Back propagation , in a nutshell

2023年4月2日　　下午 12:15

$$\vec{x} @ W \xrightarrow{f(\cdot)} \vec{y} \xrightarrow{\mathcal{E}(\cdot)} e$$

$$\frac{\partial e}{\partial \vec{y}} = \left[\frac{\partial e}{\partial y_1}, \frac{\partial e}{\partial y_2} \cdots \frac{\partial e}{\partial y_J} \right]$$

$$\vec{g} = \frac{\partial e}{\partial \vec{y}} \odot f'(\vec{x} @ W)$$

$$\frac{\partial e}{\partial W} = X^T @ \vec{g}$$

$$\frac{\partial e}{\partial \vec{x}} = \vec{g} @ W^T$$

$$\vec{x}@W \xrightarrow{f(\cdot)} \vec{y} \xrightarrow{\varepsilon(\cdot)} e = \left|\vec{y} - \vec{y^*}\right|^2$$

$$\frac{\partial e}{\partial \vec{y}} = \left[\frac{\partial e}{\partial y_1}, \frac{\partial e}{\partial y_2}, \cdots, \frac{\partial e}{\partial y_J}\right] = 2(\vec{y} - \vec{y^*})$$

$$\vec{g} = \frac{\partial e}{\partial \vec{y}} \odot f'(\vec{x}@W)$$

$$\frac{\partial e}{\partial W} = \vec{x}^T @ \vec{g}$$

$$\frac{\partial e}{\partial \vec{x}} = \vec{g} @ W^T$$

$\vec{x}@W \xrightarrow{f(\cdot)} \vec{y} \xrightarrow{\varepsilon(.)} \mathrm{e} = \left|\vec{y} - \overrightarrow{y^*}\right|^2$

$\vec{y} = f(\vec{x}@W)$

$\mathrm{e} = \left|\vec{y} - \overrightarrow{y^*}\right|^2$

$\frac{\partial \mathrm{e}}{\partial \vec{y}} = \left[\frac{\partial \mathrm{e}}{\partial y_1}, \frac{\partial \mathrm{e}}{\partial y_2}, \cdots, \frac{\partial \mathrm{e}}{\partial y_J}\right] = 2\left(\vec{y} - \overrightarrow{y^*}\right)$

$\vec{g} = \frac{\partial \mathrm{e}}{\partial \vec{y}} \odot f'(\vec{x}@W)$

$\frac{\partial \mathrm{e}}{\partial W} = \vec{x}^T@\vec{g}$

$\frac{\partial \mathrm{e}}{\partial \vec{x}} = \vec{g}@\mathrm{W}^{\mathrm{T}}$

$\frac{\partial \mathrm{e}}{\partial \vec{y}(l-1)} = \frac{\partial \mathrm{e}}{\partial \vec{x}(l)}$

$\vec{y}(0) \leftarrow inputVector$

for l= 1:L

$\vec{x}, W = \vec{y}(l-1), W(l)$

$\vec{y} = f(\vec{x}@W)$

$\vec{y}(l) = \vec{y}$

$\vec{y}(L) \rightarrow outputVector$

$\vec{y} = \vec{y}(L)$

$\mathrm{e} = \left|\vec{y} - \overrightarrow{y^*}\right|^2$

$\frac{\partial \mathrm{e}}{\partial \vec{y}} = \left[\frac{\partial \mathrm{e}}{\partial y_1}, \frac{\partial \mathrm{e}}{\partial y_2}, \cdots, \frac{\partial \mathrm{e}}{\partial y_J}\right] = 2\left(\vec{y} - \overrightarrow{y^*}\right)$

For l=L:1

$\vec{x}, W = \vec{x}(l), W(l)$

$\vec{g} = \frac{\partial \mathrm{e}}{\partial \vec{y}} \odot f'(\vec{x}@W)$

$\frac{\partial \mathrm{e}}{\partial W} = \vec{x}^T@\vec{g}$

$\frac{\partial \mathrm{e}}{\partial \vec{x}} = \vec{g}@\mathrm{W}^{\mathrm{T}}$

$\frac{\partial \mathrm{e}}{\partial \vec{y}} \leftarrow \frac{\partial \mathrm{e}}{\partial \vec{x}}$