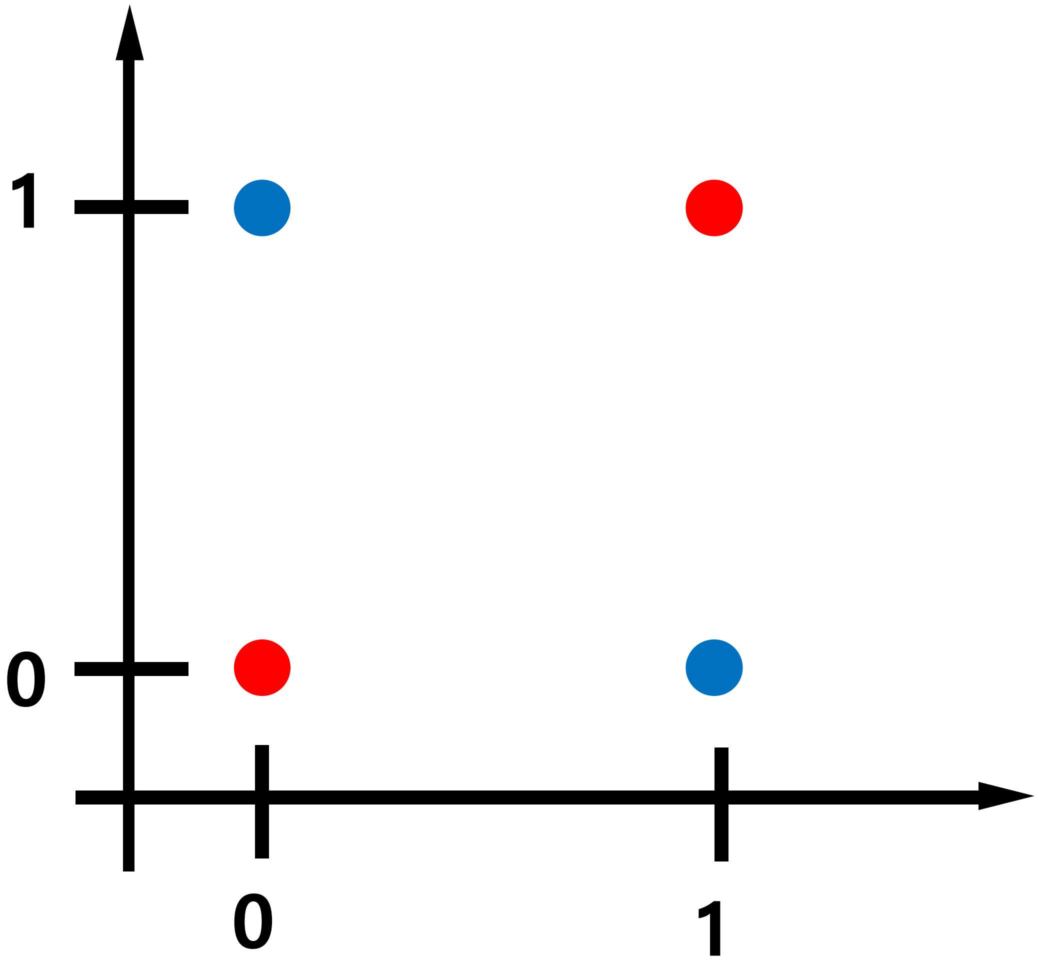
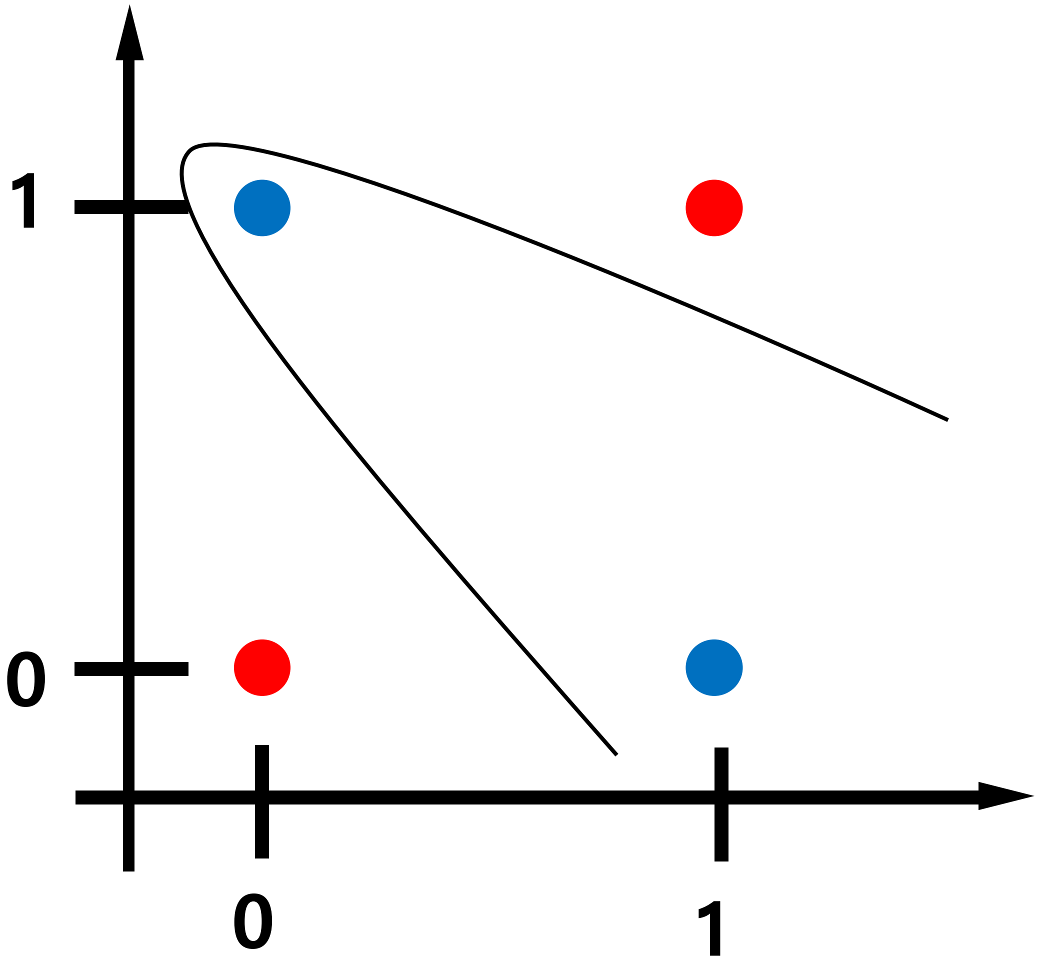
NeuralNet 101

5. Error Back Propagation

We have a problem..



We have a problem..



Add complexity on neural network

$\Rightarrow$ Let F(X) = WX + b

$\Rightarrow$ Let s(x) = sigmoid(x)

$\Rightarrow$ Then, we can set new complex function as
G(x) = s(F(s(F(s(F(x))))))

We might be able to solve the problem,
but we cannot define loss function

What is Backpropagation?

=> Another way of Differentiation

Why Backpropagation?

There is an Error in Numerical differentials that the computer progresses

Why Backpropagation?

There is an Error in Numerical differentials that the computer progresses

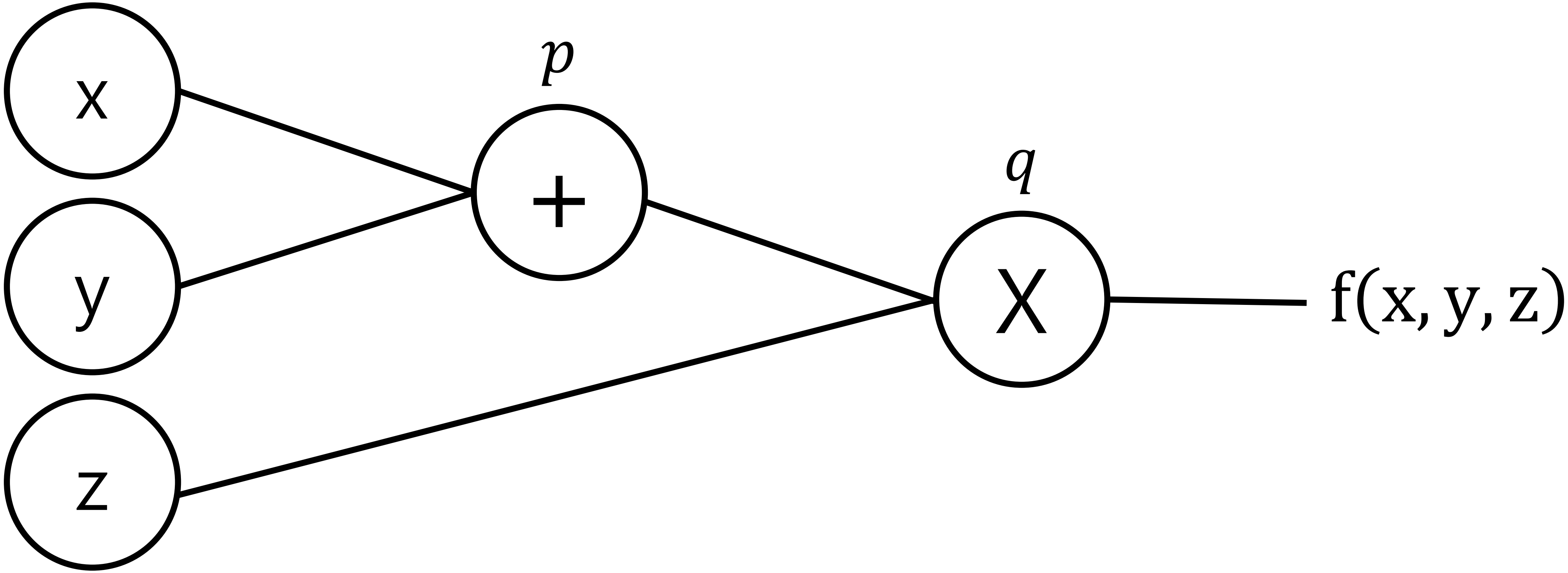
Split into several small, definite differentials with chain rules

Backpropagation - Basic

$\mathrm{f}(\mathrm{x}, \mathrm{y}, \mathrm{z}) = ((x + y)z) \quad x = 1, y = 5, z = -3$

