

$$\begin{aligned}\mathbf{W}^{(3)} &= \mathbf{W}^{(2)} - \eta \nabla E_{in} = \begin{pmatrix} -0.33 & 5 & 0 \\ 0.67 & -1 & 3 \\ -0.33 & -4 & -3 \end{pmatrix} - \begin{pmatrix} -0.73 & -2.19 & -4.38 \\ 0.73 & 2.19 & 4.38 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0.40 & 7.19 & 4.38 \\ -0.06 & -3.19 & -1.38 \\ -0.33 & -4 & -3 \end{pmatrix}\end{aligned}$$

根据式(13)得到S矩阵:

$$\begin{aligned}\mathbf{S} &= \mathbf{X}(\mathbf{W}^{(3)})^T = \\ &\begin{pmatrix} 1 & 3 & 0 \\ 1 & 3 & 6 \\ 1 & 0 & 3 \\ 1 & -3 & 0 \end{pmatrix} \begin{pmatrix} 0.40 & -0.06 & -0.33 \\ 7.19 & -3.19 & -4 \\ 4.38 & -1.38 & -3 \end{pmatrix} = \begin{pmatrix} 21.97 & -9.63 & -12.33 \\ 48.25 & -17.91 & -30.33 \\ 13.54 & -4.20 & -9.33 \\ -21.17 & 9.51 & 11.67 \end{pmatrix}\end{aligned}$$

利用Softmax得到:

$$\hat{\mathbf{Y}} = \begin{pmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \\ \hat{y}_4 \end{pmatrix} = \begin{pmatrix} 1.00 & 0.00 & 0.00 \\ 1.00 & 0.00 & 0.00 \\ 1.00 & 0.00 & 0.00 \\ 0.00 & 0.11 & 0.89 \end{pmatrix}$$

第三个样本错分, 计算 $E_{in} = (-\ln 1 - \ln 1 - \ln 0 - \ln 0.89)/4 = \infty$

第四次迭代:

我们按照式 (10) 求得梯度:

$$\begin{aligned}\nabla E_{in} &= (\hat{\mathbf{Y}} - \mathbf{Y})^T \mathbf{X} = \begin{pmatrix} 1-1 & 1-1 & 1 & 0 \\ 0 & 0 & 0-1 & 0.11 \\ 0 & 0 & 0 & 0.89-1 \end{pmatrix} \begin{pmatrix} 1 & 3 & 0 \\ 1 & 3 & 6 \\ 1 & 0 & 3 \\ 1 & -3 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 3 \\ -0.89 & -0.33 & -3 \\ -0.11 & 0.33 & 0 \end{pmatrix}\end{aligned}$$

用梯度下降法式(12)进行权系数向量更新:

$$\begin{aligned}\mathbf{W}^{(4)} &= \mathbf{W}^{(3)} - \eta \nabla E_{in} = \begin{pmatrix} 0.40 & 7.19 & 4.38 \\ -0.06 & -3.19 & -1.38 \\ -0.33 & -4 & -3 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 3 \\ -0.89 & -0.33 & -3 \\ -0.11 & 0.33 & 0 \end{pmatrix} \\ &= \begin{pmatrix} -0.60 & 7.19 & 1.38 \\ 0.83 & -2.86 & 1.62 \\ -0.22 & -4.33 & -3 \end{pmatrix}\end{aligned}$$

根据式(13)得到S矩阵: