

高级计量经济学

理论经济学博士课程 2023-2024

Lecture 3: The Linear Model

Davidson, R. & MacKinnon, J. (2009). *Econometrics Theory and Methods*. Oxford University Press.

黄嘉平

工学博士 经济学博士
深圳大学中国经济特区研究中心 讲师

办公室 粤海校区汇文楼1510
E-mail huangjp@szu.edu.cn
Website <https://huangjp.com>

线性回归模型

数据矩阵 (横截面或时间序列数据)

y x_2 x_3 x_i x_k

↓ ↓ ↓ ↓ ↓ ↓

观测值 1 →

⋮

观测值 t →

⋮

观测值 n →

country	WAGES	PRICES	GDP	EMPLOY	MONEY1	MONEY2	UNEMPLOY
Ausstralia	4.41	3.75	3.04	1.68	9.1	10.98	8.68
Austria	4.15	2.71	2.55	0.65	5.37	7.37	5.48
Belgium	3.99	2.37	2.16	0.34	5.87	-999999	8.49
Canada	3.76	2.83	2.03	1.17	6.13	8.51	9.51
Denmark	3.78	2.61	2.02	0.02	3.21	4.51	7.68
Finland	5.65	3.11	1.78	-1.06	5.97	-999999	10.44
France	3.55	2.4	2.08	0.28	5.19	4.05	10.84
Germany	4.08	2.78	2.71	0.08	9.08	7.59	6.86
Greece	14.18	13.09	2.08	0.87	14.46	14.14	8.84
Iceland	-999999	8.42	1.54	-0.13	10.73	13.64	3.07
Ireland	4.5	2.51	6.4	2.16	-999999	-999999	13.81
Italy	6.16	4.86	1.68	-0.3	4.75	4.17	10.46
Japan	2.28	1.47	2.81	1.06	5.56	4.49	2.65
Korea	12.83	6.13	7.73	2.57	-999999	17.59	2.4
Luxembourg	-999999	2.53	5.64	3.02	-999999	7.22	2.39
Mexico	-999999	30.59	2.67	4.56	-999999	32.53	3.71
Netherlands	2.59	2.22	2.86	1.88	6.31	6.61	6.41
Norway	4.31	3.12	2.98	0.36	5.98	4.66	5.07

线性回归模型

Linear Regression Model

- 回归模型是用自变量 (independent/explanatory variable) 解释因变量 (dependent variable) 的模型的总称
- 多元线性回归模型 (multiple linear regression model) 定义为

$$y_t = \beta_1 + \beta_2 x_{t2} + \cdots + \beta_k x_{tk} + u_t, \quad t = 1, \dots, n$$

其中, u_t 称为**误差项 (error term)**, 代表自变量以外所有对因变量产生影响的因素

- 如果我们假设 $E(u_t | x_{t2}, \dots, x_{tk}) = 0$, 则

$$E(y_t | x_{t2}, \dots, x_{tk}) = \beta_1 + \beta_2 x_{t2} + \cdots + \beta_k x_{tk}$$

此式称为**回归函数 (regression function)**。回归函数是研究者对变量间关系的一种假设, 因此也称为**设定 (specification)**。如果回归函数正确地描述了现实, 则称之为**正确设定 (correctly specified)**, 否则称之为**错误设定 (misspecified)**

- 在正确设定回归函数的前提下, 回归分析的目的是利用样本数据估计总体中的参数 β_0, \dots, β_k 的值。为此, 我们需要确定一种估计方法。当给定样本和估计方法时, 如果能够获得唯一的估计值, 则称该参数是**可识别的 (identifiable)**

线性模型的矩阵表达

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} \quad \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

线性回归模型 $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$

我们规定 \mathbf{X}_t 代表矩阵 \mathbf{X} 的第 t 行,
 x_i 代表矩阵 \mathbf{X} 的第 i 列。

$$\begin{aligned} y_1 &= \beta_1 x_{11} + \beta_2 x_{12} + \cdots + \beta_k x_{1k} + u_1 & y_1 &= \mathbf{X}_1 \boldsymbol{\beta} + u_1 \\ \Leftrightarrow y_2 &= \beta_1 x_{21} + \beta_2 x_{22} + \cdots + \beta_k x_{2k} + u_2 & \Leftrightarrow y_2 &= \mathbf{X}_2 \boldsymbol{\beta} + u_2 \\ &\vdots & &\vdots \\ y_n &= \beta_1 x_{n1} + \beta_2 x_{n2} + \cdots + \beta_k x_{nk} + u_n & y_n &= \mathbf{X}_n \boldsymbol{\beta} + u_n \end{aligned}$$

一般情况下我们设 \mathbf{X} 的第一列中的要素都为 1, 则 β_1 为截距。

系数的估计

矩估计

Method-of-moments Estimation

k 阶理论矩: $E[X^k]$ k 阶样本矩: $M_k = \frac{1}{n} \sum_{i=1}^n X_i^k$

矩估计就是考虑理论矩满足的方程 (组), 然后用样本矩替代理论矩, 以此求出参数的估计值。

例如, 总体均值为—阶矩 $\mu = E[X]$, 则其估计值为 $\hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}$ 。

针对 $y = X\beta + u$, 我们假设 $E[X_t^\top u_t] = \mathbf{0}$ for all t (此处共有 k 个条件), 则

$$E[x_{ti}u_t] = 0 \text{ for } i = 1, \dots, k$$

$$\begin{aligned} \mathbf{0} &= \frac{1}{n} \sum_{t=1}^n X_t^\top u_t = \frac{1}{n} \sum_{t=1}^n X_t^\top (y_t - X_t \beta) && \Rightarrow X^\top y = X^\top X \beta \\ &= \frac{1}{n} \sum_{t=1}^n X_t^\top y_t - \frac{1}{n} \sum_{t=1}^n X_t^\top X_t \beta && \Rightarrow \hat{\beta} = (X^\top X)^{-1} X^\top y \end{aligned}$$

