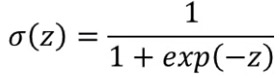
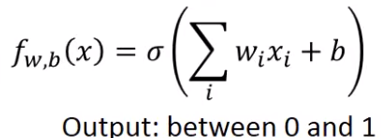
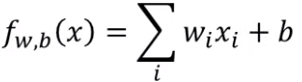
# 机器学习笔记

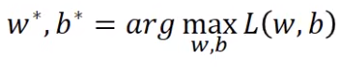
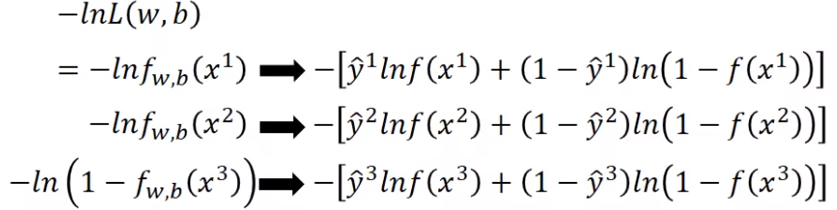
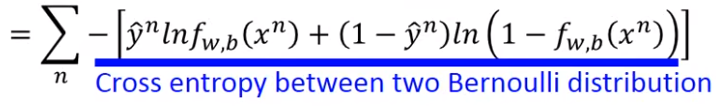
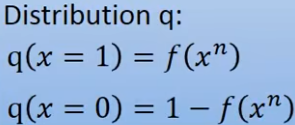
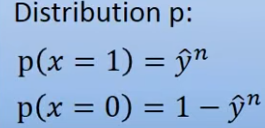
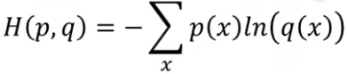
李宏毅教授机器学习(视频地址：https://www.bilibili.com/video/BV1JE411g7XF?p=3&vd\_source=0a27b8c0284e7cd3b72bc0c89478f2d7)

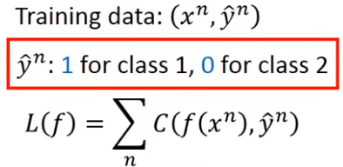
# 第三课：Logistic Regression

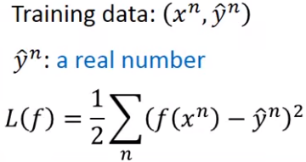
一、logistic regression与linear regression的比较

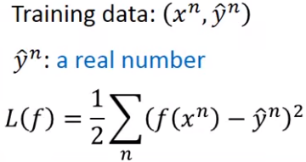
根据上一节课学习的内容我们可以知道，，我们现在将logistics regression和linear regression进行一个比较。

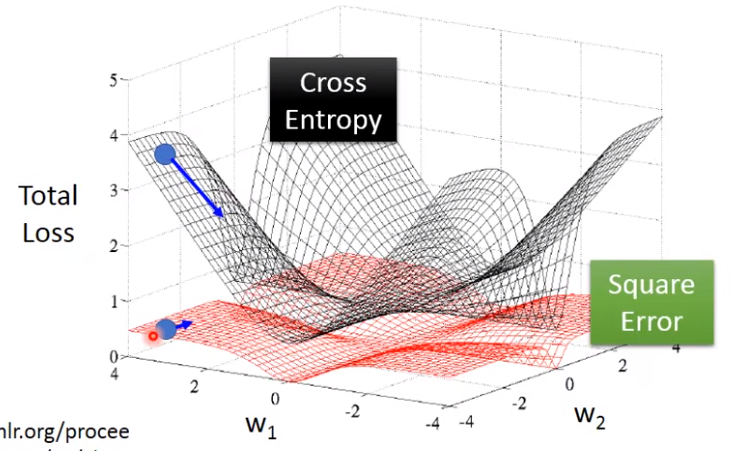
首先是第一步，function set的区别，logistic regression的function形式是，它的输出范围是0到1，而linear regression的function形式是，它的输出范围则可以是任何值。

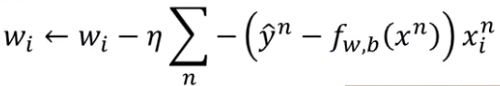
其次是第二步，Goodness of a Function的区别，之前我们需要找的是，我们现在给他进行一个对数变换，变成，就可以把乘法转化为加法，然后我们令，我们就可以把每一项写成的形式，然后我们的L(w,b)就等于，其中cross entropy代表了两个分布有多接近，假设两个分布分别是，那么他们之间的cross entropy表达式就是，这个表达式算出来的值越小代表两个分布越接近。

然后logistic regression的好坏程度就表示为，其中C是cross entropy，表达式为，

而linear regression的好坏程度则表示为。我们可以看出linear regression的loss function是用预估值与实际值差值的平方之和表示的，而logistic regression的loss function为什么不能也用平方和的形式表示呢？

如果我们的loss function的表达式是，我们将它对wi进行偏微分，得到式子是，现在在logistic regression中我们假设，当时，偏微分的值是0，这个是符合的；但是当时，偏微分的值也是0，但是这个却是不合理的。而cross entropy则不会有这样的问题。他们的图示如下图所示。

