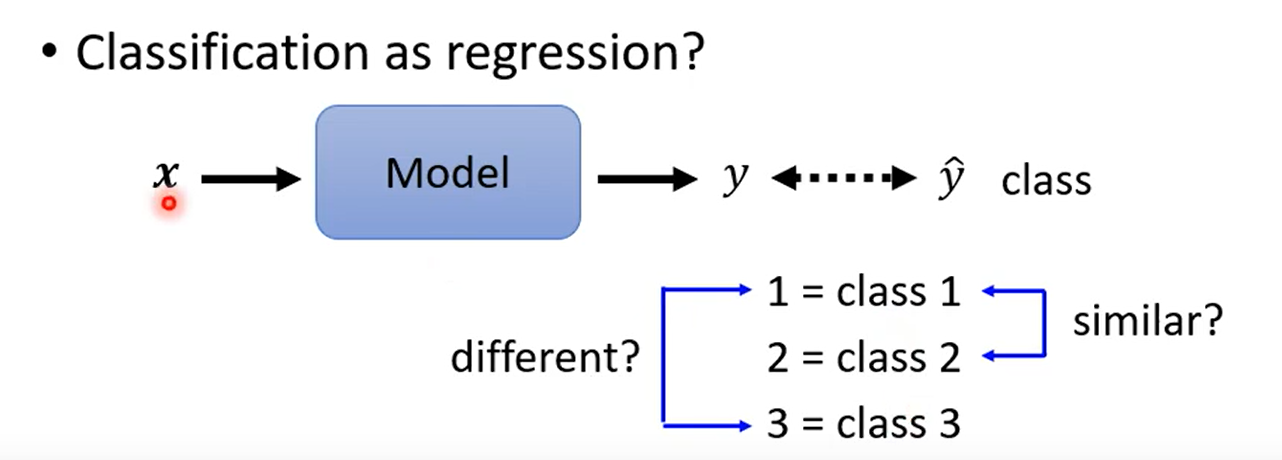
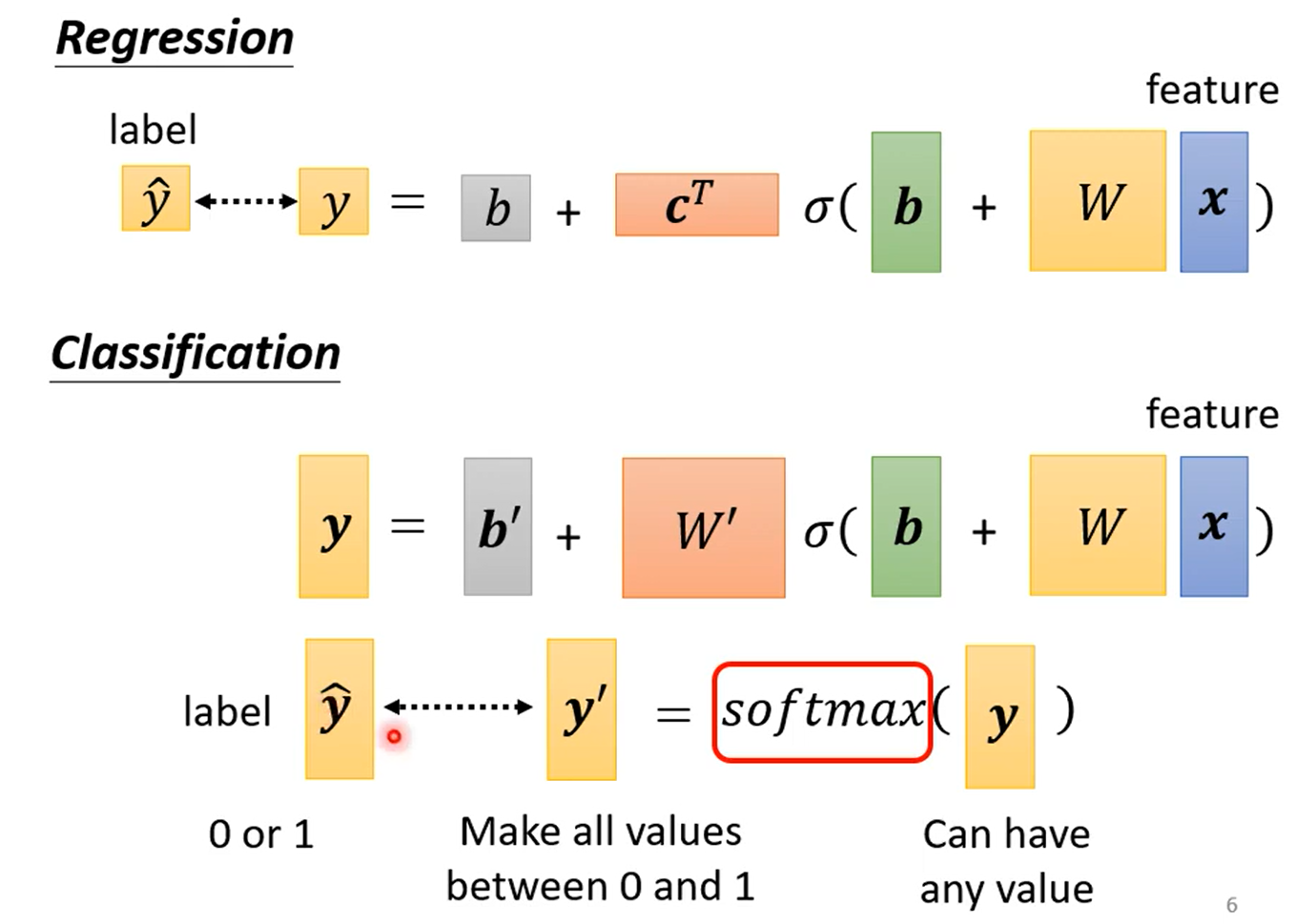
更长版本



