Plot



If we notice the error it changes in accordance to the amplitude of the input signal, hence it is not constant over time. The variation of the square error "is a quadratic function of the deviation of the weight parameters and depends only on the input signal statistics" (from lecture slide), this explains the variation of the error across time.

3. Now the same is done for the Noisy data in the file testnoisy.mat. The λ which has the least value is chosen for the order and the error is plotted over time. Because this is a noisy signal there are a lot of peaks. The figures 6,7,8 depict the errors obtained vs input over time. The figures 6,7,8 are not clear hence the figures 9&10 are given for easy understand of the system

The figure 9 compares the errors obtained for the three orders

The figure 10 plots the Average Error obtained over time for the noisy signal over the different orders

From Figure 9 and Figure 10 we can clearly understand that the system with order 4 has the least error. Hence choosing order four will have better characteristics. Because we are dealing with a convex error surface, the least error will be at the local minimum, so the order and the λ which gives the least error will perform best for the system. In the figure one we can see that the least error is obtained for order 4 so the even the noisy system performs better for order 4

For noisy data

filter order: 4, lambda: 1.000000e-01, MSE: 5.879905e+00

filter order: 8, lambda: 1.000000e-01, MSE: 2.513405e+01

filter order: 30, lambda: 1.000000e-01, MSE: 3.658688e+02

Figure 2: Table :For MSE over different hyper parameters

filter order: 4, lambda: 1.000000e-03, MSE: 1.303415e-03
filter order: 4, lambda: 2.000000e-02, MSE: 1.131925e-03
filter order: 4, lambda: 4.000000e-02, MSE: 9.657178e-04
filter order: 4, lambda: 8.000000e-02, MSE: 6.760328e-04
filter order: 4, lambda: 1.000000e-01, MSE: 5.519148e-04
filter order: 8, lambda: 1.000000e-03, MSE: 8.199672e-03
filter order: 8, lambda: 2.000000e-02, MSE: 7.912778e-03
filter order: 8, lambda: 4.000000e-02, MSE: 7.617672e-03
filter order: 8, lambda: 8.000000e-02, MSE: 7.048294e-03
filter order: 8, lambda: 1.000000e-01, MSE: 6.773844e-03
filter order: 15, lambda: 1.000000e-03, MSE: 3.950201e-02
filter order: 15, lambda: 2.000000e-02, MSE: 3.898369e-02
filter order: 15, lambda: 4.000000e-02, MSE: 3.844357e-02
filter order: 15, lambda: 8.000000e-02, MSE: 3.738001e-02
filter order: 15, lambda: 1.000000e-01, MSE: 3.685648e-02
filter order: 25, lambda: 1.000000e-03, MSE: 1.212245e-01
filter order: 25, lambda: 2.000000e-02, MSE: 1.203129e-01
filter order: 25, lambda: 4.000000e-02, MSE: 1.193597e-01
filter order: 25, lambda: 8.000000e-02, MSE: 1.174727e-01
filter order: 25, lambda: 1.000000e-01, MSE: 1.165389e-01
filter order: 30, lambda: 1.000000e-03, MSE: 1.645997e-01
filter order: 30, lambda: 2.000000e-02, MSE: 1.635118e-01
filter order: 30, lambda: 4.000000e-02, MSE: 1.623736e-01
filter order: 30, lambda: 8.000000e-02, MSE: 1.601184e-01
filter order: 30, lambda: 1.000000e-01, MSE: 1.590014e-01

