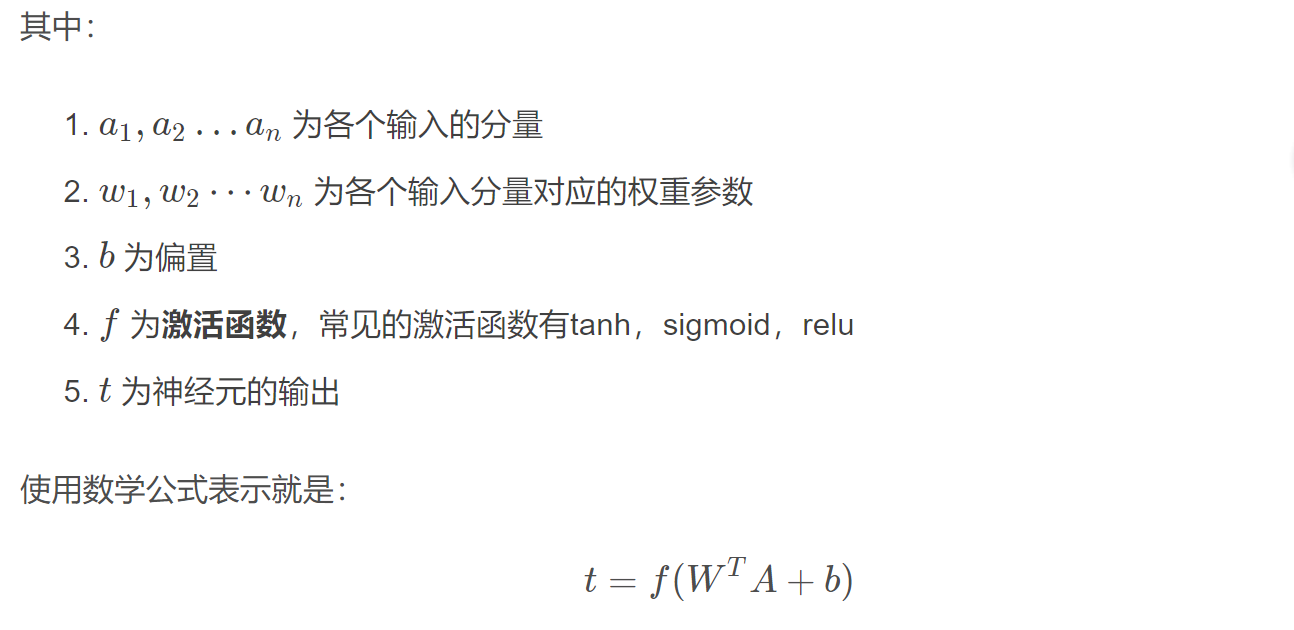
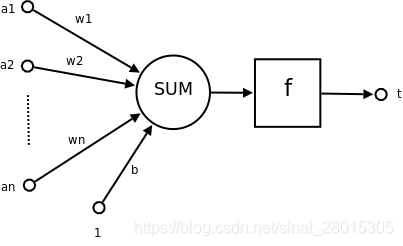
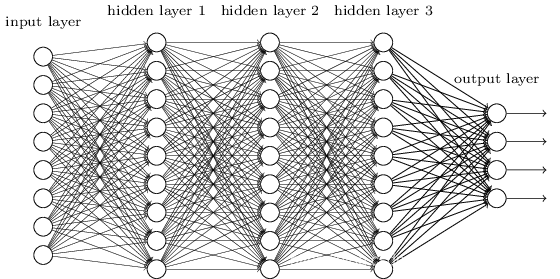
****人工神经网络****（Artificial Neural Network，ANN），简称****神经网络****（Neural Network，NN）或****类神经网络****，是一种模仿生物神经网络（动物的中枢神经系统，特别是大脑）的结构和功能的数学模型，用于对函数进行估计或近似。