## 1.把分类问题当做整数类别的回归问题，存在的问题是有时候分类问题并不是有序的



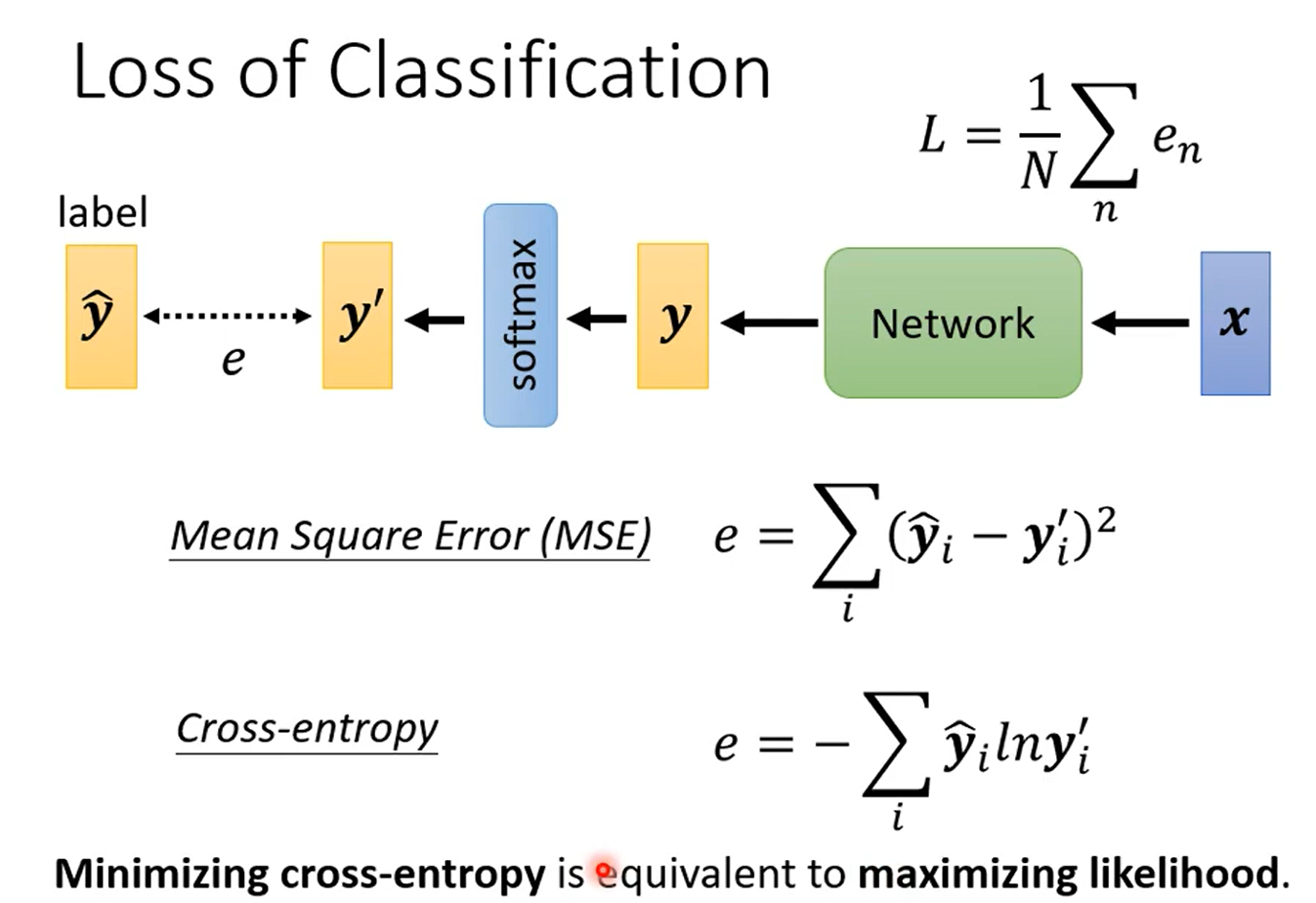
## 2.实际的方法（>2 classes）



## 3.实际的方法（classes=2）

一般不用softmax，而是sigmoid，但其实是一样的

