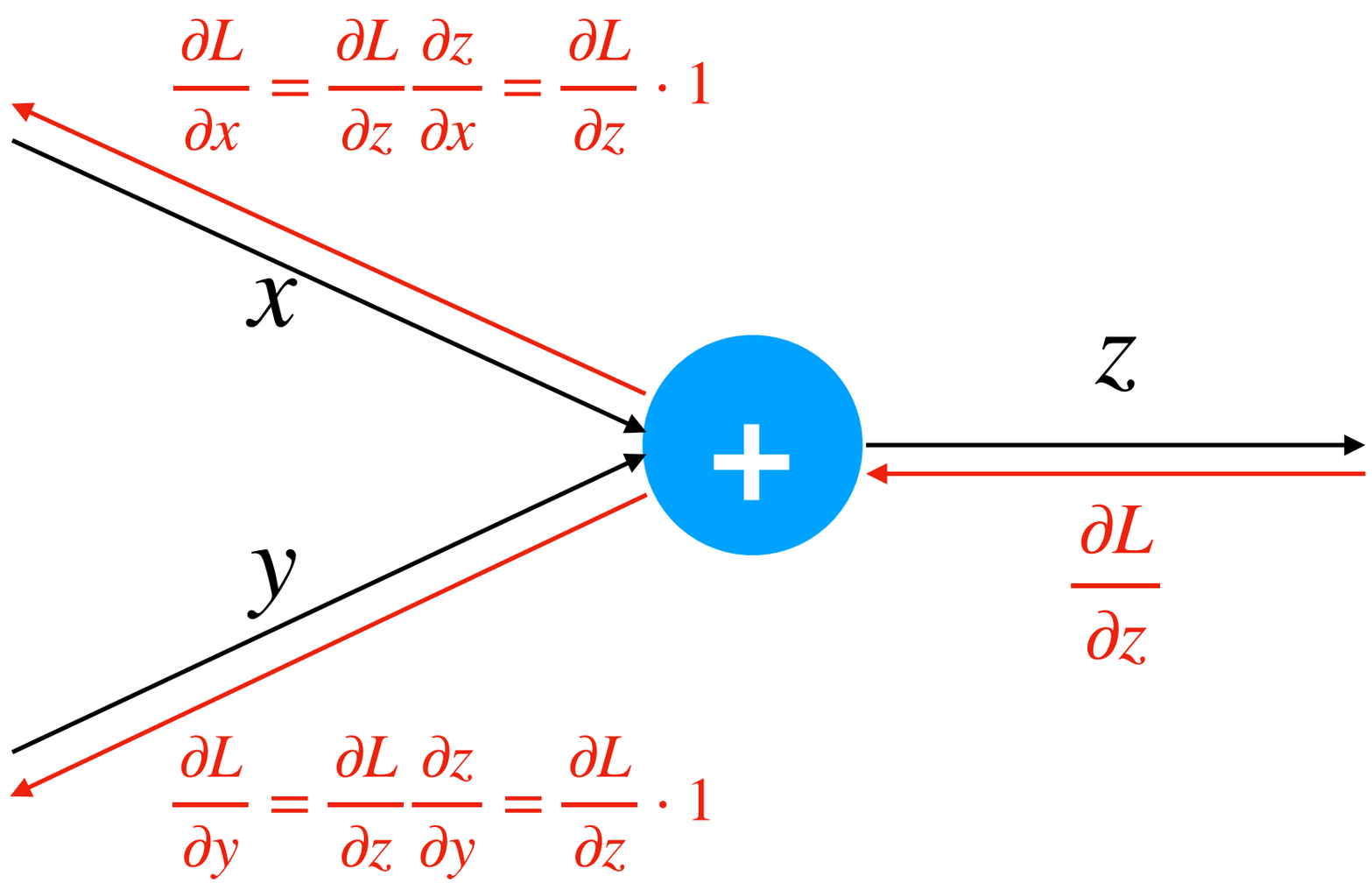
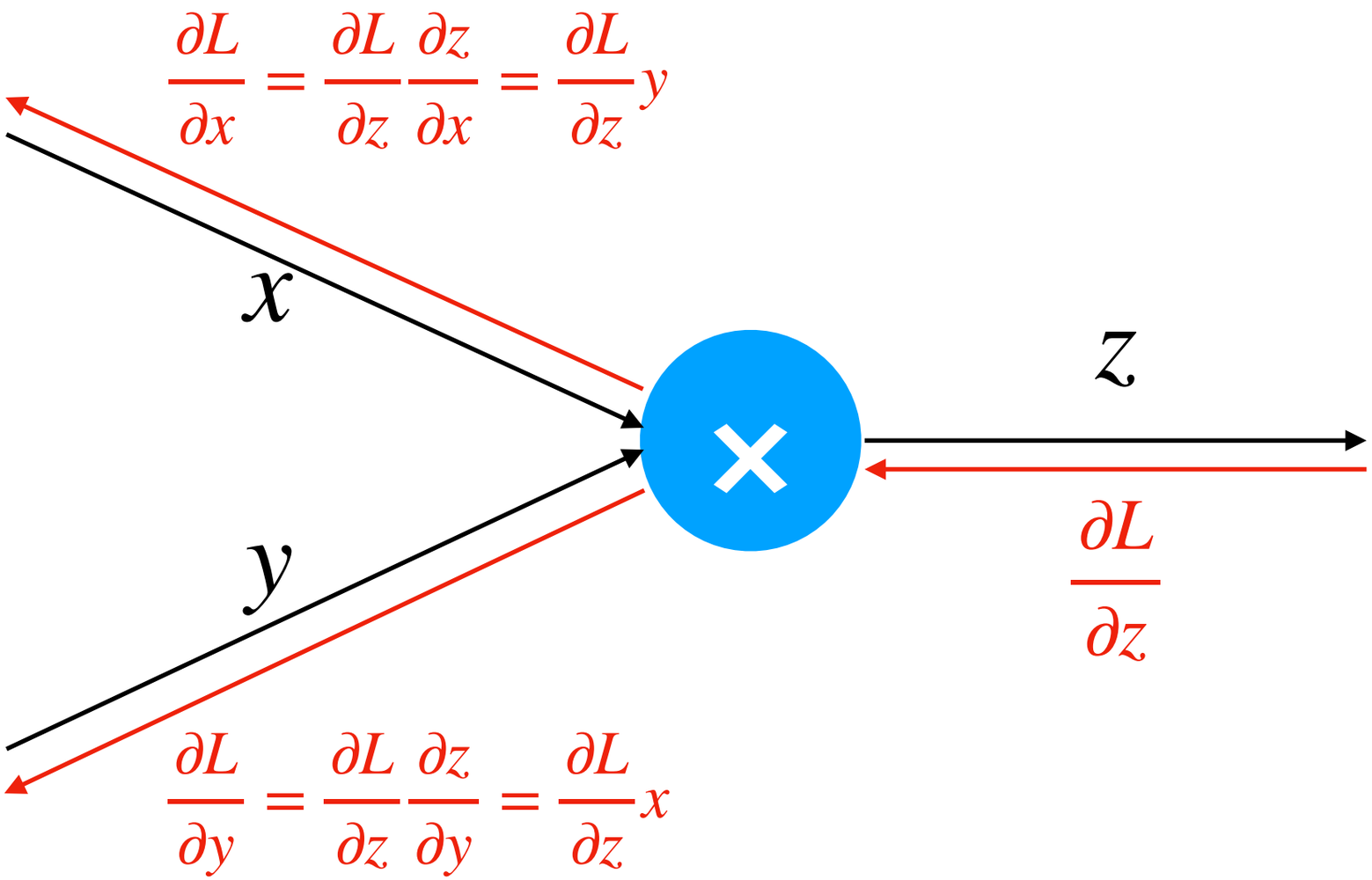
Let’s backprop