单个神经元的结构：



可见，****一个神经元的功能是求得输入向量与权向量的内积后，经一个非线性传递函数得到一个标量结果****。

多层神经网络就是由单层神经网络进行叠加之后得到的，所以就形成了****层****的概念，常见的多层神经网络有如下结构：



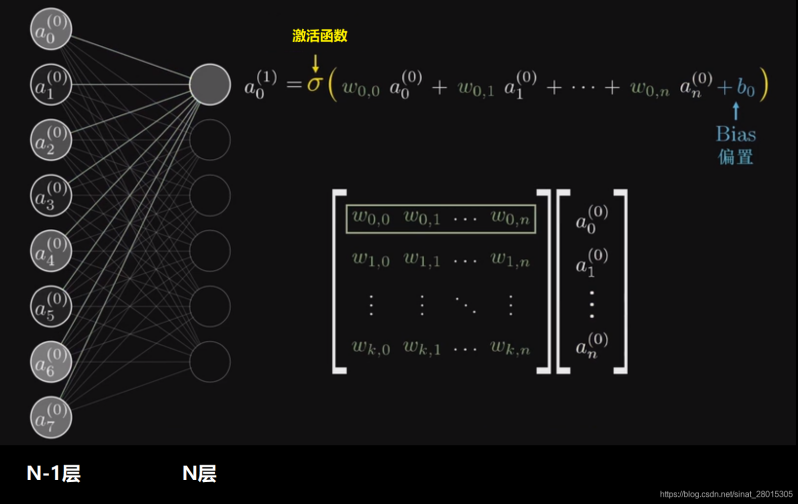
输入层（Input layer），众多神经元（Neuron）接受大量输入消息。输入的消息称为输入向量。

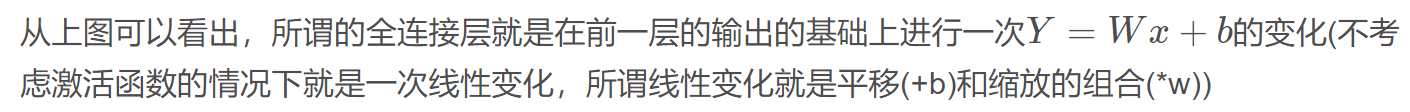
输出层（Output layer），消息在神经元链接中传输、分析、权衡，形成输出结果。输出的消息称为输出向量。

隐藏层（Hidden layer），简称“隐层”，是输入层和输出层之间众多神经元和链接组成的各个层面。隐层可以有一层或多层。隐层的节点（神经元）数目不定，但数目越多神经网络的非线性越显著，从而神经网络的强健性（robustness）更显著。

全连接层：当前一层和前一层每个神经元相互链接，我们称当前这一层为全连接层。

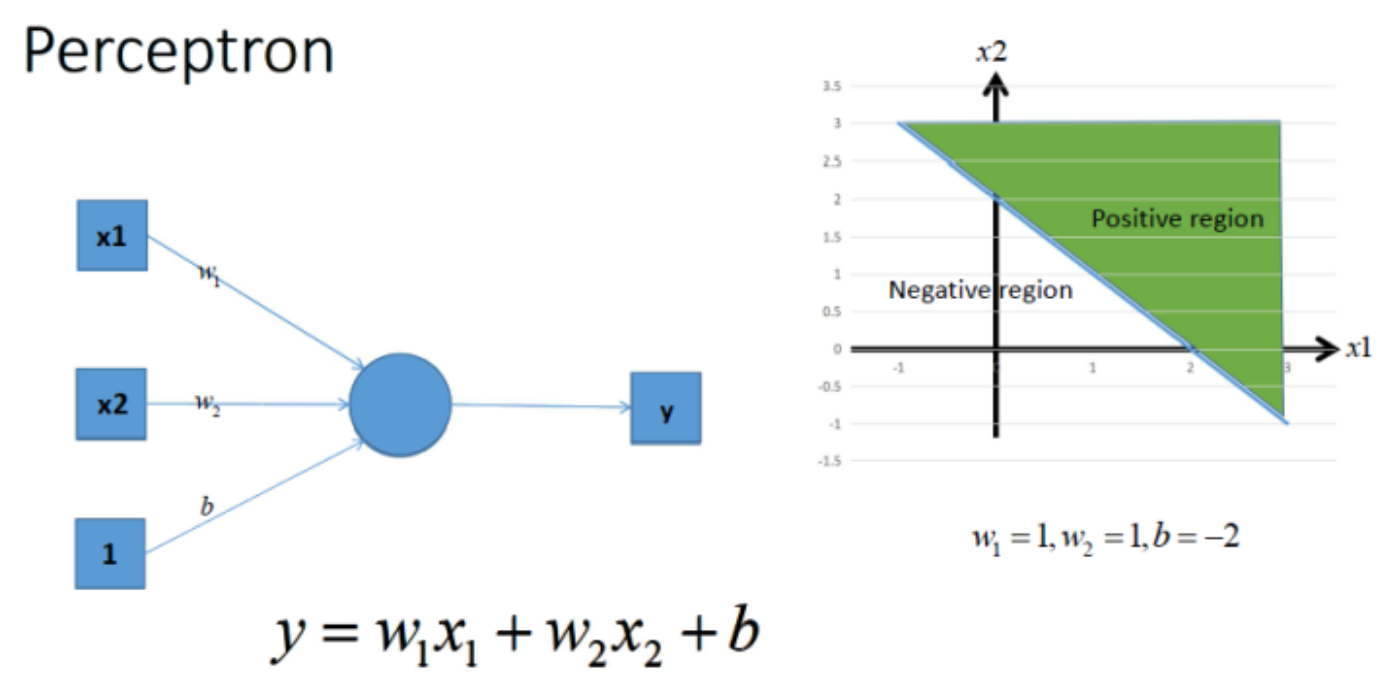
思考：假设第N-1层有m个神经元，第N层有n个神经元，当第N层是全连接层的时候，则N-1和N层之间有m\*n个参数，这些参数可以如何表示？





