1.

By Deep Learning: An Introduction for Applied Mathematicians, we have

(3.1) $\quad a^{[1]} = x \in \mathbb{R}^{[n_1]}$,

(3.2) $\quad a^{[l]} = \sigma\left(W^{[l]}a^{[l-1]} + b^{[l]}\right) \in \mathbb{R}^{[n_l]} \qquad$ for $l = 2, 3, \dots, L$

And we have $n_L = 1$.

For $l = 2$,

$$a^{[2]} = \sigma\left(W^{[2]}a^{[1]} + b^{[2]}\right) = \sigma\left(W^{[2]}x + b^{[2]}\right)$$

Let $Z^{[k]} = W^{[k]}a^{[k-1]} + b^{[k]}$. We have

$$a^{[k]} = \sigma\left(Z^{[k]}\right)$$

Then

$$\frac{\partial a^{[k]}}{\partial a^{[k-1]}} = \frac{\partial a^{[k]}}{\partial Z^{[k]}}\frac{\partial Z^{[k]}}{\partial a^{[k-1]}} = \sigma'(Z^{[k]})W^{[k]}$$

$$= \begin{bmatrix} \sigma'\left(z_1^{[k]}\right) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma'\left(z_{n_k}^{[k]}\right) \end{bmatrix} W^{[k]}$$

Let $M^{[k]} \coloneqq diag(\sigma'\left(z_1^{[k]}\right), \sigma'\left(z_2^{[k]}\right), \dots, \sigma'\left(z_{n_k}^{[k]}\right))W^{[k]} \quad$ for $k = 2, \dots, L$.

Then

$$\nabla a^{[L]}(x) = \frac{\partial a^{[L]}}{\partial x} = \frac{\partial a^{[L]}}{\partial a^{[L-1]}}\frac{\partial a^{[L-1]}}{\partial a^{[L-2]}} \cdots \frac{\partial a^{[2]}}{\partial a^{[1]}}\frac{\partial a^{[1]}}{\partial x}$$
$$= M^{[k]}M^{[k-1]} \dots M^{[2]} \cdot 1$$