NUCL 575 HMWK 2

Alex Hagen

9/26/13

Problem 1 [1, prob. 8.1] The backpropagation training algorithm is developed in Section 8.3, using the logistic function as the activation function where the derivative has the convenient form given in Equation 8.1-4. Derive the backpropagation training algorithm for the case where the activation function is an arctan function where the derivative is given by Equation 8.1-6.

For the case where the activation function is given by as

$$\Phi(I) = \frac{2}{\pi} \arctan(\alpha I) \tag{8.1-5}$$

we must find the result, given by

$$\frac{\partial \Phi(I)}{\partial I} = \frac{2}{\pi} \left[\frac{\alpha}{1+\alpha^2 I^2} \right] \tag{8.1-6}$$

This can be done by developing the backpropagation training algorithm as in section 8.3.

We must first relate the total error ε to each parameter involved, of which we have the activation function $\Phi(I)$ and the intensity αI. The derivative of this total error is proportional to the change we will make to the weight at that point, as given by

$$\Delta w_{pq.k} = -\eta_{p.q} \frac{\partial \varepsilon_q^2}{\partial w_{pq.k}} \tag{8.3-3}$$

So, the critical part of this is to find the derivative $\frac{\partial \varepsilon_q^2}{\partial w_{pq.k}}$. This can be found by the chain rule, therefore

$$\frac{\partial \varepsilon_q^2}{\partial w_{pq.k}} = \frac{\partial \varepsilon_q^2}{\partial \Phi_{q.k}} \frac{\partial \Phi_{q.k}}{\partial I_{q.k}} \frac{\partial I_{q.k}}{\partial w_{pq.k}} \tag{8.3-4}$$

Because of the solutions given in section 8.3, and the result obtained in Equation , we can put all of these pieces together to get a full algorithm.

$$\begin{aligned}
\frac{\partial \varepsilon_q^2}{\partial w_{pq.k}} &= \underbrace{\frac{\partial \varepsilon_q^2}{\partial \Phi_{q.k}}} \cdot \underbrace{\frac{\partial \Phi_{q.k}}{\partial I_{q.k}}} \cdot \underbrace{\frac{\partial I_{q.k}}{\partial w_{pq.k}}} \\
\frac{\partial \varepsilon_q^2}{\partial w_{pq.k}} &= -2\left[T_q - \Phi_{q.k}\right] \cdot \frac{2}{\pi}\left[\frac{\alpha}{1+\alpha^2 I^2}\right] \cdot \Phi_{p.j}
\end{aligned}$$

Therefore, similar to Equations 8.3-9 - 8.3-12, we have

$$\frac{\partial \varepsilon_q^2}{\partial w_{pq.k}} = -\frac{4}{\pi} \phi_{w.k} \left[T_q - \Phi_{q.k}\right] \left[\frac{\alpha}{1+\alpha^2 I^2}\right] = \delta_{pq.k} \Phi_{p.j}$$

and

$$w_{pq.k}(N-1) = w_{pq.k}(N) - \eta_{p.q} \delta_{pq.k} \Phi_{p.j}$$

Problem 2 [1, prob. 8.2] Derive the backpropagation training algorithm for the case where the neurons in the hidden layer have a logistic function for the activation function and the neurons in the output layer have linear activation functions. Compare the results with those obtained in Section 8.7 where this arrangement of activation functions are used.

In this case, there is an extra layer, r, added in, but this layer only has a linear activation function. Because there is a linear activation function, there will be no proportionality constant, as that would only modify the weights. So the output is defined as

$$T_r = \Phi_{r.k} = I_{r.k}$$

and the algorithm is given by the above chain rule, as before.

$$\begin{array}{ccccccccccc} \frac{\partial \varepsilon_q^2}{\partial w_{pq.k}} & = & \underbrace{\frac{\partial \varepsilon_r^2}{\partial \Phi_{r.k}}} & \cdot & \underbrace{\frac{\partial \Phi_{r.k}}{\partial I_{r.k}}} & \cdot & \underbrace{\frac{\partial I_{r.k}}{\partial w_{qr.k}}} & \cdot & \underbrace{\frac{\partial \Phi_{q.k}}{\partial I_{q.k}}} & \cdot & \underbrace{\frac{\partial I_{q.k}}{\partial w_{pq.k}}} \\ \frac{\partial \varepsilon_q^2}{\partial w_{pq.k}} & = & -2\left[T_r - \Phi_{r.k}\right] & \cdot & 1 & \cdot & \Phi_{q.k} & \cdot & \alpha\Phi_{q.k}\left[1-\Phi_{q.k}\right] & \cdot & \Phi_{p.k} \end{array}$$

So the overall algorithm is given by

$$\frac{\partial \varepsilon_q^2}{\partial w_{pq.k}} = -2\alpha\Phi_{q.k}^2\left[1-\Phi_{q.k}\right]\left[T_r - \Phi_{r.k}\right]\Phi_{p.j}$$

so

$$w_{qr.k}(N-1) = w_{qr.k}(N) - \delta_{qr.k}\Phi_{p.k}$$

which is identically equal to

$$\mathbf{W}_k = \left(\mathbf{\Phi}_j \mathbf{\Phi}_j^{'}\right)^{-1} \mathbf{\Phi}_j^{'} \mathbf{I}_k$$

References

[1] L. H. Tsoukalas and R. E. Uhrig. *Fuzzy and Neural Approaches in Engineering*. John Wiley & Sons, Inc., New York, 1997.