① Linear

$\vec{u} = W\vec{x} + \vec{b}^{(1)}$

$\vec{h} = ReLU(\vec{u})$

② Linear

$\vec{v} = M\vec{h} + \vec{b}^{(2)}$

$\vec{o} = softmax(\vec{v})$

③ Loss Function

$L = l(\vec{o}, \vec{y})$

↑ loss function

④ Regularization:

$s_1 = \|W\|_F^2$

$s_2 = \|M\|_F^2$

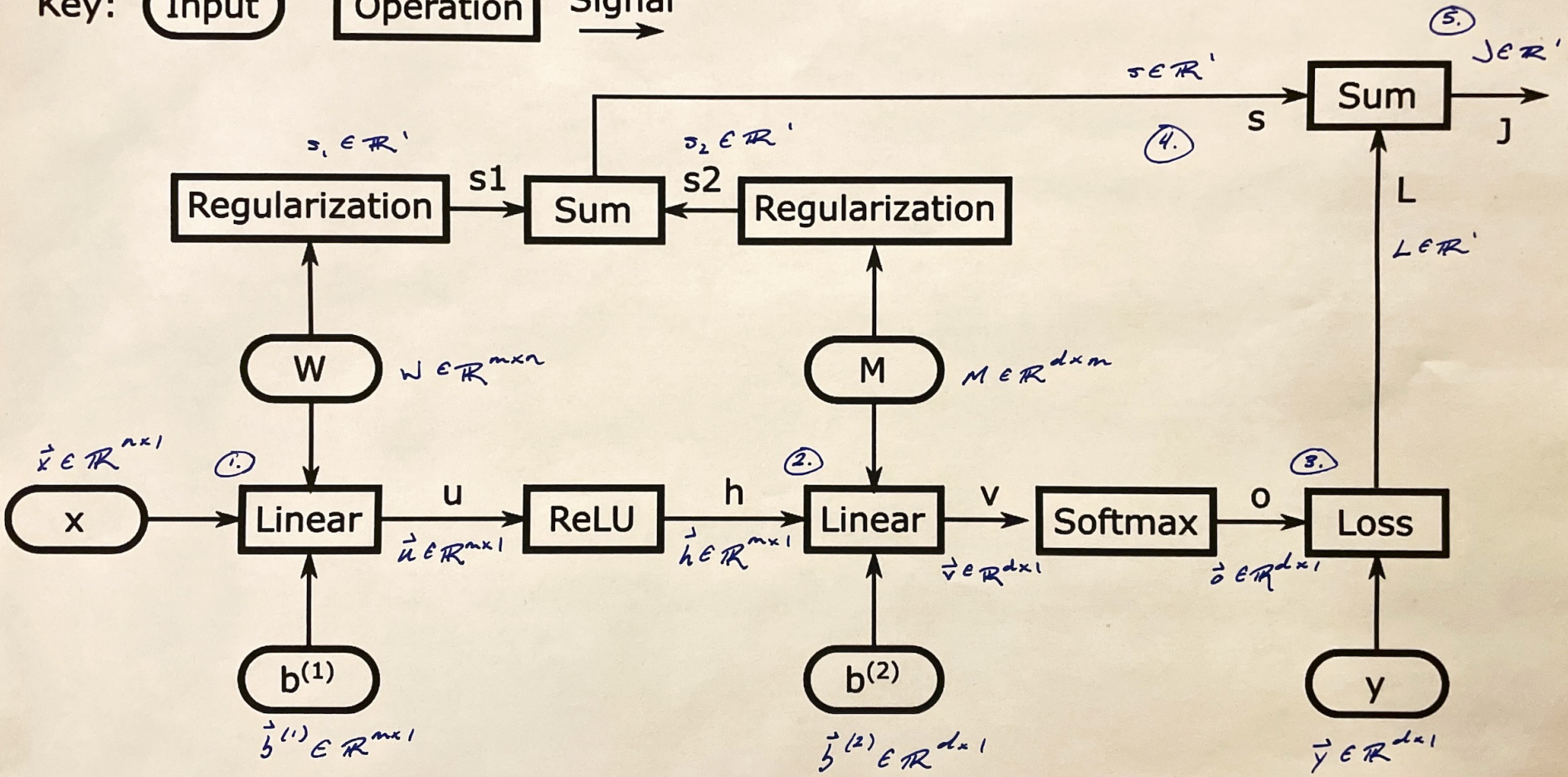
$s = \frac{\lambda}{2}(s_1 + s_2)$

Page Rank

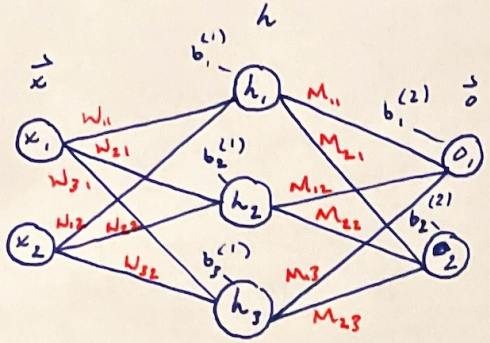
⑤ Objective Function:

$J = L + s$

Key: Input Operation Signal →



Network Architecture



input | hidden | output

weights :

$$W = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \\ W_{31} & W_{32} \end{bmatrix}_{3\times 2}$$

$$M = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \end{bmatrix}$$

① input → hidden :

$$\vec{z} = W\vec{x} + \vec{b}^{(1)}$$

where :

$W \in \mathbb{R}^{3\times 2}$

$\vec{x} \in \mathbb{R}^{2\times 1}$

$\vec{b}^{(1)} \in \mathbb{R}^{3\times 1}$

activation : ReLU

$$\vec{h} = \text{ReLU}(\vec{z})$$

where :

$\vec{h} \in \mathbb{R}^{3\times 1}$

regularization :

$$s_1 = \|W\|_F^2$$

$s_1 \in \mathbb{R}$

② hidden → output :

$$\vec{o} = M\vec{h} + \vec{b}^{(2)}$$

where :

$M \in \mathbb{R}^{2\times 3}$

$\vec{h} \in \mathbb{R}^{3\times 1}$

$\vec{b}^{(2)} \in \mathbb{R}^{2\times 1}$

$\vec{o} \in \mathbb{R}^{2\times 1}$

regularization :

$$s_2 = \|M\|_F^2$$

$s_2 \in \mathbb{R}$

③ Loss : MSE

$$L = MSE(\vec{y}, \vec{o})$$

$$L = \frac{1}{N}\sum_{i=1}^{N}(y_i - o_i)^2$$

④ regularization :

$$s = \frac{\lambda}{2}(s_1 + s_2)$$

where :

$s \in \mathbb{R}$

⑤ objective function :

$$J = L + s$$

where : $J \in \mathbb{R}$