jojonki

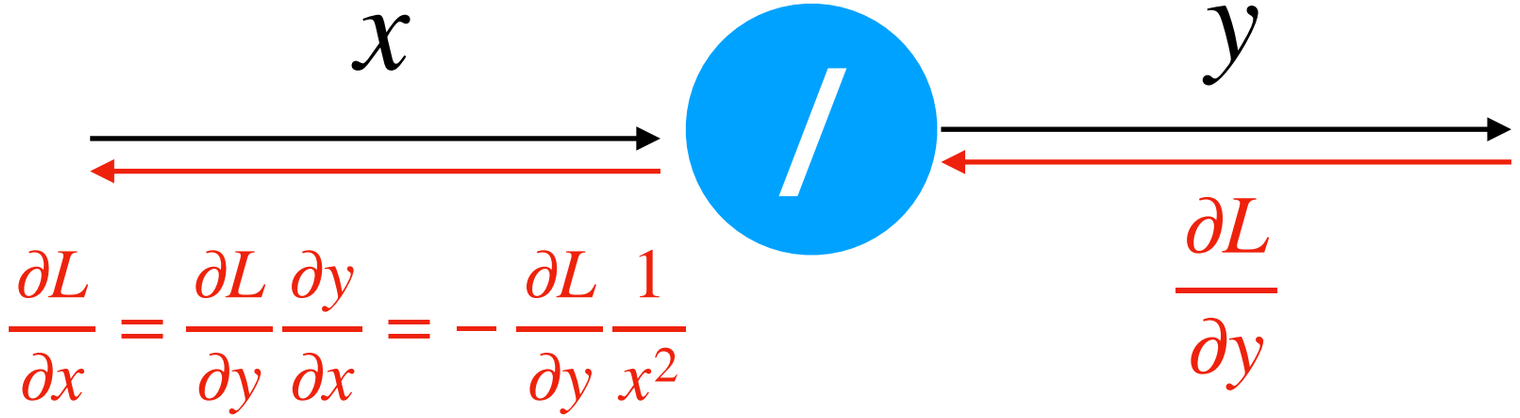
$$z = x + y$$


$\frac{\partial L}{\partial x} = \frac{\partial L}{\partial z}\frac{\partial z}{\partial x} = \frac{\partial L}{\partial z} \cdot 1$
x
y
+
z
$\frac{\partial L}{\partial z}$
$\frac{\partial L}{\partial y} = \frac{\partial L}{\partial z}\frac{\partial z}{\partial y} = \frac{\partial L}{\partial z} \cdot 1$

