一元简单线性回归

王向东*

May 2024

1 简单回归模型的定义

1.1 初始表达

简单回归模型是指只有一个自变量的回归模型，它的数学表达式为：

$$y = \beta_0 + \beta_1 x + u \tag{1}$$

其中，y 是因变量，x 是自变量，β_0 是截距项，β_1 是斜率项，u 是误差项。

1.2 模型假定

让我们先做一个简单回归模型的初始假设：

- 误差项 u 期望值为零；即
$$E(u) = 0 \tag{2}$$
- 自变量 x 与误差项 u 完全独立；即
$$p(x, u) = p_1(x)p_2(u) \tag{3}$$

根据 3式，我们可以得到：

$$E(u|x) = E(u) = 0 \tag{4}$$

这是一个关键的假设，4式又被称为**零均值假设**，它意味着自变量 x 的取值不会影响误差项 u 的期望值。也就是在进行回归分析前，需要考虑一下我们的自变量 x 的取值，会不会对不关注（忽视）的变量 u 的期望值产生影响。

另外，我们可以利用**零均值假设**得到一个很有用的推论：

$$E(y|x) = \beta_0 + \beta_1 E(x|x) + E(u|x) = \beta_0 + \beta_1 x \tag{5}$$

5式表明，给定 x 的条件下，**总体线性函数**是一个关于 x 的线性函数。即，给定 x 的条件下，y 的分布都以 $E(y|x)$ 为中心。从而我们可以将 1式看作两个部分：一个是关于 x 的线性函数，$E(y|x)$, 它被称作 y 的系统部分；误差项 u 是非系统部分，多元线性回归的章节我们将讨论这两部分的大小。

*本文档不得用于商用或贩卖。如发现错误或共同参与此项目，请联系 wangxd86@mail2.sysu.edu.cn。本讲义整理自中山大学商学院《计量经济学》课程（李广众、詹钥淞）、对外经济贸易大学 mooc《计量经济学导论》（陈志鸿、唐丹、潘红宇）、伍德里奇《计量经济学导论》。

2 最小二乘的推导

这里我们用矩估计的方法来推导 β_0 和 β_1 的估计值。这种方法不需要推导最小化目标函数的一阶条件，但两种估计方法得出的结果是等价的。根据 2式，我们可以得到：

$$cov(x,u) = E(xu) - E(x)E(u) = E(xu) = 0 \tag{6}$$

用可观测数据替代式 2、6我们可以得到：

$$E(y - \beta_0 - \beta_1 x) = 0 \tag{7}$$

$$E(x(y - \beta_0 - \beta_1 x)) = 0 \tag{8}$$

上面的式子表明，对于总体 x, y 的联合密度函数的两个限制。刚好我们需要估计的参数也是两个，运用矩估计法，我们可以得到：

$$n^{-1}\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \tag{9}$$

$$n^{-1}\sum_{i=1}^{n}x_i(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \tag{10}$$

9可以改写为 x 和 y 均值的关系：

$$\bar{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x} \tag{11}$$

进一步我们解出：

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \tag{12}$$

将 12式代入 10式，我们可以得到：

$$\sum_{i=1}^{n}x_i(y_i - \bar{y} + \hat{\beta}_1 \bar{x} - \hat{\beta}_1 x_i) = 0$$

整理得到：

$$\sum_{i=1}^{n}x_i(y_i - \bar{y}) = \hat{\beta}_1 \sum_{i=1}^{n}x_i(x_i - \bar{x})$$

根据简单的代数变换，我们可以得到 $\hat{\beta}_1$：

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{13}$$

就此，我们完成了 β_0 和 β_1 的估计。

我们可以对 β_1 做一些进一步变换：

$$\hat{\beta}_1 = \hat{\rho}_{xy}\frac{\hat{\sigma}_y}{\hat{\sigma}_x} \tag{14}$$

以上结论单位基于样本数据，不出意外，我们可以在总体中得出相同的结论。可以看出，在非实验数据的情况下，简单回归只是两个变量之间的相关性分析，在推导因果关系时要注意这一点。

