一文弄懂logistic回归为什么要用sigmoid函数

楼主有这样的疑问非常的正常,让我慢慢来解释你的问题。

一、第一个问题: 为什么要用 sigmoid 函数? 从两个角度解答

1.1 第一个角度

首先要明确的一点是:不是因为 sigmoid 有很多优秀的性质,然后在 logistic 回归模型建立的时候,从而使用 sigmoid 函数,这样是不合理的,因 为在数学当中 sigmoid 是一个函数族。具备 sigmoid 函数这样的性质的函数有很多。

在解释之前需要了解的知识: A:指数族分布 B: 广义线性模型

A: 指裁族分布

指数族分布满足下面的公式,即:

$$p(y;\eta) = b(y) \exp(\eta T(y) - a(\eta))$$

 η 为自然参数, $\Gamma(y)$ 为充分统计里,通常 $\Gamma(y) = y$, $\alpha(\eta)$ 为正则化项。

B:广义线性模型

满足下面三个假设的模型称为广义线性模型

- ① y | x;θ 满足一个以η 为参数的指数族分布
- ②给定x, 我们目标是预测y的期望值, 即 $h(x) = E(y \mid x)$
- $\Im \eta = \theta^T x$

首先在二分类中, 我们假设 ν 服从贝努利分布, 贝努利分布为指数族分布, 因为

$$p(y \mid \theta) = \phi^{T} (1 - \phi)^{T}$$

$$= \exp(y\log \phi + (1 - y)\log(1 - \phi))$$

$$= \exp(\log \frac{\phi}{1 - \phi} y + \log(1 - \phi))$$

我们令

$$\begin{cases} \eta = \log(\frac{\phi}{1-\phi}) \Rightarrow \phi = \frac{1}{1+e^{-\gamma}} \\ \alpha(\eta) = -\log(1-\phi) = \log(1+e^{\gamma}) \\ b(y) = 1 \end{cases}$$

所以可以得出贝努利分布属于指数族分布。 即贝努力分布满足广义线性模型的第一个假设,下面利用广义线性模型后面两个假设得 19645269

$$h_e(x) = E(y \mid x; \theta) = \phi = \frac{1}{1 + e^{-\tau}} = \frac{1}{1 + e^{-\theta' x}}$$

从而得到我们的预测目标即 sigmoid 函数,用来作为后验概率 p(y=1|x)

1.2 第二个角度

在解释第二个角度,我是假设楼主是比较了解贝叶斯判别。在贝叶斯判别中, 理论得出了机器学习模型的精度上界,即最大化后验概率。而 logistic 回归中 也是通过 signoid 函数来逼近后验概率 p(y=1|x)。通过贝叶斯公式得出

$$p(y=1|x) = \frac{p(x|y=1)p(y=1)}{p(x|y=0)p(y=0) + p(x|y=1)p(y=1)}$$

在这里我们假设先验服从贝努利分布,类条件概率服从高斯分布即

$$\begin{cases} p(y=1) = \phi \\ p(y=0) = 1 - \phi \\ p(x \mid y=1) = \frac{1}{(2pi)^{*/2} |\Sigma|^{1/2}} \exp(-(x - u_1)^T \Sigma^{-1} (x - u_1)) \end{cases}$$

将上式带入 p(y=1|x) en

$$\begin{split} p(y=1|x) &= \frac{p(x|y=1)p(y=1)}{p(x|y=0)p(y=0) + p(x|y=1)p(y=1)} \\ &= \frac{1}{1 + \frac{p(x|y=0)p(y=0)}{p(x|y=1)p(y=1)}} \\ &= \frac{1}{1 + \frac{1 - \phi}{\phi} \exp((u_0 - u_1)^T \Sigma^{-1} x + \frac{1}{2} (u_1^T \Sigma^{-1} u_1 - u_0^T \Sigma u_0))} \\ &= \frac{1}{1 + \exp((u_0 - u_1)^T \Sigma^{-1} x + \frac{1}{2} (u_1^T \Sigma^{-1} u_1 - u_0^T \Sigma u_0) + In(\frac{1 - \phi}{\phi}))} \end{split}$$

$$\diamondsuit$$
 , $-w_0 = \frac{1}{2} (u_1^T \Sigma^{-1} u_1 - u_0^T \Sigma u_0) + In(\frac{1-\phi}{\phi})$, $-w = \Sigma^{-1} (u_0 - u_1)$.的可以得出

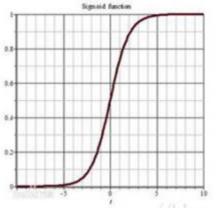
激活 Windows 转到"设置"以激活 Window

$$p(y=1|x) = \frac{1}{1+e^{-(w^2x+u_0)}}$$

不局限于类条件概率服从高斯分布,只要类条件概率服从指数族分布,都可以推出后验概率 为 sigmoid 函数的形式。从而可以看出 logistic 回归的强大的鲁榛性。

二、第二个问题: 为什么可以做成概率呢?

就是因为 sigmoid 函数这样优秀的性质,将输入变量的范围从 $(-\infty, +\infty)$,映射到 (0,1),而概率要求的正事范围在 (0,1) 之间,所以可以作为后验概率的形式。最后在配个 sigmoid 函数漂亮的曲线。



/blog.csdn.net/qq_19645269