



训练时  
向前传 forward  
向后传 backward

$$\text{Input}^{(1)} = \underline{1} = x = \underline{0}^{(1)}$$

$$\underline{I}^{(1)} = \underline{w_{1j}} \cdot \underline{x} + \underline{c_j}$$

$$\underline{O}^{(1)} = \text{sigmoid}(w_{1j} \cdot x + c_j = \text{sigmoid}(I^{(1)}))$$

$$\underline{I}^{(2)} = \underline{v_{jk}} \cdot \underline{O}^{(1)} + \underline{b_k^{(2)}}$$

$$\underline{O}^{(2)} = \underline{1}^{(2)}$$

$$E = \frac{1}{2} (O^{(2)} - y)^2 \text{ 损失函数}$$

向后传

$$E = f(w_{ij} v_{jk} c_j b_k)$$

$$\frac{\partial E}{\partial w_{ij}} \quad \frac{\partial E}{\partial v_{jk}} \quad \frac{\partial E}{\partial c_j} \quad \frac{\partial E}{\partial b_k}$$

$$\frac{\partial A x}{\partial x} = A^T$$

$$\frac{\partial E}{\partial v_{jk}} = \frac{\partial E}{\partial o^{(k)}} \cdot \frac{\partial o^{(k)}}{\partial v_{jk}} = \frac{\partial \frac{1}{2}(o^{(k)} - y)^2}{\partial o^{(k)}} \cdot \frac{\partial (v_{jk} \cdot o^{(j)} + b^{(k)})}{\partial v_{jk}} = (o^{(k)} - y) \cdot o^{(j)}$$

$$\frac{\partial E}{\partial b_k} = \frac{\partial E}{\partial o^{(k)}} \cdot \frac{\partial o^{(k)}}{\partial b_k} = \sim \cdot \frac{\partial (v_{jk} \cdot o^{(j)} + b^{(k)})}{\partial b^{(k)}} = (o^{(k)} - y)$$

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial o^{(k)}} \cdot \frac{\partial o^{(k)}}{\partial o^{(j)}} \cdot \frac{\partial o^{(j)}}{\partial w_{ij}} = (o^{(k)} - y) \cdot v_{jk} \cdot \frac{\partial o^{(j)}}{\partial w_{ij}} = (o^{(k)} - y) \cdot v_{jk} \cdot \text{sigmoid}(w_{ij} \cdot x + c_j) \cdot (1 - \text{sigmoid}(w_{ij} \cdot x + c_j)) \cdot x$$

$$\frac{\partial o^{(k)}}{\partial o^{(j)}} = \frac{\partial (v_{jk} \cdot o^{(j)} + b^{(k)})}{\partial o^{(j)}} = v_{jk}$$

$$\frac{\partial o^{(j)}}{\partial w_{ij}} = \frac{\partial \text{sigmoid}(w_{ij} \cdot x + b^j)}{\partial w_{ij}} = \text{sigmoid}(w_{ij} \cdot x + b^j) \cdot (1 - \text{sigmoid}(w_{ij} \cdot x + b^j)) \cdot x$$

$$1) \text{ } f(g(x)) = f(u) \cdot g'(x) \quad (u = g(x))$$

$$2) \text{ } \text{sigmoid}'(x) = \text{sigmoid}(x) \cdot (1 - \text{sigmoid}(x))$$

$$\frac{\partial E}{\partial c_j} = \frac{\partial E}{\partial o^{(k)}} \cdot \frac{\partial o^{(k)}}{\partial o^{(j)}} \cdot \frac{\partial o^{(j)}}{\partial c_j} =$$

$$\frac{\partial o^{(j)}}{\partial c_j} = \frac{\partial \text{sigmoid}(w_{ij} \cdot x + c_j)}{\partial c_j} = \text{sigmoid}(w_{ij} \cdot x + c_j) \cdot (1 - \text{sigmoid}(w_{ij} \cdot x + c_j)) \cdot 1$$

$$w'_{ij} = w_{ij} - \eta \frac{\partial E}{\partial w_{ij}}$$

$$v'_{jk} = v_{jk} - \eta \frac{\partial E}{\partial v_{jk}}$$

$$b'_k = b_k - \eta \frac{\partial E}{\partial b_k}$$

$$c'_j = c_j - \eta \frac{\partial E}{\partial c_j}$$