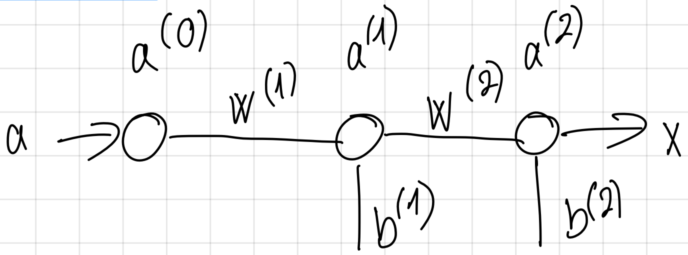
# Network 1




expected: $y$

$L = 2$

$L-1 = 1$

$L-2 = 0$

Training data: 1 example $\quad T^{(0)} = (x^{(0)}, y^{(0)})$

$$E^{(0)}(w^{(1)}, b^{(1)}, w^{(2)}, b^{(2)}) = (x^{(0)} - y^{(0)})^2$$

Want to calculate: $\nabla E^{(0)} \Rightarrow \nabla E^{(i)} \overset{\text{avg}}{\Rightarrow} \nabla E$

$$\nabla E^{(0)} = \begin{bmatrix} \partial_{w^{(1)}} E^{(0)} \\ \partial_{b^{(1)}} E^{(0)} \\ \partial_{w^{(2)}} E^{(0)} \\ \partial_{b^{(2)}} E^{(0)} \end{bmatrix}$$

backprop

## Last layer (Layer 2)

$$x^{(0)} = a^{(L)}$$

$$a^{(L)} = \sigma(z^{(L)})$$

$$z^{(L)} = w^{(L)} a^{(L-1)} + b^{(L)}$$

Reasoning:

How fast should $x^{(0)}$, therefore $a^{(L)}$ change?

$$\partial_{x^{(0)}} E^{(0)} = 2(x^{(0)} - y^{(0)})$$

let's call this $\Delta$

We know that $a^{(L)}$ should change with speed $\Delta$, so how fast should $z^{(L)}$ change? $\quad \Delta \cdot \partial_{z^{(L)}} a^{(L)} = \Delta \cdot \sigma'(z^{(L)})$

let's update $\Delta$:

$$\Delta := \Delta \cdot \sigma'(z^{(L)})$$

Now $z^{(L)}$ should change with speed $\Delta$, how fast should $w^{(L)}$ and $b^{(L)}$ change?

$$\partial_{w^{(L)}} E^{(0)} = \Delta \cdot \partial_{w^{(L)}} z^{(L)} = \underbrace{\Delta \cdot a^{(L-1)}}_{w^{(2)}} \qquad \partial_{b^{(L)}} E^{(0)} = \Delta \cdot \partial_{b^{(L)}} z^{(L)} = \Delta \cdot 1 = \underbrace{\Delta}_{b^{(2)}}$$

Let's call the current value of $\Delta$ $\Delta^{(L)}$ or $\Delta^{(2)}$.

Now: $$\nabla E^{(0)} = \begin{bmatrix} \partial_{w^{(1)}} E^{(0)} \\ \partial_{b^{(1)}} E^{(0)} \\ \Delta^{(2)} \cdot a^{(1)} \\ \Delta^{(2)} \end{bmatrix}$$

Layer 1

$$a^{(1)} = \sigma(z^{(1)})$$
$$z^{(1)} = w^{(1)} a^{(0)} + b^{(1)}$$

How fast should $a^{(1)}$ change? We can calculate using $\Delta^{(2)}$ and $z^{(2)}$:

$$\Delta^{(2)} \cdot \partial_{a^{(1)}} z^{(2)} = \Delta^{(2)} \cdot w^{(2)}$$ let's call this $\Delta^{(1)}$,

so $\Delta^{(1)} := \Delta^{(2)} \cdot w^{(2)}$

We know that $a^{(1)}$ should change with speed $\Delta^{(1)}$, so how fast should $z^{(1)}$ change? $\Delta^{(1)} \cdot \partial_{z^{(1)}} a^{(1)} = \Delta^{(1)} \cdot \sigma'(z^{(1)})$ let's update $\Delta^{(1)}$:

$\Delta^{(1)} := \Delta^{(1)} \cdot \sigma'(z^{(1)})$

Now $z^{(1)}$ should change with speed $\Delta^{(1)}$, how fast should $w^{(1)}$ and $b^{(1)}$ change?

$$\partial_{w^{(1)}} E^{(0)} = \Delta^{(1)} \cdot \partial_{w^{(1)}} z^{(1)} = \underbrace{\Delta^{(1)} \cdot a^{(0)}}_{w^{(1)}} \qquad \partial_{b^{(1)}} E^{(0)} = \Delta^{(1)} \cdot \partial_{b^{(1)}} z^{(1)} = \Delta^{(1)} \cdot 1 = \underbrace{\Delta^{(1)}}_{b^{(1)}}$$

Now: $$\nabla E^{(0)} = \begin{bmatrix} \Delta^{(1)} \cdot a^{(0)} \\ \Delta^{(1)} \\ \Delta^{(2)} \cdot a^{(1)} \\ \Delta^{(2)} \end{bmatrix}$$