

线性回归

Herbert002

2021 年 11 月 28 日

摘要

尽量通俗易懂，且不学院气质的线性回归笔记。

目录

| | | |
|---|-----------|---|
| 1 | 线性回归的通俗理解 | 1 |
| 2 | 线性回归的简要历史 | 1 |
| 3 | 一元线性回归 | 1 |
| 4 | 多元线性回归 | 1 |

1 线性回归的通俗理解

线性回归的通俗理解.

2 线性回归的简要历史

高尔顿爵士关于身高体重的实验.

3 一元线性回归

我们先考虑一元线性回归.

4 多元线性回归

多元线性回归是指回归变量个数大于等于 2 的线性回归.

回归平方和服从卡方分布的证明

$$\begin{aligned}
SSR &= \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \\
&= [\hat{\mathbf{y}} - \mathbf{1}\bar{y}]^T [\hat{\mathbf{y}} - \mathbf{1}\bar{y}] \\
&= [\mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} - \mathbf{1}(\mathbf{1}^T \mathbf{1})^{-1} \mathbf{1}^T \mathbf{y}]^T [\mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} - \mathbf{1}(\mathbf{1}^T \mathbf{1})^{-1} \mathbf{1}^T \mathbf{y}] \\
&= \mathbf{y}^T [\mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T - \mathbf{1}(\mathbf{1}^T \mathbf{1})^{-1} \mathbf{1}^T]^T [\mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} - \mathbf{1}(\mathbf{1}^T \mathbf{1})^{-1} \mathbf{1}^T \mathbf{y}]
\end{aligned} \tag{1}$$

上式中, $\mathbf{1} \in \mathbb{R}^n$, 其元素均为 1. 响应变量均值可以表示为 $\bar{y} = (\mathbf{1}^T \mathbf{1})^{-1} \mathbf{1}^T \mathbf{y}$.

接下来证明 $[\mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} - \mathbf{1}(\mathbf{1}^T \mathbf{1})^{-1} \mathbf{1}^T \mathbf{y}]$ 是实对称矩阵且是幂等的.

$$[\mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} - \mathbf{1}(\mathbf{1}^T \mathbf{1})^{-1} \mathbf{1}^T \mathbf{y}] \tag{2}$$