

# 模式识别导论第三次作业——简述题

杨登天-202028015926089-微电子研究所

1. 请对反向传播算法的训练步骤进行总结；结合三层网络给出不超过三个有关权重更新的公式，并用文字描述所述公式的含义；指出哪些因素会对网络的性能产生影响。

解答：

**任务一：总结反向传播算法的训练步骤；**

**步骤一、定义相关参数**——包括由样本变量维度定义输入层神经元个数(包含偏置的神经元)、隐含层层数、各隐含层神经元个数、输出层神经元个数、各隐含层神经元和输出神经元的激励函数与权重网络；

**步骤二、前向计算**——从输入层输入样本参数开始，经过各层输入的权重累加、激励函数推进得到输出层的输出，建立基于误差准则的目标函数并计算；

**步骤三、误差反向回传**——输出层的输出和监督信号之间的差异回传到各层之间的权重网络，依据全样本、随机单样本或者顺序单样本的方式修改权重网络。

**步骤四、计算目标函数的结果**，如果误差实现训练停止准则的要求，则退出循环；否则不断重复步骤二和步骤三直至误差满足要求。

**任务二：结合三层网络给出不超过三个有关权重更新的公式，并描述公式含义；**

**步骤一、定义相关参数**——训练数据输入输出对 $\{x_i^k, t_j^k\}$ ， $x_i^k$ 表示输入信号，其中 $i \in (1, 2 \dots d)$ ， $d$ 指的输入信号的维度，其中 $k$ 表示第 $k$ 个样本， $k \in (1, 2 \dots K)$ 。偏置为 $x_0$ ，为了方便，将偏置设置为 $x_0^k$ ； $t_j^k$ 表示输出信号的监督信号，其中 $j \in (1, 2 \dots J)$ 。输出层的结点输出 $z_j^k$ ，隐含层结点的输出 $y_h^k$ ，输入端点数目为 $d + 1$ ，输入层结点 $i$ 至隐含层结点 $h$ 的权重为 $w_{ih}$ ，隐含层结点 $h$ 到输出层结点 $j$ 的权重为 $w_{hj}$ 。其中 $h$ 为隐含层结点个数，并且

$h \in (1, 2 \dots H)$ 。目标函数采用平方误差准则函数 $J(w) = \frac{1}{2} \sum_{j=1}^J (t_j^k - z_j^k)^2$ 。对于所有的样本而言，均希望目标函数趋近于0，即 $J(w) \rightarrow 0$ 。 $\eta$ 为更新步长。 $y_h^k = f(\text{net}_h^k)$ ，其中 $\text{net}_h^k = \sum_{i=0}^d x_i^k w_{ih}$ 。 $z_j^k = g(\text{net}_j^k)$ ，其中 $\text{net}_j^k = \sum_{h=1}^H y_h^k w_{hj}$ 。

**步骤二、表达权重更新公式并描述公式含义**

2.1 从隐含层到输出层的权重更新如下：

此处仅以 $K$ 个样本批量更新为例，

$$\begin{aligned} w_{hj}(t+1) &= w_{hj}(t) + \Delta w_{hj} = w_{hj}(t) - \eta \sum_k \frac{\partial J(w)}{\partial \text{net}_j^k} \frac{\partial \text{net}_j^k}{\partial w_{hj}} = w_{hj}(t) + \eta \sum_k \Delta_j^k g'(\text{net}_j^k) y_h^k \\ &= w_{hj}(t) + \eta \sum_k \delta_j^k y_h^k \end{aligned}$$

整理得到

$$w_{hj}(t+1) = w_{hj}(t) + \eta \sum_k \delta_j^k y_h^k \quad (1)$$

公式含义：上述权重更新公式实现隐含层到输出层的权重网络更新，更新形式为批量更新，其中更新量包括： $y_h^k$ 权重所联边的起始结点的输出， $\delta_j^k$ 权重所联边的指向结点收集到的误差信号(该误差部分经由导数缩放)与 $\eta$ 步长。

2.2 从输入层到隐含层的权重更新如下：

此处仅以 $K$ 个样本批量更新为例，

$$w_{ih}(t+1) = w_{ih}(t) + \Delta w_{ih} = w_{ih}(t) - \eta \sum_{k,j} \frac{\partial J(w)}{\partial z_j^k} \frac{\partial z_j^k}{\partial net_j^k} \frac{\partial net_j^k}{\partial y_h^k} \frac{\partial y_h^k}{\partial net_h^k} \frac{\partial net_h^k}{\partial w_{ih}}$$

$$= w_{ih}(t) + \eta \sum_{k,j} \delta_j^k w_{hj} x_i^k f'(net_h^k) = w_{ih}(t) + \eta \sum_k \delta_h^k x_i^k$$

整理得到

$$w_{ih}(t+1) = w_{ih}(t) + \eta \sum_k \delta_h^k x_i^k \quad (2)$$

公式含义：上述权重更新公式实现输入层到隐含层的权重网络更新，更新形式为批量更新，其中更新量包括： $x_i^k$ 权重所联边的起始结点的输出， $\delta_h^k$ 权重所联边的指向结点收集到的误差信号(该误差部分经由导数缩放和后层权重网络收集)与 $\eta$ 步长。

从上述式子可以得到

$$w_{ih}(t+1) = w_{ih}(t) + \eta \sum_{k,j} \delta_j^k w_{hj} x_i^k f'(net_h^k) \quad (3)$$

