逻辑回归(Logistic Regression)

2019年11月7日 21:57

作者：zx青

逻辑回归算法

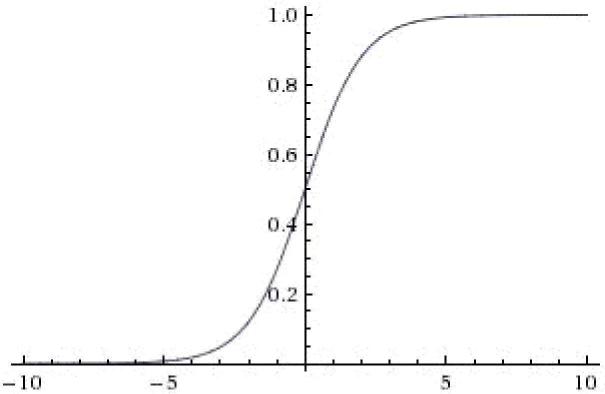
逻辑回归一般是用于解决分类问题，是将样本的特征和样本发生的概率联系起来，其中概率是一个数。
逻辑回归和线性回归同属于广义线性模型，逻辑回归就是用线性回归同属于广义线性模型，逻辑回归就是用线性回归的预测值去拟合真实标签的对数几率（一个事件的几率(odds)是指该事件发生的概率与不发生的概率之比，如果该事件发生的概率是p，那么该事件的几率是$\frac{p}{1-p}$，对数几率就是$log\frac{p}{1-p}$)
逻辑回归和线性回归本质上都是得到一条直线，不同的是，线性回归的直线是尽可能去拟合输入变量x的分布，使得训练集中所有样本点到直线的距离最短；而逻辑回归的直线是尽可能去拟合决策边界，使得训练集样本中的样本点尽可能分离开，两者的目的是不同的。

1、Sigmoid函数

Sigmoid函数公式：

$$g(x)=\frac{1}{1+e^{-(x)}}$$

当x趋近于无穷大时，g(x)趋近于1；当x趋近于无穷小时，g(x)趋近于0。



Sigmoid函数求导时有一个特性，这个特性将在下面的推导中用到，这个特性为：

$$\begin{aligned} g'(z) &= \frac{d}{dz}\frac{1}{1+e^{-z}} \\ &= \frac{1}{(1+e^{-z})^2}\left(e^{-z}\right) \\ &= \frac{1}{(1+e^{-z})}\cdot\left(1-\frac{1}{(1+e^{-z})}\right) \\ &= g(z)(1-g(z)). \end{aligned}$$

2、逻辑回归

逻辑回归本质上是线性回归，只是在特征到结果的映射中加入了一层函数映射，即先把特征线性求和，然后使用函数g(z)将最为假设函数来预测。g(z)可以将连续值映射到0到1之间。

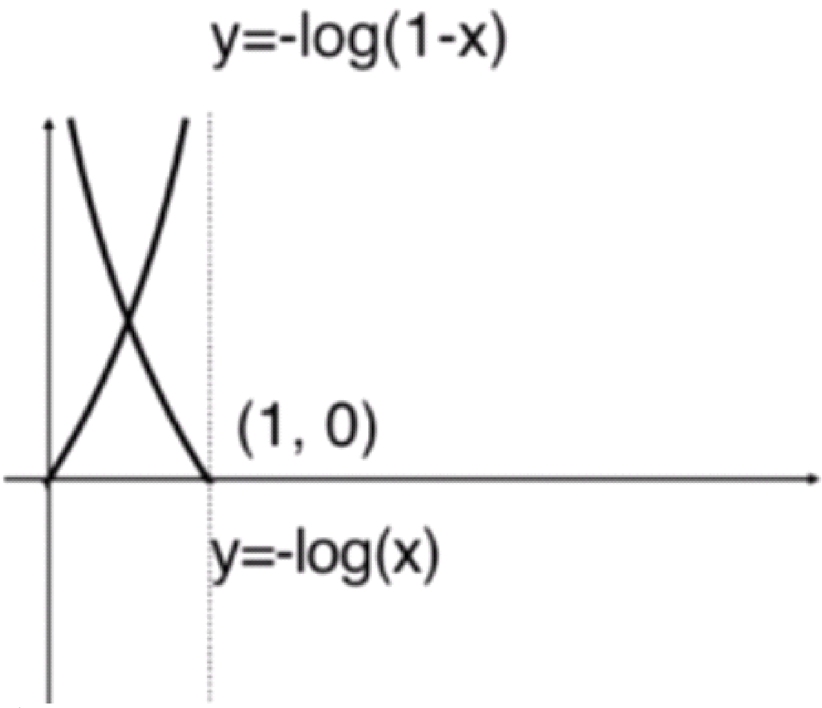
通过将线性模型和Sigmoid函数结合，就可以得到逻辑回归的公式：$h_\theta(x) = g(\theta^T x) = \frac{1}{1+e^{-(\theta^T x)}}$，$h_\theta(x)$的值域为$[0,1]$

