Assignment 2

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

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

$n_L = 1$

Solve $\nabla a^{[L]}(x)$

<Sol>

Let $z^{[l]} = W^{[l]} a^{[l-1]} + b^{[l]}$

$$a^{[L]} = \sigma(W^{[L]} a^{[L-1]} + b^{[L]}) = \sigma(z^{[L]})$$

$$\nabla_x a^{[L]} = \frac{\partial a^{[L]}}{\partial a^{[L-1]}} \cdot \frac{\partial a^{[L-1]}}{\partial x}$$

$$\frac{\partial a^{[l]}}{\partial a^{[l-1]}} = \frac{\partial a^{[l]}}{\partial z^{[l]}} \cdot \frac{\partial z^{[l]}}{\partial a^{[l-1]}}$$

$$\because 1、 \frac{\partial z^{[l+1]}}{\partial a^{[l]}} = \frac{\partial (W^{[l+1]} a^{[l]} + b^{[l+1]})}{\partial a^{[l]}} = W^{[l+1]}$$

$$2. \frac{\partial a^{[l]}}{\partial z^{[l]}} = \mathrm{diag}(\sigma'(z^{[l]})) \Rightarrow \begin{bmatrix} \sigma'(z_1^{[l]}) & 0 & \cdots & 0 \\ 0 & \sigma'(z_2^{[l]}) & & \vdots \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & \sigma'(z_{n_l}^{[l]}) \end{bmatrix}$$

$$\therefore \frac{\partial a^{[l]}}{\partial a^{[l-1]}} = \mathrm{diag}(\sigma'(z^{[l]})) \cdot W^{[l]}$$

$$\nabla a^{[L]}(x) = \frac{\partial a^{[L]}}{\partial x} = \frac{\partial a^{[L]}}{\partial a^{[L-1]}} \cdot \frac{\partial a^{[L-1]}}{\partial a^{[L-2]}} \cdot \frac{\partial a^{[L-2]}}{\partial a^{[L-3]}} \cdots \frac{\partial a^{[2]}}{\partial a^{[1]}} \cdot \underbrace{\frac{\partial a^{[1]}}{\partial x}}_{\Downarrow \; 1}$$

$$= \prod_{l=2}^{L} \left(\mathrm{diag}(\sigma'(z^{[l]})) \, W^{[l]} \right) \cdot 1 , \quad z^{[l]} = W^{[l]} a^{[l-1]} + b^{[l]}$$

※