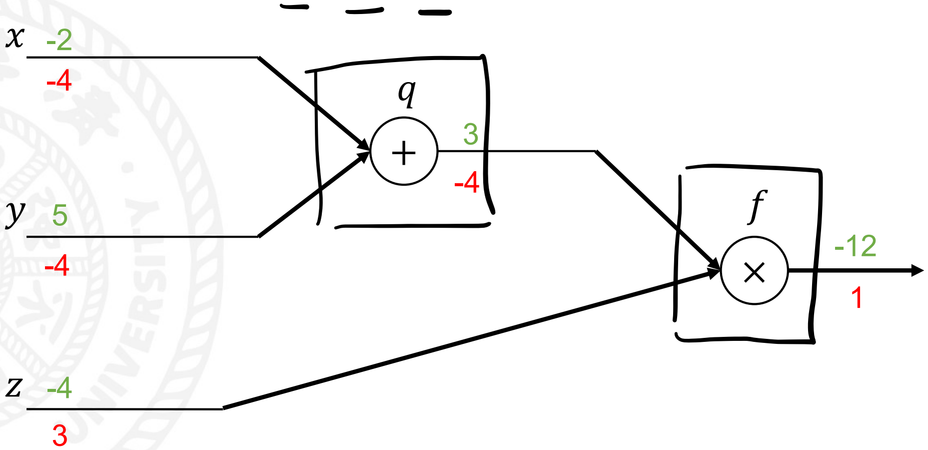
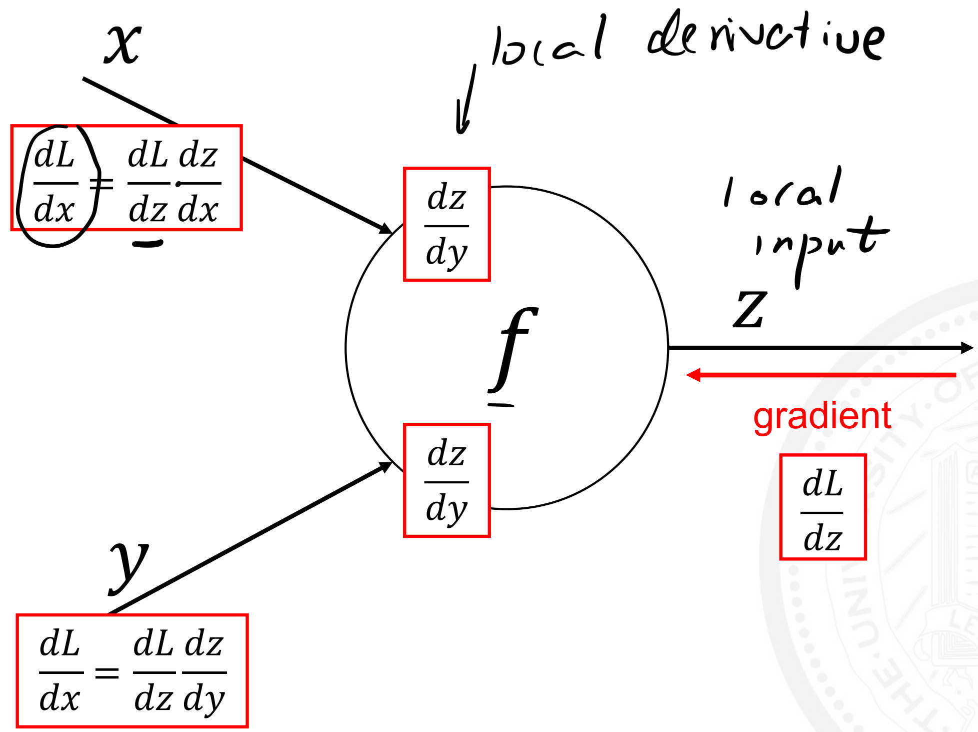
Backpropagation

$$f(x, y, z) = (x + y)z$$

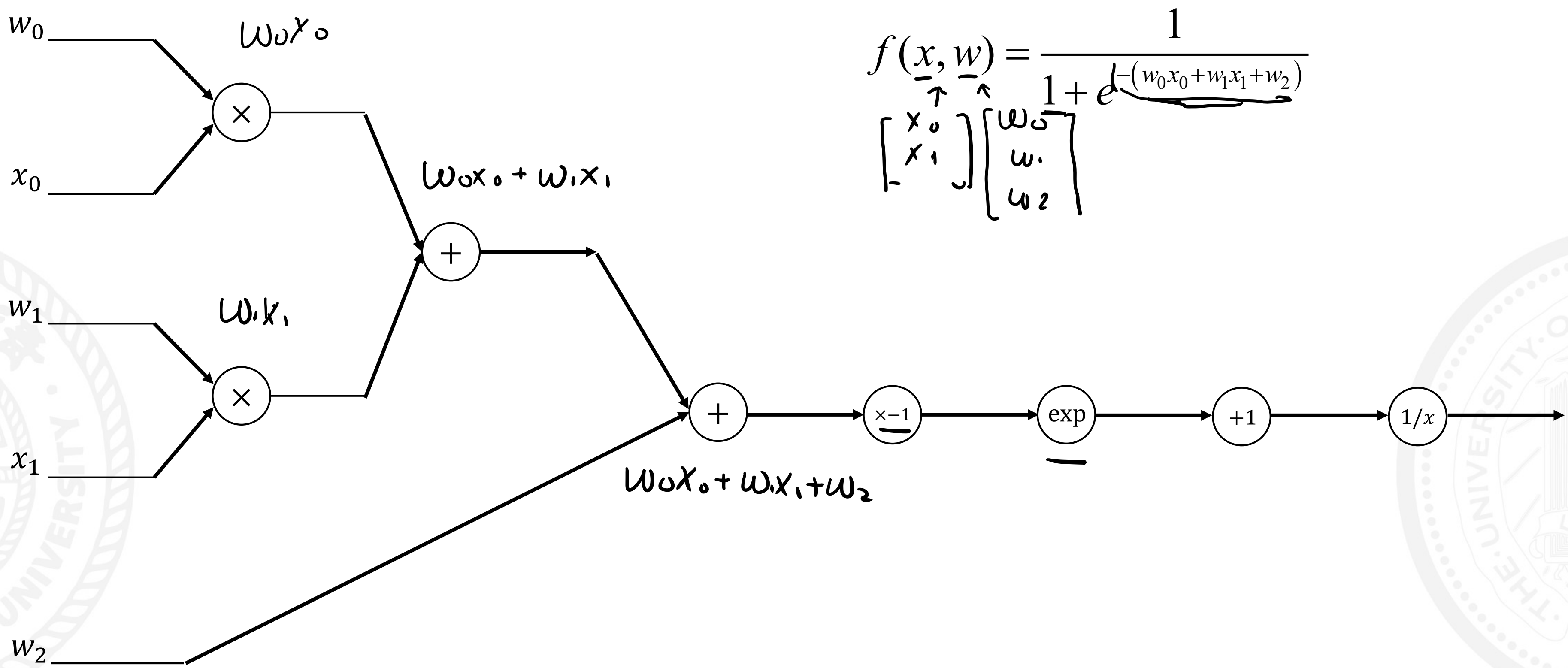
We want $\frac{df}{dx}, \frac{df}{dy}, \frac{df}{dz}$





Backpropagation

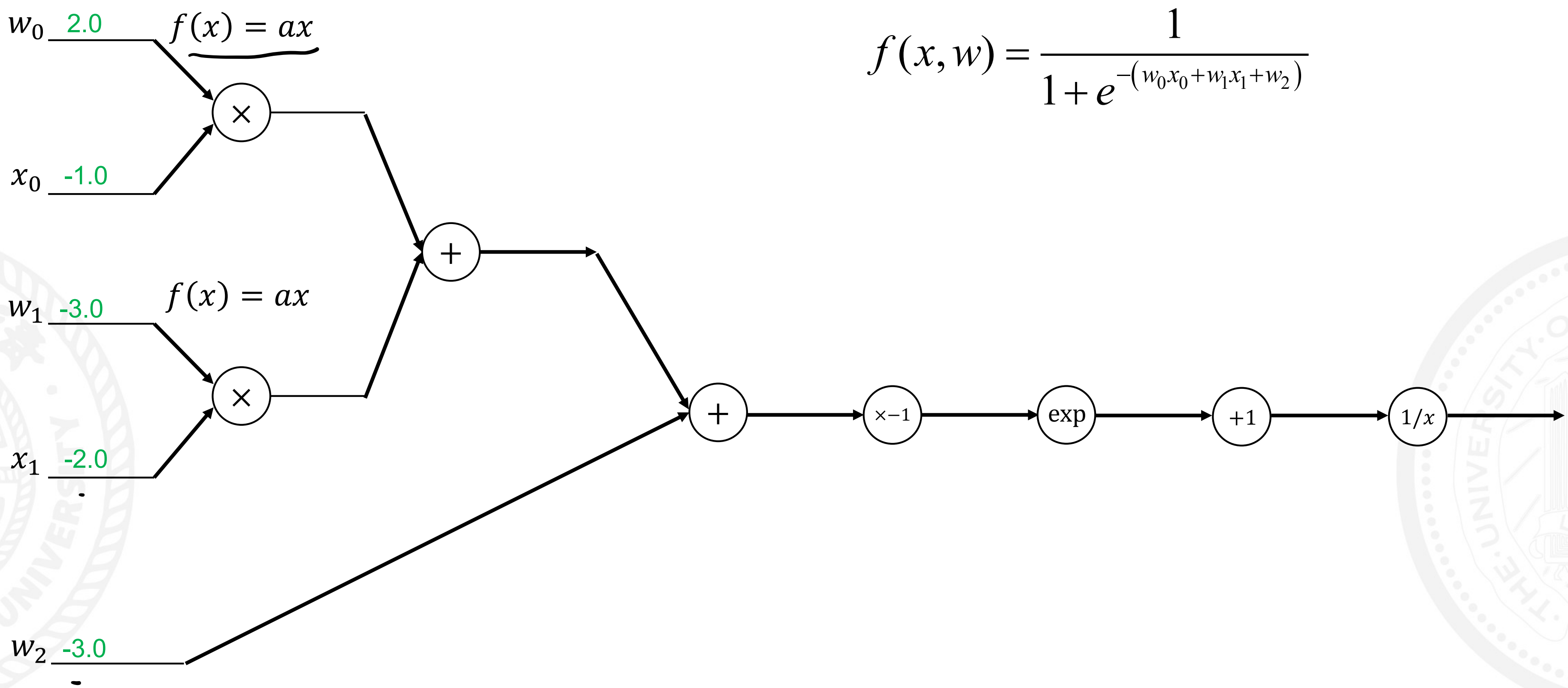
$$f(x,w)=\frac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$



