1. Let $f_\ell(z) = \sigma(W^{[\ell]}z + b^{[\ell]})$ and $W^{[\ell]}a^{[\ell-1]} + b^{[\ell]} = z^{[\ell]}$ for $\ell = 1, 2, \dots L$, then

$$a^{[L]\prime}(x) = f_L'\left(a^{[L-1]}(x)\right) \cdot a^{[L-1]\prime}(x)$$

$$= \sigma'\left(z^{[L]}\right) \cdot W^{[L]} \cdot a^{[L-1]\prime}(x)$$

note that $\sigma'(z^{[L]}) = \left(\sigma'(z_1^{[L]}), \sigma'(z_2^{[L]}) \dots\right)$

Hence $a^{[L]\prime}(x) = \sigma'(z^{[L]}) \cdot W^{[L]} \cdot \sigma'(z^{[L-1]}) W^{[L-1]}$

$\dots\dots \sigma'(z^{[2]}) W^{[2]}$

$$\nabla a^{[L]}(x) = \left(a^{[L]\prime}(x)\right)^T$$

($a^{[L]}(x)$ maps $\mathbb{R}^n$ to $\mathbb{R}$, so $a^{[L]\prime}(x)$ is a row vector $\left(\frac{\partial a^{[L]}}{\partial x_1} \dots \frac{\partial a^{[L]}}{\partial x_n}\right)$)