机器学习第二次作业

> 专业：计算机科学与技术
> 学号：2011188
> 姓名：邵琦

一、实验要求

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:

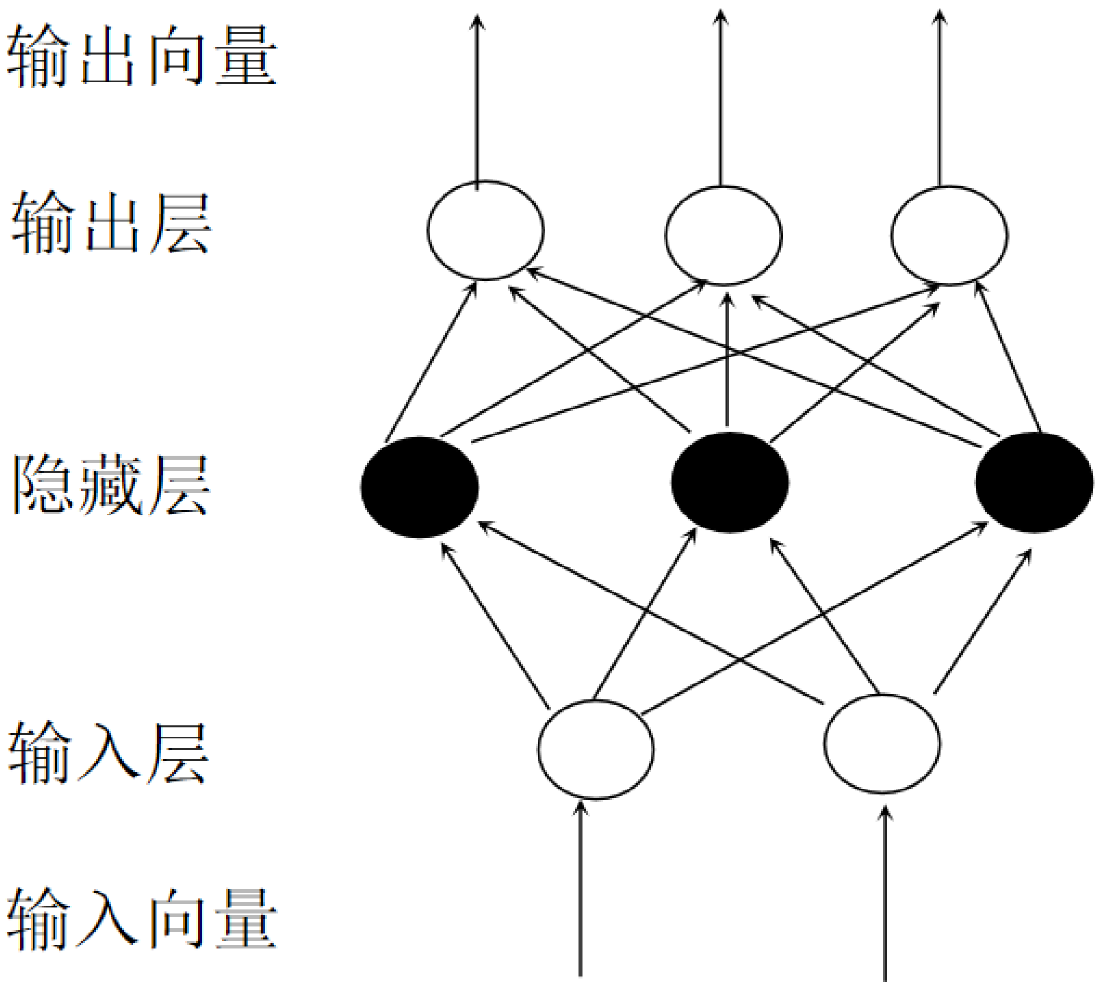
$$
\begin{aligned}
S_w &= \sum_{c=1}^{C} \sum_{\boldsymbol{y}_i^M \in c} (\boldsymbol{y}_i^M - \boldsymbol{m}_c^M)(\boldsymbol{y}_i^M - \boldsymbol{m}_c^M)^T \\
S_b &= \sum_{c=1}^{C} n_c (\boldsymbol{m}_c^M - \boldsymbol{m}^M)(\boldsymbol{m}_c^M - \boldsymbol{m}^M)^T
\end{aligned}
\tag{1}
$$

where $\boldsymbol{m}_c^M$ is the mean vector of $\boldsymbol{y}_i^M$ (the output of the ith sample from the cth 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 cth 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 jth element in the vector $\boldsymbol{y}_i^M$, $\boldsymbol{d}_{i,j}$ is the jth 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.

二、三层感知机



权重矩阵 $S_w = \sum_{c=1}^{C} \sum_{y_i^M \in c} (y_i^M - m_c^M)(y_i^M - m_c^M)^T$

偏置项矩阵 $S_b = \sum_{c=1}^{C} n_c (m_c^M - m^M)(m_c^M - m^M)^T$

误差评估函数 $E = \sum_i \sum_j \frac{1}{2}(y_{i,j}^M - d_{i,j})^2 + \frac{1}{2}\gamma(tr(S_w) - tr(S_b))$

三、反向传播算法更新MLP的权重和偏置项

输出层权重与偏置更新

权重更新

$$
\begin{aligned}
\Delta W_{ij}^2 &= \sum_{k=0}^{n-1}(-\eta\frac{\partial E}{\partial W_{ij}^2}) \\
&= \sum_{k=0}^{n-1}(-\eta\frac{\partial E}{\partial net_{ki}^2}\cdot\frac{\partial net_{ki}^2}{\partial W_{ij}^2}) \\
&= -\sum_{k=0}^{n-1}\eta\frac{\partial E}{\partial out_{ki}^2}\cdot\frac{\partial out_{ki}^2}{\partial net_{ki}^2}\cdot\frac{\partial net_{ki}^2}{\partial W_{ij}^2} \\
&= -\sum_{k=0}^{n-1}\frac{\partial E}{\partial out_{ki}^2}\cdot\eta\cdot out_{ki}^2(1-out_{ki}^2)\cdot out_{ki}^1 \\
&= -\sum_{k=0}^{n-1}\eta\epsilon_{ki}\cdot out_{ki}^2(1-out_{ki}^2)\cdot out_{ki}^1
\end{aligned}
$$

偏置更新

$$
\begin{aligned}
\Delta b_i^2 &= \sum_{k=0}^{n-1}(-\eta\frac{\partial E}{\partial b_i^2}) \\
&= \sum_{k=0}^{n-1}(-\eta\frac{\partial E}{\partial net_{ki}^2}\cdot\frac{\partial net_{ki}^2}{\partial b_i^2}) \\
&= -\sum_{k=0}^{n-1}\eta\frac{\partial E}{\partial out_{ki}^2}\cdot\frac{\partial out_{ki}^2}{\partial net_{ki}^2}\cdot\frac{\partial net_{ki}^2}{\partial b_i^2} \\
&= -\sum_{k=0}^{n-1}\frac{\partial E}{\partial out_{ki}^2}\cdot\eta\cdot out_{ki}^2(1-out_{ki}^2) \\
&= -\sum_{k=0}^{n-1}\eta\epsilon_{ki}\cdot out_{ki}^2(1-out_{ki}^2)
\end{aligned}
$$

$$
\begin{aligned}
\epsilon_{ki} &= \frac{\partial E}{\partial net_{ki}^2} \\
&= 2(out_{ki}^2-d_{ki})+\frac{1}{2}\gamma(\frac{\partial\operatorname{Tr}(S_w)}{\partial out_{ki}^2}-\frac{\partial(\operatorname{Tr}(S_b))}{\partial out_{ki}^2})
\end{aligned}
$$

而:

$$
\begin{aligned}
\frac{1}{2} \cdot \frac{\partial \operatorname{Tr}\left(S_w\right)}{\partial out_{ki}^2} &= \left(1-\frac{1}{n_c}\right)\left(out_{ki}^2-m_{ci}^M\right)+\sum_{p=0, p \neq k}^{n_c-1}-\frac{1}{n_c}\left(y_{pi}^2-m_{ci}^M\right) \\
&= out_{ki}^2-m_{ci}^M+\sum_{p=0}^{n_c-1}-\frac{1}{n_c}\left(y_{pi}^2-m_{ci}^M\right) \\
&= out_{ki}^2-m_{ci}^M+\left(-\frac{1}{n_c}\right)\left(n_c m_{ci}^M-n_c m_{ci}^M\right) \\
&= out_{ki}^2-m_{ci}^M
\end{aligned}
$$

$$
\begin{aligned}
\frac{1}{2} \cdot \frac{\partial\left(\operatorname{Tr}\left(S_b\right)\right)}{\partial out_{ki}^2} &= n_c\left(\frac{1}{n_c}-\frac{1}{n}\right)\left(m_{ci}^M-m_i^M\right)+\sum_{p=0, p \neq c}^{C-1}-\frac{n_p}{n}\left(m_{pi}^M-m_i^M\right) \\
&= \left(m_{ci}^M-m_i^M\right)+\sum_{p=0}^{c-1}-\frac{n_p}{n}\left(m_{pi}^M-m_i^M\right) \\
&= \left(m_{ci}^M-m_i^M\right)+\left(-\frac{1}{n}\right)\left(n m_i^M-n m_i^M\right) \\
&= m_{ci}^M-m_i^M
\end{aligned}
$$