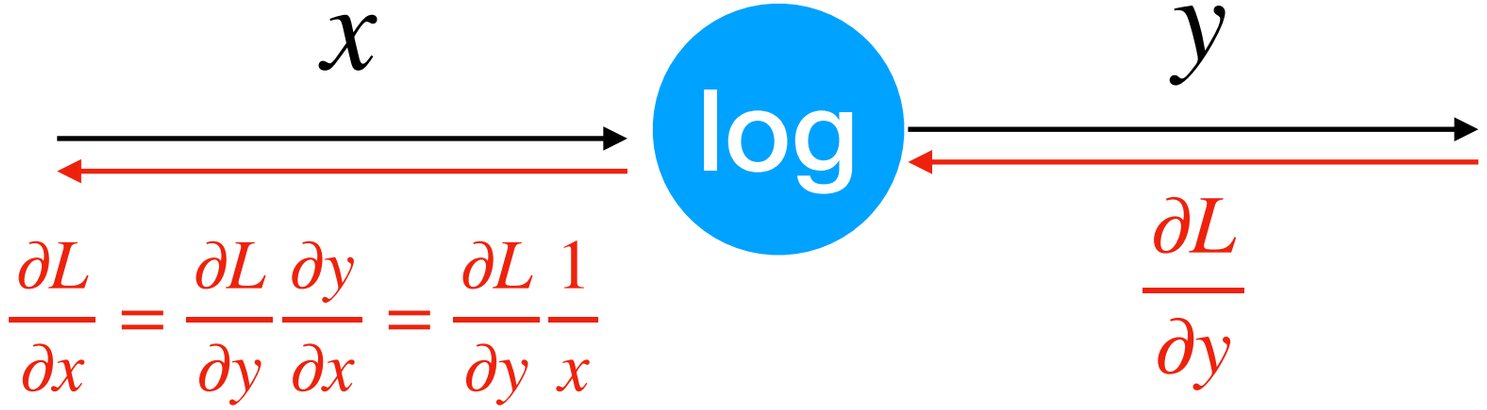
$$z = x \times y$$


$\frac{\partial L}{\partial x} = \frac{\partial L}{\partial z}\frac{\partial z}{\partial x} = \frac{\partial L}{\partial z}y$
x
y
z
$\frac{\partial L}{\partial z}$
$\frac{\partial L}{\partial y} = \frac{\partial L}{\partial z}\frac{\partial z}{\partial y} = \frac{\partial L}{\partial z}x$

