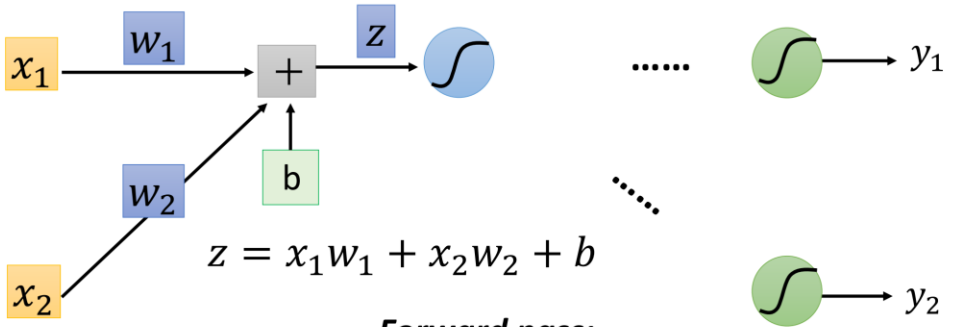
Backpropagation



Forward pass:

Compute $\partial z/\partial w$ for all parameters

$$\frac{\partial C}{\partial w} =? \quad \frac{\partial z}{\partial w}\frac{\partial C}{\partial z}$$

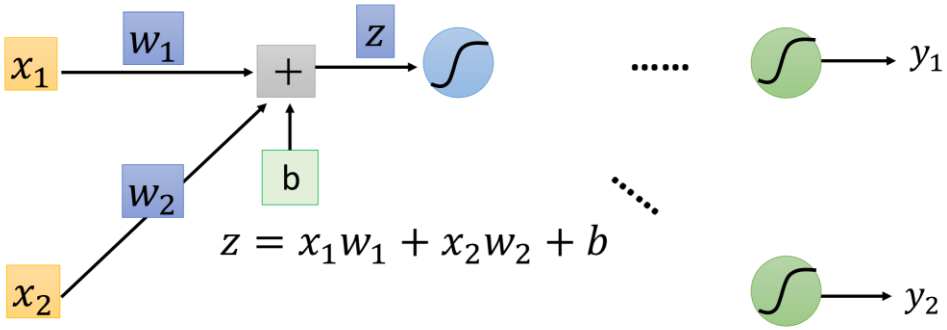
(Chain rule)

Backward pass:

