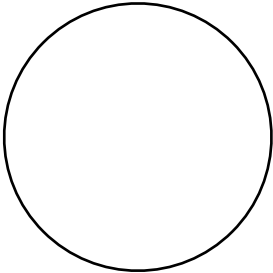
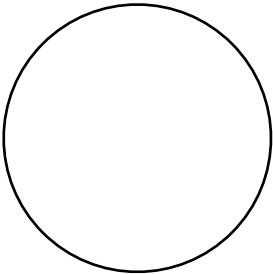


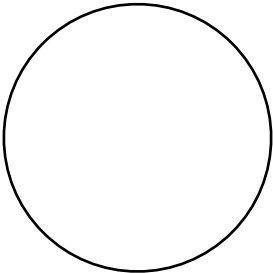
Gradient-based propagation



The most profitable application of the chain rule of all time!



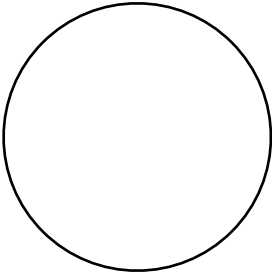






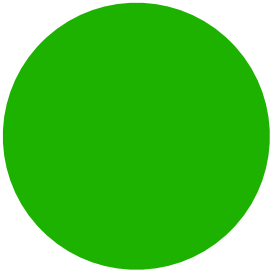


$$w_1(1)$$

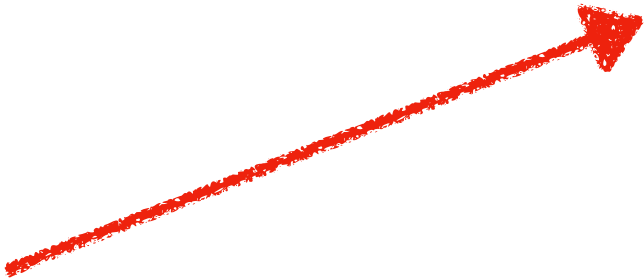


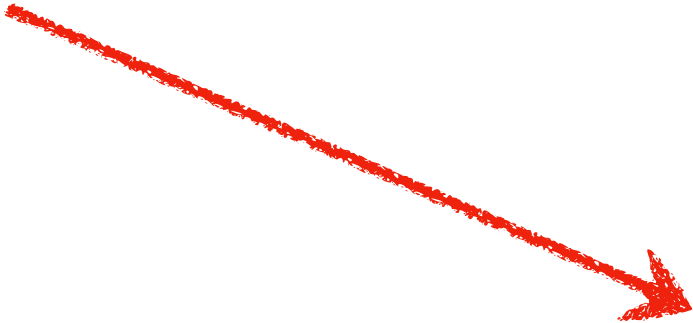
Q1)

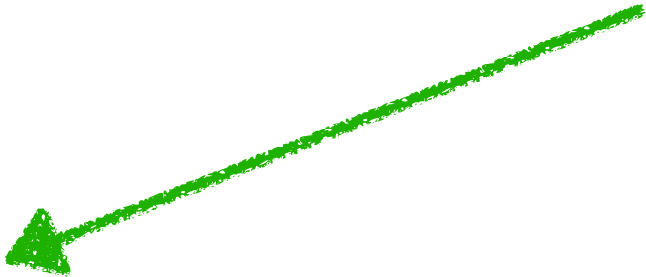
20















2(1)

$w_2(1)$









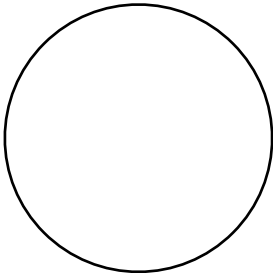
w(2)

P

v

|

x



accretion



ICE

$$\partial L_{CE}$$

$$\partial P_{y|x}$$

$$\partial P_{y|x}$$

$$\partial_Z(2)$$

$\partial z^{(2)}$



$\partial a^{(1)}$

softmax



$dz^{(1)}$

 dx_2

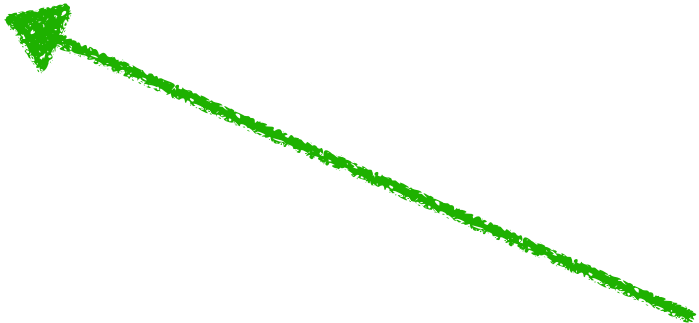
$\partial z^{(1)}$

∂x_1

$\partial a^{(1)}$

$\partial z^{(1)}$

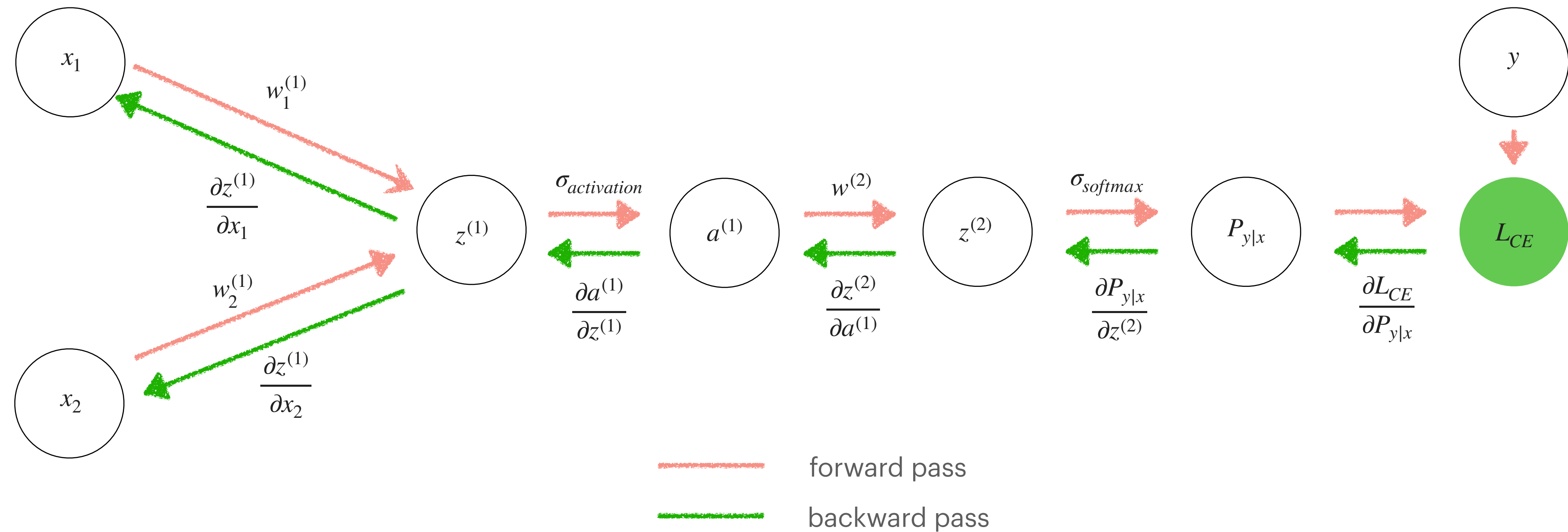




forward pass

backwards

Gradient backpropagation



The most profitable application of the chain rule of all time!

Neural network visualization