其概率，计算$E_{in} = \frac{1}{4}\sum_{n=1}^{4}(-ln\frac{1}{3}) = 1.099$

所以，我们按照式（10）求得梯度：

$$\nabla E_{in} = (\widehat{\boldsymbol{Y}} - \boldsymbol{Y})^T\boldsymbol{X} = \begin{pmatrix} \frac{1}{3}-1 & \frac{1}{3}-1 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3}-1 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3}-1 \end{pmatrix}\begin{pmatrix} 1 & 3 & 0 \\ 1 & 3 & 6 \\ 1 & 0 & 3 \\ 1 & -3 & 0 \end{pmatrix} = \begin{pmatrix} -\frac{2}{3} & -5 & -3 \\ \frac{1}{3} & 1 & 0 \\ \frac{1}{3} & 4 & 3 \end{pmatrix}$$

用梯度下降法式(12)进行权系数向量更新：

$$\boldsymbol{W}^{(1)} = \boldsymbol{W}^{(0)} - \eta\nabla E_{in} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} -\frac{2}{3} & -5 & -3 \\ \frac{1}{3} & 1 & 0 \\ \frac{1}{3} & 4 & 3 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & 5 & 3 \\ -\frac{1}{3} & -1 & 0 \\ -\frac{1}{3} & -4 & -3 \end{pmatrix}$$

根据式(13)得到S矩阵：

$$\boldsymbol{S} = \boldsymbol{X}(\boldsymbol{W}^{(1)})^{\boldsymbol{T}} = \begin{pmatrix} 1 & 3 & 0 \\ 1 & 3 & 6 \\ 1 & 0 & 3 \\ 1 & -3 & 0 \end{pmatrix}\begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ 5 & -1 & -4 \\ 3 & 0 & -3 \end{pmatrix} = \begin{pmatrix} 15.67 & -3.33 & -12.33 \\ 33.67 & -3.33 & -30.33 \\ 9.67 & -0.33 & -9.33 \\ -14.33 & 2.67 & 11.67 \end{pmatrix}$$

利用Softmax得到：

$$\widehat{\boldsymbol{Y}} = \begin{pmatrix} \vec{\hat{y}}_1 \\ \vec{\hat{y}}_2 \\ \vec{\hat{y}}_3 \\ \vec{\hat{y}}_4 \end{pmatrix} = \begin{pmatrix} 1.00 & 0.00 & 0.00 \\ 1.00 & 0.00 & 0.00 \\ 1.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 1.00 \end{pmatrix}$$

第三个样本错分，计算$E_{in} = (-ln1 - ln1 - ln0 - ln1)/4 = \infty$

第二次迭代：

我们按照式（10）求得梯度：

$$\nabla E_{in} = (\widehat{\boldsymbol{Y}} - \boldsymbol{Y})^T\boldsymbol{X} = \begin{pmatrix} 1-1 & 1-1 & 1 & 0 \\ 0 & 0 & 0-1 & 0 \\ 0 & 0 & 0 & 1-1 \end{pmatrix}\begin{pmatrix} 1 & 3 & 0 \\ 1 & 3 & 6 \\ 1 & 0 & 3 \\ 1 & -3 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 3 \\ -1 & 0 & -3 \\ 0 & 0 & 0 \end{pmatrix}$$

用梯度下降法式(12)进行权系数向量更新：