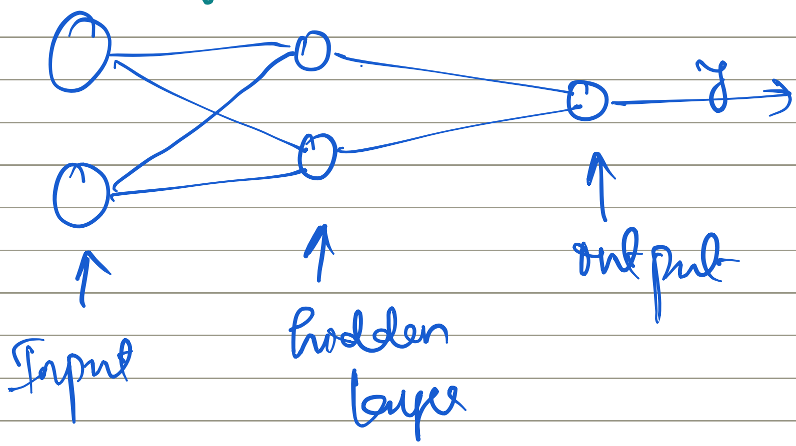
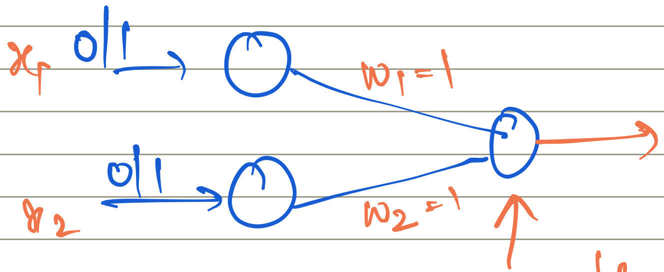
Deep Learning





$$y = \theta(1 \cdot x_1 + 1 \cdot x_2)$$

where

$$\theta(x) = \begin{cases} 1 & x \geq 1 \\ 0 & x < 1 \end{cases}$$

$$\theta(x) = \begin{cases} 1 & x \geq 2 \\ 0 & \text{otherwise} \end{cases}$$

$$y = w_1 x_1 + w_2 x_2$$

$$= \begin{bmatrix} w_1 & w_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

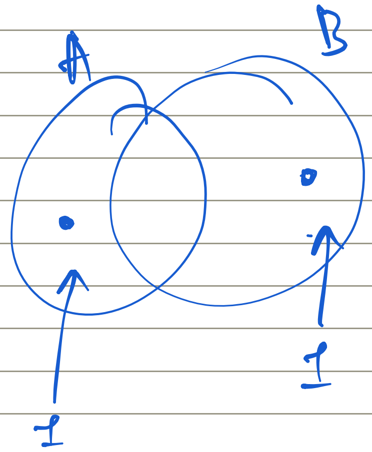
$$= W x$$

x_1	x_2	y	y
1	1	1	1
1	0	1	0
0	1	1	0
0	0	0	0
		OR	AND.

