The University of Connecticut **Spring 2021**
Dept. of ECE **KRP**

Problem Set # 5
(Due April 6, 2021).

1. Problem 11.16, Theodoridis, Page 589.

2. Suppose we have N data points $\{\underline{x}_i: i=1,2,..,N\}$. We want to find the smallest enclosing sphere for the transformed feature samples $\{\underline{\Phi}(\underline{x}_i): i=1,2,..,N)$ by solving the following optimization problem:

$$\min_{r>0,\underline{c}} r^2$$
$$\text{subject to } [\underline{\Phi}(\underline{x}_i)-\underline{c}]^T[\underline{\Phi}(\underline{x}_i)-\underline{c}] \le r^2 \ \forall i=1,2,..,N$$

Show that the optimal solution satisfies $\underline{c}=\sum_{i=1}^{N}\alpha_i\underline{\Phi}(\underline{x}_i)$ where $\{\alpha_i\}$ are the solution to the dual optimization problem:

$$\max_{\underline{\alpha}} \sum_{i=1}^{N}\alpha_i K(\underline{x}_i,\underline{x}_i)-\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j K(\underline{x}_i,\underline{x}_j)$$
$$\text{subject to}: \underline{\alpha}\ge\underline{0} \text{ and } \sum_{i=1}^{N}\alpha_i=1.$$

Compute r.

3. Problem 11.8, Theodoridis, Page 588.

4. In this problem, we use the back-propagation algorithm to solve a difficult nonlinear prediction problem and compare its performance with that of the LMS algorithm. Consider the time series modeled by

$$y(n)=w(n)+\beta\, w(n-1)w(n-2)$$

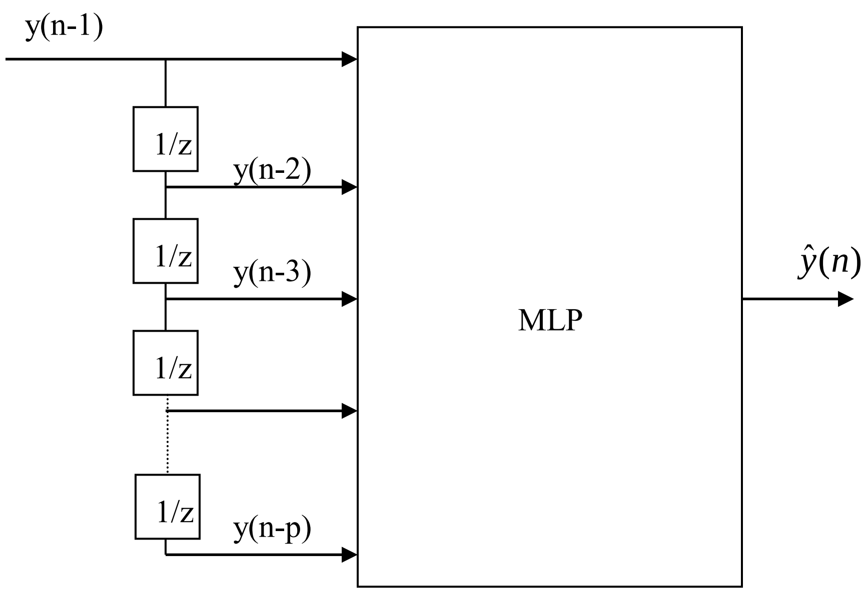
The time series has zero mean, and is uncorrelated and therefore has a white spectrum. However, the time series sample are not independent of each other, and therefore a higher-order predictor can be constructed. The variance of the model output is:

$$\sigma_y^2=\sigma_w^2(1+\beta^2\sigma_w^2)$$

where σ_w^2 is the white-noise input variance set to unity and β= ½. With these values $\sigma_y^2=1.25$.

(a) Construct a MLP with an input layer of 6 nodes, a hidden layer of 16 neurons, and a single output neuron. Fig. 1 shows the network architecture of MLP used as a predictor, where a tapped delay line is used to feed the input layer of the network. The hidden neurons use logistic nonlinearities, whereas the output neuron operates as a linear combiner. The network is trained with the heavy-ball back propagation algorithm with the following parameters: learning-rate parameter η=0.001, momentum constant μ=0.9, Total number of samples processed = 100,000, Number of samples per epoch =1,000, Total number of epochs = 100. Compute the learning curve of the nonlinear predictor, with the variance of the predictor plotted as a function of the number of training epochs.

(b) Repeat the experiment using the LMS algorithm designed to perform a linear prediction on an input of 6 samples. The learning rate of the LMS algorithm is set at 0.9.



The results should reveal that initially the back-propagation algorithm and the LMS algorithm follow a similar path, and then the back-propagation algorithm continues to improve, finally producing a prediction variance close to the ideal value. Use MATLAB NN toolbox, if necessary.