出与期望输出之间的平方误差为：

$$E=\frac{1}{2}(d_j-o_j)^2$$
$$=\frac{1}{2}[d_j-f(\boldsymbol{W}_j{}^{\mathrm{T}}\boldsymbol{X})]^2 \tag{2.21}$$

其中，误差 E 是权向量 $\boldsymbol{W}_j$ 的函数。欲使误差 E 最小，$\Delta\boldsymbol{W}_j$ 应与误差的负梯度成正比，即：

$$\Delta\boldsymbol{W}_j=-\eta\nabla E \tag{2.22}$$

式中，比例系数 η 是一个正常数。由式(2.21)，误差梯度为：

$$\nabla E=-(d_j-o_j)f'(\boldsymbol{W}_j^{\mathrm{T}}\boldsymbol{X})\boldsymbol{X} \tag{2.23}$$

将此结果代入式(2.22)，可得权值调整计算式：

$$\Delta\boldsymbol{W}_j=\eta(d_j-o_j)f'(\mathrm{net}_j)\boldsymbol{X} \tag{2.24a}$$

可以看出，上式中 η 与 $\boldsymbol{X}$ 之间的部分正是式(2.20)中定义的学习信号 δ。$\Delta\boldsymbol{W}_j$ 中每个分量的调整由下式计算：

$$\Delta w_{ij}=\eta(d_j-o_j)f'(\mathrm{net}_j)x_i \qquad i=0,1,\cdots,n \tag{2.24b}$$

δ 学习规则可推广到多层前馈网络中，权值可初始化为任意值。

下面举例说明 δ 学习规则的应用。

例 2.2 设有 3 输入单输出神经元网络，将阈值含于权向量内，故有 $w_0=T$，$x_0=-1$，学习率 $\eta=0.1$，3 个输入向量和初始权向量分别为 $\boldsymbol{X}^1=(-1,1,-2,0)^{\mathrm{T}}$，$\boldsymbol{X}^2=(-1,0,1.5,-0.5)^{\mathrm{T}}$，$\boldsymbol{X}^3=(-1,-1,1,0.5)^{\mathrm{T}}$，$d^1=-1$，$d^2=-1$，$d^3=1$，$\boldsymbol{W}(0)=(0.5,1,-1,0)^{\mathrm{T}}$。

解 设转移函数为双极性连续函数 $f(\mathrm{net})=\frac{1-\mathrm{e}^{-\mathrm{net}}}{1+\mathrm{e}^{-\mathrm{net}}}$，权值调整步骤为：

(1) 输入样本 $\boldsymbol{X}^1$，计算净输入 net^1 及权向量 $\boldsymbol{W}(1)$

$$\mathrm{net}^1=\boldsymbol{W}(0)^{\mathrm{T}}\boldsymbol{X}^1=2.5$$

$$o^1=f(\mathrm{net}^1)=\frac{1-\mathrm{e}^{-\mathrm{net}^1}}{1+\mathrm{e}^{-\mathrm{net}^1}}=0.848$$

$$f'(\mathrm{net}^1)=\frac{1}{2}[1-(o^1)^2]=0.14$$

$$\boldsymbol{W}(1)=\boldsymbol{W}(0)+\eta(d^1-o^1)f'(\mathrm{net}^1)\boldsymbol{X}^1$$
$$=(0.526,0.974,-0.948,0)^{\mathrm{T}}$$

(2) 输入样本 $\boldsymbol{X}^2$，计算净输入 net^2 及权向量 $\boldsymbol{W}(2)$

$$\mathrm{net}^2=\boldsymbol{W}(1)^{\mathrm{T}}\boldsymbol{X}^2=-1.948$$

$$o^2=f(\mathrm{net}^2)=\frac{1-\mathrm{e}^{-\mathrm{net}^2}}{1+\mathrm{e}^{-\mathrm{net}^2}}=-0.75$$

$$f'(\mathrm{net}^2)=\frac{1}{2}[1-(o^2)^2]=0.218$$

$$\boldsymbol{W}(2)=\boldsymbol{W}(1)+\eta(d^2-o^2)f'(\mathrm{net}^2)\boldsymbol{X}^2$$
$$=(0.531,0.974,-0.956,0.002)^{\mathrm{T}}$$

(3) 输入样本 $\boldsymbol{X}^3$，计算净输入 net^3 及权向量 $\boldsymbol{W}(3)$

$$\mathrm{net}^3=\boldsymbol{W}(2)^{\mathrm{T}}\boldsymbol{X}^3=-2.416$$

$$o^3=f(\mathrm{net}^3)=\frac{1-\mathrm{e}^{-\mathrm{net}^3}}{1+\mathrm{e}^{-\mathrm{net}^3}}=-0.842$$