则：

$$
\begin{aligned}
\varepsilon_{ki} = & 2\left(out_{ki}^2-d_{ki}\right)+\gamma\left(out_{ki}^2-m_{ci}^M\right)-\gamma\left(m_{ci}^M-m_i^M\right) \\
= & (2+\gamma) out_{ki}^2-2 d_{ki}-2 \gamma m_{ci}^M+\gamma m_i^M
\end{aligned}
$$

相关变量（隐藏层同）：

学习率： η

第n层由k个向量计算得到的输出向量的第i个值(未经过激活函数)： net_{ki}^n

第n层由k个向量计算得到的输出向量的第i个值(经过激活函数)： out_{ki}^n

输出结果中属于c类的向量个数：n_c

$$
m_{ci}^M=\frac{\sum_{k=0, y_k \in c}^{n_c-1} out_{ki}^2}{n_c}
$$

$$
\begin{aligned}
m_i^M &= \frac{\sum_{k=0}^{n-1} out_{ki}^2}{n} \\
&= \frac{\sum_{p=0}^{c-1} n_p m_{pi}^M}{n}
\end{aligned}
$$

隐藏层权重与偏置更新

权重更新

$$
\begin{aligned}
\Delta W_{ij}^1 &= \sum_{k=0}^{n-1}(-\eta\frac{\partial E}{\partial W_{ij}^1}) \\
&= \sum_{k=0}^{n-1}(-\eta\frac{\partial E}{\partial net_{ki}^1}\cdot\frac{\partial net_{ki}^1}{\partial W_{ij}^1}) \\
&= -\sum_{k=0}^{n-1}\eta\frac{\partial E}{\partial out_{ki}^1}\cdot\frac{\partial out_{ki}^1}{\partial net_{ki}^1}\cdot\frac{\partial net_{ki}^1}{\partial W_{ij}^1} \\
&= -\sum_{k=0}^{n-1}\eta\sum_{p=0}^{N_2-1}\delta_{kp}^2\frac{\partial net_{kp}^2}{\partial out_{ki}^1}out_{ki}^0 out_{ki}^1(1-out_{ki}^1) \\
&= -\eta\sum_{k=0}^{n-1}out_{ki}^0 out_{ki}^1(1-out_{ki}^1)\sum_{p=0}^{N_2-1}\delta_{kp}^2 W_{ij}^2
\end{aligned}
$$

偏置更新

$$
\begin{aligned}
\Delta b_i^1 &= \sum_{k=0}^{n-1}(-\eta\frac{\partial E}{\partial b_i^1}) \\
&= \sum_{k=0}^{n-1}(-\eta\frac{\partial E}{\partial net_{ki}^1}\cdot\frac{\partial net_{ki}^1}{\partial b_i^1}) \\
&= -\sum_{k=0}^{n-1}\eta\frac{\partial E}{\partial out_{ki}^1}\cdot\frac{\partial out_{ki}^1}{\partial net_{ki}^1}\cdot\frac{\partial net_{ki}^1}{\partial b_i^1} \\
&= -\sum_{k=0}^{n-1}\eta\sum_{p=0}^{N_2-1}\delta_{kp}^2\frac{\partial net_{kp}^2}{\partial out_{ki}^1}out_{ki}^1(1-out_{ki}^1) \\
&= -\eta\sum_{k=0}^{n-1}out_{ki}^1(1-out_{ki}^1)\sum_{p=0}^{N_2-1}\delta_{kp}^2 W_{ij}^2
\end{aligned}
$$

其中:

$$
\delta_{ki}^2 = \varepsilon_{ki} out_{ki}^2(1-out_{ki}^2)
$$

更新方式

依据上面计算所得，我们可以如下更新权重和偏置：

$$W_{ij}^{n'} = W_{ij}^{n} + \Delta W_{ij}^{n}$$

$$b_i^{n'} = b_i^{n} + \Delta b_i^{n}$$