XOR

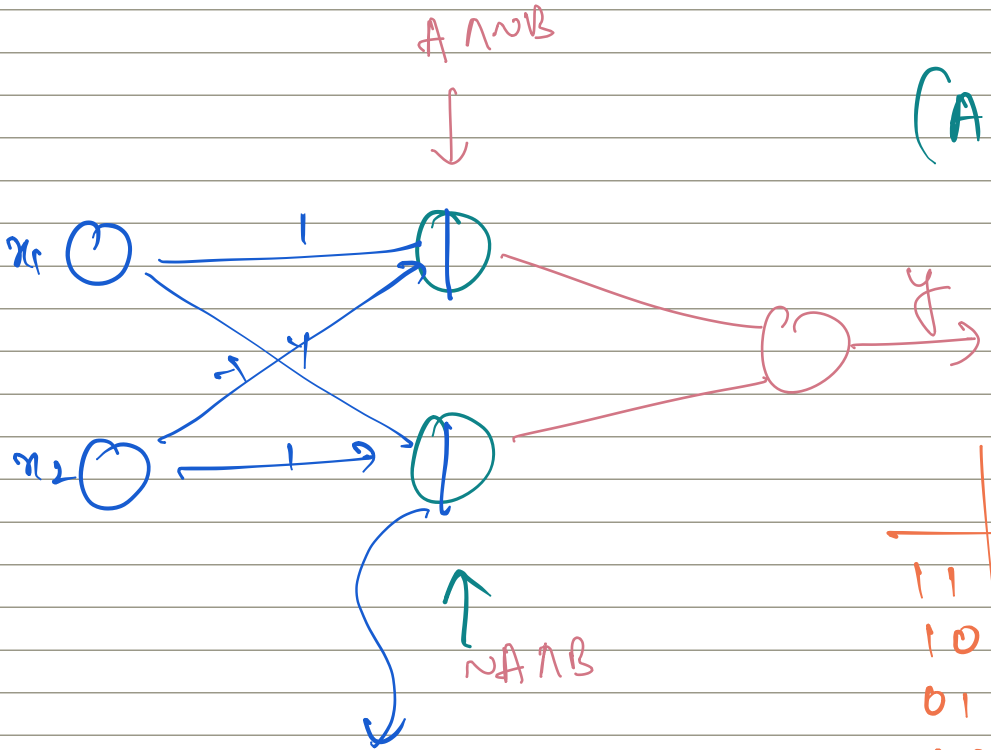
$x_1 \rightarrow 0$

$x_2 \rightarrow 0$



$(A \cap B^c) \cup (A^c \cap B)$

$(A \wedge \sim B) \vee (\sim A \wedge B)$



$\theta_1(x) = \begin{cases} 1 & x > \\ \end{cases}$

	$A \wedge \sim B$
1 1	0
1 0	1
0 1	0
0 0	0

12×13

$12(10+3)$

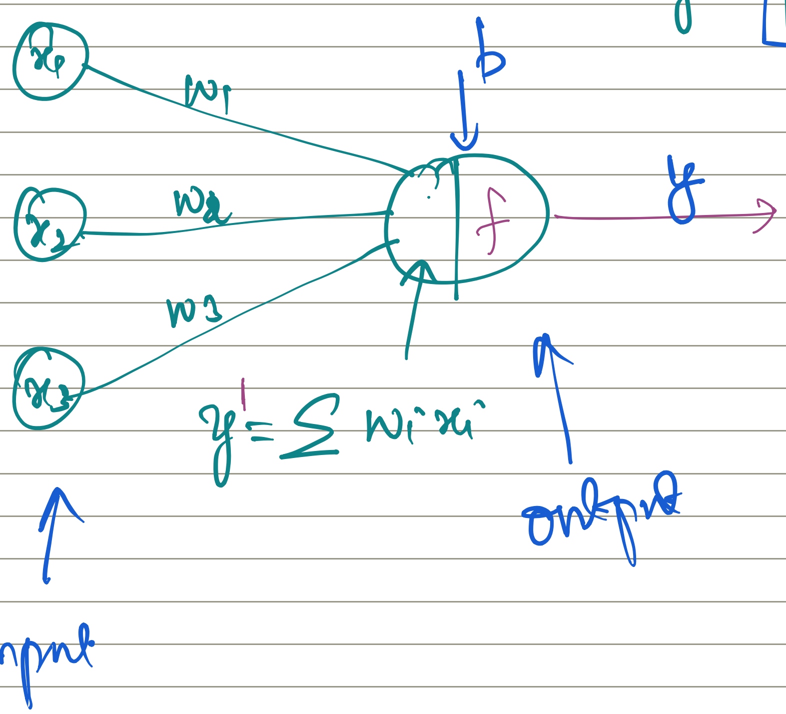
$$\begin{array}{r} 12 \\ 13 \\ \hline \\ \hline \end{array}$$

