

深度学习 训练营

一步一个脚印，掌握深度学习



第7周，误差反向传播算法

第7周，课程核心内容：

7.1-计算图和求导的链式法则

7.2-反向传播算法和计算图

7.3-基于链式法则的梯度计算

7.4-误差反向传播的通用公式

7.5-矩阵表示反向传播算法

7.6-实验，实现通用的神经网络

7.7-实验，实现反向传播算法

7.8-实验，小批量梯度下降算法

7.9-实验，手写数字识别实验(上)

7.10-实验，手写数字识别实验(下)

$$\delta_j^{(k)} = \frac{\partial E}{\partial z_j^{(k)}} = \sum_{i=1}^n \frac{\partial E}{\partial z_i^{(k+1)}} \frac{\partial z_i^{(k+1)}}{\partial a_j^{(k)}} \frac{\partial a_j^{(k)}}{\partial z_j^{(k)}}$$

$$= \sum_{i=1}^n \delta_i^{(k+1)} w_{ij}^{(k+1)} (\text{sigmoid}'(z_j^{(k)}))$$

预习指导问题

同学们在视频学习前，请尝试回答如下问题，并在视频学习后，将答案进行归纳与整理：

- 1.什么是计算图?如何使用计算图表示线性回归模型的计算过程?
- 2.给定复合函数 $f(g(x))$ ，如何使用链式法则，求该函数的导数?
- 3.反向传播算法的基本原理是什么?它是如何利用链式法则进行求导的?
- 4.误差反向传播的通用公式是什么?它是如何推导出来的?
- 5.误差变量 δ 是什么?在误差反向传播中，为什么要引入这个误差变量 δ ?
- 6.神经网络的误差 E 关于参数 w 和 b 的偏导数与误差变量 δ ，有什么样的关系?
- 7.什么是矩阵的哈达玛积运算，它与矩阵乘法有什么不同?
- 8.如何将误差反向传播的通用公式，表示为矩阵形式?
- 9.小批量梯度下降算法与标准的梯度下降算法有何不同?为什么要使用小批量梯度下降算法?
- 10.在手写数字识别实验中，神经网络的输入层有多少个神经元，与什么有关?

课后编程练习

同学们在视频学习后，请尝试如下编程练习，并在直播课中，跟着老师一起完成全部编程作业。

- 1.请使用python，实现一个任意结构的神经网络，其中神经网络的层数和每层神经元的个数，可以通过用户的输入来指定。
- 2.请使用python，在不使用pytorch的前提下，实现神经网络的反向传播算法，并使用该算法训练一个3层网络拟合正弦函数。
- 3.请使用python，实现一个小批量梯度下降算法，迭代逻辑回归模型。
- 4.请设计并实现一个神经网络，用于解决MNIST数据集的手写数字识别问题。
- 5.请基于pytorch，训练一个神经网络，用于解决手写数字识别问题，并在测试集，测试该模型的效果。

最后将所有问题，都直接与小黑黑老师讨论清楚吧！

微信号:xhh890921