Figure 3: For Filter order 4

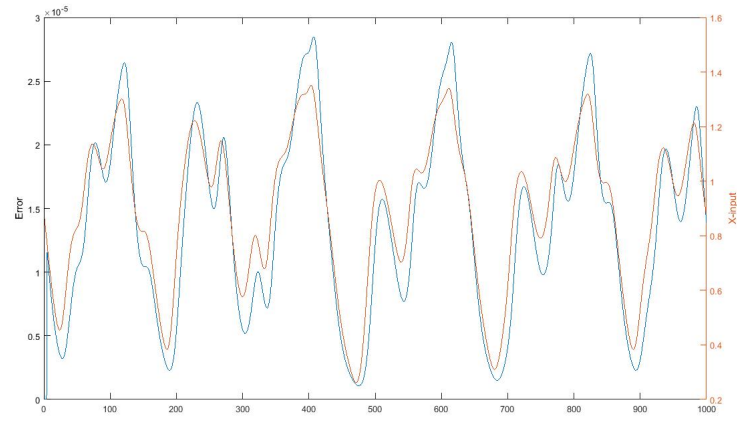


Figure 4: For Filter order 8

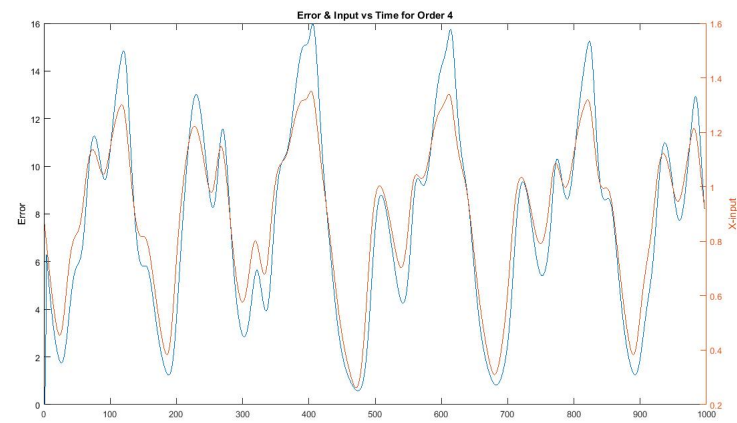


Figure 5: For Filter order 30

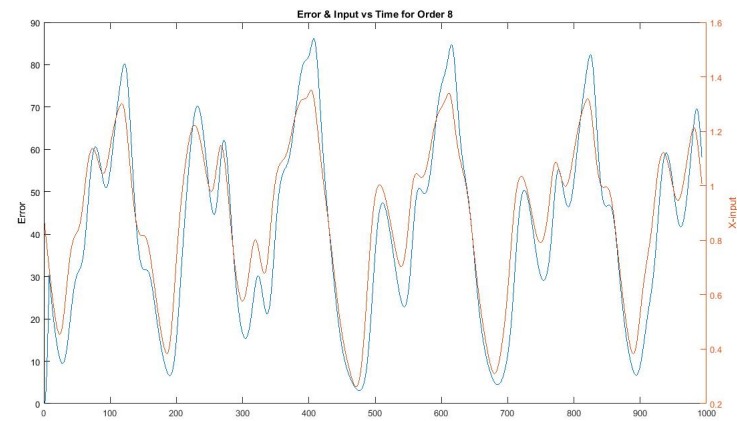


Figure 6: For Filter order 4 - Noisy Signal

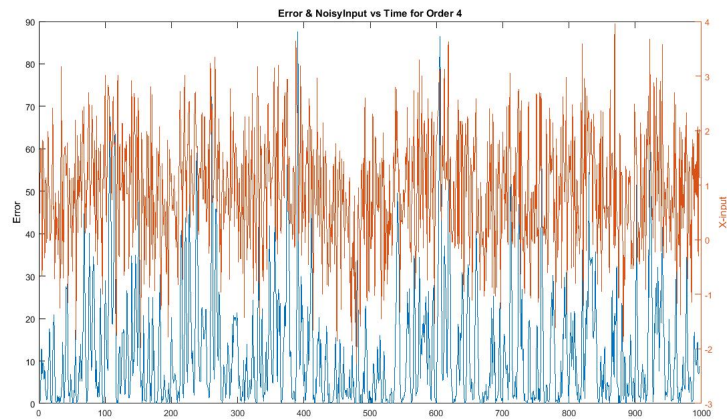


Figure 7: For Filter order 8 - For noisy Signal

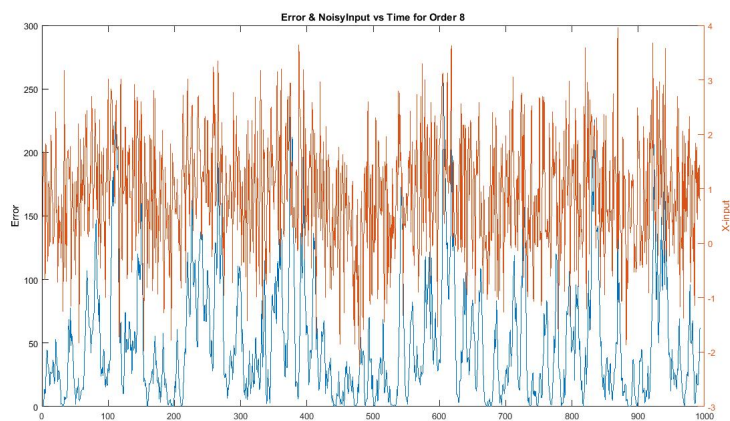


Figure 8: For Filter order 30 - For noisy Signal

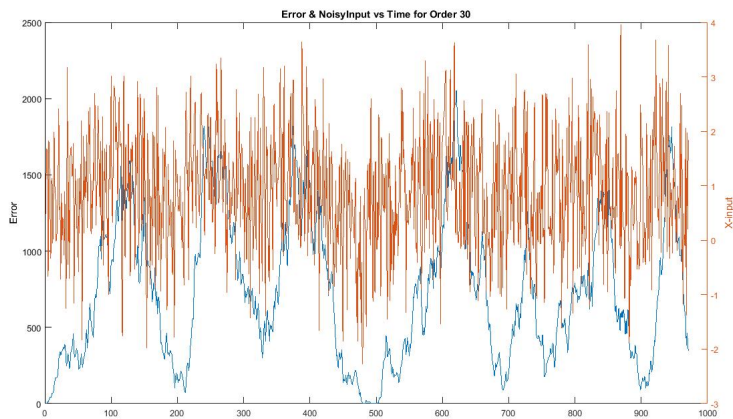


Figure 9: Comparison of Errors for the Different Orders on the Noisy Signal

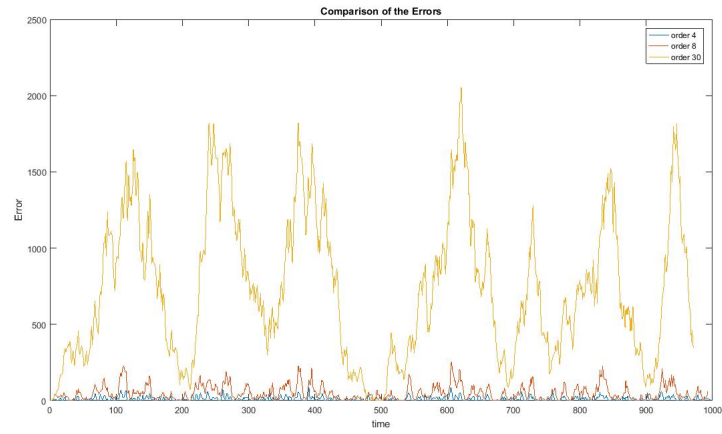


Figure 10: Avg Error over the time range for the noisy signal for different orders