下面介绍一些新的变量定义：

- 拟合值：$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$
- 残差：$\hat{u}_i = y_i - \hat{y}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$

- 残差平方和：$SSR = \sum_{i=1}^{n} \hat{u}_i^2$

需要指出的是，我们务必要区分残差 $\hat{u}$ 和误差 u 的定义，残差是我们估计值与实际值之间的差异，而误差是我们无法观测到的。到这里并未结束，因为这些结果并不能揭示充足的因果关系。计量经济学中或许更加关注 OLS 统计学性质，比如无偏性、一致性、有效性等。在讨论它们之前让我们先讨论一下 OLS 对任意样本的性质。

3 OLS 对任一样本数据的性质

3.1 OLS 统计量的代数性质

以下性质推导过程略去，但请牢记，后面的证明过程将大量重复使用这些性质。

代数性质：$\sum_{i=1}^{n}(x_i - \bar{x}) = 0$、$\sum_{i=1}^{n}(x_i - \bar{x})x_i = \sum_{i=1}^{n}(x_i - \bar{x})^2$、$\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n}(x_i - \bar{x})y_i = \sum_{i=1}^{n}(y_i - \bar{x})x_i$。

统计量性质：

$$\sum_{i=1}^{n} \hat{u}_i = 0 \tag{15}$$

$$\sum_{i=1}^{n} x_i \hat{u}_i = 0 \Leftrightarrow cov(x, \hat{u}) = 0 \tag{16}$$

$$\sum_{i=1}^{n} \hat{y}_i \hat{u}_i = 0 \Leftrightarrow cov(\hat{y}, \hat{u}) = 0 \tag{17}$$

引入一些新的估计量：

- $SST = \sum_{i=1}^{n}(y_i - \bar{y})^2$
- $SSE = \sum_{i=1}^{n} \hat{u}_i^2$
- $SSR = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$

3.2 OLS 拟合优度

我们不难得出：

$$SST = SSE + SSR \tag{18}$$

证明过程如下：

$$
\begin{aligned}
SST &= \sum_{i=1}^{n}(y_i-\bar{y})^2 \\
&= \sum_{i=1}^{n}[(y_i-\hat{y}_i)+(\hat{y}_i-\bar{y})]^2 \\
&= \sum_{i=1}^{n}[\hat{u}_i+(\hat{y}_i-\bar{y})]^2 \\
&= \sum_{i=1}^{n}\hat{u}_i^2+\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2+2\sum_{i=1}^{n}\hat{u}_i(\hat{y}_i-\bar{y}) \\
&= SSE+SSR+2\sum_{i=1}^{n}\hat{u}_i(\hat{y}_i-\bar{y})
\end{aligned}
$$

很容易证明 $\sum_{i=1}^{n}\hat{u}_i(\hat{y}_i-\bar{y})=0$，因此 18式成立。证明如下：

$$
\begin{aligned}
\sum_{i=1}^{n}\hat{u}_i(\hat{y}_i-\bar{y}) &= \sum_{i=1}^{n}\hat{u}_i\hat{y}_i-\bar{y}\sum_{i=1}^{n}\hat{u}_i \\
&= 0
\end{aligned}
$$

于是拟合优度 R^2 的定义为：

$$
R^2=\frac{SSR}{SST}=1-\frac{SSE}{SST} \tag{19}
$$

R^2 的取值范围为 $[0,1]$。需要指出的是，我们可以将式 1简单 OLS 回归模型看作是一个常数回归加一个过原点回归的组合，过原点回归的 R^2 值是有可能为负数的，此时如果加上一个常数项 β_0，β_1 的估计值将会等于 0，R^2 值就会变为 0。可见，R^2 是一个相对常数回归的评价指标。

