Chapter 11: Backpropagation

Evangelos Stamos

2024.10.22

E11.7

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

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

and

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

1

We have the input target pair $p_1 = 1$, $t_1 = 2$. Let's 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

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

Thus, our error is:

$$e = t_1 - a_2 = 2 - 0 = 2$$

2

To compute sensitivities, we first need to find the derivatives of the transformations using the chain rule:

$$F_2(n_2) = \frac{d}{dn} n_2 = 1, \quad F_1(n_1) = \frac{d}{dn} n_1 = 1$$

Thus,

$$s_2 = -2F_2(n_2)(t-a) = -2(1)(2) = -4$$

$$s_1 = F_1(n_1) W_2 s_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \begin{bmatrix} -4 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \end{bmatrix}$$

3

We can compute the partial derivative $\frac{\partial e^2}{\partial W_1}$ as follows:

$$\frac{\partial e^2}{\partial W_1} = 2e \frac{\partial}{\partial W_1}(t - a_2) = 2e \frac{\partial}{\partial W_1}(t - W_2(W_1 p + b_1) + b_2) = 2eW_2 p$$

Thus,

$$2eW_2 p = 2(2) \begin{bmatrix} -1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \end{bmatrix} = \begin{bmatrix} -4 \\ 4 \end{bmatrix}$$