神经网络作业解答:

1 PPT 选择题答案：3 4 4

2，（1）因为是 RGB 图像，所以共有 3 个通道，全连接情况下包含的参数是：300*300（图像大小）*3（颜色通道数）*100（第一层神经元个数）+1*100（每个神经元都要与输入层的常数项连接）=27000100

（2）卷积操作时，第一层的参数是由卷积滤波器大小和常数项共同确定的，因此其包含的参数为：（5*5（滤波器大小）*3（颜色通道数）+1（常数项））*100（滤波器个数）=7600；

第一层神经元个数：由于我们没有讲像素填充问题，所以，习题不考虑像素填充情况，那么，卷积后第一层神经元的个数分两种情况：第一种情况是滤波器为 5*5*3，这时神经元个数为（(300（图像长或宽）-5（滤波器大小）/1（移动步长））+1）^2*100（滤波器个数）=8761600

第二种情况是滤波器为 5*5*1，这时要考虑通道数（(300（图像长或宽）-5（滤波器大小）/1（移动步长））+1）^2*3（颜色通道数）*100（滤波器个数）=26284800。但请大家注意的是当前对彩色图像用滤波器卷积时一般用的都是 5*5*3，也就是第一种情况。

3，假设训练样本集有 N 个样本$\{\vec{x}_1, \dots \vec{x}_n, \dots \vec{x}_N\}$，每个样本有 d 维特征，写成增广向量后是 d+1 维，$\vec{x}_n = (1, x_{n1}, \dots x_{nd})^T$，将神经网络的输入层当第 0 层，所以写为：$\vec{x}_n^{(0)} = (1, x_{n1}^{(0)}, \dots x_{nd}^{(0)})^T$，当 d=2 时，$\vec{x}_n^{(0)} = (1, x_{n1}^{(0)}, x_{n2}^{(0)})^T$

假设第一层有两个神经元，第二层有三个神经元，第三层有一个神经元。

第一层、第二层和第三层的权系数矩阵分别为：

$$\mathbf{w}^{(1)}=\begin{pmatrix} w_{01}^{(1)} & w_{02}^{(1)} \\ w_{11}^{(1)} & w_{12}^{(1)} \\ w_{21}^{(1)} & w_{22}^{(1)} \end{pmatrix}，\ \mathbf{w}^{(2)}=\begin{pmatrix} w_{01}^{(2)} & w_{02}^{(2)} & w_{03}^{(2)} \\ w_{11}^{(2)} & w_{12}^{(2)} & w_{13}^{(2)} \\ w_{21}^{(2)} & w_{22}^{(2)} & w_{23}^{(2)} \end{pmatrix}，\ \mathbf{w}^{(3)}=\begin{pmatrix} w_{01}^{(3)} \\ w_{11}^{(3)} \\ w_{21}^{(3)} \\ w_{31}^{(3)} \end{pmatrix}$$

则第一层神经元的输入为：

$$\begin{pmatrix} s_1^{(1)} \\ s_2^{(1)} \end{pmatrix}=(\mathbf{w}^{(1)})^T\vec{x}_n^{(0)}$$

假设第一层神经元的激活函数为ReLU，即：$x^{(1)}=\max(0,s^{(1)})$，则：

$$\begin{pmatrix} x_1^{(1)} \\ x_2^{(1)} \end{pmatrix}=\begin{pmatrix} \max(0,s_1^{(1)}) \\ \max(0,s_2^{(1)}) \end{pmatrix}$$

第二层神经元的输入为：

$$\begin{pmatrix} s_1^{(2)} \\ s_2^{(2)} \\ s_3^{(2)} \end{pmatrix}=(\mathbf{w}^{(2)})^T\begin{pmatrix} 1 \\ x_1^{(1)} \\ x_2^{(1)} \end{pmatrix}$$

假设第二层神经元的激活函数为ReLU，即：$x^{(2)}=\max(0,s^{(2)})$，则：

$$\begin{pmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{pmatrix}=\begin{pmatrix} \max(0,s_1^{(2)}) \\ \max(0,s_2^{(2)}) \\ \max(0,s_3^{(2)}) \end{pmatrix}$$

则第三层的输入为：

$$\mathrm{s}_1^{(3)}=(\mathbf{w}^{(3)})^T\begin{pmatrix} 1 \\ x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{pmatrix}$$

