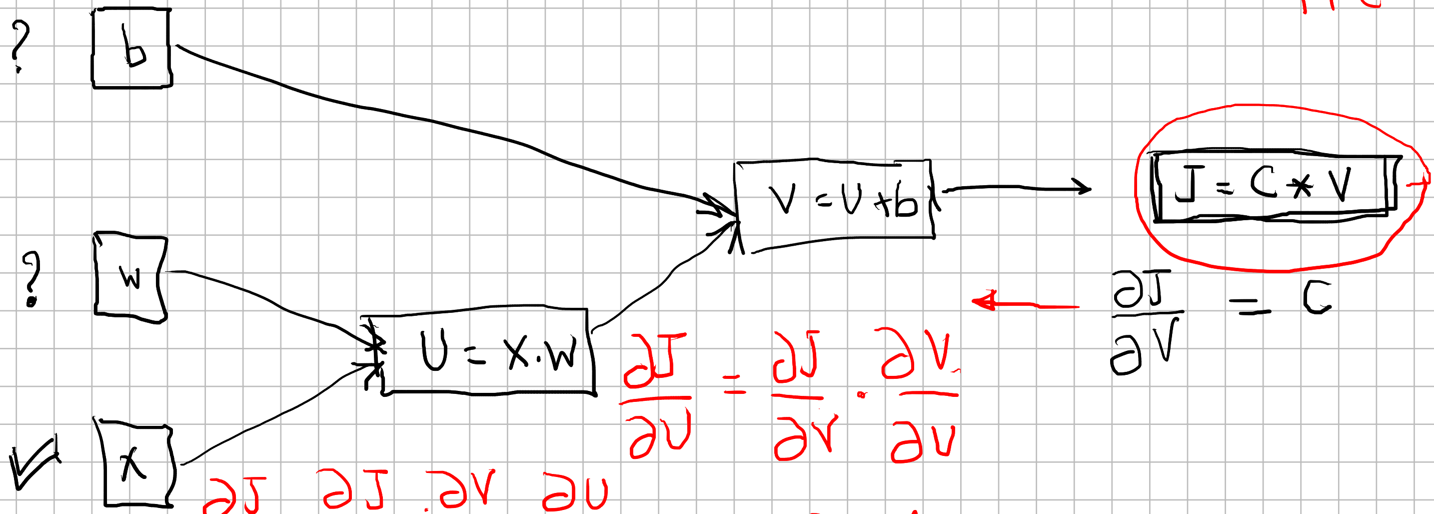
$\hat{y} = x.w + b \Rightarrow$ Cost = $J = C \times \hat{y}$

$$\sigma = \frac{1}{1+e^{-z}}$$

$$J = \mathcal{L}(a, y)$$

$$= -\left[y \cdot \log a + (1-y) \cdot \log(1-a) \right]$$



$$\frac{\partial J}{\partial V} = C$$

$$\frac{\partial J}{\partial U} = \frac{\partial J}{\partial V} \cdot \frac{\partial V}{\partial U}$$

$$= C \times 1$$

$$\frac{\partial J}{\partial x} = \frac{\partial J}{\partial V} \cdot \frac{\partial V}{\partial U} \cdot \frac{\partial U}{\partial x}$$

$$= C \times W$$

x_1, w_1, x_2, w_2, b $\rightarrow$ Agg. fn: $\boxed{z = x_1 \cdot w_1 + x_2 \cdot w_2 + b}$ $\rightarrow$ $\boxed{\hat{y} = a = \sigma(z)}$ $\rightarrow$ $\boxed{\mathcal{L}(a, y)}$

$$-\left[y \cdot \log(a) + (1-y) \cdot \log(1-a)\right]$$

$$\boxed{dz = \frac{\partial L}{\partial a} \cdot \frac{\partial a}{\partial z}}$$

$$\frac{\partial a}{\partial z} = a \cdot (1-a)$$

$$\frac{\partial L}{\partial a} = da = -\left[\frac{y}{a} - \frac{1-y}{1-a}\right]$$

$$= \left[-\frac{y}{a} + \frac{1-y}{1-a}\right] \cdot a \cdot (1-a)$$

$$= -\frac{y \cdot (1-a) \cdot \cancel{a}}{\cancel{a}} + \frac{(1-y)}{\cancel{1-a}} \cdot a \cdot \cancel{(1-a)}$$

$$= -y + \cancel{ya} + a - \cancel{ya}$$

$$\boxed{dz = (a-y)}$$

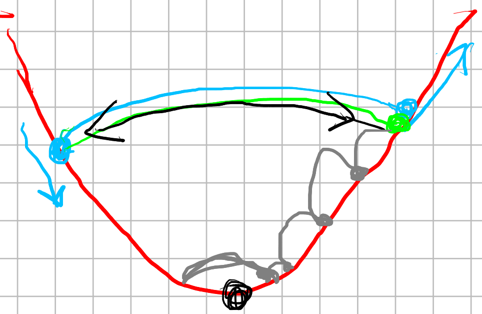
$$dw_1 = x_1 \cdot dz$$
$$dw_2 = x_2 \cdot dz$$
$$db = dz$$

$(a-y)$

$$w_1 = w_1 - \alpha * x_1 \cdot dz$$
$$w_2 = w_2 - \alpha \cdot x_2 \cdot dz$$
$$b = b - \alpha \cdot dz$$

$(a-y)^2$

$2 * (a-y)$



$$\mathrm{I} \qquad\qquad \mathrm{II}$$

$$x \rightarrow A \,|\, g \rightarrow A \,|\, \sigma \rightarrow \hat{y}$$

$w_1 \,|\, b_1$

$$\mathrm{I} \qquad\qquad \mathrm{II}$$

$$\frac{a^{[0]}}{x} \rightarrow \boxed{Z_1 = x w_1 + b_1} \rightarrow \boxed{a^{[1]} = g[z_1]} \rightarrow \boxed{Z_2 = a^{[1]} \cdot W_2 + b_2} \rightarrow \boxed{a^{[2]} = \sigma(z_2)} \rightarrow [(a^{[2]}, y)]$$

$$\frac{\partial L}{\partial z_1} = dz_1 = \underbrace{\frac{\partial L}{\partial z_2} \cdot \frac{\partial z_2}{da^{[1]}}}_{da^{[1]}} \cdot \frac{da^{[1]}}{dz_1}$$

$$\partial z_2 = a^{[2]} - y$$

$$\partial W_2 = a^{[1]} \cdot dz_2$$

$$\partial b_2 = \partial z_2$$

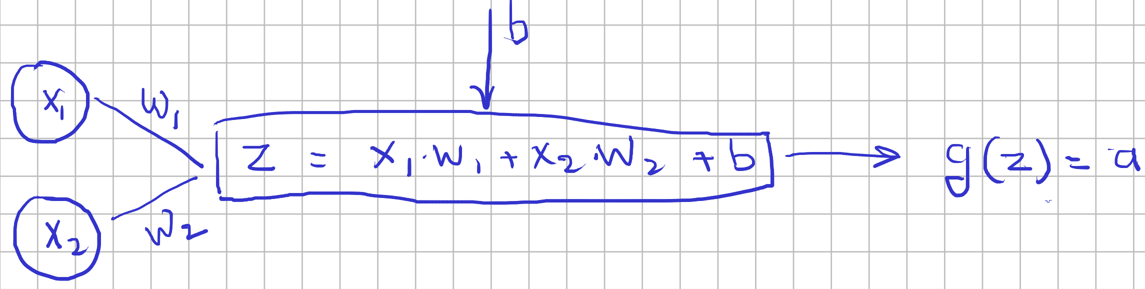
$$\boxed{\partial a^{[1]}} = W_2 \cdot dz_2$$

$$\frac{\partial a^{[1]}}{\partial z_1} = \frac{\partial\left[g(z_1)\right]}{\partial z_1}$$

$$= g'(z_1)$$

$$\partial W_1 = \underset{a^{[0]}}{x} \cdot dz_1$$

$$\partial b_1 = dz_1$$

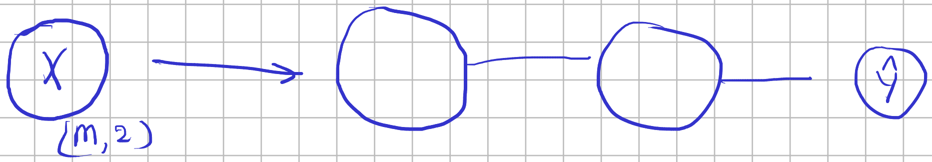


$$z = \left[x_1, x_2\right] \cdot \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} + b$$

$$X = \underset{(m,2)}{\begin{bmatrix} x_1^{(1)}, & x_2^{(1)} \\ x_1^{(2)}, & x_2^{(2)} \\ x_1^{(3)}, & x_2^{(3)} \\ \vdots & \\ x_1^{(m)}, & x_2^{(m)} \end{bmatrix}} \cdot \underset{(2,1)}{\begin{bmatrix} w_1 \\ w_2 \end{bmatrix}} + b = \underset{(m,1)}{Z} \qquad \underset{(m,1)}{a} = g(z)$$

$$\boxed{(m,1)}$$

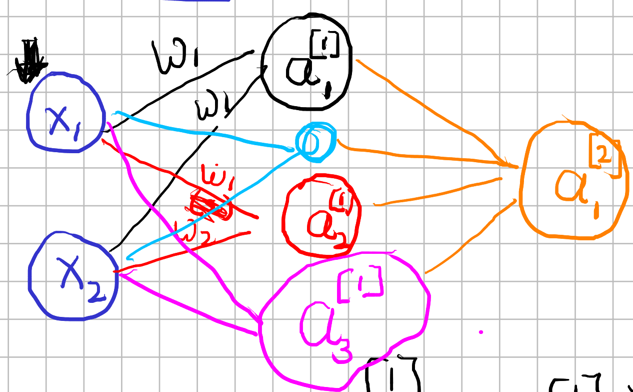
$$a = \sigma\left[X_{(m,2)} \cdot W_{(2,1)} + b_{(1,1)}\right]$$



$$a^{[1]} = g\left(X_{m,2} \cdot W^{[1]}_{(2,1)} + b^{[1]}_{(1,1)}\right)$$

$$a^{[2]} = \sigma\left(a^{[1]}_{(m,1)} \cdot W^{[2]}_{(1,1)} + b^{[2]}_{(1,1)}\right)$$

Neural Network



$$a_1^{[1]} = g\left(X \cdot W_1^{[1]} + b_1^{[1]}\right)$$

$$a_2^{[1]} = g\left(X \cdot W_2^{[1]} + b_2^{[1]}\right)$$

$$a_3^{[1]} = g\left(X \cdot W_3^{[1]} + b_3^{[1]}\right)$$

$$a^{[1]} = \left[a_1^{[1]}, a_2^{[1]}, a_3^{[1]}, a_4^{[1]} \cdots a_n^{[1]}\right]$$

$$a^{[1]} = g\left[X \cdot \left[W_1^{[1]}, W_2^{[1]}, W_3^{[1]}, W_4^{[1]} \cdots\right] + \left[b_1^{[1]}, b_2^{[1]}, b_3^{[1]}, b_4^{[1]}\right]_{(1,4)}\right]$$

$$\begin{bmatrix} w_1^{[1]}, & w_1, & w_1, & w_1 \\ w_2, & w_2, & w_2, & w_2 \end{bmatrix}$$

$(2,4)$

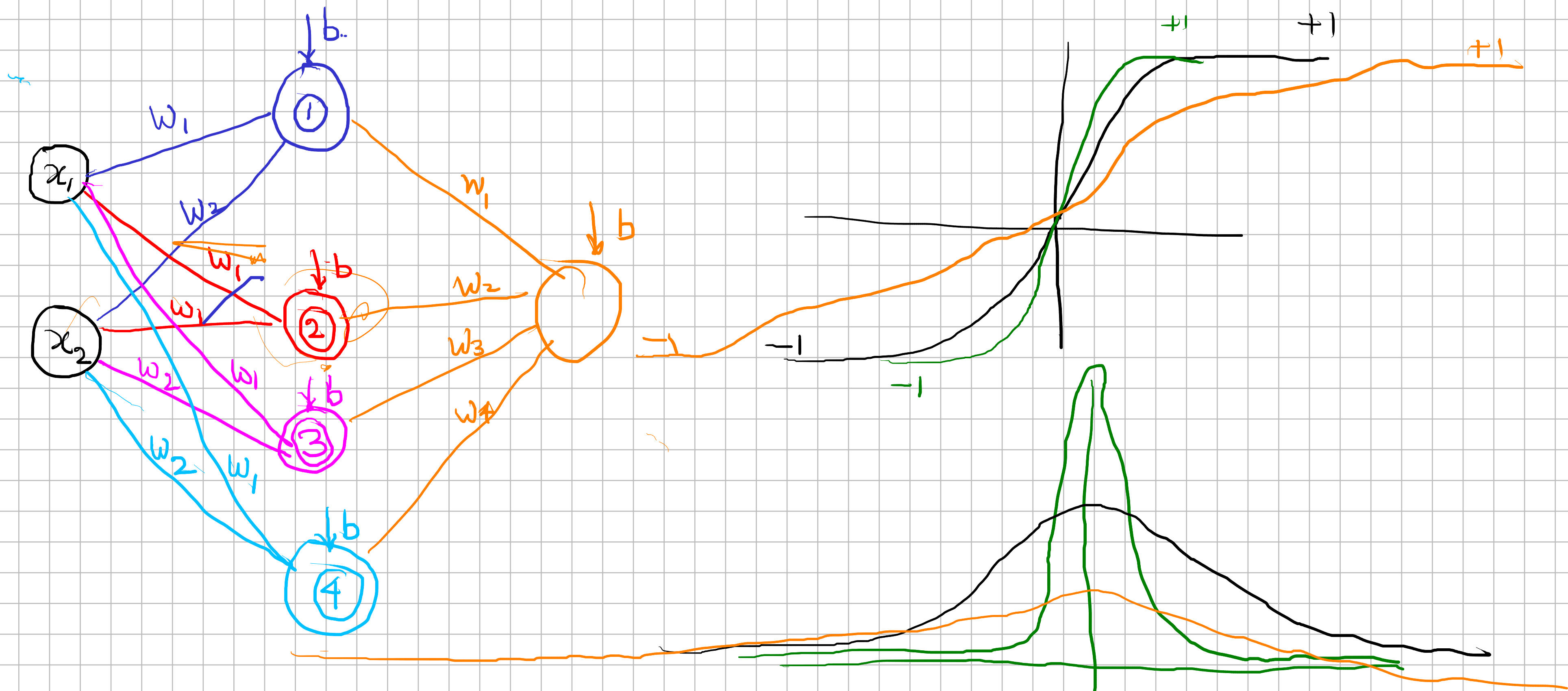
$4 \quad 4$

$$\underset{(m,4)}{a^{[1]}} = X_{m,2} \cdot W^{[1]}_{(2,4)} + b^{[1]}_{(1,4)}$$

$$a^{[2]} = \sigma\left[a^{[1]}_{(m,4)} \cdot W^{[2]}_{(4,1)} + b^{[2]}_{(1,1)}\right]$$

$$L(a^{[2]}_i, y_i)$$

$$J(\hat{y}, y) = \text{Cost} = \frac{1}{m} \sum_{i=0}^{m} L(a^{[2]}_i, y_i)$$



$$
\begin{aligned}
\text{Loss} &= f_1\left(a^{[2]}\right) \\
&= f_2\left(Z_2\right) \\
&= f_3\left(a^{[1]}\right) \\
&= f_4\left(z_1\right)
\end{aligned}
$$

$$
\frac{\partial \text{Loss}}{\partial z_1} = \frac{\partial \text{Loss}}{\partial a^{[2]}} \cdot \frac{\partial a^{[2]}}{\partial z_2} \cdot \frac{\partial z_2}{\partial a^{[1]}} \cdot \frac{\partial a^{[1]}}{\partial z_1}
$$

$$dz_2 = a_2 - y$$

$$dW_2 = \frac{1}{m} a_1 \cdot dz_2$$

$$db_2 = \frac{1}{m} dz_2 (\text{Sum})$$

$$dz_1 = dz_2 \cdot W_2 \cdot g'(z_1)$$

$$dW_1 = \frac{1}{m} \cdot x \cdot dz_1$$

$$db_1 = \frac{1}{m} \times dz_1 (\text{Sum})$$