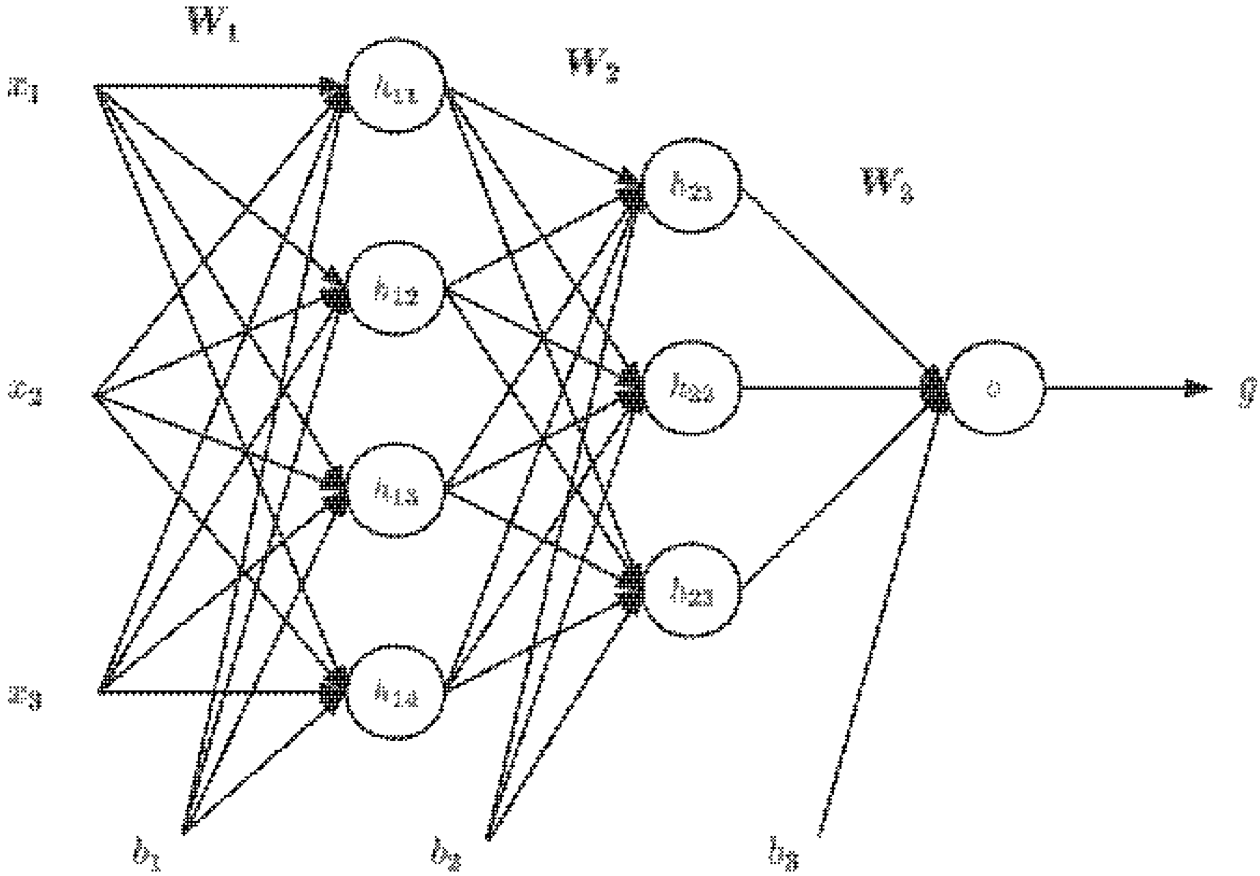
题目 1

考虑含两组隐层单元的三层前馈神经网络，如图所示。



$$\boldsymbol{W_1} = \begin{bmatrix} 0.1 & 0.1 & 0.3 \\ 0.2 & 0.3 & 0.2 \\ 0.1 & 0.2 & 0.1 \\ 0.3 & 0.1 & 0.1 \end{bmatrix}$$