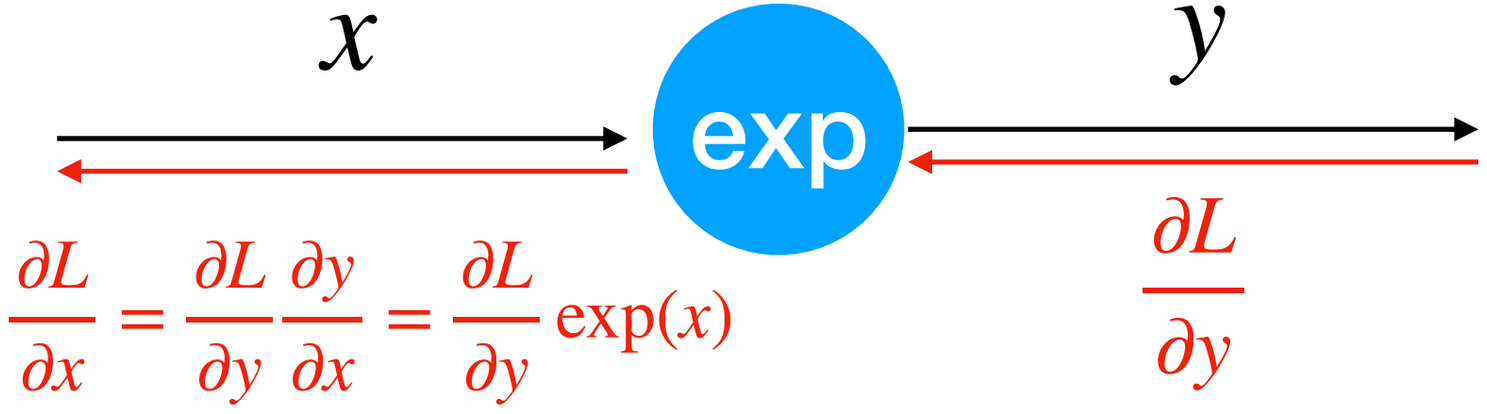
$$y = 1/x$$



$$y = \log(x)$$

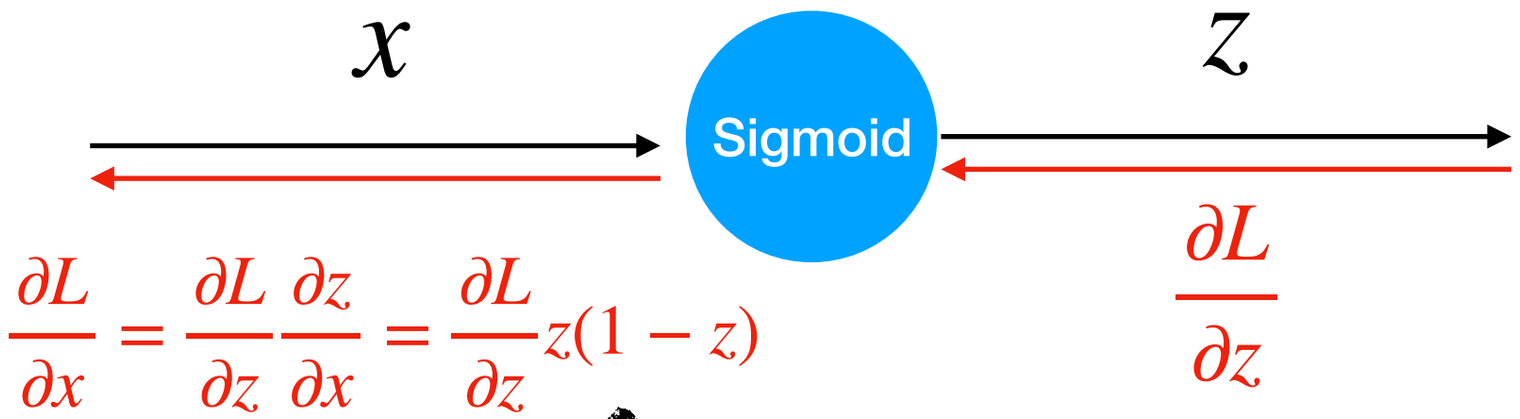

x
log
y
$\frac{\partial L}{\partial x} = \frac{\partial L}{\partial y}\frac{\partial y}{\partial x} = \frac{\partial L}{\partial y}\frac{1}{x}$
