# DNN的公式推导

## 符号说明

|  |  |
| --- | --- |
|  | 第层神经元的个数 |
|  | 神经元的激活函数 |
|  | 第层到第层的权重矩阵 |
|  | 为中的元素，表示第层第个神经元到第层第个神经元的连接权重 |
|  | 等于，表示第层到第层的偏置 |
|  | 等于，表示第层神经元的状态值（输入值） |
|  | 等于，表示第层神经元的激活值（输出值） |
| **\*** | 矩阵乘积 |
|  | 哈达玛乘积 |

**信息向前传播**

即当前层的状态值和激活值可由前一层的状态值和激活值求出：

**误差后向传播**

假设训练数据为{(),……. (),……()}共N个，其中x代表输入值，y代表理想输出值，输出数据为维的，即 **=** ,则对于某一个训练数据()来说，其代价函数可以写为：

其中，表示第i个训练集对应的第L层（输出层）网络，第k个神经元的激活值。则所有训练集的总体平均代价可写为：

目标：调节和，使得最小。

方法：梯度下降法。

即：

其中代表学习因子，控制每次更新参数的步长。接下来只需确定和，就能完成我们的目标。接下来，我们考虑单次的代价函数，即和**。**为便于表达，将其简化为和

## 一、输出层的权重更新

假设神经网络共L层，则第L层就为输出层，对进行展开:

带入得:

引入符号，记:

则:

同理可推出的增量公式，最终，输出层主要公式的向量形式为：

## 二、隐藏层的权重更新

设为第层神经元的编号，为第层神经元的编号,为层神经元的编号，

与输出层同理可推得:

其中

同理可推出的增量公式，最终，隐藏层主要公式的向量形式为：

## 三、总结

综合输出层和隐藏层的主要公式，可得BP神经网络算法的核心公式为：

向量形式为：

## 四、编程注意事项

当(第一层的状态值)为 x （m代表单次训练集中元素的个数，n代表训练集的个数）的矩阵时，和的权重更新可以直接写为：

为便于编程，同时各向量可改写为：

其中为 x 的单位矩阵，**ONE**为 x 元素都为1的矩阵。