Chapter 11: Backpropagation

Brandon Morgan

1/21/2021

E11.7

We are given the two layer network linear network in Figure E11.5, along with the following initial weight and bias matrices:

$$w_1(0) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, b_1(0) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

and

$$w_2(0) = \begin{bmatrix} -1 & 1 \end{bmatrix}, b_2(0) = \begin{bmatrix} 3 \end{bmatrix}$$

1

We have the input target pair $p_1 = 1, t_1 = 2$, lets compute the error. Because our transfer function is linear, $a_i = n_i$:

First Layer:

$$a_1 = n_1 = W_1 p_1 + b_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

Second Layer:

$$a_2 = n_2 = W_2 p_2 + b_2 = \begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 0 \end{bmatrix} + \begin{bmatrix} 2 \end{bmatrix} = \begin{bmatrix} -1 \end{bmatrix}$$

Thus, our error is $e = t_1 - a_2 = 2 - (-1) = 3$.

2

To compute sensitivites, we first need to find the first derivatives of the transformations by use of chain rule.

$$F_2(n_2) = \frac{d}{dn} n = 1$$

$$F_1(n_1) = \frac{d}{dn} n = 1$$

Thus,

$$s_2 = -2F_2(n_2)(t-a) = -2(1)(3) = -6$$

$$s_1 = F_1(n_1)W_2 s_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} [-6] = \begin{bmatrix} 6 \\ -6 \end{bmatrix}$$

3

We can compute the partial derivative $\frac{\partial e^2}{\partial w_1}$

$$\frac{\partial e^2}{\partial w_1} = 2e\frac{\partial e}{\partial w_1} = 2e\frac{\partial}{\partial w_1}(t-a_2) = 2e\frac{\partial}{\partial w_1}(t - w_2(w_1 p + b_1) + b_2) = 2ew_2 p$$

$$2ew_2 p = 2(3) \begin{bmatrix} -1 \\ 1 \end{bmatrix} [1] = \begin{bmatrix} -6 \\ 6 \end{bmatrix}$$