

# 问题

逻辑回归相比于线性回归，有何异同

## 问题背景

逻辑回归与线性回归是机器学习中非常常用和基础的模型，面试当中也会经常被问到，而两者有着很多共同点，就连名字都差不多一样，所以会被问道两者的区别也是太正常不过了。我们都知道，回归是对连续值的预测，那么“逻辑回归”是回归问题吗？

我们首先聊聊**逻辑回归这个名字的由来**：对逻辑回归公式进行整理，我们可以得到  $\log \frac{p}{1-p} = \omega^T x + b$ ，其中  $p = P(y = 1|x)$  也就是将给定输入  $x$  预测为正类的概率。如果将一个事件的几率(odd)定义为该事件发生的概率与不发生的概率的比值  $\frac{p}{1-p}$ ，那么逻辑回归就可以看做是对  $y = 1|x$  这一事件的对数几率的线性回归，于是“逻辑回归”这个称谓就延续了下来。

## 问题解答

**两者最本质的区别**：逻辑回归处理的是分类问题，而线性回归处理的是回归问题。

**两者最大的区别**：逻辑回归将  $\hat{y}$  视为因变量而不是  $\frac{p}{1-p}$ ，而  $\hat{y}$  是类别标签，如二分类中  $\hat{y}$  可以是 0 或 1，是一个离散值，但线性回归中的因变量是连续值。

**两者相同之处**：

1. 都是使用**极大似然估计法来估计模型参数**：这里可能会有点疑问是说线性回归中明明是使用均方误差即最小二乘法来估计的，不是极大似然估计呀。其实实际上在预测值与真实值之间的误差满足正态分布的假设下，最小化均方误差与极大似然估计在本质上是相同的。
2. 二者在求解超参数的过程中，**都可以使用梯度下降的方法**，这也是监督学习中一个常见的相似之处

以上内容参考自《百面机器学习》的逻辑回归一节

## 拓展

这里的拓展主要是线性回归与逻辑回归的模型定义与求解过程。

### 线性回归

线性回归试图学得

$$f(x_i) = \omega x_i + b, \text{ 使得 } f(x_i) \approx y_i \quad (1)$$

其中  $y_i$  为真实值，而  $f(x_i)$  为预测值

那么怎么去确定参数  $\omega$  和  $b$  呢，显然关键在于怎么去衡量  $f(x_i)$  与  $y_i$  之间的距离，而在回归任务中一般使用均方误差，因此我们的目标变为让均方误差最小化：

$$(\omega^*, b) = \operatorname{argmin}_{(\omega, b)} \sum_{i=1}^m (f(x_i) - y_i)^2 \quad (2)$$

接下来的模型参数求解可以对  $\omega$  与  $b$  进行求导或者使用梯度下降法。

## 逻辑回归

在线性回归的基础上，逻辑回归可以理解为我们只需找到一个单调可微函数将线性模型的预测值与把该输入归类为正类的概率  $p = P(y = 1|x)$  联系起来即可，即将线性回归得到的连续值映射为二分类的类别0或1，因此我们自然而然就想到了sigmoid函数，它的函数形式如下：

$$y = \frac{1}{1 + e^{-z}} \quad (3)$$

我们令  $z = \omega^T x + b$ ，令  $y$  为把该输入  $x$  归类为为正类的概率  $p$ ，便可以得到：

$$p = \frac{1}{1 + e^{-(\omega^T x + b)}}$$

对以上公式稍做变换就可以得到

$$\begin{aligned} p &= P(y = 1|x) = \frac{e^{\omega^T x + b}}{1 + e^{\omega^T x + b}} \\ 1 - p &= P(y = 0|x) = \frac{1}{1 + e^{\omega^T x + b}} \\ \log \frac{p}{1 - p} &= \omega^T x + b \end{aligned} \quad (4)$$

下面使用极大似然估计来估计模型参数，



以上内容参考周志华的西瓜书《机器学习》和李航的《统计学习方法》

By Yee

2020.05.08