$$\boldsymbol{W_2} = \boldsymbol{W_2^T}$$

$$\boldsymbol{W_3} = [0.2 \quad 0.2 \quad 0.4]$$

$$[b_1 \quad b_2 \quad b_3] = [0.5 \quad 0.5 \quad 0.2]$$

各参数的初始值如上所示。定义该网络的隐藏层单元的激活函数为 $h = \cos(z)$ ，输出单元为 Logistic 函数。

a) 当输入为 $(x_1, x_2, x_3) = (0.05, 0.10, 0.05)$ 时，计算该神经网络输出 y 的值。请写明必要的计算过程。

b) 在 a) 的基础上，若 $y = 0.95$,采用最小化均方误差作为优化准则，请根据 BP 算法计算参数 W_3 的梯度。

c) 在 b) 的基础上，若采用梯度下降更新参数，且学习率设置为 0.1，写出更新后的参数 W_3 。

解：

a）

隐藏层 Layer 1：

$$a^{[0]} = x$$

$$z^{[1]} = \boldsymbol{W_1} a^{[0]} + b_1 = \begin{bmatrix} 0.1 & 0.1 & 0.3 \\ 0.2 & 0.3 & 0.2 \\ 0.1 & 0.2 & 0.1 \\ 0.3 & 0.1 & 0.1 \end{bmatrix} \begin{bmatrix} 0.05 \\ 0.10 \\ 0.05 \end{bmatrix} + \begin{bmatrix} 0.5 \\ 0.5 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 0.53 \\ 0.55 \\ 0.53 \\ 0.53 \end{bmatrix}$$

隐藏层 Layer 2：

$$a^{[1]} = h(z^{[1]}) = \begin{bmatrix} 0.8628 \\ 0.8525 \\ 0.8628 \\ 0.8628 \end{bmatrix}$$

$$z^{[2]} = \boldsymbol{W_2} a^{[1]} + b_2 = \begin{bmatrix} 0.1 & 0.2 & 0.1 & 0.3 \\ 0.1 & 0.3 & 0.2 & 0.1 \\ 0.3 & 0.2 & 0.1 & 0.1 \end{bmatrix} \begin{bmatrix} 0.8628 \\ 0.8525 \\ 0.8628 \\ 0.8628 \end{bmatrix} + \begin{bmatrix} 0.5 \\ 0.5 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 1.1019 \\ 1.1009 \\ 1.1019 \end{bmatrix}$$

隐藏层 Layer 3：

$$a^{[2]} = h(z^{[2]}) = \begin{bmatrix} 0.4519 \\ 0.4528 \\ 0.4519 \end{bmatrix}$$

$$z^{[3]} = \boldsymbol{W_3} a^{[2]} + b_3 = \begin{bmatrix} 0.2 & 0.2 & 0.4 \end{bmatrix} \begin{bmatrix} 0.4519 \\ 0.4528 \\ 0.4519 \end{bmatrix} + 0.2 = 0.5617$$

输出单元为 Logistic 函数：

$$\hat{y} = \sigma(z^{[3]}) = \frac{1}{1 + e^{-z^{[3]}}} = 0.6368$$

b）已知 $y = 0.95$，$\hat{y} = 0.6368$，

损失函数

$$e = (y - \hat{y})^2$$

模型输出

$$\hat{y} = a^{[3]} = \sigma(z^{[3]}) = \frac{1}{1+e^{-z^{[3]}}}$$

而

$$z^{[3]} = \boldsymbol{W_3} a^{[2]} + b_3$$

$$\frac{\partial e}{\partial \boldsymbol{W_3}} = \frac{\partial e}{\partial y} \frac{\partial y}{\partial z^{[3]}} \frac{\partial z^{[3]}}{\partial \boldsymbol{W_3}} = 2(y - \hat{y})$$

因此，链式法则分别求导：

$$\begin{aligned}
\frac{\partial e}{\partial \hat{y}} &= 2(\hat{y} - y) \\
\frac{\partial \hat{y}}{\partial z^{[3]}} &= \frac{e^{-z^{[3]}}}{(1+e^{-z^{[3]}})^2} = \hat{y}(1-\hat{y}) \\
\frac{\partial z^{[3]}}{\partial \boldsymbol{W_3}} &= a^{[2]^T} \\
\Rightarrow \frac{\partial e}{\partial \boldsymbol{W_3}} &= \frac{\partial e}{\partial y} \cdot \frac{\partial y}{\partial z^{[3]}} \cdot \frac{\partial z^{[3]}}{\partial \boldsymbol{W_3}} = 2(\hat{y} - y)\hat{y}(1-\hat{y})a^{[2]^T} \\
&= \begin{bmatrix} -0.0655 & -0.0656 & -0.0655 \end{bmatrix}
\end{aligned}$$

