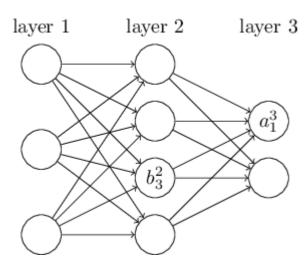
2023/4/6 下午2:16 mlp_bp



$$egin{aligned} z_j &= \sum_i x_i \cdot w_{ij} + b_j \ y_j &= f\left(z_j
ight) \ e &= rac{1}{2} \sum_j \left(y_j - y_j^*
ight)^2 \ rac{\partial e}{\partial y_j} &= y_j - y_j^* \ \delta_j &\equiv rac{\partial e}{\partial z_j} &= rac{\partial e}{\partial y_j} \cdot f'(z_j) \ rac{\partial e}{\partial w_{ij}} &= x_i \delta_j \ rac{\partial e}{\partial b_j} &= \delta_j \ rac{\partial e}{\partial x_i} &= \sum_j \delta_j w_{ji}^T \end{aligned}$$

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

• 使用 $abla_{ec{y}}(e)$

2023/4/6 下午2:16 mlp_bp

$$egin{aligned} z_j &= \sum_i x_i \cdot w_{ij} + b_j \ y_j &= f\left(z_j
ight) \ e &= rac{1}{2} \sum_j \left(y_j - y_j^*
ight)^2 \
abla_{y_j} &\equiv rac{\partial e}{\partial y_j} = y_j - y_j^* \
abla_{z_j} &\equiv rac{\partial e}{\partial z_j} =
abla_{y_j} \cdot f'(z_j) \
abla_{w_{ij}} &= x_i
abla_{z_j} \
abla_{w_{ij}} &= x_i
abla_{z_j} \
abla_{y_j} \cdot f'(z_j) \
abla_{w_{ij}} &= \nabla_{z_j} \
abla_{z_j} \cdot \nabla_{z_j}
abla_{z_j}
abla_{z_j}$$

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

$$\begin{split} \vec{z} &= \vec{x}@W + \vec{b} \\ \vec{y} &= f\left(\vec{z}\right) \\ e &= \frac{1}{2} ||\vec{y} - \vec{y}^*||^2 \\ \nabla_{\vec{y}} &\equiv \frac{\partial e}{\partial \vec{y}} = \vec{y} - \vec{y}^* \\ \nabla_{\vec{z}} &= \nabla_{\vec{y}} \odot f'(\vec{z}) \\ \nabla_{W} &= \vec{x}^T \nabla_{\vec{z}} \\ \nabla_{\vec{b}} &= \nabla_{\vec{z}} \\ \nabla_{\vec{v}} &\leftarrow \nabla_{\vec{x}} \end{split}$$

ver.2

2023/4/6 下午2:16 mlp_bp

$$egin{aligned} ec{x}@W + ec{b} &
ightarrow ec{z} \ f\left(ec{z}
ight)
ightarrow ec{y} \ ec{y}
ightarrow ec{x}, \ ext{(to Next layer)} \ E(ec{y}; ec{y}^*) &\equiv rac{1}{2} \| ec{y} - ec{y}^* \|^2
ightarrow e \
abla_{ec{y}} &\equiv rac{\partial e}{\partial ec{y}} = ec{y} - ec{y}^*, \quad ext{(at Last layer)} \
abla_{ec{y}} & \odot f'(ec{z})
ightarrow
abla_{ec{z}} \
abla_{ec{y}} & \nabla_{ec{z}}
ightarrow
abla_{ec{y}} \
abla_{ec{y}} & \nabla_{ec{z}}
ightarrow
abla_{ec{y}} \
abla_{ec{y}} & \nabla_{ec{y}}
ightarrow
abla_{ec{y}} \

abla_{ec{y}} \
abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abla_{ec{y}} \

abl$$

localhost:7765 3/3