

$$\begin{aligned} \mathbf{W}^{(2)} &= \mathbf{W}^{(1)} - \eta \nabla E_{in} = \begin{pmatrix} \frac{2}{3} & 5 & 3 \\ -\frac{1}{3} & -1 & 0 \\ \frac{1}{3} & -4 & -3 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 3 \\ -1 & 0 & -3 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} -0.33 & 5 & 0 \\ 0.67 & -1 & 3 \\ -0.33 & -4 & -3 \end{pmatrix} \end{aligned}$$

根据式(13)得到S矩阵:

$$\mathbf{S} = \mathbf{X}(\mathbf{W}^{(2)})^T = \begin{pmatrix} 1 & 3 & 0 \\ 1 & 3 & 6 \\ 1 & 0 & 3 \\ 1 & -3 & 0 \end{pmatrix} \begin{pmatrix} -0.33 & 0.67 & -0.33 \\ 5 & -1 & -4 \\ 0 & 3 & -3 \end{pmatrix} = \begin{pmatrix} 14.67 & -2.33 & -12.33 \\ 14.67 & 15.67 & -30.33 \\ -0.33 & 9.67 & -9.33 \\ -15.33 & 3.67 & 11.27 \end{pmatrix}$$

利用Softmax得到:

$$\hat{\mathbf{Y}} = \begin{pmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \\ \hat{y}_4 \end{pmatrix} = \begin{pmatrix} 1.00 & 0.00 & 0.00 \\ 0.27 & 0.73 & 0.00 \\ 0.00 & 1.00 & 0.00 \\ 0.00 & 0.00 & 1.00 \end{pmatrix}$$

第二个样本错分, 计算 $E_{in} = (-\ln 1 - \ln 0.27 - \ln 1 - \ln 1)/4 = 0.33$

第三次迭代:

我们按照式 (10) 求得梯度:

$$\begin{aligned} \nabla E_{in} &= (\hat{\mathbf{Y}} - \mathbf{Y})^T \mathbf{X} = \begin{pmatrix} 1-1 & 0.27-1 & 0 & 0 \\ 0 & 0.73 & 1-1 & 0 \\ 0 & 0 & 0 & 1-1 \end{pmatrix} \begin{pmatrix} 1 & 3 & 0 \\ 1 & 3 & 6 \\ 1 & 0 & 3 \\ 1 & -3 & 0 \end{pmatrix} \\ &= \begin{pmatrix} -0.73 & -2.19 & -4.38 \\ 0.73 & 2.19 & 4.38 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

用梯度下降法式(12)进行权系数向量更新: