逻辑回归

模型表示

先回顾一下线性回归, 线性回归通过历史数据拟合出一条直线, 用这条直线对新的数据进行预测, 线性回归公式如下:

$$y = \theta \cdot x + b$$

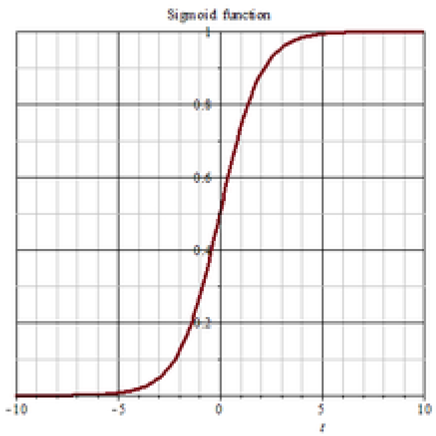
而对于逻辑回归来说, 其思想也是基于线性回归. 其公式如下:

$$h_\theta(x) = \frac{exp(\theta \cdot x)}{1 + exp(\theta \cdot x)} = \frac{1}{1 + exp(-\theta \cdot x)}$$

其中,

$$y = \frac{1}{1 + exp(-x)} = sigmoid(x) = \sigma(x)$$

被称为sigmoid函数.



sigmoid函数的值域为[0,1], 正好把线性输出变成概率, $h_\theta(x)$正好表示预测为positive的概率.

当$h_\theta x$>0.5时, 预测为positive

当$h_\theta x$<0.5时, 预测为nagtive

对于输出为positive和nagtive的概率分别为:

$$P(y=1|x;\theta) = h_\theta(x) = \sigma(\theta \cdot x)$$

$$P(y=0|x;\theta) = 1 - h_\theta(x) = 1 - \sigma(\theta \cdot x)$$

模型参数估计-极大似然估计

对于给定的训练数据集$T=(x_1,y_x),(x_2,y_2),...,(x_N,y_N)$, 可以应用极大似然估计法估计模型参数$\theta$, 从而得到逻辑回归.

设

$$P(Y=1|x)=h_\theta(x)=\pi(x)$$

$$P(Y=0|x)=1-\pi(x)$$

那么

$$\pi(x)=\frac{exp(\theta\cdot x)}{1+exp(\theta\cdot x)}$$

$$1-\pi(x)=\frac{1}{1+exp(\theta\cdot x)}$$

一个事件发生和不发生的概率的比值称为该时间的几率, 则预测为positive的对数几率为:

$$log(\frac{\pi(x)}{1-\pi(x)})=\theta\cdot x$$

似然函数为:

$$L(\theta)=\prod_{i=1}^{N}P(y^{(i)}|x^{(i)};\theta)$$

也就是:

$$L(\theta)=\prod_{i=1}^{N}[\pi(x_i)]^{y_i}[1-\pi(x_i)]^{1-y_i}$$

对数似然函数为:

$$l(\theta)=log(L(\theta))=\sum_{i=1}^{N}[y_i log(\pi(x_i))+(1-y_i)log(1-\pi(x_i))]$$

$$l(\theta)=\sum_{i=1}^{N}[y_i log(\frac{\pi(x_i)}{1-\pi(x_i)})+log(1-\pi(x_i))]$$

$$l(\theta) = \sum_{i=1}^{N}[y_i(\theta \cdot x) - log(1 + exp(\theta \cdot x_i)]$$

对$l(\theta)$求极大值, 可以转化为求$J(\theta)$:

$$J(\theta) = -\frac{1}{N}l(\theta)$$

的极小值。可以使用梯度下降法进行求解.