MM估计量

最小二乘估计

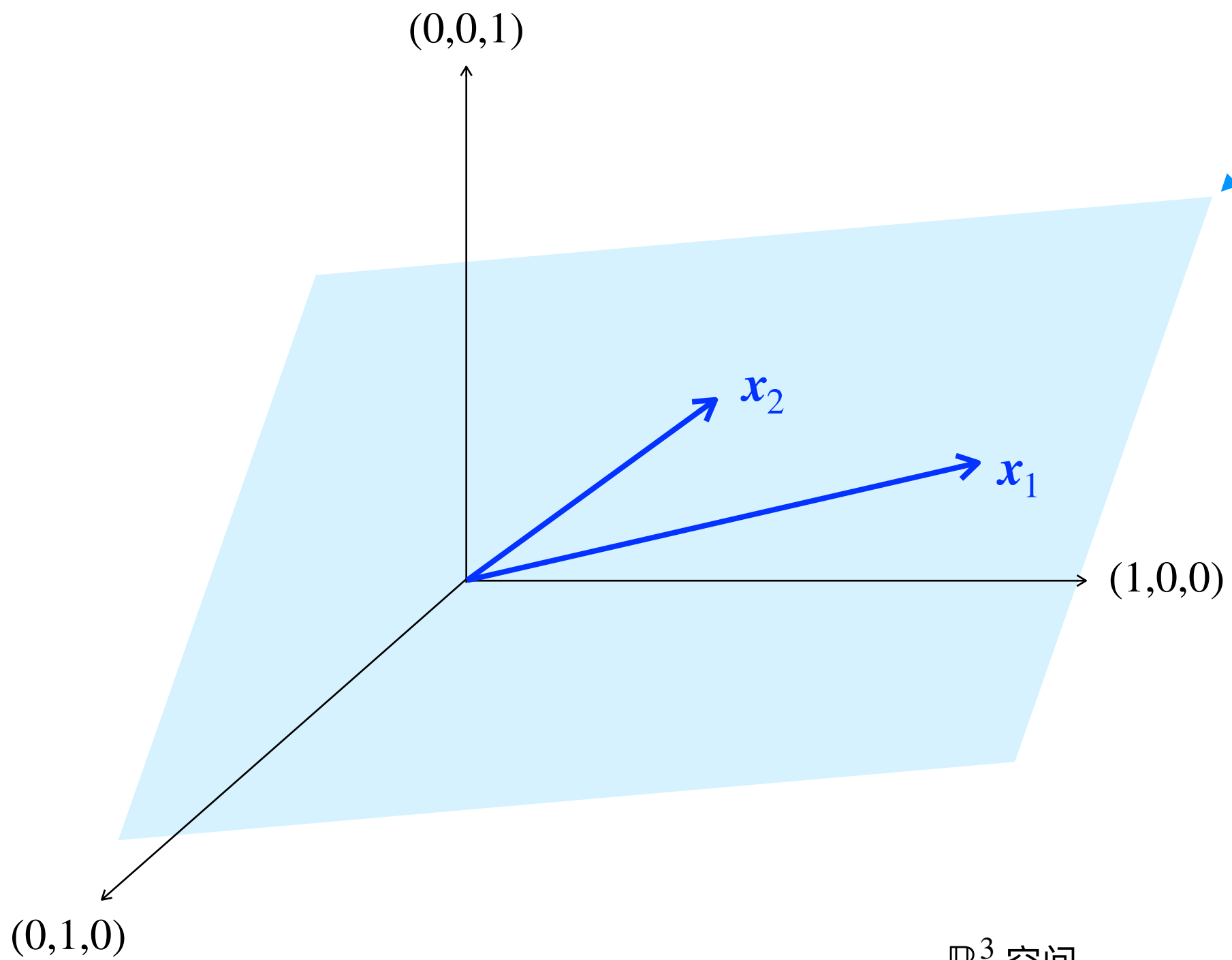
Ordinary Least Squares Estimation

针对任意 β ，我们称 $y - X\beta$ 为残差 (residuals)。残差平方和 (sum of squared residuals) 定义为

