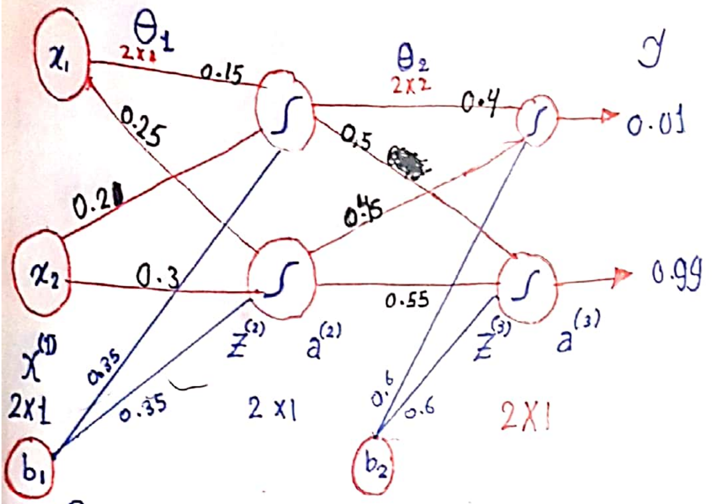
Back Propagation



⇨ Given ⇨ $x_1 = 0.05$, $x_2 = 0.1$, $b_1 = 1$, $b_2 = 1$, $\alpha = 0.5$

Do Back Propagation to make $y_1 = 0.01$, $y_2 = 0.99$ for one iteration.

Solution

⇨ Forward Propagation

$$Z^{(2)} = \theta_1 \cdot x^{(1)} = \begin{bmatrix} 0.15 & 0.20 & 0.35 \\ 0.25 & 0.3 & 0.35 \end{bmatrix} \cdot \begin{bmatrix} 0.05 \\ 0.1 \\ 1 \end{bmatrix}$$

$$Z^{(2)} = \begin{bmatrix} 0.3775 \\ 0.3925 \end{bmatrix}$$

$$a^{(2)} = \frac{1}{1 + e^{-Z}} = \begin{bmatrix} 0.5932699921 \\ 0.5968843783 \end{bmatrix}$$

$$Z^{(3)} = \theta_2 \cdot a^{(2)} = \begin{bmatrix} 0.4 & 0.45 & 0.6 \\ 0.5 & 0.55 & 0.6 \end{bmatrix} \cdot \begin{bmatrix} 0.5932699 \\ 0.59688 \\ 1 \end{bmatrix} = \begin{bmatrix} 1.105906 \\ 1.224921 \end{bmatrix}$$

b

$$a^{(3)} = \frac{1}{1 + e^{-Z}} = \begin{bmatrix} 0.7513651 \\ 0.7729284 \end{bmatrix}$$

⇨ Back Propagation for Output

[1] Error Calc

$$err^{(3)} = a^{(3)} - y = \begin{bmatrix} 0.7513651 \\ 0.7729284 \end{bmatrix} - \begin{bmatrix} 0.01 \\ 0.99 \end{bmatrix}$$

$$err^{(3)} = \begin{bmatrix} 0.7413651 \\ -0.217072 \end{bmatrix}$$

[2] Layer Differentiation

$$dg(Z^{(3)}) = a^{(3)}(1 - a^{(3)})$$

$$dg(Z^{(3)}) = \begin{bmatrix} 0.7513651 \\ 0.7729284 \end{bmatrix} * \begin{bmatrix} 1- \\ 1- \end{bmatrix}$$

$$dg(Z^{(3)}) = \begin{bmatrix} 0.1868156 \\ 0.1755101 \end{bmatrix}$$

[3] $\delta^{(3)} = err^{(3)} * dg(Z^{(3)})$

$$\delta^{(3)} = \begin{bmatrix} 0.7413651 \\ -0.217072 \end{bmatrix} \cdot \begin{bmatrix} 0.1868156 \\ 0.1755101 \end{bmatrix}$$

$$\delta^{(3)} = \begin{bmatrix} 0.1384986 \\ -0.038098 \end{bmatrix}$$

[4] Update weights

$$d\theta^{(2)} = \alpha \cdot \delta^{(3)} \cdot a^{(2)T}$$

$$d\theta^{(2)} = 0.5 \cdot \begin{bmatrix} 0.1384986 \\ -0.038098 \end{bmatrix} \cdot \begin{bmatrix} 0.5932699 & 0.5968844 \end{bmatrix}$$

$$d\theta^{(2)} = \begin{bmatrix} 0.041 & 0.041[illegible] \\ -0.011 & -0.011 \end{bmatrix}$$

$$\theta^{(2)} = \theta^{(2)} - d\theta^{(2)} = \begin{bmatrix} 0.4 & 0.45 \\ 0.5 & 0.55 \end{bmatrix} - \begin{bmatrix} 0.041 & 0.0418 \\ -0.011 & -0.011 \end{bmatrix}$$

$$\theta^{(2)} = \begin{bmatrix} 0.359 & 0.409 \\ 0.511 & 0.561 \end{bmatrix}$$

⇨ Back Propagation for Hidden Layer

[1] Error Calculation

$err^{(2)} = \Theta_T^{(2)} \cdot \delta^{(3)}$

$err^{(2)} = \begin{bmatrix} 0.4 & 0.5 \\ 0.45 & 0.55 \end{bmatrix} \cdot \begin{bmatrix} 0.1384986 \\ -0.038098 \end{bmatrix}$

$err^{(2)} = \begin{bmatrix} 0.03635044 \\ 0.04137047 \end{bmatrix}$

[2] Layer differentiation

$dg(z^{(2)}) = a^{(2)} [1 - a^{(2)}]$

$dg(z^{(2)}) = \begin{bmatrix} 0.593269921 \\ 0.5968843783 \end{bmatrix} * \begin{bmatrix} 1 - \\ 1 - \end{bmatrix}$

$dg(z^{(2)}) = \begin{bmatrix} 0.24130483 \\ 0.24061341 \end{bmatrix}$

[3] $\delta^{(2)} = err^{(2)} \cdot dg(z^{(2)})$

$\delta^{(2)} = \begin{bmatrix} 0.03635044 \\ 0.04137047 \end{bmatrix} \cdot \begin{bmatrix} 0.24130483 \\ 0.24061341 \end{bmatrix}$

$\delta^{(2)} = \begin{bmatrix} 0.00877 \\ 0.00995 \end{bmatrix}$

[4] Update Weights

$d\theta^{(1)} = \alpha \cdot \delta^{(2)} \cdot X_T^{(1)}$

$= 0.5 \cdot \begin{bmatrix} 0.0087 \\ 0.00995 \end{bmatrix} \cdot \begin{bmatrix} 0.05 & 0.1 \end{bmatrix}$

$d\theta^{(1)} = \begin{bmatrix} 0.0002175 & 0.000435 \\ 0.00024875 & 0.0004975 \end{bmatrix}$

$\theta^{(1)} = \theta^{(1)} - d\theta^{(1)}$

$= \begin{bmatrix} 0.15 & 0.2 \\ 0.25 & 0.3 \end{bmatrix} - \begin{bmatrix} 0.00 \\ \end{bmatrix}$

$\theta^{(1)} = \begin{bmatrix} 0.1497825 & 0.1999565 \\ 0.24975125 & 0.2995025 \end{bmatrix}$