$\frac{\partial L}{\partial y}$

$$y = \exp(x)$$



Sigmoid

$$z = \frac{1}{1 + \exp(-x)}$$



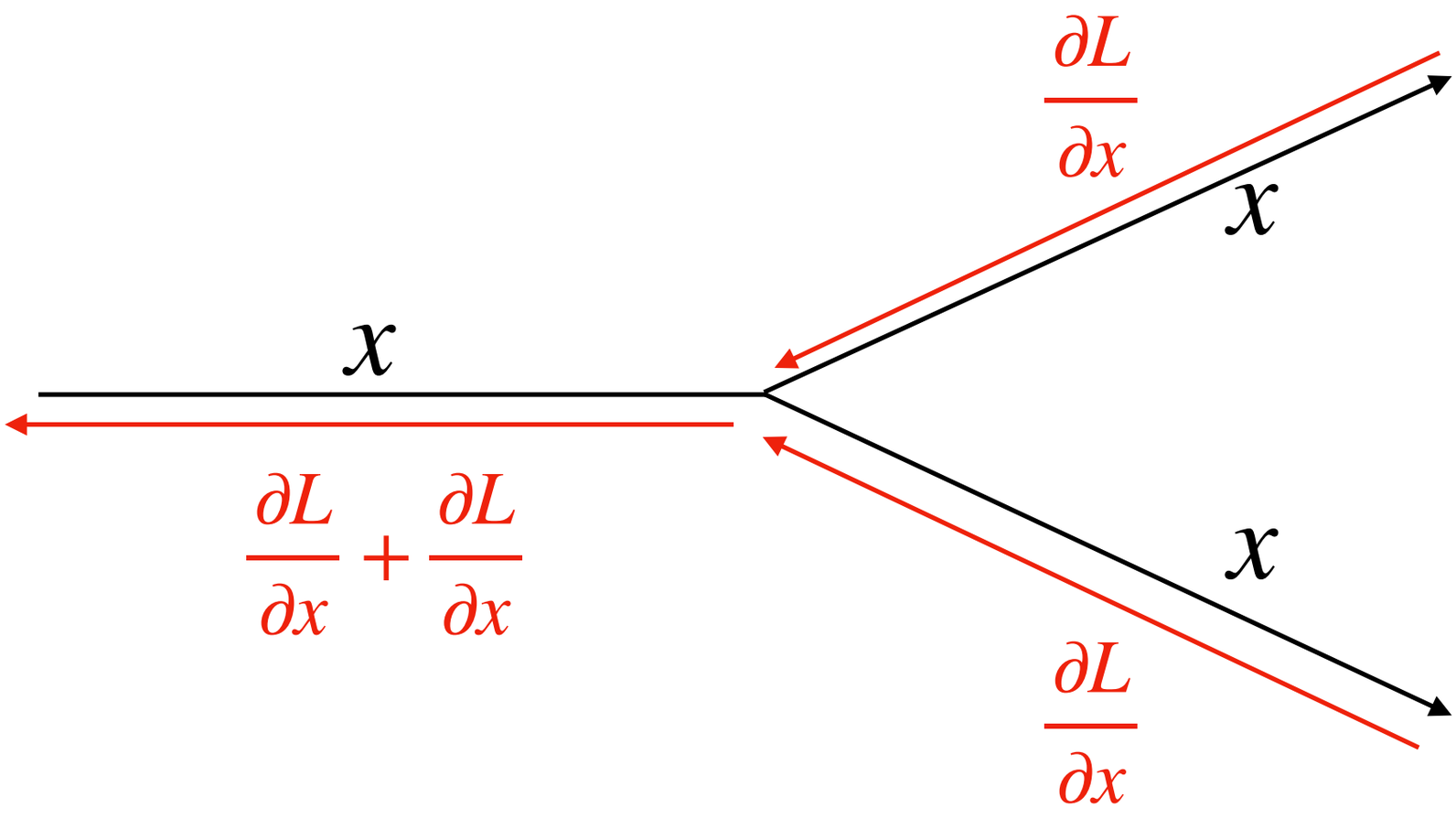
$$\left\{\frac{1}{1+\exp(-x)}\right\}' = \left\{(1+\exp(-x))^{-1}\right\}'$$

