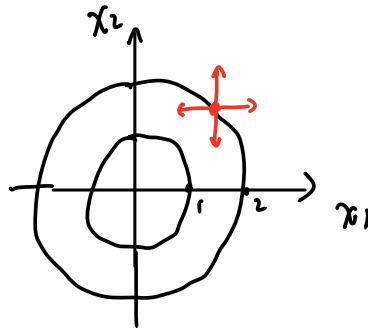


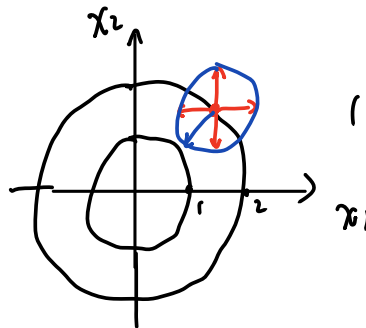
1. Goal : $\min_x f(x)$

Eg $f(x) = x_1^2 + x_2^2$



Contour line

x_1	x_2	f	
2	2	8	
1	2	5	★
3	2	13	
2	1	5	★
2	3	13	



$$\left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$$

$$\Rightarrow (-1, -1)$$

$$f'(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} \Rightarrow (1, 1)$$

$$x = x - \alpha f'(x) \Rightarrow (-\alpha, -\alpha)$$

2. $f(x)$

$$J(w, b | X, y)$$

Logistic regression

$$\hat{y} = \sigma(w_1 x_1 + w_2 x_2 + b)$$

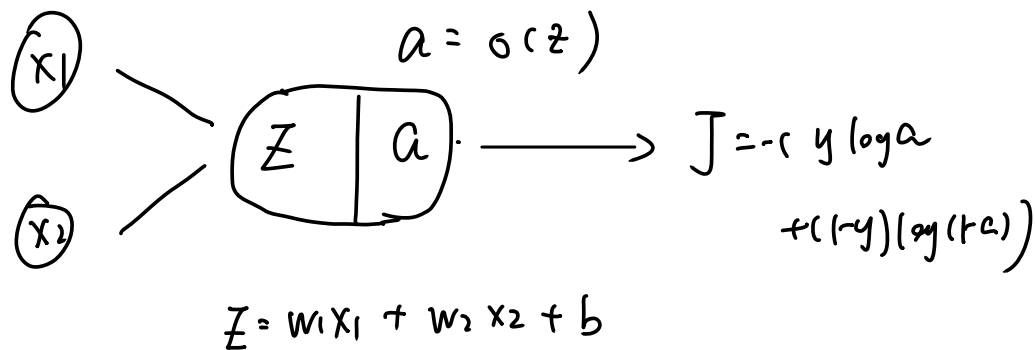
$$\text{cost} = J = - \sum y \log \hat{y} + (1 - y) \log (1 - \hat{y})$$

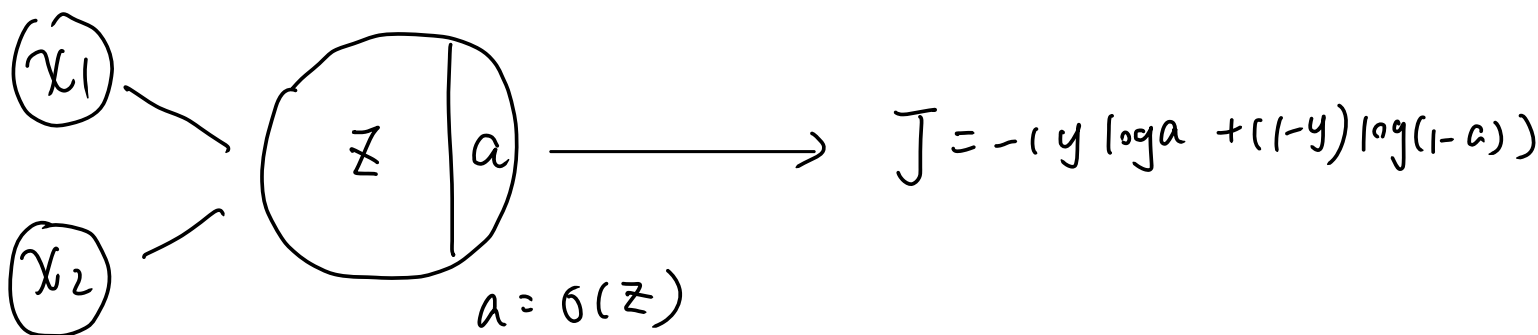
$$\text{Goal: } \min_{w, b} J(w, b | X, y)$$

$$\Rightarrow w_1 = w_1 - \alpha \cdot \frac{\partial J}{\partial w_1}$$

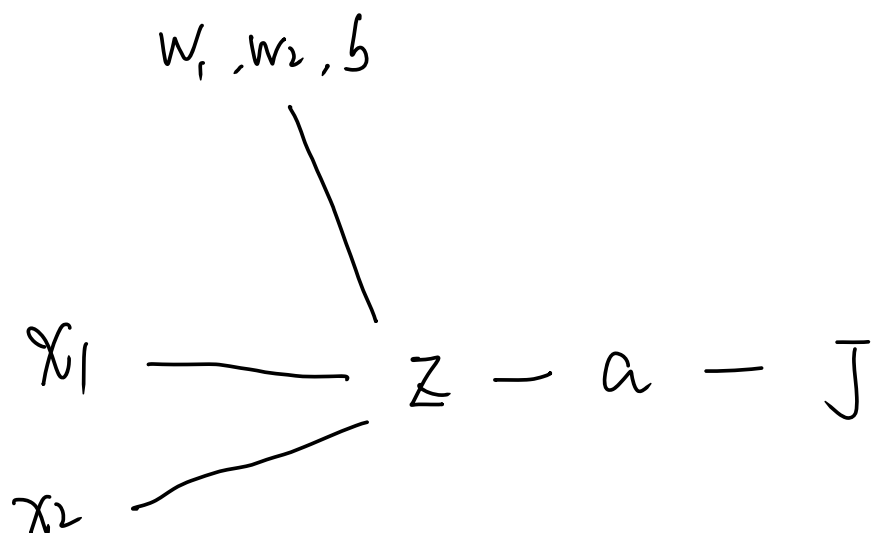
$$w_2 = w_2 - \alpha \cdot \frac{\partial J}{\partial w_2}$$

