南　开　大　学

网　络　空　间　安　全　学　院

机器学习第二次实验报告

BP 算法推导

许健

学号：2013018

专业：信息安全

指导教师：谢晋

2022 年 11 月 24 日

一、 BP 算法推导

(一) 符号约定

1. M: 层数，输入约定为第 0 层
2. N_i: 第 i 层神经元个数，N_0 为输入特征个数
3. $y_i^{(l)}$: 第 l 层第 i 个神经元输出值，$y_i^{(1)} = x_i$
4. $net_i^{(l)}$: 第 l 层第 i 个神经元净输入值，以 sigmoid 激活函数为例：

$$y_i^{(l)} = \frac{1}{1 + \exp\left[-net_i^{(l)}\right]}$$

5. $w_{ij}^{(l)}$: 连接第 l 层第 i 个神经元与第 1-1 层第 j 个神经元的权值
6. $\boldsymbol{y}_i^M$: 第 i 个样本的输出列向量，这是实验题目中给的定义，和前面的 $y_i^{(l)}$ 维度有差别
7. $\boldsymbol{y}_{i,j}^M$: 第 i 个样本的第 j 个元素 (预测值)
8. $\boldsymbol{d}_{i,j}$: 第 i 个样本的第 j 个元素 (真实值)
9. m_c^M: 第 c 类样本的均值向量
10. m^M: 所有样本的均值向量
11. n_c: 第 c 类样本的个数
12. 权重矩阵 $S_w = \sum_{c=1}^C \sum_{\boldsymbol{y}_i^M \in c} \left(\boldsymbol{y}_i^M - \boldsymbol{m}_c^M\right)\left(\boldsymbol{y}_i^M - \boldsymbol{m}_c^M\right)^T$
13. 偏置矩阵 $S_b = \sum_{c=1}^C n_c \left(\boldsymbol{m}_c^M - \boldsymbol{m}^M\right)\left(\boldsymbol{m}_c^M - \boldsymbol{m}^M\right)^T$
14. 多层感知机 MLP: $E = \sum_i \sum_j \frac{1}{2}\left(\boldsymbol{y}_{i,j}^M - \boldsymbol{d}_{i,j}\right)^2 + \frac{1}{2}\gamma\left(\operatorname{tr}(S_w) - \operatorname{tr}(S_b)\right)$

(二) 推导过程

对 $w_{ij}^{(l)}$ 求偏导，得到如下的公式

$$\Delta w_{ij}^{(l)} = -\eta \frac{\partial E}{\partial w_{ij}^{(l)}} = -\eta \frac{\partial E}{\partial net_i^{(l)}} \cdot \frac{\partial net_i^{(l)}}{\partial w_{ij}^{(l)}} \tag{1}$$

其中

$$net_i^{(l)} = \sum_{k=0}^{n} \left(w_{i,k}^{(l)} \cdot y_k^{(l-1))} + B_{i,k}^{(l)}\right) \tag{2}$$

$$\frac{\partial net_i^{(l)}}{\partial w_{ij}^{(l)}} = y_j^{(l-1)} \tag{3}$$

对 (1) 式中的第二项定义如下

$$\delta_i^{(l)} = -\frac{\partial E}{\partial net_i^{(l)}} \tag{4}$$

因此

$$\Delta w_{ij}^{(l)} = -\eta \frac{\partial E}{\partial net_i^{(l)}} \cdot \frac{\partial net_i^{(l)}}{\partial w_{ij}^{(l)}} = \eta \delta_i^{(l)} y_j^{(l-1)} \tag{5}$$

对 $b_{ij}^{(l)}$ 求偏导，得到如下的公式，R 代表全 1 的行向量

$$\begin{aligned}\Delta b_{i,j}^{(l)} &= -\eta \frac{\partial E}{\partial b_{i,j}^{(l)}} \\ &= -\eta \frac{\partial E}{\partial net_i^{(l)}} \cdot \frac{\partial net_i^{(l)}}{\partial b_{i,j}^{(l)}} \\ &= -\eta \frac{\partial E}{\partial net_i^{(l)}} \cdot R^T \\ &= \eta \delta_i^{(l)} \cdot R^T\end{aligned} \tag{6}$$

