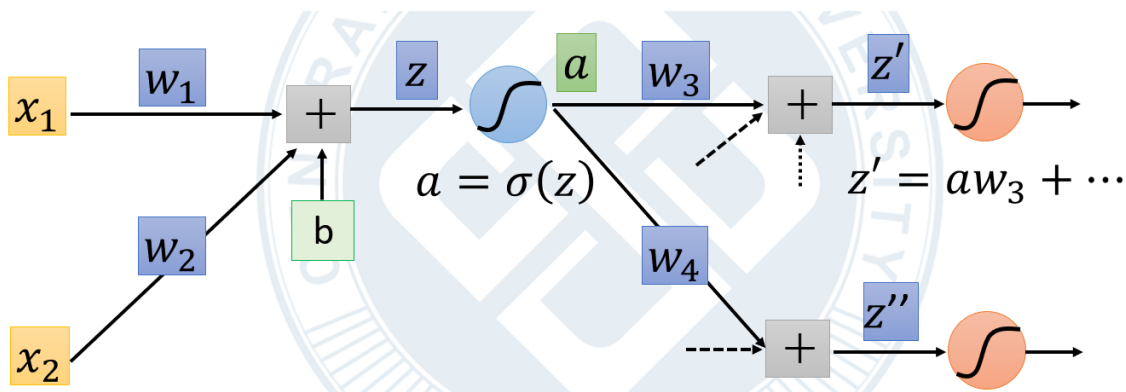


杂谈勾股定理

无名氏

2024 年 1 月 20 日

1 公式测试



已知： z 为激活函数输入， a 为激活函数输出和下一层的输入， l 为损失函数。

$$z = w_1 x_1 + w_2 x_2 + b$$

$$a = \sigma(z)$$

每个权重参数的梯度计算方式为

$$\begin{aligned} \frac{\partial l}{\partial w_1} &= \frac{\partial z}{\partial w_1} \cdot \frac{\partial l}{\partial z} \\ &= \frac{\partial z}{\partial w_1} \cdot \frac{\partial a}{\partial z} \cdot \frac{\partial l}{\partial a} \\ &= x_1 \cdot \frac{\partial a}{\partial z} \cdot \frac{\partial l}{\partial a} \\ &= x_1 \cdot \sigma'(z) \cdot \frac{\partial l}{\partial a} \end{aligned} \tag{1}$$

发现只有 $\frac{\partial l}{\partial a}$ 是未知的，需要继续计算。

$$\begin{aligned} \frac{\partial l}{\partial a} &= \frac{\partial z'}{\partial a} \cdot \frac{\partial l}{\partial z'} + \frac{\partial z''}{\partial a} \cdot \frac{\partial l}{\partial z''} \\ &= w_3 \cdot \frac{\partial l}{\partial z'} + w_4 \cdot \frac{\partial l}{\partial z''} \\ &= w_3 \cdot \sigma'(z') \cdot \frac{\partial l}{\partial a'} + w_4 \cdot \sigma'(z'') \cdot \frac{\partial l}{\partial a''} \\ &= \dots \end{aligned} \tag{2}$$

发现从前向后计算梯度时，需要后续的所有的 $\frac{\partial l}{\partial z}$ 或 $\frac{\partial l}{\partial a}$ 的具体数值；而从后向前计算梯度，可以直接计算出梯度，并且计算结果可以在计算前层时直接使用，因此采用从后向前计算梯度的方式，即反向传播。

发现从前向后计算梯度时，需要后续的所有的 $\frac{\partial l}{\partial z}$ 或 $\frac{\partial l}{\partial a}$ 的具体数值；而从后向前计算梯度，可以直接计算出梯度，并且计算结果可以在计算前层时直接使用，因此采用从后向前计算梯度的方式，即反向传播。