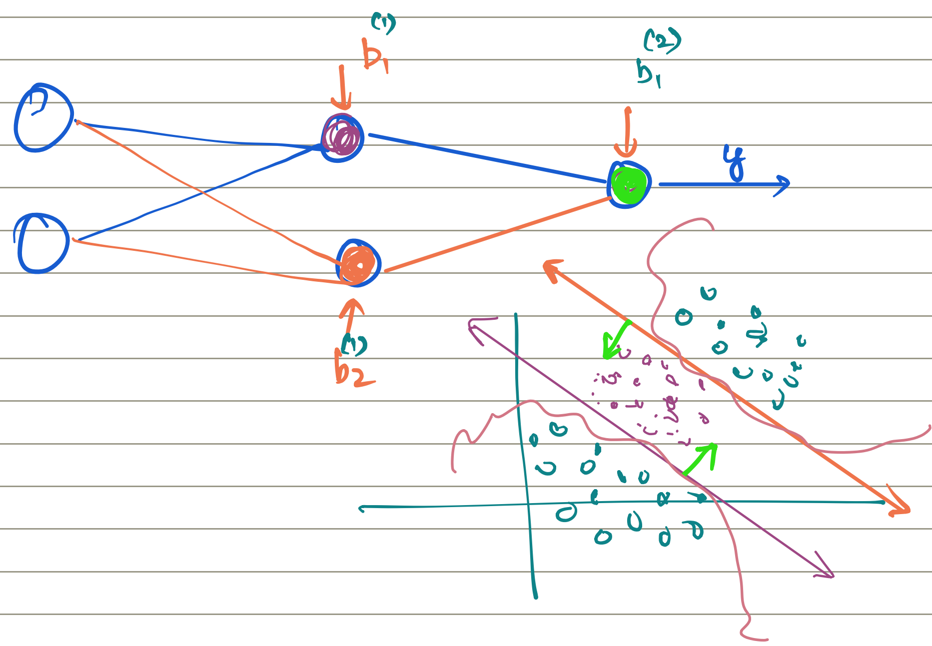
	x_1	x_2	x_3	y
$R_1 \rightarrow$				
$R_2 \rightarrow$				
⋮				
$R_N \rightarrow$				

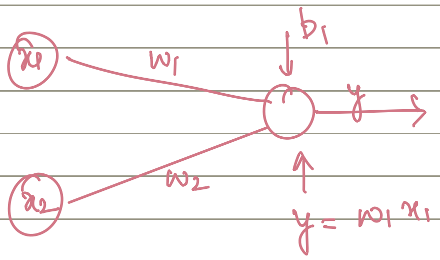
y: c/ numerical — Linear Regression.

binary — Logistic Regression

$$y = \boxed{\beta_0} + \beta_1 x_1 + \beta_2 x_2$$







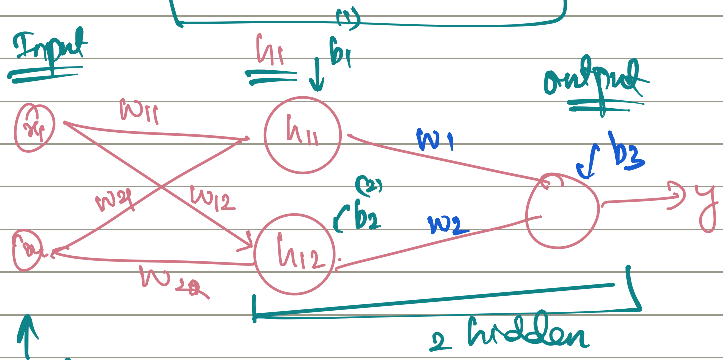
$$W = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$y = w_1 x_1 + w_2 x_2 + b_1$$

$$= [w_1 \; w_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + b_1$$

$$\boxed{y = W^T x + b_1}$$



Input layer.

$$h_{11} = [w_{11} \; w_{12}]\, x + b_1$$

$$h_{12} = [w_{21} \; w_{22}]\, x + b_2$$

$$\begin{bmatrix} h_{11} \\ h_{12} \end{bmatrix} = \underbrace{\begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix}}_{W_1^T} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \underbrace{\begin{bmatrix} b_1^{(1)} \\ b_2^{(1)} \end{bmatrix}}_{b_1}$$

$$y = w_1 h_{11} + w_2 h_{12} + \boxed{b_3}$$

b_2

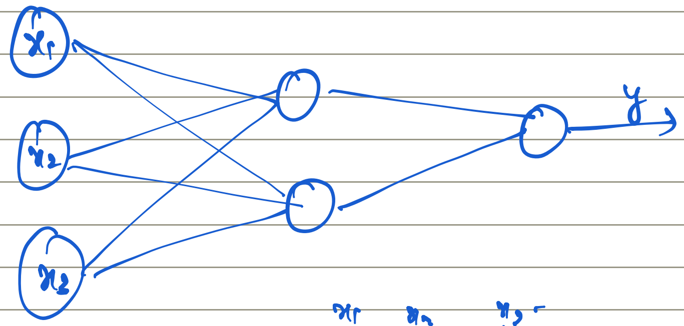
$$= \underbrace{[w_1 \; w_2]}_{W_2^T} \begin{bmatrix} h_{11} \\ h_{12} \end{bmatrix} + b_3$$

