5/9/10

Error gradiant:

Error gradient is the rate of change of activation function used; multiplied by the error at iteration 'p'. It is represented by $\delta(p)$.

If $y_k(p)$ is the activation function, then $\delta_k(p) = \frac{\partial y_k(p)}{\partial x_k(p)} \cdot e_k(p)$

Error gradiant of a backpropogation Neural Network:

In backpropagation neural network, activation function,

$$y_k(p) = \frac{1}{1+e^{-x_k(p)}} \quad \text{——— (1)}$$

Deducting 1 from both sides —

$$\therefore \; 1 - y_k(p) = 1 - \frac{1}{1+e^{-x_k(p)}} = \frac{1+e^{-x_k(p)}-1}{1+e^{-x_k(p)}}$$

$$= \frac{e^{-x_k(p)}}{1+e^{-x_k(p)}} \quad \text{——— (2)}$$

Differentiating equ^n (1) with r. to. x —

$$\frac{\partial y_k(p)}{\partial x_k(p)} = \frac{\partial}{\partial x_k p}\left[\frac{1}{1+e^{-x_k(p)}}\right]$$

$$= \frac{\left[1+e^{-x_k(p)}\right]\cdot 0 - 1\left[0-e^{-x_k(p)}\right]}{\left[1+e^{-x_k(p)}\right]^2}$$

$$= \frac{e^{-x_k(p)}}{\left[1+e^{-x_k(p)}\right]^2}$$

$$= \frac{1}{\left[1+e^{-x_k(p)}\right]} \cdot \frac{e^{-x_k(p)}}{\left[1+e^{-x_k(p)}\right]}$$

$$= y_k(p) \cdot \left[1 - y_k(p)\right]$$ [using equ^n (1) and (2)]

So, according to definition error gradiant is —

$$\delta_k(p) = \frac{\partial y_k(p)}{\partial x_k(p)} \cdot e_k(p)$$

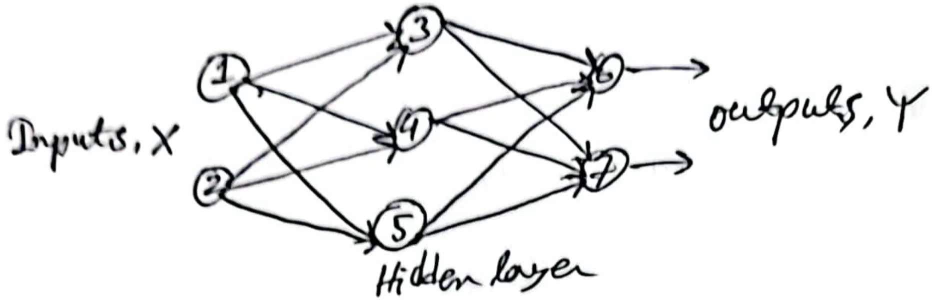
$$= y_k(p) \cdot [1 - y_k(p)] \cdot e_k(p) \quad [\text{using eqn } 3]$$

So,

$$\delta_k(p) = y_k(p)[1 - y_k(p)] \cdot e_k(p)$$

(derived)

Q) For the following NN, $w13 = 0.3$, $w14 = -0.4$, $w15 = 0.6$, $w23 = -0.9$, $w24 = -0.2$, $w25 = -0.3$, $w36 = -0.4$, $w37 = -0.1$, $w46 = 0.3$, $w47 = -0.3$, $w56 = 0.4$, $w57 = 0.8$, learning rate $= 0.2$
If $x1 = 1$, $x2 = 0$ and $X = [1, 0]$ and $Y = [0, 1]$, find the updated weights after one iteration.



Soln: Given input set $X = [1, 0]$, desired output set $Y = [0, 1]$.

① Let us at first find out the output at hidden layer:

$$y_3^{(1)} = \text{sigmoid}(x_1 w_{13} + x_2 w_{23} - \theta_3)$$
$$= \text{sigmoid}(1 \times 0.3 + 0 - 0)$$
$$= \text{sigmoid}(0.3) = \frac{1}{1 + e^{-0.3}} = 0.574$$

$$y_5^{(1)} = \text{sigmoid}(x_1 w_{15} + x_2 w_{25} - \theta_5)$$
$$= \text{sigmoid}(1 \times 0.6 + 0 - 0)$$
$$= \text{sigmoid}(0.6) = \frac{1}{1 + e^{-0.6}} = 0.646$$

$$y_4^{(1)} = \text{sigmoid}(x_1 w_{14} + x_2 w_{24} - \theta_4)$$
$$= \text{sigmoid}(1 \times (-0.4) + 0 - 0)$$
$$= \text{sigmoid}(-0.4) = \frac{1}{1 + e^{0.4}} = 0.40$$