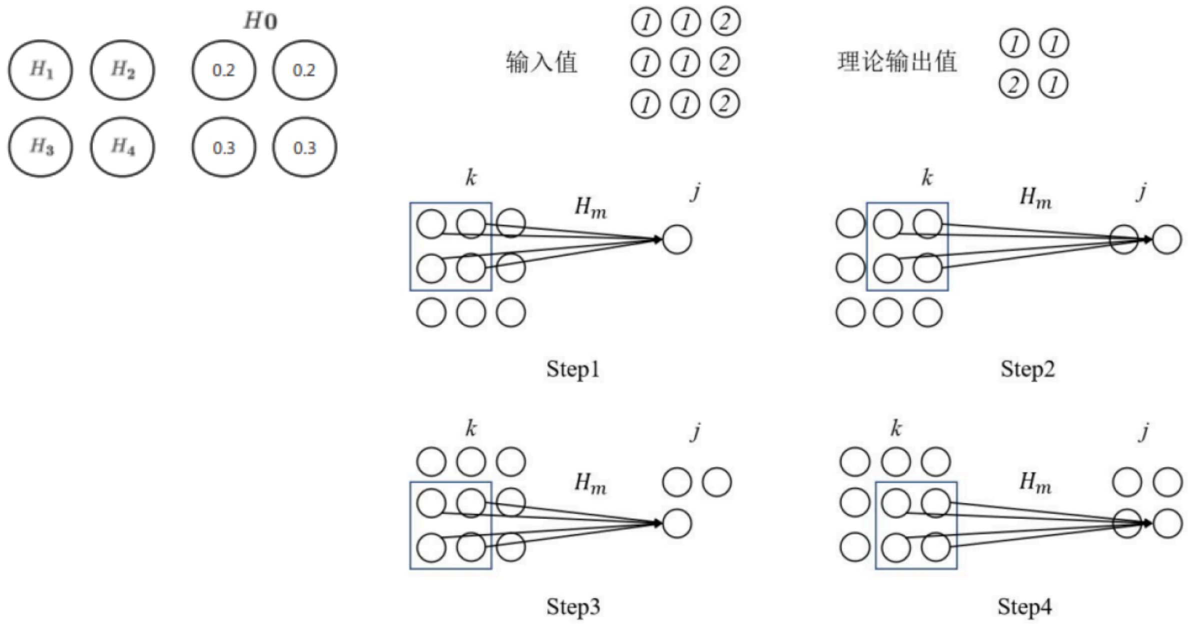
c）

$$\boldsymbol{W_3} := \boldsymbol{W_3} - \eta \frac{\partial e}{\partial \boldsymbol{W_3}}$$

$$\boldsymbol{W_3} = \begin{bmatrix} 0.2065 & 0.2066 & 0.4065 \end{bmatrix}$$

题目 3

尝试对单卷积层的神经网络权值 H_m 进行更新。初始权向量 H_0，输入值与理论输出值如下图所示。



神经网络损失函数为 $L=\frac{1}{2}(y-d)^2$，学习率 $\alpha=0.5$。其中 y 为网络输出，d 为理论输出。如图，卷积层将分四步扫描输入值，并返回四个输出。请根据输入值与初始权值，计算该网络输出值，并根据理论输出值更新权值参数。

解：

$$X=\begin{bmatrix}1&1&2\\1&1&2\\1&1&2\end{bmatrix}\quad H_0=\begin{bmatrix}0.2&0.2\\0.3&0.3\end{bmatrix}\quad d=\begin{bmatrix}1&1\\2&1\end{bmatrix}$$

$$\hat{y}_{11}=\begin{bmatrix}1&1\\1&1\end{bmatrix}*\begin{bmatrix}0.2&0.2\\0.3&0.3\end{bmatrix}=1\quad \hat{y}_{12}=\begin{bmatrix}1&2\\1&2\end{bmatrix}*\begin{bmatrix}0.2&0.2\\0.3&0.3\end{bmatrix}=1.5$$

$$\hat{y}_{21}=\begin{bmatrix}1&1\\1&1\end{bmatrix}*\begin{bmatrix}0.2&0.2\\0.3&0.3\end{bmatrix}=1\quad \hat{y}_{22}=\begin{bmatrix}1&2\\1&2\end{bmatrix}*\begin{bmatrix}0.2&0.2\\0.3&0.3\end{bmatrix}=1.5$$

$$\therefore\ \hat{y}=\begin{bmatrix}1&1.5\\1&1.5\end{bmatrix}$$

$$L=\frac{1}{2}(\hat{y}-d)^2\quad\Rightarrow\quad\frac{\partial L}{\partial\hat{y}}=\hat{y}-d=\begin{bmatrix}0&0.5\\-1&0.5\end{bmatrix}$$