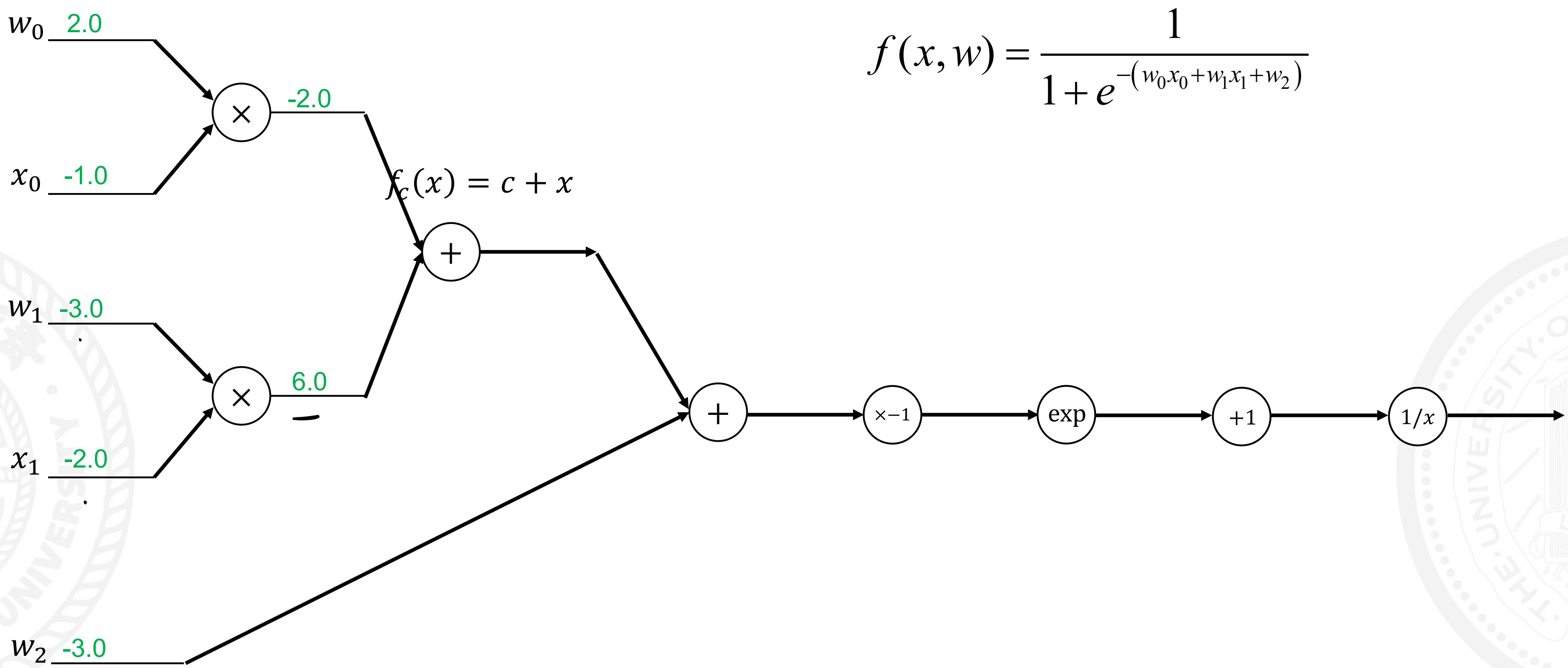
Backpropagation

TBSI 清华-伯克利深圳学院 Tsinghua-Berkeley Shenzhen Institute

$$f(x,w) = \frac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$



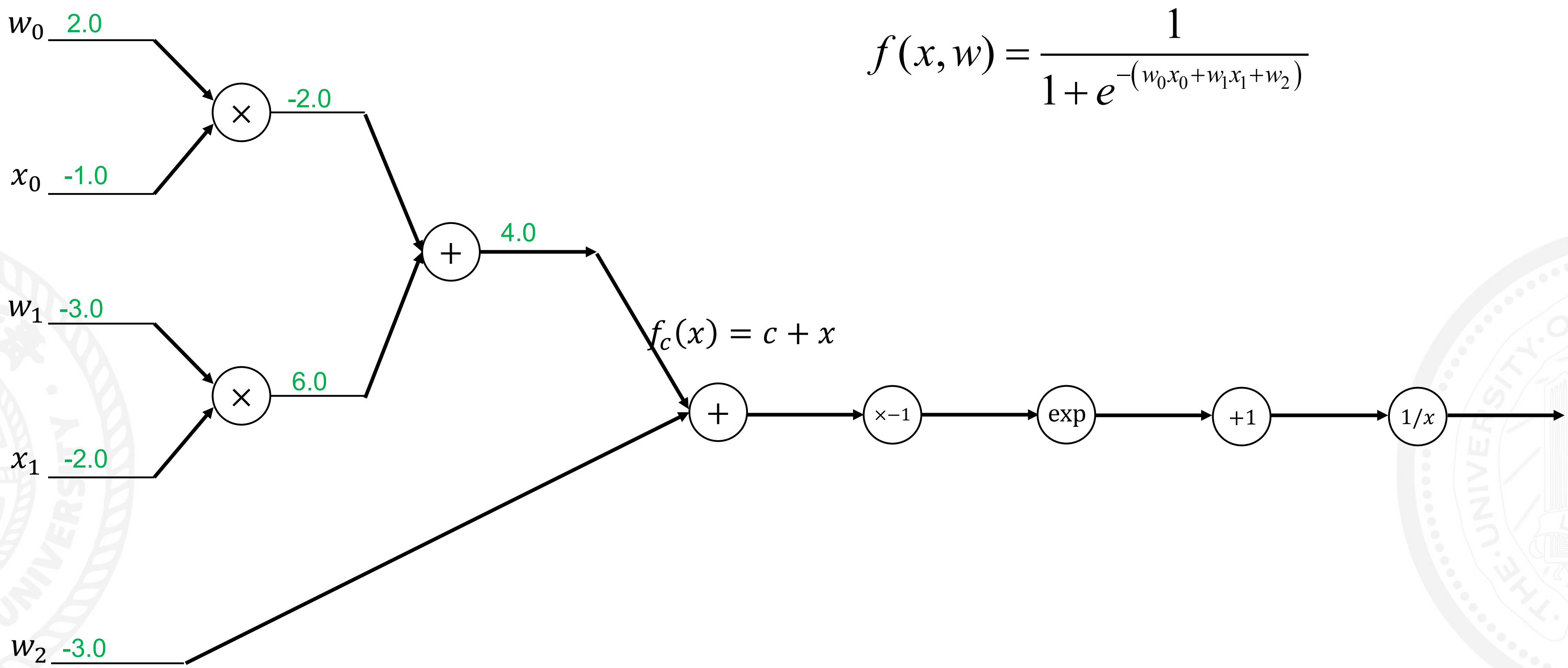
Backprojection



$$f(x,w) = \frac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$

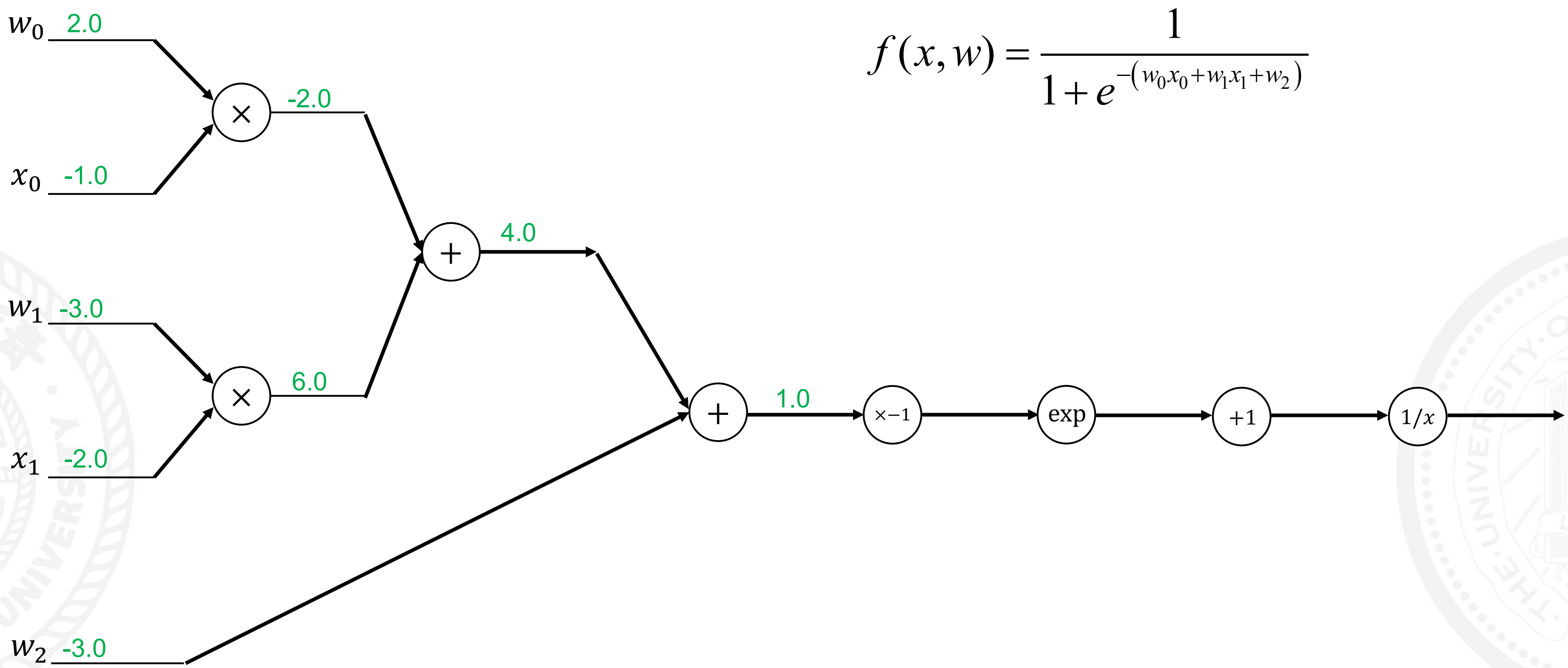
Backpropagation

$$f(x,w) = \frac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$



Backpropagation

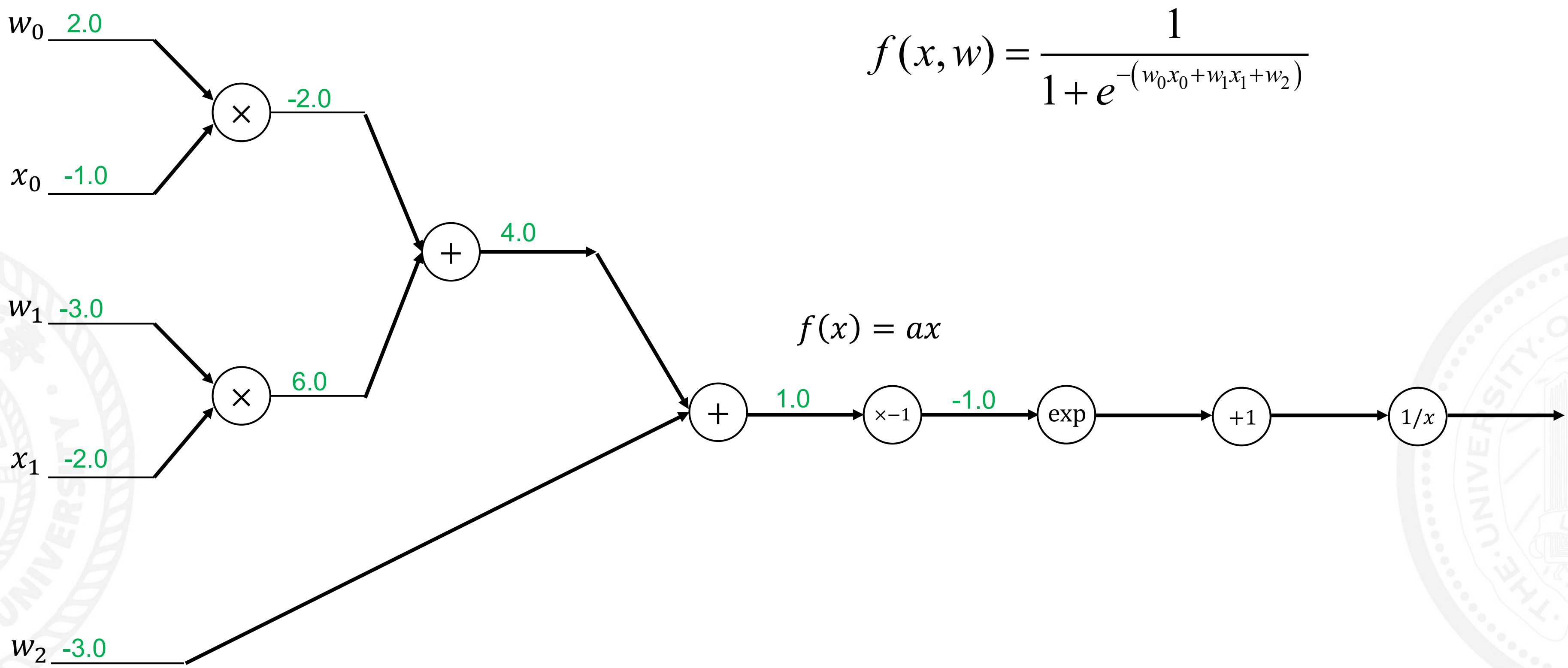
$$f(x,w) = \frac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$



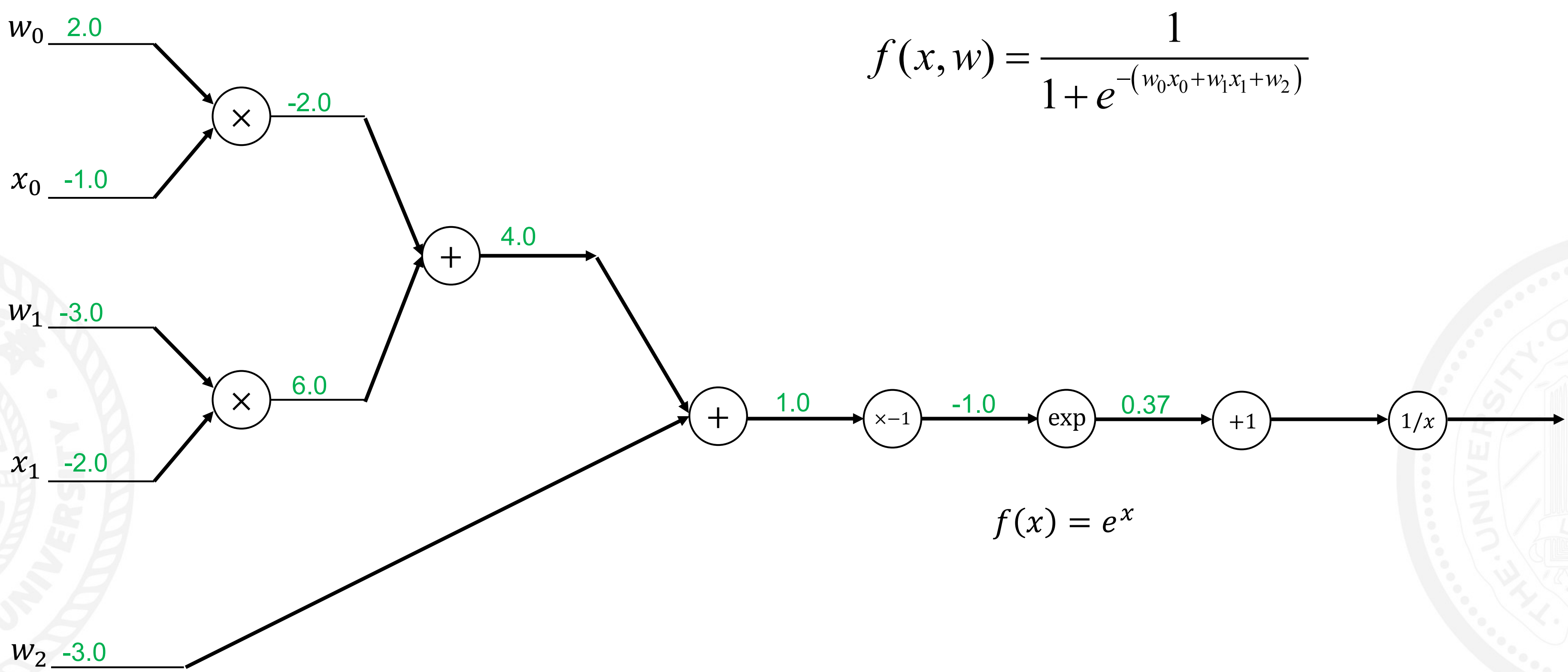
Backpropagation

$$f(x,w) = \frac{1}{1+e^{-(w_0 x_0 + w_1 x_1 + w_2)}}$$

$$f(x) = ax$$

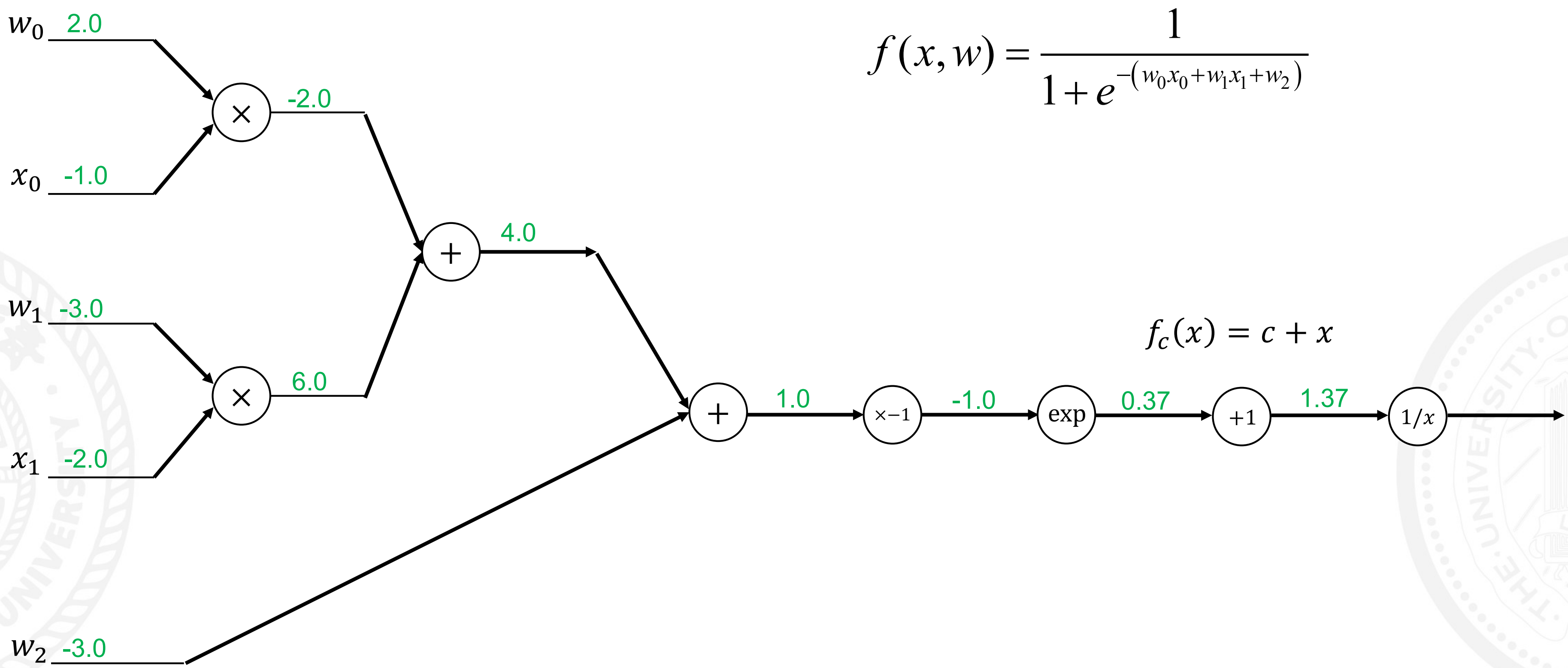


Backpropagation



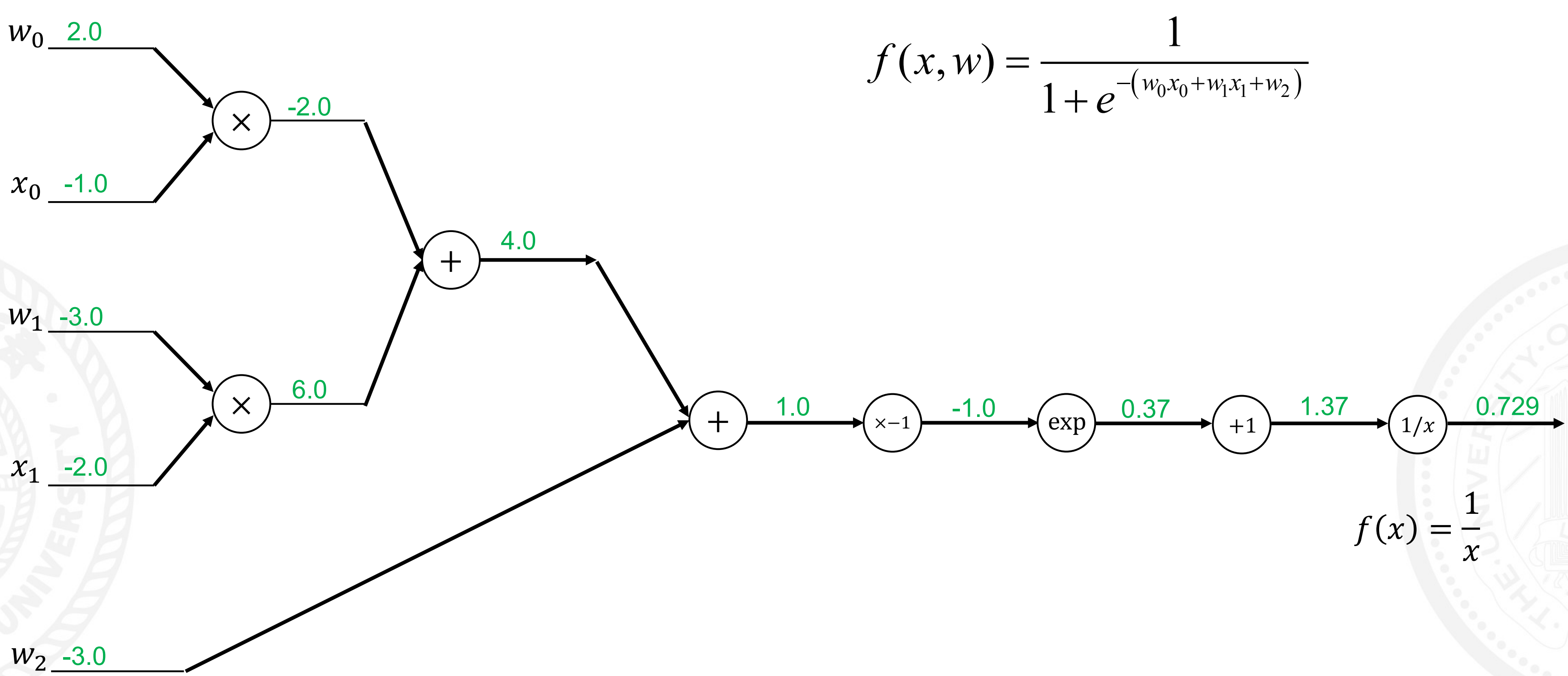
Backpropagation

$$f(x,w) = \frac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$



$$f_c(x) = c + x$$

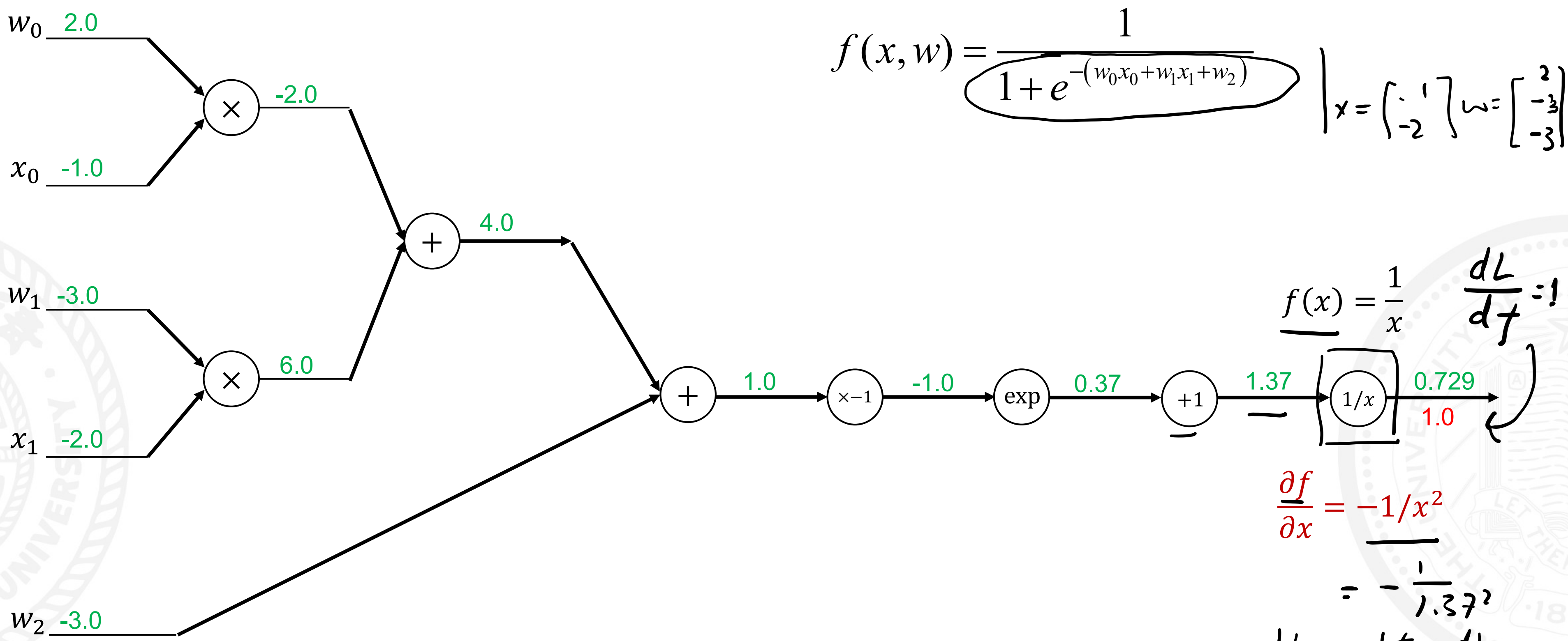
Backprojection



Backpropagation

TBSI 清华-伯克利深圳学院 Tsinghua-Berkeley Shenzhen Institute

$$f(x,w) = \frac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$

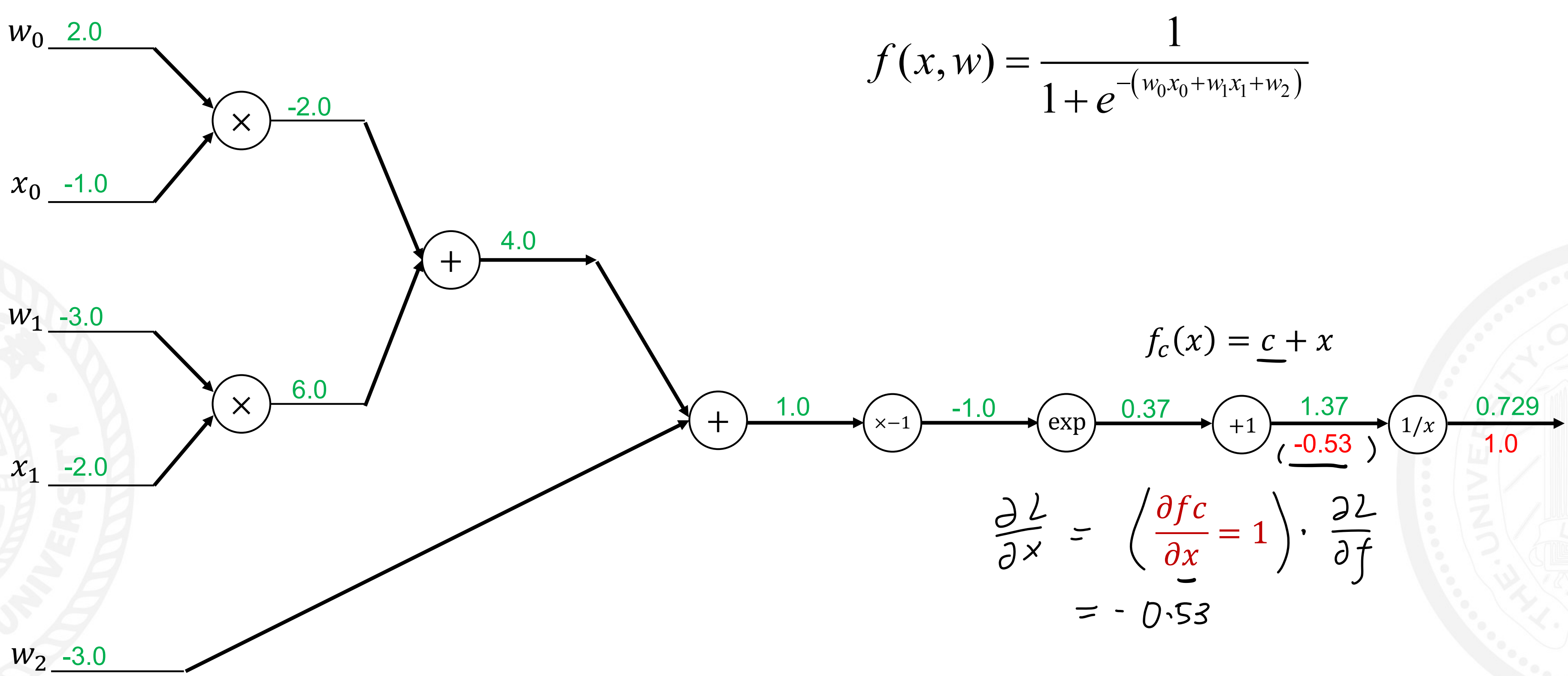


$$f(x) = \frac{1}{x}$$

$$\frac{\partial f}{\partial x} = -1/x^2$$

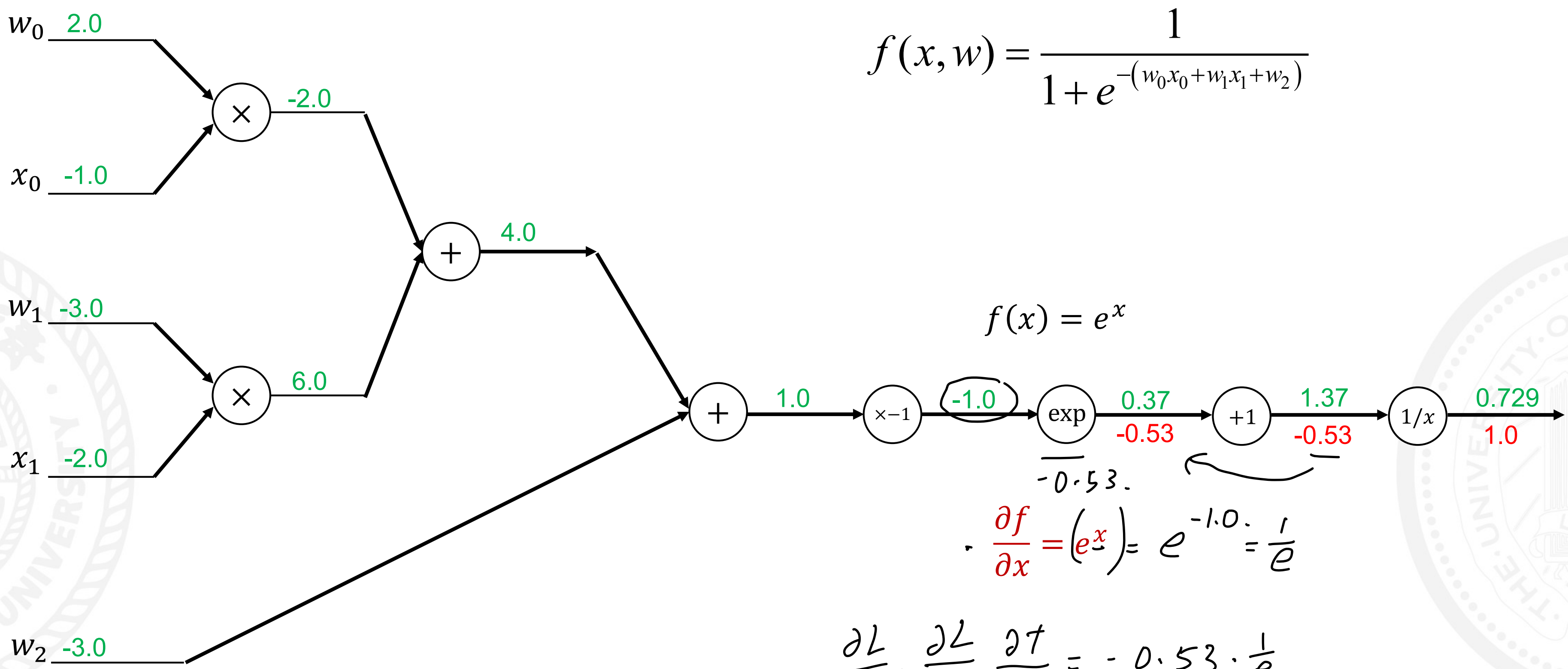
Backpropagation

$$f(x,w) = \frac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$



Backpropagation

$$f(x,w) = \frac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$



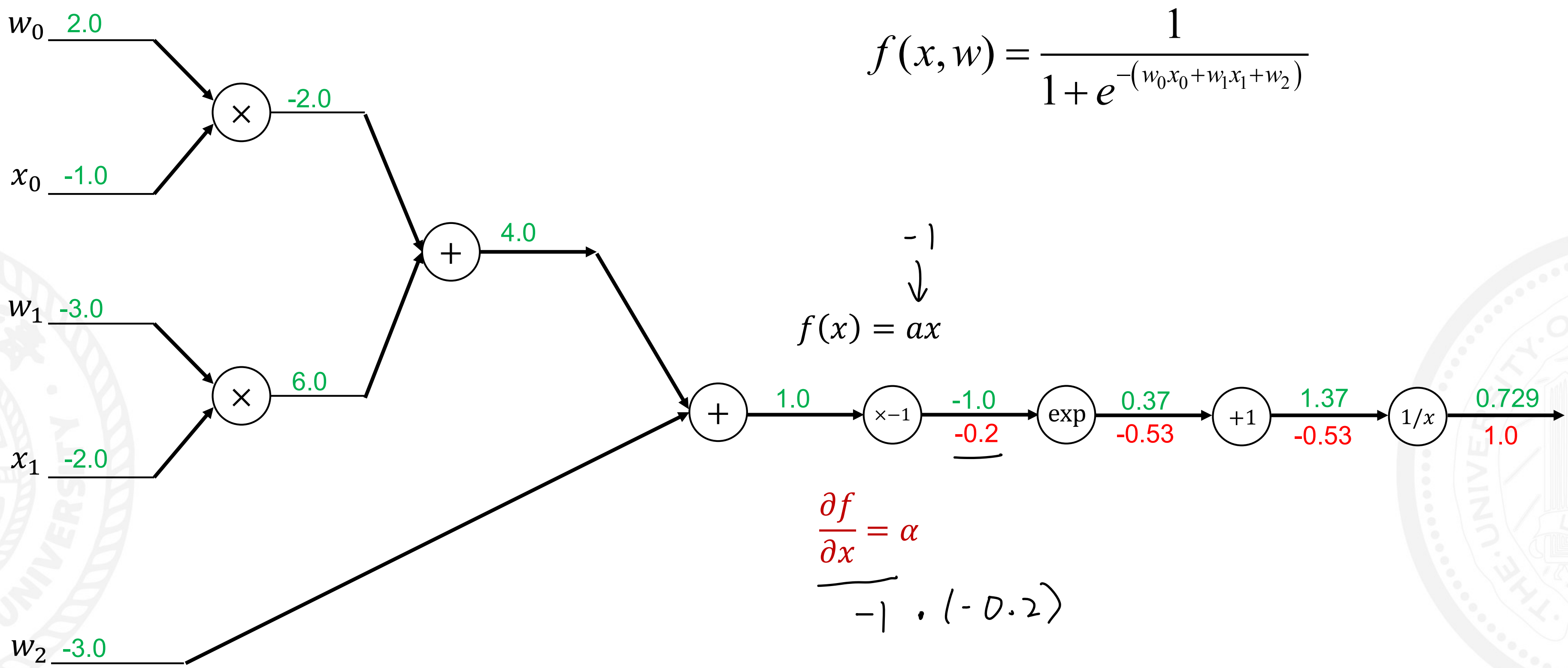
$$f(x) = e^x$$

$$\frac{\partial f}{\partial x} = (e^x) = e^{-1.0} = \frac{1}{e}$$

$$\frac{\partial L}{\partial x} = \frac{\partial L}{\partial f}\frac{\partial f}{\partial x} = -0.53 \cdot \frac{1}{e}.$$

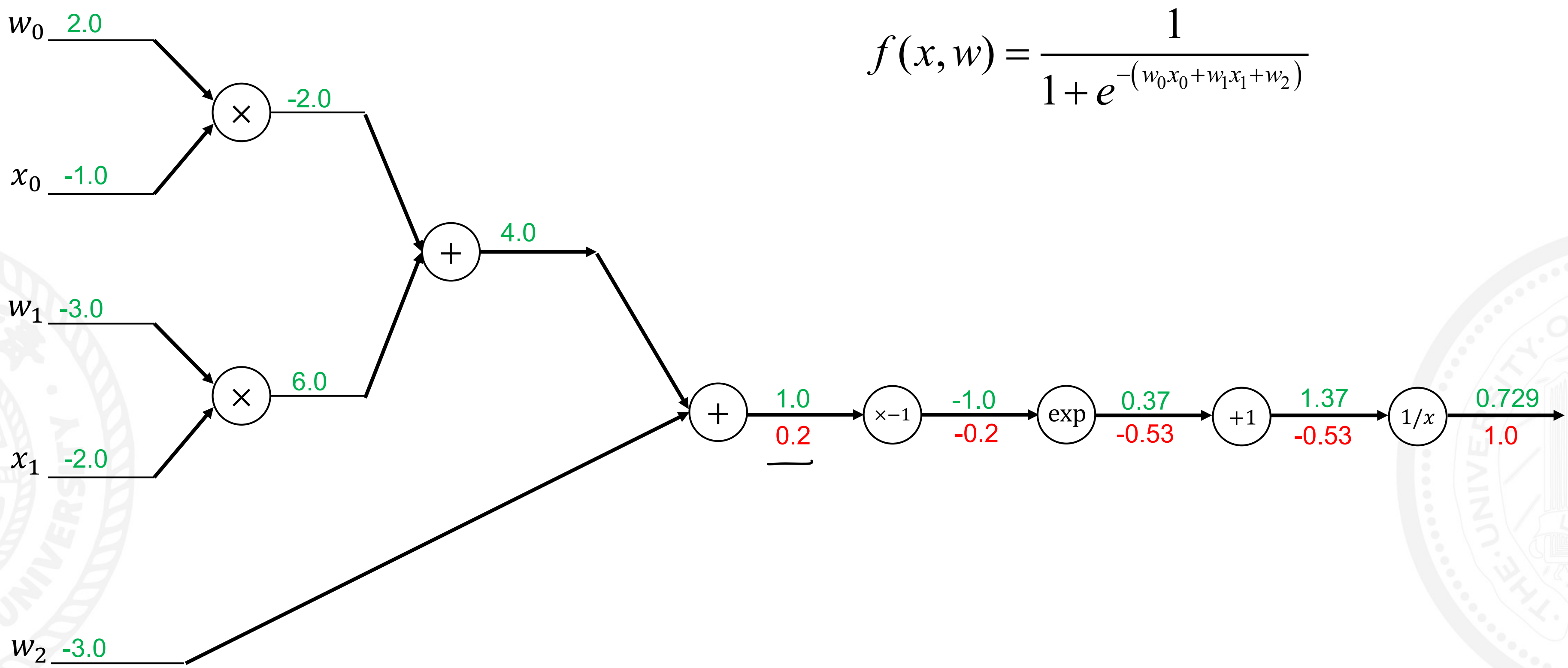
Backpropagation

$$f(x,w) = \frac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$

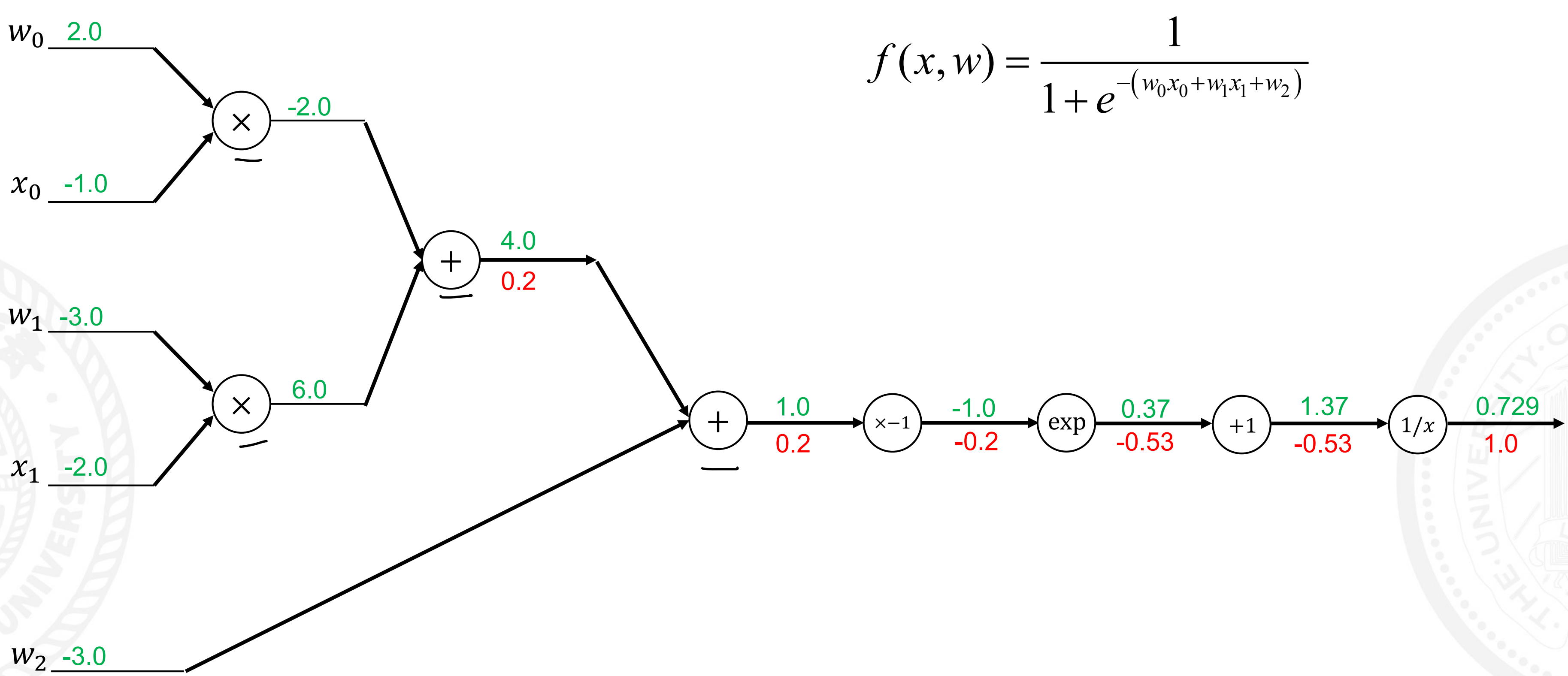


Backpropagation

$$f(x,w) = \frac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$

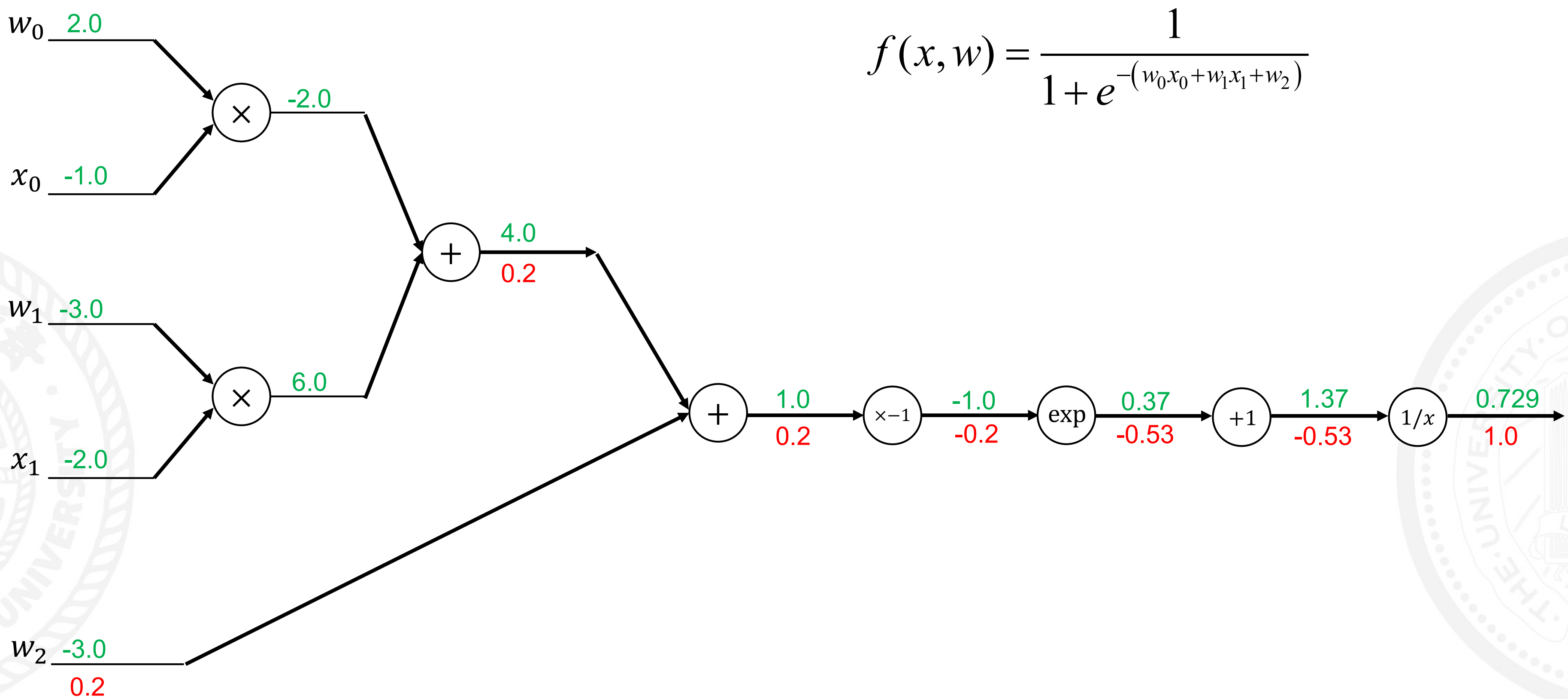


Backpropagation



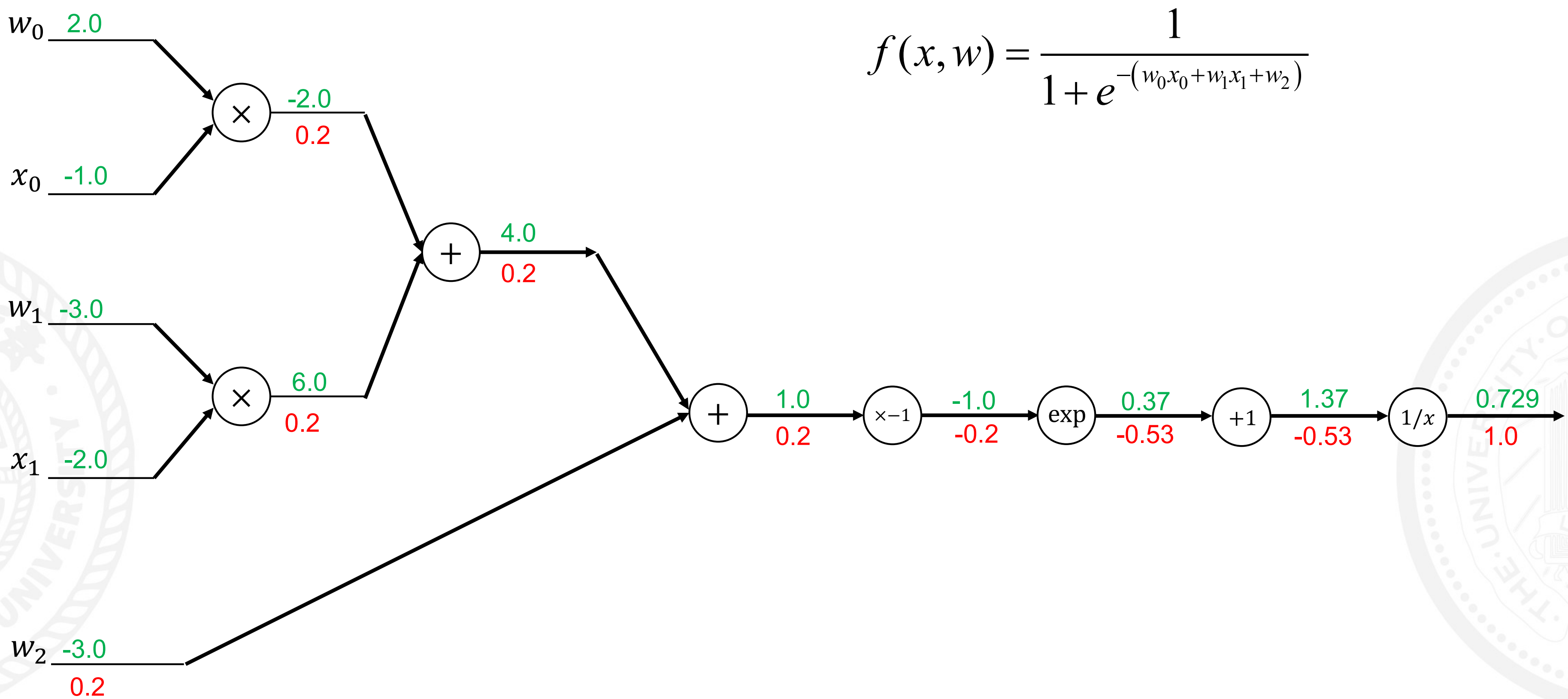
Backpropagation

$$f(x,w) = \frac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$



Backpropagation

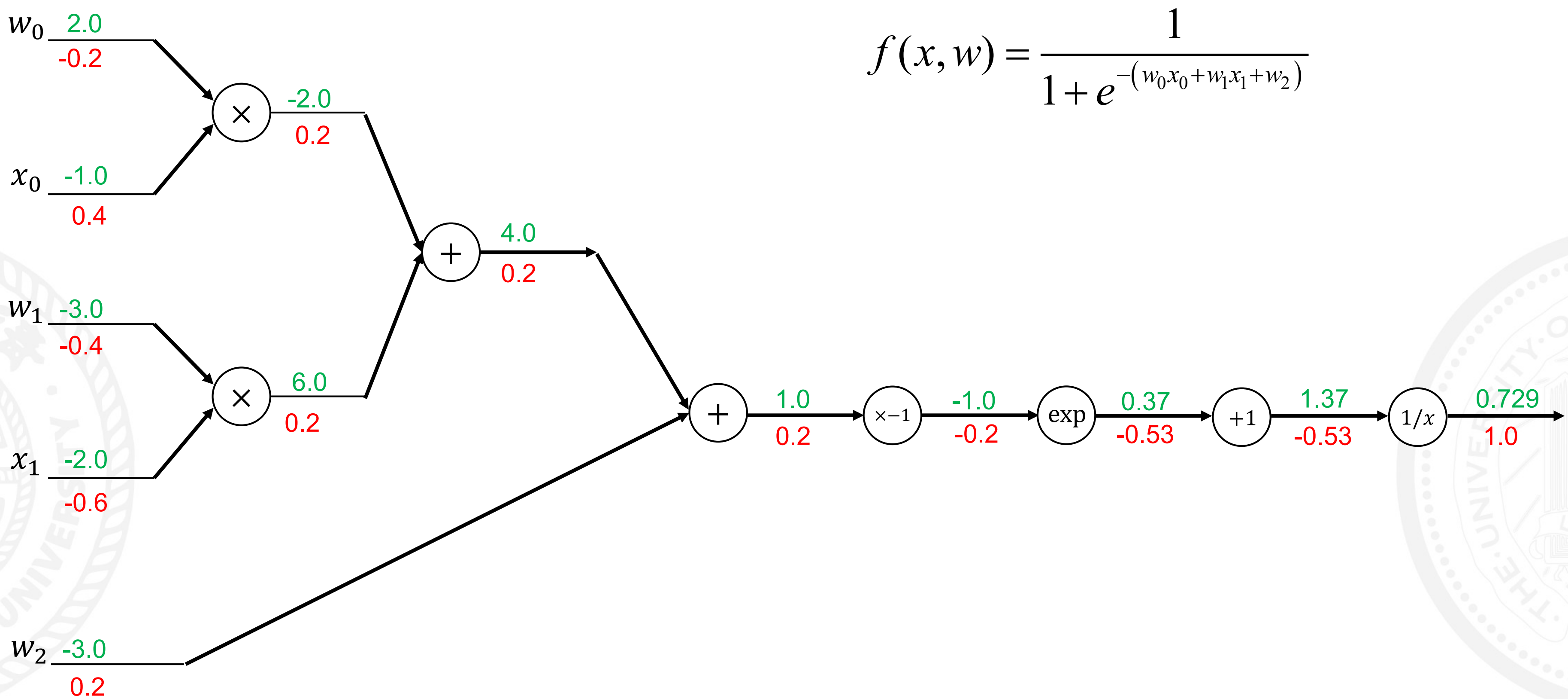
$$f(x,w)=\frac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$



