$$a^{[1]} = x \in \mathbb{R}^{n_1} \quad \text{—— (3.1)}$$
$$a^{[l]} = \sigma(W^{[l]}a^{[l-1]} + b^{[l]}) \in \mathbb{R}^{n_l}, \quad \text{for } l = 2 \sim L. \quad \text{—— (3.3)}$$

Define $z^{[l]} = W^{[l]}a^{[l-1]} + b^{[l]}$. Then we have $a^{[L]} = \sigma(z^{[L]})$.
Define $\delta^{[l]} = \dfrac{\partial a^{[L]}}{\partial z^{[l]}}$. Then $\delta^{[L]} = \sigma'(z^{[L]})$.

$$\Rightarrow \delta^{[l]} = \frac{\partial a^{[L]}}{\partial z^{[l+1]}} \cdot \frac{\partial z^{[l+1]}}{\partial z^{[l]}}, \quad \text{for } l = L-1$$
$$= \delta^{[L]} \cdot W^{[L]} \operatorname{diag}(\sigma'(z^{[L-1]}))$$

For $l = L-2, \dots, 2$

$$\Rightarrow \delta^{[l]} = \frac{\partial a^{[L]}}{\partial z^{[l+1]}} \cdot \frac{\partial z^{[l+1]}}{\partial z^{[l]}} = \sigma'(z^{[l]}) \cdot (W^{[l+1]})^T \delta^{[l+1]}$$

For $n_1 = 1$, $z^{[2]} = W^{[2]}a^{[1]} + b^{[2]}$
Then $\nabla_a a^{[L]}(x) = \dfrac{\partial a^{[L]}}{\partial x} = \dfrac{\partial a^{[L]}}{\partial z^{[2]}} \cdot \dfrac{\partial z^{[2]}}{\partial x}$

$$= \delta^{[2]} (W^{[2]})^T$$