因为第三层是线性操作，即输出$\hat{y}=\mathrm{s}_1^{(3)}$

对于输入样本$\vec{x}_n$，假设其标签为y_n，采用平方误差函数。即：$e_n=(y_n-\hat{y}_n)^2$

$$\delta_1^{(3)}=-2(y_n-\mathrm{s}_1^{(3)})$$

运用反向传播法，于是： $\delta_j^{(2)} = \sum_k (\delta_k^{(3)})(w_{jk}^{(3)})(x_j^{(2)})'$

对于ReLU来说，其导数为：$(x_j^{(L)})' = [\![s_j^{(L-1)} \geq 0]\!]$

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

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

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：

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

假定t表示迭代次数，η为学习步长，利用梯度下降法进行权系数更新：

$$\mathbf{w}_{t+1}^{(1)} = \mathbf{w}_t^{(1)} - \eta \vec{x}_n^{(0)} (\overrightarrow{\delta}^{(1)})^T = \mathbf{w}_t^{(1)} - \eta \begin{pmatrix} 1 \\ x_{n1}^{(0)} \\ x_{n2}^{(0)} \end{pmatrix} \left(\delta_1^{(1)}, \delta_2^{(1)} \right)$$

$$= \begin{pmatrix} w_{01}^{(1)} & w_{02}^{(1)} \\ w_{11}^{(1)} & w_{12}^{(1)} \\ w_{21}^{(1)} & w_{22}^{(1)} \end{pmatrix} - \eta \begin{pmatrix} \delta_1^{(1)} & \delta_2^{(1)} \\ x_{n1}^{(0)} \delta_1^{(1)} & x_{n1}^{(0)} \delta_2^{(1)} \\ x_{n2}^{(0)} \delta_1^{(1)} & x_{n2}^{(0)} \delta_2^{(1)} \end{pmatrix}$$

$$\mathbf{w}_{t+1}^{(2)} = \mathbf{w}_t^{(2)} - \eta \vec{x}_n^{(1)} (\overrightarrow{\delta}^{(2)})^T = \mathbf{w}_t^{(2)} - \eta \begin{pmatrix} 1 \\ x_1^{(1)} \\ x_2^{(1)} \end{pmatrix} \left(\delta_1^{(2)}, \delta_2^{(2)}, \delta_3^{(2)} \right)$$

$$= \begin{pmatrix} w_{01}^{(2)} & w_{02}^{(2)} & w_{03}^{(2)} \\ w_{11}^{(2)} & w_{12}^{(2)} & w_{13}^{(2)} \\ w_{21}^{(2)} & w_{22}^{(2)} & w_{23}^{(2)} \end{pmatrix} - \eta \begin{pmatrix} \delta_1^{(2)} & \delta_2^{(2)} & \delta_3^{(2)} \\ x_1^{(1)} \delta_1^{(2)} & x_1^{(1)} \delta_2^{(2)} & x_1^{(1)} \delta_3^{(2)} \\ x_2^{(1)} \delta_1^{(2)} & x_2^{(1)} \delta_2^{(2)} & x_2^{(1)} \delta_3^{(2)} \end{pmatrix}$$

$$\mathbf{w}_{t+1}^{(3)} = \mathbf{w}_t^{(3)} - \eta \vec{x}_n^{(2)} (\overrightarrow{\delta}^{(3)})^T = \mathbf{w}_t^{(3)} - \eta \begin{pmatrix} 1 \\ x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{pmatrix} \delta_1^{(3)} = \begin{pmatrix} w_{01}^{(3)} \\ w_{11}^{(3)} \\ w_{21}^{(3)} \\ w_{31}^{(3)} \end{pmatrix} - \eta \begin{pmatrix} \delta_1^{(3)} \\ x_1^{(2)} \delta_1^{(3)} \\ x_2^{(2)} \delta_1^{(3)} \\ x_3^{(2)} \delta_1^{(3)} \end{pmatrix}$$

反复迭代至T次。

习题解答:

有训练样本集为: $\mathrm{D}=\{(\vec{x}_1,y_1)=((1,1)^T,1),(\vec{x}_2,y_2)=((-1,-1)^T,1),(\vec{x}_3,y_3)=((-1,1)^T,-1),(\vec{x}_4,y_4)=((1,-1)^T,-1)\}$, 假设某神经网络结构为第一层有两个神经元，第二层有三个神经元，第三层有一个神经元，前两层每个神经元的激活函数为ReLU，第三层为线性输出，即$\hat{y}=\mathrm{s}_1^{(3)}$，假设误差函数为: $E_{in}=\frac{1}{N}\sum_n(y_n-\hat{y}_n)^2$，学习率为0.01。

t=0

$$\mathbf{w}_0^{(1)}=\begin{pmatrix}1&1\\1&1\\1&1\end{pmatrix},\ \mathbf{w}_0^{(2)}=\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix},\ \mathbf{w}_0^{(3)}=\begin{pmatrix}1\\1\\1\\1\end{pmatrix}$$

t=1时，对于第一个样本$\vec{x}_1=(1,1)^T$，则第一层神经元的输入为:

$$\begin{pmatrix}s_1^{(1)}\\s_2^{(1)}\end{pmatrix}=(\mathbf{w}^{(1)})^T\vec{x}_n^{(0)}=\begin{pmatrix}1&1&1\\1&1&1\end{pmatrix}\begin{pmatrix}1\\1\\1\end{pmatrix}=\begin{pmatrix}3\\3\end{pmatrix}$$

则: $\begin{pmatrix}x_1^{(1)}\\x_2^{(1)}\end{pmatrix}=\begin{pmatrix}\max(0,s_1^{(1)})\\\max(0,s_2^{(1)})\end{pmatrix}=\begin{pmatrix}3\\3\end{pmatrix}$

第二层神经元的输入为:

$$\begin{pmatrix}s_1^{(2)}\\s_2^{(2)}\\s_3^{(2)}\end{pmatrix}=(\mathbf{w}^{(2)})^T\begin{pmatrix}1\\x_1^{(1)}\\x_2^{(1)}\end{pmatrix}=\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}\begin{pmatrix}1\\3\\3\end{pmatrix}=\begin{pmatrix}7\\7\\7\end{pmatrix}$$

则: $\begin{pmatrix}x_1^{(2)}\\x_2^{(2)}\\x_3^{(2)}\end{pmatrix}=\begin{pmatrix}7\\7\\7\end{pmatrix}$

则第三层的输入为:

$$\mathrm{s}_1^{(3)}=(\mathbf{w}^{(3)})^T\begin{pmatrix}1\\x_1^{(2)}\\x_2^{(2)}\\x_3^{(2)}\end{pmatrix}=\begin{pmatrix}1&1&1&1\end{pmatrix}\begin{pmatrix}1\\7\\7\\7\end{pmatrix}=22$$

即输出$\hat{y}=\mathrm{s}_1^{(3)}=22$

对于样本$\vec{x}_1$，其标签为1，采用平方误差函数：$e_n=(y_n-\hat{y}_n)^2$，则：

$$\delta_1^{(3)}=-2\left(y_n-\mathrm{s}_1^{(3)}\right)=-2(1-22)=42$$

运用反向传播法，于是：

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

即：$\vec{\delta}^{(2)}=\begin{pmatrix}\delta_1^{(2)}\\ \delta_2^{(2)}\\ \delta_3^{(2)}\end{pmatrix}=\begin{pmatrix}\delta_1^{(3)}w_{11}^{(3)}\left[\!\left[s_1^{(2)}\geq 0\right]\!\right]\\ \delta_1^{(3)}w_{21}^{(3)}\left[\!\left[s_2^{(2)}\geq 0\right]\!\right]\\ \delta_1^{(3)}w_{31}^{(3)}\left[\!\left[s_3^{(2)}\geq 0\right]\!\right]\end{pmatrix}=\begin{pmatrix}42*1*1\\ 42*1*1\\ 42*1*1\end{pmatrix}=\begin{pmatrix}42\\ 42\\ 42\end{pmatrix}$

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：

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

由此可以得到：

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

$$=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 1 & 1\\ 1 & 1 & 1\end{pmatrix}\begin{pmatrix}42\\ 42\\ 42\end{pmatrix}=\begin{pmatrix}126\\ 126\end{pmatrix}$$

令$\eta=0.01$，利用梯度下降法进行权系数更新：

$$\mathbf{w}_1^{(1)}=\mathbf{w}_0^{(1)}-\eta\vec{x}_n^{(0)}\overrightarrow{(\delta^{(1)})}^T=\mathbf{w}_t^{(1)}-\eta\begin{pmatrix}1\\ x_{n1}^{(0)}\\ x_{n2}^{(0)}\end{pmatrix}\left(\delta_1^{(1)},\delta_2^{(1)}\right)$$

$$=\begin{pmatrix}w_{01}^{(1)} & w_{02}^{(1)}\\ w_{11}^{(1)} & w_{12}^{(1)}\\ w_{21}^{(1)} & w_{22}^{(1)}\end{pmatrix}-\eta\begin{pmatrix}\delta_1^{(1)} & \delta_2^{(1)}\\ x_{n1}^{(0)}\delta_1^{(1)} & x_{n1}^{(0)}\delta_2^{(1)}\\ x_{n2}^{(0)}\delta_1^{(1)} & x_{n2}^{(0)}\delta_2^{(1)}\end{pmatrix}$$

$$=\begin{pmatrix}1 & 1\\ 1 & 1\\ 1 & 1\end{pmatrix}-0.01\begin{pmatrix}126 & 126\\ 1*126 & 1*126\\ 1*126 & 1*126\end{pmatrix}=\begin{pmatrix}-0.26 & -0.26\\ -0.26 & -0.26\\ -0.26 & -0.26\end{pmatrix}$$

$$\mathbf{w}_1^{(2)}=\mathbf{w}_0^{(2)}-\eta\vec{x}_n^{(1)}\overrightarrow{(\delta^{(2)}}) ^T=\mathbf{w}_0^{(2)}-\eta\begin{pmatrix}1\\x_1^{(1)}\\x_2^{(1)}\end{pmatrix}\left(\delta_1^{(2)},\delta_2^{(2)},\delta_3^{(2)}\right)$$

$$=\begin{pmatrix}w_{01}^{(2)}&w_{02}^{(2)}&w_{03}^{(2)}\\w_{11}^{(2)}&w_{12}^{(2)}&w_{13}^{(2)}\\w_{21}^{(2)}&w_{22}^{(2)}&w_{23}^{(2)}\end{pmatrix}-\eta\begin{pmatrix}\delta_1^{(2)}&\delta_2^{(2)}&\delta_3^{(2)}\\x_1^{(1)}\delta_1^{(2)}&x_1^{(1)}\delta_2^{(2)}&x_1^{(1)}\delta_3^{(2)}\\x_2^{(1)}\delta_1^{(2)}&x_2^{(1)}\delta_2^{(2)}&x_2^{(1)}\delta_3^{(2)}\end{pmatrix}$$

$$=\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}-0.01\begin{pmatrix}42&42&42\\3*42&3*42&3*42\\3*42&3*42&3*42\end{pmatrix}$$

$$=\begin{pmatrix}0.58&0.58&0.58\\-0.26&-0.26&-0.26\\-0.26&-0.26&-0.26\end{pmatrix}$$

$$\mathbf{w}_1^{(3)}=\mathbf{w}_0^{(3)}-\eta\vec{x}_n^{(2)}\overrightarrow{(\delta^{(3)}}) ^T=\mathbf{w}_0^{(3)}-\eta\begin{pmatrix}1\\x_1^{(2)}\\x_2^{(2)}\\x_3^{(2)}\end{pmatrix}\delta_1^{(3)}=\begin{pmatrix}1\\1\\1\\1\end{pmatrix}-0.01\begin{pmatrix}42\\7*42\\7*42\\7*42\end{pmatrix}$$

$$=\begin{pmatrix}0.58\\-1.94\\-1.94\\-1.94\end{pmatrix}$$

t=2，对于第二个样本$\vec{x}_2=(-1,-1)^T$，则第一层神经元的输入为：

$$\begin{pmatrix}s_1^{(1)}\\s_2^{(1)}\end{pmatrix}=(\mathbf{w}^{(1)})^T\vec{x}_n^{(0)}=\begin{pmatrix}-0.26&-0.26&-0.26\\-0.26&-0.26&-0.26\end{pmatrix}\begin{pmatrix}1\\-1\\-1\end{pmatrix}=\begin{pmatrix}0.26\\0.26\end{pmatrix}$$

则：$\begin{pmatrix}x_1^{(1)}\\x_2^{(1)}\end{pmatrix}=\begin{pmatrix}\max(0,s_1^{(1)})\\\max(0,s_2^{(1)})\end{pmatrix}=\begin{pmatrix}0.26\\0.26\end{pmatrix}$

第二层神经元的输入为：

$$\begin{pmatrix}s_1^{(2)}\\s_2^{(2)}\\s_3^{(2)}\end{pmatrix}=(\mathbf{w}^{(2)})^T\begin{pmatrix}1\\x_1^{(1)}\\x_2^{(1)}\end{pmatrix}=\begin{pmatrix}0.58&-0.26&-0.26\\0.58&-0.26&-0.26\\0.58&-0.26&-0.26\end{pmatrix}\begin{pmatrix}1\\0.26\\0.26\end{pmatrix}=\begin{pmatrix}0.44\\0.44\\0.44\end{pmatrix}$$

则：$\begin{pmatrix}x_1^{(2)}\\x_2^{(2)}\\x_3^{(2)}\end{pmatrix}=\begin{pmatrix}0.44\\0.44\\0.44\end{pmatrix}$

则第三层的输入为：

$$s_1^{(3)} = (\mathbf{w}^{(3)})^T \begin{pmatrix} 1 \\ x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{pmatrix} = (0.58 \quad -1.94 \quad -1.94 \quad -1.94) \begin{pmatrix} 1 \\ 0.44 \\ 0.44 \\ 0.44 \end{pmatrix} = -1.98$$

即输出$\hat{y} = s_1^{(3)} = -1.98$

对于样本$\vec{x}_2$，其标签为1，采用平方误差函数：$e_n = (y_n - \hat{y}_n)^2$，则：

$$\delta_1^{(3)} = -2\left(y_n - s_1^{(3)}\right) = -2\big(1 - (-1.98)\big) = -5.96$$

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：$\vec{\delta}^{(2)} = \begin{pmatrix} \delta_1^{(2)} \\ \delta_2^{(2)} \\ \delta_3^{(2)} \end{pmatrix} = \begin{pmatrix} \delta_1^{(3)} w_{11}^{(3)} [\![s_1^{(2)} \geq 0]\!] \\ \delta_1^{(3)} w_{21}^{(3)} [\![s_2^{(2)} \geq 0]\!] \\ \delta_1^{(3)} w_{31}^{(3)} [\![s_3^{(2)} \geq 0]\!] \end{pmatrix} = \begin{pmatrix} -5.96 * (-1.94) * 1 \\ -5.96 * (-1.94) * 1 \\ -5.96 * (-1.94) * 1 \end{pmatrix} =$