Compute $\partial C/\partial z$ for all activation function inputs z

Backpropagation – Forward pass

Compute $\partial z/\partial w$ for all parameters



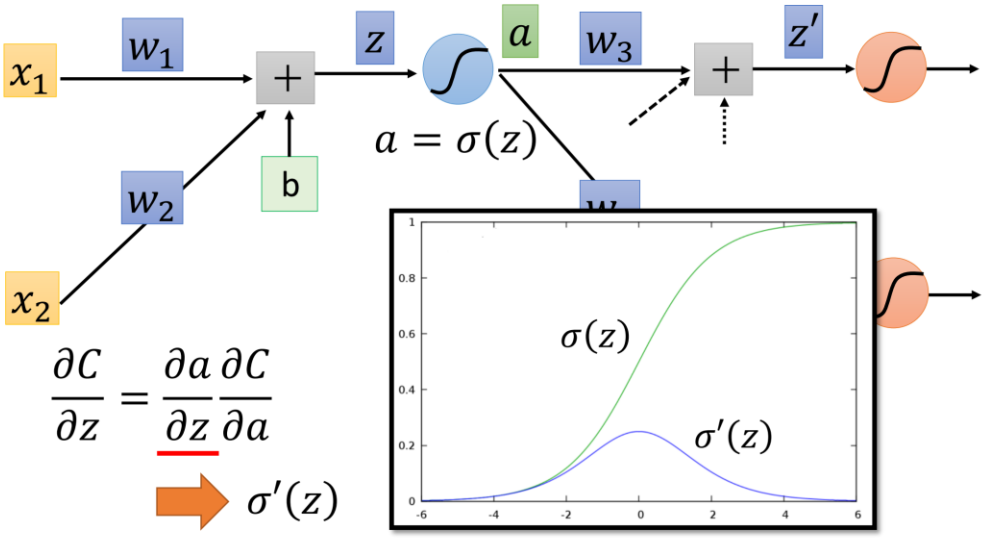
$\partial z/\partial w_1 =?\ x_1$

$\partial z/\partial w_2 =?\ x_2$

The value of the input connected by the weight

Backpropagation – Backward pass

Compute $\partial C/\partial z$ for all activation function inputs z

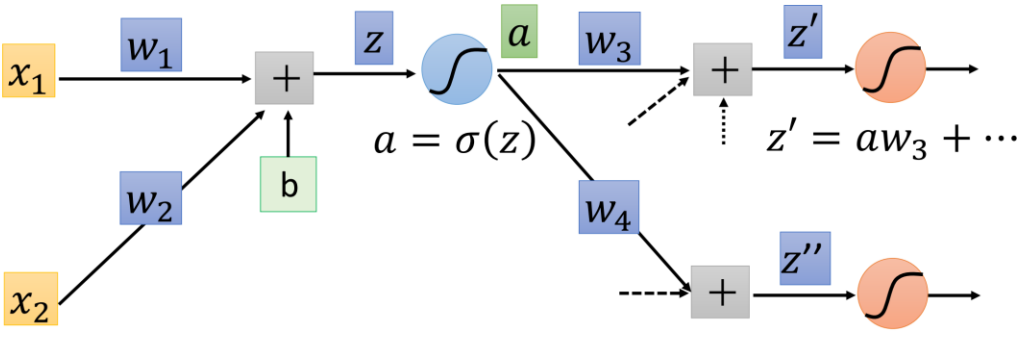


$$\frac{\partial C}{\partial z} = \frac{\partial a}{\partial z}\frac{\partial C}{\partial a}$$

$\sigma'(z)$

Backpropagation – Backward pass

Compute $\partial C/\partial z$ for all activation function inputs z



$$\frac{\partial C}{\partial z} = \frac{\partial a}{\partial z}\frac{\partial C}{\partial a}$$

$$\frac{\partial C}{\partial a} = \frac{\partial z'}{\partial a}\frac{\partial C}{\partial z'} + \frac{\partial z''}{\partial a}\frac{\partial C}{\partial z''}$$

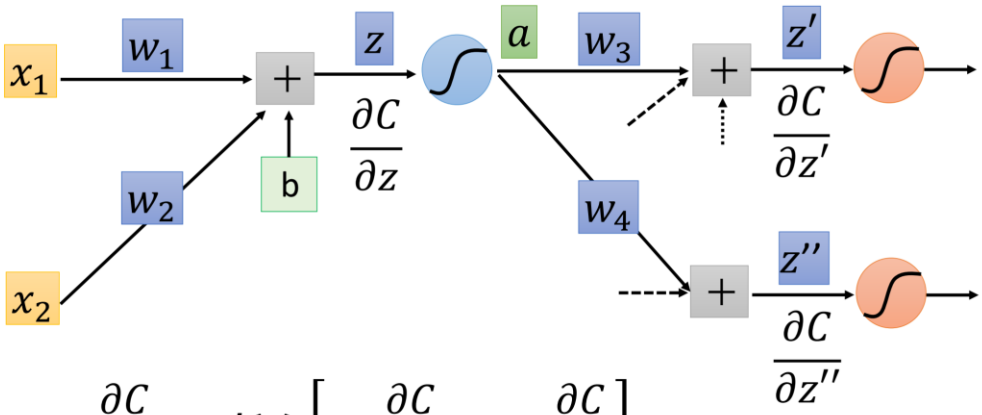
(Chain rule)

w_3 ? w_4 ?

Assumed it's known

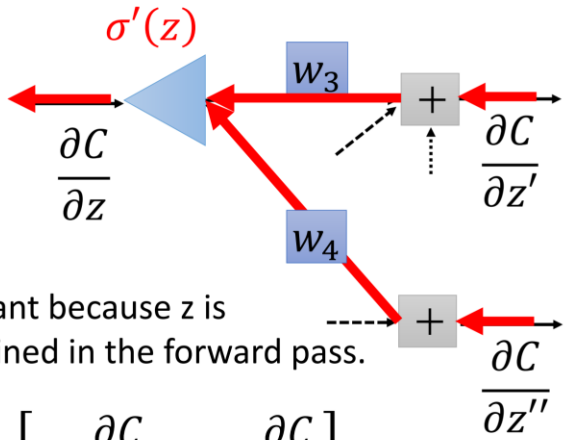
Backpropagation – Backward pass

Compute $\partial C/\partial z$ for all activation function inputs z



$$\frac{\partial C}{\partial z} = \sigma'(z)\left[w_3\frac{\partial C}{\partial z'} + w_4\frac{\partial C}{\partial z''}\right]$$

Backpropagation – Backward pass



$\sigma'(z)$ is a constant because z is already determined in the forward pass.

$$\frac{\partial C}{\partial z} = \sigma'(z)\left[w_3\frac{\partial C}{\partial z'} + w_4\frac{\partial C}{\partial z''}\right]$$