## 4.分类的损失

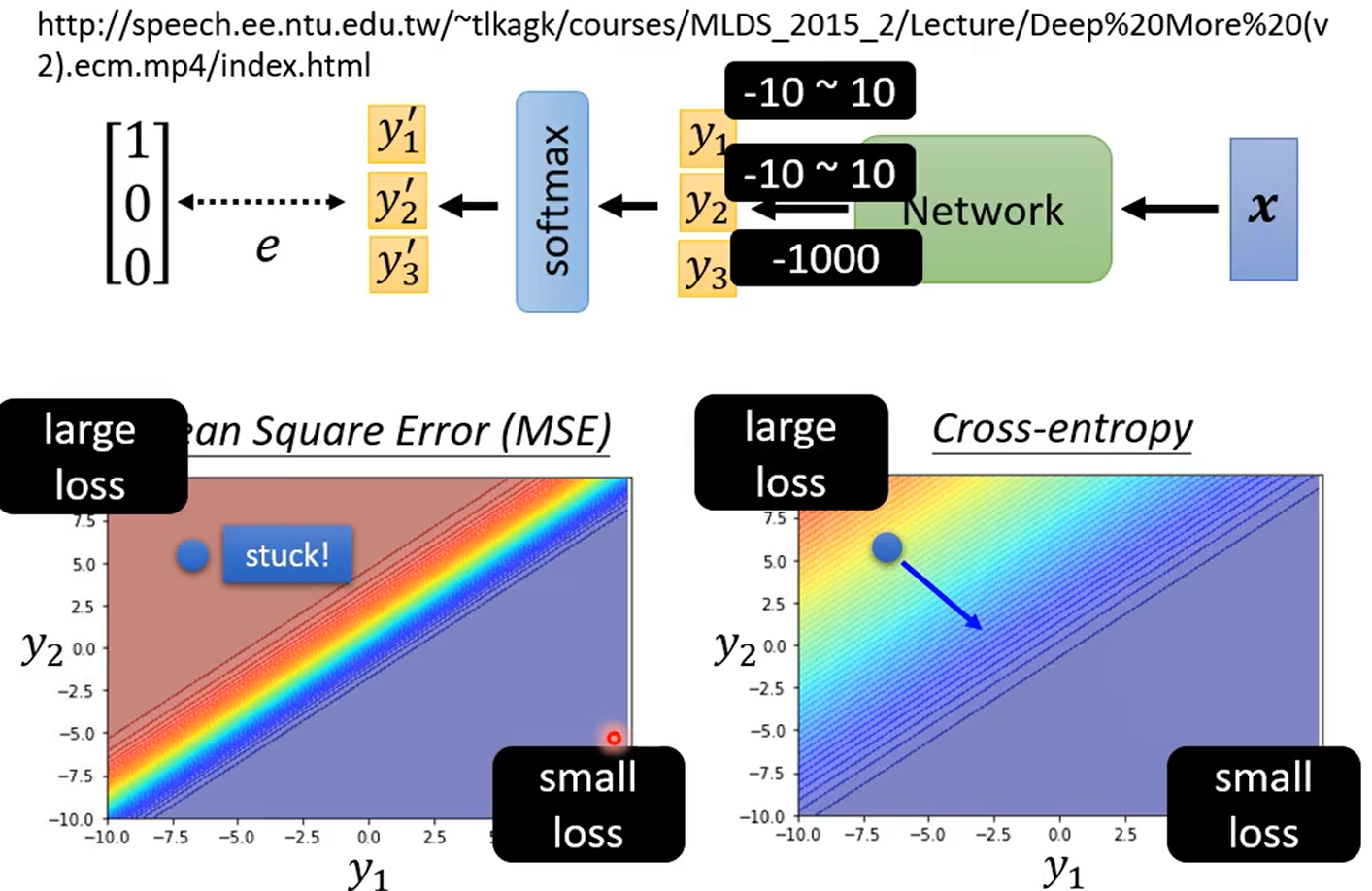


最小化交叉熵，就是最大化似然函数

其实在2019的视频中，说到了为什么MSE损失函数不能用作分类的损失函数

另外在pytorch的交叉熵损失函数，已经内置了softmax

## 5.为什么交叉熵比MSE更适合分类损失函数



MSE远离最优点的时候，损失函数很平坦