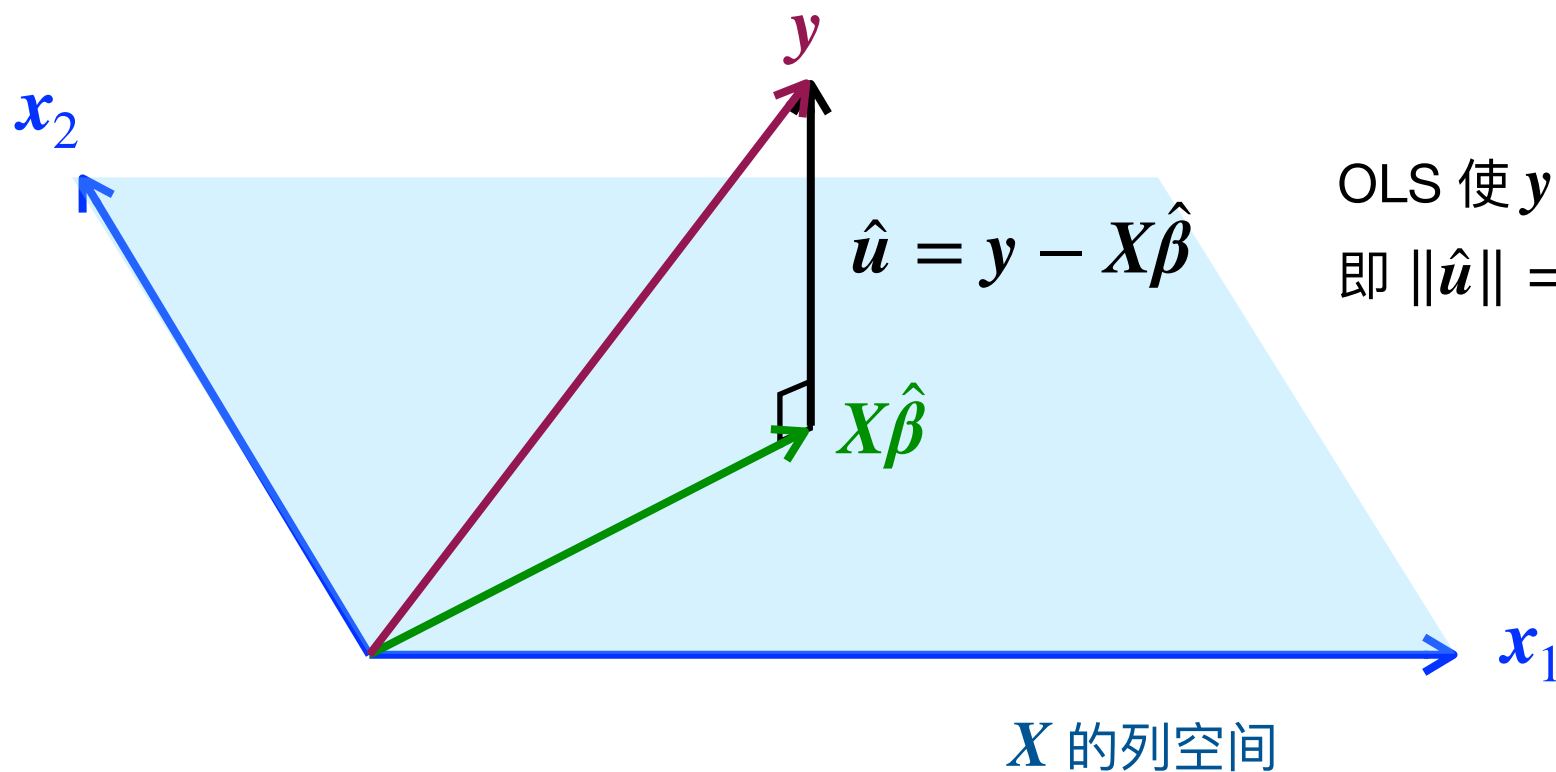
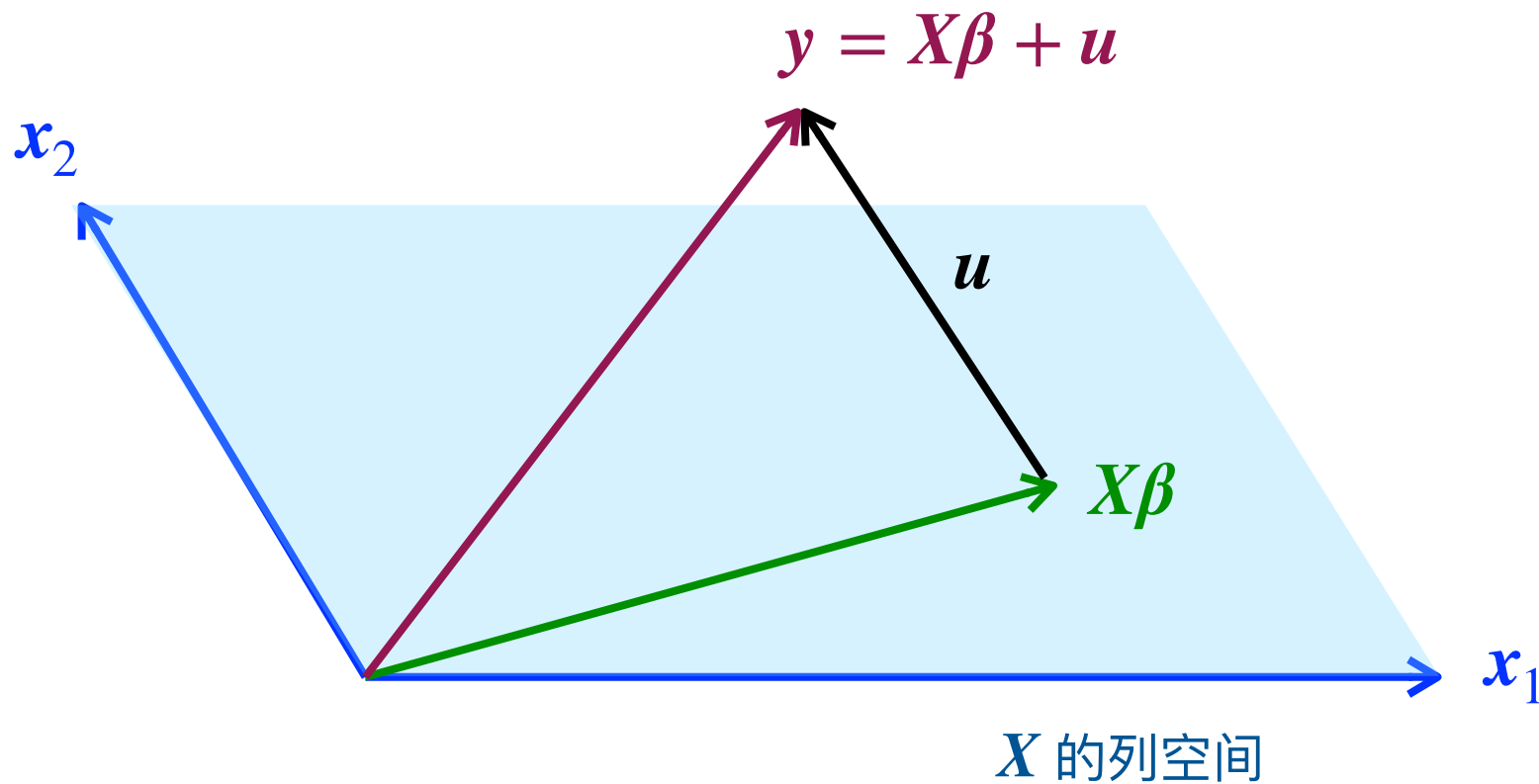
$$SSR(\beta) = \sum_{t=1}^n (y_t - X_t\beta)^2 = (y - X\beta)^\top (y - X\beta)$$

