



$$z_j = \sum_i x_i \cdot w_{ij} + b_j$$

$$y_j = f(z_j)$$

$$e = \frac{1}{2} \sum_j (y_j - y_j^*)^2$$

$$\frac{\partial e}{\partial y_j} = y_j - y_j^*$$

$$\delta_j \equiv \frac{\partial e}{\partial z_j} = \frac{\partial e}{\partial y_j} \cdot f'(z_j)$$

$$\frac{\partial e}{\partial w_{ij}} = x_i \delta_j$$

$$\frac{\partial e}{\partial b_j} = \delta_j$$

$$\frac{\partial e}{\partial x_i} = \sum_j \delta_j w_{ji}^T$$

$$\frac{\partial e}{\partial y_j} \leftarrow \frac{\partial e}{\partial x_i}$$

- 使用 $\nabla_{\vec{y}}(e)$

$$z_j = \sum_i x_i \cdot w_{ij} + b_j$$

$$y_j = f(z_j)$$

$$e = \frac{1}{2} \sum_j (y_j - y_j^*)^2$$

$$\nabla_{y_j} \equiv \frac{\partial e}{\partial y_j} = y_j - y_j^*$$

$$\nabla_{z_j} \equiv \frac{\partial e}{\partial z_j} = \nabla_{y_j} \cdot f'(z_j)$$

$$\nabla_{w_{ij}} = x_i \nabla_{z_j}$$

$$\nabla_{b_j} = \nabla_{z_j}$$

$$\nabla_{x_i} = \sum_j \nabla_{z_j} w_{ji}^T$$

$$\nabla_{y_j} \leftarrow \nabla_{x_i}$$

- 使用 vector notation, \vec{x} , \vec{y} , \vec{z} , all are the row vectors

$$\vec{z} = \vec{x} @ W + \vec{b}$$

$$\vec{y} = f(\vec{z})$$

$$e = \frac{1}{2} \|\vec{y} - \vec{y}^*\|^2$$

$$\nabla_{\vec{y}} \equiv \frac{\partial e}{\partial \vec{y}} = \vec{y} - \vec{y}^*$$

$$\nabla_{\vec{z}} = \nabla_{\vec{y}} \odot f'(\vec{z})$$

$$\nabla_W = \vec{x}^T \nabla_{\vec{z}}$$

$$\nabla_{\vec{b}} = \nabla_{\vec{z}}$$

$$\nabla_{\vec{x}} = \nabla_{\vec{z}} @ W^T$$

$$\nabla_{\vec{y}} \leftarrow \nabla_{\vec{x}}$$

- ver.2

$$\vec{x} @ W + \vec{b} \rightarrow \vec{z}$$

$$f(\vec{z}) \rightarrow \vec{y}$$

$$\vec{y} \rightarrow \vec{x},$$

(to Next layer)

$$E(\vec{y}; \vec{y}^*) \equiv \frac{1}{2} \|\vec{y} - \vec{y}^*\|^2 \rightarrow e$$

$$\nabla_{\vec{y}} \equiv \frac{\partial e}{\partial \vec{y}} = \vec{y} - \vec{y}^*, \quad (\text{at Last layer})$$

$$\nabla_{\vec{y}} \odot f'(\vec{z}) \rightarrow \nabla_{\vec{z}}$$

$$\left\{ \begin{array}{l} \nabla_{\vec{z}} \rightarrow \nabla_{\vec{b}} \\ \vec{x}^T \nabla_{\vec{z}} \rightarrow \nabla_W \\ \nabla_{\vec{z}} @ W^T \rightarrow \nabla_{\vec{x}} \end{array} \right.$$

$$\nabla_{\vec{x}} \rightarrow \nabla_{\vec{y}},$$

(to Previous layer)