关键是求 $\delta_i^{(l)}$，以下是对$\delta_i^{(l)}$的推导

如果是输出层 l=M

$$\begin{aligned}\delta_i^{(M)} &= -\frac{\partial E}{\partial net_i^{(M)}} \\ &= -\frac{\partial E}{\partial y_i^{(M)}} \cdot \frac{\partial y_i^{(M)}}{\partial net_i^{(1)}} \\ &= -\frac{\partial E}{\partial y_i^{(M)}} \cdot y_i^{(M)} \cdot \left[1 - y_i^{(M)}\right]\end{aligned} \tag{7}$$

如果是隐藏层 l<M

$$\begin{aligned}\delta_i^{(l)} &= -\frac{\partial E}{\partial net_i^{(l)}} \\ &= -\sum_{k=0}^{N_l-1} \frac{\partial E}{\partial net_k^{(l+1)}} \frac{\partial net_k^{(l+1)}}{\partial net_i^{(l)}} \\ &= \sum_{k=0}^{N_l-1} \delta_k^{(l+1)} \frac{\partial net_k^{(l+1)}}{\partial net_i^{(l)}} \\ &= \sum_{k=0}^{N_l-1} \delta_k^{(l)} \frac{\partial net_k^{(l+1)}}{\partial y_i^{(l)}} \frac{\partial y_i^{(l)}}{\partial net_i^{(l)}} \\ &= \sum_{k=0}^{N_{l+1}-1} \delta_k^{(l+1)} \frac{\partial net_k^{(l+1)}}{\partial y_i^{(l)}} y_i^{(l)} \left[1 - y_i^{(l)}\right]\end{aligned} \tag{8}$$

其中

$$net_k^{(l+1)} = \sum_{j=0}^{N_l} w_{kj}^{(l+1)} y_j^{(l)} \tag{9}$$

$$\frac{\partial net_k^{(l+1)}}{\partial y_i^{(l)}} = w_{ki}^{(l+1)} \tag{10}$$

因此

$$\begin{aligned}\delta_i^{(l)} &= \sum_{k=0}^{N_{l+1}-1} \delta_k^{(l+1)} w_{ki}^{(l+1)} y_i^{(l)} \left[1 - y_i^{(l)}\right] \\ &= y_i^{(l)} \left[1 - y_i^{(l)}\right] \sum_{k=0}^{N_{l+1}-1} \delta_k^{(l+1)} w_{ki}^{(l+1)}\end{aligned} \tag{11}$$

只要计算出 (7) 式中的 $\frac{\partial E}{\partial y_i^{(M)}}$ 即可，下面给出一种求 $\frac{\partial E}{\partial y_i^{(M)}}$ 的方式。因为实验中给的 $\boldsymbol{y}_i^M$ 和我们之前的 $y_i^{(l)}$ 有些许不同，下面我们将使用 $\boldsymbol{y}_i^M$ 对正则化项求偏导。

定义本次实验的输出向量 y 维度为 C*N，有 N 个样本，每个样本有 C 类。根据标量对向量的求导法则 $\frac{\partial E}{\partial y_i^M} = [\frac{\partial E}{\partial y_{i,1}^M}, \frac{\partial E}{\partial y_{i,2}^M}, ..., \frac{\partial E}{\partial y_{i,C}^M}]^T$，对其中任一元素 $\frac{\partial E}{\partial y_{i,j}^M}$ 有:

$$\frac{\partial E}{\partial y_{i,j}^M} = \left(y_{i,j}^M - d_{i,j}\right) + \frac{1}{2}\gamma\left[\frac{\partial \operatorname{tr}(S_w)}{\partial y_{i,j}^M} - \frac{\partial \operatorname{tr}(S_b)}{\partial y_{i,j}^M}\right] \tag{12}$$

想要求出 $\frac{\partial E}{\partial y_{i,j}^M}$，我们需要知道 $\frac{\partial tr(S_w)}{\partial y_{i,j}^M}$ 和 $\frac{\partial tr(S_b)}{\partial y_{i,j}^M}$，因此下面计算正则化项的偏导