对式子变形可得：$\log\frac{y}{1-y} = \theta^T x$（对数几率公式）

二项Logistic回归：

多项Logistic回归：

3、逻辑回归算法的损失函数



$$cost \begin{cases} 如果y = 1，h_\theta(x)越小，cost越大 \\ 如果y = 0，h_\theta(x)越大，cost越大 \end{cases}$$

$$cost \begin{cases} -\log(h_\theta(x)), & if\ \ y = 1 \\ -\log(1-h_\theta(x)), & if\ y = 0 \end{cases} \quad => cost = -ylog(h_\theta(x)) - (1-y)\log(1-h_\theta(x))$$

对于m个样本，总的损失函数为

$$J_{(\theta)} = -\frac{1}{m}\sum_{i=1}^{m} y_i log\left(h_\theta(x_i)\right) - (1-y_i)\log\left(1-h_\theta(x_i)\right)$$

m是样本数，y是标签，取值为0或1，i表示第i个样本，$h_\theta(x_i)$*表示预测的输出。*

4、逻辑回归算法损失函数求解

$$J_{(\theta)} = -\frac{1}{m}\sum_{i=1}^{m} y_i log\left(h_\theta(x_i)\right) - (1-y_i)\log\left(1-h_\theta(x_i)\right) \quad log\left(h_\theta(x_i)\right)' = 1 - h_\theta(x_i)$$

$$\nabla J(\theta)=\begin{bmatrix}\frac{\partial J_{(\theta)}}{\partial\theta_0}\\ \frac{\partial J_{(\theta)}}{\partial\theta_1}\\ \cdots\\ \frac{\partial J_{(\theta)}}{\partial\theta_n}\end{bmatrix}$$

$$J_{(\theta)}=-\frac{1}{m}\sum_{i=1}^{m}y_i log\left(g\left(\theta^T x_i\right)\right)-(1-y_i)\log\left(g\left(\theta^T x_i\right)\right)$$

$$(log\left(g\left(\theta^T x_i\right)\right))'=1-g\left(\theta^T x_i\right)$$

$$\frac{d(y_i log\left(g\left(\theta^T x_i\right)\right)}{d\theta_j}=y_i(1-g\left(\theta^T x_i\right))x_j$$

$$\frac{\Delta J(\theta)}{\Delta\theta_j}=\frac{1}{m}\sum_{i=1}^{m}\left(g\left(\theta^T x_i\right)-y_i\right)x_j=\frac{1}{m}\sum_{i=1}^{m}\left(\hat{y}_i-y_i\right)x_j=\frac{1}{m}X^T\left(g\left(\theta^T x_i\right)-y\right)$$

5、逻辑回归算法的python代码实现

https://github.com/zx-qing/Machine_Learning/blob/master/Logistic_Regression/logicRegression.ipynb