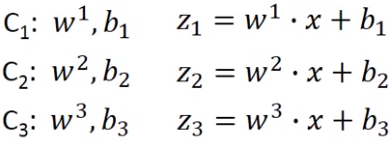


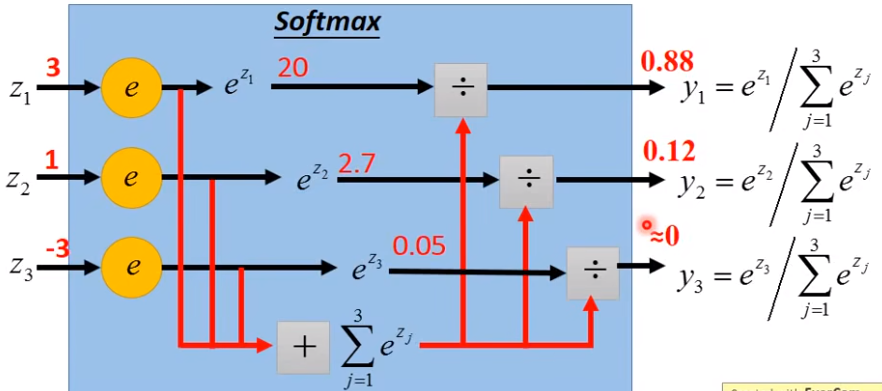
最后是第三步，Find the best function，数学部分就不具体描述了，我们只需要知道w的update过程是的。这个部分是表达式在logistic regression和linear regression中是一模一样的，唯一区别就在于的取值范围。

二、discriminative与generative的比较

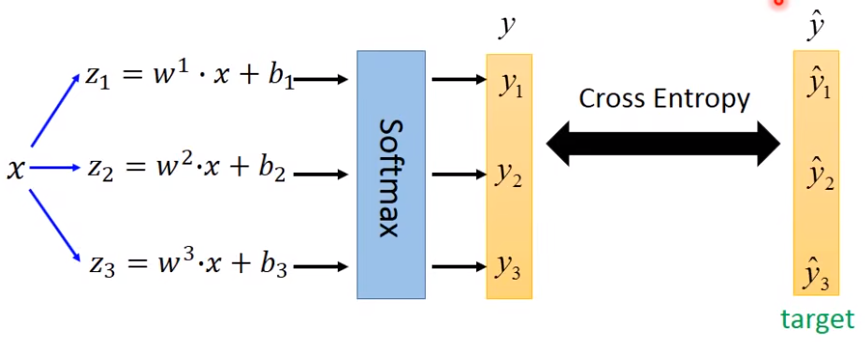
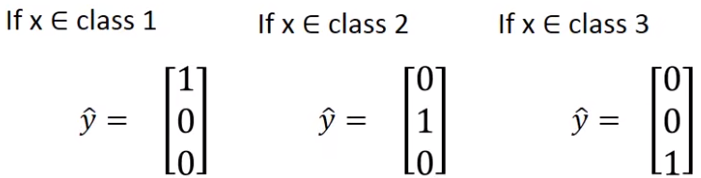
logistic regression的方法我们称之为discriminative的方法，而用gaussian来描述的方法我们称之为generative的方法。他们的model是一样的，都是，但是我们用两种方法找出来的w和b一般是不一样的。Generative中我们对model进行了一些假设，而discriminative中我们并没有对model进行任何的假设。多数情况下discriminative的model预测效果是比generative的model更好的，但是并不是任何情况下都更好，如果在data量很少的情况下，generative model的预测效果则会更好一些，因为generative model是已经进行过假设了的，它受data量的影响不是很大，而discriminative model没有做过任何假设，所以它受data量的影响是比较大的。

三、Multi-class Classification

对于多个class的分类来说，我们假设，我们将其进行一下指数变换，如图所示，

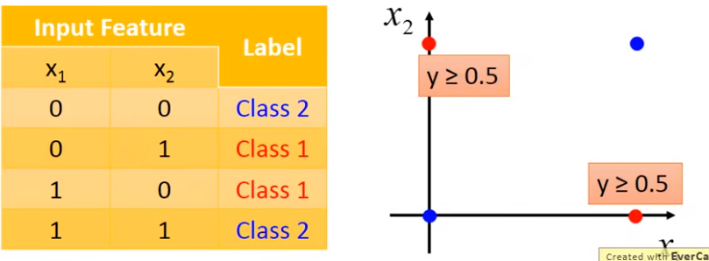


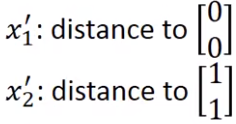
这个过程称为softmax，它的特点是每个yi的取值范围都是(0,1)，同时，对应的值也就是每个yi对应的概率，我们通过softmax之后，再做一个cross entropy，如下图所示。

，

四、logistic regression的限制

在logistic regression中，对于图示所给的输入，我们是无法对其进行分类的，因为logistic regression的类型的分界线是一条直线，而我们无法做到用一条直线使得红色点都在一边，蓝色点都在另一边。



但是我们可以通过feature transformation来解决这个问题，我们令，然后将feature转化一下，我们就可以将图进行转化，如图所示，我们将左边的图转化为了右边的图。

