

① Linear

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

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

② Linear

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

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

③ Loss Function

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

Loss Function

④ Regularization:

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

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

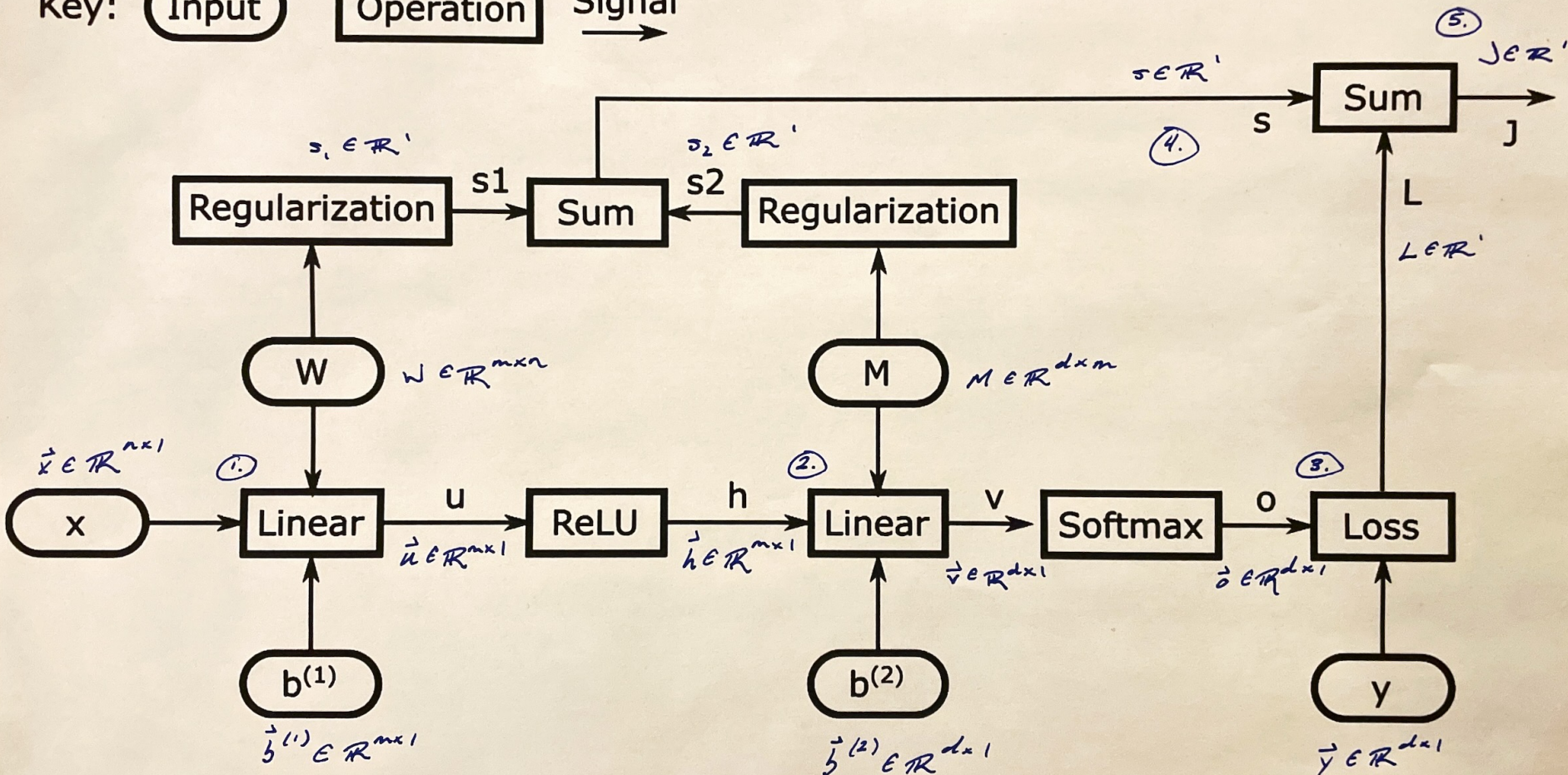
$$s = \frac{\lambda}{2} (s_1 + s_2)$$

Page Romyrk

⑤ Objective Function:

