$$f'(\mathrm{net}^3)=\frac{1}{2}[1-(o^3)^2]=0.145$$

$$\boldsymbol{W}(3)=\boldsymbol{W}(2)+\eta(d^3-o^3)f'(\mathrm{net}^3)\boldsymbol{X}^3$$
$$=(0.505,0.947,-0.929,0.016)^{\mathrm{T}}$$

2.4.4 LMS 学习规则

1962 年，Bernard Widrow 和 Marcian Hoff 提出了 Widrow-Hoff 学习规则。因为它能使神经元实际输出与期望输出之间的平方差最小，所以又称为最小均方规则（LMS）。LMS 学习规则的学习信号为：

$$r=d_j-\boldsymbol{W}_j{}^{\mathrm{T}}\boldsymbol{X} \tag{2.25}$$

权向量调整量为：

$$\Delta\boldsymbol{W}_j=\eta(d_j-\boldsymbol{W}_j^{\mathrm{T}}\boldsymbol{X})\boldsymbol{X} \tag{2.26a}$$

$\Delta\boldsymbol{W}_j$ 的各分量为：

$$\Delta w_{ij}=\eta(d_j-\boldsymbol{W}_j{}^{\mathrm{T}}\boldsymbol{X})x_j \qquad i=0,1,\cdots,n \tag{2.26b}$$

实际上，如果在 δ 学习规则中假定神经元转移函数为 $f(\boldsymbol{W}_j{}^{\mathrm{T}}\boldsymbol{X})=\boldsymbol{W}_j{}^{\mathrm{T}}\boldsymbol{X}$，则有 $f'(\boldsymbol{W}_j{}^{\mathrm{T}}\boldsymbol{X})=1$，此时式(2.20) 与式(2.25) 相同。因此，LMS 学习规则可以看成是 δ 学习规则的一个特殊情况。该学习规则与神经元采用的转移函数无关，因而不需要对转移函数求导数，不仅学习速度较快，而且具有较高的精度。权值可初始化为任意值。

2.4.5 Correlation 学习规则

Correlation（相关）学习规则规定学习信号为：

$$r=d_j \tag{2.27}$$

易得出 $\Delta\boldsymbol{W}_j$ 及 Δw_{ij} 分别为：

$$\Delta\boldsymbol{W}_j=\eta\, d_j\boldsymbol{X} \tag{2.28a}$$

$$\Delta w_{ij}=\eta\, d_jx_i \qquad i=0,1,\cdots,n \tag{2.28b}$$

该规则表明，当 d_j 是 x_i 的期望输出时，相应的权值增量 Δw_{ij} 与两者的乘积 d_jx_i 成正比。

如果 Hebb 学习规则中的转移函数为二进制函数，且有 $o_j=d_j$，则相关学习规则可看作 Hebb 规则的一种特殊情况。应当注意的是，Hebb 学习规则是无导师学习，而相关学习规则是有导师学习。这种学习规则要求将权值初始化为零。

2.4.6 Winner-Take-All 学习规则

Winner-Take-All（胜者为王）学习规则是一种竞争学习规则，用于无导师学习。一般将网络的某一层确定为竞争层，对于一个特定的输入 X，竞争层的所有 p 个神经元均有输出响应，其中响应值最大的神经元为在竞争中获胜的神经元，即：

$$\boldsymbol{W}_{j^*}^{\mathrm{T}}\boldsymbol{X}=\max_{i=1,2,\cdots,p}(\boldsymbol{W}_i{}^{\mathrm{T}}\boldsymbol{X}) \tag{2.29}$$

只有获胜神经元才有权调整其权向量 $\boldsymbol{W}_m$，调整量为：

$$\Delta\boldsymbol{W}_{j^*}=\alpha(\boldsymbol{X}-\boldsymbol{W}_{j^*}) \tag{2.30}$$

式中，$\alpha\in(0,1]$，是一个小的学习常数，一般其值随着学习的进展而减小。由于两个向量的点积越大，表明两者越近似，所以调整获胜神经元权值的结果是使 $\boldsymbol{W}_m$ 进一步接近当前输入 $\boldsymbol{X}$。显然，当下次出现与 $\boldsymbol{X}$ 相像的输入模式时，上次获胜的神经元更容易获胜。在反复的竞争学习过程中，竞争层的各神经元所对应的权向量被逐渐调整为输入样本空间的聚类中心。