第 c 类样本的均值向量:

$$m_c^M = \frac{1}{n_c}\sum_{y_i^m \in c} y_i^M \tag{13}$$

所有样本的均值向量:

$$m^M = \frac{1}{n}\sum_{i=1}^{n} y_i^M \tag{14}$$

给出求 $\frac{\partial tr(S_w)}{\partial y_{i,j}^M}$ 的推导过程:

$$\begin{aligned} S_w &= \sum_{c=1}^{c}\sum_{y_i^m \in c}\left(y_i^m - m_c^M\right)\left(y_i^M - m_c^M\right)^{\top} \\ &= \sum_{c=1}^{c}\sum_{y_i^m \in c}\left(\frac{n_c-1}{n}y_i^m - \frac{a_i}{n_c}\right)\left(\frac{n_c-1}{n_c}y_i^m - \frac{a_i}{n_c}\right)^{\top} \end{aligned} \tag{15}$$

a_i 表示当前类别除 i 以外的所有样本向量之和:

$$a_i = \sum_{y_i^M \in c, i\neq j} y_i^M \tag{16}$$

S_w 矩阵的迹 $tr\left(S_w\right)$:

$$\operatorname{tr}\left(S_w\right) = \sum_{c=1}^{c}\sum_{y_i^m \in c}\sum_{j=1}^{c}\left(\frac{n_c-1}{n_c}y_{i,j}^M - \frac{a_{i,j}}{n_c}\right)^2 \tag{17}$$

$tr\left(S_w\right)$ 对 $y_{i,j}^M$ 的偏导:

$$\frac{\partial \operatorname{tr}\left(S_w\right)}{\partial y_{i,j}^M} = 2\frac{n_c-1}{n_c}\left(\frac{n_c-1}{n_c}y_{i,j}^M - \frac{a_{i,j}}{n_c}\right) \tag{18}$$

给出求 $\frac{\partial tr(S_b)}{\partial y_{i,j}^M}$ 的推导过程:

$$S_b = \sum_{c=1}^{C} n_c\left(m_c^M - m^M\right)\left(m_c^M - m^M\right)^{\top} \tag{19}$$

S_b 矩阵的迹 $tr\left(S_b\right)$:

$$\begin{aligned} tr\left(S_b\right) &= \sum_{c=1}^{C}\sum_{j=1}^{C} n_c\left(\frac{1}{n_c}\sum_{y_i^M \in c} y_{i,j}^M - \frac{1}{n}\sum_{i=1}^{n} y_{i,j}^M\right)^2 \\ &= \sum_{i=1}^{C}\sum_{j=1}^{C} n_c\left(m_{c,j}^M - m_j^M\right)^2 \end{aligned} \tag{20}$$

$tr\left(S_b\right)$ 对 $y_{i,j}^M$ 的偏导:

$$
\begin{aligned}
\frac{\partial tr\left(S_b\right)}{\partial y_{i,j}^M} &= \frac{\partial \sum_{c=1}^{C} n_c \left[\left(m_{c,i}^M - m_i^M\right)^2\right]}{\partial y_{i,j}^M} \\
&= \frac{\partial \left[n_p \sum_{j=1}^{C} \left(m_{i,j}^M - m_j^M\right)^2\right]}{\partial y_{i,j}^M} + \frac{\sum_{c=1,c\neq p}^{C} \sum_{j=1}^{C} \partial n_c \left[\left(m_{c,i}^M - m_i^M\right)^2\right]}{\partial y_{i,j}^M} \\
&= 2\frac{n-n_p}{n}\left(m_{i,j}^M - m_j^M\right) - 2\sum_{c=1,c\neq p}^{C}\sum_{j=1}^{C}\frac{n_c}{n}\left(m_{c,j}^M - m_j^M\right)
\end{aligned}
\tag{21}
$$

公式中的 p 表示当前 i 样本所在的类别，本应该使用 n_i，为了防止混淆改为 n_p