# 机器学习——BP 算法证明

学号：2011428
姓名：王天行
专业：密码科学与技术

2022 年 11 月 25 日

# 1 作业要求

在这个练习中，你需要以三层感知机为例，使用反向传播算法更新 MLP 的权重和偏置项，并将推导过程以报告的形式提交。MLP 以及权重、偏置项的定义如下：

Define $S_w$ and $S_b$ as:

$$S_w = \sum_{c=1}^{C} \sum_{\boldsymbol{y}_i^M \in c} (\boldsymbol{y}_i^M - \boldsymbol{m}_c^M)(\boldsymbol{y}_i^M - \boldsymbol{m}_c^M)^T$$
$$S_b = \sum_{c=1}^{C} n_c (\boldsymbol{m}_c^M - \boldsymbol{m}^M)(\boldsymbol{m}_c^M - \boldsymbol{m}^M)^T \tag{1}$$

where $\boldsymbol{m}_c^M$ is the mean vector of $\boldsymbol{y}_i^M$ (the output of the $i$th sample from the $c$th class), $\boldsymbol{m}^M$ is the mean vector of the output $\boldsymbol{y}_i^M$ from all classes, $n_c$ is the number of samples from the $c$th class. Define the discriminative regularization term $tr(S_w) - tr(S_b)$ and incorporate it into the objective function of the MLP:

$$E = \sum_i \sum_j \frac{1}{2} (\boldsymbol{y}_{i,j}^M - \boldsymbol{d}_{i,j})^2 + \frac{1}{2} \gamma (tr(S_w) - tr(S_b)). \tag{2}$$

where $\boldsymbol{y}_{i,j}^M$ is the $j$th element in the vector $\boldsymbol{y}_i^M$, $\boldsymbol{d}_{i,j}$ is the $j$th element in the label vector $\boldsymbol{d}_i$, $tr$ denotes the trace of the matrix. Use the BP algorithm to update parameters $\boldsymbol{W}$ and $\boldsymbol{b}$ of the MLP.

# 2 证明过程

将 MLP 的输入、神经元的输出（y 矩阵）和神经元的输入（n 矩阵）分别视为一个矩阵，且矩阵中的一个列向量即为一个输入样本。且 y 和 n 的一行均代表所有样本对应在该神经元的输出。定义 sum 为输入的每个样本的特征数，p 为样本数。

对于每一层每一个神经元，其输出 $y_{i,j}^{(l)}$ 为

$$y_{i,j}^{(l)} = \frac{1}{1 + exp[n_{i,j}^{(l)}]} \tag{1}$$

其中 i=1，2，……，n；j=1，2，……，p

$y_{i,j}$ 为第 j 个样本的输出向量的第 i 个元素。对于每个神经元的输入与上一个神经元输出的关系为

$$n_i^{(l)} = \sum_{k=0}^{sum} (w_{i,k}^{(l)} \cdot y_k^{(l-1)} + B_{i,k}^{(l)}) \tag{2}$$

其中 i=1，2，……，n

（2）式中，$B_{i,k}$ 为含有 p 个元素的行向量，且每个值均为 $b_{i,k}^{(l)}$ 。因此，权重

项 $w_{i,j}$ 为

$$\begin{aligned}\triangle w_{i,j}^{(l)} &= -\eta\frac{\partial E}{\partial w_{i,j}^{(l)}} \\ &= -\eta\frac{\partial E}{\partial n_i^{(l)}}\cdot\frac{\partial n_i^{(l)}}{\partial w_{i,j}^{(l)}} \\ &= -\eta\frac{\partial E}{\partial n_i^{(l)}}\cdot[y_i^{l-i}]^T\end{aligned} \tag{3}$$

其中 i=1，2，……，n-1；j=1，2，……，n-1

定义 $\delta_i^l = -\frac{\partial E}{\partial n_i^{(l)}}$ , 那么权重项 $b_{i,j}$ 为

$$\begin{aligned}\triangle b_{i,j}^{(l)} &= -\eta\frac{\partial E}{\partial b_{i,j}^{(l)}} \\ &= -\eta\frac{\partial E}{\partial n_{i,}^{(l)}}\cdot\frac{\partial E}{\partial b_{i,j}^{(l)}} \\ &= -\eta\frac{\partial E}{\partial n_{i,}^{(l)}}\cdot I^T \\ &= \eta\delta_i^{(l)}\cdot R^T\end{aligned} \tag{4}$$

