Chapter 11: Backpropagation

Brandon Morgan

1/21/2021

E11.12

We are given the two layer network linear network in Figure E11.10, along with the following initial weight and bias matrices:

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

and

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

In addition, we are given the following transfer function:

$$f_1(n) = n^3,\ f_2(n) = n$$

1

We have the input target pair:

$$p_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, t_1 = \begin{bmatrix} -1 \end{bmatrix}$$

Therefore, let $a_0 = p_1$. Lets now perform one iteration of steepest descent backpropagation with learning rate $\alpha = 0.5$.

Calculating Error

First, lets compute the error:

$$n_1 = w_1 a_0 + b_1 = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$a_1 = n_1^3 = (\begin{bmatrix} -1 \\ 1 \end{bmatrix})^3 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$a_2 = n_2 = w_2 a_1 + b_2 = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} + [1] = 1$$

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

Calculating Sensitivites

First, we must calculate the partial derivatives

$$f_2'(n_2) = 1$$

$$f_1'(n_1) = 3n_1^2$$

Now we backpropage the sensitivites:

$$s_2 = -2f_2'(n_2)(t - a_2) = -2(1)(-2) = 4$$

$$s_1 = f_1(n_1)w_2^T s_2 = \begin{bmatrix} 3(-1)^2 & 0 \\ 0 & 3(1)^2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} (4) = \begin{bmatrix} 12 \\ 12 \end{bmatrix}$$

Updating weights

$$w_2(1) = w_2(0) - \alpha s_2 (a_1)^T = \begin{bmatrix} 1 & 1 \end{bmatrix} - 0.5(4) \begin{bmatrix} -1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -1 \end{bmatrix}$$

$$w_1(1) = w_1(0) - \alpha s_1 (a_0)^T = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} - 0.5 \begin{bmatrix} 12 \\ 12 \end{bmatrix} \begin{bmatrix} -1 & 1 \end{bmatrix} = \begin{bmatrix} 7 & -7 \\ 7 & 6 \end{bmatrix}$$

Updating Biases

$$b_2(1) = b_2(0) - \alpha s_2 = [1] - 0.5(4) = -1$$

$$b_1(1) = b_1(0) - \alpha s_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} - 0.5 \begin{bmatrix} 12 \\ 12 \end{bmatrix} = \begin{bmatrix} -5 \\ -4 \end{bmatrix}$$

This completes one iteration of steepest descent backpropagation.