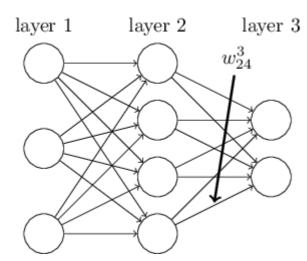
## **Derive backpropagation**

## **1 Forward Propagation**



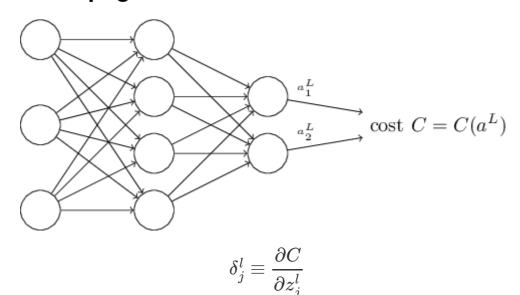
 $w_{jk}^l$  is the weight from the  $k^{\rm th}$  neuron in the  $(l-1)^{\rm th}$  layer to the  $j^{\rm th}$  neuron in the  $l^{\rm th}$  layer

$$oldsymbol{a}^l = \sigma(oldsymbol{z}^l) \ oldsymbol{z}^l = oldsymbol{w}^l oldsymbol{a}^{l-1} + oldsymbol{b}^l$$

non-vectored form

$$a_j^l = \sigma(z_j^l) \ z_j^l = \sum_k w_{jk}^l a_k^{l-1} + b_j^l$$

## 2 Backward Propagation



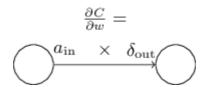
$$\boldsymbol{\delta}^{L} = \frac{\partial C}{\partial \boldsymbol{a}^{L}} \odot \sigma'(\boldsymbol{z}^{L}) = \nabla_{a} C \odot \sigma'(\boldsymbol{z}^{L}) \tag{1}$$

$$oldsymbol{\delta}^l = ((oldsymbol{w}^{l+1})^T oldsymbol{\delta}^{l+1}) \odot \sigma'(oldsymbol{z}^l)$$
 (2)

$$\frac{\partial C}{\partial b^l} = \delta^l \tag{3}$$

$$\frac{\partial C}{\partial w_{jk}^l} = \delta_j^l a_k^{l-1} \tag{4}$$

a(in)是输入给权重w的神经元的激活值, $\delta(out)$ 是输出自权重w的神经元的误差,那么代价函数对参数w的偏导如下:



## 3 Proof and Derivation

• 3.1 The first equation

$$\begin{split} \delta_j^L &= \sum_k \frac{\partial C}{\partial a_k^L} \cdot \frac{\partial a_k^L}{\partial z_j^L} \\ &= \frac{\partial C}{\partial a_j^L} \cdot \frac{\partial a_j^L}{\partial z_j^L} \\ &= \frac{\partial C}{\partial a_j^L} \sigma'(z_j^L) \end{split}$$

当j不等于k时, a(k)对z(j)偏导为0

• 3.2 The second equation

$$egin{aligned} \delta_j^l &= rac{\partial C}{\partial z_j^l} \ &= rac{\partial C}{\partial a_j^l} \cdot rac{\partial a_j^l}{\partial z_j^l} \ &= rac{\partial C}{\partial a_j^l} \sigma'(z_j^l) \ &= \sum_k rac{\partial C}{\partial z_k^{l+1}} \cdot rac{\partial z_k^{l+1}}{\partial a_j^l} \cdot \sigma'(z_j^l) \ &= \sum_k \delta_k^{l+1} w_{kj} \sigma'(z_j^l) \end{aligned}$$

• 3.3 The third equation

$$rac{\partial C}{\partial b_j^l} = rac{\partial C}{\partial z_j^l} \cdot rac{\partial z_j^l}{\partial b_j^l} = rac{\partial C}{\partial z_j^l} \cdot 1 = \delta_j^l$$

• 3.4 The fourth equation

$$\begin{split} \frac{\partial C}{\partial w_{jk}^l} &= \frac{\partial C}{\partial z_j^l} \cdot \frac{\partial z_j^l}{\partial w_{jk}^l} \\ &= \delta_j^l \cdot \frac{\partial z_j^l}{\partial w_{jk}^l} \\ &= \delta_j^l \cdot \frac{\partial (\sum_k a_k^{l-1} w_{jk}^l + b_j^l)}{\partial w_{jk}^l} \\ &= \delta_j^l a_k^{l-1} \end{split}$$