其中 i=1，2，……，n-1；j=1，2，……，n-1；I 代表全 1 的行向量，大小为 1*n。

若为输出层

$$\begin{aligned}\delta_i^l &= -\frac{\partial E}{\partial n_i^{(m)}} \\ &= -\frac{\partial E}{\partial y_i^{(m)}}\cdot\frac{\partial y_i^{(m)}}{\partial n_i^{(m)}} \\ &= -\frac{\partial E}{\partial y_i^{(m)}}\cdot S^{(m)}\end{aligned} \tag{5}$$

其中 i=1，2，……，n-1；m 代表输出层；矩阵 $S^{(l)}$ 为 n*n 的矩阵，且只有主对角线上的元素不为 0，$S_{k,k}^{(l)} = y_{i,k}^{(l)}\cdot[1-y_{i,k}^{(l)}]$

若不为输出层

$$
\begin{aligned}
\delta_i^l &= -\frac{\partial E}{\partial n_i^{(l)}} \\
&= -\sum_{k=0}^{sum-1} \frac{\partial E}{\partial n_k^{(l+1)}} \cdot \frac{\partial n_k^{(l+1)}}{\partial n_i^{(l)}} \\
&= \sum_{k=0}^{sum-1} \delta_k^{(l+1)} \cdot \frac{\partial n_k^{(l+1)}}{\partial n_i^{(l)}} \\
&= \sum_{k=0}^{sum-1} \delta_k^{(l+1)} \cdot \frac{\partial n_k^{(l+1)}}{\partial y_i^{(l)}} \cdot \frac{\partial y_i^{(l)}}{\partial n_i^{(l)}} \\
&= \sum_{k=0}^{sum-1} \delta_k^{(l+1)} \cdot w_{ki} \cdot E \cdot S^{(l)}
\end{aligned} \tag{6}
$$

其中 i=1，2，……，n-1；E 为 n*n 的单位矩阵。因此只需要计算出 $\frac{\partial E}{\partial y_i^{(m)}}$ 即可

由于 $\frac{\partial E}{\partial y_i^{(m)}} = y_i^{(m)}$，有

$$\frac{\partial E}{\partial y_{i,j}^M} = (y_{i,j}^M - d_{j,i}) + \frac{1}{2}\gamma[\frac{\partial tr(S_w)}{\partial y_{i,j}^M} - \frac{\partial tr(S_b)}{\partial y_{i,j}^M}] \tag{7}$$

$$
\begin{aligned}
tr(S_w) &= \sum_{c=1}^{C}\sum_{j\in c}\sum_{i=0}^{sum-1}(y_{i,j}^M - m_{c,i}^M)^2 \\
&= \sum_{c=1}^{C}\sum_{j\in c}\sum_{i=0}^{sum-1}(\frac{n_c-1}{n_c}y_{i,j}^M - \frac{a}{n_c})^2
\end{aligned} \tag{8}
$$

a 代表除 j 对应的样本外属于当前类的所有样本第 i 个元素之和。于是有

$$
\begin{aligned}
\frac{\partial tr(S_w)}{\partial y_{i,j}^M} &= \frac{\partial \sum_{x=0}^{n_c-1}(\frac{n_c-1}{n_c}y_{x,j}^M - \frac{a}{n_c})^2}{\partial y_{i,j}^M} + \frac{\partial \sum_{x=0}^{n_c-1}\sum_{z\in c, z\neq j}(y_{x,z}^M - \frac{y_{x,j}^M}{n_c} - \frac{a}{n_c})^2}{\partial y_{i,j}^M} \\
&= 2\frac{n_c-1}{n_c}(y_{i,j}^M - m_{c,i}^M) - \frac{2}{n_c}\sum_{z\in c, z\neq j}(y_{i,z}^M - m_{c,i}^M)
\end{aligned} \tag{9}
$$

$$
\begin{aligned}
\frac{\partial tr(S_b)}{\partial y_{i,j}^M} &= \frac{\partial \sum_{c=1}^{C} n_c[(m_{c,i}^M - m_i^M)^2]}{\partial y_{i,j}^M} \\
&= \frac{\partial n_p[(m_{p,i}^M - m_i^M)^2]}{\partial y_{i,j}^M} + \sum_{c=1,c\neq p}^{C} \frac{\partial n_c[(m_{c,i}^M - m_i^M)^2]}{\partial y_{i,j}^M} \\
&= 2\frac{n-n_p}{n}(m_{p,i}^M - m_i^M) - \sum_{c=1,c\neq p}^{C} \frac{2n_c}{n}(m_{c,i}^M - m_i^M)
\end{aligned}
\tag{10}
$$