必须需要强调的是，R^2 值并非衡量 OLS 估计结果优劣的金标准。R^2 大并不代表 OLS 结果成立，R^2 小也不代表 OLS 估计结果没用。

4 度量单位和函数形式

改变自变量或因变量的单位，会改变 OLS 估计值，这之间的转换相当直接，我们略过。下面我们对变量进行函数变换。它对 OLS 估计值的影响如下：

- 对变量取自然对数：取微分后可得，取对数变量的增长率与其他变量的关系。
- 对变量取平方或其他函数转化：可得到非线性关系，但是 OLS 估计值仍然是线性的。

可见简单一元线性回归中的线性是指参数的线性，而不是变量的线性。

5 OLS 估计量的统计性质

或许你已经知道 OLS 的良好性质，它是一个 BLUE（best linear unbiased estimator）。本章节将会证明这一点。这对将来回归系数的显著性检验和置信区间的构建是非常重要的。

在正式开始证明之前，我们先引入一个新的概念：**高斯-马尔科夫假定**。这是 OLS 估计量的一个重要前提，它包括以下几个假定：

- **SLR.1 线性于参数**：模型是一个线性于参数 β_0、β_1 的模型。

- **SLR.2 随机性**：样本从整体中随机抽取。
- **SLR.3 波动性**：因变量的样本结果不是常数。这是一条很弱的假设，但很有必要。
- **SLR.4 零条件均值**：$E(u|x) = 0$，它在最初的假设中已经提到。
- **SLR.5 同方差性**：$Var(u|x) = \sigma^2$，即误差项的方差不随 x 而变化，也就是不会因为抽样的改变而改变。

为了防止困惑，**接下来的估计量的统计性质都是基于已知 x，推导得出的**。你也可以把这一点视作下面的统计性质都是基于条件的，如条件期望、条件方差。简言之，就是将初始模型 (1) 中的 β_0、β_1、x 看作常数项，将 u、y 看作随机变量。

还需要对 $\hat{\beta}_1$ 做一些代数变换：

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})y_i}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})y_i}{SST_x} = \frac{\sum_{i=1}^n (x_i - \bar{x})(\beta_0 + \beta_1 x_i + u_i)}{SST_x}$$

$$\sum_{i=1}^n (x_i - \bar{x})(\beta_0 + \beta_1 x_i + u_i) = \beta_0 \sum_{i=1}^n (x_i - \bar{x}) + \beta_1 \sum_{i=1}^n (x_i - \bar{x})x_i + \sum_{i=1}^n (x_i - \bar{x})u_i$$

经过代数变化进一步化简：

$$\begin{aligned}
&\beta_0 \sum_{i=1}^n (x_i - \bar{x}) + \beta_1 \sum_{i=1}^n (x_i - \bar{x})x_i + \sum_{i=1}^n (x_i - \bar{x})u_i \\
&= 0 + \beta_1 \sum_{i=1}^n (x_i - \bar{x})^2 + \sum_{i=1}^n (x_i - \bar{x})u_i \\
&= \beta_1 SST_x + \sum_{i=1}^n (x_i - \bar{x})u_i
\end{aligned}$$

代回 $\hat{\beta}_1$ 的表达式：

$$\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^n (x_i - \bar{x})u_i}{SST_x} \quad = \beta_1 + \frac{\sum_{i=1}^n d_i u_i}{SST_x} \tag{20}$$

5.1 期望与无偏性

该性质的证明仅需要用到 SLR.1-4 假设。证明如下：

$$\begin{aligned}
E(\hat{\beta}_1) &= E(\beta_1) + E\left(\frac{\sum_{i=1}^n (x_i - \bar{x})u_i}{SST_x}\right) \\
&= \beta_1 + \frac{1}{SST_x} E\left(\sum_{i=1}^n (x_i - \bar{x})u_i\right) \\
&= \beta_1 + \frac{1}{SST_x} \sum_{i=1}^n (x_i - \bar{x})E(u_i) \\
&= \beta_1
\end{aligned}$$

下面证明 $\hat{\beta}_0$ 的无偏性需要用到这一点。证明如下：

$$\begin{aligned}E(\hat{\beta}_0) &= E(\bar{y} - \hat{\beta}_1\bar{x}) \\ &= E(\bar{y}) - \bar{x}E(\hat{\beta}_1) \\ &= E(\beta_0 + \beta_1\bar{x} + \bar{u}) - \bar{x}E(\hat{\beta}_1) \\ &= \beta_0 + \beta_1\bar{x} + E(\bar{u}) - \beta_1\bar{x} \\ &= \beta_0\end{aligned}$$