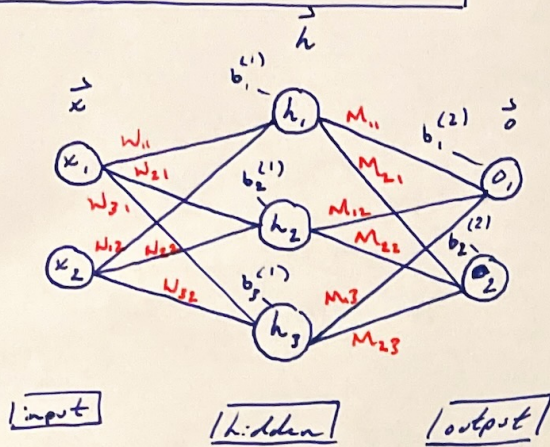
$$J = L + s$$

Key: **Input** **Operation** **Signal**





# NETWORK ARCHITECTURE



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  $\rightarrow$  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  $\rightarrow$  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

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

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

④ regularization:

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

where:

$$s \in \mathbb{R}$$

⑤ objective function:

$$J = L + s$$

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

# TEST CASES FORWARD PROPAGATION

## CASE 1

$\lambda = 0 \Rightarrow$  no regularization

values:

$$\vec{x} = \begin{bmatrix} 0.5 \\ -5.0 \end{bmatrix}_{2 \times 1} \quad W = \begin{bmatrix} 2.0 & 0.5 \\ 3.0 & 0.25 \\ 4.0 & 0.1 \end{bmatrix}_{3 \times 2} \quad \vec{b}_1 = \begin{bmatrix} 1.5 \\ 5.0 \\ 10.0 \end{bmatrix}_{3 \times 1}$$

$$M = \begin{bmatrix} 5.5 & 2.0 & 0.5 \\ 1.5 & 1.0 & 1.0 \end{bmatrix}_{2 \times 3} \quad \vec{b}_2 = \begin{bmatrix} 2.0 \\ 5.0 \end{bmatrix}_{2 \times 1}$$

① input  $\rightarrow$  hidden:

$$\vec{z} = W\vec{x} + \vec{b}_1$$

$$\vec{z} = \begin{bmatrix} 2.0 & 0.5 \\ 3.0 & 0.25 \\ 4.0 & 0.1 \end{bmatrix} \begin{bmatrix} 0.5 \\ -5.0 \end{bmatrix} + \vec{b}_1 = \begin{bmatrix} (2)0.5 + (0.5)(-5) \\ (3)0.5 + (0.25)(-5) \\ (4)0.5 + (0.1)(-5) \end{bmatrix} + \vec{b}_1$$

$$= \begin{bmatrix} -1.5 \\ 0.25 \\ 1.5 \end{bmatrix} + \begin{bmatrix} 1.5 \\ 5.0 \\ 10.0 \end{bmatrix} = \begin{bmatrix} 0 \\ 5.25 \\ 11.5 \end{bmatrix}_{3 \times 1}$$

$\Rightarrow$  activation: ReLU

$$\vec{h} = \text{ReLU}(\vec{z}) = \begin{bmatrix} 0 \\ 5.25 \\ 11.5 \end{bmatrix}_{3 \times 1}$$

② hidden  $\rightarrow$  output:

$$\vec{o} = M\vec{h} + \vec{b}_2$$

$$= \begin{bmatrix} 5.5 & 2.0 & 0.5 \\ 1.5 & 1.0 & 1.0 \end{bmatrix} \begin{bmatrix} 0 \\ 5.25 \\ 11.5 \end{bmatrix} + \vec{b}_2$$

$$= \begin{bmatrix} (5.5)0 + (2)5.25 + (0.5)11.5 \\ (1.5)0 + (1.0)5.25 + (1.0)11.5 \end{bmatrix} + \vec{b}_2$$

$$= \begin{bmatrix} 14.25 \\ 16.75 \end{bmatrix} + \begin{bmatrix} 2.0 \\ 5.0 \end{bmatrix} = \begin{bmatrix} 16.25 \\ 21.75 \end{bmatrix}_{2 \times 1}$$

