逻辑回归

数据：$(X,Y),X\in R^n,Y\in\{0,1\}$ 服从伯努利分布

推导sigmoid函数：

odd:

$P(y_i=1|x_i)=p,P(y_i=0|x_i)=1-p,odd=\frac{p}{1-p}\in[0,+\infty)$

而特征的线性组合$WX\in(-\infty,+\infty)$,对odd求对数$logit(p)=log(odd)=log\frac{p}{1-p}\in(-\infty,+\infty)$

令$logit(p)=WX=Z,log\frac{p}{1-p}=Z,e^Z=\frac{p}{1-p},p=\frac{1}{1-e^{-Z}}$

即$p=simoid(WX)$

模型：$P_1=P(y_i=1|x_i)=\frac{1}{1+e^{-Wx_i}}$

$P_0=P(y_i=0|x_i)=\frac{e^{-Wx_i}}{1+e^{-Wx_i}}$

$P(y_i|x_i)=P_1^{y_i}P_0^{1-y_i}$

求解MLE：

$$\begin{aligned}\hat{w}&=\arg\max_w logP(Y|X)\\&=\arg\max_w log\prod P(y_i|x_i)\\&=\arg\max_w\sum logP(y_i|x_i)\\&=\arg\max_w\sum logP_1^{y_i}P_0^{1-y_i}\\&=\arg\max_w\sum(y_ilogP_1+(1-y_i)P_0)\end{aligned}$$

得到交叉熵损失：

$loss=-\sum(y_ilogP_1+(1-y_i)P_0)$

梯度下降法求解参数

$\frac{\partial L}{\partial W}=-\sum y_ix_i(1-\frac{1}{1+e^{-WX}})-(1-y_i)x_i\frac{1}{1+e^{-WX}}$

梯度下降跟新：

$w=w-\eta\frac{\partial L}{\partial W}$