

## Forward Propagation

$$\vec{u} = W\vec{x} + \vec{b}$$

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

$$\vec{v} = M\vec{h} + \vec{c}$$

$$L = \frac{1}{2} \|\vec{y} - \vec{v}\|_F^2$$

$$s_1 = \frac{1}{2} \|W\|_F^2 \quad s_2 = \frac{1}{2} \|M\|_F^2$$

$$s = s_1 + s_2$$

$$J = L + s$$

## Backward Propagation

$$\frac{\partial J}{\partial L} = 1 \quad \frac{\partial J}{\partial s} = 1 \quad \frac{\partial J}{\partial s_1} = 1 \quad \frac{\partial J}{\partial s_2} = 1$$

$$\frac{\partial J}{\partial \vec{v}} = \text{prod}\left(\frac{\partial J}{\partial L}, \frac{\partial L}{\partial \vec{v}}\right) = \frac{\partial L}{\partial \vec{v}}$$

$$\frac{\partial s_1}{\partial W} = W \quad \frac{\partial s_2}{\partial M} = M \quad \frac{\partial s}{\partial W} = W \quad \frac{\partial s}{\partial M} = M$$

$$\frac{\partial J}{\partial M} = \text{prod}\left(\frac{\partial J}{\partial \vec{v}}, \frac{\partial \vec{v}}{\partial M}\right) + \text{prod}\left(\frac{\partial J}{\partial s}, \frac{\partial s}{\partial M}\right) = \frac{\partial J}{\partial \vec{v}} \vec{h}^T + M$$

$$\frac{\partial J}{\partial \vec{h}} = \text{prod}\left(\frac{\partial J}{\partial \vec{v}}, \frac{\partial \vec{v}}{\partial \vec{h}}\right) = M^T \frac{\partial J}{\partial \vec{v}} \quad \frac{\partial J}{\partial \vec{u}} = \text{prod}\left(\frac{\partial J}{\partial \vec{h}}, \frac{\partial \vec{h}}{\partial \vec{u}}\right) = \frac{\partial J}{\partial \vec{h}} \odot \text{ReLU}'(\vec{u})$$

$$\frac{\partial J}{\partial W} = \text{prod}\left(\frac{\partial J}{\partial \vec{u}}, \frac{\partial \vec{u}}{\partial W}\right) + \text{prod}\left(\frac{\partial J}{\partial s}, \frac{\partial s}{\partial W}\right) = \frac{\partial J}{\partial \vec{u}} \vec{x}^T + W$$

# Forward Propagation Examples

$$\vec{x} = \begin{bmatrix} 6 \\ 3 \end{bmatrix} \quad W = \begin{bmatrix} .3 & .5 \\ .2 & .7 \\ .1 & .4 \end{bmatrix} \quad \vec{b} = \begin{bmatrix} .1 \\ .2 \\ .3 \end{bmatrix}$$

$$\vec{u} = W\vec{x} + \vec{b}$$

$$\begin{bmatrix} .3 & .5 \\ .2 & .7 \\ .1 & .4 \end{bmatrix} \begin{bmatrix} 6 \\ 3 \end{bmatrix} = \begin{bmatrix} .3 \times 6 + .5 \times 3 \\ .2 \times 6 + .7 \times 3 \\ .1 \times 6 + .4 \times 3 \end{bmatrix} = \begin{bmatrix} 3.3 \\ 3.3 \\ 1.8 \end{bmatrix} \quad \begin{bmatrix} 3.3 \\ 3.3 \\ 1.8 \end{bmatrix} + \begin{bmatrix} .1 \\ .2 \\ .3 \end{bmatrix} = \begin{bmatrix} 3.4 \\ 3.5 \\ 2.1 \end{bmatrix}$$

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

$$\text{ReLU}\left(\begin{bmatrix} 3.4 \\ 3.5 \\ 2.1 \end{bmatrix}\right) = \begin{bmatrix} 3.4 \\ 3.5 \\ 2.1 \end{bmatrix}$$

$$M = \begin{bmatrix} .3 & .5 & .1 \\ .2 & .7 & .8 \end{bmatrix} \quad \vec{c} = \begin{bmatrix} .3 \\ .5 \end{bmatrix}$$

$$\vec{y} = M\vec{h} + \vec{c}$$

$$\begin{bmatrix} .3 & .5 & .1 \\ .2 & .7 & .8 \end{bmatrix} \begin{bmatrix} 3.4 \\ 3.5 \\ 2.1 \end{bmatrix} = \begin{bmatrix} .3 \times 3.4 + .5 \times 3.5 + .1 \times 2.1 \\ .2 \times 3.4 + .7 \times 3.5 + .8 \times 2.1 \end{bmatrix} = \begin{bmatrix} 2.98 \\ 4.81 \end{bmatrix}$$

$$\begin{bmatrix} 2.98 \\ 4.81 \end{bmatrix} + \begin{bmatrix} .3 \\ .5 \end{bmatrix} = \begin{bmatrix} 3.28 \\ 5.31 \end{bmatrix}$$

$$L = \frac{1}{2} \|\vec{y} - \vec{v}\|_F^2 \quad \vec{y} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$$

$$\frac{1}{2} \left( \sum \left( \left( \begin{bmatrix} 3 \\ 5 \end{bmatrix} - \begin{bmatrix} 3.28 \\ 5.31 \end{bmatrix} \right)^2 \right) \right) = \frac{1}{2} \sum \begin{bmatrix} 0.0784 \\ 0.0961 \end{bmatrix} = 0.08725$$

$$S_1 = \frac{1}{2} \|W\|_F^2 = \frac{1}{2} \sum \left( \begin{bmatrix} .3 & .5 \\ .2 & .7 \\ .1 & .4 \end{bmatrix}^2 \right) = 0.52$$

$$S_2 = \frac{1}{2} \|M\|_F^2 = \frac{1}{2} \sum \left( \begin{bmatrix} .3 & .5 & .1 \\ .2 & .7 & .8 \end{bmatrix}^2 \right) = 0.76$$

$$S = S_1 + S_2 = 0.52 + 0.76 = 1.28$$

$$J = L + S = 0.08725 + 1.28 = 1.36725$$