Backpropagation

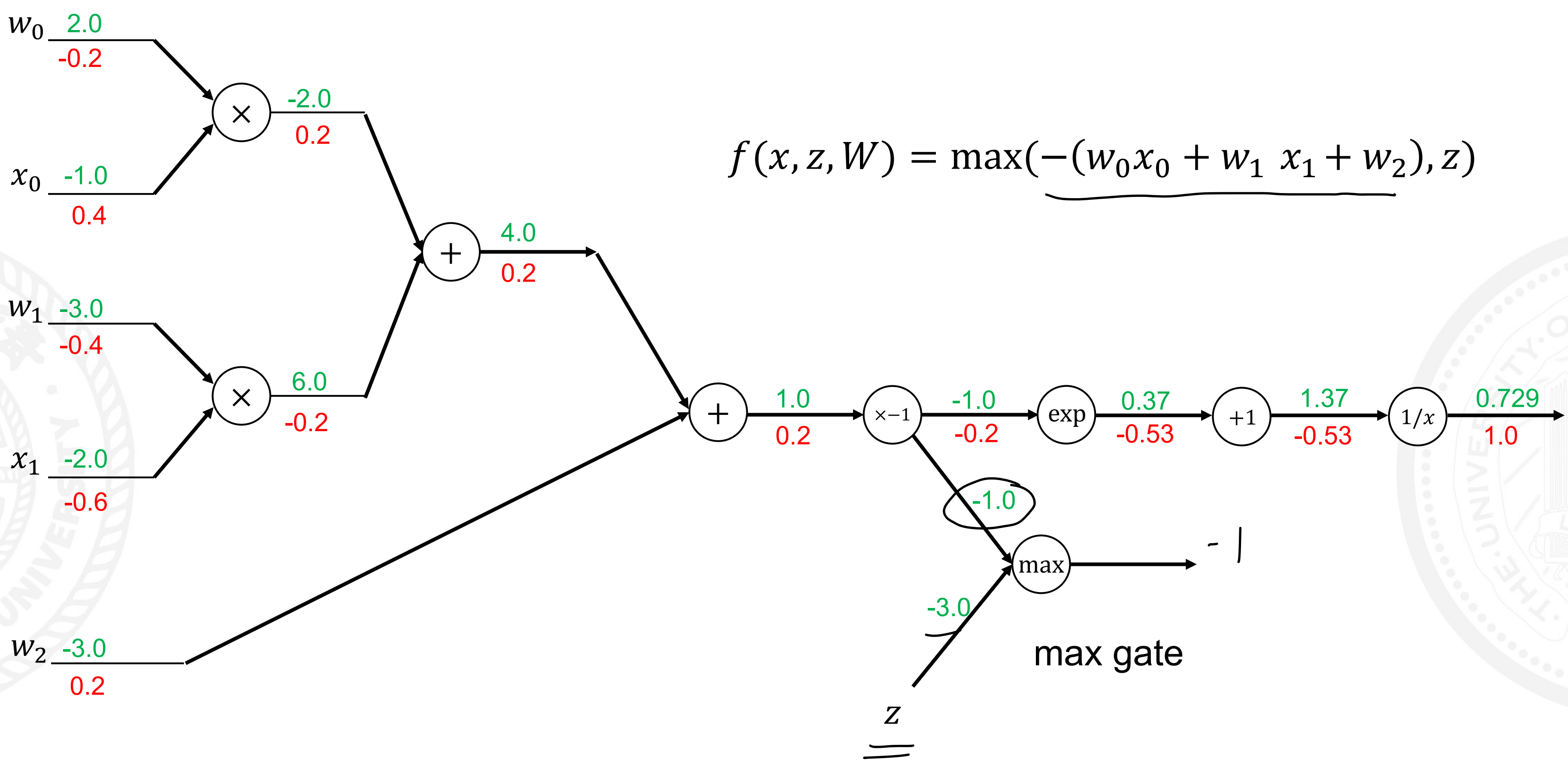
TBSI 清华-伯克利深圳学院 Tsinghua-Berkeley Shenzhen Institute

