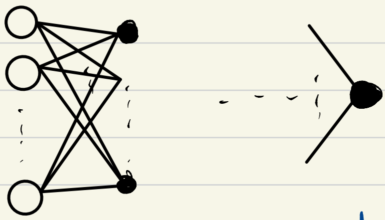
1.



Consider n_l : is the number of layer l for $l = 1, \dots L-1$, $n_L = 1$

Assume $z^{(l)} = W^{(l)} a^{(l-1)} + b^{(l)}$ $\forall\ l = 1, \dots, L$
where $z^{(l)} \in \mathbb{R}^{n_l}$ is the output of layer $l-1$,
$W^{(l)} \in \mathbb{R}^{n_l \times n_{l-1}}$: weight, $b^{(l)} \in \mathbb{R}^{n_l}$: bias

$$\Rightarrow \begin{cases} a^{(1)} = x \in \mathbb{R}^{n_1} \\ a^{(l)} = \sigma(z^{l}) = \sigma\left(W^{(l)} a^{(l-1)} + b^{(l)}\right) \end{cases}$$

Define $\delta^{(l)} := \frac{\partial a^{(L)}}{\partial z^{(l)}} = \frac{\partial a^{(L)}}{\partial a^{(l)}} \cdot \frac{\partial a^{(l)}}{\partial z^{(l)}}$

$\Rightarrow \delta^{(L)} = \frac{\partial a^{(L)}}{\partial z^{(L)}} = \sigma'(z^{(L)})$ $(\because\ a^{(L)} \in \mathbb{R}^1)$

2.

Nope!