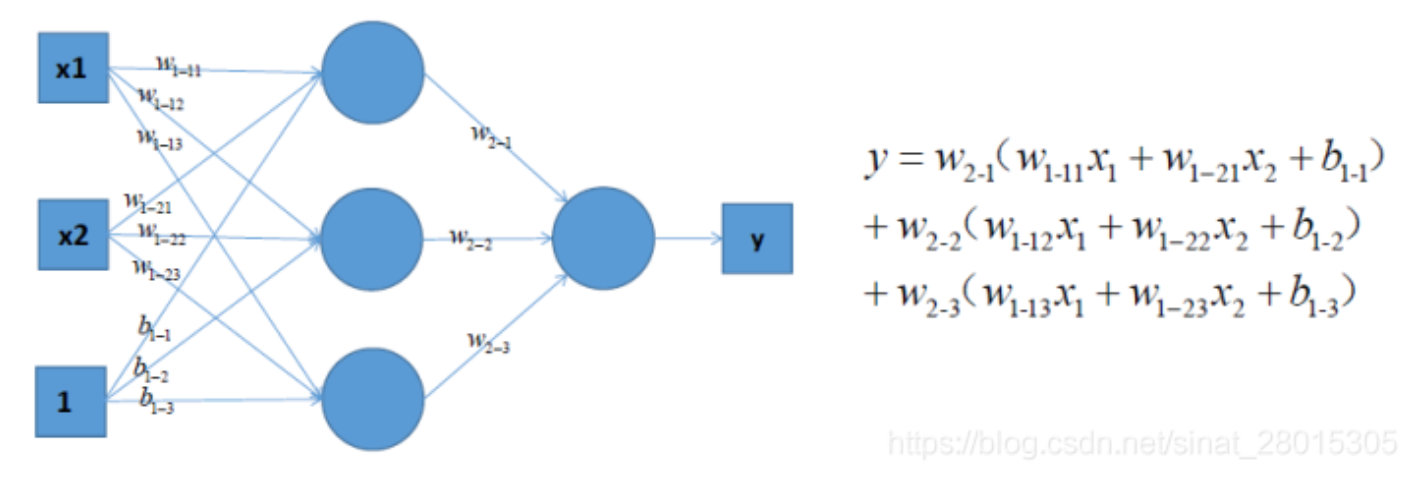
**激活函数的作用：**

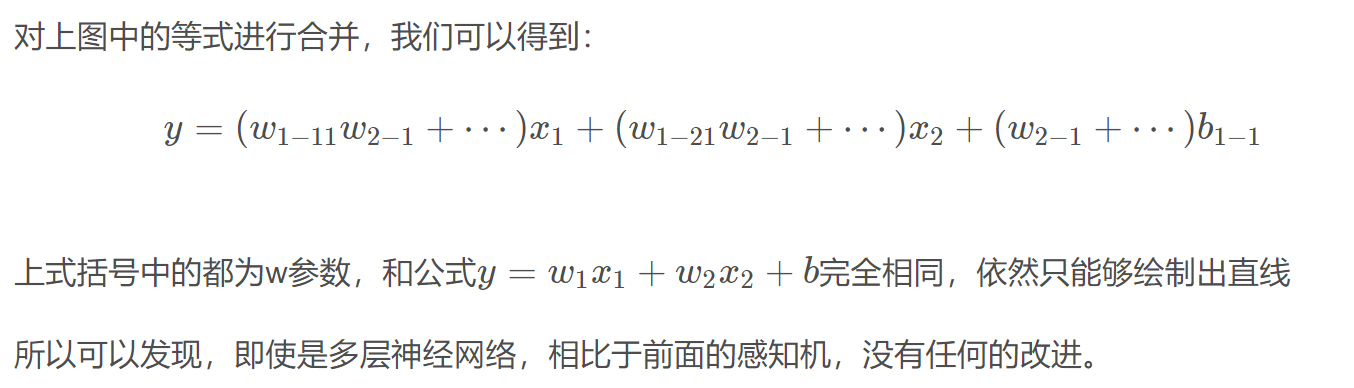
通过不带激活函数的感知机模型（只有输入输出层）我们可以划出一条线, 把平面分割开：