$$f(x,w)=\frac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$



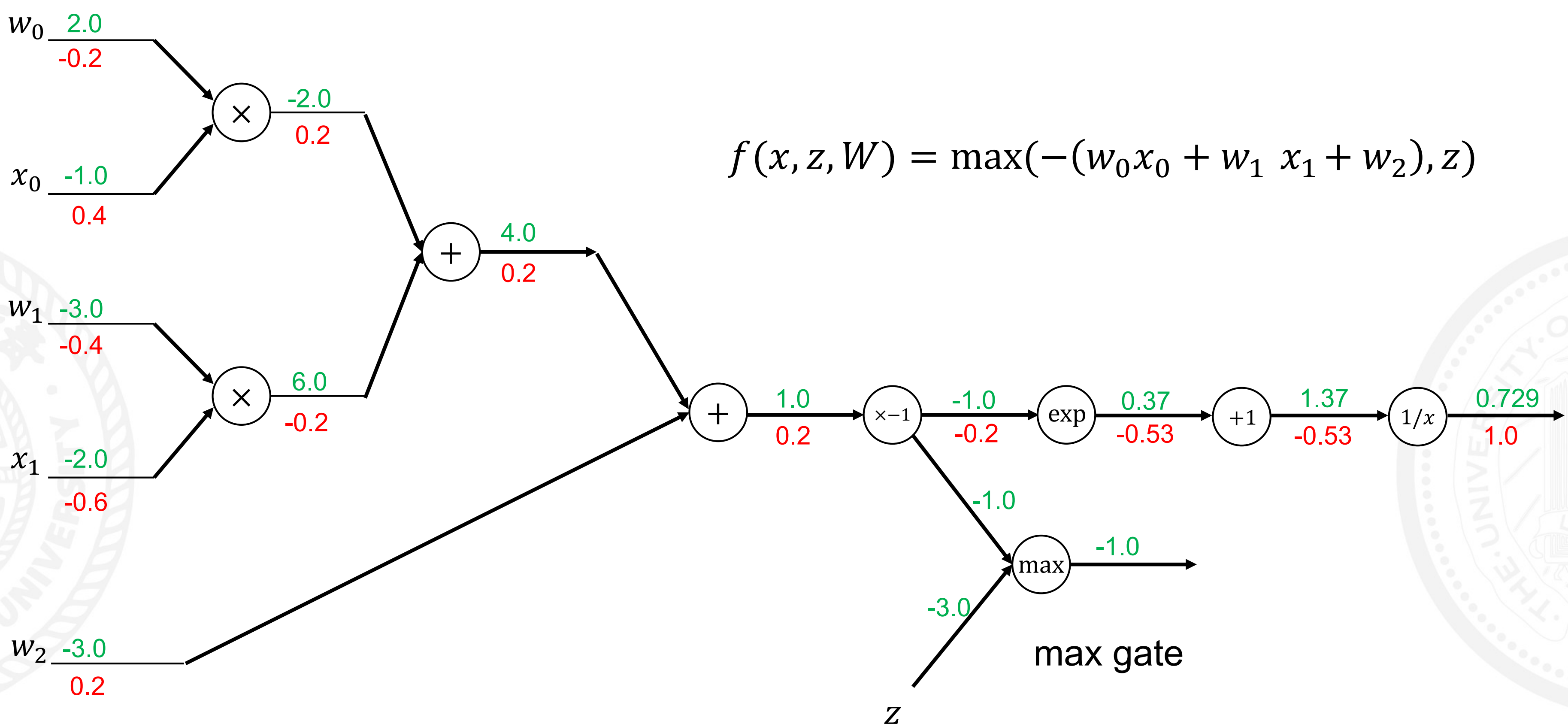
Backpropagation

$$f(x, z, W) = \max(-(w_0 x_0 + w_1 x_1 + w_2), z)$$



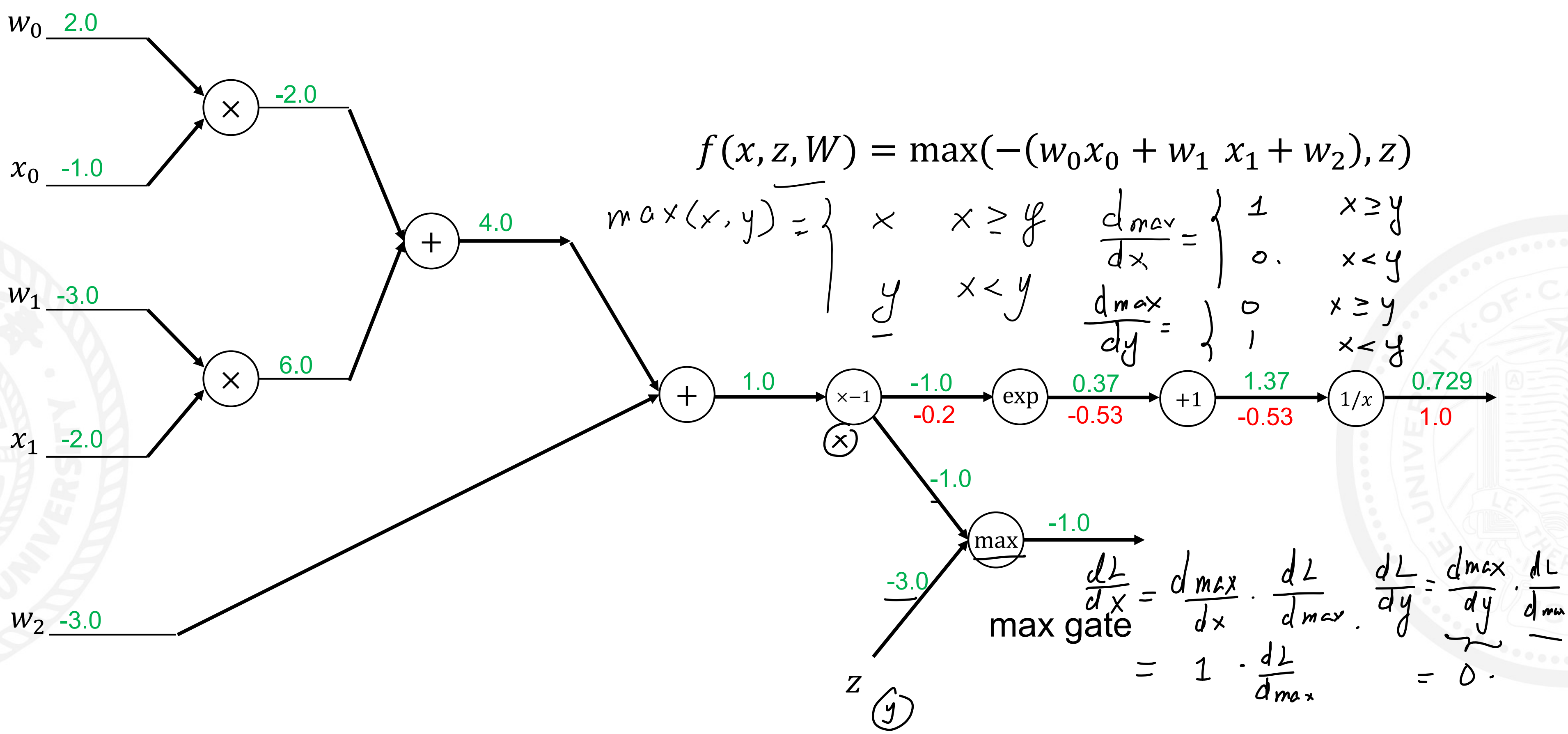
Backpropagation

$$f(x, z, W) = \max(-(w_0 x_0 + w_1\ x_1 + w_2), z)$$



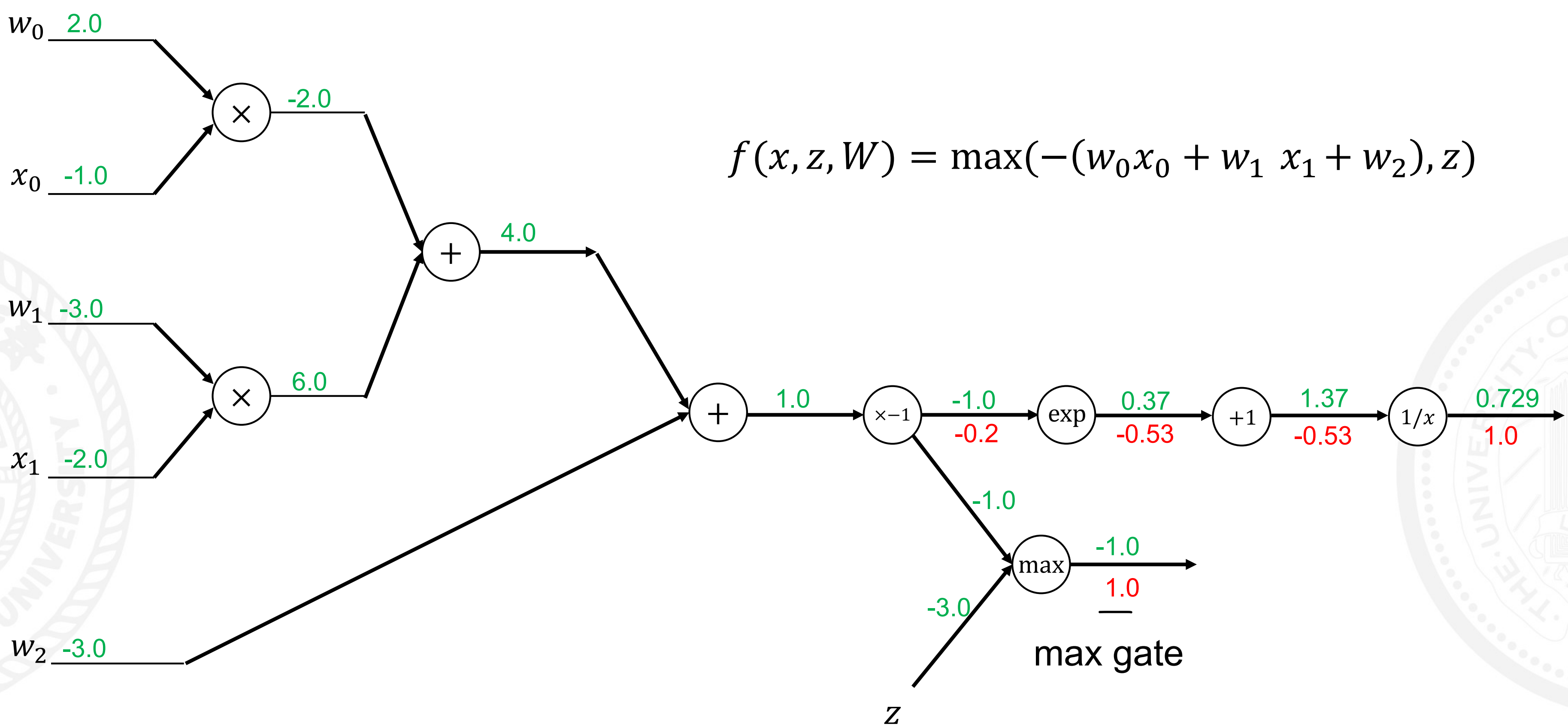
Backpropagation

$$f(x, z, W) = \max(-(w_0 x_0 + w_1 x_1 + w_2), z)$$



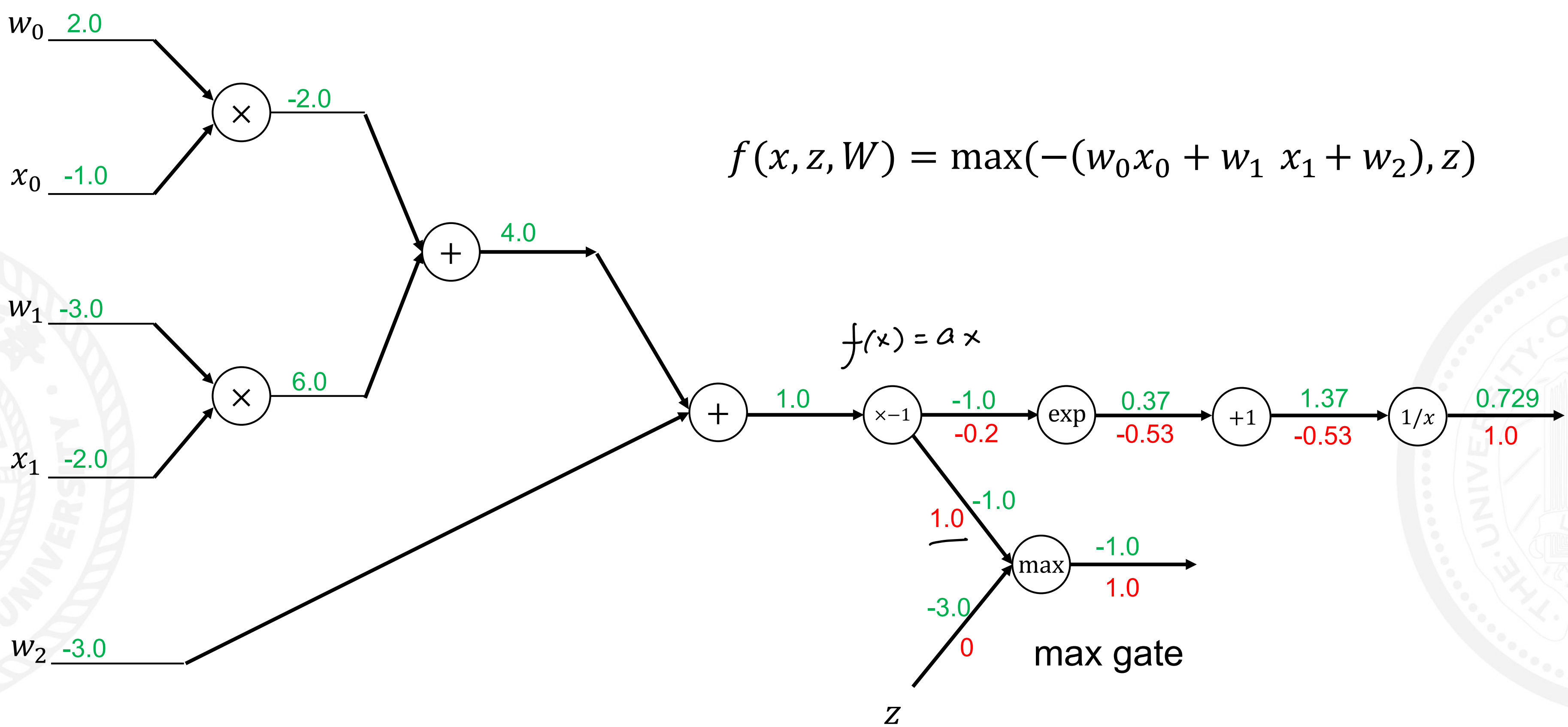
Backpropagation

$$f(x, z, W) = \max(-(w_0 x_0 + w_1\ x_1 + w_2), z)$$