② Actual outputs at output level:

$$y_6 = \text{sigmoid}(y_3 w_{36} + y_4 w_{46} + y_5 w_{56} - \theta_6)$$
$$= \text{sigmoid}(0.574(-0.4) + 0.3 \times 0.4 + 0.646 \times 0.4 - 0)$$
$$= \text{sigmoid}(-0.23 + 0.12 + 0.2584 - 0)$$
$$= \text{sigmoid}(0.1484) = \frac{1}{1 + e^{-0.1484}} = 0.537$$

$$y_7 = \text{sigmoid}(y_3 w_{37} + y_4 w_{47} + y_5 w_{57} - \theta_7)$$

$$= \text{sigmoid}(0.574(-0.1) + 0.4(0.3) + 0.646 \times 0.8 - 0)$$

$$= \text{sigmoid}(-0.0574 - 0.12 + 0.5168)$$

$$= \text{sigmoid}(0.3394)$$

$$= \frac{1}{1 + e^{-0.3394}} = 0.584 \leftarrow$$

(3) Now calculating error at output level:

$$e_6 = y_{d6} - y_6 = 0 - 0.537 = -0.537$$

$$e_7 = y_{d7} - y_7 = 1 - 0.584 = 0.416$$

(4) Now, calculating error gradients,

$$\delta_6 = y_6 [1 - y_6] \times e_6$$

$$= 0.537 [1 - 0.537] (-0.537) = -0.1335 \leftarrow$$

$$\delta_7 = y_7 [1 - y_7] \times e_7$$

$$= 0.584 [1 - 0.584] \times 0.416 = 0.101 \leftarrow$$

Now updating weights betn o/p and hidden layer:

(5) $\Delta w_{36} = \alpha \times y_3 \times \delta_6 = 0.2 \times 0.574 (-0.1335)$

$$= -0.0153 \leftarrow$$

$$\Delta w_{37} = \alpha \times y_3 \times \delta_7 = 0.2 \times 0.574 \times (0.101)$$

$$= 0.012 \leftarrow$$

$$\Delta w_{46} = \alpha \times y_4 \times \delta_6 = 0.2 \times 0.4 \times (-0.1335)$$

$$= -0.01 \leftarrow$$

$$\Delta w_{47} = \alpha \times y_4 \times \delta_7 = 0.2 \times 0.4 \times 0.101$$

$$= 0.0081 \leftarrow$$

②

$\Delta w_{56} = \alpha \times y_5 \times \delta_6 = 0.2 \times 0.646 \times (-0.1337)$

$= -0.017 \leftarrow$

$\Delta w_{57} = \alpha \times y_5 \times \delta_7 = 0.2 \times 0.646 \times 0.101$

$= 0.013 \leftarrow$

Therefore,

(6) $w_{36}(2) = w_{36}(1) + \Delta w_{36} = -0.4 + (-0.0153) = -0.4153 \leftarrow$

$w_{37}(2) = w_{37}(1) + \Delta w_{37} = -0.1 + 0.012 = -0.088 \leftarrow$

$w_{46}(2) = 0.3 - 0.01 = 0.29 \leftarrow$

$w_{47}(2) = w_{47}(1) + \Delta w_{47} = -0.3 + 0.0081 = -0.2919 \leftarrow$

$w_{56}(2) = w_{56}(1) + \Delta w_{56} = 0.4 - 0.017 = 0.383 \leftarrow$

$w_{57}(2) = 0.8 + 0.013 = 0.813 \leftarrow$

Error gradients at hidden layer:

(7) $\delta_3 = y_3 \cdot [1 - y_3] \times [w_{36} \times \delta_6 + w_{37} \times \delta_7]$

$= 0.574 \times 0.426 \times [-0.4 \times (-0.1335) + (-0.1)(0.101)]$

$= 0.244 [0.0534 - 0.0101]$

$= 0.244 \times 0.0433 = 0.0105 \leftarrow$

$\delta_4 = y_4 [1 - y_4] \times [w_{46} \times \delta_6 + w_{47} \times \delta_7]$

$= 0.4 [1 - 0.4] \times [0.3 (-0.1335) + (-0.3)(0.101)]$

$= 0.24 \times [-0.04005 - 0.0303]$

$= 0.24 (-0.07035) = -0.017 \leftarrow$

$\delta_5 = y_5 [1 - y_5] \times [w_{56} \times \delta_6 + w_{57} \times \delta_7]$

$= 0.646 \times [1 - 0.646] \times [0.4 \times (-0.1335) + 0.8 \times 0.101]$

$= 0.2287 [-0.0534 + 0.0808]$

$= 0.2287 \times 0.0274 = 0.0063 \leftarrow$

(8) $\Delta w_{13} = \alpha \times x_1 \times \delta_3 = 0.2 \times 1 \times 0.0105 = 0.0021$

$\Delta w_{14} = \alpha \times x_1 \times \delta_4 = 0.2 \times 1 \times 0.0063 = 0.00126$

$\Delta w_{23} = \alpha \times x_2 \times \delta_3 = 0.2 \times 0 \times 0.0105 = 0$

$\Delta w_{24} = \alpha \times x_2 \times \delta_4 = 0.2 \times 0 \times 0.0105 = 0$

$\Delta w_{25} = \alpha \times x_2 \times \delta_5 = 0.2 \times 0 \times 0.0063 = 0$

Therefore,

(9) $w_{13}(2) = w_{13}(1) + \Delta w_{13} = 0.3 + 0.0021 = 0.3021$

$w_{14}(2) = w_{14} + \Delta w_{14} = -0.4 - 0.0034 = -0.4034$

$w_{15}(2) = w_{15} + \Delta w_{15} = 0.6 + 0.00126 = 0.60126$

$w_{23}(2) = w_{23} + \Delta w_{23} = -0.9 + 0 = -0.9$

$w_{24}(2) = w_{24} + \Delta w_{24} = -0.2 + 0 = -0.2$

$w_{25}(2) = w_{25} + \Delta w_{25} = -0.3 + 0 = -0.3$