

$$a^{[n]} = \sigma(w^{[n]} a^{[n-1]} + b^{[n]}), \quad n=2, \dots, L$$

$$\Rightarrow \text{let } z^{[n]} = w^{[n]} a^{[n-1]} + b^{[n]}, \quad n=2, \dots, L$$

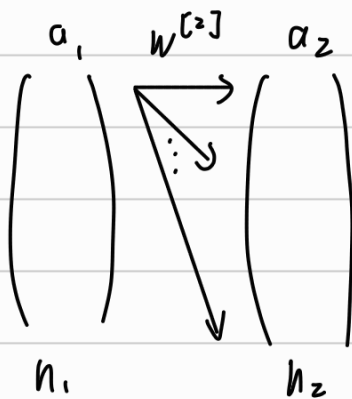
$$\delta^{[L]} = \frac{\partial a^{[L]}}{\partial z^{[L]}} = \sigma'(z^{[L]})$$

$$\delta^{[L-1]} = \frac{\partial a^{[L]}}{\partial z^{[L-1]}} = \frac{\partial a^{[L]}}{\partial z^{[L]}} \cdot \frac{\partial z^{[L]}}{\partial z^{[L-1]}} = \sigma'(z^{[L-1]}) \circ \left[(w^{[L]})^T \delta^{[L]} \right]$$

\vdots

$$\delta^{[n]} = \sigma'(z^{[n]}) \circ \left[(w^{[n+1]})^T \delta^{[n+1]} \right], \quad n=2, \dots, (L-2), \quad \text{where } \delta^{[L]} = \frac{\partial a^{[L]}}{\partial z^{[L]}}$$

$$\Rightarrow \nabla a^{[L]}(x) = \frac{\partial a^{[L]}}{\partial a^{[1]}} = \frac{\partial a^{[L]}}{\partial z^{[2]}} \cdot \frac{\partial z^{[2]}}{\partial a^{[1]}} = (w^{[2]})^T \delta^{[2]}$$



\dots

