Artificial Neural Networks

Exercise Set 3

Antonio Peters

March 30, 2016

1 Proof of Convergence of the Perceptron Learning Rule

Assume the matrix $P = [p_1, \ldots, p_m]$ as a set of input patterns and the matrix $T = [t_1, \ldots, t_m]$ as a set of targets in a perceptron network with s layers. Each column of P, p_i is activated and the result $a_i \in A$ is compared to the corresponding column of T, t_i. The activation can be seen as

$$\begin{aligned} &a_i = \text{hardlim}(W * p_i + b) \\ &\text{Or for the entire matrix} \\ &A = \text{hardlim}(W * P + b) \end{aligned} \tag{1}$$

We can reorder the activation by grouping the weighting and bias and adding a row of ones to P as in Equation 2

$$V = [Wb] \quad \text{and} \quad q_i = \begin{bmatrix} p_i \\ 1 \end{bmatrix} \quad \text{or} \quad Q = \begin{bmatrix} P \\ 1, \ldots, 1 \end{bmatrix} \tag{2}$$

This is then updated to adjust for the error between A and T, E with $e_i \in E$, $e_i = t_i - a_i$ and the matrix V is updated for each element of Q such that

$$V_{j+1} = V_j + e_i q_i \tag{3}$$

Where V_{j+1} is the updated V to be used with V_0 being randomly set and in this particular case we prove for $V_0 \neq 0$. We assume that $\forall j$, $V_{j+1} \neq V_j$