$$b = b - \alpha \cdot \frac{\partial J}{\partial b}$$





$$z = w_1 x_1 + w_2 x_2 + b$$



$$\frac{\partial J}{\partial w_1} = \frac{\partial J}{\partial a} \frac{\partial a}{\partial z} \frac{\partial z}{\partial w_1}$$

$$\frac{\partial J}{\partial a} = -\frac{y}{a} + \frac{1-y}{1-a}$$

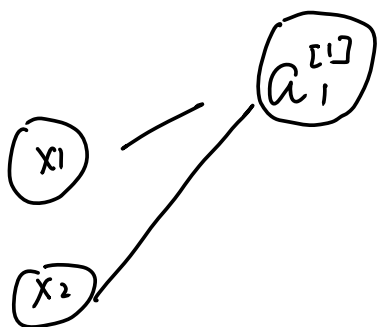
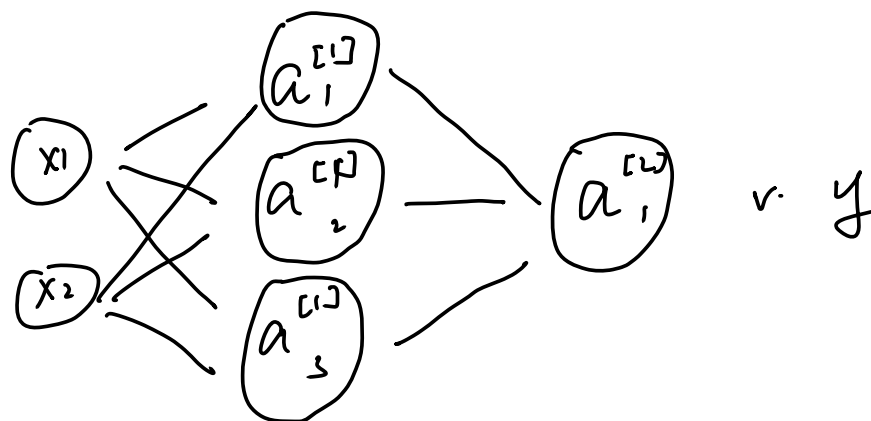
$$\frac{\partial a}{\partial z} = a(1-a)$$

$$\frac{\partial z}{\partial w_1} = x_1$$

$$\left. \begin{array}{l} \frac{\partial J}{\partial a} = -\frac{y}{a} + \frac{1-y}{1-a} \\ \frac{\partial a}{\partial z} = a(1-a) \end{array} \right] \frac{\partial J}{\partial z} = -(1-a)y + a(1-y) = a - y$$

$$\left\{ \begin{array}{l} \frac{\partial J}{\partial w_1} = (a - y) x_1 \\ \frac{\partial J}{\partial w_2} = (a - y) x_2 \\ \frac{\partial J}{\partial b} = (a - y) \end{array} \right\} \frac{\partial J}{\partial w} = (a - y) x$$

3. Shallow Neural Net



Logistic Regression

$$z = w^T x + b$$

$$a = g(z) \leftarrow \text{ReLU, tanh}$$

PCA · x_1, x_2, x_3, x_4

$\Rightarrow y_1, y_2, y_3, y_4$

$$y_4 = w^T x$$

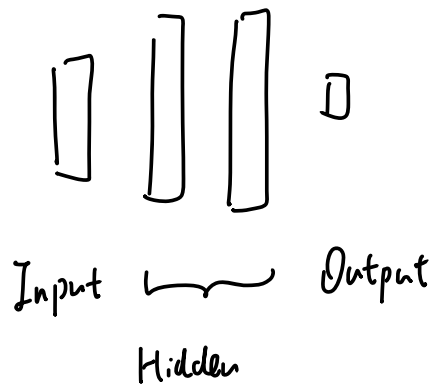
$$\sigma(y_4)$$

w, b

$$w^{[1]} \quad 2 \times 3 \quad \begin{bmatrix} w_{1,1} & w_{1,2} \\ w_{2,1} & w_{2,2} \\ w_{3,1} & w_{3,2} \end{bmatrix} \quad \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

$x_1 \quad a_1$
 $\quad \quad a_2$
 $x_2 \quad a_3$

$$w^{[2]} \quad \begin{matrix} a_1^{[1]} \\ a_2^{[1]} \\ a_3^{[1]} \end{matrix} \quad \begin{matrix} \diagup \\ \diagdown \end{matrix} \quad a_1^{[2]} \quad \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \quad 3 \times 1 \quad [b]$$



x

$$a^{[0]} \quad a^{[1]} \quad a^{[2]} \quad a^{[3]}$$

$$z^{[1]} = W^{[1]T} a^{[0]} + b^{[1]}$$

$$a^{[1]} = \sigma(z^{[1]})$$

$$W^{[1]} : \frac{15}{2} \# \text{ nodes} \times \frac{4}{2} \# \text{ nodes}$$

$$a^{[0]} : \frac{15}{2} \# \text{ nodes} \times m$$

$$z^{[1]}, b^{[1]} : \frac{4}{2} \# \text{ nodes} \times m$$