Oct. 3, 2022, Mon

$$\begin{array}{ccc} x_1 & y_1 & \hat{y}_1 \\ \vdots & \vdots & \vdots \\ x_N & y_N & \hat{y}_N \end{array}$$

ground truth label   prediction

goal: $\hat{y}_i$ the same as y_i

$\hat{y}_i = f(x_i)$

linear regression $\quad \hat{y}_i = w^T x_i$

logistic regression $\quad \hat{y}_i = \sigma(w^T x_i)$

loss function

linear regression $J(w) = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2$ squared loss

logistic regression

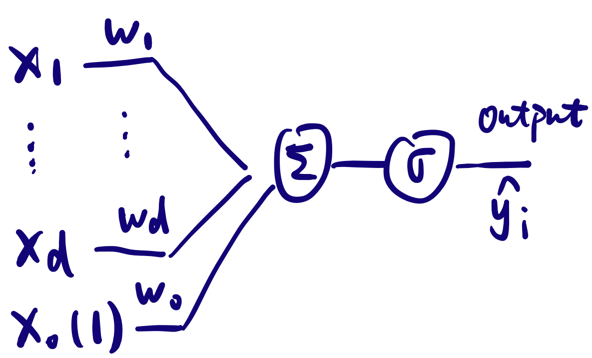
$$J(w) = \sum_{i=1}^{N} -y_i \log \hat{y}_i - (1-y_i)\log(1-\hat{y}_i)$$ cross entropy loss

$$J(w) = \sum_{i=1}^{N} |y_i - \hat{y}_i| + \frac{1}{2}\lambda w^T w$$

Gradient Descent

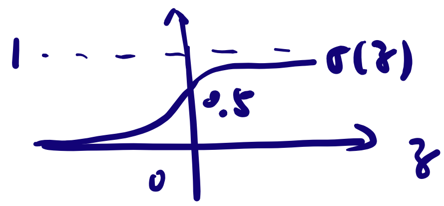
$\min J(w)$

$$w_k = w_{k-1} + \eta \frac{\partial J(w)}{\partial w}$$



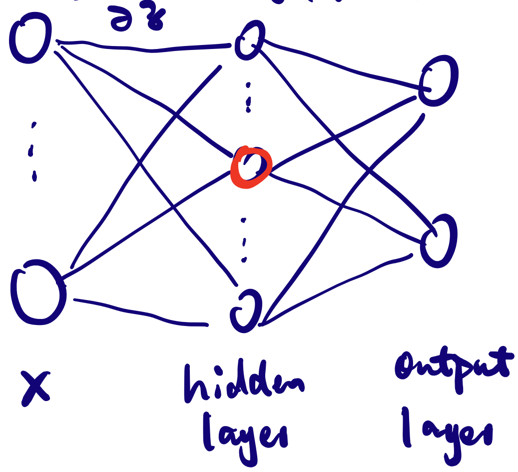
activation function

$\boxed{\text{sigmoid}}$ function



$$\sigma(z) = \frac{1}{1+\exp(-z)}$$

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

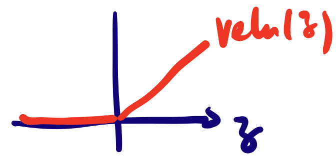


$$h = w_h^T x \qquad o = w_o^T h = w_o^T(w_h^T x) = w^T x$$

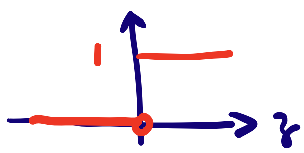
$$h = \sigma(w_h^T x) \qquad o = \sigma(w_o^T h)$$

non-linear activation function

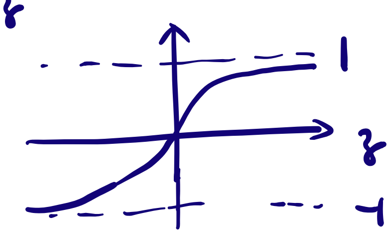
$\boxed{\text{relu}}$ $\quad \text{relu}(z) = \max(0, z)$



