

# Exercise2-MLP\_BP

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:

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

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

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

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

作业提交要求：

1. 推导过程需用公式编辑器进行编辑，可以使用word自带的公式编辑器，也可以使用overleaf中的latex模板。
2. 作业以pdf格式提交，命名方式为：**第二次作业-学号-姓名**”（若未按该命名方式命名，该项作业小分扣2分），发送至邮箱nk\_ml\_2022@163.com。
3. 作业截止日期：2022年11月25日24:00。