请注意，无偏性规定的是随机变量 $\hat{\beta}_0$ 和 $\hat{\beta}_1$ 分布的期望值等于真实值。这并没有告诉我们从一个随机的样本得到的估计值具体是什么。

今后要做的时间序列分析将会放松假设 SLR.2，但对于横截面数据，放松 SLR.2 假设意味着过度取样，将来我们会对非随机抽样问题进行讨论。

如果放松 SLR.4 假设，那么意味着误差项 u 中，有一部分因素既会影响自变量，也会影响因变量，那么 $\hat{\beta}_1$ 将不再是无偏的，并导致伪相关问题。不过将来我们会研究如何确定这种偏误的大小和方向。

5.2 方差与有效性

5.2.1 估计量方差

估计量的有效性是指估计量的方差最小。因此我们需要首先推导 $\hat{\beta}_1$ 的方差, 将 x 用抽样样本估计可得：

$$\begin{aligned}Var(\hat{\beta}_1) &= Var\left(\frac{\sum_{i=1}^{n}(x_i - \bar{x})u_i}{SST_x}\right) \\ &= \frac{1}{SST_x^2}Var\left(\sum_{i=1}^{n}(x_i - \bar{x})u_i\right) \\ &= \frac{1}{SST_x^2}\sum_{i=1}^{n}Var((x_i - \bar{x})u_i) \\ &= \frac{1}{SST_x^2}\sum_{i=1}^{n}(x_i - \bar{x})^2 Var(u_i)\end{aligned}$$

这时让我们把面板数据中常用的假设，SLR.5，纳入进来以简化计算：

$$\begin{aligned}&\frac{1}{SST_x^2}\sum_{i=1}^{n}(x_i - \bar{x})^2 Var(u_i) \\ &= \frac{\sigma^2}{SST_x^2}\sum_{i=1}^{n}(x_i - \bar{x})^2 \\ &= \frac{\sigma^2}{SST_x}\end{aligned}$$

如果 SLR.5 不成立, 则意味着误差项方差会随着 x 的变动而变动，应该表示为：

$$\begin{aligned}&\frac{1}{SST_x^2}\sum_{i=1}^{n}(x_i - \bar{x})^2 Var(u_i) \\ &= \frac{1}{SST_x^2}\sum_{i=1}^{n}\sigma_i^2(x_i - \bar{x})^2\end{aligned}$$

结合 SLR.4 和 SLR.5，我们可以得到 y 的条件均值和条件方差很有用的结论：

$$E(y|x) = \beta_0 + \beta_1 x \tag{21}$$

$$Var(y|x) = \sigma^2 \tag{22}$$

现在我们来推导 $\hat{\beta}_0$ 的方差（基于 SLR.5 成立的条件）：

$$\begin{aligned}
Var(\hat{\beta}_0) &= Var(\bar{y} - \hat{\beta}_1\bar{x}) \\
&= Var(\bar{y}) + \bar{x}^2 Var(\hat{\beta}_1) \\
&= Var(\beta_0 + \beta_1\bar{x} + \bar{u}) + \bar{x}^2\frac{\sigma^2}{SST_x} \\
&= Var(\bar{u}) + \bar{x}^2\frac{\sigma^2}{SST_x} \\
&= \frac{\sigma^2}{N} + \frac{(\sum_{i=1}^n x_1)^2\sigma^2}{N^2(\sum_{i=1}^n x_1^2 - \frac{1}{N}(\sum_{i=1}^n x_i)^2)} \\
&= \frac{N\sigma^2\sum_{i=1}^n x_i^2}{N(N\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2)} \\
&= \frac{\sigma^2\sum_{i=1}^n x_i^2}{N(\sum_{i=1}^n x_i^2 - \frac{1}{N}(\sum_{i=1}^n x_i)^2)} \\
&= \frac{N^{-1}\sigma^2\sum_{i=1}^n x_i^2}{SST_x}
\end{aligned}$$

