课本中相关章节的证明过程

第 2 章有关的证明过程

2.1 一元线性回归模型

有一元线性回归模型为：$y_t = \beta_0 + \beta_1 x_t + u_t$

上式表示变量 y_t 和 x_t 之间的真实关系。其中 y_t 称被解释变量（因变量），x_t 称解释变量（自变量），u_t 称随机误差项，β_0 称常数项，β_1 称回归系数（通常未知）。上模型可以分为两部分。**（1）回归函数部分，$E(y_t) = \beta_0 + \beta_1 x_t$,**

（2）随机部分，u_t 。

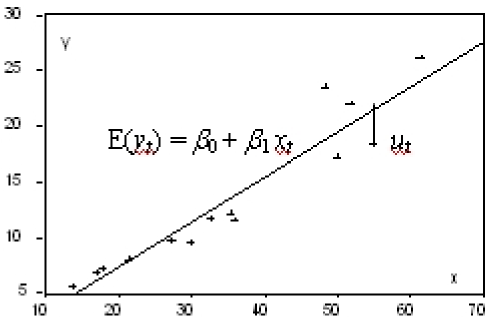

图 2.8　真实的回归直线

这种模型可以赋予各种实际意义，收入与支出的关系；如脉搏与血压的关系；商品价格与供给量的关系；文件容量与保存时间的关系；林区木材采伐量与木材剩余物的关系；身高与体重的关系等。

以收入与支出的关系为例。

假设固定对一个家庭进行观察，随着收入水平的不同，与支出呈线性函数关系。但实际上数据来自各个家庭，来自各个不同收入水平，使其他条件不变成为不可能，所以由数据得到的散点图不在一条直线上（不呈函数关系），而是散在直线周围，服从统计关系。随机误差项 u_t 中可能包括家庭人口数不同，消费习惯不同，不同地域的消费指数不同，不同家庭的外来收入不同等因素。所以，**在经济问题上“控制其他因素不变”实际是不可能的。**

回归模型的随机误差项中一般包括如下几项内容，（1）非重要解释变量的省略，（2）人的随机行为，（3）数学模型形式欠妥，（4）归并误差（粮食的归并）（5）测量误差等。

回归模型存在两个特点。（1）建立在某些假定条件不变前提下抽象出来的回归函数不能百分之百地再现所研究的经济过程。（2）也正是由于这些假定与抽象，才使我们能够透过复杂的经济现象，深刻认识到该经济过程的本质。

通常，线性回归函数 $E(y_t) = \beta_0 + \beta_1 x_t$ 是**观察不到**的，利用样本得到的只是对 $E(y_t) = \beta_0 + \beta_1 x_t$ 的估计，即对β_0 和β_1 的估计。

在对回归函数进行估计之前应该对随机误差项 u_t 做出如下假定。

(1) u_t 是一个随机变量，u_t 的取值服从概率分布。

(2) $E(u_t) = 0$。

(3) $D(u_t) = E[u_t - E(u_t)]^2 = E(u_t)^2 = \sigma^2$。称 u_i 具有同方差性。

(4) u_t 为正态分布（根据中心极限定理）。以上四个假定可作如下表达：$u_t \sim N(0, \sigma^2)$。

(5) $Cov(u_i, u_j) = E[(u_i - E(u_i))(u_j - E(u_j))] = E(u_i, u_j) = 0, \quad (i \neq j)$。含义是不同观测值所

对应的随机项相互独立。称为 u_i 的非自相关性。

(6) x_i 是非随机的。

(7) $\mathrm{Cov}(u_i, x_i) = \mathrm{E}[(u_i - \mathrm{E}(u_i))(x_i - \mathrm{E}(x_i))] = \mathrm{E}[u_i (x_i - \mathrm{E}(x_i))] = \mathrm{E}[u_i x_i - u_i \mathrm{E}(x_i)] = \mathrm{E}(u_i x_i) = 0.$ u_i 与 x_i 相互独立。否则，分不清是谁对 y_t 的贡献。

(8) 对于多元线性回归模型，解释变量之间不能完全相关或高度相关（非多重共线性）。

在假定（1），（2）成立条件下有 $\mathrm{E}(y_t) = \mathrm{E}(\beta_0 + \beta_1 x_t + u_t) = \beta_0 + \beta_1 x_t$ 。

2.2 最小二乘估计（OLS）

对于所研究的经济问题，通常真实的回归直线是观测不到的。收集样本的目的就是要对这条真实的回归直线做出估计。

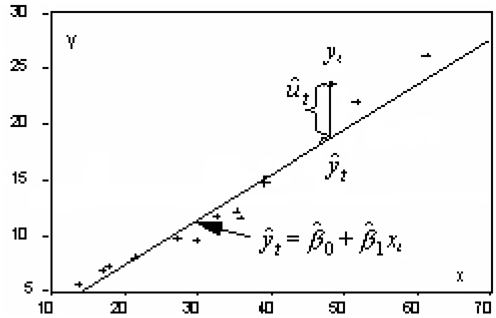

图 2.9

怎样估计这条直线呢？显然综合起来看，这条直线处于样本数据的中心位置最合理。怎样用数学语言描述“处于样本数据的中心位置”？设估计的直线用

$$\hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1 x_t$$

表示。其中 $\hat{y}_t$ 称 y_t 的拟合值（fitted value），$\hat{\beta}_0$ 和 $\hat{\beta}_1$ 分别是 β_0 和 β_1 的估计量。观测值到这条直线的纵向距离用 $\hat{u}_t$ 表示，称为残差。

$$y_t = \hat{y}_t + \hat{u}_t = \hat{\beta}_0 + \hat{\beta}_1 x_t + \hat{u}_t$$

称为估计的模型。假定样本容量为 T。（1）用“残差和最小”确定直线位置是一个途径。但很快发现计算“残差和”存在相互抵消的问题。（2）用“残差绝对值和最小”确定直线位置也是一个途径。但绝对值的计算比较麻烦。（3）最小二乘法的原则是以“残差平方和最小”确定直线位置。用最小二乘法除了计算比较方便外，得到的估计量还具有优良特性（这种方法对异常值非常敏感）。设残差平方和用 Q 表示，

$$Q = \sum_{i=1}^{T} \hat{u}_t^{\,2} = \sum_{i=1}^{T} (y_t - \hat{y}_t)^2 = \sum_{i=1}^{T} (y_t - \hat{\beta}_0 - \hat{\beta}_1 x_t)^2 ,$$

则通过 Q 最小确定这条直线，即确定 $\hat{\beta}_0$ 和 $\hat{\beta}_1$ 的估计值。以 $\hat{\beta}_0$ 和 $\hat{\beta}_1$ 为变量，把 Q 看作是 $\hat{\beta}_0$ 和 $\hat{\beta}_1$ 的函数，这是一个求极值的问题。求 Q 对 $\hat{\beta}_0$ 和 $\hat{\beta}_1$ 的偏导数并令其为零，得正规方程，

$$\left\{ \frac{\partial Q}{\partial \hat{\beta}_0} = 2 \sum_{i=1}^{T} (y_t - \hat{\beta}_0 - \hat{\beta}_1 x_t)(-1) = 0 \right. \qquad (2.7)$$

$$\frac{\partial Q}{\partial \hat{\beta}_1} = 2\sum_{i=1}^{T}(y_t - \hat{\beta}_0 - \hat{\beta}_1 x_t)(-x_t) = 0 \tag{2.8}$$

下面用代数和矩阵两种形式推导计算结果。

首先用代数形式推导。由（2.7）、（2.8）式得，

$$\begin{cases} \sum_{i=1}^{T}(y_t - \hat{\beta}_0 - \hat{\beta}_1 x_t) = 0 & (2.9) \\ \sum_{i=1}^{T}(y_t - \hat{\beta}_0 - \hat{\beta}_1 x_t)\,x_t = 0 & (2.10) \end{cases}$$

