BP 神经网络的反向传播计算

Sunday, September 29, 2024 4:52 PM

对BP神经网络，由于总的误差为：

$e=\frac{1}{2}\sum(d_o-y_o)^2$，我们希望得到误差对梯度

导数：

$$\frac{\partial e}{\partial w_{ho}}=\frac{\partial e}{\partial y_i}\cdot\frac{\partial y_i}{\partial w}$$

（y_i：输出，w：输入）

其中：$\frac{\partial e}{\partial y_i}=\frac{\partial(\frac{1}{2}\sum(d_o-y_o)^2)}{\partial y_i}=(d_o-y_o)\cdot-1\cdot\frac{\partial y_o}{\partial y_i}$

考虑到：

$y_o=f_2(y_i)$，则有：

$$\frac{\partial e}{\partial y_i}=-(d_o-y_o)\cdot f_2'(y_i)$$

又：$y_i=\sum w_{ho}\cdot h_o+b_{ho}$

而：对：$\frac{\partial y_i}{\partial w_{ho}}=h_o$ $\longrightarrow$ 得：$\frac{\partial e}{\partial w_{ho}}=-(d_o-y_o)f_2'(y_i)\cdot h_o$

取 $\delta_o=(d_o-y_o)f_2'(y_i)$，则：

$$\frac{\partial e}{\partial w_{ho}}=-\delta_o h_o \quad ☆$$

其中：δ_o为梯度

② 隐藏层误差：由：$\frac{\partial e}{\partial w_{ih}}=\frac{\partial e}{\partial h_i}\cdot\frac{\partial h_i}{\partial w_{ih}}$，有：

其中：$\frac{\partial e}{\partial h_i}=\frac{\partial e}{\partial h_o}\frac{\partial h_o}{\partial h_i}=\frac{\partial e}{\partial y_i}\cdot\frac{\partial y_i}{\partial h_o}\cdot\frac{\partial h_o}{\partial h_i}=-\delta_o\cdot w_{ho}\cdot\frac{\partial h_o}{\partial h_i}$

又：$h_o=f_1(h_i)$，则：$\frac{\partial e}{\partial h_i}=-\delta_o\cdot w_{ho}f_1'(h_i)$

$\Rightarrow$ 取：$\delta_h=\delta_o w_o f'(h_i)$ 为梯度，则：$\frac{\partial e}{\partial h_i}=-\delta_h$

而：$\frac{\partial h_i}{\partial w_{ih}}=x_i$ 代入，有：

$$\frac{\partial e}{\partial w_{ih}}=-\delta_o w_{ho}f_1'(h_i)\cdot x_i=-\delta_i x_i \quad ☆$$

$\frac{\partial e}{\partial w_{ih}} = -\delta_h x_i$ ☆

最终有：

① 隐含层更新公式：

由 $w^{N+1} = w^N - \frac{\partial e}{\partial w^N}$，有：

$$\begin{cases} w_{ho}^{N+1} = w_{ho}^N + \eta_1 \cdot \delta_o h_o \\ w_{ih}^{N+1} = w_{ih}^N + \eta_2 \delta_h x_i \end{cases}$$

为迭代公式 ✡