

Homework 1

Question 1. 实现正向传播和反向传播的推导
在进行证明前，需要先定义一些符号：

表 1. 符号说明

符号含义	含义
$a^{(l)}$	第 1 层经过激活函数激活后的输出
$a^{(L)}$	最后一层输出，这时 $y = a^{(L)}$
x	输入数据, 并且 $x = a^{(0)}$
$w_{ij}^{(l)}$	权重, l 代表第 l 层, i 代表第 l 层的第 i 个神经元, j 代表第 $l-1$ 层的第 j 个神经元
$b^{(l)}$	第 l 层的偏置
$z^{(l)}$	对第 $l-1$ 层的输出进行仿射变化得到结果, 也就是 $z^{(l)} = w^{(l)}a^{(l-1)} + b^{(l)}$
$\delta_{(l)}$	第 l 层的激活函数, 并且有 $a^{(l)} = \delta_{(l)}(z^{(l)})$
t	输出的监督值
$E(y, t)$	损失函数
θ	权重, 就是 w 的总称

正向传播和方向传播都是为了求解得到损失函数关于权重以及偏置的梯度，采用的都是链式法则。

(a) 反向传播公式推导：

由链式法则，可以得到：

$$\begin{aligned}
 (1a) \quad & \left\{ \frac{\partial E}{\partial w_{ij}^{(l)}} = \frac{\partial E}{\partial z^{(l)}} \frac{\partial z^{(l)}}{\partial w_{ij}^{(l)}} \right. \\
 (1b) \quad & \left. \frac{\partial E}{\partial b_j^{(l)}} = \frac{\partial E}{\partial z^{(l)}} \frac{\partial z^{(l)}}{\partial b_j^{(l)}} \right.
 \end{aligned}$$

(b) 正向传播公式推导：

Question 2. 对于一个神经网络 $y = DNN(x, \theta)$, 试求解 $\frac{\partial y}{\partial x}$