（2.9）式两侧用除 T，并整理得，

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \tag{2.11}$$

把（2.11）式代入（2.10）式并整理，得，

$$\sum_{i=1}^{T}[(y_t - \bar{y}) - \hat{\beta}_1 (x_t - \bar{x})]\,x_t = 0 \tag{2.12}$$

$$\sum_{i=1}^{T}(y_t - \bar{y})x_t - \hat{\beta}_1 \sum_{i=1}^{T}(x_t - \bar{x})x_t = 0 \tag{2.13}$$

$$\hat{\beta}_1 = \frac{\sum x_t (y_t - \bar{y})}{\sum (x_t - \bar{x})x_t} \tag{2.14}$$

因为 $\sum_{i=1}^{T}\bar{x}(y_t - \bar{y}) = 0$，$\sum_{i=1}^{T}\bar{x}(x_t - \bar{x}) = 0$，[采用离差和为零的结论：$\sum_{i=1}^{T}(x_t - \bar{x}) = 0$，$\sum_{i=1}^{T}(y_t - \bar{y}) = 0$]。

所以，通过配方法，分别在（2.14）式的分子和分母上减 $\sum_{i=1}^{T}\bar{x}(y_t - \bar{y})$ 和 $\sum_{i=1}^{T}\bar{x}(x_t - \bar{x})$ 得，

$$\hat{\beta}_1 = \frac{\sum x_t (y_t - \bar{y}) - \sum \bar{x}(y_t - \bar{y})}{\sum (x_t - \bar{x})x_t - \sum \bar{x}(x_t - \bar{x})} \tag{2.15}$$

$$= \frac{\sum (x_t - \bar{x})(y_t - \bar{y})}{\sum (x_t - \bar{x})^2} \tag{2.16}$$

即有结果：

$$\begin{cases} \hat{\beta}_1 = \dfrac{\sum (x_t - \bar{x}_t)(y_t - \bar{y}_t)}{\sum (x_t - \bar{x})^2} & (2.17) \\ \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} & \end{cases}$$

这是观测值形式。如果以离差形式表示，就更加简洁好

记。

$$\begin{cases} \hat{\beta}_1 = \dfrac{\sum x_t y_t}{\sum x_t^{\,2}} \\ \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \end{cases}$$

矩阵形式推导计算结果：

由正规方程，

$$\begin{cases} \dfrac{\partial Q}{\partial \hat{\beta}_0} = 2\sum_{i=1}^{T}(y_t - \hat{\beta}_0 - \hat{\beta}_1 x_t)(-1) = 0 \\ \dfrac{\partial Q}{\partial \hat{\beta}_1} = 2\sum_{i=1}^{T}(y_t - \hat{\beta}_0 - \hat{\beta}_1 x_t)(-x_t) = 0 \end{cases}$$

$$\begin{cases} \hat{\beta}_0 T + \hat{\beta}_1 \left(\sum_{i=1}^{T} x_t\right) = \sum_{i=1}^{T} y_t \\ \hat{\beta}_0 \sum_{i=1}^{T} x_t + \hat{\beta}_1 \left(\sum_{i=1}^{T} x_t^{\,2}\right) = \sum_{i=1}^{T} x_t y_t \end{cases}$$

$$\begin{bmatrix} T & \sum x_t \\ \sum x_t & \sum x_t^{\,2} \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = \begin{bmatrix} \sum y_t \\ \sum x_t y_t \end{bmatrix}$$

$$\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = \begin{bmatrix} T & \sum x_t \\ \sum x_t & \sum x_t^{\,2} \end{bmatrix}^{-1} \begin{bmatrix} \sum y_t \\ \sum x_t y_t \end{bmatrix}$$

$$= \frac{1}{T\sum x_t^{\,2} - (\sum x_t)^2} \begin{bmatrix} \sum x_t^{\,2} & -\sum x_t \\ -\sum x_t & T \end{bmatrix} \begin{bmatrix} \sum y_t \\ \sum x_t y_t \end{bmatrix} = \begin{pmatrix} \dfrac{\sum x_t^{\,2} \sum y_t - \sum x_t \sum x_t y_t}{T\sum x_t^{\,2} - (\sum x_t)^2} \\ \dfrac{T\sum x_t y_t - \sum x_t \sum y_t}{T\sum x_t^{\,2} - (\sum x_t)^2} \end{pmatrix}$$

注意：关键是求逆矩阵$\begin{bmatrix} T & \sum x_t \\ \sum x_t & \sum x_t^{\,2} \end{bmatrix}^{-1}$。它等于其伴随阵除以其行列式，伴随阵是其行列式对应的代数余子式构成的方阵的转置。

写成观测值形式。

$$\begin{cases} \hat{\beta}_1 = \dfrac{\sum (x_t - \bar{x}_t)(y_t - \bar{y}_t)}{\sum (x_t - \bar{x})^2} \\ \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \end{cases}$$

如果，以离式形式表示更为简洁：

$$\begin{cases} \hat{\beta}_1 = \dfrac{\sum x_t y_t}{\sum {x_t}^2} \\ \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \end{cases}$$

2.3 一元线性回归模型的特性

1．线性特性（将结果离差转化为观测值表现形式）

$$\hat{\beta}_2=\frac{\sum x_i y_i}{\sum x_i^2}=\frac{\sum x_i(Y_i-\bar{Y})}{\sum x_i^2}$$

$$=\sum\frac{x_i}{\sum x_i^2}Y_i-\frac{\bar{Y}\sum x_i}{\sum x_i^2}=\sum K_iY_i$$

$$\hat{\beta}_1=\bar{Y}-\hat{\beta}_2\bar{X}=\bar{Y}-\bar{X}\sum K_iY_i$$

$$=\frac{1}{n}\sum Y_i-\sum K_i\bar{X}Y_i=\sum\left(\frac{1}{n}-K_i\bar{X}\right)Y_i$$

2．无偏性

$$\hat{\beta}_2=\sum K_iY_i=\sum K_i(\beta_1+\beta_2X_i+u_i)$$

$$=\sum K_i\beta_1+\sum K_i\beta_2X_i+\sum K_iu_i$$

$$=\beta_1\sum K_i+\beta_2\sum K_iX_i+\sum K_iu_i$$

其中：

$$\sum K_i=\sum\frac{x_i}{\sum x_i^2}=\frac{\sum x_i}{\sum x_i^2}=\frac{\sum(X_i-\bar{X})}{\sum x_i^2}=0$$

$$\sum K_iX_i=\sum\frac{x_i}{x_i^2}X_i=\frac{\sum x_i(X_i-\bar{X}+\bar{X})}{\sum x_i^2}$$

$$=\frac{\sum x_i(X_i-\bar{X})+\sum x_i\bar{X}}{\sum x_i^2}$$

$$=\frac{\sum x_i^2+\bar{X}\sum x_i}{\sum x_i^2}=\frac{\sum x_i^2}{\sum x_i^1}=1$$

故有：$\hat{\beta}_2=\beta_2+\sum K_iu_i$

$$E\hat{\beta}_2=E(\beta_2+\sum K_iu_i)=\beta_2+\sum K_iEu_i=\beta_2$$

$$\hat{\beta}_1 = \sum\left(\frac{1}{n} - K_i\overline{X}\right)Y_i$$

$$= \sum\left(\frac{1}{n} - K_i\overline{X}\right)\left(\beta_1 + \beta_2 X_i + u_i\right)$$

$$= \sum\frac{\beta_1}{n} + \sum\frac{\beta_2 X_i}{n} + \sum\frac{u_i}{n}$$

$$-\sum\beta_1 K_i\overline{X} - \sum\beta_2 K_i\overline{X}X_i - \sum K_i\overline{X}u_i$$

$$= \beta_1 + \beta_2\overline{X} + \overline{u} - \beta_1\overline{X}\sum K_i - \beta_2\overline{X}\sum K_i X_i - \overline{X}\sum K_i u_i$$

$$= \beta_1 + \sum(\frac{1}{n} - \overline{X}K_i)u_i$$

$$\therefore E\hat{\beta}_1 = \beta_1 + \sum(\frac{1}{n} - K_i\overline{X})Eu_i = \beta_1$$

3． 有效性

首先讨论参数估计量的方差。

$$Var(\hat{\beta}_2) = E(\hat{\beta}_2 - E(\hat{\beta}_2))^2$$

$$= E(\hat{\beta}_2 - \beta_2)^2 = E((\beta_2 + \sum K_i u_i) - \beta_2)^2 = E(\sum K_i u_i)^2$$

$$\because \left(\sum K_i u_i\right)^2 = (K_1u_1 + K_2u_2 + \ldots + K_nu_n)(K_1u_1 + K_2u_2 + \ldots + K_nu_n)$$

$$= \sum(K_iu_i)^2 + \sum_{i\neq j}\sum K_iK_ju_iu_j$$

$$\therefore E(\sum K_iu_i)^2 = E\sum(K_iu_i)^2 + E\sum_{i\neq j}\sum K_iK_ju_iu_j$$

$$= \sum K_i^2 Eu_i^2 = o^2\sum\left(\frac{x_i}{\sum x_i^2}\right)^2 = \frac{o^2}{\sum x_i^2}$$

即：

$$Var(\hat{\beta}_2) = \frac{o^2}{\sum x_i^2}$$

同理有：

$$Var(\hat{\beta}_1) = o^2 \frac{\sum X_i^2}{n\sum x_i^2}$$

$$Var(\hat{\beta}_1) = E(\hat{\beta}_1 - E(\hat{\beta}_1))^2 = E(\sum\left(\frac{1}{n} - K_i\overline{X}\right)u_i)^2$$

$$\left(\sum\left(\frac{1}{n} - K_i\overline{X}\right)u_i\right)^2 = \sum\left(\frac{1}{n} - K_i\overline{X}\right)^2 u_i^2$$

$$+\sum\sum_{i\neq j}\left(\frac{1}{n}K_i\overline{X}\right)\left(\frac{1}{n} - K_j\overline{X}\right)u_i u_j$$

$$Var(\hat{\beta}_1) = o^2\sum\left(\frac{1}{n} - K_i\overline{X}\right)^2$$

$$= o^2\sum(\frac{1}{n^2} - 2\frac{K_i\overline{X}}{n} + K_i^2\overline{X}^2)$$

$$= \frac{o^2}{n} - \frac{2o^2\overline{X}}{n}\sum K_i + o^2\overline{X}^2\sum K_i^2$$

$$= \frac{o^2}{n} + \frac{o^2(\sum X_i)^2}{n^2\sum x_i^2}$$

$$= \frac{o^2}{n^2}\left(\frac{n(\sum x_i^2) + (\sum X_i)^2}{\sum x_i^2}\right)$$

$$= \frac{no^2\left((\sum X_i^2 - n\overline{X}^2) + \frac{1}{n}(\sum X_i)^2\right)}{n^2\sum x_i^2}$$

$$= \sigma^2 \frac{\sum X_i^2}{n\sum x_i^2}$$

显然各自的标准误差为：

$$se(\hat{\beta}_2) = \frac{\sigma}{\sqrt{\sum x_i^2}}, \quad se(\hat{\beta}_1) = \sigma\sqrt{\frac{\sum X_i^2}{n\sum x_i^2}}$$

标准差的作用：衡量估计值的精度。

由于σ为总体方差，也需要用样本进行估计。

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{n-2}$$

证明过程如下：

▯▯ $:Y_i = \beta_1 + \beta_2 X_i + u_i$

因此有： $\overline{Y} = \beta_1 + \beta_2 \overline{X} + \overline{u}$

那么：$(Y_i - \overline{Y}) = y_i = (\beta_1 + \beta_2 X_i + u_i) - (\beta_1 + \beta_2 \overline{X} + \overline{u})$

$$= \beta_2 x_i + (u_i - \overline{u})$$

根据定义：$e_i = y_i - \hat{\beta}_2 x_i$,

（实际观测值与样本回归线的差值）

则有：

$$e_i = (\beta_2 x_i + (u_i - \overline{u})) - \hat{\beta}_2 x_i = (u_i - \overline{u}) - (\hat{\beta}_2 - \beta_2) x_i$$

两边平方，再求和：

$$\sum e_i^2 = \sum (u_i - \overline{u})^2 - \sum 2(u_i - \overline{u})(\hat{\beta}_2 - \beta_2) x_i + \sum ((\hat{\beta}_2 - \beta_2) x_i)^2$$

$$= (\hat{\beta}_2 - \beta_2)^2 \sum x_i^2 + \sum (u_i - \overline{u})^2 - 2(\hat{\beta}_2 - \beta_2) \sum (u_i - \overline{u}) x_i$$

对上式两边取期望有：

$$E(\sum e_i^2) = \sum x_i^2 E(\hat{\beta}_2 - \beta_2)^2$$

$$+ E(\sum (u_i - \overline{u})^2) - 2E\left[(\hat{\beta}_2 - \beta_2)\sum (u_i - \overline{u}) x_i\right]$$

$$= A + B - C$$

其中：$A = \sum x_i^2 \frac{o^2}{\sum x_i^2} = o^2$

$$B = E\sum u_i^2 - nE\overline{u}^2 = no^2 - nE\frac{1}{n^2}(\sum u_i)^2$$

$$= no^2 - \frac{1}{n}E(\sum u_i^2 + \sum_{i \neq j}\sum u_i u_j)$$

$$= no^2 - \frac{1}{n}(no^2) = (n-1)o^2$$

$$C = 2E\left[\frac{\sum x_i u_i}{\sum x_i^2}\right]\left(\sum u_i x_i - \overline{u}\sum x_i\right)$$

$$= 2E\left[\frac{\left(\sum x_i u_i\right)^2}{\sum x_i^2}\right] = 2E(\hat{\beta}_2 - \beta_2)^2 \sum x_i^2$$

$$= 2\frac{o^2}{\sum x_i^2}\sum x_i^2$$

$$= 2o^2$$

故有：$E\sum e_i^2 = (n-1)o^2$

即有：$o^2 = E\left[\frac{\sum e_i^2}{n-2}\right]$,

令 $\hat{o}^2 = \frac{\sum e_i^2}{n-2}$，则问题得证。

关于 $\sum e_i^2$ 的计算：

$$\sum e_i^2 = \sum y_i^2 - \hat{\beta}_2^2 \sum x_i^2 = \sum y_i^2 - \hat{\beta}_2 \sum x_i y_i$$

关于 $\overline{R}^2 \leq R^2$ 的证明：

$$\overline{R}^2 = 1 - \left(1 - R^2\right)\frac{n-1}{n-k} = 1 - a \times \left(1 - R^2\right)$$，其中：$a \geq 1$。

当 $k = 1 \Rightarrow a = 1$

$$\overline{R}^2 = 1 - \left(1 - R^2\right) \times \frac{n-1}{n-1} = 1 - \left(1 - R^2\right) = R^2$$

当 $k > 1 \Rightarrow a > 1$，当 $0 \leq R^2 \leq 1$ 时，有：

$$\begin{aligned} R^2 - \overline{R}^2 &= R^2 - \left[1 - \left(1 - R^2\right) \times a\right] \\ &= R^2 - 1 + a - aR^2 \\ &= a - 1 - R^2(a-1) \\ &= (a-1)\left(1 - R^2\right) > 0 \\ &\Rightarrow R^2 \geq \overline{R}^2 \end{aligned}$$

Q.E.D.

关于 R^2 可能小于 0 的证明。

设：$Y_t = \beta_2 X_t + u_t$

则有：

$$J = \min_{\hat{\beta}_2} \sum e_t^2 = \min_{\hat{\beta}_2} \sum \left(Y_t - \hat{\beta}_2 X_t\right)^2$$

那么 $\dfrac{\partial J}{\partial \hat{\beta}_2} = 0$

$$= -2\sum \left(Y_t - \hat{\beta}_2 X_t\right) X_t \Rightarrow \sum X_t e_t = 0$$

但：$\sum e_t \neq 0$，因为没有 $\dfrac{\partial J}{\partial \hat{\beta}_1}=0$ 存在。

同时，还有：

$$\bar{Y}=\hat{\beta}_2\bar{X}+\bar{e}$$

$$\begin{aligned}Y_t-\bar{Y}&=\hat{\beta}_2X_t-\bar{Y}+e_t\\&=\hat{\beta}_2X_t-\left(\hat{\beta}_2\bar{X}+\bar{e}\right)+e_t\\&=\hat{\beta}_2\left(X_t-\bar{X}\right)+\left(e_t-\bar{e}\right)\end{aligned}$$

$$\begin{aligned}TSS&=\sum\left(Y_t-\bar{Y}\right)^2=\sum Y_t^2-n\bar{Y}^2\\&=\sum\left(\hat{\beta}_2\left(X_t-\bar{X}\right)+\left(e_t-\bar{e}\right)\right)^2\\&=\sum\left(\hat{\beta}_2\left(X_t-\bar{X}\right)\right)^2+\sum\left(e_t-\bar{e}\right)^2+2\hat{\beta}_2\sum\left(X_t-\bar{X}\right)\left(e_t-\bar{e}\right)\end{aligned}$$

其中：

$$\begin{aligned}\sum\left(X_t-\bar{X}\right)\left(e_t-\bar{e}\right)&=\sum X_t\left(e_t-\bar{e}\right)-\bar{X}\sum\left(e_t-\bar{e}\right)\\&=\sum X_te_t-\bar{e}\sum X_t-0\end{aligned}$$

$$\because \sum\left(e_t-\bar{e}\right)=\sum e_t-n\bar{e}=\sum e_t-n\frac{1}{n}\sum e_t=0,$$

和 $\sum X_te_t=0$

$$\therefore \sum\left(X_t-\bar{X}\right)\left(e_t-\bar{e}\right)=-n\bar{X}\bar{e}$$

则：

$$\begin{aligned}TSS&=\hat{\beta}_2^2\sum\left(X_t-\bar{X}\right)^2+\sum\left(e_t-\bar{e}\right)^2-2\hat{\beta}_2n\bar{X}\bar{e}\\&=\hat{\beta}_2^2\sum X_t^2-n\hat{\beta}_2^2\bar{X}^2+\sum e_t^2-n\bar{e}^2-2\hat{\beta}_2n\bar{X}\bar{e}\\&=\hat{\beta}_2^2\sum X_t^2+\sum e_t^2-n\bar{e}^2-2\hat{\beta}_2n\bar{X}\bar{e}-n\hat{\beta}_2^2\bar{X}^2\\&=\hat{\beta}_2^2\sum X_t^2+\sum e_t^2-n\left(\hat{\beta}_2^2\bar{X}^2+2\hat{\beta}_2\bar{X}\bar{e}+\bar{e}^2\right)\end{aligned}$$

考虑到：

$$n\overline{Y}^2 = n\left(\hat{\beta}_2\overline{X} + \overline{e}\right)^2 = n\left(\hat{\beta}_2^2\overline{X}^2 + 2\hat{\beta}_2\overline{X}\overline{e} + \overline{e}^2\right)$$

$$\sum Y_t^2 = \sum\left(\hat{\beta}_2 X_t + e_t\right)^2 = \hat{\beta}_2^2\sum X_t^2 + 2\sum\hat{\beta}_2 X_t e_t + \sum e_t^2$$

$$= \hat{\beta}_2^2\sum X_t^2 + \sum e_t^2$$

若定义

$$TSS = \sum Y_t^2 - n\overline{Y}^2 = \hat{\beta}_2^2\sum X_t^2 + \sum e_t^2 - n\left(\hat{\beta}_2^2\overline{X}^2 + 2\hat{\beta}_2\overline{X}\overline{e} + \overline{e}^2\right)$$

$$RSS = TSS - \hat{\beta}_2^2\sum X_t^2$$

$$= \sum e_t^2$$

$$RSS - TSS = n\left(\hat{\beta}_2^2\overline{X}^2 + 2\hat{\beta}_2\overline{X}\overline{e} + \overline{e}^2\right) - \hat{\beta}_2^2\sum X_t^2$$

$$= n\left(\hat{\beta}_2^2\left(\frac{1}{n}\sum X_t\right)^2 + 2\hat{\beta}_2\overline{X}\overline{e} + \overline{e}^2\right) - \hat{\beta}_2^2\sum X_t^2$$

$$= n\hat{\beta}_2^2\left(\sum X_t\right)^2 + n\left(2\hat{\beta}_2\overline{X}\overline{e} + \overline{e}^2\right) - \hat{\beta}_2^2\sum X_t^2$$

$$= n\hat{\beta}_2^2\left(\sum X_t^2 + \sum_{t\neq s}\sum X_t X_s\right) + n\left(2\hat{\beta}_2\overline{X}\overline{e} + \overline{e}^2\right) - \hat{\beta}_2^2\sum X_t^2$$

$$= (n-1)\hat{\beta}_2^2\sum X_t^2 + n\hat{\beta}_2^2\sum_{t\neq s}\sum X_t X_s + n\left(2\hat{\beta}_2\overline{X}\overline{e} + \overline{e}^2\right)$$

可能小于 0。

参考书：

Dennis J. Aigner Basic Econometrics, Prentice-Hall, Englewood Cliffs, N. J. 1971,pp85-88

第二章

2.1 简单线性回归最小二乘估计最小方差性质的证明

对于 OLS 估计式 $\hat{\beta}_1$ 和 $\hat{\beta}_2$，已知其方差为

$$Var(\hat{\beta}_1) = \sigma^2 \frac{\sum X_i^2}{N\sum x_i^2}$$

$$Var(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_i^2}$$

这里只证明 $Var(\hat{\beta}_2)$ 最小，$Var(\hat{\beta}_1)$ 最小的证明可以类似得出。

设 β_2 的另一个线性无偏估计为 β_2^*，即

$$\beta_2^* = \sum w_i Y_i$$

其中

$$w_i \neq k_i, k_i = \frac{x_i}{\sum x_i^2}$$

$$E(\beta_2^*) = E(\sum w_i Y_i)$$

$$= E[\sum w_i(\beta_1 + \beta_2 X_i + u_i)]$$

$$= \beta_1 \sum w_i + \beta_2 \sum w_i X_i$$

因为 β_2^* 也是 β_2 的无偏估计，即 $E(\beta_2^*) = \beta_2$，必须有

$$\sum w_i = 0, \quad \sum w_i X_i = 1$$

同时

$$Var(\beta_2^*) = Var(\sum w_i Y_i)$$

$$= \sum w_i^2 Var(Y_i)$$

$$= \sigma^2 \sum w_i^2 \qquad [\text{因为} Var(Y_i) = \sigma^2]$$

$$= \sigma^2 \sum (w_i - k_i + k_i)^2$$

$$= \sigma^2 \sum (w_i - k_i)^2 + \sigma^2 \sum k_i^2 + 2\sigma^2 \sum (w_i - k_i) k_i$$

$$= \sigma^2 \sum (w_i - k_i)^2 + \sigma^2 \sum k_i^2 + 2\sigma^2 (\sum w_i k_i - \sum k_i^2)$$

上式最后一项中

$$\sum w_i k_i - \sum k_i^2 = \frac{\sum w_i x_i}{\sum x_i^2} - \frac{\sum x_i^2}{(\sum x_i^2)^2}$$

$$=\frac{\sum w_i(X_i-\bar{X})}{\sum x_i^2}-\frac{1}{\sum x_i^2}$$

$$=\frac{\sum w_i X_i-\bar{X}\sum w_i}{\sum x_i^2}-\frac{1}{\sum x_i^2}$$

$$=0$$ （因为 $\sum w_i=0$， $\sum w_i X_i=1$

）

所以 $$Var(\beta_2^*)=\sigma^2\sum(w_i-k_i)^2+\sigma^2\sum[\frac{x_i^2}{(\sum x_i^2)^2}]$$

$$=\sigma^2\sum(w_i-k_i)^2+\frac{\sigma^2}{\sum x_i^2}$$

$$=\sigma^2\sum(w_i-k_i)^2+Var(\hat{\beta}_2)$$

而 $\sigma^2\geq 0$，因为 $w_i\neq k_i$，则有 $(w_i-k_i)^2\geq 0$，为此

$$Var(\beta_2^*)\geq Var(\hat{\beta}_2)$$

只有 $w_i=k_i$ 时，$Var(\beta_2^*)=Var(\hat{\beta}_2)$，由于 β_2^* 是任意设定的 β_2 的线性无偏估计式，这表明 β_2 的 OLS 估计式具有最小方差性。

2.2 σ^2 最小二乘估计的证明

用离差形式表示模型时

$$y_i=Y_i-\bar{Y}$$

$$=(\beta_1+\beta_2 X_i+u_i)-(\beta_1+\beta_2\bar{X}+\bar{u})$$

$$=(u_i-\bar{u})+\beta_2 x_i$$

而且 $$\hat{y}_i=\hat{Y}_i-\bar{Y}$$

$$=(\hat{\beta}_1+\hat{\beta}_2 X_i)-(\hat{\beta}_1+\hat{\beta}_2\bar{X})$$

$$=\hat{\beta}_2 x_i$$

因此 $$e_i=y_i-\hat{y}_i=(u_i-\bar{u})-(\hat{\beta}_2-\beta_2)x_i$$

则有 $$\sum e_i^2=\sum[(u_i-\bar{u})-(\hat{\beta}_2-\beta_2)x_i]^2$$

$$=\sum(u_i-\bar{u})^2+(\hat{\beta}_2-\beta_2)^2\sum x_i^2-2(\hat{\beta}_2-\beta_2)\sum(u_i-\bar{u})x_i$$

取 $\sum e_i^2$ 的期望

$$E(\sum e_i^2)=E[\sum(u_i-\bar{u})^2]+\sum x_i^2E(\hat{\beta}_2-\beta_2)^2-2E[(\hat{\beta}_2-\beta_2)\sum(u_i-\bar{u})x_i]$$

式中 (1) $E[\sum(u_i-\bar{u})^2]=E[\sum u_i^2-n(\bar{u})^2]$

$$=\sum E(u_i^2)-\frac{1}{n}E(\sum u_i)^2$$

$$=\sum\sigma^2-\frac{1}{n}E(u_1^2+u_2^2+\cdots+u_n^2+2u_1u_2+\cdots+2u_{n-1}u_n)$$

$$=\sum\sigma^2-\frac{1}{n}E(u_1^2+u_2^2+\cdots+u_n^2)$$

$$=\sum\sigma^2-\frac{1}{n}n\sigma^2=(n-1)\sigma^2$$

(2) $\sum x_i^2E(\hat{\beta}_2-\beta_2)^2=\sum x_i^2\frac{\sigma^2}{\sum x_i^2}=\sigma^2$

(3) $-2E[(\hat{\beta}_2-\beta_2)\sum(u_i-\bar{u})x_i]=-2E[\frac{\sum x_iu_i}{\sum x_i^2}(\sum x_iu_i-\bar{u}\sum x_i)]$

$$=-2E[\frac{(\sum x_iu_i)^2}{\sum x_i^2}]$$

$$=-2E[(\hat{\beta}_2-\beta_2)^2\sum x_i^2]$$

$$=-2\sum x_i^2E(\hat{\beta}_2-\beta_2)^2=-2\sigma^2$$

所以 $E(\sum e_i^2)=(n-1)\sigma^2+\sigma^2-2\sigma^2=(n-2)\sigma^2$

如果定义 $\hat{\sigma}^2=\frac{\sum e_i^2}{n-2}$

其期望值为 $E(\hat{\sigma}^2)=E[\frac{\sum e_i^2}{n-2}]=\sigma^2$

这说明 $\hat{\sigma}^2=\frac{\sum e_i^2}{n-2}$ 是 σ^2 的无偏估计。

第三章

3.1　多元线性回归最小二乘估计无偏性的证明

因为
$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\beta}+\mathbf{U})$$
$$= (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X})\boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{U}$$
$$= \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{U}$$

对两边取期望，
$$E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'[E(\mathbf{U})]$$
$$= \boldsymbol{\beta} \qquad \text{［由假定 1: } E(\mathbf{U}) = \mathbf{0}\text{］}$$

即 $\hat{\boldsymbol{\beta}}$ 是 $\boldsymbol{\beta}$ 的无偏估计。

3.2　多元线性回归最小二乘估计最小方差性的证明

设 $\boldsymbol{\beta}^*$ 为 $\boldsymbol{\beta}$ 的另一个关于 **Y** 的线性无偏估计式，可知

$$\boldsymbol{\beta}^* = \mathbf{AY} \qquad \text{（}\mathbf{A}\text{ 为常数矩阵）}$$

由无偏性可得
$$E(\boldsymbol{\beta}^*) = E(\mathbf{AY}) = E[\mathbf{A}(\mathbf{X}\boldsymbol{\beta}+\mathbf{U})]$$
$$= E(\mathbf{AX}\boldsymbol{\beta}) + \mathbf{A}E(\mathbf{U})$$
$$= \mathbf{AX}E(\boldsymbol{\beta}) = \boldsymbol{\beta}$$

所以必须有
$$\mathbf{AX} = \mathbf{I}$$

要证明最小二乘法估计式的方差 $Var(\hat{\beta})$ 小于其他线性去偏估计式的方差 $Var(\beta^*)$，只要证明协方差矩阵之差

$$E[(\boldsymbol{\beta}^* - \boldsymbol{\beta})(\boldsymbol{\beta}^* - \boldsymbol{\beta})'] - E[(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})']$$

为半正定矩阵，则称最小二乘估计 $\hat{\boldsymbol{\beta}}$ 是 $\boldsymbol{\beta}$ 的最小方差线性无偏估计式。

因为
$$\boldsymbol{\beta}^* - \boldsymbol{\beta} = \mathbf{AY} - \boldsymbol{\beta} = \mathbf{A}(\mathbf{X}\boldsymbol{\beta}+\mathbf{U}) - \boldsymbol{\beta}$$
$$= \mathbf{AX}\boldsymbol{\beta} + \mathbf{AU} - \boldsymbol{\beta}$$
$$= \boldsymbol{\beta} + \mathbf{AU} - \boldsymbol{\beta} = \mathbf{AU}$$

所以
$$E[(\boldsymbol{\beta}^* - \boldsymbol{\beta})(\boldsymbol{\beta}^* - \boldsymbol{\beta})'] = E[(\mathbf{AU})(\mathbf{AU})'] = E(\mathbf{AUU}'\mathbf{A}')$$
$$= \mathbf{AE}(\mathbf{UU}')\mathbf{A}' = \mathbf{AA}'\sigma^2$$

由于
$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{U}$$

$$E[(\hat{\boldsymbol\beta}-\boldsymbol\beta)(\hat{\boldsymbol\beta}-\boldsymbol\beta)'] = E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{U}][(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{U}]'$$

$$= E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}''\mathbf{U}][\mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}]$$

$$= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\mathbf{U}\mathbf{U}')\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$

$$= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\sigma^2 = (\mathbf{X}'\mathbf{X})^{-1}\sigma^2$$

所以 $$E[(\boldsymbol\beta^*-\boldsymbol\beta)(\boldsymbol\beta^*-\boldsymbol\beta)'] - E[(\hat{\boldsymbol\beta}-\boldsymbol\beta)(\hat{\boldsymbol\beta}-\boldsymbol\beta)'] = \mathbf{A}\mathbf{A}'\sigma^2 - (\mathbf{X}'\mathbf{X})^{-1}\sigma^2$$

$$= [\mathbf{A}\mathbf{A}' - (\mathbf{X}'\mathbf{X})^{-1}]\sigma^2$$

由于

$$[\mathbf{A}-(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'][\mathbf{A}-(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']' = [\mathbf{A}-(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'][\mathbf{A}'-\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}]$$

$$= \mathbf{A}\mathbf{A}' - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{A}' - \mathbf{A}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$

$$= \mathbf{A}\mathbf{A}' - (\mathbf{X}'\mathbf{X})^{-1}$$

由线性代数知，对任一非奇异矩阵$\mathbf{C}$，$\mathbf{C}\mathbf{C}'$为半正定矩阵。如果令$[\mathbf{A}-(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'] = \mathbf{C}$

则 $$\mathbf{C}\mathbf{C}' = [\mathbf{A}-(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'][\mathbf{A}-(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']' = \mathbf{A}\mathbf{A}' - (\mathbf{X}'\mathbf{X})^{-1}$$

由于半正定矩阵对角线元素非负，因此有$\mathbf{A}\mathbf{A}' - (\mathbf{X}'\mathbf{X})^{-1} \geq \mathbf{0}$

即 $$E(\beta_j^* - \beta_j)^2 - E(\hat\beta_j - \beta_j) \geq 0 \qquad (j = 1, 2, \cdots k)$$

这证明了β_j的最小二乘估计$\hat\beta_j$在β_j的所有无偏估计中是方差最小的估计式。

3.3 残差平方和$\sum e_i^2$的均值为$(n-k)\sigma^2$的证明

由残差向量的定义及参数的最小二乘估计式，有

$$\mathbf{e} = \mathbf{Y} - \hat{\mathbf{Y}} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol\beta}$$

$$= \mathbf{Y} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

$$= [\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{Y}$$

可以记$\mathbf{P} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$，则

$$\mathbf{e} = \mathbf{P}\mathbf{Y} = [\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'][\mathbf{X}\boldsymbol\beta + \mathbf{U}]$$

$$= \mathbf{X}\boldsymbol\beta - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\boldsymbol\beta + \mathbf{P}\mathbf{U}$$

$$= \mathbf{P}\mathbf{U}$$

容易验证，P 为对称等幂矩阵，即

$$\mathbf{P} = \mathbf{P}'$$

$$\mathbf{P}^2 = \mathbf{PP} = \mathbf{P}$$

残差向量的协方差矩阵为

$$Var(\mathbf{e}) = E(\mathbf{ee}') = E[\mathbf{PU}(\mathbf{PU})']$$
$$= E[\mathbf{P}(\mathbf{UU}')\mathbf{P}']$$
$$= \mathbf{P}[E(\mathbf{UU}')]\mathbf{P}'$$
$$= \mathbf{P}(\sigma^2\mathbf{I})\mathbf{P}'$$
$$= \mathbf{PP}'\sigma^2 = \mathbf{P}\sigma^2$$

利用矩阵迹的性质，有

$$\sum e_i^2 = \mathbf{e}'\mathbf{e} = tr(\mathbf{ee}')$$

两边取期望得

$$E(\sum e_i^2) = E(\mathbf{e}'\mathbf{e}) - E[tr(\mathbf{ee}')]$$
$$= tr[E(\mathbf{e}'\mathbf{e})] = tr[\mathbf{P}\sigma^2]$$
$$= \sigma^2 tr[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']$$
$$= \sigma^2\{tr(\mathbf{I}) - tr[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}]\}$$
$$= \sigma^2[n - tr(\mathbf{I})]$$
$$= (n-k)\sigma^2$$

第五章

5.1 在异方差性条件下参数估计统计性质的证明

1、参数估计的无偏性仍然成立

设模型为 $$Y_i = \beta_1 + \beta_2 X_i + v_i, \qquad i = 1,2,\cdots,n \tag{1}$$

用离差形式表示 $$y_i = \beta_2 x_i + u_i \qquad (\text{其中} u_i = v_i - \bar{v}) \tag{2}$$

参数 β_2 的估计量 $\hat{\beta}_2$ 为

$$\hat{\beta}_2 = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{\sum x_i(\beta_2 x_i + u_i)}{\sum x_i^2} = \frac{\beta_2\sum x_i^2 + \sum x_i u_i}{\sum x_i^2}$$
$$= \beta_2 + \frac{\sum x_i u_i}{\sum x_i^2} \tag{3}$$

$$E(\hat{\beta}_2) = \beta_2 + E(\frac{\sum x_i u_i}{\sum x_i^2}) = \beta_2 + \frac{\sum E(x_i u_i)}{\sum x_i^2} = \beta_2 \tag{4}$$

在证明中仅用到了假定 $E(x_i u_i)=0$ 。

2、参数估计的有效性不成立

假设（1）式存在异方差，且 $\mathrm{var}(u_i)=\sigma_i^2=\sigma^2 X_i^2$ ，则参数 β_2 的估计 $\hat{\beta}_2$ 的方差为

$$Var(\hat{\beta}_2^*)=E\left[\hat{\beta}_2-E(\hat{\beta}_2)\right]^2=E\left(\hat{\beta}_2-\beta_2\right)^2=E\left(\beta_2+\frac{\sum x_i u_i}{\sum x_i^2}-\beta_2\right)^2$$

$$=E\left(\frac{\sum x_i u_i}{\sum x_i^2}\right)^2=E\left(\frac{\sum_{i=j} x_i^2 u_i^2+2\sum_{i\neq j} x_i x_j u_i u_j}{(\sum x_i^2)^2}\right)=\frac{\sum_{i=j} x_i^2 E(u_i^2)+2\sum_{i\neq j} x_i x_j E(u_i u_j)}{(\sum x_i^2)^2}$$

$$=\frac{\sum_{i=j} x_i^2 E(u_i^2)}{(\sum x_i^2)^2}=\frac{\sum_{i=j} x_i^2\sigma_i^2}{(\sum x_i^2)^2}=\frac{\sigma^2\sum x_i^2 X_i^2}{(\sum x_i^2)^2}=\frac{\sigma}{\sum x_i^2}\cdot\frac{\sum x_i^2 X_i^2}{\sum x_i^2} \tag{5}$$

在上述推导中用了假定 $E(u_i u_j)=0, i\neq j$ 。

下面对（2）式运用加权最小二乘法（WLS）。设权数为 $w_i=\frac{1}{z_i}$ ，对（2）式变换为

$$\frac{y_i}{z_i}=\beta_2\frac{x_i}{z_i}+\frac{u_i}{z_i} \tag{6}$$

可求得参数的估计 $\hat{\beta}_2$，根据本章第四节变量变换法的讨论，这时新的随机误差项 $\frac{u_i}{z_i}$ 为同方差，即 $\mathrm{var}(\frac{u_i}{z_i})=\sigma^2$ ，而 $\hat{\beta}_2$ 的方差为

$$\mathrm{var}(\hat{\beta}_2)_{wls}=\frac{\sigma^2}{\sum\left(\frac{x_i}{z_i}\right)^2} \tag{7}$$

为了便于区别，用 $(\hat{\beta}_2)_{\mathrm{wls}}$ 表示加权最小二乘法估计的 β_2，用 $(\hat{\beta}_2)_{\mathrm{ols}}$ 表示 OLS 法估计的 β_2 。

比较（5）式与（7）式，即在异方差下用 OLS 法得到参数估计的方差与用 WLS 法得到参数估计的方差相比较为

$$\frac{\operatorname{var}(\hat{\beta}_2)_{wls}}{\operatorname{var}(\hat{\beta}_2)_{ols}}=\frac{\dfrac{\sigma^2}{\sum\left(\dfrac{x_i}{z_i}\right)^2}}{\dfrac{\sum x_i^2\sigma_i^2}{\left(\sum x_i^2\right)^2}}=\frac{\dfrac{\sigma^2}{\sum\left(\dfrac{x_i}{z_i}\right)^2}}{\dfrac{\sum x_i^2\sigma^2 z_i^2}{\left(\sum x_i^2\right)^2}}=\frac{\left(\sum x_i^2\right)^2}{\sum\left(\dfrac{x_i}{z_i}\right)^2\left(\sum x_i^2 z_i^2\right)} \tag{8}$$

令 $\frac{x_i}{z_i}=a_i, z_i x_i=b_i$ ，由初等数学知识有 $\frac{\left(\sum ab\right)^2}{\sum a^2\sum b^2}\le 1$ ，因此（10）式右端有

$$\frac{\left(\sum x_i^2\right)^2}{\sum\left(\dfrac{x_i}{z_i}\right)^2\left(\sum x_i^2 z_i^2\right)}\le 1 \tag{9}$$

从而，有

$$\operatorname{var}(\hat{\beta}_2)_{wls}\le \operatorname{var}(\hat{\beta}_2)_{ols}$$

这就证明了在异方差下，仍然用普通最小二乘法所得到的参数估计值的方差不再最小。

5.2 对数变换后残差为相对误差的证明

事实上，设样本回归函数为

$$Y_i=\hat{\beta}_1+\hat{\beta}_2 X_i+e_i \tag{10}$$

其中 $e_i=Y_i-\hat{Y}$ 为残差，取对数后的样本回归函数为

$$\ln Y=\hat{\alpha}_1+\hat{\alpha}_2\ln X+e^* \tag{11}$$

其中残差为 $e^*=\ln Y-\ln\hat{Y}$ ，因此

$$e^*=\ln Y-\ln\hat{Y}=\ln(\frac{Y}{\hat{Y}})=\ln(\frac{\hat{Y}+Y-\hat{Y}}{\hat{Y}})=\ln(1+\frac{Y-\hat{Y}}{\hat{Y}}) \tag{12}$$

对（12）式的右端，依据泰勒展式

$$\ln(1+X)=X-\frac{X^2}{2}+\frac{X^3}{3}-\frac{X^4}{4}+\cdots+(-1)^{n-1}\frac{X^n}{n}+\cdots \tag{13}$$

将（13）式中的 X 用 $\frac{Y-\hat{Y}}{\hat{Y}}$ 替换，则 e^* 可近似地表示为

$$e^*\approx\frac{Y-\hat{Y}}{\hat{Y}} \tag{14}$$

即表明（11）式中的误差项为相对误差。

第六章：

6.1：存在自相关时参数估计值方差的证明

$$Var(\hat{\beta}_2) = E(\hat{\beta}_2 - \beta_2)^2$$

$$= E\left(\frac{\Sigma x_t u_t}{\Sigma x_t^2}\right)^2$$

$$= \left(\frac{1}{\Sigma x_t^2}\right)^2 E(x_1 u_1 + x_2 u_2 + \cdots + x_n u_n)^2$$

$$= \left(\frac{1}{\Sigma x_r^2}\right)^2 E[(x_1^2 u_1^2 + x_2^2 u_2^2 + \cdots + x_n^2 u_n^2)$$

$$+ 2(x_1 x_2 u_1 u_2 + x_1 x_3 u_1 u_3 + \cdots + x_{n-1} x_n u_{n-1} u_n)]$$

$$= \left(\frac{1}{\Sigma x_t^2}\right)^2 [(x_1^2 E(u_1^2) + x_2^2 E(u_2^2) + \cdots + x_n^2 E(u_n^2)$$

$$+ 2\left[x_1 x_2 E(u_1 u_2) + x_1 x_3 E(u_1 u_3) + \cdots + x_{n-1} x_n E(u_{n-1} u_n)\right]$$

$$= \frac{\sigma_u^2}{\Sigma x_t^2} + \frac{2}{(\Sigma x_t^2)^2}[x_1 x_2 \rho \sigma_u^2 + x_1 x_3 \rho^2 \sigma_u^2 + \cdots + x_{n-1} x_n \rho \sigma_u^2]$$

$$= \frac{\sigma_u^2}{\sum_{t=1}^{n} x_t^2}\left(1 + 2\rho \frac{\sum_{t=1}^{n-1} x_t x_{t+1}}{\sum_{t=1}^{n} x_t^2} + 2\rho^2 \frac{\sum_{t=1}^{n-2} x_t x_{t+2}}{\sum_{t=1}^{n} x_t^2} + \cdots + 2\rho^{n-1} \frac{x_1 x_n}{\sum_{t=1}^{n} x_t^2}\right)$$

第九章

9.1 $\hat{\alpha}_2$概率极限性质的证明

$$\underset{n\to\infty}{p\lim}\hat{\alpha}_2=\underset{n\to\infty}{p\lim}\beta_2+\underset{n\to\infty}{p\lim}\beta_3\frac{\sum x_{2i}x_{3i}}{\sum x_{2i}^2}+\underset{n\to\infty}{p\lim}\frac{\sum x_{2i}(u_i-\bar{u})}{\sum x_{2i}^2}$$

$$=\beta_2+\beta_3\frac{\underset{n\to\infty}{p\lim}\frac{1}{n}\sum x_{2i}x_{3i}}{\underset{n\to\infty}{p\lim}\frac{1}{n}\sum x_{2i}^2}+\frac{\underset{n\to\infty}{p\lim}\frac{1}{n}\sum x_{2i}(u_i-\bar{u})}{\underset{n\to\infty}{p\lim}\frac{1}{n}\sum x_{2i}^2}$$

$$=\beta_2+\beta_3\frac{Cov\left(X_{2i},X_{3i}\right)}{Var\left(X_{2i}\right)}+\frac{Cov\left(X_{2i},u_i\right)}{Var\left(X_{2i}\right)}$$

其中：$\frac{1}{n}\sum x_{2i}^2$为 X_2 的样本方差，$\frac{1}{n}\sum x_{2i}x_{3i}$为 X_2 和 X_3 的样本协方差，$\frac{1}{n}\sum x_{2i}(u_i-\bar{u})$为 X_2 和u_i的样本协方差。

9.2 参数$\hat{\alpha}_2$一致性的证明

$$\underset{n\to\infty}{p\lim}\hat{\alpha}_2=\underset{n\to\infty}{p\lim}\beta_2+\underset{n\to\infty}{p\lim}\left(\frac{(\sum x_{3i}^2)(\sum x_{2i}(u_i-\bar{u}))-(\sum x_{2i}x_{3i})(\sum x_{3i}(u_i-\bar{u}))}{\sum x_{2i}^2\sum x_{3i}^2-(\sum x_{2i}\sum x_{3i})^2}\right)$$

$$=\beta_2+\left(\frac{\underset{n\to\infty}{p\lim}\left[(\sum x_{3i}^2)(\sum x_{2i}(u_i-\bar{u}))-(\sum x_{2i}x_{3i})(\sum x_{3i}(u_i-\bar{u}))\right]}{\underset{n\to\infty}{p\lim}\left[\sum x_{2i}^2\sum x_{3i}^2-(\sum x_{2i}\sum x_{3i})^2\right]}\right)$$

$$=\beta_2+\left(\frac{\underset{n\to\infty}{p\lim}(\sum x_{3i}^2)\underset{n\to\infty}{p\lim}(\sum x_{2i}(u_i-\bar{u}))-\underset{n\to\infty}{p\lim}(\sum x_{2i}x_{3i})\underset{n\to\infty}{p\lim}(\sum x_{3i}(u_i-\bar{u}))}{\underset{n\to\infty}{p\lim}\left(\sum x_{2i}^2\sum x_{3i}^2-(\sum x_{2i}\sum x_{3i})^2\right)}\right)$$

$$=\beta_2+\left(\frac{Var\left(X_{3i}\right)Cov\left(X_{2i},u_i\right)-Cov\left(X_{2i},X_{3i}\right)Cov\left(X_{3i},u_i\right)}{Var\left(X_{2i}\right)Var\left(X_{3i}\right)-\underset{n\to\infty}{p\lim}\left(\frac{1}{n}\sum x_{2i}\sum x_{3i}\right)^2}\right)$$

$$\overset{Cov(X_{2i},u_i)=0,Cov(X_{3i},u_i)=0}{=}\beta_2+\left(\frac{Var\left(X_{3i}\right)\times 0-Cov\left(X_{2i},X_{3i}\right)\times 0}{Var\left(X_{2i}\right)Var\left(X_{3i}\right)-\underset{n\to\infty}{p\lim}\left(\frac{1}{n}\sum x_{2i}\sum x_{3i}\right)^2}\right)=\beta_2$$

同理，可证$E(\hat{\alpha}_1)=\beta_1$，$E(\hat{\alpha}_3)=\beta_3=0$；$\underset{n\to\infty}{p\lim}\hat{\alpha}_1=\beta_1$和$\underset{n\to\infty}{p\lim}\hat{\alpha}_3=\beta_3=0$。

9.3 有测量误差模型参数估计结果的推导

$$\hat{\beta} = \frac{\sum x_i y_i}{\sum x_i^2}$$

$$= \frac{\sum [x_i^* + (\omega_i - \overline{\omega})][y_i^* + (\varepsilon_i - \overline{\varepsilon})]}{\sum [x_i^* + (\omega_i - \overline{\omega})]^2}$$

$$= \frac{\sum x_i^* y_i^* + \sum (\omega_i - \overline{\omega}) y_i^* + \sum x_i^* (\varepsilon_i - \overline{\varepsilon}) + \sum (\omega_i - \overline{\omega})(\varepsilon_i - \overline{\varepsilon})}{\sum x_i^{*2} + \sum (\omega_i - \overline{\omega})^2 + 2\sum x_i^* (\omega_i - \overline{\omega})}$$

$$= \frac{\sum x_i^* \left(\beta x_i^* + \left(u_i - \overline{u}\right)\right) + \sum (\omega_i - \overline{\omega}) \left(\beta x_i^* + \left(u_i - \overline{u}\right)\right) + \sum x_i^* (\varepsilon_i - \overline{\varepsilon}) + \sum (\omega_i - \overline{\omega})(\varepsilon_i - \overline{\varepsilon})}{\sum x_i^{*2} + \sum (\omega_i - \overline{\omega})^2 + 2\sum x_i^* (\omega_i - \overline{\omega})}$$

$$= \frac{\beta \sum x_i^{*2} + \sum x_i^* (u_i - \overline{u}) + \beta \sum (\omega_i - \overline{\omega}) x_i^* + \sum (\omega_i - \overline{\omega})(u_i - \overline{u}) + \sum x_i^* (\varepsilon_i - \overline{\varepsilon}) + \sum (\omega_i - \overline{\omega})(\varepsilon_i - \overline{\varepsilon})}{\sum x_i^{*2} + \sum (\omega_i - \overline{\omega})^2 + 2\sum x_i^* (\omega_i - \overline{\omega})}$$

因此，有测量误差模型参数估计的概率极限为

$$p\lim \hat{\beta} = \frac{p\lim \left[\beta \sum x_i^{*2} + \sum x_i^* (u_i - \overline{u}) + \beta \sum (\omega_i - \overline{\omega}) x_i^* + \sum (\omega_i - \overline{\omega})(u_i - \overline{u}) + \sum x_i^* (\varepsilon_i - \overline{\varepsilon}) + \sum (\omega_i - \overline{\omega})(\varepsilon_i - \overline{\varepsilon})\right]}{p\lim \left[\sum x_i^{*2} + \sum (\omega_i - \overline{\omega})^2 + 2\sum x_i^* (\omega_i - \overline{\omega})\right]}$$

$$= \frac{\beta Var\left(X_i^*\right) + Cov\left(X_i^*, u_i\right) + \beta Cov\left(X_i^*, \omega_i\right) + Cov\left(\omega_i, u_i\right) + Cov\left(X_i^*, \varepsilon_i\right) + Cov\left(\omega_i, \varepsilon_i\right)}{Var\left(X_i^*\right) + Var\left(\omega_i\right) + 2Cov\left(X_i^*, \omega_i\right)}$$

$$= \frac{\beta Var X_i^*}{Var X_i^* + \sigma_\omega^2} = \frac{\beta}{1 + \frac{\sigma_\omega^2}{\sigma_{X^*}^2}}$$

第 11 章

11.1 联立方程偏倚的证明

例如，设联立方程模型为

$$C_t=\beta_0+\beta_1 Y_t+u_t \tag{1}$$

$$Y_t=C_t+I_t \tag{2}$$

对(1)式 β_1 的 OLS 估计为：

$$\hat{\beta}_1=\frac{\sum c_t y_t}{\sum y_t^2}=\frac{\sum C_t y_t}{\sum y_t^2}=\frac{\sum(\beta_0+\beta_1 Y_t+u_t)y_t}{\sum y_t^2}=\beta_1+\frac{\sum u_t y_t}{\sum y_t^2} \tag{3}$$

其中利用了 $\sum y_t=0$ 和 $\sum Y_t y_t\Big/\sum y_t^2=1$。对上式两边取期望，得

$$E(\hat{\beta}_1)=\beta_1+E\left(\frac{\sum u_t y_t}{\sum y_t^2}\right)$$

这里的 $E\left(\sum u_t y_t\Big/\sum y_t^2\right)\neq 0$，则 $E(\hat{\beta}_1)\neq\beta_1$，$\hat{\beta}_1$ 是 β_1 的有偏估计。

对（3）式取概率极限，得

$$p\lim(\hat{\beta}_1)=p\lim(\beta_1)+p\lim(\frac{\sum u_t y_t}{\sum y_t^2})=p\lim(\beta_1)+\frac{p\lim(\frac{1}{n}\sum u_t y_t)}{p\lim(\sum y_t^2)} \tag{4}$$

其中：$(\sum u_t y_t)\big/n$ 是 Y 与 u 的样本协方差，其总体协方差为

$$p\lim(\frac{1}{n}\sum u_t y_t)=Cov(Y_t,u_t)=\frac{\sigma^2}{1-\beta_1}$$

$(\sum y_t^2)\big/n$ 是 Y 的样本方差，其总体方差为

$$p\lim(\frac{1}{n}\sum y_t^2)=\sigma_Y^2$$

因此

$$p\lim(\hat{\beta}_1)=\beta_1+\frac{1}{1-\beta_1}\frac{\sigma^2}{\sigma_Y^2}$$

因为 $\frac{\sigma^2}{\sigma_Y^2}\neq 0$，则 $p\lim(\hat{\beta}_1)\neq\beta_1$，这说明 $\hat{\beta}_1$ 不是 β_1 的一致估计。