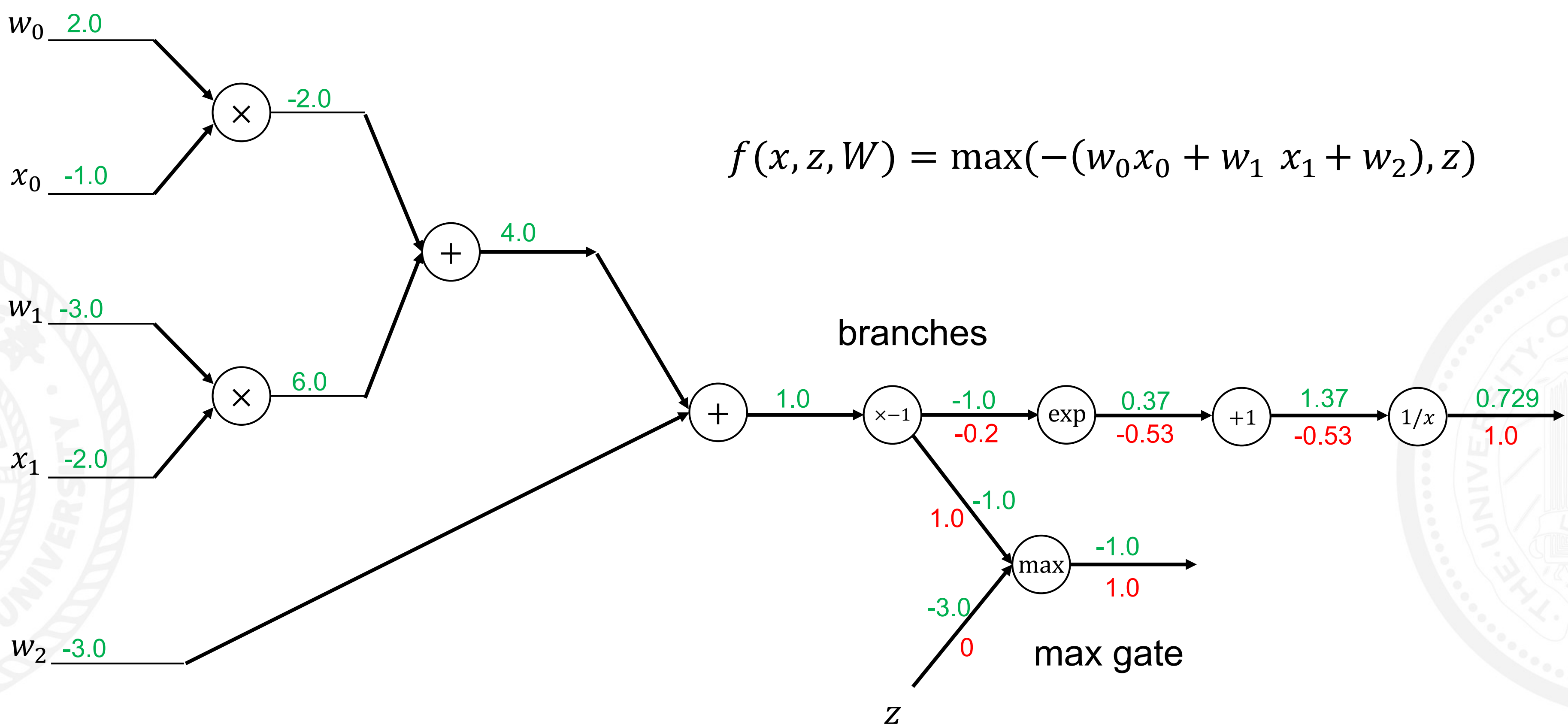
Backpropagation

$$f(x, z, W) = \max(-(w_0 x_0 + w_1\ x_1 + w_2), z)$$



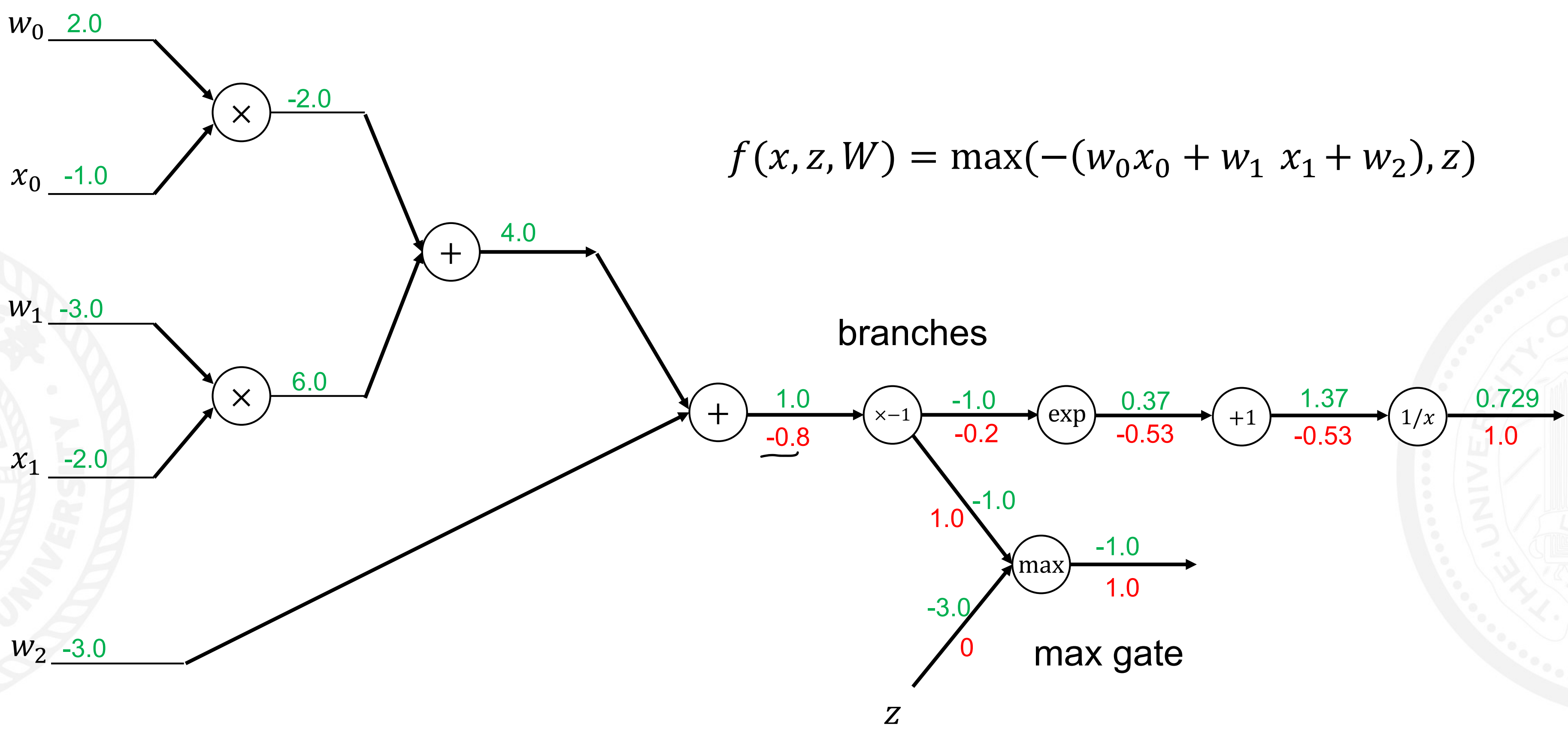
Backpropagation

$$f(x, z, W) = \max(-(w_0 x_0 + w_1\ x_1 + w_2), z)$$



Backpropagation

$$f(x, z, W) = \max(-(w_0 x_0 + w_1\ x_1 + w_2), z)$$



Backpropagation

Vectorized example

$$f(\mathbf{x}, \mathbf{W}) = \|\mathbf{W} \cdot \mathbf{x}\|^2$$

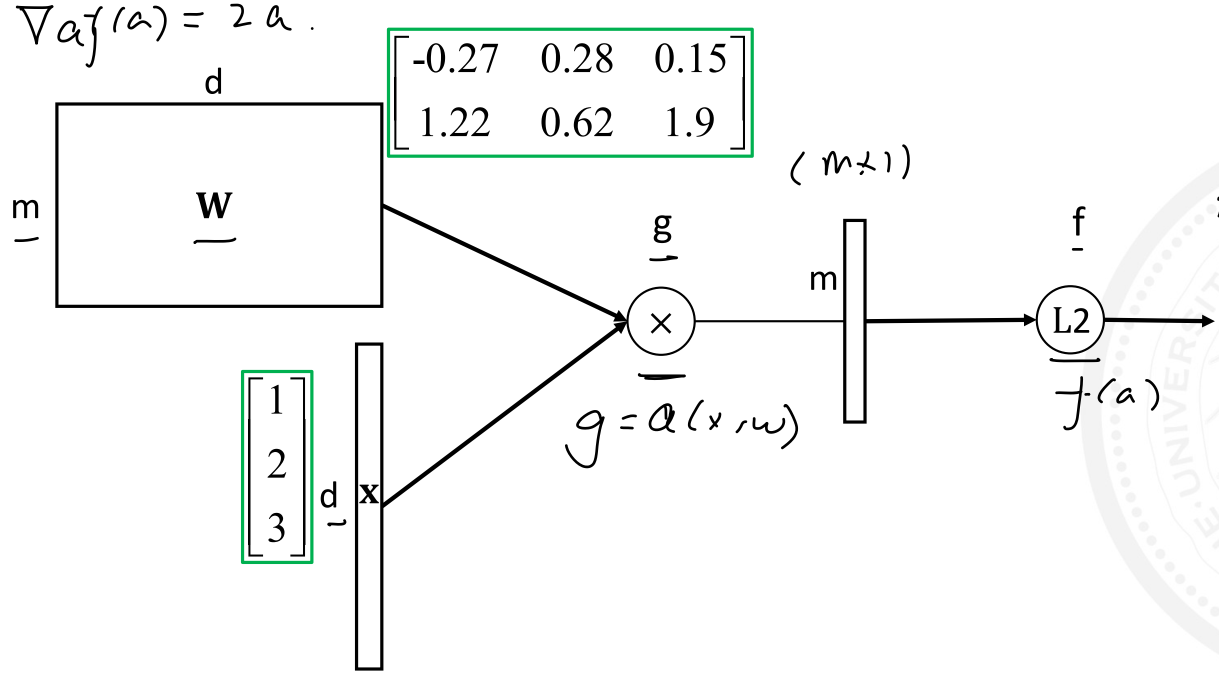
① $a(x, w) = Xw$

② $f(a) = \|a\|^2 = a^T a$

$\nabla_a f(a) = 2a$.

