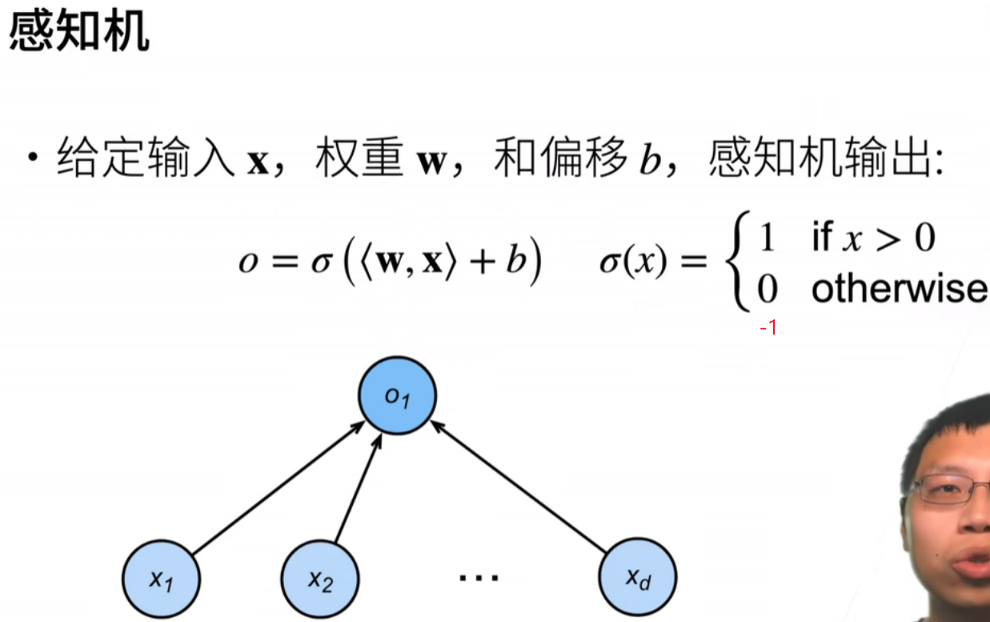
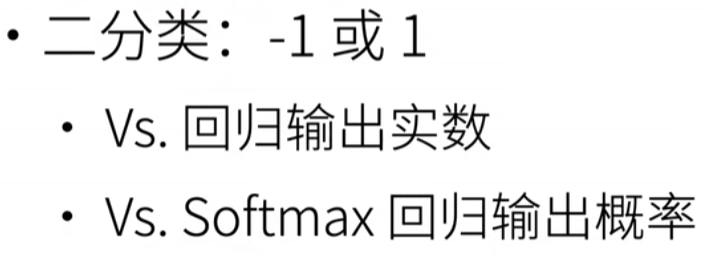
多层感知机