$$= [w_1 \; w_2] \left(W_1^T x + b_1 \right) + b_2$$

$$\boxed{y = W_2^T \left(W_1^T x + b_1 \right) + b_2}$$

known known

Eg:



$$
x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}
\qquad
W_0 = \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \\ \downarrow & \downarrow & \downarrow \end{array} \\ \begin{bmatrix} 1 & 3 & 1 \\ 2 & -5 & 3 \end{bmatrix}_{2\times 3} \end{array}
$$

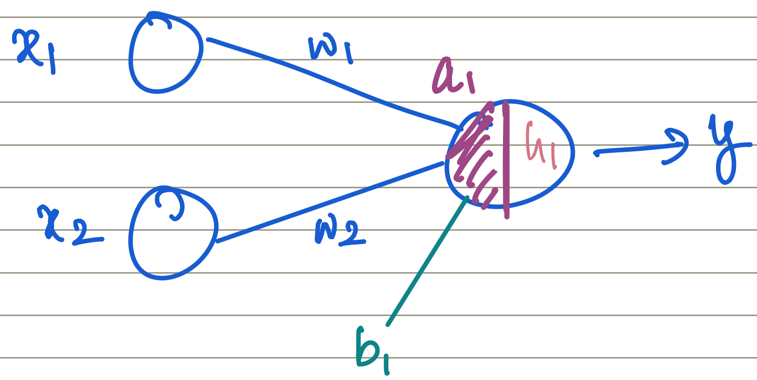
$$
b^{(1)} = \begin{bmatrix} 1/2 \\ 0 \end{bmatrix}
$$

$$
w_1 = \begin{bmatrix} 2 & 2 \end{bmatrix}_{1\times 2}
$$

$$
b^{(2)} = \begin{bmatrix} 5 \end{bmatrix}_{1\times 1}
$$

Q

Compute y. (assume the linear perceptron).



$$a_1 = w_1 x_1 + w_2 x_2 + b_1$$

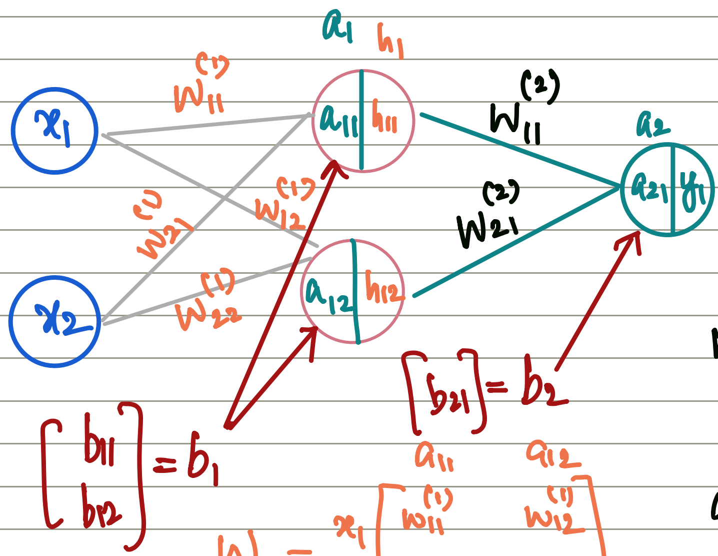
$$h_1 = g(a_1)$$

$$g(x) = x$$

$$g(x) = \frac{1}{1+e^{-x}}$$

$$g(x) = \begin{cases} 0 & x < 0 \\ x & x \geq 0 \end{cases}$$

Rectified Linear Unit.
"Relu"



$$W_2 = \begin{bmatrix} w_{11} \\ w_{21} \end{bmatrix}$$

$$\begin{aligned} a_2 &= W_2^T h_1 + b_2 \\ &= \underset{1\times 2}{\begin{bmatrix} w_{11} & w_{21} \end{bmatrix}} \underset{2\times 1}{\begin{bmatrix} h_{11} \\ h_{12} \end{bmatrix}} + b_{21} \end{aligned}$$