$\begin{pmatrix} 11.56 \\ 11.56 \\ 11.56 \end{pmatrix}$

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：

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

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -0.26 & -0.26 & -0.26 \\ -0.26 & -0.26 & -0.26 \end{pmatrix} \begin{pmatrix} 11.56 \\ 11.56 \\ 11.56 \end{pmatrix} = \begin{pmatrix} -9.02 \\ -9.02 \end{pmatrix}$$

令$\eta = 0.01$，利用梯度下降法进行权系数更新：

$$\mathbf{w}_2^{(1)} = \mathbf{w}_1^{(1)} - \eta \vec{x}_n^{(0)} (\overrightarrow{\delta}^{(1)})^T = \mathbf{w}_1^{(1)} - \eta \begin{pmatrix} 1 \\ x_{n1}^{(0)} \\ x_{n2}^{(0)} \end{pmatrix} \left(\delta_1^{(1)}, \delta_2^{(1)}\right)$$

$$= \begin{pmatrix} w_{01}^{(1)} & w_{02}^{(1)} \\ w_{11}^{(1)} & w_{12}^{(1)} \\ w_{21}^{(1)} & w_{22}^{(1)} \end{pmatrix} - \eta \begin{pmatrix} \delta_1^{(1)} & \delta_2^{(1)} \\ x_{n1}^{(0)}\delta_1^{(1)} & x_{n1}^{(0)}\delta_2^{(1)} \\ x_{n2}^{(0)}\delta_1^{(1)} & x_{n2}^{(0)}\delta_2^{(1)} \end{pmatrix}$$

$$= \begin{pmatrix} -0.26 & -0.26 \\ -0.26 & -0.26 \\ -0.26 & -0.26 \end{pmatrix} - 0.01 \begin{pmatrix} -9.02 & -9.02 \\ (-1)*(-9.02) & (-1)*(-9.02) \\ (-1)*(-9.02) & (-1)*(-9.02) \end{pmatrix}$$

$$= \begin{pmatrix} -0.17 & -0.17 \\ -0.35 & -0.35 \\ -0.35 & -0.35 \end{pmatrix}$$

$$\mathbf{w}_2^{(2)} = \mathbf{w}_1^{(2)} - \eta \vec{x}_n^{(1)} (\overrightarrow{\delta}^{(2)})^T = \mathbf{w}_1^{(2)} - \eta \begin{pmatrix} 1 \\ x_1^{(1)} \\ x_2^{(1)} \end{pmatrix} \left(\delta_1^{(2)}, \delta_2^{(2)}, \delta_3^{(2)}\right)$$

$$= \begin{pmatrix} w_{01}^{(2)} & w_{02}^{(2)} & w_{03}^{(2)} \\ w_{11}^{(2)} & w_{12}^{(2)} & w_{13}^{(2)} \\ w_{21}^{(2)} & w_{22}^{(2)} & w_{23}^{(2)} \end{pmatrix} - \eta \begin{pmatrix} \delta_1^{(2)} & \delta_2^{(2)} & \delta_3^{(2)} \\ x_1^{(1)}\delta_1^{(2)} & x_1^{(1)}\delta_2^{(2)} & x_1^{(1)}\delta_3^{(2)} \\ x_2^{(1)}\delta_1^{(2)} & x_2^{(1)}\delta_2^{(2)} & x_2^{(1)}\delta_3^{(2)} \end{pmatrix}$$

$$= \begin{pmatrix} 0.58 & 0.58 & 0.58 \\ -0.26 & -0.26 & -0.26 \\ -0.26 & -0.26 & -0.26 \end{pmatrix}$$

$$- 0.01 \begin{pmatrix} 11.56 & 11.56 & 11.56 \\ 0.26*11.56 & 0.26*11.56 & 0.26*11.56 \\ 0.26*11.56 & 0.26*11.56 & 0.26*11.56 \end{pmatrix}$$

$$= \begin{pmatrix} 0.46 & 0.46 & 0.46 \\ -0.29 & -0.29 & -0.29 \\ -0.29 & -0.29 & -0.29 \end{pmatrix}$$

$$\mathbf{w}_2^{(3)} = \mathbf{w}_1^{(3)} - \eta \vec{x}_n^{(2)} (\overrightarrow{\delta}^{(3)})^T = \mathbf{w}_1^{(3)} - \eta \begin{pmatrix} 1 \\ x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{pmatrix} \delta_1^{(3)}$$

$$= \begin{pmatrix} 0.58 \\ -1.94 \\ -1.94 \\ -1.94 \end{pmatrix} - 0.01 \begin{pmatrix} -5.96 \\ 0.44*(-5.96) \\ 0.44*(-5.96) \\ 0.44*(-5.96) \end{pmatrix} = \begin{pmatrix} 0.64 \\ -1.91 \\ -1.91 \\ -1.91 \end{pmatrix}$$

t=3，对于第三个样本$\vec{x}_3 = (-1,1)^T$，则第一层神经元的输入为：

$$\begin{pmatrix} s_1^{(1)} \\ s_2^{(1)} \end{pmatrix} = (\mathbf{w}^{(1)})^T \vec{x}_n^{(0)} = \begin{pmatrix} -0.17 & -0.35 & -0.35 \\ -0.17 & -0.35 & -0.35 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -0.17 \\ -0.17 \end{pmatrix}$$

则：$\begin{pmatrix} x_1^{(1)} \\ x_2^{(1)} \end{pmatrix} = \begin{pmatrix} \max(0, s_1^{(1)}) \\ \max(0, s_2^{(1)}) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

第二层神经元的输入为：

$$\begin{pmatrix} s_1^{(2)} \\ s_2^{(2)} \\ s_3^{(2)} \end{pmatrix} = (\mathbf{w}^{(2)})^T \begin{pmatrix} 1 \\ x_1^{(1)} \\ x_2^{(1)} \end{pmatrix} = \begin{pmatrix} 0.46 & -0.29 & -0.29 \\ 0.46 & -0.29 & -0.29 \\ 0.46 & -0.29 & -0.29 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.46 \\ 0.46 \\ 0.46 \end{pmatrix}$$

则：$\begin{pmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{pmatrix} = \begin{pmatrix} 0.46 \\ 0.46 \\ 0.46 \end{pmatrix}$

则第三层的输入为：

$$\mathrm{s}_1^{(3)} = (\mathbf{w}^{(3)})^T \begin{pmatrix} 1 \\ x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{pmatrix} = \begin{pmatrix} 0.64 & -1.91 & -1.91 & -1.91 \end{pmatrix} \begin{pmatrix} 1 \\ 0.46 \\ 0.46 \\ 0.46 \end{pmatrix} = -2.00$$

即输出$\hat{y} = \mathrm{s}_1^{(3)} = -2.00$

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：

$$\delta_1^{(3)} = -2\left(y_n - \mathrm{s}_1^{(3)}\right) = -2\left(-1 - (-2.00)\right) = -2.00$$

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：

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

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

令$\eta=0.01$，利用梯度下降法进行权系数更新：

$$\mathbf{w}_3^{(1)}=\mathbf{w}_2^{(1)}-\eta\vec{x}_n^{(0)}\overrightarrow{(\delta}^{(1)})^T=\mathbf{w}_2^{(1)}-\eta\begin{pmatrix}1\\ x_{n1}^{(0)}\\ x_{n2}^{(0)}\end{pmatrix}\left(\delta_1^{(1)},\delta_2^{(1)}\right)$$

$$=\begin{pmatrix}w_{01}^{(1)} & w_{02}^{(1)}\\ w_{11}^{(1)} & w_{12}^{(1)}\\ w_{21}^{(1)} & w_{22}^{(1)}\end{pmatrix}-\eta\begin{pmatrix}\delta_1^{(1)} & \delta_2^{(1)}\\ x_{n1}^{(0)}\delta_1^{(1)} & x_{n1}^{(0)}\delta_2^{(1)}\\ x_{n2}^{(0)}\delta_1^{(1)} & x_{n2}^{(0)}\delta_2^{(1)}\end{pmatrix}$$

$$=\begin{pmatrix}-0.17 & -0.17\\ -0.35 & -0.35\\ -0.35 & -0.35\end{pmatrix}-0.01\begin{pmatrix}0 & 0\\ (-1)*0 & (-1)*0\\ (+1)*0 & (+1)*0\end{pmatrix}$$

$$=\begin{pmatrix}-0.17 & -0.17\\ -0.35 & -0.35\\ -0.35 & -0.35\end{pmatrix}$$

$$\mathbf{w}_3^{(2)}=\mathbf{w}_2^{(2)}-\eta\vec{x}_n^{(1)}\overrightarrow{(\delta}^{(2)})^T=\mathbf{w}_2^{(2)}-\eta\begin{pmatrix}1\\ x_1^{(1)}\\ x_2^{(1)}\end{pmatrix}\left(\delta_1^{(2)},\delta_2^{(2)},\delta_3^{(2)}\right)$$