综上，如果样本 x 的异质性越强，扰动项（噪音）u 的波动越小且样本容量越大，就可以估计出更精确的 β_0、β_1。

5.2.2 估计量的有效性

知道了 $\hat{\beta}_0$、$\hat{\beta}_1$ 的方差, 我们现在可以证明它在线性无偏估计量中是最优（有效）的, 证明过程请参考讲义 4.4（3）“minimum variance” 并自行证明。

由此，在 5 个高斯马尔可夫条件下可以得出的**高斯马尔可夫定理**，即 OLS 估计量是最优线性无偏估计量（BLUE）。

5.3 误差方差的无偏估计

同方差假设中给定了误差方差 $Var(u|x) = \sigma^2$，但在现实应用中，鲜有已知误差项 (不可观测因素) 方差的情况，所以还需要对误差方差进行无偏的估计, 来度量 $\hat{\beta}_1$ 对 β_1 估计情况的波动程度。步骤如下：

$$\begin{aligned}
\hat{u}_i &= y_i - \hat{y}_i \\
&= \beta_0 + \beta_1 x_i + u_i - \hat{\beta}_0 - \hat{\beta}_1 x_i \\
&= (\beta_0 - \hat{\beta}_0) + (\beta_1 - \hat{\beta}_1)x_i + u_i
\end{aligned}$$

等式两边累加取均值：

$$\sum \hat{u}_i = \sum[(\beta_0 - \hat{\beta}_0) + (\beta_1 - \hat{\beta}_1)x_i + u_i]$$
$$0 = (\beta_0 - \hat{\beta}_0) + (\beta_1 - \hat{\beta}_1)\bar{x} + \bar{u}$$

与原方程做差：

$$\hat{u}_i = (\beta_1 - \hat{\beta}_1)(x_i - \bar{x}) + (u_i - \bar{u})$$

两边平方、求和、取期望:

$$E(\sum \hat{u}_i^2) = E[\sum(\hat{\beta}_1 - \beta_1)^2(x_i - \bar{x})^2] - 2E[\sum(\hat{\beta}_1 - \beta_1)(x_i - \bar{x})(u_i - \bar{u})] + E[\sum(u_i - \bar{u})^2]$$

其中, 第一项:

$$E[\sum(\hat{\beta}_1 - \beta_1)^2(x_i - \bar{x})^2] = \sum(x_i - \bar{x})^2 E(\beta_1 - \hat{\beta}_1)^2 = \frac{\sigma^2}{\sum(x_i - \bar{x})^2}\sum(x_i - \bar{x})^2 = \sigma^2$$

第二项中:

$$\begin{aligned} E[\sum(\beta_1 - \hat{\beta}_1)(x_i - \bar{x})(u_i - \bar{u})] &= \frac{E[\sum(u_i - \bar{u}(x_i - \bar{x}))\sum(x_i - \bar{x})u_i]}{\sum(x_i - \bar{x})^2} \\ &= \frac{E[\sum(x_i - \bar{x})u_i]^2}{\sum(x_i - \bar{x})^2} \\ &= \frac{\sum(x_i - \bar{x})^2 Var(u_i)}{\sum(x_i - \bar{x})^2} \\ &= \sigma^2 \end{aligned}$$

tips: 从第二步到第三步，分子中的化简用到了 $E(u_i) = 0$、$E(u_i)^2 = Var(u) = \sigma$ 和 $E(u_i u_j) = 0$，$i \neq j$。这是从 SLR.2（随机抽样）、SLR.4（零条件均值）、SLR.5（同方差）假设中得出的。

第三项:

$$\begin{aligned} E[\sum(u_i - \bar{u})^2] &= \sum E(u_i - \bar{u})^2 \\ &= \sum E(u_i^2) - 2E(\bar{u}\sum u_i) + NE(\bar{u}^2) \\ &= n\sigma^2 - 2\sigma^2 + \sigma^2 \\ &= (n-1)\sigma^2 \end{aligned}$$

