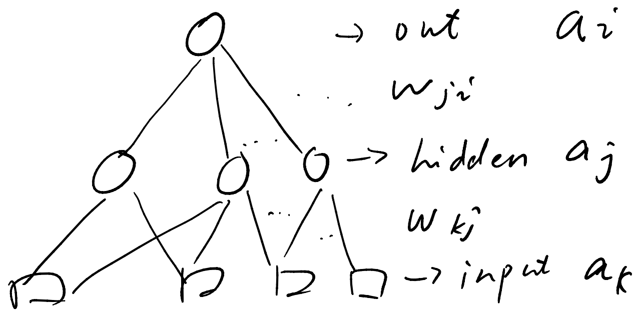
Back-propagation deriviation

Thursday, 9 April 2020 7:03 AM



$$a_i = g(x_i)$$
$$x_i = \sum_j a_j w_{ji}$$
$$a_j = g(x_j)$$
$$x_j = \sum_k a_k w_{kj}$$

forward

g: activation function

$$E = \frac{1}{2}\sum_i (y_i - a_i)^2 = \frac{1}{2}\left[(y_0-a_0)^2 + \cdots (y_i - a_i)^2 + \cdots + (y_n - a_n)^2\right]$$

1. Backward partial derivitives

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial Err_i} \cdot \frac{\partial Err_i}{\partial a_i} \cdot \frac{\partial a_i}{\partial g} \cdot \frac{\partial g}{\partial x_i} \cdot \frac{\partial x_i}{\partial w_{ji}}$$

$$= (y_i - a_i) \cdot (-1) \cdot 1 \cdot g'(x_i) \cdot a_j$$

use the derivatives of the activation function

$$= \underbrace{(a_i - y_i) \cdot g'(x_i)}_{\delta_i} \cdot a_j$$

$$= a_j \cdot \delta_i$$

2. hidden → input

$$\frac{\partial E}{\partial w_{kj}} = \frac{\partial E}{\partial Err} \cdot \frac{\partial Err}{\partial w_{kj}} = \sum_i (a_i - y_i) \frac{\partial Err}{\partial w_{kj}}$$

$$= \sum_i (a_i - y_i) \cdot g'(x_i) \cdot \frac{\partial x_i}{\partial w_{kj}} = \sum_i (a_i - y_i) \cdot g'(x_i) \cdot \frac{\partial x_i}{\partial a_j} \cdot \frac{\partial a_j}{\partial w_{kj}}$$

$$= \sum_i \underbrace{(a_i - y_i) \cdot g'(x_i)}_{\delta_i} \cdot w_{ji} \cdot \frac{\partial a_j}{\partial w_{kj}}$$

$$= \left(\sum_i \delta_i w_{ji}\right) \cdot g'(x_j) \cdot \frac{\partial x_j}{\partial w_{kj}}$$

$$= \underbrace{\left(\sum_i \delta_i w_{ji}\right) \cdot g'(x_j)}_{\delta_j} \cdot a_k$$