$$=\begin{pmatrix}w_{01}^{(2)} & w_{02}^{(2)} & w_{03}^{(2)}\\ w_{11}^{(2)} & w_{12}^{(2)} & w_{13}^{(2)}\\ w_{21}^{(2)} & w_{22}^{(2)} & w_{23}^{(2)}\end{pmatrix}-\eta\begin{pmatrix}\delta_1^{(2)} & \delta_2^{(2)} & \delta_3^{(2)}\\ x_1^{(1)}\delta_1^{(2)} & x_1^{(1)}\delta_2^{(2)} & x_1^{(1)}\delta_3^{(2)}\\ x_2^{(1)}\delta_1^{(2)} & x_2^{(1)}\delta_2^{(2)} & x_2^{(1)}\delta_3^{(2)}\end{pmatrix}$$

$$=\begin{pmatrix}0.46 & 0.46 & 0.46\\ -0.29 & -0.29 & -0.29\\ -0.29 & -0.29 & -0.29\end{pmatrix}-0.01\begin{pmatrix}3.82 & 3.82 & 3.82\\ 0*3.82 & 0*3.82 & 0*3.82\\ 0*3.82 & 0*3.82 & 0*3.82\end{pmatrix}$$

$$=\begin{pmatrix}0.42 & 0.42 & 0.42\\ -0.29 & -0.29 & -0.29\\ -0.29 & -0.29 & -0.29\end{pmatrix}$$

$$\mathbf{w}_3^{(3)} = \mathbf{w}_2^{(3)} - \eta \vec{x}_n^{(2)} \overrightarrow{(\delta^{(3)})}^T = \mathbf{w}_2^{(3)} - \eta \begin{pmatrix} 1 \\ x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{pmatrix} \delta_1^{(3)}$$

$$= \begin{pmatrix} 0.64 \\ -1.91 \\ -1.91 \\ -1.91 \end{pmatrix} - 0.01 \begin{pmatrix} -2.00 \\ 0.46 * (-2.00) \\ 0.46 * (-2.00) \\ 0.46 * (-2.00) \end{pmatrix} = \begin{pmatrix} 0.66 \\ -1.90 \\ -1.90 \\ -1.90 \end{pmatrix}$$

t=4，对于第四个样本$\vec{x}_2 = (1, -1)^T$，则第一层神经元的输入为：

$$\begin{pmatrix} s_1^{(1)} \\ s_2^{(1)} \end{pmatrix} = (\mathbf{w}^{(1)})^T \vec{x}_n^{(0)} = \begin{pmatrix} -0.17 & -0.35 & -0.35 \\ -0.17 & -0.35 & -0.35 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -0.17 \\ -0.17 \end{pmatrix}$$

则：$\begin{pmatrix} x_1^{(1)} \\ x_2^{(1)} \end{pmatrix} = \begin{pmatrix} \max(0, s_1^{(1)}) \\ \max(0, s_2^{(1)}) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

第二层神经元的输入为：

$$\begin{pmatrix} s_1^{(2)} \\ s_2^{(2)} \\ s_3^{(2)} \end{pmatrix} = (\mathbf{w}^{(2)})^T \begin{pmatrix} 1 \\ x_1^{(1)} \\ x_2^{(1)} \end{pmatrix} = \begin{pmatrix} 0.42 & -0.29 & -0.29 \\ 0.42 & -0.29 & -0.29 \\ 0.42 & -0.29 & -0.29 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.42 \\ 0.42 \\ 0.42 \end{pmatrix}$$

则：$\begin{pmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{pmatrix} = \begin{pmatrix} 0.42 \\ 0.42 \\ 0.42 \end{pmatrix}$

则第三层的输入为：

$$\mathrm{s}_1^{(3)} = (\mathbf{w}^{(3)})^T \begin{pmatrix} 1 \\ x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{pmatrix} = \begin{pmatrix} 0.66 & -1.90 & -1.90 & -1.90 \end{pmatrix} \begin{pmatrix} 1 \\ 0.42 \\ 0.42 \\ 0.42 \end{pmatrix} = -1.73$$

即输出$\hat{y} = \mathrm{s}_1^{(3)} = -1.73$

对于样本$\vec{x}_4$，其标签为-1，采用平方误差函数：$e_n = (y_n - \hat{y}_n)^2$，则：

$$\delta_1^{(3)} = -2\left(y_n - \mathrm{s}_1^{(3)}\right) = -2\left(-1 - (-1.73)\right) = -1.46$$

运用反向传播法，于是：

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

即：$\vec{\delta}^{(2)}=\begin{pmatrix}\delta_1^{(2)}\\ \delta_2^{(2)}\\ \delta_3^{(2)}\end{pmatrix}=\begin{pmatrix}\delta_1^{(3)}w_{11}^{(3)}\left[\!\left[s_1^{(2)}\geq 0\right]\!\right]\\ \delta_1^{(3)}w_{21}^{(3)}\left[\!\left[s_2^{(2)}\geq 0\right]\!\right]\\ \delta_1^{(3)}w_{31}^{(3)}\left[\!\left[s_3^{(2)}\geq 0\right]\!\right]\end{pmatrix}=\begin{pmatrix}(-1.46)*(-1.90)*1\\ (-1.46)*(-1.90)*1\\ (-1.46)*(-1.90)*1\end{pmatrix}=$