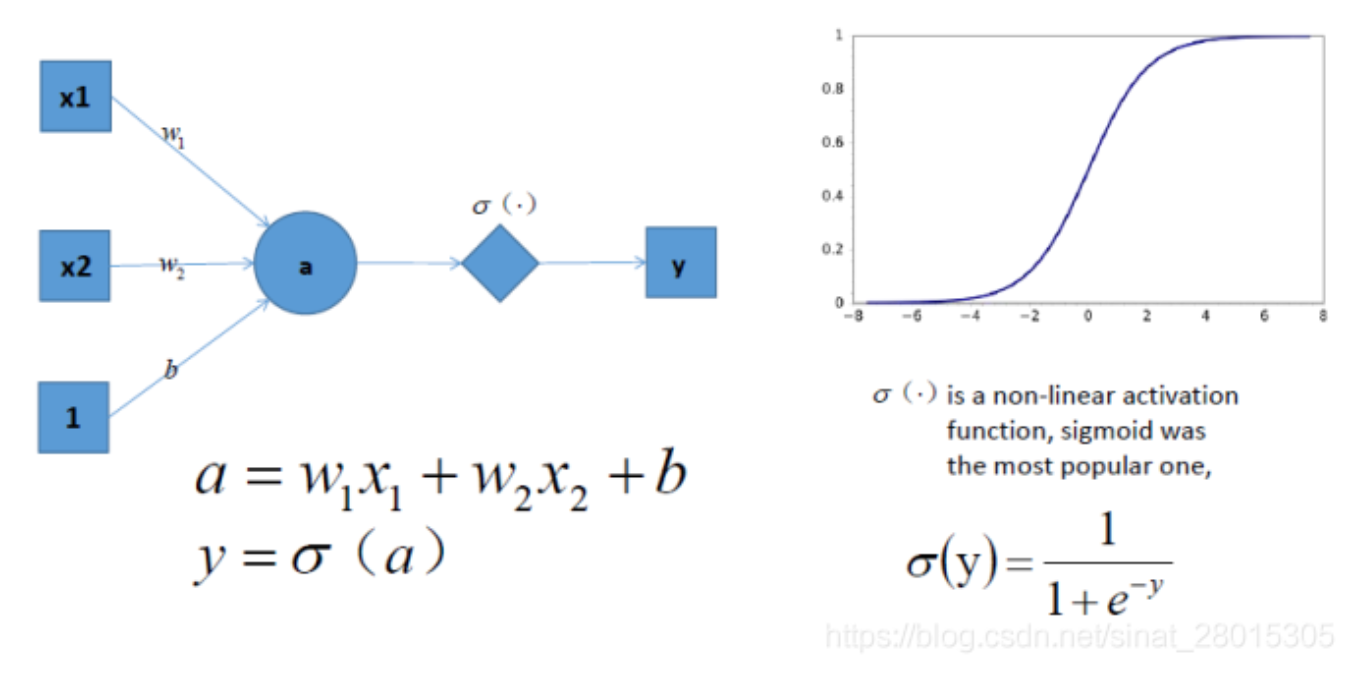
只有一条直线的分割效果并不好，那么这个时候该怎么办？

可以考虑使用多层神经网络来进行尝试，比如****在前面的感知机模型中再增加一层****





但是如果此时，我们在前面感知机的基础上加上****非线性的激活函数****之后，输出的结果就不再是一条直线

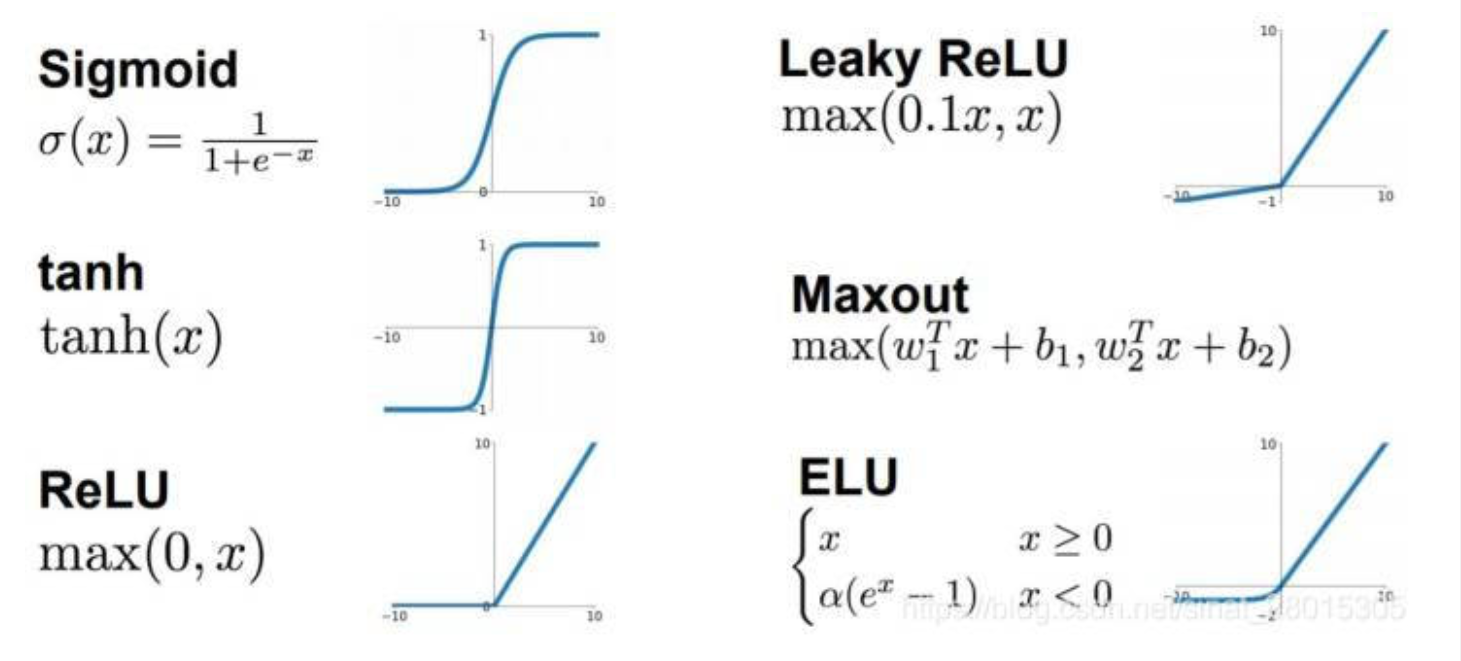


如上图，右边是sigmoid函数，对感知机的结果，通过sigmoid函数进行处理

如果给定合适的参数w和b，就可以得到合适的曲线，能够完成对最开始问题的非线性分割

所以激活函数很重要的一个****作用****就是****增加模型的非线性分割能力****

常见的激活函数有：



看图可知：

sigmoid 只会输出正数，以及靠近0的输出变化率最大

tanh和sigmoid不同的是，tanh输出可以是负数

Relu是输入只能大于0,如果你输入含有负数，Relu就不适合，如果你的输入是图片格式，Relu就挺常用的，因为图片的像素值作为输入时取值为[0,255]。

激活函数的作用除了前面说的****增加模型的非线性分割能力****外，还有

* ****提高模型鲁棒性****
* ****缓解梯度消失问题****
* ****加速模型收敛等****