

深度学习入门

基于Python的理论与实现

Stephen CUI 

March 14, 2023

Chapter 1

误差反向传播法

通过数值微分计算了神经网络的权重参数的梯度（严格来说，是损失函数关于权重参数的梯度）。数值微分虽然简单，也容易实现，但缺点是计算上比较费时间。本章我们将学习一个能够高效计算权重参数的梯度的方法——误差反向传播法。

1.1 计算图

计算图将计算过程用图形表示出来。这里说的图形是数据结构图，通过多个节点和边表示（连接节点的直线称为“边”）。

计算图通过节点和箭头表示计算过程。节点用 \bigcirc 表示， \bigcirc 中是计算的内容。将计算的中间结果写在箭头的上方，表示各个节点的计算结果从左向右传递。

虽然Figure 1.1中把“ $\times 2$ ”“ $\times 1.1$ ”等作为一个运算整体用 \bigcirc 括起来了，不过只用 \bigcirc 表示乘法运算“ \times ”也是可行的。此时，如Figure 1.2所示，可以将“2”和“1.1”分别作为变量“苹果的个数”和“消费税”标在 \bigcirc 外面。

用计算图解题的情况下，需要按如下流程进行。

1. 构建计算图。
2. 在计算图上，从左向右进行计算。

这里的第2步“从左向右进行计算”是一种正方向上的传播，简称为**正向传播**（forward propagation）。正向传播是从计算图出发点 to 结束点的传播。既然有正向传播这个名称，当然也可以考虑反向（从图上看的话，就是从右向左）的传播。实际上，这种传播称为**反向传播**（backward propagation）。反向传播反向传播将在接下来的导数计算中发挥重要作用。

1.1.1 局部计算

计算图的特征是可以通过传递“局部计算”获得最终结果。“局部”这个词的意思是“与自己相关的某个小范围”。局部计算是指，无论全局发生了什么，都能只根据与自己相关的信息输出接下来的结果。

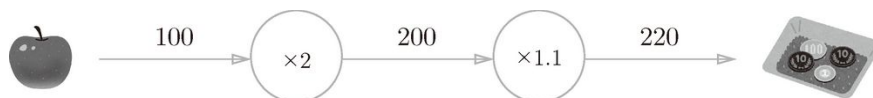


Figure 1.1: Based on the calculation graph to solve the answer

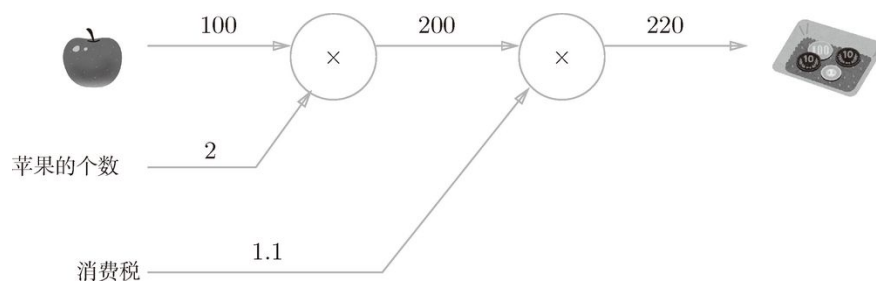


Figure 1.2: Based on the calculation graph to solve the answer2