

第六章：神经网络

第一部分：计算与证明

1. Consider a three-layer network for classification with n_H nodes in hidden layer, and c nodes in output layer. The patterns (also say samples) are in d dimensional space. The activation function (or transfer function) for the nodes in the hidden layer is the sigmoid function. Differently, the nodes in the output layer will employ the following softmax operation as their activation function:

$$z_j = \frac{e^{net_j}}{\sum_{m=1}^c e^{net_m}}, \quad j=1,2,\dots,c,$$

where net_j stands for the weighted sum at the j -th node in the output layer.

- (a) Derive the learning rule under the back propagation framework if the criterion function for each sample is the sum of the squared errors, that is （即分析每一层权重的更新方法）：

$$J(\mathbf{w}) = \frac{1}{2} \sum_{j=1}^c (t_j - z_j)^2,$$

Where t_j is the known target value for the sample at the j -th node in the output layer.

- (b) 结合课堂所学知识，对反向传播算法进行总结。

注意：本题只需要推导出单个样本对权重更新的贡献即可（因为多个样本只是简单地相加）

2. 请描述自组织算法的计算步骤，给出训练算法的框图。
3. 拟考虑对 400×400 大小的图像数据集运用卷积神经网络。假定共有 4 个卷积层。第一隐含层采用 5×5 大小的滤波器，其它卷积层采用 3×3 大小的滤波器。第 1 至第 4 卷积层分别包含的图像数目为 20, 30, 20, 10。Pooling 操作均采用在 2×2 大小的局部窗口内取最大值(即 max pooling)，同时假定在 max pooling 之后完成激励操作。假定连接到最后一个 pooling 操作之后的前向神经网络为单层前向神经网络，且输出层的结点个数为 10。
(1) 请按层指出该网络需要计算的权重数量；相对于全连接和非权值共享，请指出当同时采用权值共享和局部连接时所减少的权重数量；(2) 请指出在遇到 max pooling 操作时，错误如何传？(3) 请讨论你所能想到的对网络结构的改变。

第二部分：计算机编程

本题使用的数据如下：

第一类 10 个样本（三维空间）：

[1.58, 2.32, -5.8], [0.67, 1.58, -4.78], [1.04, 1.01, -3.63],
[-1.49, 2.18, -3.39], [-0.41, 1.21, -4.73], [1.39, 3.16, 2.87],
[1.20, 1.40, -1.89], [-0.92, 1.44, -3.22], [0.45, 1.33, -4.38],
[-0.76, 0.84, -1.96]

第二类 10 个样本（三维空间）：

[0.21, 0.03, -2.21], [0.37, 0.28, -1.8], [0.18, 1.22, 0.16],
[-0.24, 0.93, -1.01], [-1.18, 0.39, -0.39], [0.74, 0.96, -1.16],
[-0.38, 1.94, -0.48], [0.02, 0.72, -0.17], [0.44, 1.31, -0.14],
[0.46, 1.49, 0.68]

第三类 10 个样本（三维空间）：

[-1.54, 1.17, 0.64], [5.41, 3.45, -1.33], [1.55, 0.99, 2.69],
[1.86, 3.19, 1.51], [1.68, 1.79, -0.87], [3.51, -0.22, -1.39],
[1.40, -0.44, -0.92], [0.44, 0.83, 1.97], [0.25, 0.68, -0.99],
[0.66, -0.45, 0.08]

1. 请编写两个通用的三层前向神经网络反向传播算法程序，一个采用批量方式更新权重，另一个采用单样本方式更新权重。其中，隐含层结点的激励函数采用双曲正切函数，输出层的激励函数采用 sigmoid 函数。目标函数采用平方误差准则函数。
2. 请利用上面的数据验证你写的程序，分析如下几点：
 - (a) 隐含层不同结点数目对训练精度的影响；
 - (b) 观察不同的梯度更新步长对训练的影响，并给出一些描述或解释；
 - (c) 在网络结构固定的情况下，绘制出目标函数随着迭代步数增加的变化曲线。