$$W(1)=W(0)+\eta f(\mathrm{net}^1)X^1=(1.905,-2.81,1.357,0.5)^{\mathrm{T}}$$

(2) $\mathrm{net}^2=W(1)^{\mathrm{T}}X^2=-0.154$

$$o^2=f(\mathrm{net}^2)=\frac{1-\mathrm{e}^{-\mathrm{net}}}{1+\mathrm{e}^{-\mathrm{net}}}=-0.077$$

$$W(2)=W(1)+\eta f(\mathrm{net}^2)X^2=(1.828,-2.772,1.512,0.616)^{\mathrm{T}}$$

(3) $\mathrm{net}^3=W(2)^{\mathrm{T}}X^3=-3.36$

$$o^3=f(\mathrm{net}^3)=\frac{1-\mathrm{e}^{-\mathrm{net}}}{1+\mathrm{e}^{-\mathrm{net}}}=-0.932$$

$$W(3)=W(2)+\eta f(\mathrm{net}^3)X^3=(1.828,-3.70,2.44,-0.785)^{\mathrm{T}}$$

比较两种权值调整结果可以看出，两种转移函数下的权值调整方向是一致的，但采用连续转移函数时，权值调整力度减弱。

2.4.2 Perceptron 学习规则

1958 年，美国学者 Frank Rosenblatt 首次定义了一个具有单层计算单元的神经网络结构，称为 Perceptron（感知器）。感知器的学习规则规定，学习信号等于神经元期望输出（教师信号）与实际输出之差：

$$r=d_j-o_j \tag{2.17}$$

式中，d_j 为期望的输出，$o_j=f(W_j{}^{\mathrm{T}}X)$。感知器采用了符号函数作为转移函数，其表达为：

$$f(W_j^{\mathrm{T}}X)=\mathrm{sgn}(W_j^{\mathrm{T}}X)=\begin{cases}1, & W_j^{\mathrm{T}}X\geqslant 0\\ -1, & W_j^{\mathrm{T}}X<0\end{cases} \tag{2.18}$$

因此，权值调整公式应为：

$$\Delta W_j=\eta[d_j-\mathrm{sgn}(W_j^{\mathrm{T}}X)]X \tag{2.19a}$$

？特指？

$$\Delta w_{ij}=\eta[d_j-\mathrm{sgn}(W_j^{\mathrm{T}}X)]x_i \qquad i=0,1,\cdots,n \tag{2.19b}$$

← 变量数

式中，当实际输出与期望值相同时，权值不需要调整；在有误差存在情况下，由于 d_j 和 $\mathrm{sgn}(W_j^{\mathrm{T}}X)\in\{-1,1\}$，权值调整公式可简化为：

$$\Delta W_j=\pm 2\eta X \tag{2.19c}$$

感知器学习规则只适用于二进制神经元，初始权值可取任意值。

感知器学习规则代表一种有导师学习。由于感知器理论是研究其他神经网络的基础，该规则对于神经网络的有导师学习具有极为重要的意义。

2.4.3 δ 学习规则

1986 年，认知心理学家 McClelland 和 Rumelhart 在神经网络训练中引入了 δ（Delta）规则，该规则亦可称为连续感知器学习规则，与上述离散感知器学习规则并行。δ 规则的学习信号规定为：

$$\begin{aligned} r&=[d_j-f(W_j^{\mathrm{T}}X)]f'(W_j^{\mathrm{T}}X)\\ &=(d_j-o_j)f'(\mathrm{net}_j)\end{aligned} \tag{2.20}$$

上式定义的学习信号称为 δ。式中，$f'(W_j^{\mathrm{T}}X)$ 是转移函数 $f(\mathrm{net}_j)$ 的导数。显然，δ 规则要求转移函数可导，因此只适用于有导师学习中定义的连续转移函数，如 Sigmoid 函数。

事实上，δ 规则很容易由输出值与期望值的最小平方误差条件推导出来。定义神经元输