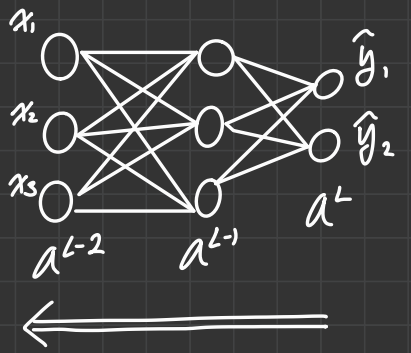
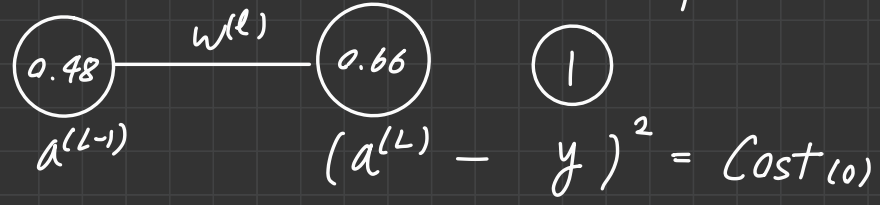
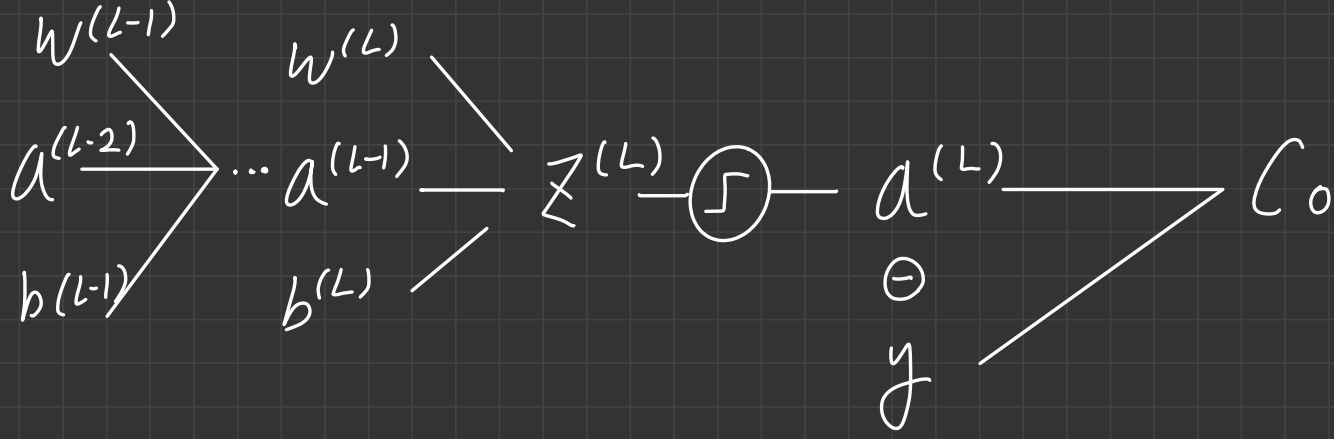
* Back propagation

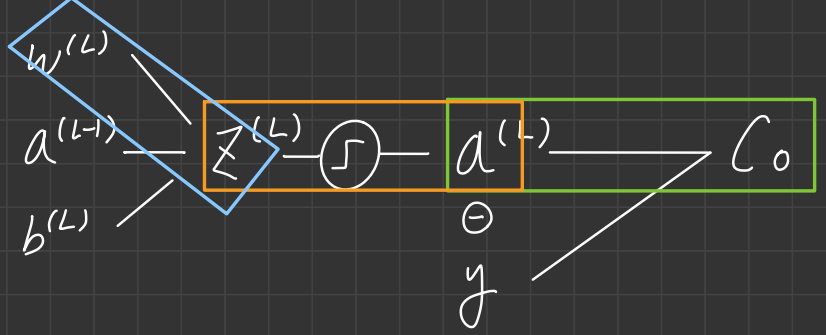




$W(l) \cdot a^{(L-1)} + bias = z^{(L)}$

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





$$(a^{L}-y)^2$$

$$\frac{\partial C_0}{\partial a^{(L)}} = 2(a^{(L)}-y)$$

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

$$\frac{\partial z^{(L)}}{\partial w^{(L)}} = a^{(L-1)}$$

Suppose $w^{(L)}$ $\partial w^{(L)}$

$z^{(L)}$ $\partial z^{(L)}$

the chain rule : small change of w → to Cost.

Cost의 변화 ↑ w의 변화량

$$\frac{\partial C_0}{\partial w^{(L)}} = \frac{\partial z^{(L)}}{\partial w^{(L)}} \times \frac{\partial a^{(L)}}{\partial z^{(L)}} \times \frac{\partial C_0}{\partial a^{(L)}}$$

what we want

$\frac{\partial z^{(L)}}{\partial w^{(L)}}$: $w \to z$

$\frac{\partial a^{(L)}}{\partial z^{(L)}}$: $z \to a$

$\frac{\partial C_0}{\partial a^{(L)}}$: $a \to C_0$