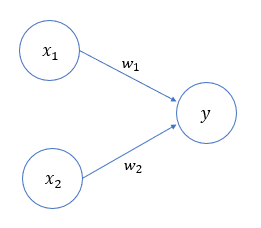
大模型技术之深度学习核心

# 2.神经网络基础

## 神经网络的的构成

### 感知机

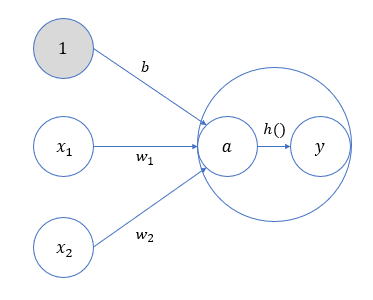
二分类模型，接受多个信号，输出一个信号，且只有0和1的取值；



是输入信号，是输出信号，是权重，○ 称为神经元或节点激活函数

参数，被称为 **偏置；**偏置则可以用来控制神经元被激活的容易程度

### 激活函数



可以将输入信号的加权总和转换为输出信号，起到“激活神经元”的作用，所以被称为 **激活函数**。

如果定义一个函数：

（2.2）

那么式（2.1）就可以简化为：

（2.3）

为了更明显地表示出两个处理步骤，可以进一步写成：

（2.4）

（2.5）

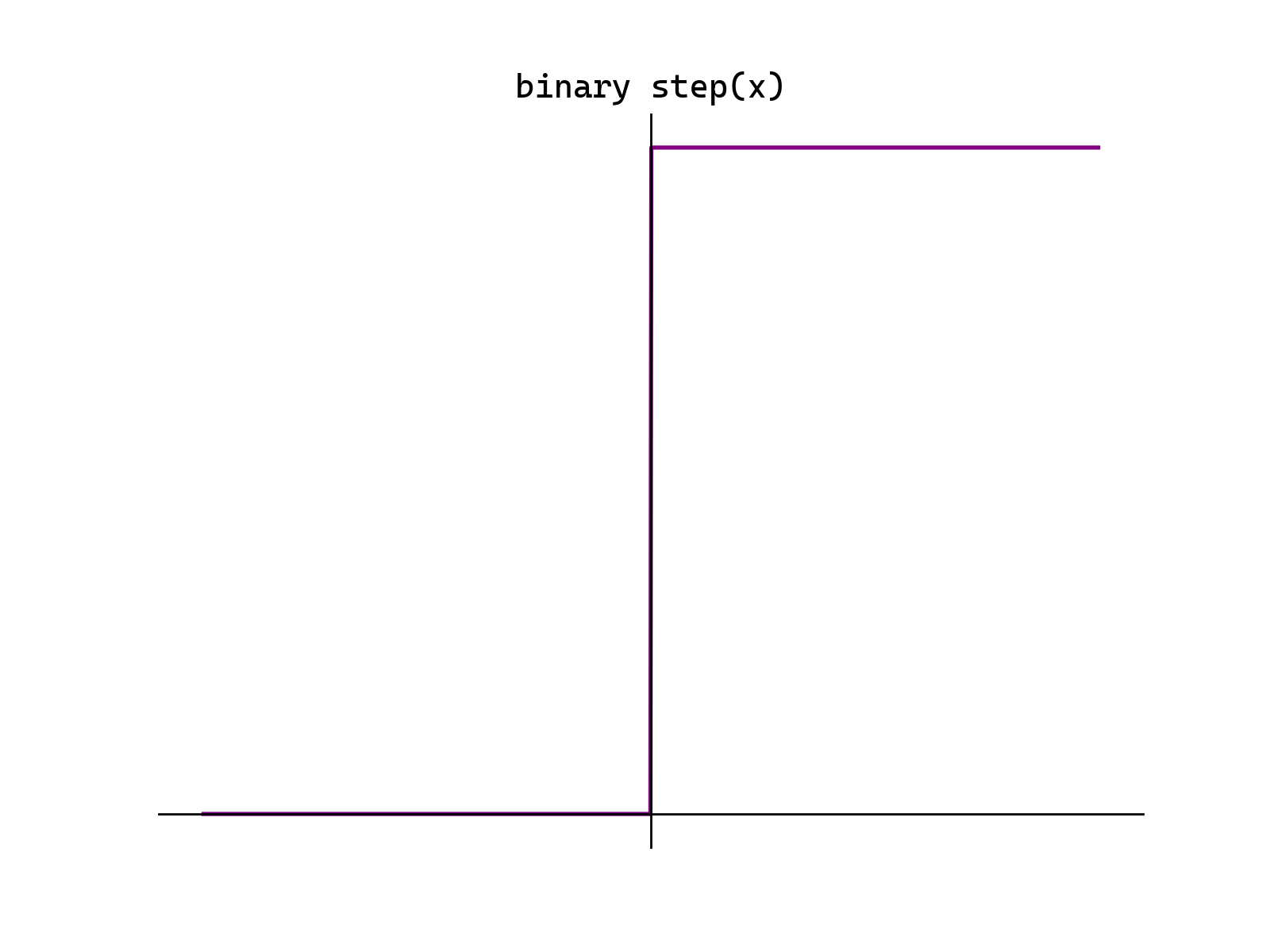
#### 激活函数的作用

a. 激活函数式连接感知机和神经网络的桥梁

b. 激活函数，整个神经网络就等效于单层线性变换，不论如何加深层数，总是存在与之等效的“无隐藏层的神经网络”;

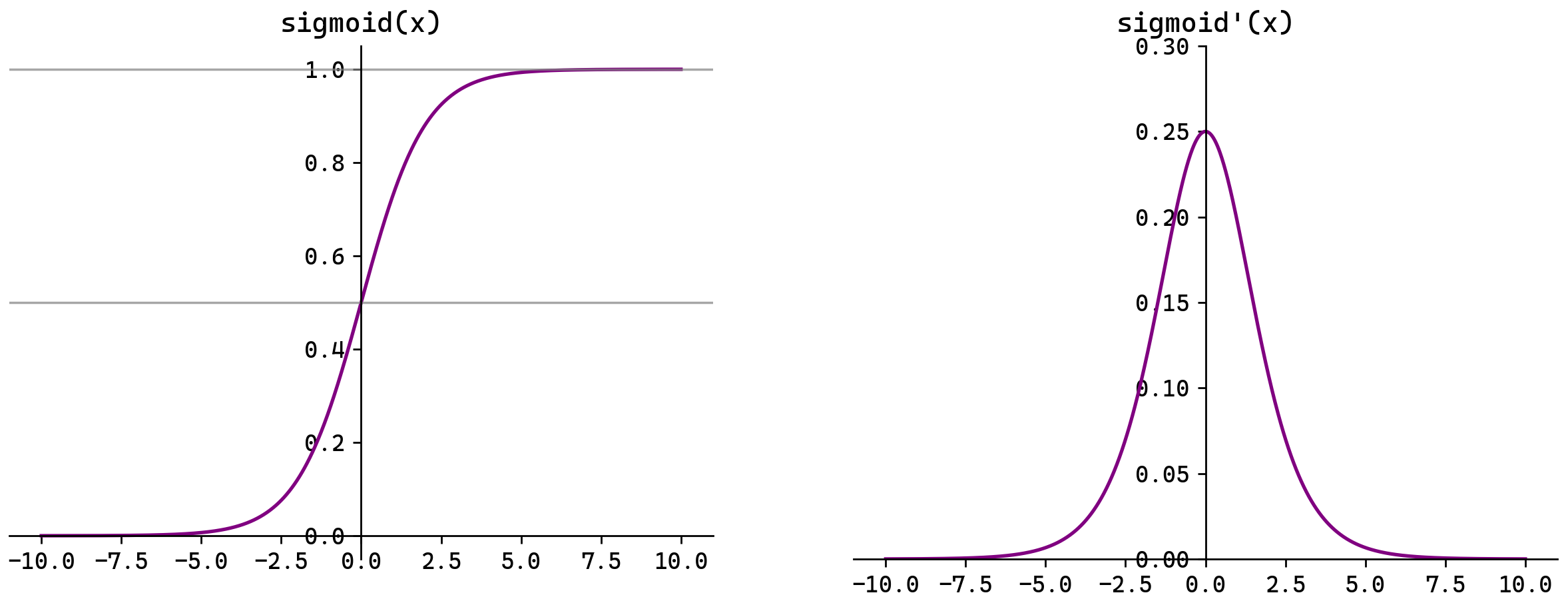
c. 激活函数必须是非线性函数，也正是激活函数的存在为神经网络引入了非线性，使得神经网络能够学习和表示复杂的非线性关系