OLS 的目的是找到使 $SSR(\beta)$ 取值最小的 $\hat{\beta}$ ，即

$$\hat{\beta} = \arg \min_{\beta \in \mathbb{R}^k} SSR(\beta)$$



由 $ax_1 + bx_2$ 定义的平面
平面上的任意一点都可以写成
 $\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = Xa$ 的形式，
因此也称作 X 的列空间
(column space)。



OLS 使 y 和 $X\hat{\beta}$ 间的距离最小，
 即 $\|\hat{u}\| = \sqrt{\hat{u}^\top \hat{u}} \leq \sqrt{u^\top u} = \|u\|$

OLS 估计量

$$\begin{aligned}
 SSR(\boldsymbol{\beta}) &= (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^\top (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \\
 &= \mathbf{y}^\top \mathbf{y} - \mathbf{y}^\top \mathbf{X}\boldsymbol{\beta} - \boldsymbol{\beta}^\top \mathbf{X}^\top \mathbf{y} + \boldsymbol{\beta}^\top \mathbf{X}^\top \mathbf{X}\boldsymbol{\beta} \\
 &= \mathbf{y}^\top \mathbf{y} - 2\mathbf{y}^\top \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}^\top \mathbf{X}^\top \mathbf{X}\boldsymbol{\beta}
 \end{aligned}$$

若 A 为对称矩阵, 则 $\frac{\partial \mathbf{x}^\top A \mathbf{x}}{\partial \mathbf{x}} = 2A\mathbf{x}$ 。 $\mathbf{X}^\top \mathbf{X}$ 为对称矩阵。

一阶条件为

$$\frac{\partial SSR(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \begin{bmatrix} \frac{\partial SSR(\boldsymbol{\beta})}{\partial \beta_1} \\ \frac{\partial SSR(\boldsymbol{\beta})}{\partial \beta_2} \\ \vdots \\ \frac{\partial SSR(\boldsymbol{\beta})}{\partial \beta_k} \end{bmatrix} = -2\mathbf{X}^\top \mathbf{y} + 2\mathbf{X}^\top \mathbf{X}\boldsymbol{\beta} = \mathbf{0}$$

$$\Rightarrow \mathbf{X}^\top \mathbf{y} = \mathbf{X}^\top \mathbf{X}\hat{\boldsymbol{\beta}}$$

$$\Rightarrow \hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}$$

OLS估计量 = MM估计量

$$\text{若 } \mathbf{a}^\top \mathbf{x} = \sum_{i=1}^n a_i x_i, \text{ 则 } \frac{\partial \mathbf{a}^\top \mathbf{x}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial \mathbf{a}^\top \mathbf{x}}{\partial x_1} \\ \vdots \\ \frac{\partial \mathbf{a}^\top \mathbf{x}}{\partial x_n} \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \mathbf{a}。$$