$\begin{pmatrix}2.77\\ 2.77\\ 2.77\end{pmatrix}$

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：

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

由此可以得到：

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

$$=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}-0.29 & -0.29 & -0.29\\ -0.29 & -0.29 & -0.29\end{pmatrix}\begin{pmatrix}2.77\\ 2.77\\ 2.77\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix}$$

令$\eta=0.01$，利用梯度下降法进行权系数更新：

$$\mathbf{w}_4^{(1)}=\mathbf{w}_3^{(1)}-\eta\vec{x}_n^{(0)}\overrightarrow{(\delta}^{(1)})^T=\mathbf{w}_3^{(1)}-\eta\begin{pmatrix}1\\ x_{n1}^{(0)}\\ x_{n2}^{(0)}\end{pmatrix}\left(\delta_1^{(1)},\delta_2^{(1)}\right)$$

$$=\begin{pmatrix}w_{01}^{(1)} & w_{02}^{(1)}\\ w_{11}^{(1)} & w_{12}^{(1)}\\ w_{21}^{(1)} & w_{22}^{(1)}\end{pmatrix}-\eta\begin{pmatrix}\delta_1^{(1)} & \delta_2^{(1)}\\ x_{n1}^{(0)}\delta_1^{(1)} & x_{n1}^{(0)}\delta_2^{(1)}\\ x_{n2}^{(0)}\delta_1^{(1)} & x_{n2}^{(0)}\delta_2^{(1)}\end{pmatrix}$$

$$=\begin{pmatrix}-0.17 & -0.17\\ -0.35 & -0.35\\ -0.35 & -0.35\end{pmatrix}-0.01\begin{pmatrix}0 & 0\\ (+1)*0 & (+1)*0\\ (-1)*0 & (-1)*0\end{pmatrix}$$

$$=\begin{pmatrix}-0.17 & -0.17\\ -0.35 & -0.35\\ -0.35 & -0.35\end{pmatrix}$$

$$\mathbf{w}_4^{(2)} = \mathbf{w}_3^{(2)} - \eta \vec{x}_n^{(1)} (\overrightarrow{\delta}^{(2)})^T = \mathbf{w}_3^{(2)} - \eta \begin{pmatrix} 1 \\ x_1^{(1)} \\ x_2^{(1)} \end{pmatrix} \left(\delta_1^{(2)}, \delta_2^{(2)}, \delta_3^{(2)} \right)$$

$$= \begin{pmatrix} w_{01}^{(2)} & w_{02}^{(2)} & w_{03}^{(2)} \\ w_{11}^{(2)} & w_{12}^{(2)} & w_{13}^{(2)} \\ w_{21}^{(2)} & w_{22}^{(2)} & w_{23}^{(2)} \end{pmatrix} - \eta \begin{pmatrix} \delta_1^{(2)} & \delta_2^{(2)} & \delta_3^{(2)} \\ x_1^{(1)}\delta_1^{(2)} & x_1^{(1)}\delta_2^{(2)} & x_1^{(1)}\delta_3^{(2)} \\ x_2^{(1)}\delta_1^{(2)} & x_2^{(1)}\delta_2^{(2)} & x_2^{(1)}\delta_3^{(2)} \end{pmatrix}$$

$$= \begin{pmatrix} 0.42 & 0.42 & 0.42 \\ -0.29 & -0.29 & -0.29 \\ -0.29 & -0.29 & -0.29 \end{pmatrix} - 0.01 \begin{pmatrix} 2.77 & 2.77 & 2.77 \\ 0 * 2.77 & 0 * 2.77 & 0 * 2.77 \\ 0 * 2.77 & 0 * 2.77 & 0 * 2.77 \end{pmatrix}$$

$$= \begin{pmatrix} 0.39 & 0.39 & 0.39 \\ -0.29 & -0.29 & -0.29 \\ -0.29 & -0.29 & -0.29 \end{pmatrix}$$

$$\mathbf{w}_4^{(3)} = \mathbf{w}_3^{(3)} - \eta \vec{x}_n^{(2)} (\overrightarrow{\delta}^{(3)})^T = \mathbf{w}_3^{(3)} - \eta \begin{pmatrix} 1 \\ x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{pmatrix} \delta_1^{(3)}$$

$$= \begin{pmatrix} 0.66 \\ -1.90 \\ -1.90 \\ -1.90 \end{pmatrix} - 0.01 \begin{pmatrix} -1.46 \\ 0.42 * (-1.46) \\ 0.42 * (-1.46) \\ 0.42 * (-1.46) \end{pmatrix} = \begin{pmatrix} 0.67 \\ -1.89 \\ -1.89 \\ -1.89 \end{pmatrix}$$