NOTE:

$\lambda = 0$  so no regularization

## CASE 2

$\lambda = 0.1$  (regularization)

values:

$$\vec{x} = \begin{bmatrix} 1.0 \\ 2.0 \end{bmatrix}_{2 \times 1} \quad W = \begin{bmatrix} 1.0 & 0.5 \\ 0.5 & -1.0 \\ 2.0 & 5.0 \end{bmatrix}_{3 \times 2} \quad \vec{b}_1 = \begin{bmatrix} 2.0 \\ 1.0 \\ -2.0 \end{bmatrix}_{3 \times 1}$$

$$M = \begin{bmatrix} 5.0 & 0.5 & 1.0 \\ 0.5 & 1.0 & 2.0 \end{bmatrix}_{2 \times 3} \quad \vec{b}_2 = \begin{bmatrix} 4.0 \\ -5.0 \end{bmatrix}_{2 \times 1}$$

① input  $\rightarrow$  hidden:

$$\vec{z} = W\vec{x} + \vec{b}_1$$

$$= \begin{bmatrix} 1.0 & 0.5 \\ 0.5 & -1.0 \\ 2.0 & 5.0 \end{bmatrix} \begin{bmatrix} 1.0 \\ 2.0 \end{bmatrix} + \vec{b}_1$$

$$= \begin{bmatrix} (1)(1) + (0.5)(2) \\ (0.5)(1) + (-1)(2) \\ (2)(1) + (5)(2) \end{bmatrix} + \vec{b}_1 = \begin{bmatrix} 2.0 \\ -1.5 \\ 12 \end{bmatrix} + \begin{bmatrix} 2.0 \\ 1.0 \\ -2.0 \end{bmatrix}$$

$$\vec{z} = \begin{bmatrix} 4.0 \\ -0.5 \\ 10.0 \end{bmatrix}_{3 \times 1}$$

$\Rightarrow$  activation: ReLU

$$\vec{h} = \text{ReLU}(\vec{z}) = \begin{bmatrix} 4.0 \\ 0.0 \\ 10.0 \end{bmatrix}$$

$\Rightarrow$  regularization step:

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

$$= (1)^2 + (0.5)^2 + (0.5)^2 + (-1)^2 + (2)^2 + (5)^2$$

$$s_1 = \underline{31.5}$$

② hidden  $\rightarrow$  output:

$$\vec{o} = M\vec{h} + \vec{b}_2$$

$$= \begin{bmatrix} 5.0 & 0.5 & 1.0 \\ 0.5 & 1.0 & 2.0 \end{bmatrix} \begin{bmatrix} 4.0 \\ 0 \\ 10.0 \end{bmatrix} + \vec{b}_2$$

$$= \begin{bmatrix} (5)(4) + (0.5)0 + (1)(10) \\ (0.5)(4) + (1.0)0 + (2.0)(10) \end{bmatrix} + \vec{b}_2 = \begin{bmatrix} 30.0 \\ 22.0 \end{bmatrix} + \begin{bmatrix} 4.0 \\ -5.0 \end{bmatrix}$$

$$\vec{o} = \begin{bmatrix} 34.0 \\ 17.0 \end{bmatrix}$$

$\Rightarrow$  regularization step:

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

$$= (5)^2 + (0.5)^2 + (1)^2 + (0.5)^2 + (1)^2 + (2)^2$$

$$= \underline{31.5}$$

③ Regularization:

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

$$= \frac{0.1}{2} (31.5 + 31.5)$$

$$s = \underline{3.15}$$