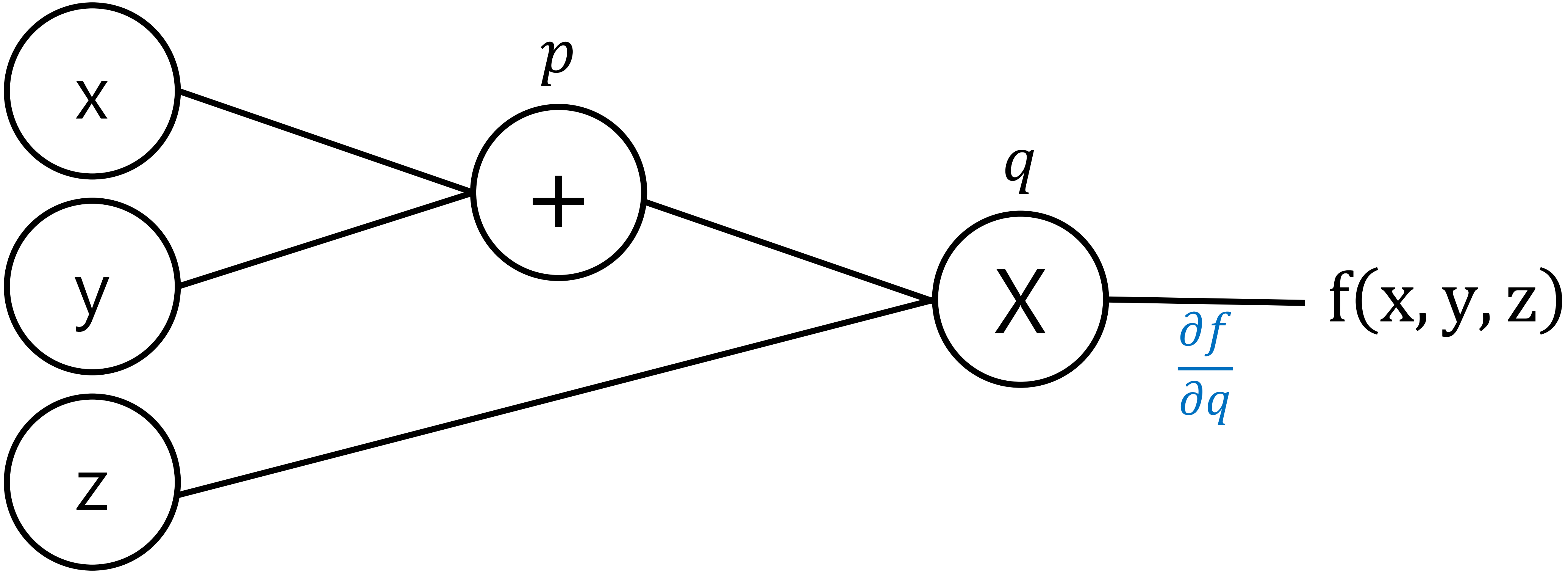
Backpropagation - Basic

$\mathrm{f(x, y, z)} = ((x + y)z) \quad x = 1, y = 5, z = -3$



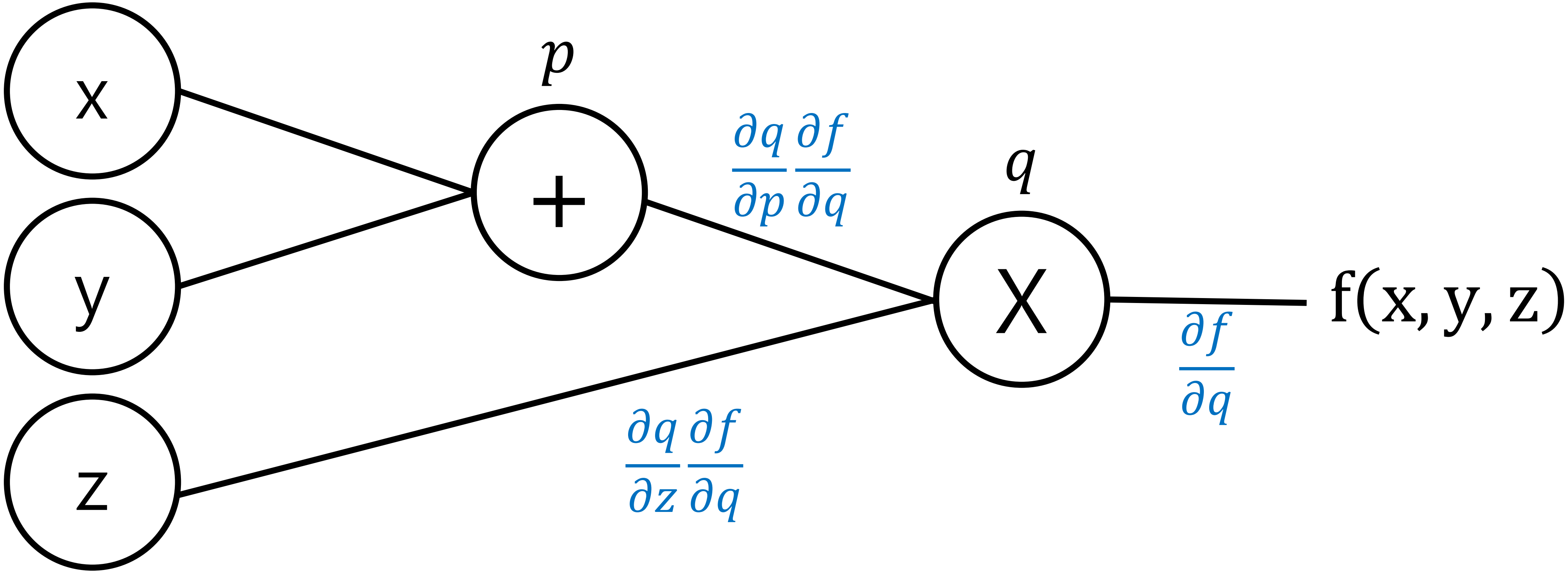
Backpropagation - Basic

$\mathrm{f(x, y, z)} = ((x + y)z) \quad x = 1, y = 5, z = -3$



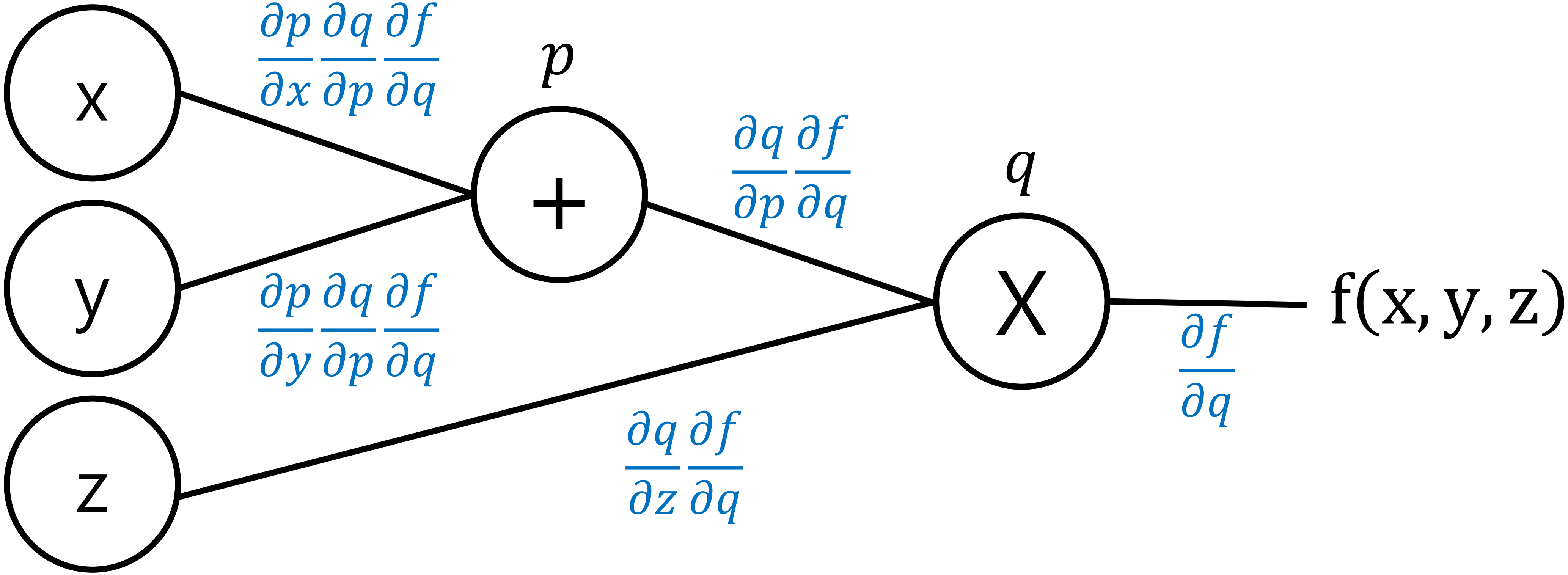
Backpropagation - Basic

$\mathrm{f(x,y,z)} = ((x+y)z) \quad x = 1, y = 5, z = -3$


p
x
+
y
$\frac{\partial q}{\partial p}\frac{\partial f}{\partial q}$
q
X
$\mathrm{f(x,y,z)}$
$\frac{\partial f}{\partial q}$
z
$\frac{\partial q}{\partial z}\frac{\partial f}{\partial q}$

