1 逻辑回归

逻辑回归模型属于对数线性模型. 二项逻辑回归模型是如下的条件概率分布:

$$P(Y=1|x)=\frac{\exp(wx+b)}{1+\exp(wx+b)}=\frac{1}{1+\exp(-(wx+b))}$$

$$P(Y=0|x)=\frac{1}{1+\exp(wx+b)}=\frac{\exp(-(wx+b))}{1+\exp(-(wx+b)}$$

可以简化成:

$$P(Y=1|x)=\frac{\exp(wx)}{1+\exp(wx)}=\frac{1}{1+\exp(-wx)}$$

$$P(Y=0|x)=\frac{1}{1+\exp(wx)}=\frac{\exp(-wx))}{1+\exp(-wx)}$$

1.1 对数几率

逻辑回归模型的特点: 一个时间的几率 (odds) 是指该事件发生的概率与该事件不发生的概率的比值. 如果事件发生的概率是 p, 那么该事件的几率是 $\frac{p}{1-p}$, 该事件的对数几率 (log odds) 或 logit 函数是

$$logit(p)=\log\frac{p}{1-p}$$

对逻辑回归而言:

$$\log\frac{P(Y=1|x)}{1-P(Y=1|x)}=wx$$

也就是说, 在逻辑回归模型中, 输出 $Y=1$ 的对数几率是输入 x 的线性函数, 或者说, 输出 $Y=1$ 的对数几率是由输入 x 的线性函数表示的模型. 线性函数的值越接近正无穷, 概率值就越接近 1; 线性函数的值越接近负无穷, 概率值就越接近 0.

1.2 模型参数估计

可以应用极大似然估计法估计模型参数.

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

似然函数为

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

对数似然函数为:

$$
\begin{aligned}
L(w) &= \sum_{i=1}^{N}[y_i \log \pi(x_i) + (1-y_i)\log(1-\pi(x_i))] \\
&= \sum_{i=1}^{N}[y_i \log \frac{\pi(x_i)}{1-\pi(x_i)} + \log(1-\pi(x_i))] \\
&= \sum_{i=1}^{N}[y_i(wx_i) - \log(1+\exp(wx_i))]
\end{aligned}
$$

对 $L(w)$ 求极大值, 得到 w 的估计值.
逻辑回归学习中通常采用的方法是梯度下降法及拟牛顿法.

1.3 多项逻辑回归模型

$$
P(Y=k|x) = \frac{\exp(w_k x)}{1+\sum_{k=1}^{K-1}\exp(w_k x)}, k = 1, 2, ..., K-1
$$

$$
P(Y=K|x) = \frac{1}{1+\sum_{k=1}^{K-1}\exp(w_k x)}
$$