$$W_1 = \begin{array}{c} \\ x_1 \\ x_2 \end{array} \begin{array}{c} \begin{array}{cc} a_{11} & a_{12} \end{array} \\ \begin{bmatrix} w_{11}^{(1)} & w_{12}^{(1)} \\ w_{21}^{(1)} & w_{22}^{(1)} \end{bmatrix} \end{array}$$

$$a_{11} = W_{11} x_1 + W_{21} x_2 \quad = w_{11} h_{11} + w_{21} h_{12}$$

$$a_{12} = W_{12} x_1 + W_{22} x_2$$

$$\begin{bmatrix} a_{11} \\ a_{12} \end{bmatrix} = \begin{bmatrix} W_{11} & W_{21} \\ W_{12} & W_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + b_1$$

$$= W_1^T X + b_1$$

we shall follow this case.

$$\boxed{\begin{bmatrix} a_{11} & a_{12} \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix}}$$

$$h_1 = g(a_1)$$

$$= g\left(w_1^T x + b_1 \right)$$

Let $\boxed{y = 1}$ Assump:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad W_1 = \begin{bmatrix} 0.5 & 0.25 \\ 0.3 & 0.4 \end{bmatrix} \qquad b_1 = \begin{bmatrix} 0 \\ 0.1 \end{bmatrix}$$

Assume $g(x) = \frac{1}{1+e^{-x}}$ $\qquad W_2 = \begin{bmatrix} 0.5 \\ 0.25 \end{bmatrix}$

Compute a_1:

$$b_2 = [0.4]$$

$$a_1 = \begin{bmatrix} a_{11} \\ a_{12} \end{bmatrix} = W^T x + b_1$$

$$= \underbrace{\begin{bmatrix} 0.5 & 0.25 \\ 0.3 & 0.4 \end{bmatrix}^T}_{2\times 2} \underbrace{\begin{bmatrix} 1 \\ 2 \end{bmatrix}}_{1\times 2} + \begin{bmatrix} 0 \\ 0.1 \end{bmatrix}$$

$$\boxed{a_1 = \begin{bmatrix} 1.1 \\ 1.15 \end{bmatrix}}$$

$$g(a_1) = \begin{bmatrix} g(1.1) \\ g(1.15) \end{bmatrix} = \begin{bmatrix} \frac{1}{1+e^{-1.1}} \\ \frac{1}{1+e^{-1.15}} \end{bmatrix}$$

$$\begin{bmatrix} h_{11} \\ h_{12} \end{bmatrix} = \begin{bmatrix} 0.75 \\ 0.76 \end{bmatrix}$$

$$a_2 = W_2^T h_1 + b_2$$

$$= \begin{bmatrix} 0.5 & 0.25 \end{bmatrix} \begin{bmatrix} 0.75 \\ 0.76 \end{bmatrix} + [0.4]$$

$$= 0.965$$

$$g_2(x) = \begin{cases} 0 & x \le 0 \\ x & x > 0 \end{cases}$$

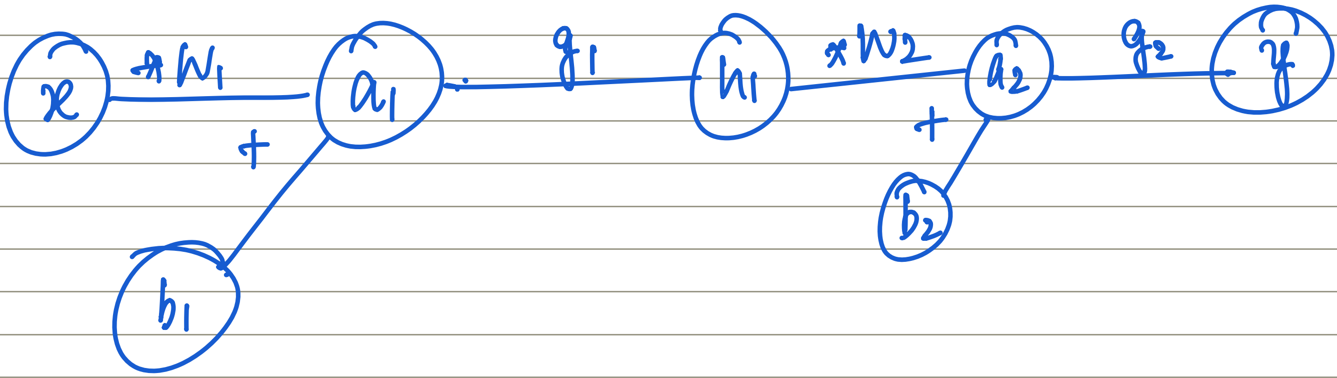
$$\hat{y} = g_2(a_2)$$

$$= 0.965$$

What is the loss? (Based on absolute error)