其他必要条件

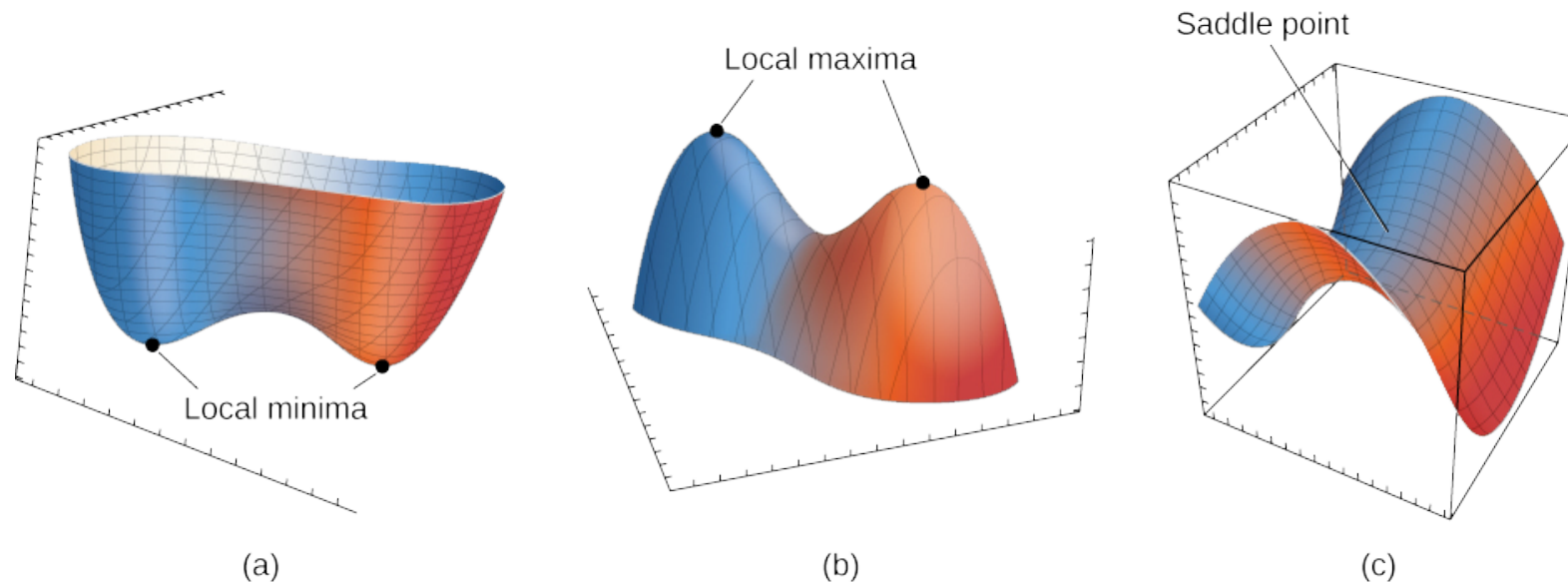
- OLS 解的存在条件

为保证 $\hat{\beta} = (X^T X)^{-1} X^T y$ 存在，我们需要保证 $(X^T X)^{-1}$ 存在，即 $\text{rank}(X^T X) = k$ (full rank 满秩)。该条件可以通过假设 $\text{rank}(X) = k$ 满足。

OLS假设之一，若在 $n > k$ 时 $\text{rank}(X) < k$ ，就会出现共线性问题。

- OLS 解真的使 SSR 最小化

一般情况下一阶条件不能保证最优解：



图片地址 https://math.libretexts.org/Courses/University_of_California_Davis/UCD_Mat_21C:_Multivariate_Calculus/13:_Partial_Derivatives/13.7:_Extreme_Values_and_Saddle_Points

我们需要讨论二阶条件，即 $SSR(\beta)$ 是否为凸函数（海赛矩阵为正定矩阵）

二阶条件

Hessian Matrix

$$\frac{\partial^2 \text{SSR}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} = \begin{bmatrix} \frac{\partial^2 \text{SSR}(\boldsymbol{\beta})}{\partial \beta_1^2} & \frac{\partial^2 \text{SSR}(\boldsymbol{\beta})}{\partial \beta_1 \partial \beta_2} & \cdots & \frac{\partial^2 \text{SSR}(\boldsymbol{\beta})}{\partial \beta_1 \partial \beta_k} \\ \frac{\partial^2 \text{SSR}(\boldsymbol{\beta})}{\partial \beta_2 \partial \beta_1} & \frac{\partial^2 \text{SSR}(\boldsymbol{\beta})}{\partial \beta_2^2} & \cdots & \frac{\partial^2 \text{SSR}(\boldsymbol{\beta})}{\partial \beta_2 \partial \beta_k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \text{SSR}(\boldsymbol{\beta})}{\partial \beta_k \partial \beta_1} & \frac{\partial^2 \text{SSR}(\boldsymbol{\beta})}{\partial \beta_k \partial \beta_2} & \cdots & \frac{\partial^2 \text{SSR}(\boldsymbol{\beta})}{\partial \beta_k^2} \end{bmatrix} = 2\mathbf{X}^\top \mathbf{X} \text{ 是正定矩阵 (positive definite)}$$

\mathbf{A} 是正定矩阵 $\Leftrightarrow \mathbf{x}^\top \mathbf{A} \mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$

如果 $n \times k$ 矩阵 \mathbf{A} 满足 $n > k$, $\text{rank}(\mathbf{A}) = k$, 则 $\mathbf{A}^\top \mathbf{A}$ 是正定矩阵, $\mathbf{A} \mathbf{A}^\top$ 是半正定矩阵 (positive semidefinite)

证明:

因为 $\text{rank}(\mathbf{A}) = k$, $\mathbf{z} = \mathbf{A} \mathbf{x} \neq \mathbf{0}$, 因此

$$\mathbf{x}^\top (\mathbf{A}^\top \mathbf{A}) \mathbf{x} = (\mathbf{A} \mathbf{x})^\top (\mathbf{A} \mathbf{x}) = \mathbf{z}^\top \mathbf{z} > 0$$