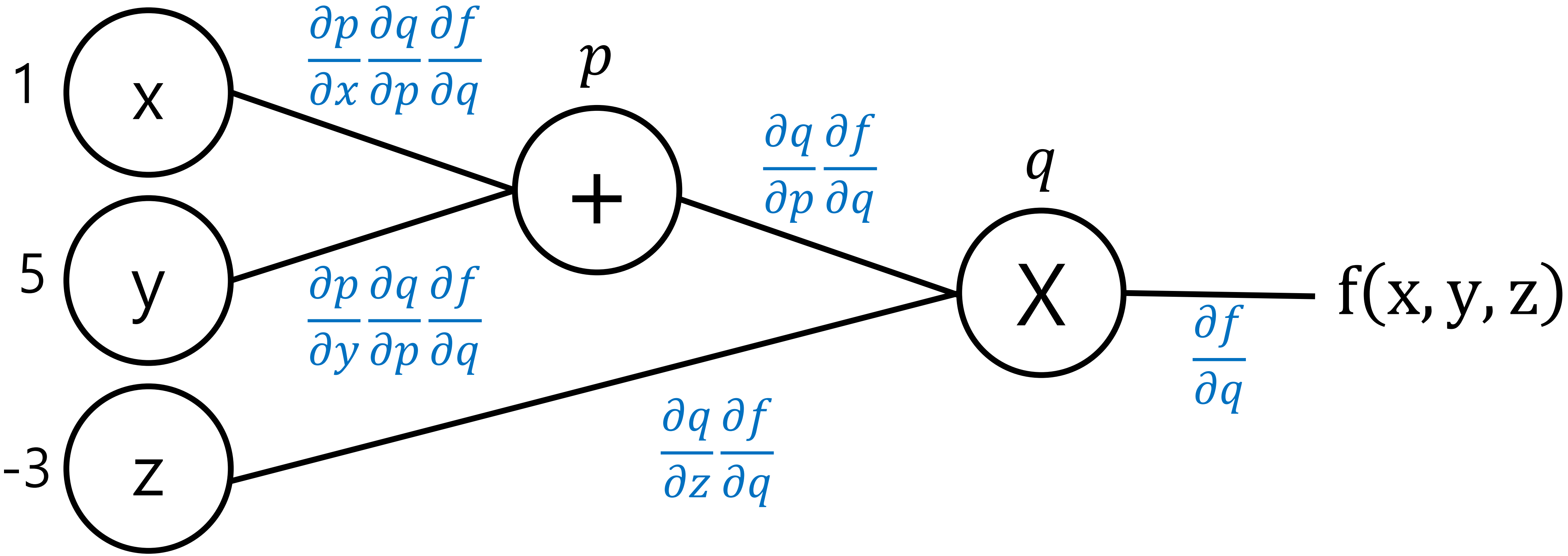
Backpropagation - Basic

$f(\mathrm{x,y,z}) = ((x+y)z) \quad x = 1, y = 5, z = -3$


$\frac{\partial p}{\partial x}\frac{\partial q}{\partial p}\frac{\partial f}{\partial q}$
p
x
+
$\frac{\partial q}{\partial p}\frac{\partial f}{\partial q}$
q
y
X
f(x, y, z)
$\frac{\partial p}{\partial y}\frac{\partial q}{\partial p}\frac{\partial f}{\partial q}$
$\frac{\partial f}{\partial q}$
z
$\frac{\partial q}{\partial z}\frac{\partial f}{\partial q}$

Backpropagation - Basic

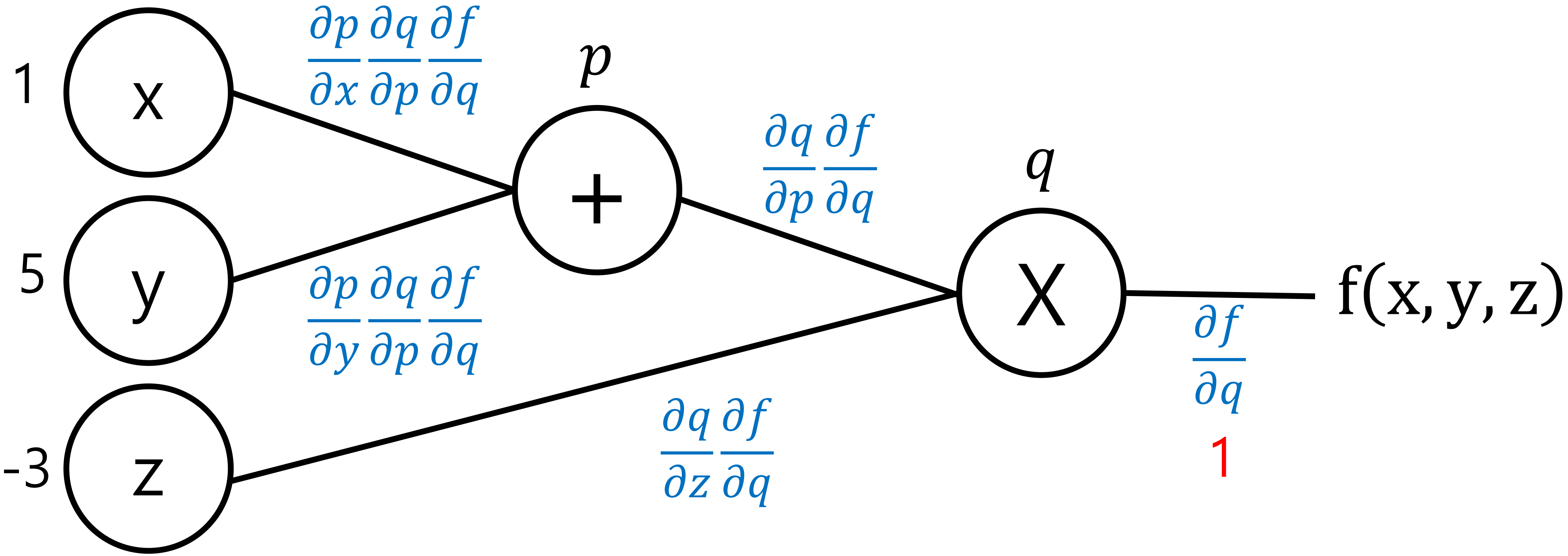
$f(x, y, z) = ((x + y)z) \quad x = 1, y = 5, z = -3$



