则第三层的输入为：

$$s_1^{(3)} = (\mathbf{w}^{(3)})^T \begin{pmatrix} 1 \\ x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{pmatrix} = (0.64 \quad -1.91 \quad -1.91 \quad -1.91) \begin{pmatrix} 1 \\ 0.46 \\ 0.46 \\ 0.46 \end{pmatrix} = -2.00$$

即输出$\hat{y} = s_1^{(3)} = -2.00$

对于样本$\vec{x}_3$，其标签为-1，采用平方误差函数：$e_n = (y_n - \hat{y}_n)^2$，则：

$$\delta_1^{(3)} = -2\left(y_n - s_1^{(3)}\right) = -2\left(-1 - (-2.00)\right) = -2.00$$

运用反向传播法，于是：

$$\delta_j^{(2)} = \sum\nolimits_k (\delta_k^{(3)})(w_{jk}^{(3)})[\![s_j^{(2)} \geq 0]\!] = \delta_1^{(3)} w_{j1}^{(3)} [\![s_j^{(2)} \geq 0]\!]$$

即：$\vec{\delta}^{(2)} = \begin{pmatrix} \delta_1^{(2)} \\ \delta_2^{(2)} \\ \delta_3^{(2)} \end{pmatrix} = \begin{pmatrix} \delta_1^{(3)} w_{11}^{(3)} [\![s_1^{(2)} \geq 0]\!] \\ \delta_1^{(3)} w_{21}^{(3)} [\![s_2^{(2)} \geq 0]\!] \\ \delta_1^{(3)} w_{31}^{(3)} [\![s_3^{(2)} \geq 0]\!] \end{pmatrix} = \begin{pmatrix} (-2.00) * (-1.91) * 1 \\ (-2.00) * (-1.91) * 1 \\ (-2.00) * (-1.91) * 1 \end{pmatrix} = \begin{pmatrix} 3.82 \\ 3.82 \\ 3.82 \end{pmatrix}$

继续运用反向传播法，于是：$\delta_j^{(1)} = \sum_k (\delta_k^{(2)})(w_{jk}^{(2)})(x_j^{(1)})'$，所以：

$$\delta_j^{(1)} = \sum\nolimits_k (\delta_k^{(2)})(w_{jk}^{(2)})[\![s_j^{(1)} \geq 0]\!] = (\delta_1^{(2)} w_{j1}^{(2)} + \delta_2^{(2)} w_{j2}^{(2)} + \delta_3^{(2)} w_{j3}^{(2)})[\![s_j^{(1)} \geq 0]\!]$$

由此可以得到：

$$\vec{\delta}^{(1)} = \begin{pmatrix} \delta_1^{(1)} \\ \delta_2^{(1)} \end{pmatrix} = \begin{pmatrix} [\![s_1^{(1)} \geq 0]\!] & 0 \\ 0 & [\![s_2^{(1)} \geq 0]\!] \end{pmatrix} \begin{pmatrix} w_{11}^{(2)} & w_{12}^{(2)} & w_{13}^{(2)} \\ w_{21}^{(2)} & w_{22}^{(2)} & w_{23}^{(2)} \end{pmatrix} \begin{pmatrix} \delta_1^{(2)} \\ \delta_2^{(2)} \\ \delta_3^{(2)} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} -0.29 & -0.29 & -0.29 \\ -0.29 & -0.29 & -0.29 \end{pmatrix} \begin{pmatrix} 3.82 \\ 3.82 \\ 3.82 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

令$\eta = 0.01$，利用梯度下降法进行权系数更新：

版权归SMART-AIA-HUST所有，严禁复制