因为 $n > k$, 且 $\text{rank}(\mathbf{A}^\top) = \text{rank}(\mathbf{A}) = k$, 所以 $\mathbf{s} = \mathbf{A}^\top \mathbf{x}$ 可以为 $\mathbf{0}$ 。因此

$$\mathbf{x}^\top (\mathbf{A} \mathbf{A}^\top) \mathbf{x} = (\mathbf{A}^\top \mathbf{x})^\top (\mathbf{A}^\top \mathbf{x}) = \mathbf{s}^\top \mathbf{s} \geq 0$$

□

$\Rightarrow 2\mathbf{X}^\top \mathbf{X}$ 是正定矩阵, 满足二阶条件

一个特例的推导

$$y_t = \beta_1 x_{t1} + \beta_2 x_{t2} + u_t$$

$$\text{求解 } \min_{\beta_1, \beta_2} \sum [y_i - (\beta_1 x_{t1} - \beta_2 x_{t2})]^2$$

一阶条件为：

$$\frac{\partial S}{\partial \beta_1} = -2 \sum x_{t1} y_i + 2\beta_1 \sum x_{t1}^2 + 2\beta_2 \sum x_{t1} x_{t2} = 0$$

$$\frac{\partial S}{\partial \beta_2} = -2 \sum x_{t2} y_i + 2\beta_1 \sum x_{t1} x_{t2} + 2\beta_2 \sum x_{t2}^2 = 0$$

解一阶条件，得

$$\hat{\beta}_1 = \frac{(\sum x_{t1} y_i)(\sum x_{t2}^2) - (\sum x_{t2} y_i)(\sum x_{t1} x_{t2})}{(\sum x_{t1}^2)(\sum x_{t2}^2) - (\sum x_{t1} x_{t2})^2}$$

$$\hat{\beta}_2 = \frac{(\sum x_{t2} y_i)(\sum x_{t1}^2) - (\sum x_{t1} y_i)(\sum x_{t1} x_{t2})}{(\sum x_{t1}^2)(\sum x_{t2}^2) - (\sum x_{t1} x_{t2})^2}$$

课后练习：

1. 确认左面的结果和公式

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

得到的结果一致。

2. 确认二阶条件成立。

y 的分解: $TSS = ESS + SSR$

将 OLS 估计量 $\hat{\beta}$ 代入回归模型, 可得

$$y = X\hat{\beta} + \hat{u}, \quad \hat{u} = y - X\hat{\beta} \text{ 是 OLS 残差}$$

此时,

$$\begin{aligned} y^T y &= (X\hat{\beta} + \hat{u})^T (X\hat{\beta} + \hat{u}) \\ &= (X\hat{\beta})^T X\hat{\beta} + \hat{u}^T X\hat{\beta} + (X\hat{\beta})^T \hat{u} + \hat{u}^T \hat{u} \\ &= \hat{\beta}^T X^T X\hat{\beta} + 2\hat{u}^T X\hat{\beta} + \hat{u}^T \hat{u} \end{aligned}$$

已知 $\hat{u}^T X = (y - X\hat{\beta})^T X = y^T X - \hat{\beta}^T X^T X = \mathbf{0}^T$ 由 $\hat{\beta}$ 表达式可得

