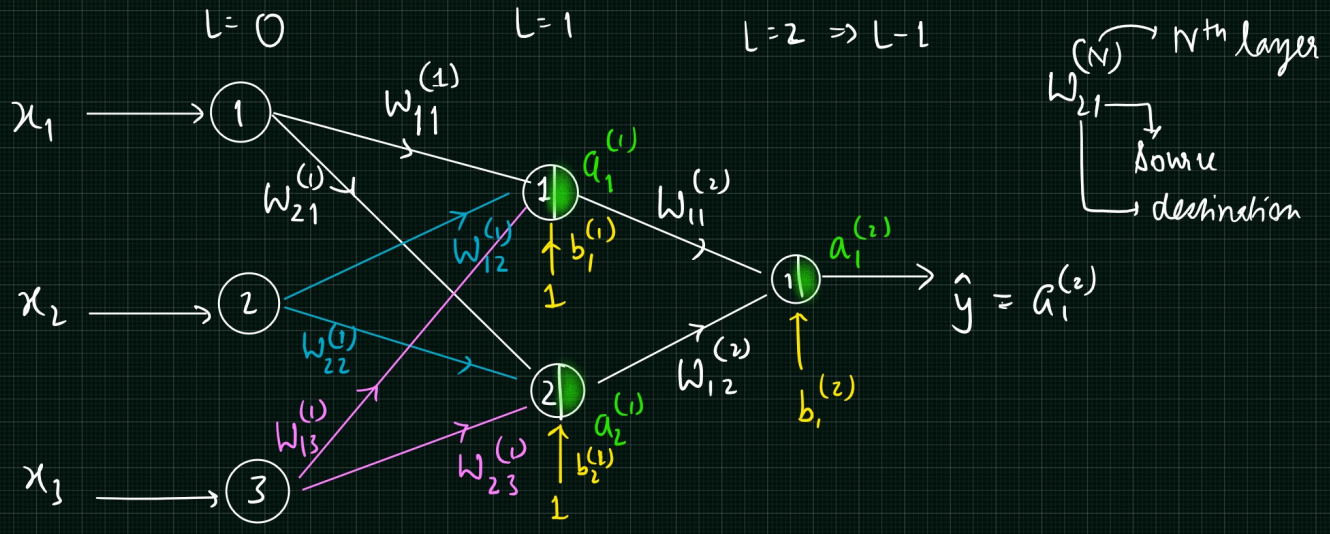
Forward

PROPAGATION



at layer 1,

① $$\begin{cases} z_1^{(1)} = W_{11}^{(1)} x_1 + W_{12}^{(1)} x_2 + W_{13}^{(1)} x_3 + b_1^{(1)} \\ a_1^{(1)} = \sigma(z_1^{(1)}) \end{cases}$$

② $$\begin{cases} z_2^{(1)} = W_{21}^{(1)} x_1 + W_{22}^{(2)} x_2 + W_{23}^{(1)} x_3 + b_2^{(1)} \\ a_2^{(1)} = \sigma(z_2^{(1)}) \end{cases}$$

final layer { o/p layer }

③ $$\begin{cases} z_1^{(2)} = W_{11}^{(2)} a_1^{(1)} + W_{12}^{(2)} a_2^{(1)} + b_1^{(2)} \\ a_1^{(2)} = \sigma(z_1^{(2)}) \longrightarrow \hat{y} \end{cases}$$

$$\text{error} = f(\hat{y}, y) = C(\hat{y}, y)$$

Weight update rule (Generic)

$$W = W + \Delta W$$

$$\Delta W = -\eta \frac{\partial C}{\partial W}$$

$\frac{\partial C}{\partial W}$ → Change in C w.r.t small change in W

from ① & ② for layer 1.

$$\begin{bmatrix} w_{11}^{(1)} & w_{12}^{(1)} & w_{13}^{(1)} \\ w_{21}^{(1)} & w_{22}^{(1)} & w_{23}^{(1)} \end{bmatrix}_{2\times 3} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{3\times 1} + \begin{bmatrix} b_1^{(1)} \\ b_2^{(1)} \end{bmatrix}_{2\times 1} = \begin{bmatrix} z_1^{(1)} \\ z_2^{(1)} \end{bmatrix}_{2\times 1}$$

$$\xrightarrow{\text{act}} \begin{bmatrix} \sigma(z_1^{(1)}) \\ \sigma(z_2^{(1)}) \end{bmatrix}_{2\times 1} = \begin{bmatrix} a_1^{(1)} \\ a_2^{(1)} \end{bmatrix}_{2\times 1}$$

at final layer

$$\begin{bmatrix} w_{11}^{(2)} & w_{12}^{(2)} \end{bmatrix}_{1\times 2} \begin{bmatrix} a_1^{(1)} \\ a_2^{(1)} \end{bmatrix}_{2\times 1} + \begin{bmatrix} b_1^{(2)} \end{bmatrix}_{1\times 1} = \begin{bmatrix} z_1^{(2)} \end{bmatrix}_{1\times 1} \xrightarrow{\text{act}^{n}} \begin{bmatrix} \sigma(z_1) \end{bmatrix}$$

$$\hat{y} \longleftarrow \begin{bmatrix} a_1^{(2)} \end{bmatrix}$$