,

CS 6673 fall 2018

PROJECT 2 (100 points max), Due October 23, 2018
Submit a hardcopy before the start of class

Problem 1 (100 points max)

Consider an ADALINE with multiple bipolar output units. Train it using the delta (LMS) rule for the following training set:

(class 1)

$$\mathbf{s}^{(1)} = \begin{bmatrix} 1 & 1 & -1 \end{bmatrix},\ \mathbf{s}^{(2)} = \begin{bmatrix} 1 & 2 & -1 \end{bmatrix} \quad \text{with targets} \quad \begin{bmatrix} -1 & -1 \end{bmatrix},$$

(class 2)

$$\mathbf{s}^{(3)} = \begin{bmatrix} 2 & -1 & 1 \end{bmatrix},\ \mathbf{s}^{(4)} = \begin{bmatrix} 2 & 0 & 1 \end{bmatrix},\ \mathbf{s}^{(5)} = \begin{bmatrix} 1 & -2 & 1 \end{bmatrix},\ \mathbf{s}^{(6)} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \quad \text{with targets} \quad \begin{bmatrix} -1 & 1 \end{bmatrix},$$

(class 3)

$$\mathbf{s}^{(7)} = \begin{bmatrix} -1 & 2 & 1 \end{bmatrix},\ \mathbf{s}^{(8)} = \begin{bmatrix} -2 & 1 & 1 \end{bmatrix} \quad \text{with targets} \quad \begin{bmatrix} 1 & -1 \end{bmatrix},$$

(class 4)

$$\mathbf{s}^{(9)} = \begin{bmatrix} -1 & -1 & -1 \end{bmatrix},\ \mathbf{s}^{(10)} = \begin{bmatrix} -2 & -2 & -1 \end{bmatrix},\ \mathbf{s}^{(11)} = \begin{bmatrix} -2 & -1 & -1 \end{bmatrix} \quad \text{with targets} \quad \begin{bmatrix} 1 & 1 \end{bmatrix}.$$

Assume zero initial weights and biases, find numerically the weights and biases as accurately as you can without using any kind of inverses (or solving a set of linear equations). You will have to experiment with different constant values of the learning rate α, as well as different scheduling schemes for gradually decreasing α. Run a preset number of steps each time.

Can all the training patterns be correctly classified? What are the weights and biases that you get? How many cycles through the training set did you make? What is the value of α at the end?

If we define the mean square error as

$$\frac{1}{Q}\sum_{q=1}^{Q}\sum_{j=1}^{M}\left(y_j - t_j^{(q)}\right)^2,$$

how large is this error that you find at the end of the training process?

[Note that if all the training vectors are classified correctly, then this mean square error obtained using the weights and biases that you found, and by applying the transfer function, must be exactly zero.]

Using the best convergence method that you found above, consider a training where $\mathbf{s}^{(8)} = \begin{bmatrix} x & 1 & 1 \end{bmatrix}$. But instead of $x = -2$, consider increasing x. Is there a value above which something noteworthy happens? Describe the situation and why it happens.

Comment on all your findings as much as you can. Include a copy of you computer program and all the key results that you find.