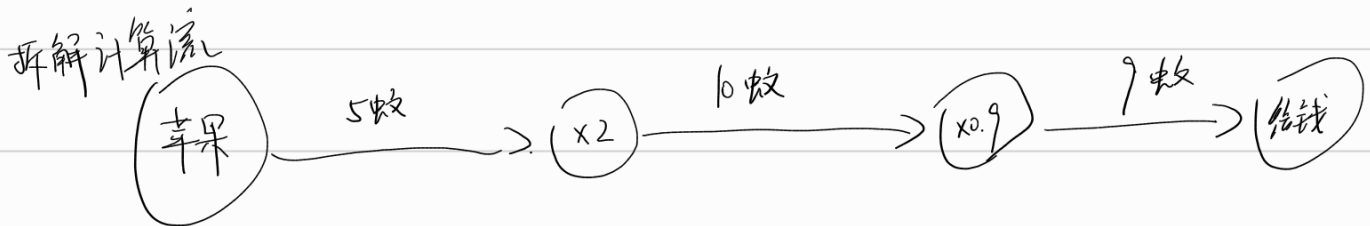


Q: 为什么要有这个?

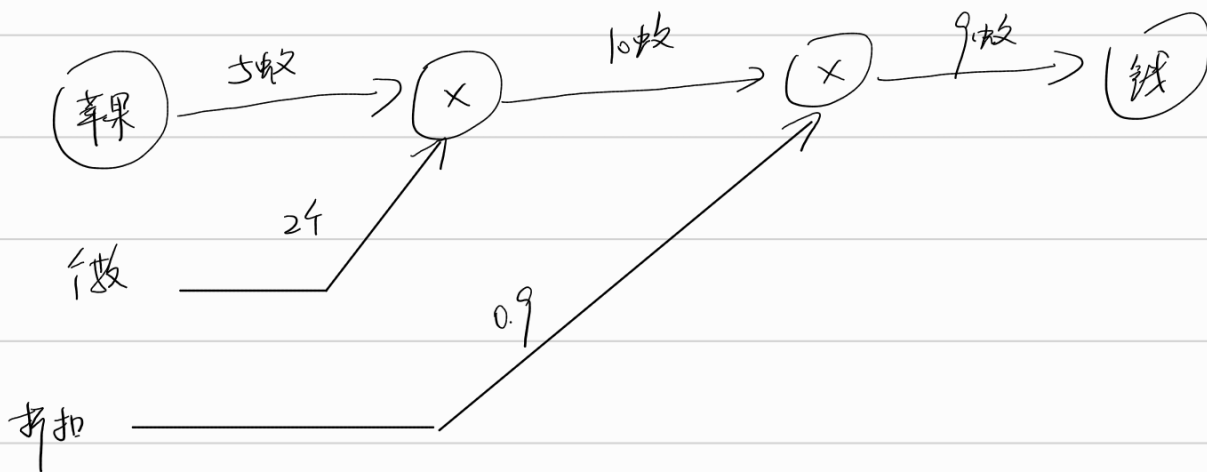
A: 正向传播为从左 (input layer) 到右 (output layer), 依次计算每个节点的值, 直到算出output值之后拿计算值和真实值去算loss (比如均方差)。反向传播为从右到左, 以loss为起点, 求各个点的偏导数, 以此来更新w。

注意, 他们是相互配合的, 不是字面上的替代品!!!

构造场景: 小明去水果店买苹果, 一个5块, 买了两个, 然后被告知可以打九折



拆解过程和数值



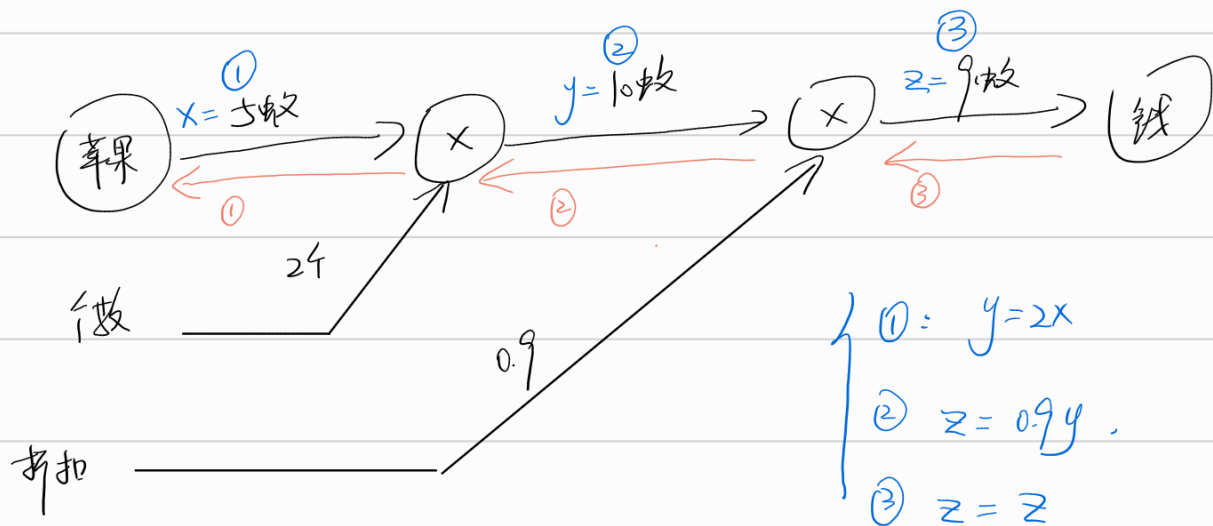
正常人都是从左向右看, 这就是正向传播

forward propagation

如果你右往左看, 那就是反向传播

backward propagation

比如我要求支付金额关于苹果价格的导数



从反向的角度, 我们从右往左看, 得到.