**感知机**：感知机是人工智能最早的模型，是一个有监督的学习算法，两分类任务，只有一个元素输出，如下，w和x做内积后加上b再过sigma函数，其值大于0输出1，值小于等于0输出0【也可以输出-1】

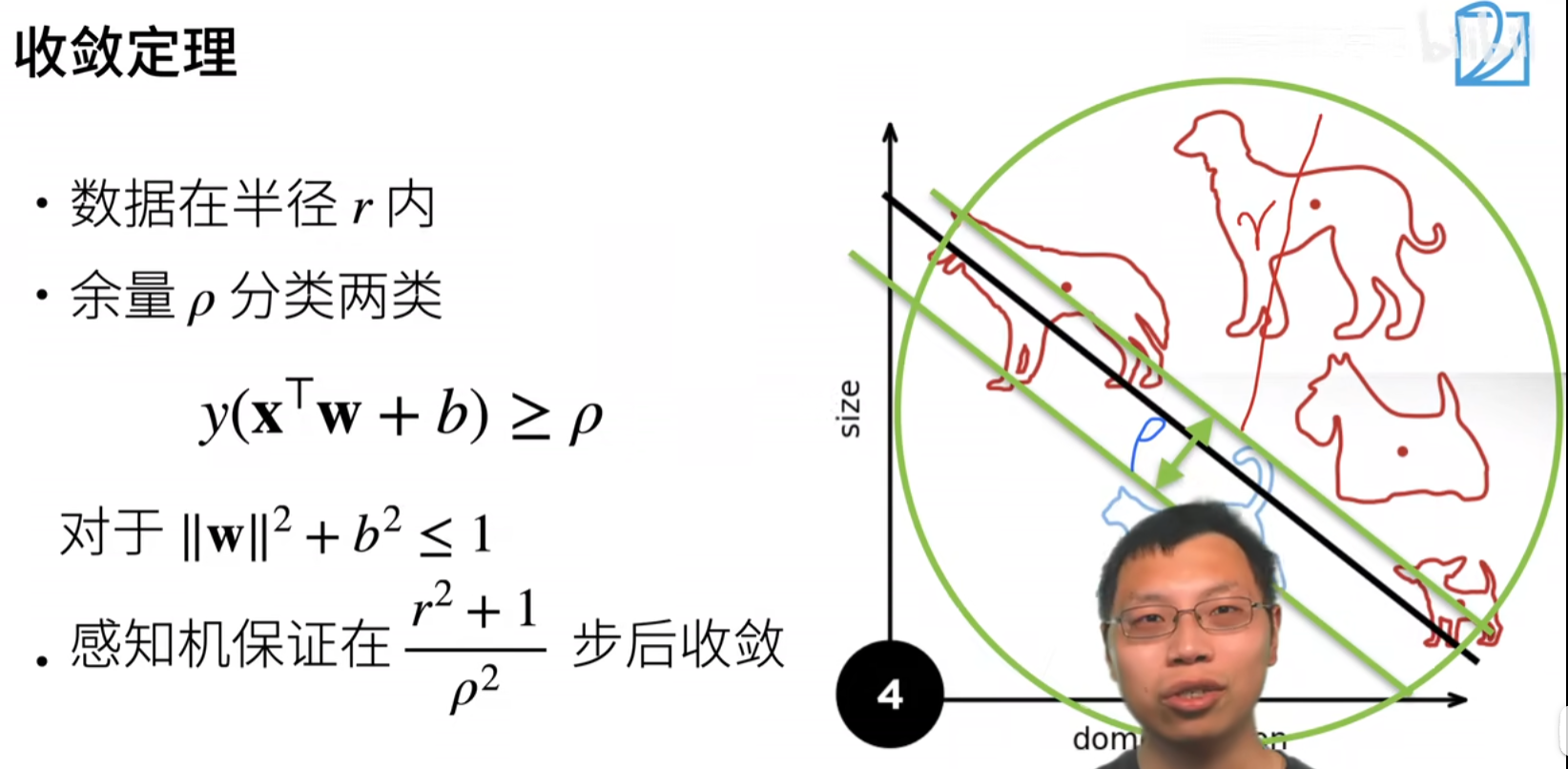


**与线性回归的不同**：线性回归输出的是一个实数而感知机输出的是一个离散的类。  
**与softmax的区别**：在有n个类的情况下，softmax会输出n个元素，所以可以是一个多分类的问题，但是感知机这里只输出一个元素，说明感知机最多只能做一个二分类问题 。



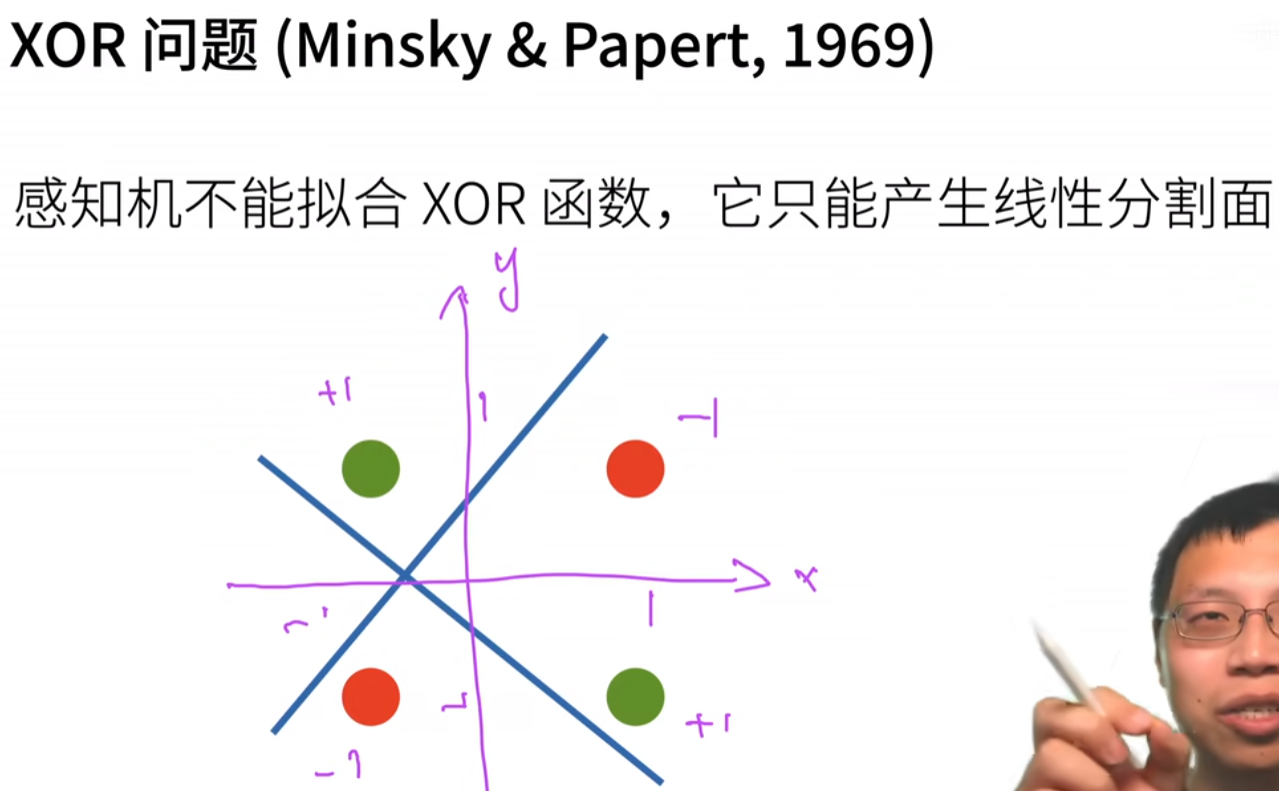
预测值大于0应该是正类，预测是小于等于0应该是负类，当预测正确的时候，真实值y乘以预测值<w,x>+b是正值大于0，当**预测失败**的时候，**真实值y乘以预测值<w,x>+b是负值小于等于0**，对每一个样本都做这样的判断，直到所有样本都分类正确。

猫狗类别划分，根据每一个类别是否划分正确，更新权重。



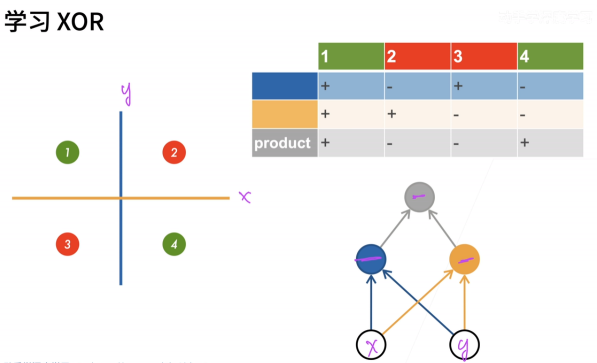
做两个假设，数据区域半径r，余量 ρ，存在一个分截面，使得分截面对所有分类都是正确的，而且有一定余量 ρ（≥0）。r平方是数据的大小，ρ是数据是不是很好，两个分类是否能很好分开。

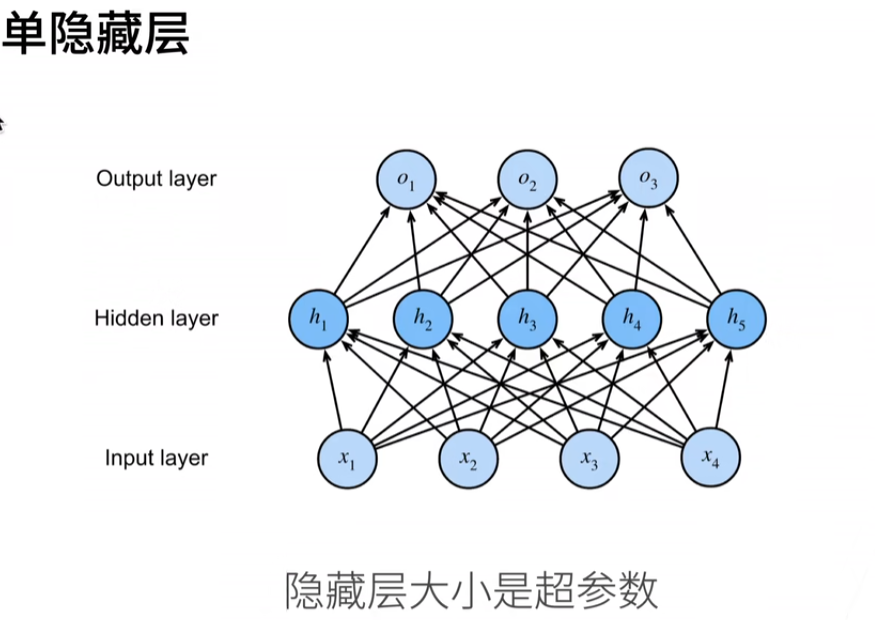
在任意二维坐标系内，无法找出一个线性分割面能将下图的点分割开，无法做分类。多层感知机能解决这个问题。



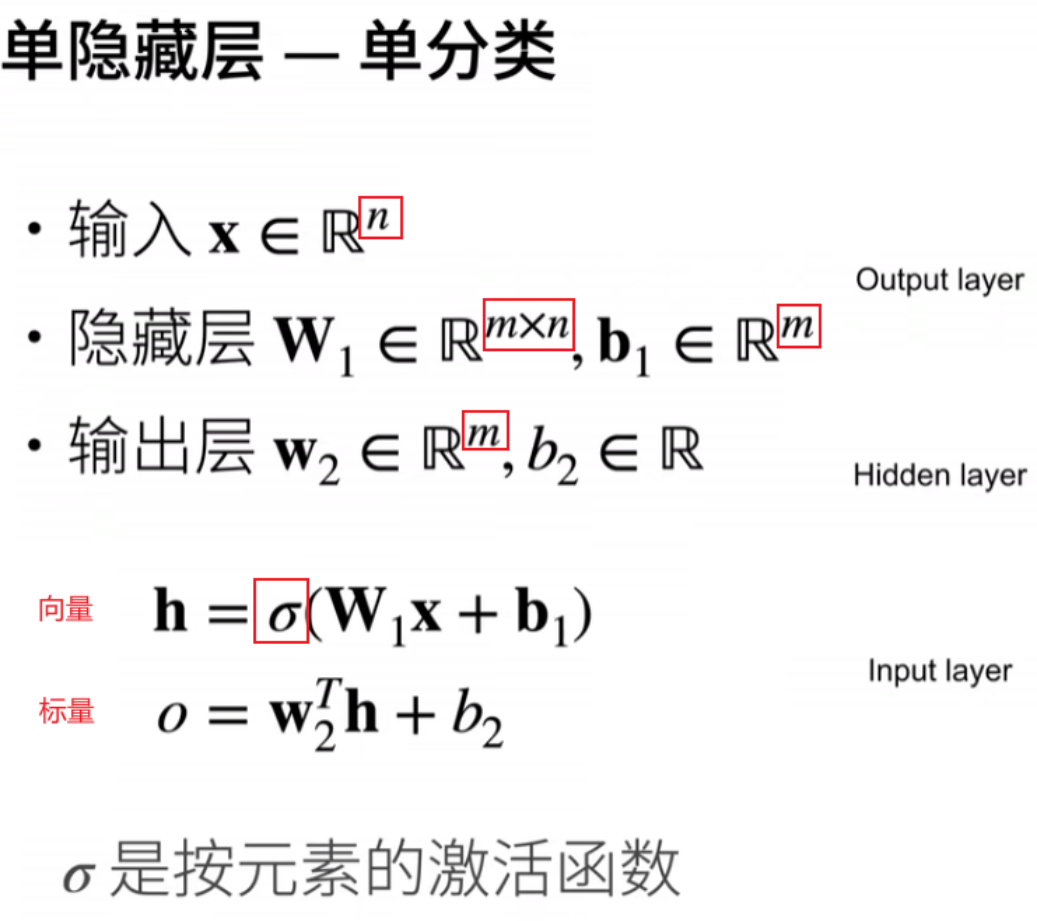
**多层感知机（MLP）**是一种由多个层组成的前馈神经网络，主要用于解决监督学习问题。它通过多个隐藏层的非线性变换来学习输入与输出之间的复杂映射关系。

第一步学习蓝色的线，第二步学习黄色的线，复合两个函数，对两个结果做异或操作得出结果

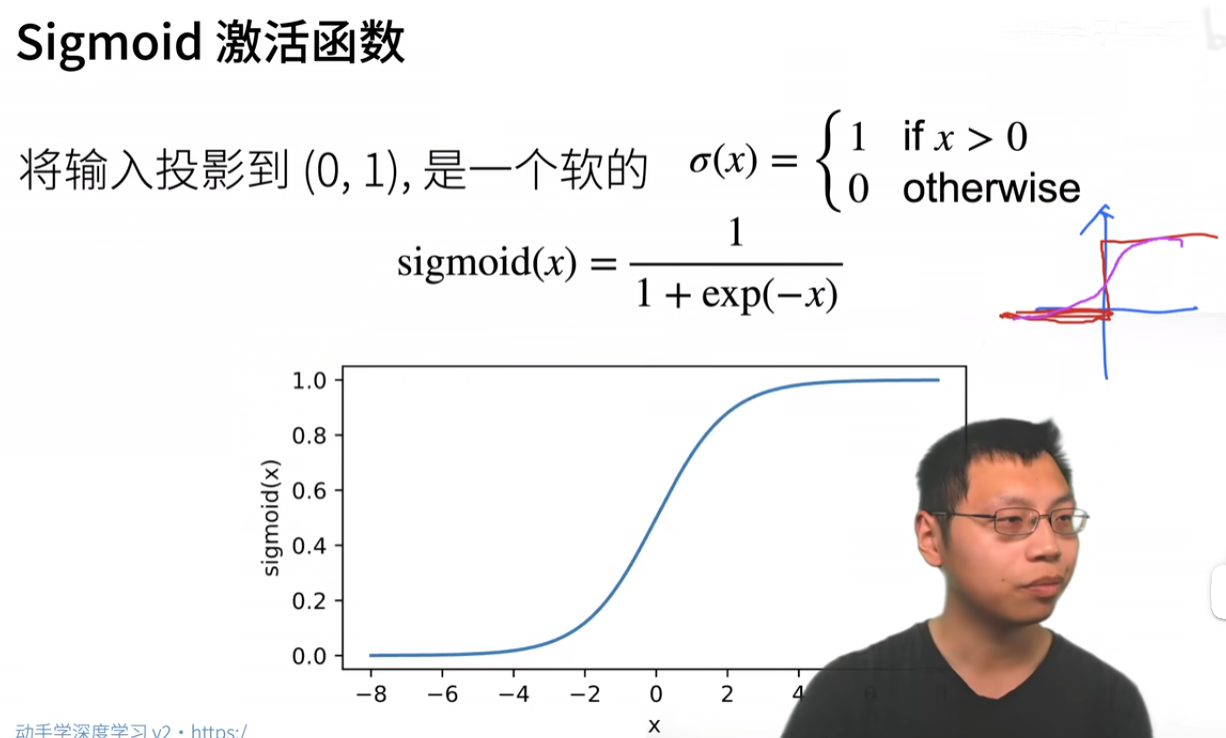


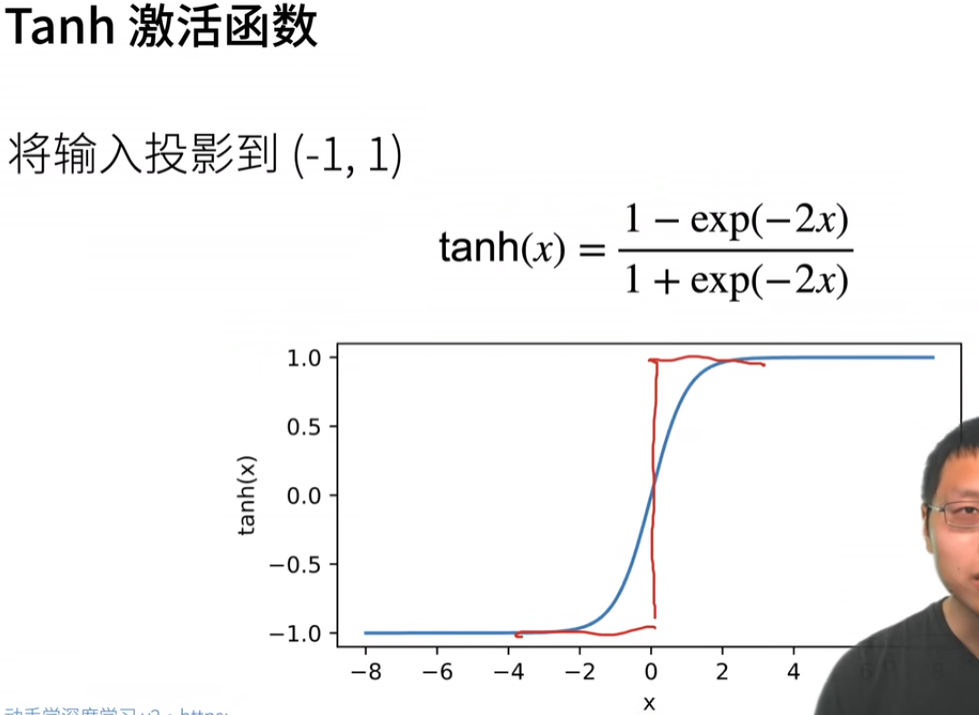


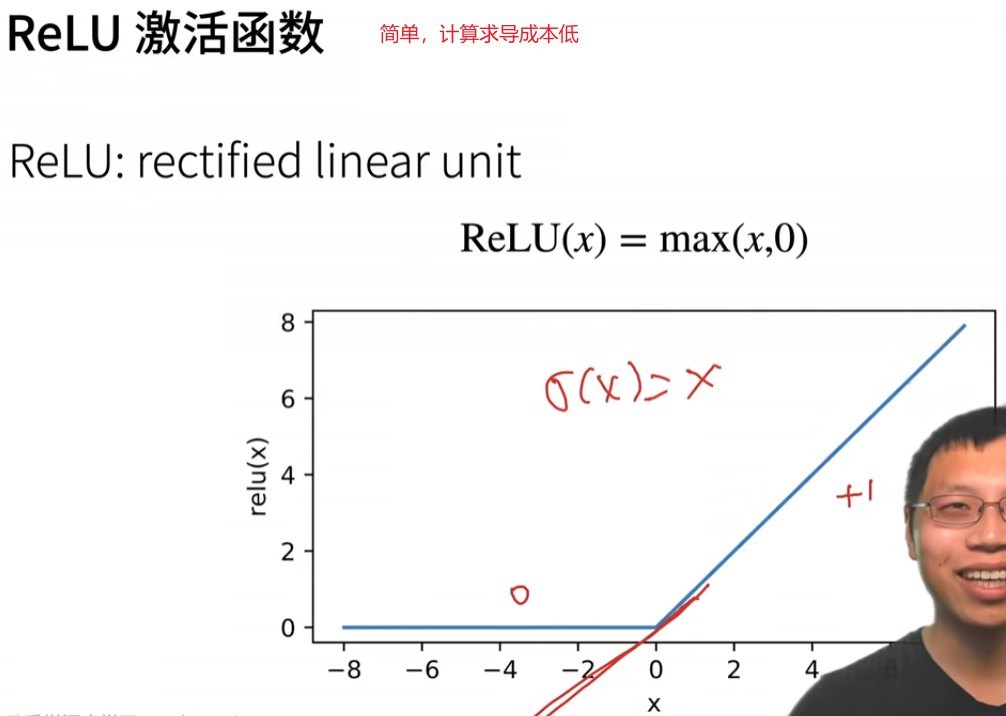
输入输出层大小是由数据决定的，数据有多大，数据有多少类。隐藏层大小是可以自行设置的



注：sigma函数不能是线性的，否则等价于单层的感知机，引入激活函数就是为了引入非线性，使得神经网络能够学习复杂的模式和特征。



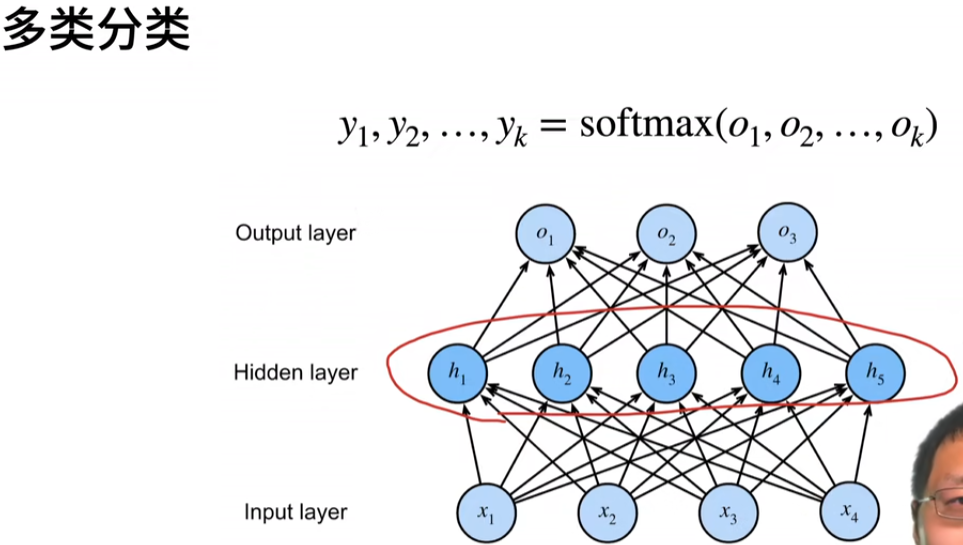


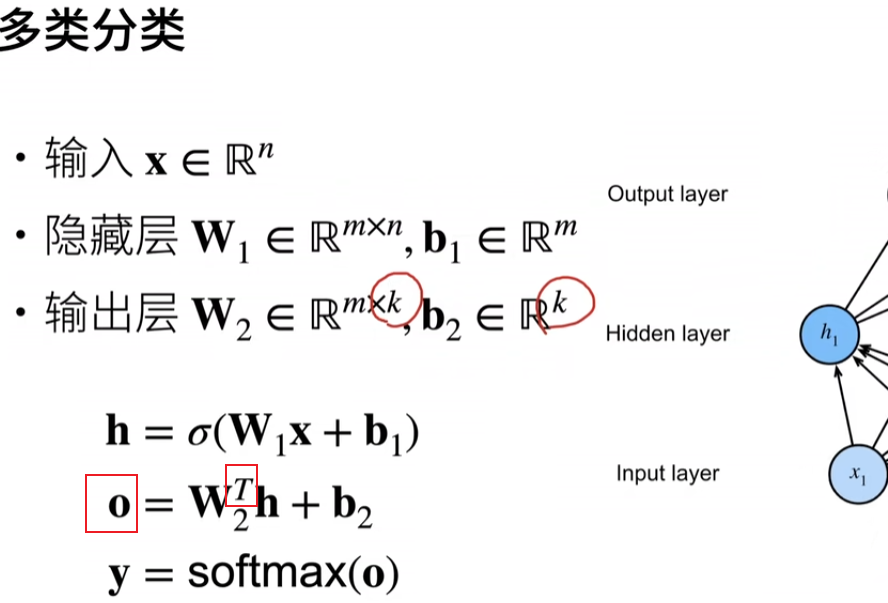


在cpu上计算指数函数比较贵，计算一次指数计算等于计算几百次乘法运算的花销。所以常用ReLU激活函数【计算简单】。

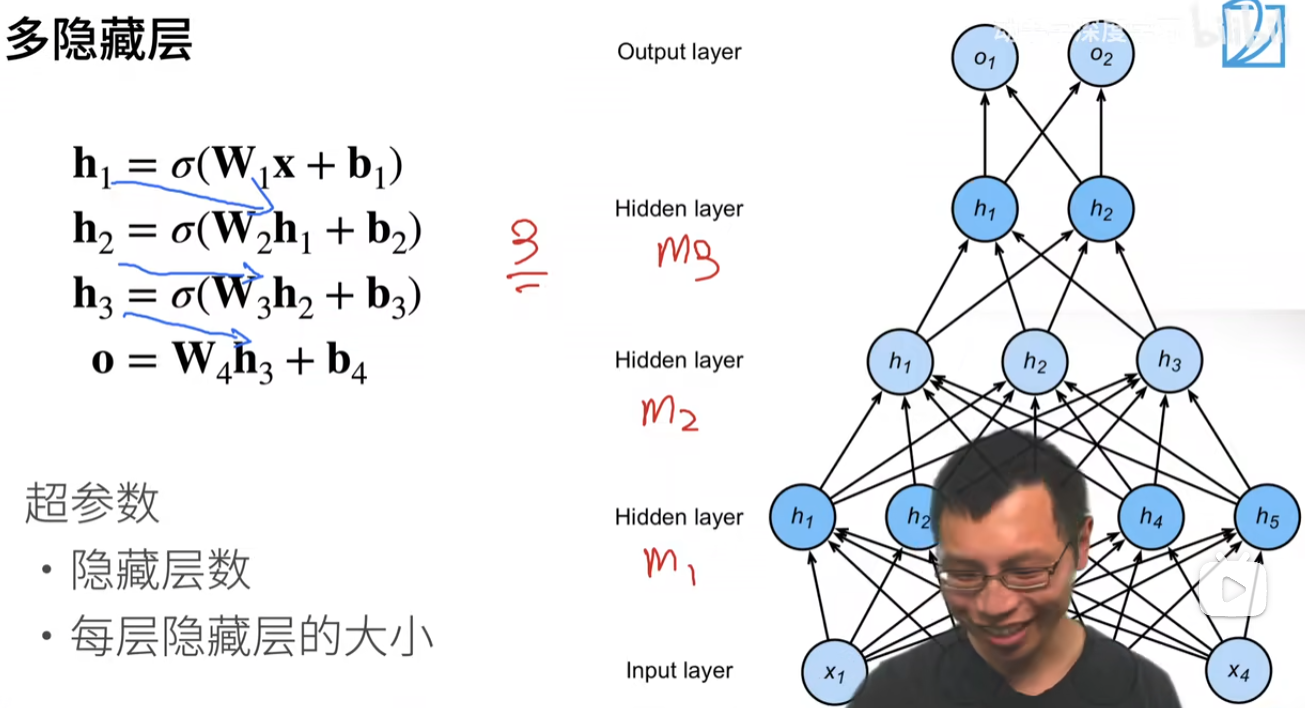
softmax做的事情：把所有输出单元拉到0-1之间，且所有分类的值等于1，表示为概率的形式

多类分类感知机和softmax回归没有本质区别：没加隐藏层就是softmax回归，加了隐藏层就是多层感知机，只是名称上的变化





和前面单分类的区别，就是在输出层，因为有k个输出，所以输出层权重w2是m\*k的，偏移也是k维向量。同时对输出要做一次softmax



激活函数不能少，少了层数会减一，主要是为了避免层数的塌陷，最后一层输出层不需要激活函数。

超参数有两个：有几个隐藏层，每个隐藏层是多大。

配置好每一层隐藏层都有哪些东西、长什么样子（技术经验）。

根据输入数据复杂度的多少（维度高低），考虑模型（大小和复杂度）。

两个选择：

1. 模型做的不深，但是每一层都很大
2. 每一层不大，但是模型比较深

假设输入数据是128维（高），输出是5类（低），把高维数据压缩到低维。机器学习本质上是做压缩，把图片或者其他信息压缩成某一个类。

压缩的三个做法：

1. 逐层慢慢压缩

2. 第一层隐藏层可以比输入层稍微大一点，先扩张再压缩

3. 卷积神经–先压缩再扩张的模型，避免overfitting。

多层感知机代码：比softmax回归多加了隐藏层

**深度学习的好处，模型变化很大（MLP cnn等多模型），但是从代码实现的角度，只是改了一点点(模型结构的代码) 所以用MLP–多层感知机 Multi layer Perceptron比较多，而SVM–支持向量机Support Vector Machine用的比较少(因为要改很多东西)**