#### 阶跃函数



阶跃函数的导数恒为零，所以应用不多；

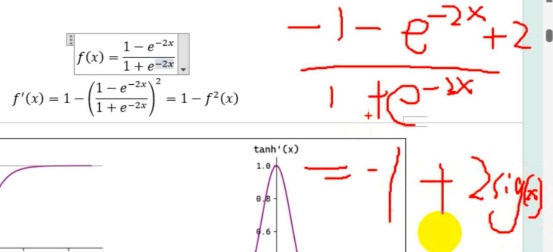
#### Sigmoid函数

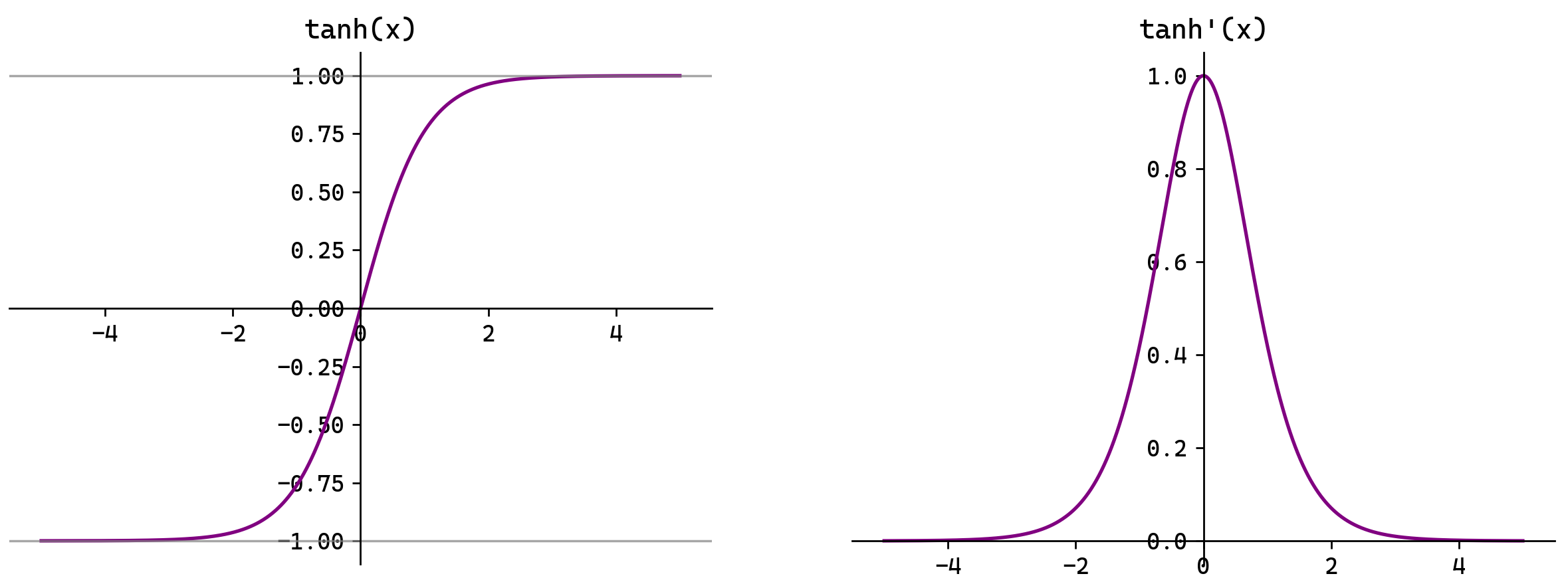


Sigmoid的导数范围为(0,0.25)，梯度较小。当输入在[-6,6]之外时，导数接近0，此时网络参数的更新将会极其缓慢。使用Sigmoid作为激活函数，可能出现梯度消失（在逐层反向传播时，梯度会呈指数级衰减），一般只作为输出层，或者只有两三层的中间层（隐藏层）；

#### Tanh（双曲正切）函数

在sigmod函数上先乘以2再减去1之后，得到的函数；





Tanh的输出以0为中心，且其梯度相较于Sigmoid更大，收敛速度相对更快。但同样也存在梯度消失现象；