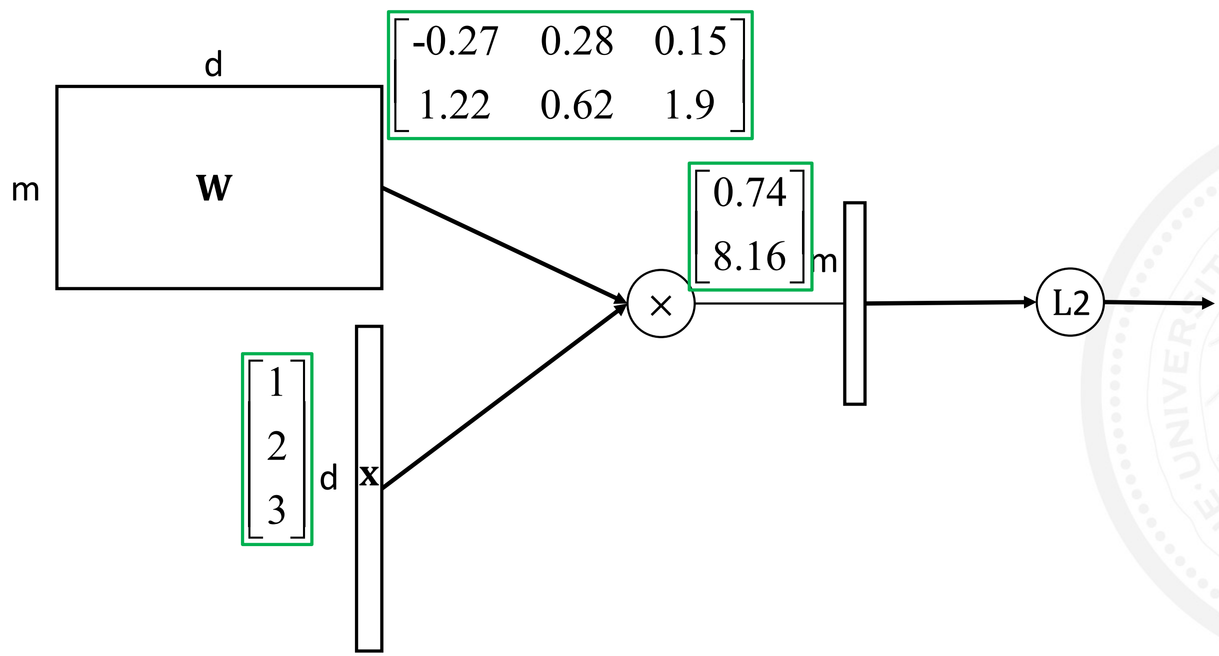
$\nabla_w a(x, w) = X^T$

$\nabla_x a(x, w) = W$



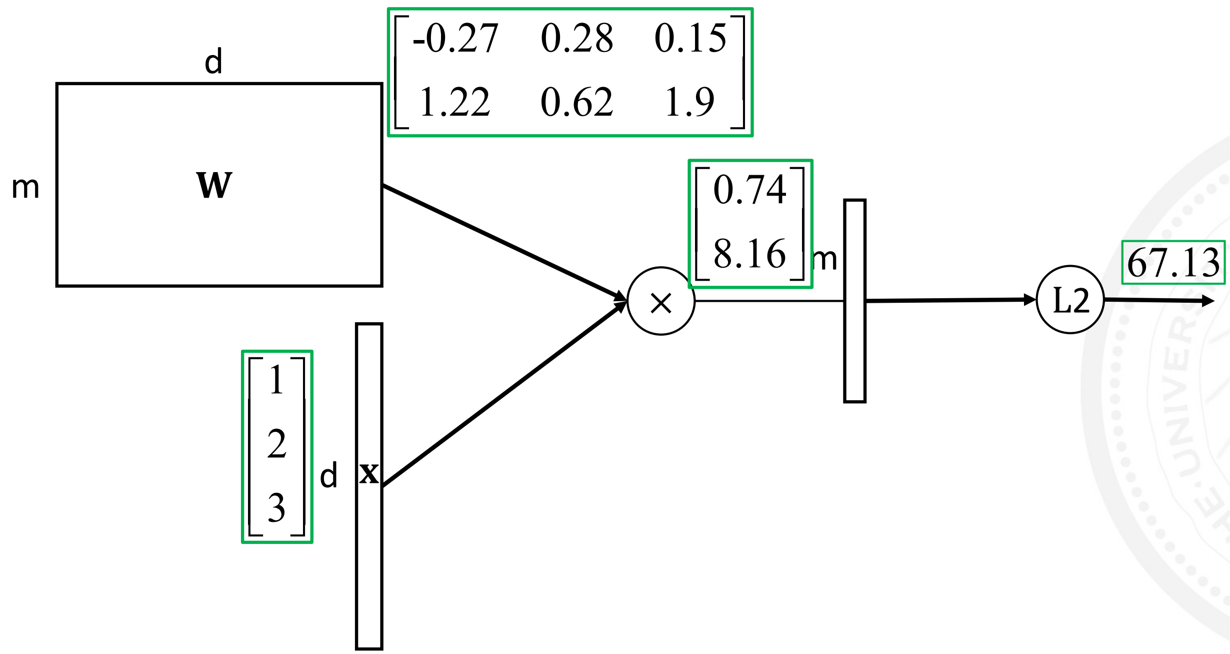
Vectorized example

$$f(\mathbf{x}, \mathbf{W}) = \|\mathbf{W} \cdot \mathbf{x}\|^2$$



Vectorized example

$$f(\mathbf{x}, \mathbf{W}) = \|\mathbf{W} \cdot \mathbf{x}\|^2$$



Backpropagation

Vectorized example

$$f(\mathbf{x}, \mathbf{W}) = \|\mathbf{W} \cdot \mathbf{x}\|^2$$

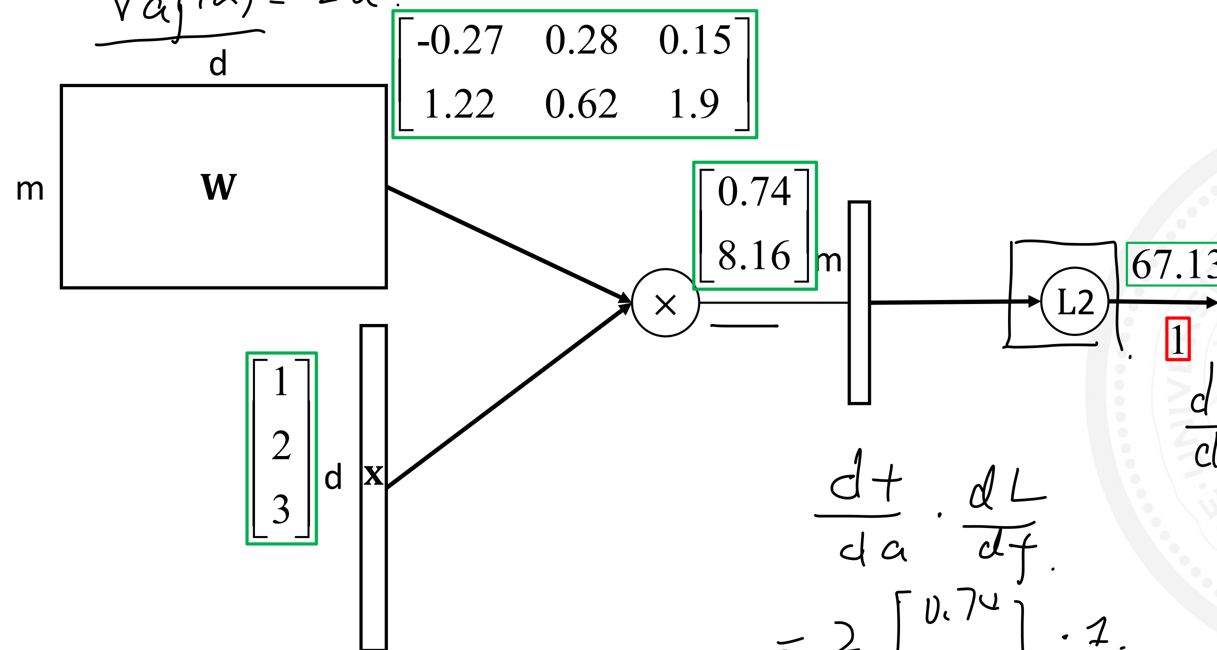
① $a(x, w) = Xw$

② $f(a) = \|a\|^2 = a^T a$

$\nabla_a f(a) = 2a$

$\nabla_w a(x, w) = X^T$

$\nabla_x a(x, w) = W$



$\frac{dL}{df} = 1$

$\frac{df}{da} \cdot \frac{dL}{df}$

$= 2 \begin{bmatrix} 0.74 \\ 8.16 \end{bmatrix} \cdot 1$

Backpropagation

Vectorized example

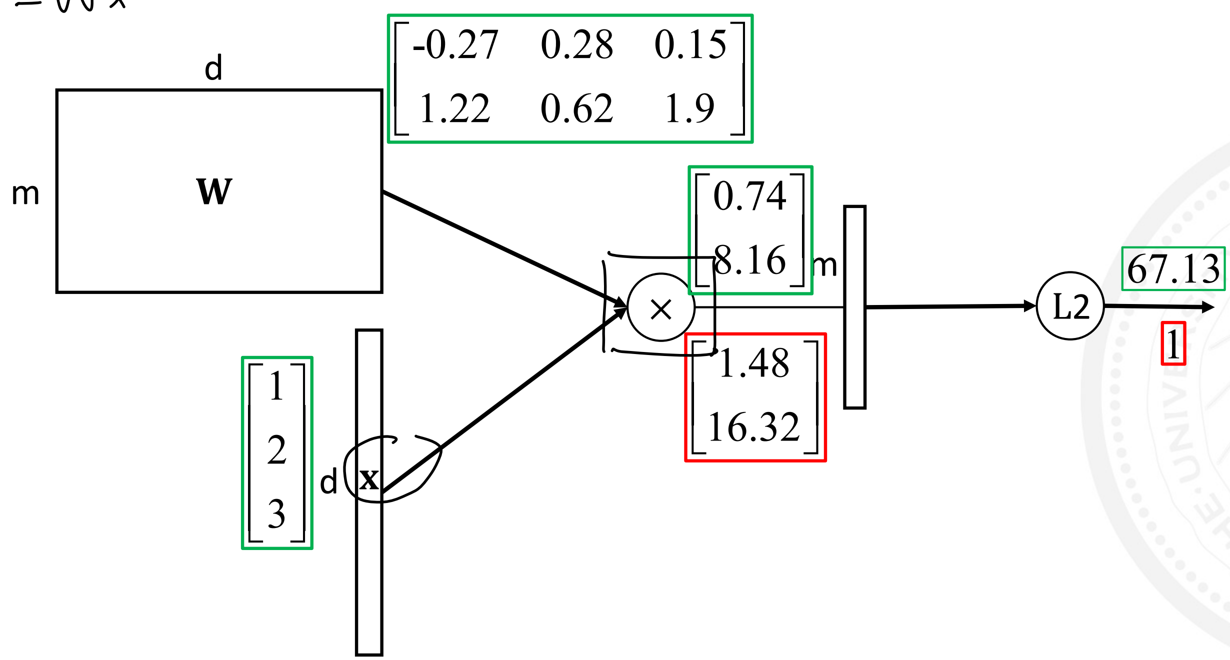
$$f(\mathbf{x}, \mathbf{W}) = \|\mathbf{W} \cdot \mathbf{x}\|^2$$

$$a(W, x) = Wx$$

$$\frac{df}{dw} = \left(\frac{df}{da}\right)\left(\frac{da}{dw}\right)$$

$$= \begin{pmatrix} 1.48 \\ 16.32 \end{pmatrix} x^T$$

$$= \begin{pmatrix} 1.48 \\ 16.32 \end{pmatrix} (1 \quad 2 \quad 3).$$



Backpropagation

Vectorized example

$$f(\mathbf{x}, \mathbf{W}) = \|\mathbf{W} \cdot \mathbf{x}\|^2$$