

$$\frac{\partial J(\theta)}{\partial \theta_{ij}^{(l)}} = \frac{\partial J(\theta)}{\partial a_i^{(l+1)}} \times \frac{\partial a_i^{(l+1)}}{\partial z_i^{(l+1)}} \times \frac{\partial z_i^{(l+1)}}{\partial \theta_{ij}^{(l)}}$$

$$\frac{\partial a_i^{(l+1)}}{\partial z_i^{(l+1)}} = g'(z_i^{(l+1)})$$

$$\frac{\partial z_i^{(l+1)}}{\partial \theta_{ij}^{(l)}} = \frac{\partial ((\theta^{(l)} \times a^{(l)})_i)}{\partial \theta_{ij}^{(l)}} = a_j^{(l)}$$

$$\text{Definition: } \delta_i^{(l)} = \frac{\partial J(\theta)}{\partial z_i^{(l)}} \quad \delta^{(l)} = \frac{\partial J(\theta)}{\partial z^{(l)}}$$

$$\begin{aligned} \frac{\partial J(\theta)}{\partial a_i^{(l+1)}} &= \sum_{k=1}^{s_{l+2}} \left(\frac{\partial J(\theta)}{\partial z_k^{(l+2)}} \times \frac{\partial z_k^{(l+2)}}{\partial a_i^{(l+1)}} \right) = \sum_{k=1}^{s_{l+2}} \left(\frac{\partial J(\theta)}{\partial z_k^{(l+2)}} \times \theta_{\textcolor{red}{k}i}^{(l+1)} \right) \\ &= ((\theta^{(l+1)})^T \times \frac{\partial J(\theta)}{\partial z^{(l+2)}})_i = ((\theta^{(l+1)})^T \times \delta^{(l+2)})_i \end{aligned}$$

$$\text{Hence, } \frac{\partial J(\theta)}{\partial \theta_{ij}^{(l)}} = a_j^{(l)} \times \delta_i^{(l+1)} \quad \text{and} \quad \delta_i^{(l)} = \left((\theta^{(l)})^T \times \delta^{(l+1)} \right) \times g'(z^{(l)})$$