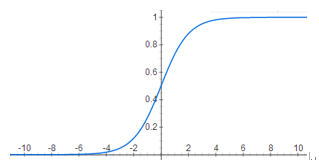
**逻辑斯蒂回归**

 在LR中，将线性回归的结果通过sigmod函数映射到0到1之间，映射的结果刚好可以看做是数据样本点属于某一类的概率，如果结果越接近0或者1，说明分类结果的可信度越高。对于线性不可分的数据，可以对非线性函数进行线性加权，得到一个不是超平面的分割面。

(1)sigmoid函数形式为：

C:\Users\jojo\AppData\Local\YNote\data\m15230512298@163.com\2cc6b65ead984f34ab1be16ab1bbcd60\clipboard.png



(2)二项逻辑斯蒂回归模型

条件概率分布P(Y|X)表示，随机变量X取值为实数，随机变量Y取值为1或0。模型的条件概率分布

 P(Y=1|X)= exp(w·x + b) / (1 + exp(w·x + b))

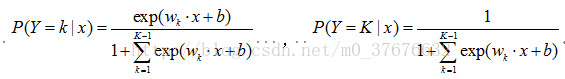
 P(Y=0|X)= 1 / (1 + exp(w·x + b))

 在进行分类时，通过比较上面两式的大小来将输入实例分配到概率值大的那一类。

    也就是说，LR将输入实例x的线性回归的加权结果w·x+b通过sigmoid函数映射到0~1之间，加权结果的值越接近正无穷，概率值就越接近1，反之则越接近0。

 多项LR回归模型

LR模型也可以推广到解决多分类问题，模型如下：



模型参数估计

 LR的参数估计使用最大似然估计。LR损失函数为对数损失函数，所以这里最大似然估计可以理解为损失函数为对数损失函数的经验风险最小化。

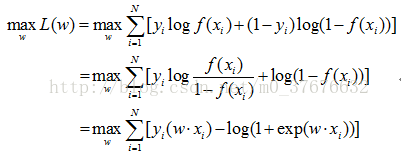
假设：

C:\Users\jojo\AppData\Local\YNote\data\m15230512298@163.com\04bbfd3f1f5443b5824675a5406c04e2\clipboard.png

那么似然函数为 :

C:\Users\jojo\AppData\Local\YNote\data\m15230512298@163.com\d6d0b267f1e54d86b9e44b5ca6e50d0c\clipboard.png

对上式取对数后得到最大值



 损失函数

LR损失函数为对数损失函数。对数损失函数的定义是

C:\Users\jojo\AppData\Local\YNote\data\m15230512298@163.com\8d04788d06314666b1f8c1b5a880da99\327093310020.png

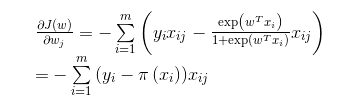
在参数模型估计中的对数似然函数极大化即对应着损失函数极小化。故代价函数定义为：

C:\Users\jojo\AppData\Local\YNote\data\m15230512298@163.com\8d30ee77afba4acc9bd314104d7e81b1\clipboard.png

梯度下降法求参数

逻辑斯蒂回归学习中通常采用的方法是梯度下降法及拟牛顿法，下面介绍逻辑斯蒂回归中梯度下降法求参数。

求偏导：



因此，在使用梯度下降法求解时，每次迭代地更新公式为：

C:\Users\jojo\AppData\Local\YNote\data\m15230512298@163.com\86b72d58b1a94c1a8df6b5c8e791aa93\clipboard.png