1
x
5
y
-3
z
$\frac{\partial p}{\partial x}\frac{\partial q}{\partial p}\frac{\partial f}{\partial q}$
$\frac{\partial p}{\partial y}\frac{\partial q}{\partial p}\frac{\partial f}{\partial q}$
p
+
$\frac{\partial q}{\partial p}\frac{\partial f}{\partial q}$
$\frac{\partial q}{\partial z}\frac{\partial f}{\partial q}$
q
X
$\frac{\partial f}{\partial q}$
f(x, y, z)

Backpropagation - Basic

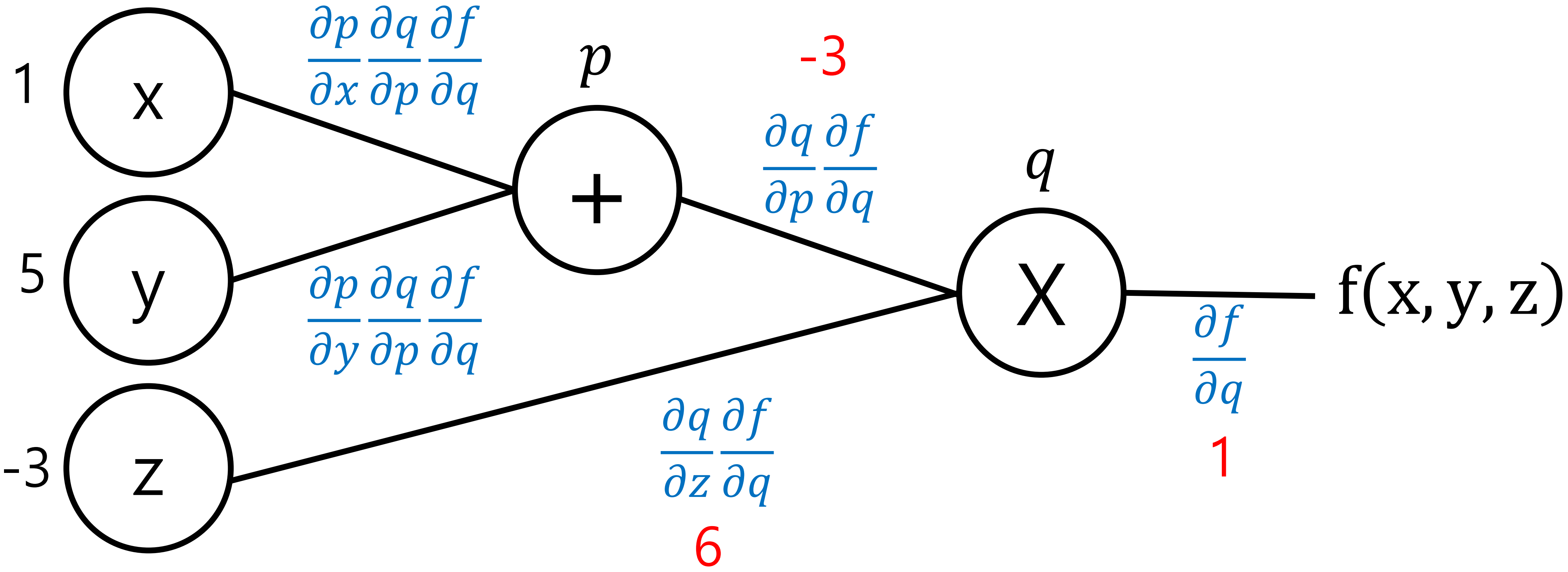
$\mathrm{f}(\mathrm{x},\mathrm{y},\mathrm{z}) = ((x+y)z) \qquad x = 1, y = 5, z = -3$



1
5
-3
x
y
z
p
q
+
X
f(x, y, z)
$\frac{\partial p}{\partial x}\frac{\partial q}{\partial p}\frac{\partial f}{\partial q}$
$\frac{\partial p}{\partial y}\frac{\partial q}{\partial p}\frac{\partial f}{\partial q}$
$\frac{\partial q}{\partial p}\frac{\partial f}{\partial q}$
$\frac{\partial q}{\partial z}\frac{\partial f}{\partial q}$
$\frac{\partial f}{\partial q}$
1