第二步到第三步同样用到了上面的 tips，需要一些技巧性的代数变换, 这里不再赘述。

综上，我们可以得到:

$$E(\sum \hat{u}_i^2) = \Sigma^2 - 2\sigma^2 + (n-1)\sigma^2 = (n-2)\sigma^2$$

所以对 σ^2 的无偏估计为:

$$\hat{\sigma}^2 = \frac{1}{n-2}\sum \hat{u}_i^2 \tag{23}$$

据此，再引入两个新的统计量:

- 回归标准误（SER）: $SER = \sqrt{\hat{\sigma}^2}$ 它虽然不是 σ 的无偏估计，但是在大样本情况下，它是 σ 的渐近无偏估计 (一致性成立)。
- 标准误（standard error）: $se(\hat{\beta}_1) = \frac{\hat{\sigma}}{\sqrt{SST_x}} = \frac{\hat{\sigma}}{(\sum(x_i - \bar{x})^2)^{1/2}}$ 它是 $\hat{\beta}$ 标准差 $sd(\hat{\beta}_1) = \sigma/\sqrt{SST_x}$ 的一个自然估计量。

5.4 OLS 估计量的抽样分布

为了得到 OLS 回归估计量的抽样分布，我们需要在高斯-马尔可夫假定之外添加另一个假设：假定 u 独立于 x, 且 u 服从期望为 0 方差为 σ^2 的正态分布:

$$u \sim N(0, \sigma^2) \tag{24}$$

小样本估计量性质	大样本估计量性质	小样本假设	大样本假设
无偏性	一致性	$E(u\|x)=0$	E(u)=0、$Cov(u,x)=0$
正态性	渐进正态性	u 服从正态分布	u 满足中心极限定理假设
有效性	渐进有效性		

表 1: 大样本 OLS 与小样本 OLS 对比

高斯-马尔可夫假定 + 正态性假定被称为经典线性模型假定 (CLM)。在大样本情形时，在某些条件下我们可以放弃 u 的正态性假设。因为相比于误差项的分布，我们更关心估计量 $\hat{\beta}_j$ 的分布, 而它可以由中心极限定理得到。

根据（20）式，我们知道 $\hat{\beta}_j$ 是 u 的线性组合。所以，在 CLM 假定下，给定 x，$\hat{\beta}_j$ 服从正态分布：

$$\hat{\beta}_j \sim N(\beta_j, \frac{\sigma^2}{SST_x}) \tag{25}$$

标准化后，我们可以得到：

$$\frac{\hat{\beta}_j - \beta_j}{sd(\hat{\beta}_j)} \sim N(0,1) \tag{26}$$

6 OLS 的大样本性质

6.1 一致性

证明[1]：对（20）略作变形：

$$\hat{\beta}_1 = \beta_1 + \frac{n^{-1}\sum\left(x_i-\bar{x}\right)u_i}{n^{-1}s_x^2} \tag{27}$$

由于 $Cov(u,x)=0$,

$$\operatorname{Plim}\hat{\beta}_1 = \beta_1 + \frac{\operatorname{Cov}(x,u)}{\operatorname{Var}(x)} = \beta_1$$

在 SLR.1-4 下，该结论成立

6.2 渐进正态性

根据（27）式, 我们可以知道第二项分子依分布收敛到正态分布，分母依概率收敛到一个常数，所以 $\hat{\beta}_1$ 是一个正态分布 $N(\hat{\beta}, \hat{\sigma}^2_{\hat{\beta}_1})$, 其中 $\hat{\sigma}^2_{\hat{\beta}_1}$ 为[2]：

$$\sigma^2_{\hat{\beta}_1} = \frac{1}{n}\frac{\operatorname{Var}\left[\left(x_i-\bar{x}\right)u_i\right]}{\left[\operatorname{Var}\left(x_i\right)\right]^2}$$

6.3 渐近有效性

许多估计方法包括 OLS 得到的估计量都是一致的。然而，在 5 条高斯-马尔可夫假定下，OLS 估计量有最小渐近方差，即 OLS 估计量是渐近有效的。注意如果不满足同方差性，该结论不成立。

---

[1]本章节涉及证明较复杂，牢记结论和假设即可

[2]同方差情况下可以将 u_i 拿出来