$$= -\frac{-\exp(-x)}{(1+\exp(-x))^2}$$

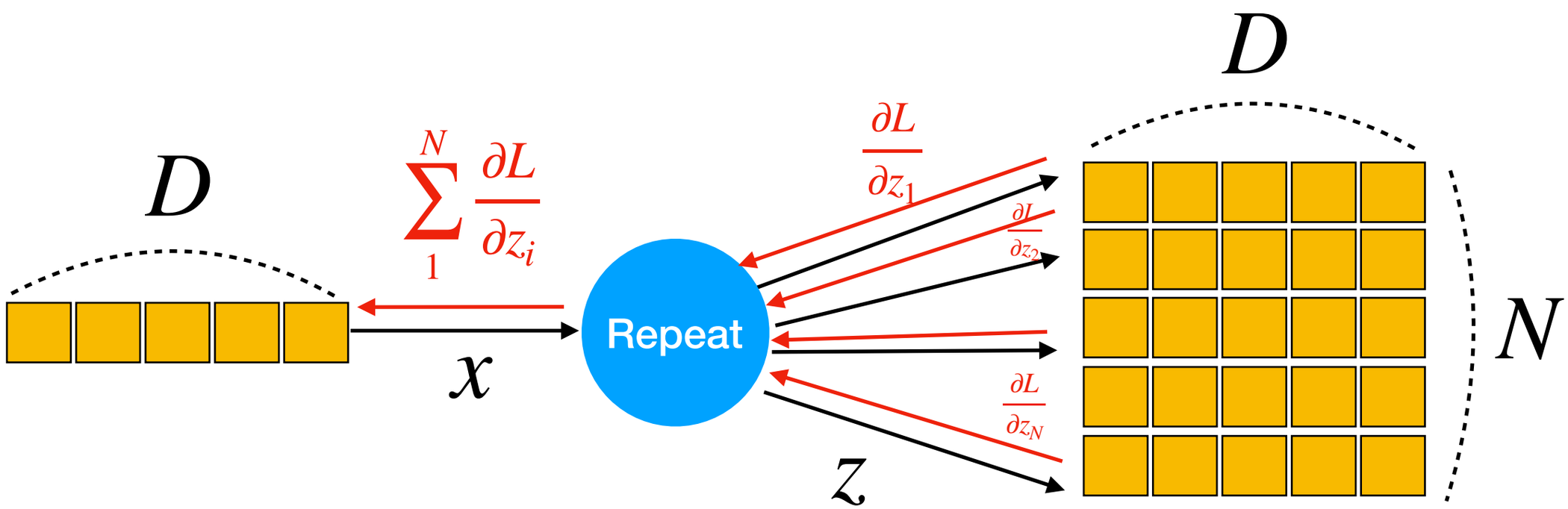
$$= z^2\frac{1-z}{z}$$

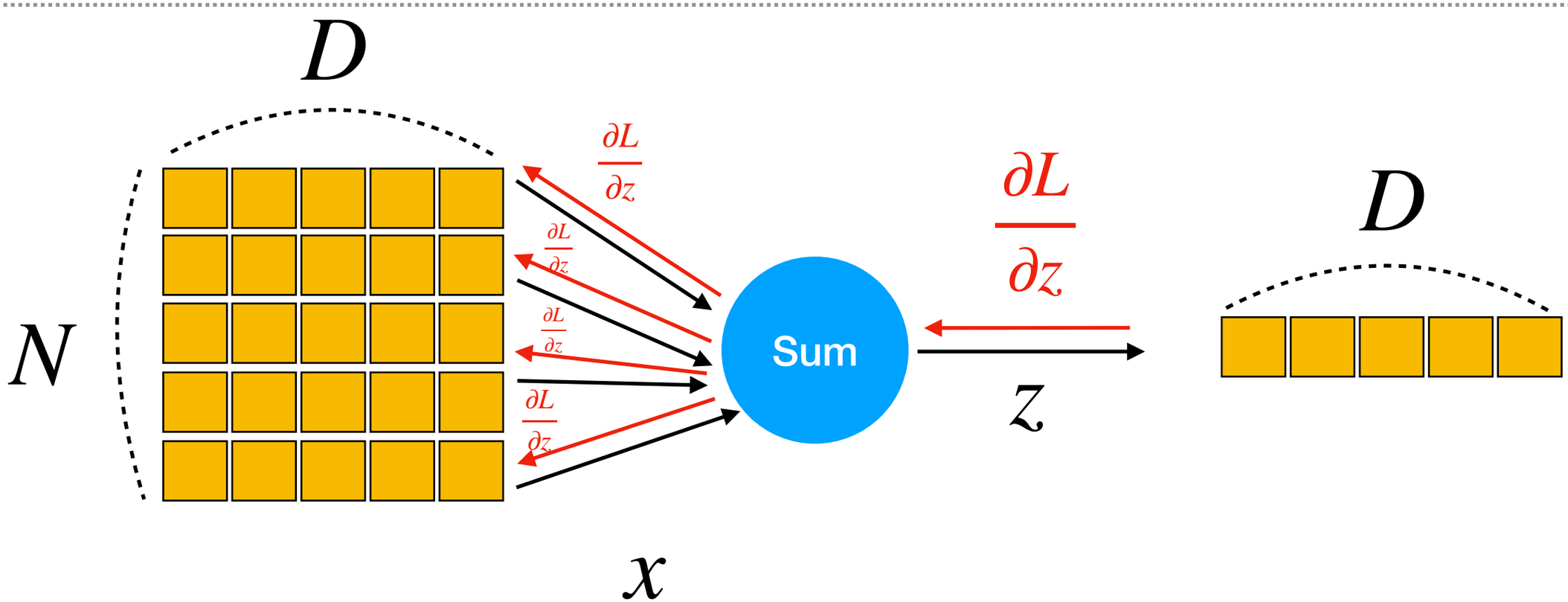
$$= z(1-z)$$

Branch


$\frac{\partial L}{\partial x}$
x
x
$\frac{\partial L}{\partial x} + \frac{\partial L}{\partial x}$
x
$\frac{\partial L}{\partial x}$

