

# 逻辑回归

## 场景描述

逻辑回归可以说是机器学习领域最基础也是最常用的模型；

## 问题 1：逻辑回归相比于线性回归，有何异同

- 逻辑回归处理分类问题，线性回归处理回归问题；
- 逻辑回归中，因变量取值是一个二元分布，模型学习得出的是  $E[y|x;\theta]$ ，即给定自变量和超参数后，得到因变量的期望，并基于此期望来处理预测分类问题。而线性回归中实际上求解的是  $y' = \theta^T x$ ，是对真实值的一个近似。实际上将逻辑回归式子整理得到：

$$\log \frac{p}{1-p} = \theta^T x,$$

- 在关于逻辑回归的讨论中，我们均认为  $y$  是因变量，而非  $\frac{p}{1-p}$ ，这便引出了逻辑回归

于线性回归最大的区别，即逻辑回归中因变量为**离散值**，而线性回归中因变量是**连续的**；

- 相同点：
  - 二者都可以使用极大似然估计来对训练样本进行建模。线性回归使用最小二乘法，实际上就是在自变量  $x$  与超参数  $\theta$  确定，因变量  $y$  服从正太分布的假设下，使用极大似然估计的一个化简；而逻辑回归中通过对似然函数

$$L(\theta) = \prod_{i=1}^n p(y_i | x_i; \theta) = \prod_{i=1}^n (\pi(x_i))^{y_i} (1-\pi(x_i))^{1-y_i}$$

的学习，得到最佳参数  $\theta$ ；

- 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。

问题 2：当使用逻辑回归处理多标签的分类问题时，有哪些常见的做法，分别应用于哪些场景，它们之间又有怎样的关系？

$$h_{\theta}(x) = \begin{bmatrix} p(y=1|x;\theta) \\ p(y=2|x;\theta) \\ \vdots \\ p(y=k|x;\theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^k e^{\theta_j^T x}} \begin{bmatrix} e^{\theta_1^T x} \\ e^{\theta_2^T x} \\ \vdots \\ e^{\theta_k^T x} \end{bmatrix};$$