Weight, Bias

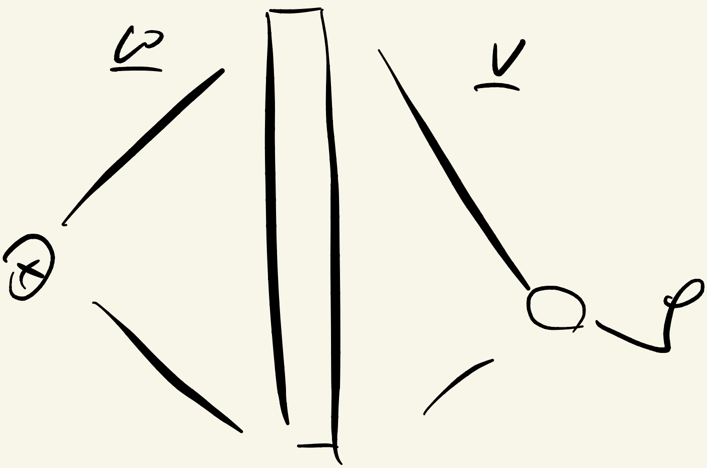
$$\begin{bmatrix} | & | & & | \\ \underline{w}x_1 & \underline{w}x_2 & \cdots & \underline{w}x_k \\ | & | & & | \end{bmatrix} + [\underline{b}, \underline{b}, \ldots]$$

$$\Downarrow$$

$$\sigma\left(\begin{bmatrix} | & | & & | \\ \underline{w}x_1 + \underline{b}, & \underline{w}x_2 + b, & \cdots, & \underline{w}x_k + b \\ | & | & & | \end{bmatrix}\right)$$

$$\Downarrow$$

$$\Rightarrow V \begin{bmatrix} | & & | \\ \sigma(\underline{w}x_1 + b), & \cdots, & \sigma(\underline{w}x_k + b) \\ | & & | \end{bmatrix}$$

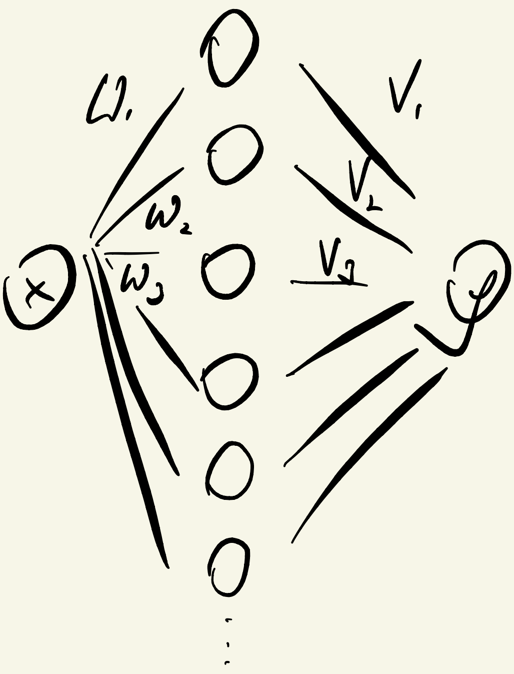


$$[f_1, f_2, \dots, f_k] = V^T \begin{bmatrix} | & | & & | \\ \sigma(\underline{w}x_1 + b) & \sigma(\underline{w}x_2 + b) & \dots & \sigma(\underline{w}x_k + b) \\ | & | & & | \end{bmatrix}$$

$$y^T = V^T \sigma(\underline{W} x^T + \underline{b})$$

$$x \longrightarrow \sigma(Wx + b)$$

$$\begin{cases} \sigma_{ReLu} = \max(x, 0) \\ \sigma_{sigmoid} = \dfrac{1}{1+e^{-x}} \end{cases}$$



$$\sigma\left(\begin{bmatrix} w_1 x \\ w_2 x \\ \vdots \\ w_n x \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \right)$$

$$\sigma(\underline{v}) = \begin{bmatrix} \sigma(v_1) \\ \sigma(v_2) \\ \vdots \\ \sigma(v_n) \end{bmatrix}$$

$$a = \sigma(\underline{w}x + b)$$

$$y = v_1 a_1 + v_2 a_2 + \dots + v_n a_n$$

$$= v \cdot a$$

$$y = v^T a$$

$$y = v^T(\sigma(\underline{w}x + \underline{b}))$$

$$y_1 = v^T \sigma(\underline{w}x_1 + \underline{b})$$

$$y_2 = v^T \sigma(\underline{w}x_2 + \underline{b})$$

$$x = [1, 2, \dots, 100]$$

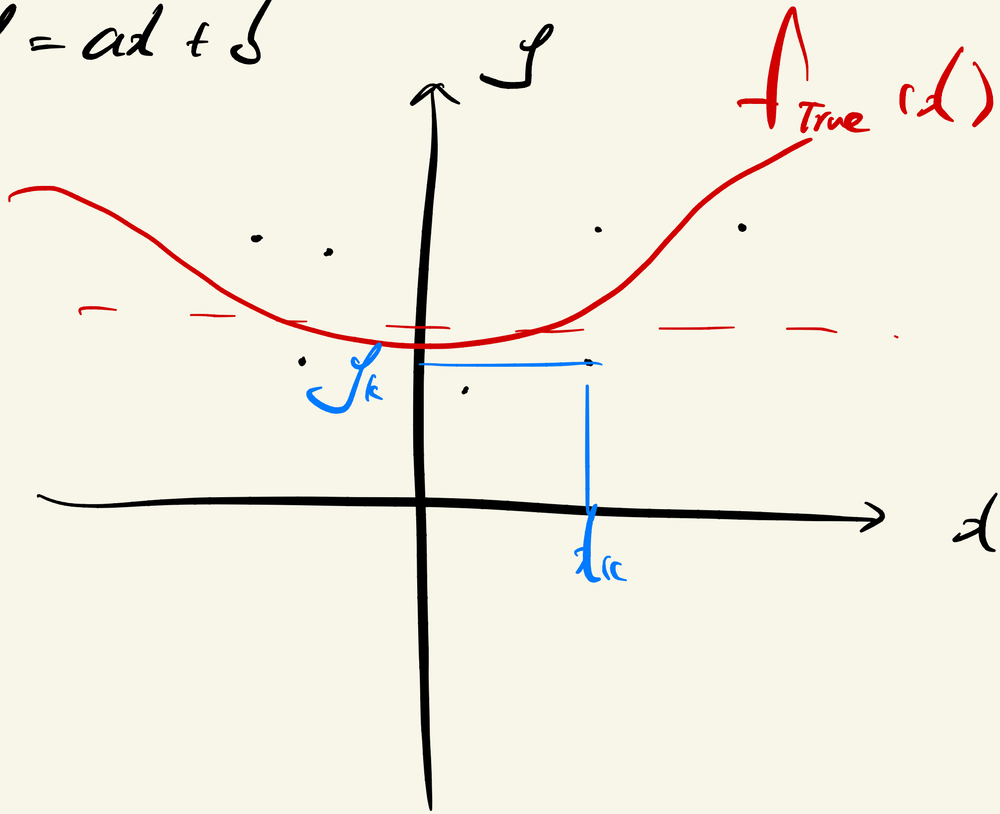
$$y = [0, \dots, 0]$$

$$y_i = \underline{v}^T(\underline{w}x_i + \underline{b})$$

$$\underline{w}\,\underline{x}^T \Rightarrow \underset{n \times 1}{\begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}} \underset{1 \times k}{\begin{bmatrix} x_1 & \cdots & x_k \end{bmatrix}}$$

Linear Regression : Line of best fit

$y = ax + b$



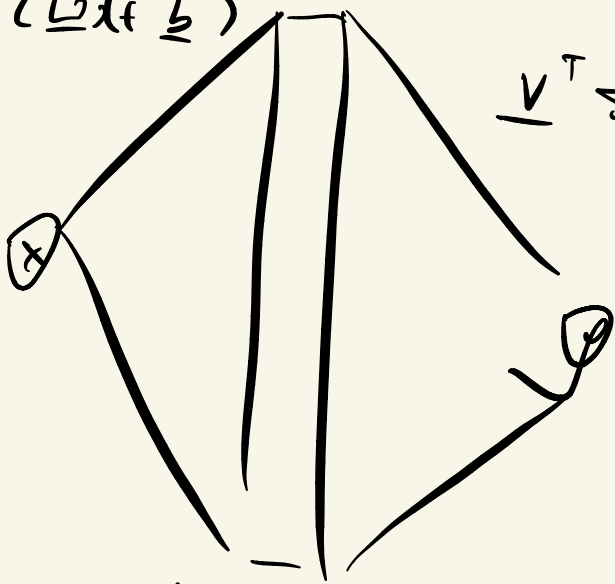
NN Regression : Best NN

$$\hat{y} = f_{NN}(x_i; \theta)$$

$$\theta = w, b, v$$

$\sigma(\underline{w}x + \underline{b})$

$\underline{v}^T \sigma(\underline{w}x + b)$



$$\min_{w, b, v} Loss = \frac{1}{N} \sum_{i=1}^{N} \left| f_{NN}(x_i; \theta) - y_i \right|^2$$ should be small