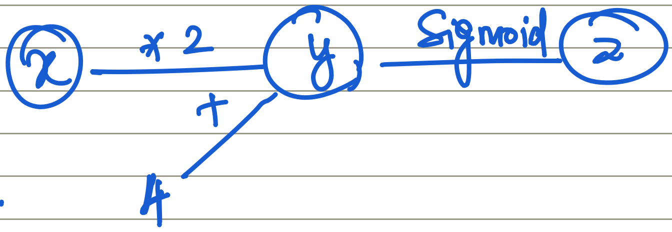
$$|1 - 0.965| = \boxed{0.035}$$

Computational Graph



Computing Partial Derivatives.

Prob: 01. Consider the Computational graph



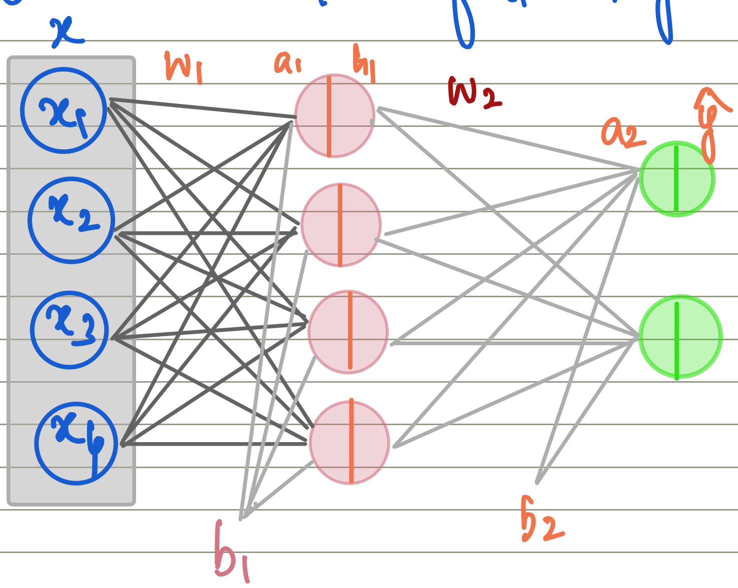
find $\frac{\partial z}{\partial x}$.

Hint:

$$z = \text{sigmoid}(y) = \frac{1}{1+e^{-y}}$$

$$y = 2x + 4.$$

Prob: Consider the following network.



Use the following initial values to compute a forward pass. (that is $\hat{y}$)

Input	True output	Loss Function
$x = [2 \;\; 5 \;\; 3 \;\; 3]$	$y = [1 \;\; 0]$	$L[y, \hat{y}] = \sum_i (y_i - \hat{y}_i)^2$

$$W_1 = \begin{array}{c} \\ a_{11} \\ a_{12} \\ a_{13} \\ a_{14} \end{array} \begin{array}{c} \begin{array}{cccc} x_1 & x_2 & x_3 & x_4 \end{array} \\ \begin{bmatrix} 0.1 & -0.3 & -0.3 & 0.2 \\ 0.3 & -0.2 & 0 & 0.5 \\ 0.8 & 0.5 & 0.5 & -0.9 \\ -0.4 & 0.5 & 0.4 & 0.7 \end{bmatrix} \end{array} \qquad b_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$g_1(a_1) \leftarrow \text{Sigmoid}(a_1)$$

$$W_2 = \begin{array}{c} \\ a_{21} \\ a_{22} \\ a_{23} \\ a_{24} \end{array} \begin{array}{c} \begin{array}{cc} h_{11} & h_{12} \end{array} \\ \begin{bmatrix} 0.5 & 0.5 \\ 0.8 & 0.2 \\ 0.2 & 0.3 \\ 0.4 & -0.5 \end{bmatrix} \end{array} \qquad b_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad g_2(a_2) \leftarrow \text{Softmax}(a_2)$$

Prob: update the weight $w_{21}^{(2)}$

Prob: update the weight $w_{12}^{(1)}$