1、 Let $a_i^{[l]} = \sigma(z_i^{[l]})$ for $i = 1 \cdots n_l$

since $\frac{\partial a_i^{[l]}}{\partial z_j^{[l]}} = \frac{\partial \sigma(z_i^{[l]})}{\partial z_j^{[l]}} = \begin{cases} \sigma'(z_j^{[l]}), & \text{if } i=j \\ 0, & \text{if } i \neq j \end{cases}$

$\Rightarrow \frac{\partial a^{[l]}}{\partial z^{[l]}} = \mathrm{diag}(\sigma'(z^{[l]}))$

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

$$\nabla_a a^{[L]}(x) = \frac{\partial a^{[L]}}{\partial a^{[L-1]}} \cdot \frac{\partial a^{[L-1]}}{\partial a^{[L-2]}} \cdots \frac{\partial a^{[2]}}{\partial a^{[1]}} = \left(\mathrm{diag}(\sigma'(z^{[L]}))\, W^{[L]}\right) \cdot \left(\mathrm{diag}(\sigma'(z^{[L-1]}))\, W^{[L-1]}\right) \cdots \cdot \left(\mathrm{diag}(\sigma'(z^{[2]}))\, W^{[2]}\right)$$

2、 Is there any specific example about approximation theory ?