Repeat / Sum

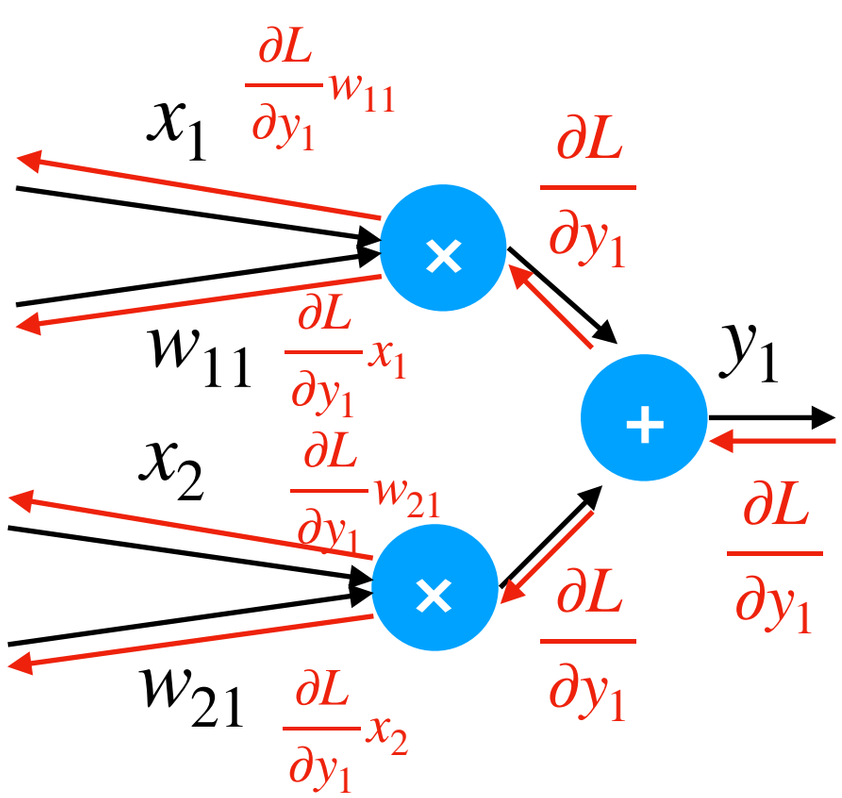

D
$\sum_1^N \frac{\partial L}{\partial z_i}$
x
Repeat
$\frac{\partial L}{\partial z_1}$
$\frac{\partial L}{\partial z_2}$
$\frac{\partial L}{\partial z_N}$
z
D
N


D
N
$\frac{\partial L}{\partial z}$
$\frac{\partial L}{\partial z}$
$\frac{\partial L}{\partial z}$
$\frac{\partial L}{\partial z}$
x
Sum
$\frac{\partial L}{\partial z}$
z
D

$y = xW$

$$y = (x_1, x_2)\begin{pmatrix} w_{11}, w_{12}, w_{13} \\ w_{21}, w_{22}, w_{23} \end{pmatrix}$$

$$= (x_1 w_{11} + x_2 w_{21}, x_1 w_{12} + x_2 w_{22}, x_1 w_{13} + x_2 w_{23})$$



$$\frac{\partial L}{\partial x_i} = \sum_j \frac{\partial L}{y_j} w_{ij}$$

$$= \left(\frac{\partial L}{\partial y_1}, \frac{\partial L}{\partial y_2}, \frac{\partial L}{\partial y_3}\right)\begin{pmatrix} w_{i1} \\ w_{i2} \\ w_{i3} \end{pmatrix}$$

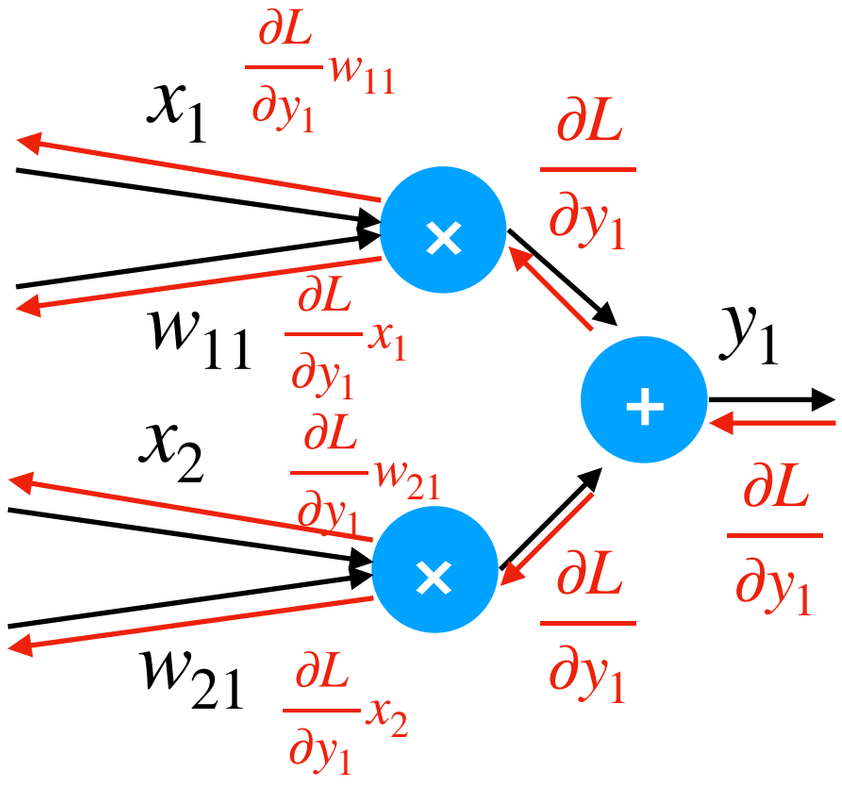
$$\frac{\partial L}{\partial x} = \left(\frac{\partial L}{\partial y_1}, \frac{\partial L}{\partial y_2}, \frac{\partial L}{\partial y_3}\right)\begin{pmatrix} w_{11}, w_{21} \\ w_{12}, w_{22} \\ w_{13}, w_{23} \end{pmatrix}$$

$$= \frac{\partial L}{\partial y} w^T$$

$y = xW$

$$y = (x_1, x_2)\begin{pmatrix} w_{11}, w_{12}, w_{13} \\ w_{21}, w_{22}, w_{23} \end{pmatrix}$$