$$\frac{\partial L}{\partial h_{ij}}=\sum_{m,n}\frac{\partial L}{\partial\hat{y}_{m,n}}\cdot\frac{\partial\hat{y}_{m,n}}{\partial h_{ij}}$$

$$\frac{\partial\hat{y}}{\partial h_{11}}=\begin{bmatrix}1&1\\1&1\end{bmatrix}\quad\frac{\partial\hat{y}}{\partial h_{12}}=\begin{bmatrix}1&2\\1&2\end{bmatrix}$$

$$\frac{\partial\hat{y}}{\partial h_{21}}=\begin{bmatrix}1&1\\1&1\end{bmatrix}\quad\frac{\partial\hat{y}}{\partial h_{22}}=\begin{bmatrix}1&2\\1&2\end{bmatrix}$$

$$\frac{\partial L}{\partial h_{11}} = \frac{\partial L}{\partial \hat{y}_{11}} \cdot \frac{\partial \hat{y}_{11}}{\partial h_{11}} + \frac{\partial L}{\partial \hat{y}_{12}} \cdot \frac{\partial \hat{y}_{12}}{\partial h_{11}} + \frac{\partial L}{\partial \hat{y}_{21}} \cdot \frac{\partial \hat{y}_{21}}{\partial h_{11}} + \frac{\partial L}{\partial \hat{y}_{22}} \cdot \frac{\partial \hat{y}_{22}}{\partial h_{11}}$$

$$= 0 \cdot 1 + 0.5 \cdot 1 - 1 \cdot 1 + 0.5 \cdot 1 = 0$$

$$\frac{\partial L}{\partial h_{12}} = \frac{\partial L}{\partial \hat{y}_{11}} \cdot \frac{\partial \hat{y}_{11}}{\partial h_{12}} + \frac{\partial L}{\partial \hat{y}_{12}} \cdot \frac{\partial \hat{y}_{12}}{\partial h_{12}} + \frac{\partial L}{\partial \hat{y}_{21}} \cdot \frac{\partial \hat{y}_{21}}{\partial h_{12}} + \frac{\partial L}{\partial \hat{y}_{22}} \cdot \frac{\partial \hat{y}_{22}}{\partial h_{12}}$$

$$= 0 \cdot 1 + 0.5 \cdot 2 - 1 \cdot 1 + 0.5 \cdot 2 = 1$$

$$\frac{\partial L}{\partial h_{21}} = \frac{\partial L}{\partial \hat{y}_{11}} \cdot \frac{\partial \hat{y}_{11}}{\partial h_{21}} + \frac{\partial L}{\partial \hat{y}_{12}} \cdot \frac{\partial \hat{y}_{12}}{\partial h_{21}} + \frac{\partial L}{\partial \hat{y}_{21}} \cdot \frac{\partial \hat{y}_{21}}{\partial h_{21}} + \frac{\partial L}{\partial \hat{y}_{22}} \cdot \frac{\partial \hat{y}_{22}}{\partial h_{21}}$$

$$= 0 \cdot 1 + 0.5 \cdot 1 - 1 \cdot 1 + 0.5 \cdot 1 = 0$$

$$\frac{\partial L}{\partial h_{22}} = \frac{\partial L}{\partial \hat{y}_{11}} \cdot \frac{\partial \hat{y}_{11}}{\partial h_{22}} + \frac{\partial L}{\partial \hat{y}_{12}} \cdot \frac{\partial \hat{y}_{12}}{\partial h_{22}} + \frac{\partial L}{\partial \hat{y}_{21}} \cdot \frac{\partial \hat{y}_{21}}{\partial h_{22}} + \frac{\partial L}{\partial \hat{y}_{22}} \cdot \frac{\partial \hat{y}_{22}}{\partial h_{22}}$$

$$= 0 \cdot 1 + 0.5 \cdot 2 - 1 \cdot 1 + 0.5 \cdot 2 = 1$$

$$\text{then } \frac{\partial L}{\partial H_0} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$$

$$\text{we have } H_1 \leftarrow H_0 - \alpha \frac{\partial L}{\partial H_0} = \begin{bmatrix} 0.2 & 0.2 \\ 0.3 & 0.3 \end{bmatrix} - 0.5 \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.2 & -0.3 \\ 0.3 & -0.2 \end{bmatrix}$$