公式含义：上述权重更新公式实现输入层到隐含层的权重网络更新，该表达式是考虑从输出层经过隐含层到输入层的完整表达式，更新形式为批量更新，其中更新量包括： $x_i^k$ 权重所联边的起始结点的输出， $\delta_j^k$ 权重所联边的指向结点收集到的误差信号(该误差部分经由导数缩放)， $f'(net_h^k)$ 表示隐含层激励函数的导函数（放缩因子）， $w_{hj}$ 表示隐含层结点 $h$ 到输出层结点 $j$ 的权重与 $\eta$ 步长。

### 任务三：哪些因素会对网络性能产生影响；

因素 1、激励函数——激励函数建议选取单调函数，否则误差函数会包含额外的局部极小值点，从而增加训练难度。

因素 2、隐含层数——对于分类问题，隐含层的个数决定了网络的表达能力，决定决策面的复杂程度；同时如果各节点均使用 Sigmoid 函数，一个隐含层足以实现任意判别分类问题，两个隐含层则足以实现输入向量的任意输出函数。

因素 3、结点个数——隐含层结点数太少，网络难以建立复杂的判别界面；隐含层结点太多，判决界面仅包含训练样本点而失去推广能力。

因素 4、初始权重——在批处理权重更新算法中，初始权重的更新值 $\Delta w_{hj} \neq 0$ ，网络失去学习能力。此外，通常从一个均匀分布中随机选择初始值： $-w_0 < w < w_0$ 。如果 $w_0$ 太小，隐含层的网络加权和就会很小，网络则类似于线性网络。

因素 5、正则化技术——可以防止网络出现过拟合的一种有效方法。

因素 6、学习率——学习率 $\eta$ 太小，则收敛较慢；过大则不稳定。

因素 7、附加冲量项——可以使得迭代序列更平滑，实现逃离饱和区的可能性。

因素 8、训练停止准则——如果过度训练网络（即使训练样本都能正确的分类）有可能会产生一个过拟合问题，也就是对新样本的分类能力不一定高。

解答完毕

2. 请描述自组织映射网络的构造原理，给出自组织算法的计算步骤(即网络训练)。

解答：

任务一：描述自组织映射网络的构造原理；

构造原理从生物学角度看：自组织映射网络是对大脑皮层各功能分区的近似实现，同时也模拟了功能分区内同层神经元之间的相互作用。

构造原理从结构上看：输入层的每一个单元与输出层的每个单元相联，但输出层相邻神

经元之间有相互作用。

**构造原理从功能上看：**输入向量连接到某个结点的权重组成该结点的权重向量；一个结点对输入向量的响应程度可以用欧氏距离或者内积表示；对于一个输入样本，输出层的所有结点中对其响应最大的结点是获胜神经元。

**构造原理从作用上看：**学习过程中，输出层获胜神经元要调整自身权重向量，它周围的神经元在其影响下也要不同程度地调整权重。邻近结点相互作用随着时间增长而减弱，减弱体现在作用区域的缩小和作用权重改变程度的下降，最终结果是获胜神经元仅作用其本身。

自组织映射网络的构造原理从生物学角度出发，基于上述结构、功能和作用共同实现。

**任务二：给出自组织算法的计算步骤(网络训练)；**

此处以单样本更新权重向量为例

**步骤一、网络初始化**

训练数据输入输出对 $\{x_i^k, z_j^k\}$ ， $x_i^k$ 表示输入信号，其中 $i \in (1, 2 \dots d)$ ， $d$ 指的输入信号的维度，其中 $k$ 表示第 $k$ 个样本， $k \in (1, 2 \dots K)$ ， $z_j^k$ 表示输出信号，其中 $j \in (1, 2 \dots J)$ 。初始化权重网络，对于 $x_i^k$ 和 $z_j^k$ 之间的权重用 $w_{ij}$ 表示。

**步骤二、输入向量**

输入训练样本 $x^k = \{x_1^k, x_2^k \dots x_d^k\}$ ，并已准备好 $K$ 个样本。

**步骤三、计算映射层的权重向量和输入向量的欧式距离**

计算权重向量和输入向量之间的欧氏距离 $z_j^k$

$$z_j^k = \sqrt{\sum_{i=1}^d (x_i^k - w_{ij})^2}$$

**步骤四、选择与权重向量的距离最小的神经元**

通过步骤三得到 $J$ 个输出神经元的欧氏距离 $z^k$

$$z^k = \{z_1^k, z_2^k \dots z_J^k\}$$

根据胜者为王策略确定获胜神经元 $z_{success}^k$

$$z_{success}^k = \min z^k = \min\{z_1^k, z_2^k \dots z_J^k\}$$

给出该获胜神经元的邻接神经元集合

$$h = \{\dots success - 1, success + 1, success + 2 \dots\}$$

**步骤五、调整权重**

调整获胜神经元和临近神经元的权重，按照下式更新：

$$\Delta w_{ij} = \eta h(success - i, success) \times (x_i^k - w_{ij})$$

$$w_{ij}(t + 1) = w_{ij}(t) + \Delta w_{ij}$$

**步骤六、检查是否达到预先设定的训练停止准则**

如果满足训练停止准则，则退出程序；否则返回步骤二，进入下一轮学习。

解答完毕