LR与最大熵

LR回归

[toc]

问题引入

分类问题，比如iris分类，mnist分类

模型定义

$$ \begin{align} P(Y=1|x)=\frac{exp(wx*+b)}{1+exp(w*x)+b}\ P(Y=0|x)=\frac{1}{1+exp(w*x)+b} \end{align} $$

其中$x \in \backslash\mathrm{bold}R^n$ 是输入，$Y \in 0,1$ 是输出，$w \in \backslash\mathrm{bold}R^n$ 和$b \in \backslash\mathrm{bold}R$ 是参数 有时为了方便，将权值向量和输入向量加以扩充，记$w=(w^1,w^2,...,w^n,b)^T,x=(x^1,x^2,...,x^n,1)^T$,这时，LR模型如下:

$$ \begin{align} P(Y=1|x)=\frac{exp(wx*)}{1+exp(w*x)}\ P(Y=0|x)=\frac{1}{1+exp(w*x)} \end{align} $$

几率(odds)和对数几率(log odds)

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

对LR回归而言,

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

模型求解

最大似然估计求解 给定数据集$T=(x_1,y_1),(x_2,y_2),...,(x_n,y_n)$,其中$x_i \in \backslash\mathrm{bold}R^n$,$y_i \in 0,1$,应用**最大似然估计法**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

$$ \begin{equation} \begin{align} L(w)&=\sum_{i=1}^{N}\left[y_ilog\pi(x_i)+(1-y_i)log(1-\pi(x_i))\right]\ &=\sum_{i=1}^N\left[y_ilog\frac{\pi(x_i)}{1-\pi(x_i)}+log(1-\pi(x_i))\right]\ &=\sum_{i=1}^N\left[y_i(wx*_i)-log(1+exp(w*x_i))\right] \end{align} \end{equation} $$

代码

MAXENT

问题引入

模型定义

模型求解

代码

附录

迭代尺度法，坐标下降法，梯度下降法，共轭梯度法，牛顿法，拟牛顿法