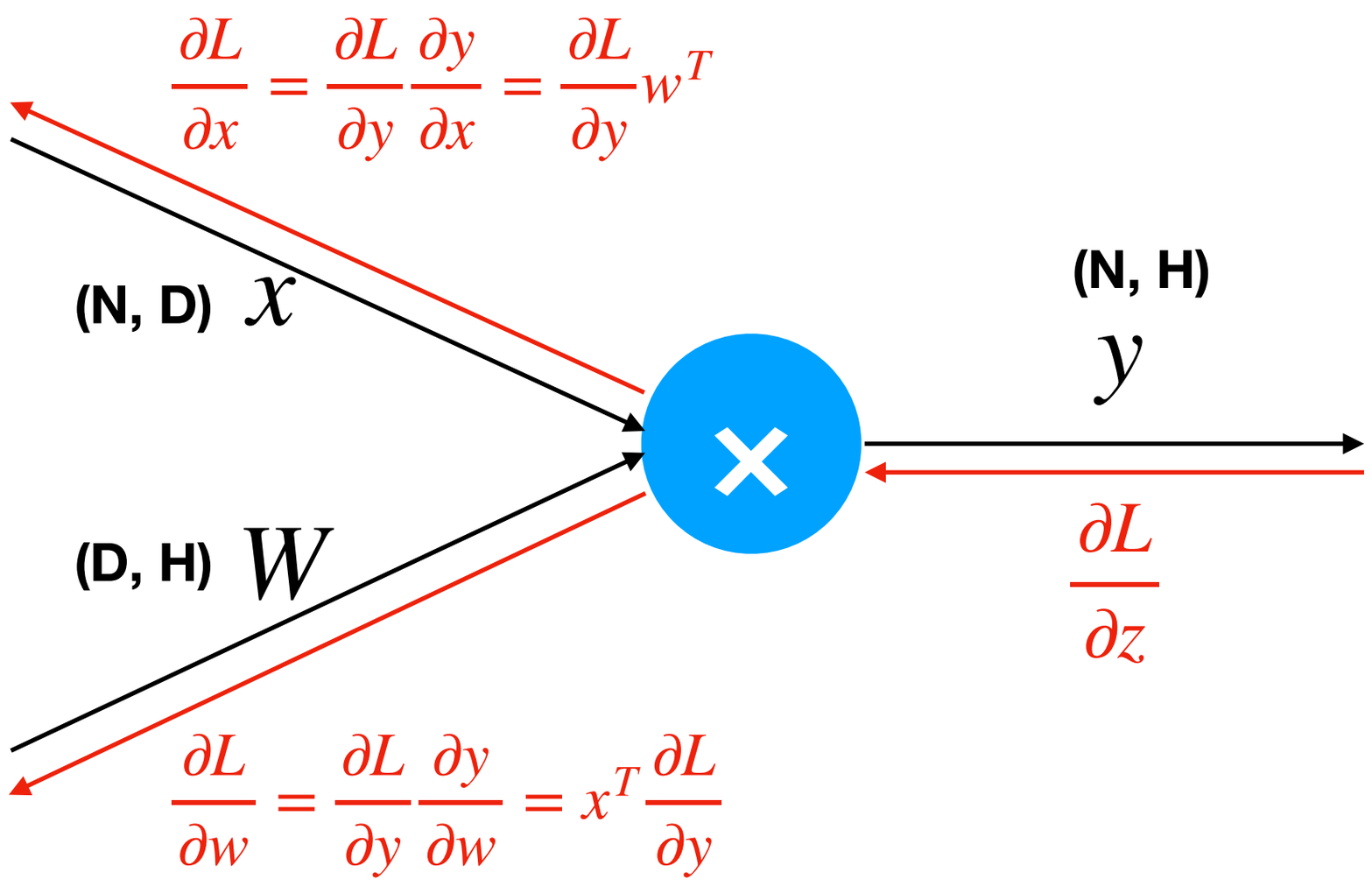
$$= (x_1 w_{11} + x_2 w_{21}, x_1 w_{12} + x_2 w_{22}, x_1 w_{13} + x_2 w_{23})$$



$$\begin{pmatrix} \frac{\partial L}{\partial w_{11}}, \frac{\partial L}{\partial w_{12}}, \frac{\partial L}{\partial w_{13}} \\ \frac{\partial L}{\partial w_{21}}, \frac{\partial L}{\partial w_{22}}, \frac{\partial L}{\partial w_{23}} \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \left(\frac{\partial L}{\partial y_1}, \frac{\partial L}{\partial y_2}, \frac{\partial L}{\partial y_3} \right)$$

$$\frac{\partial L}{\partial w} = x^T \frac{\partial L}{\partial y}$$

$y = xW$ 3/3


$\frac{\partial L}{\partial x} = \frac{\partial L}{\partial y}\frac{\partial y}{\partial x} = \frac{\partial L}{\partial y}w^T$
(N, D) x
(D, H) W
(N, H)
y
$\frac{\partial L}{\partial z}$
$\frac{\partial L}{\partial w} = \frac{\partial L}{\partial y}\frac{\partial y}{\partial w} = x^T\frac{\partial L}{\partial y}$

Softmax with Loss