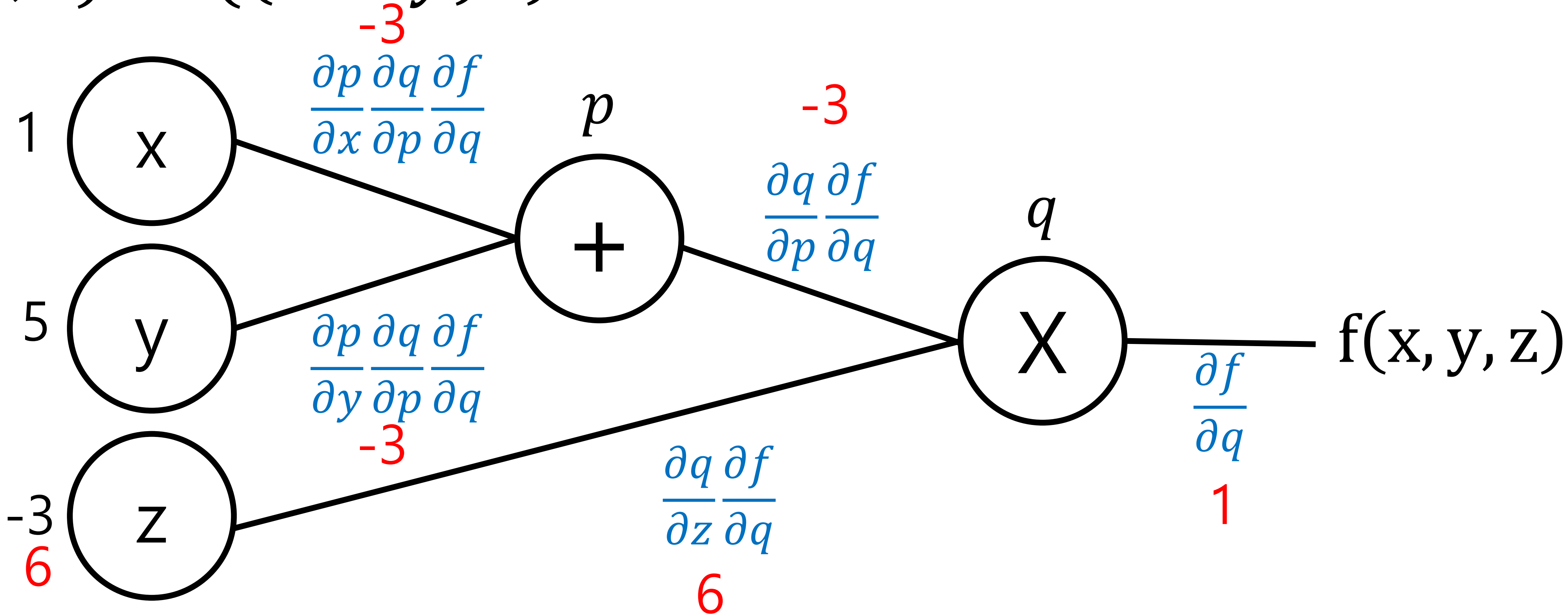
Backpropagation - Basic

$$f(x, y, z) = ((x + y)z) \qquad x = 1, y = 5, z = -3$$


1
x
5
y
-3
z
$\frac{\partial p}{\partial x}\frac{\partial q}{\partial p}\frac{\partial f}{\partial q}$
$\frac{\partial p}{\partial y}\frac{\partial q}{\partial p}\frac{\partial f}{\partial q}$
p
+
-3
$\frac{\partial q}{\partial p}\frac{\partial f}{\partial q}$
q
X
$\frac{\partial q}{\partial z}\frac{\partial f}{\partial q}$
6
f(x, y, z)
$\frac{\partial f}{\partial q}$
1

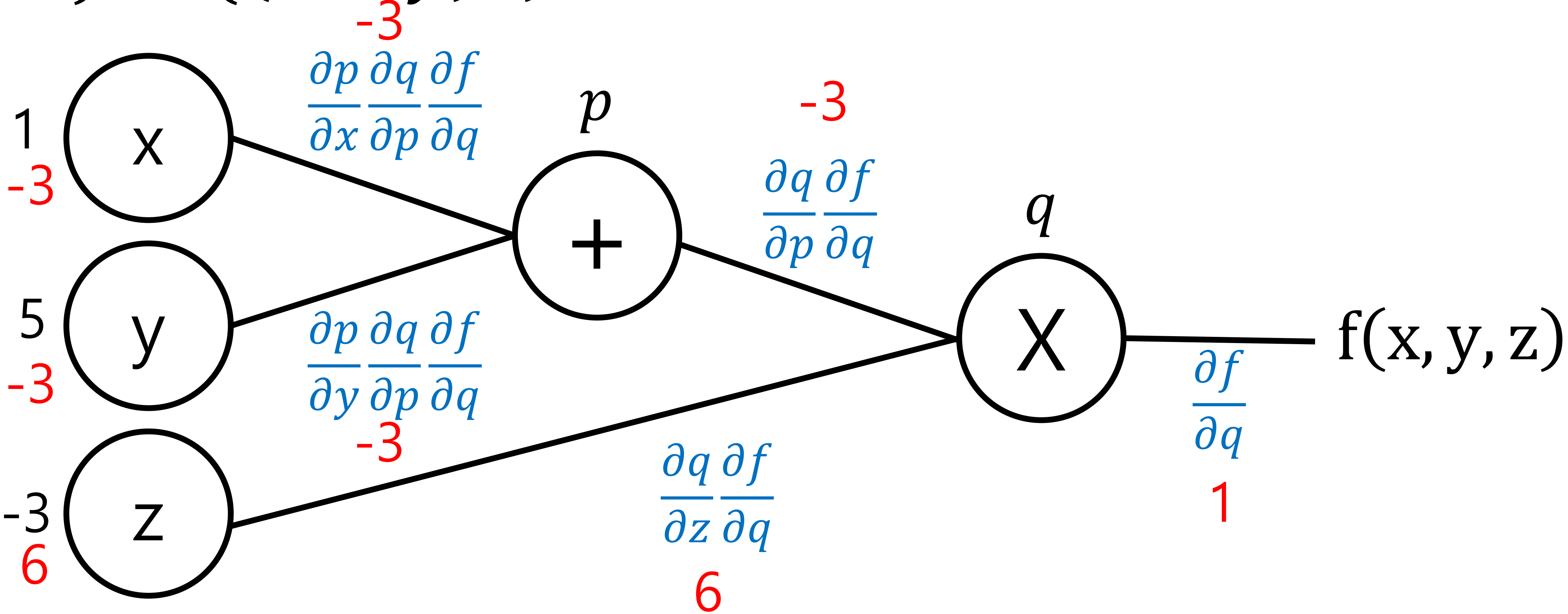
Backpropagation - Basic

$f(x, y, z) = ((x + y)z) \quad x = 1, y = 5, z = -3$


-3
$\frac{\partial p}{\partial x}\frac{\partial q}{\partial p}\frac{\partial f}{\partial q}$
1
x
p
-3
$\frac{\partial q}{\partial p}\frac{\partial f}{\partial q}$
+
q
5
y
X
f(x, y, z)
$\frac{\partial p}{\partial y}\frac{\partial q}{\partial p}\frac{\partial f}{\partial q}$
-3
$\frac{\partial f}{\partial q}$
1
-3
6
z
$\frac{\partial q}{\partial z}\frac{\partial f}{\partial q}$
6

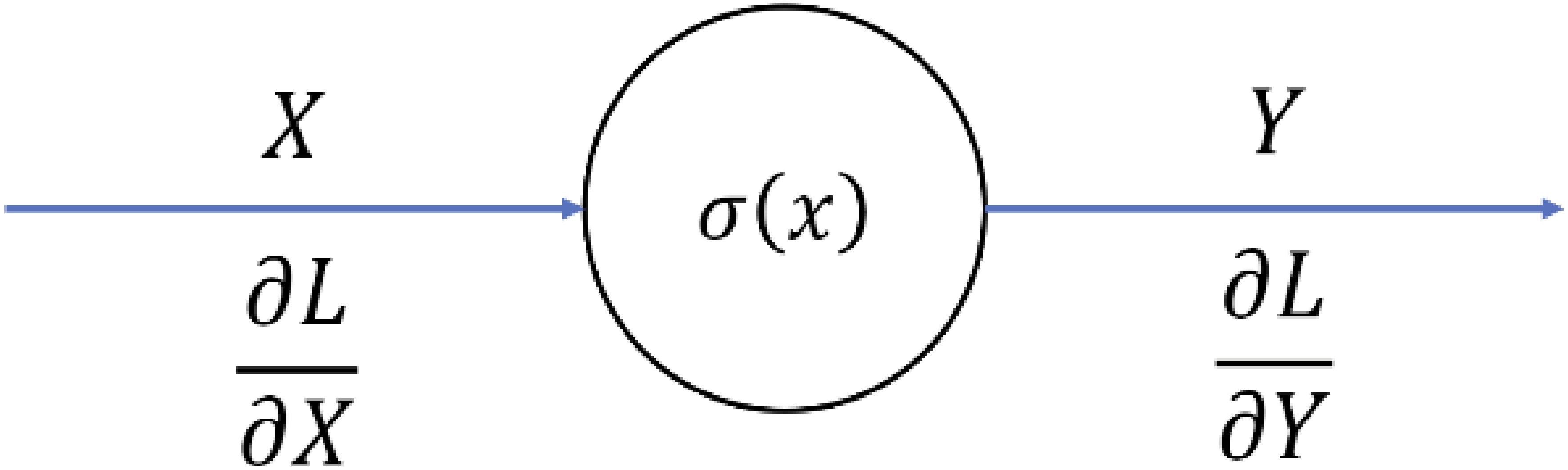
Backpropagation - Basic

$$\mathrm{f(x, y, z)} = ((x + y)z) \qquad x = 1, y = 5, z = -3$$



-3
$\frac{\partial p}{\partial x}\frac{\partial q}{\partial p}\frac{\partial f}{\partial q}$
p
-3
$\frac{\partial q}{\partial p}\frac{\partial f}{\partial q}$
q
1
-3
x
5
-3
y
+
X
$\mathrm{f(x, y, z)}$
$\frac{\partial p}{\partial y}\frac{\partial q}{\partial p}\frac{\partial f}{\partial q}$
-3
$\frac{\partial f}{\partial q}$
1
-3
6
z
$\frac{\partial q}{\partial z}\frac{\partial f}{\partial q}$
6