$$\frac{\partial\, \text{relu}(z)}{\partial z}$$

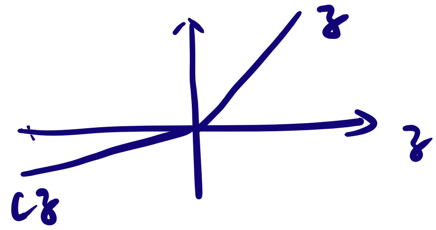


$\boxed{\tanh}$ $\quad \tanh(z) = \dfrac{e^z - e^{-z}}{e^z + e^{-z}}$

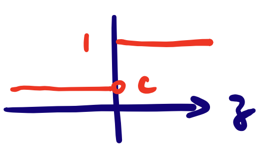


$$\text{leakyrelu}(z) = \max(cz, z)$$

c : small number



$$\frac{\partial\, \text{leakyrelu}(z)}{\partial z}$$



$$J(w) = \sum_{i=1}^{N} J_i = \frac{1}{2} \sum_{i=1}^{N} \sum_{l=1}^{k} (y_{il} - o_{il})^2$$

$$o_l = \sigma(w_l^T h) \qquad z = w_l^T h$$

$$\frac{\partial J}{\partial w_{lj}} = \frac{\partial J}{\partial o_l} \cdot \frac{\partial o_l}{\partial w_{lj}}$$

$$= \sum_{i=1}^{N} \underbrace{(y_{il} - o_{il})}_{\frac{\partial J}{\partial o_{il}}} \cdot \underbrace{o_{il}(1 - o_{il})}_{\frac{\partial \sigma}{\partial z}} \cdot \underbrace{h_j}_{\frac{\partial z}{\partial w_{lj}}}$$

Chain Rule

$$f(x, y, z) = (x + y) \cdot z = \underline{q \cdot z} \qquad q = x + y \qquad \frac{\partial q}{\partial x} = 1$$

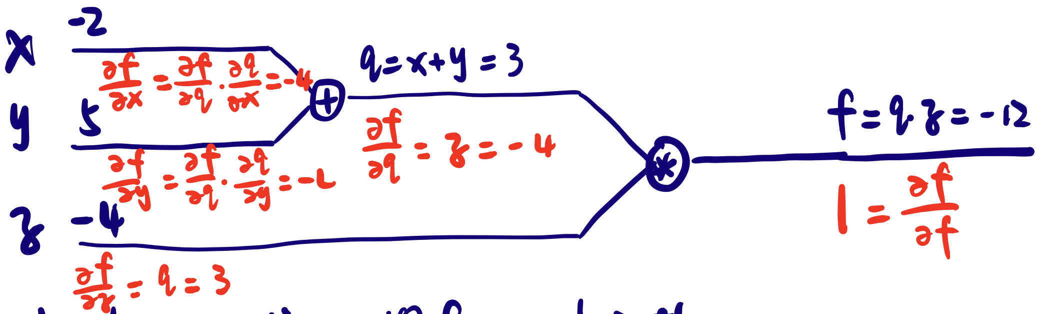
e.g. $\quad x = -2, \quad y = 5, \quad z = -4 \qquad \frac{\partial q}{\partial y} = 1$

$$f(-2, 5, -4) = (-2 + 5) \cdot (-4) = -12 \qquad \frac{\partial f}{\partial q} = z$$

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial q} \cdot \frac{\partial q}{\partial x} + \frac{\partial f}{\partial z} \cdot \frac{\partial z}{\partial x} = z \cdot 1 + q \cdot 0 = z \qquad \frac{\partial f}{\partial z} = q$$

$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial q} \cdot \frac{\partial q}{\partial y} + 0 = z$$

$$\frac{\partial f}{\partial z} = q = x + y$$



backpropogation : ① forward path

② backward path