****Logistic回归****为概率型非线性回归模型，是研究二分类观察结果IMG_256与一些影响因素IMG_257之间关系的一种多

变量分析方法。通常的问题是，研究某些因素条件下某个结果是否发生，比如医学中根据病人的一些症状来判断它是

否患有某种病。

在讲解****Logistic回归****理论之前，我们先从LR分类器说起。LR分类器，即****Logistic Regression Classifier。****

在分类情形下，经过学习后的LR分类器是一组权值IMG_258，当测试样本的数据输入时，这组权值与测试数据按

照线性加和得到

IMG_259

这里IMG_260是每个样本的IMG_261个特征。

之后按照****sigmoid函数****的形式求出

IMG_262

由于****sigmoid函数****的定义域为IMG_263，值域为IMG_264，因此最基本的LR分类器适合对两类目标进行分类。

所以****Logistic回归****最关键的问题就是研究如何求得IMG_265这组权值。这个问题是用****极大似然估计****来做的。

下面正式地来讲Logistic回归模型。

考虑具有IMG_266个独立变量的向量IMG_267，设条件慨率IMG_268为根据观测量相对于某事件IMG_269发生的

概率。那么****Logistic回归****模型可以表示为

IMG_270

这里IMG_271称为****Logistic函数。****其中IMG_272

那么在IMG_273条件下IMG_274不发生的概率为

IMG_275

所以事件发生与不发生的概率之比为

IMG_276

这个比值称为事件的发生比（the odds of experiencing an event），[简记](https://www.baidu.com/s?wd=%E7%AE%80%E8%AE%B0&tn=24004469_oem_dg&rsv_dl=gh_pl_sl_csd" \t "https://blog.csdn.net/pql925/article/details/_blank)为odds。

对odds取对数得到

IMG_277

可以看出Logistic回归都是围绕一个Logistic函数来展开的。接下来就讲如何用****极大似然估计****求分类器的参数。

假设有IMG_278个观测样本，观测值分别为IMG_279，设IMG_280为给定条件下得到IMG_281的概率，同样地，

IMG_282的概率为IMG_283，所以得到一个观测值的概率为IMG_284。

因为各个观测样本之间相互独立，那么它们的联合分布为各边缘分布的乘积。得到似然函数为

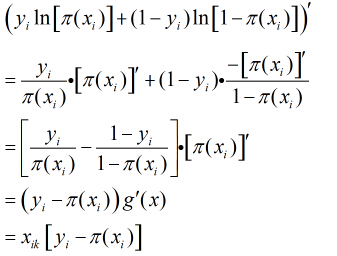
IMG_285

然后我们的目标是求出使这一似然函数的值最大的参数估计，最大似然估计就是求出参数IMG_286，使得IMG_287

取得最大值，对函数IMG_288取对数得到

IMG_289

继续对这IMG_290个IMG_291分别求偏导，得到IMG_292个方程，比如现在对参数IMG_293求偏导，由于



所以得到

IMG_295

这样的方程一共有IMG_296个，所以现在的问题转化为解这IMG_297个方程形成的方程组。

上述方程比较复杂，一般方法似乎不能解之，所以我们引用了****牛顿-**[拉菲](https://www.baidu.com/s?wd=%E6%8B%89%E8%8F%B2&tn=24004469_oem_dg&rsv_dl=gh_pl_sl_csd" \t "https://blog.csdn.net/pql925/article/details/_blank)**森迭代****方法求解。

利用牛顿迭代求多元函数的最值问题以后再讲。。。

简单牛顿迭代法：[http://zh.m.wikipedia.org/wiki/%E7%89%9B%E9%A1%BF%E6%B3%95](http://zh.m.wikipedia.org/wiki/%E7%89%9B%E9%A1%BF%E6%B3%95" \t "https://blog.csdn.net/pql925/article/details/_blank)

实际上在上述似然函数求最大值时，可以用梯度上升算法，一直迭代下去。梯度上升算法和牛顿迭代相比，收敛速度

慢，因为梯度上升算法是一阶收敛，而牛顿迭代属于二阶收敛。