Making Sigmoid Layer



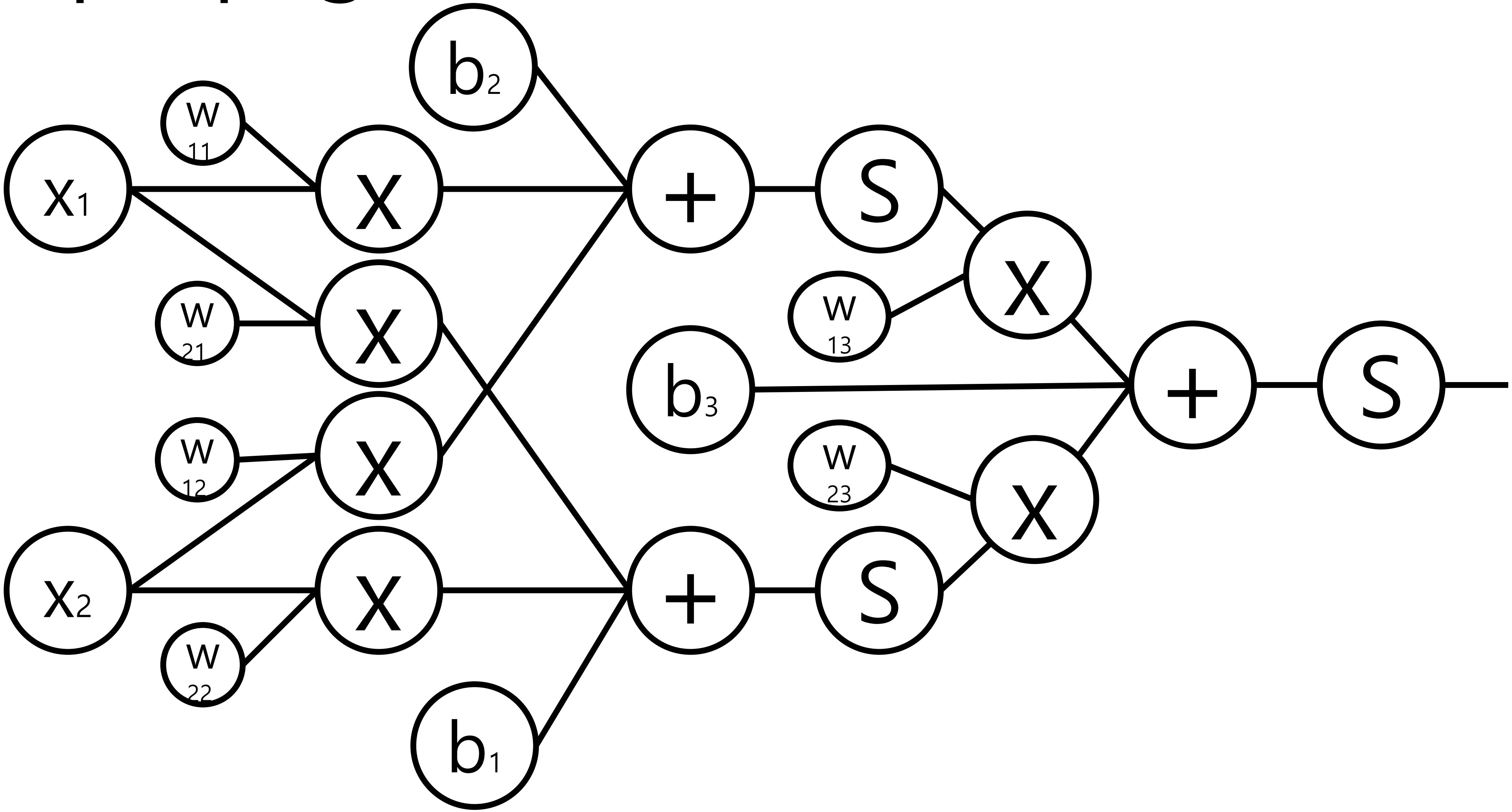
X $\sigma(x)$ Y

$\frac{\partial L}{\partial X}$ $\frac{\partial L}{\partial Y}$

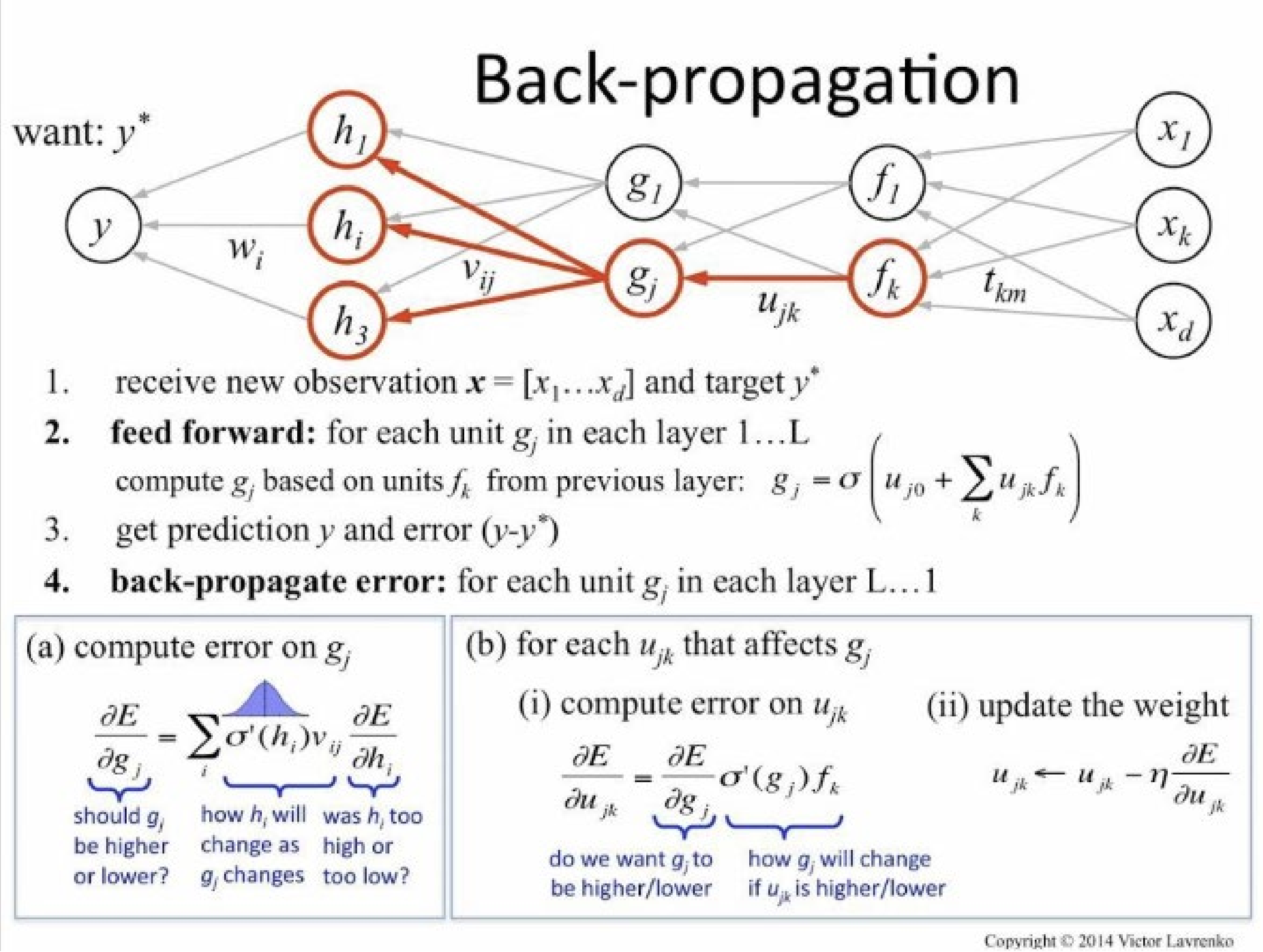
$$\sigma(x) = \frac{1}{1 + e^{-x}} \qquad \frac{\partial \sigma(x)}{\partial x} = (1 - \sigma(x))\sigma(x)$$

$$\frac{\partial L}{\partial X} = \frac{\partial L}{\partial Y} \cdot \frac{\partial Y}{\partial X} = \frac{\partial L}{\partial Y}(1 - Y)Y$$

Backpropagation - Real



Backpropagation - Real



Back-propagation

1. receive new observation $\mathbf{x} = [x_1 \ldots x_d]$ and target y^*
2. **feed forward:** for each unit g_j in each layer 1…L
   compute g_j based on units f_k from previous layer: $g_j = \sigma\left(u_{j0} + \sum_k u_{jk} f_k\right)$
3. get prediction y and error $(y\text{-}y^*)$
4. **back-propagate error:** for each unit g_j in each layer L…1

(a) compute error on g_j

$$\underbrace{\frac{\partial E}{\partial g_j}}_{\text{should } g_j \text{ be higher or lower?}} = \sum_i \underbrace{\sigma'(h_i) v_{ij}}_{\text{how } h_i \text{ will change as } g_j \text{ changes}} \underbrace{\frac{\partial E}{\partial h_i}}_{\text{was } h_i \text{ too high or too low?}}$$

(b) for each u_{jk} that affects g_j

(i) compute error on u_{jk}

$$\frac{\partial E}{\partial u_{jk}} = \underbrace{\frac{\partial E}{\partial g_j}}_{\text{do we want } g_j \text{ to be higher/lower}} \underbrace{\sigma'(g_j) f_k}_{\text{how } g_j \text{ will change if } u_{jk} \text{ is higher/lower}}$$

(ii) update the weight

$$u_{jk} \leftarrow u_{jk} - \eta \frac{\partial E}{\partial u_{jk}}$$

Sigmoid, bias, weight added

Lab Session

vlab-kaist/NN101_23S/lab/week5

$$f(x, y) = \sin(x)(x + y)(x^3 + 1)$$