$$\begin{cases} ③ = \frac{\partial z}{\partial z} \\ ② = \frac{\partial z}{\partial y} \\ ① = \frac{\partial z}{\partial x} \end{cases}$$

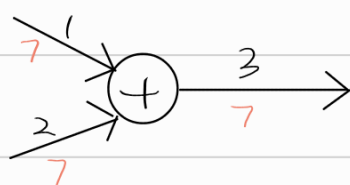
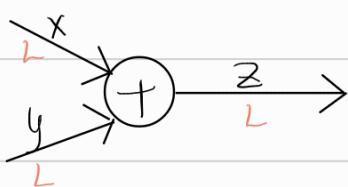
tips: 根据高数中的链式法则 (chain rule). 若 $z = (x+y)^2$, 我们求 z 关于 x 的导数, 我们可得.

$$\begin{cases} z = t^2 \\ t = x+y \end{cases} \Rightarrow \frac{\partial z}{\partial x} = \frac{\partial z}{\partial t} \cdot \frac{\partial t}{\partial x} = 2t = 2(x+y)$$

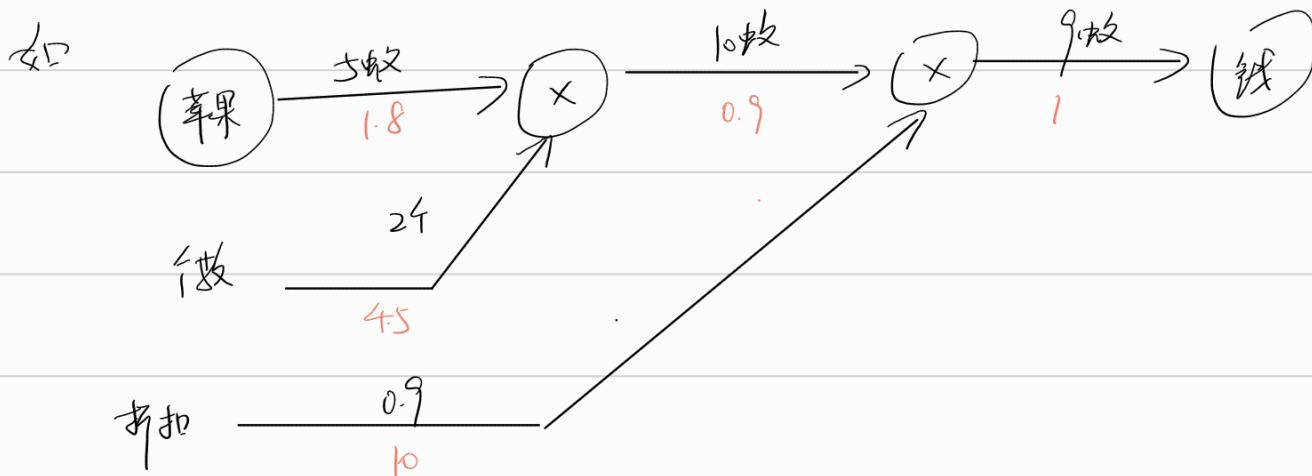
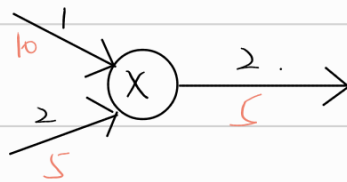
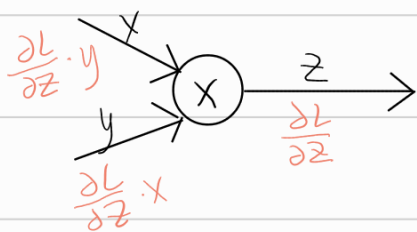
接上可得

$$\begin{cases} ③ = \frac{\partial z}{\partial z} = 1 \\ ② = \frac{\partial z}{\partial y} = \frac{\partial z}{\partial z} \cdot \frac{\partial z}{\partial y} = (0.9y)' = 0.9 \\ ① = \frac{\partial z}{\partial x} = \frac{\partial z}{\partial z} \cdot \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial x} = 0.9 \cdot (2x)' = 1.8 \end{cases}$$

其中, 对于符号为 "+" 的, 该节点前后的数值相同, 比如:



对于“x”乘法，则要乘以对应的值，即将输入信号~~反向传播~~后来以下游的值。



对于激活函数。

$$\text{ReLU}: y = \begin{cases} x & x > 0 \\ 0 & x \leq 0 \end{cases}$$

$$\frac{\partial y}{\partial x} = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases}$$

因为其实在数学上，会用数值微分法去算偏导，因其简单，错少，但巨慢

$$\frac{df(x)}{dx} = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

所以，为什么要有反向，就是不想每次计算都从左算到右，而是可以通过最后的 loss，直接获得梯度值，以此更新 w 的值。

我懒得证明为什么反向传播是对的，这个我猜也不会考。

但是要是工程上，可以数值微分法去check这玩意算的到底对不对，这个过程就比较简单了。感兴趣问我要代码。