

# 20220527-机器学习

---

## 1.学习内容

### 1.1 机器学习

周志华《机器学习》

线性模型

## 2.结果描述

## 1.学习内容

### 1.1 机器学习

周志华《机器学习》

线性模型

$$f(\mathbf{x}) = w_1x_1 + w_2x_2 + \dots + w_dx_d + b, \quad (3.1)$$

一般用向量形式写成

$$f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b, \quad (3.2)$$

$$\begin{aligned} (w^*, b^*) &= \arg \min_{(w, b)} \sum_{i=1}^m (f(x_i) - y_i)^2 \\ &= \arg \min_{(w, b)} \sum_{i=1}^m (y_i - wx_i - b)^2. \end{aligned}$$

$$\begin{aligned} \frac{\partial E_{(w, b)}}{\partial w} &= 2 \left( w \sum_{i=1}^m x_i^2 - \sum_{i=1}^m (y_i - b) x_i \right), \\ \frac{\partial E_{(w, b)}}{\partial b} &= 2 \left( mb - \sum_{i=1}^m (y_i - wx_i) \right), \end{aligned}$$

$$w = \frac{\sum_{i=1}^m y_i (x_i - \bar{x})}{\sum_{i=1}^m x_i^2 - \frac{1}{m} \left( \sum_{i=1}^m x_i \right)^2} ,$$

$$b = \frac{1}{m} \sum_{i=1}^m (y_i - wx_i) ,$$

更一般的形式：

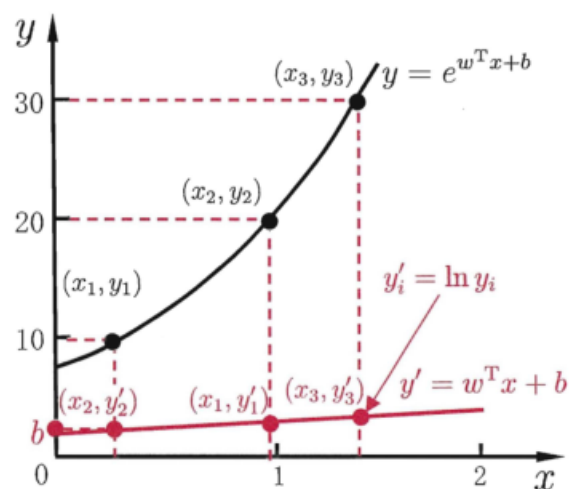
$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1d} & 1 \\ x_{21} & x_{22} & \dots & x_{2d} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{md} & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{x}_1^T & 1 \\ \mathbf{x}_2^T & 1 \\ \vdots & \vdots \\ \mathbf{x}_m^T & 1 \end{pmatrix}$$

$$\hat{\mathbf{w}}^* = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} ,$$

对数线性回归：

$$y = \mathbf{w}^T \mathbf{x} + b .$$

$$\ln y = \mathbf{w}^T \mathbf{x} + b .$$



更一般地, 考虑单调可微函数  $g(\cdot)$ , 令

$$y = g^{-1}(w^T x + b), \quad (3.15)$$

这样得到的模型称为“广义线性模型”(generalized linear model), 其中函数  $g(\cdot)$  称为“联系函数”(link function). 显然, 对数线性回归是广义线性模型在  $g(\cdot) = \ln(\cdot)$  时的特例.

对数线性回归:

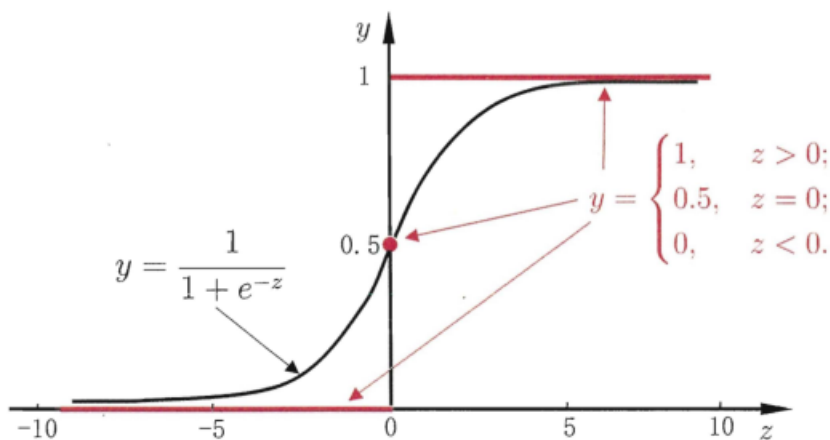


图 3.2 单位阶跃函数与对数几率函数

$$y = \frac{1}{1 + e^{-z}}.$$

$$\ln \frac{y}{1-y} = \mathbf{w}^T \mathbf{x} + b .$$

若将  $y$  视为样本  $\mathbf{x}$  作为正例的可能性, 则  $1-y$  是其反例可能性, 两者的比值

$$\frac{y}{1-y} \tag{3.20}$$

称为“几率”(odds), 反映了  $\mathbf{x}$  作为正例的相对可能性. 对几率取对数则得到“对数几率”(log odds, 亦称 logit)

$$\ln \frac{y}{1-y} . \tag{3.21}$$

## 2.结果描述

临近离校感觉没啥心思学习, 只草草看了一点线性回归的内容。深刻地认识到, 一切学科的尽头都是数学。感觉这本书一个不太友好地地方便是, 一上来就亮公式, 也不告诉你来龙去脉, 对初学者确实不友好。当然, 要兼顾广度和深度本身也不是一件易事, 对一本概述型的教材不能寄予太多期望。之后等返家了再接着学。