EEL 6504

Due April 25, 2019

Project 2

The goal is to implement the QKLMS trained with both MSE and MCC using steepest descent learning to predict the next sample of the sun spot time series data, which is available in course website. This data set records the sun spot activity yearly since 1700.

Conventional prediction. Normally we would delay the input to the model by one sample and the current sample will be used for the desired response. Use an input layer with 6 delays for the embedding to map the data to RKHS. You should also compare the performance of the QKLMS with the FIR filter trained with the same two cost functions. More precisely, I would like you to compare the prediction results across linear, nonlinear, MSE and MCC.

Please provide an analysis of the three free parameters of QKLMS, using the learning curve for both systems trained with MSE and MCC, the final prediction error power in a test set, and the tradeoff computation accuracy. Please create the histogram of the errors in the test set for the different systems. Remember to select appropriately the kernel size in MCC (the third parameter). Test the performance of the conventional predictor for 10 and 20 steps ahead prediction and summarize in a table the important results. You can substitute the single delay between the desired and the input by z-10 and z-20, or simply estimate the prediction error 10 and 20 times in the future with the single delay.

Learning the trajectory. The final test that you will be doing is to **generate** with the different trained models (LMS, QKLMS) the sun spot time series using a single point in the trajectory as the initial condition. Iterative prediction trains the model in the same way as it is going to be used in the test set.

To accomplish this, you have to input as the initial condition to the trained the models six samples of the time series to fill the memory in the input layer. After this, instead of continuing to feed the test time series as done above, you feedback the output of the model to its input delayed by one sample. The error is still created between the system output and the next sample. This is called iterative prediction. As you can expect the system will require many more iterations for training, because the initial parameters in the FIR have random values, so initially the output is very different from the next sample. To help the system learn, many researchers pre train the model in the conventional way, and then switched to the new mode (the switching can also be annealed (1-a(n))x(n)+a(n)y(n-1), where a(n) starts small and approaches 1 across n). If you are brave, you can also adapt the parameter a(n)…. but you have a recurrent model.

In a table show the number of samples (horizon of predictability) that each of the different trained systems (conventional and autonomous) can predict, with an instantaneous normalized error < 0.3, the continuation of the time series from which you picked the 6 initial samples. Use 50 different initial conditions to obtain an average number of predicted samples. Explain the reason for the results obtained. This is a project, so I expect that you prepare your report as a 7 pages scientific paper, using the format for IEEE Trans. Signal Processing as the template.