

- [logistic回归损失函数](#)
 - [什么是损失](#)
 - [成本函数](#)

logistic回归损失函数

为了训练logistic回归函数的参数 w 以及 b ，需要定义一个成本函数.而回归损失函数，可以用于衡量算法的运行情况

什么是损失

从数学上讲，损失可以定义为 $L(\hat{y}, y) = (\hat{y} - y)^2$ 或者 $\frac{1}{2}(\hat{y} - y)^2$

但是在**logistic**回归中,通常并不会这么做,因为在之后的优化问题中,多数是非凸的,最后得到的结果是很多个局部最优解

误差平方看似合理,但是在梯度下降法中不太好用,所以在**logistic**回归中,我们会定义一个不同的损失函数,起着与误差平方相似的作用,可以为我们提供一个凸的优化问题,很容易去优化

$$L(\hat{y}, y) = -(y \ln \hat{y}) + (1 - y) \ln (1 - \hat{y})$$

对于这个**logistic**回归的损失函数,我们依然想让它尽可能地小

所以接下来会举例存在的两种情况

如果 $y = 1$,则会有 $L(\hat{y}, y) = -\ln \hat{y}$

也就是说 $-\ln \hat{y}$ 应当足够小,也就是 \hat{y} 应当足够大,但是由于**sigmoid**函数的作用, \hat{y} 不可能大于1,所以 \hat{y} 应当尽可能趋近于1

同理,如果 $y = 0$,则会有 $L(\hat{y}, y) = -\ln (1 - \hat{y})$

\hat{y} 就应当尽可能趋近于0

成本函数

损失函数是在单个训练样本中定义的,它衡量了在单个训练样本上的表现.但是如果要衡量在全体训练样本上的表现,则需要定义一个**成本函数**

$$J(w, b) = \frac{1}{m} \sum_{i=1}^m L(\hat{y}^{(i)}, y^{(i)}) \quad (\hat{y} \text{ 是利用 } w \text{ 和 } b \text{ 通过 } \textit{logistic} \text{ 回归算法得出的预测值})$$

所以我们需要找到合适的参数 w 和 b 使得 $J(w, b)$ 尽可能小