1、

$$a^{[L]} = \sigma\left(w^{[L]} a^{[L-1]} + b^{[L]}\right)$$

$$\because \frac{\partial a^{[L]}}{\partial x} = \frac{\partial a^{[L]}}{\partial a^{[L-1]}} \cdot \frac{\partial a^{[L-1]}}{\partial a^{[L-2]}} \cdots \frac{\partial a^{[2]}}{\partial a^{[1]}}, \quad \frac{\partial a^{[L]}}{\partial a^{[L-1]}} = D^{[L]} W^{[L]}$$

$$\Rightarrow \nabla a^{[L]}(x)^T = \left(D^{[L]} W^{[L]}\right)\left(D^{[L-1]} W^{[L-1]}\right) \cdots \left(D^{[2]} W^{[2]}\right)$$

$$\therefore \nabla a^{[L]}(x) = \left[\left(D^{[L]} W^{[L]}\right) \cdots \left(D^{[2]} W^{[2]}\right)\right]^T, \text{ when } n_L = 1$$