因此可得出

$$y^T y = \hat{\beta}^T X^T X\hat{\beta} + \hat{u}^T \hat{u}$$

↑

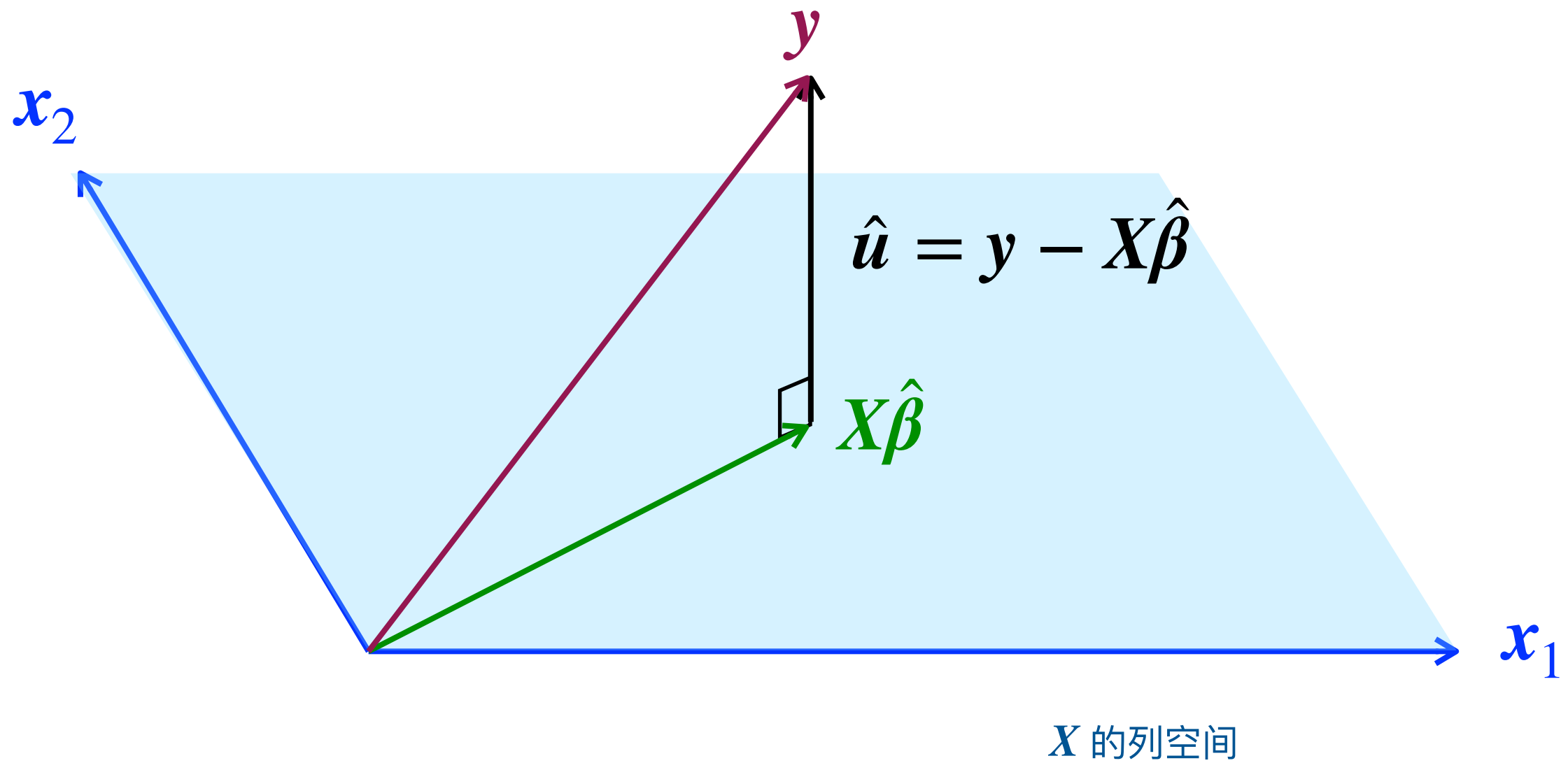
Explained Sum of Squares

Total Sum of Squares

Sum of Squared Residuals

$$\mathbf{y}^\top \mathbf{y} = \hat{\boldsymbol{\beta}}^\top \mathbf{X}^\top \mathbf{X} \hat{\boldsymbol{\beta}} + \hat{\mathbf{u}}^\top \hat{\mathbf{u}}$$

$$\Leftrightarrow \|\mathbf{y}\|^2 = \|\mathbf{X}\hat{\boldsymbol{\beta}}\|^2 + \|\hat{\mathbf{u}}\|^2 \quad (\text{勾股定理})$$



正交向量与 OLS

Orthogonal Vectors and OLS

若向量 a 与 b 满足 $a \cdot b = a^\top b = b^\top a = 0$ (即内积为零), 则称 a 与 b 正交, 写为 $a \perp b$ 。

OLS 估计中, 残差 \hat{u} 与 $X\hat{\beta}$ 正交, 因此可得

$$(X\hat{\beta})^\top \hat{u} = 0$$

$$\Leftrightarrow (X\hat{\beta})^\top (y - X\hat{\beta}) = 0$$

$$\Leftrightarrow \hat{\beta}^\top (X^\top y - X^\top X\hat{\beta}) = 0$$

假设 $\hat{\beta} \neq \mathbf{0}$, 则 $X^\top y = X^\top X\hat{\beta} \Rightarrow \hat{\beta} = (X^\top X)^{-1} X^\top y$

