

§ 2.3 多元线性回归模型的参数估计

Estimation of Multiple Linear Regression Model

- 一、多元线性回归模型
- 二、多元线性回归模型的参数估计
- 三、OLS估计量的统计性质
- 四、参数估计量的方差-协方差矩阵和随机误差项 σ^2 方差的估计
- 五、样本容量问题
- 六、多元线性回归模型实例

一、多元线性回归模型

1、多元线性回归模型的形式

- 由于：
在实际经济问题中，一个变量往往受到多个原因变量的影响；
“从一般到简单”的建模思路。
- 所以，在线性回归模型中的解释变量有多个，至少开始是这样。这样的模型被称为多元线性回归模型。
- 多元线性回归模型参数估计的原理与一元线性回归模型相同，只是计算更为复杂。

- 多元线性回归模型的一般形式为：

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \mu_i \quad i=1,2,\cdots,n$$

(2.3.1)

其中：k 为解释变量的数目；

习惯上把常数项看成为一个虚变量的系数，在参数估计过程中该虚变量的样本观测值始终取1。这样：

模型中解释变量的数目为 **(k+1)**。

- 多元线性回归模型的矩阵表达式为：

$$\mathbf{Y} = \mathbf{XB} + \mathbf{N} \quad (2.3.2)$$

其中

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{k1} \\ 1 & x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{kn} \end{bmatrix}_{n \times (k+1)}$$

$$\mathbf{B} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}_{(k+1) \times 1} \quad \mathbf{N} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix}_{n \times 1}$$

2、多元线性回归模型的基本假定

模型(2.3.1)或(2.3.2)在满足下述所列的基本假设的情况下，可以采用普通最小二乘法（OLS）估计参数。

关于经典回归模型的假定

标量符号

1、解释变量 X_1, X_2, \dots, X_k 是非随机的或固定的；而且各 X 之间互不相关（无多重共线性(**no multicollinearity**)）

矩阵符号

1、 $n \times (k + 1)$ 矩阵 \mathbf{X} 是非随机的；且 \mathbf{X} 的秩 $\rho(\mathbf{X}) = k + 1$ ，此时 $\mathbf{X}^T \mathbf{X}$ 也是满秩的

标量符号

2、随机误差项具有零均值、同方差及不序列相关

$$E(\mu_i) = 0 \quad i = 1, 2, \dots, n$$

$$Var(\mu_i) = E(\mu_i^2) = \sigma^2 \quad i = 1, 2, \dots, n$$

$$Cov(\mu_i, \mu_j) = E(\mu_i \mu_j) = 0 \quad i \neq j$$

矩阵符号

2、 $E(N) = 0, \quad E(NN^T) = \sigma^2 I$

$$E(N) = E \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = \begin{pmatrix} E(\mu_1) \\ \vdots \\ E(\mu_n) \end{pmatrix} = 0$$

$$E(NN^T) = E \left(\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} (\mu_1 \quad \cdots \quad \mu_n) \right) = E \begin{pmatrix} \mu_1^2 & \cdots & \mu_1 \mu_n \\ \vdots & \ddots & \vdots \\ \mu_n \mu_1 & \cdots & \mu_n^2 \end{pmatrix} = \begin{pmatrix} \sigma^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma^2 \end{pmatrix} = \sigma^2 I$$

标量符号

3、解释变量与随机项不相关

$$\text{Cov}(X_{ji}, \mu_i) = 0 \quad i = 1, 2, \dots, n$$

矩阵符号

3、 $E(X^T N) = 0$ ，即

$$E \begin{pmatrix} \sum \mu_i \\ \sum X_{1i} \mu_i \\ \vdots \\ \sum X_{Ki} \mu_i \end{pmatrix} = \begin{pmatrix} \sum E(\mu_i) \\ \sum X_{1i} E(\mu_i) \\ \vdots \\ \sum X_{Ki} E(\mu_i) \end{pmatrix} = 0$$

标量符号

4、（为了假设检验），随机扰动项服从正态分布

$$\mu_i \sim N(0, \sigma^2) \quad i = 1, 2, \dots, n$$

矩阵符号

4、向量 N 为一多维正态分布，即

$$N \sim N(0, \sigma^2 I)$$

二、多元线性回归模型的参数估计

1、普通最小二乘估计

- 普通最小二乘估计

随机抽取被解释变量和解释变量的 n 组样本观测值:

$$(Y_i, X_{ji}), \quad i = 1, 2, \dots, n, \quad j = 0, 1, 2, \dots, k$$

如果模型的参数估计值已经得到, 则有:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \dots + \hat{\beta}_{ki} X_{Ki} \quad i=1, 2, \dots, n$$

(2.3.3)

根据最小二乘原理，参数估计值应该是下列方程组的解：

$$\begin{cases} \frac{\partial}{\partial \hat{\beta}_0} Q = 0 \\ \frac{\partial}{\partial \hat{\beta}_1} Q = 0 \\ \frac{\partial}{\partial \hat{\beta}_2} Q = 0 \\ \vdots \\ \frac{\partial}{\partial \hat{\beta}_k} Q = 0 \end{cases} \quad (2.3.4)$$

其中

$$\begin{aligned} Q &= \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \\ &= \sum_{i=1}^n (Y_i - (\hat{\beta}_0 + \hat{\beta}_1 Y_{1i} + \hat{\beta}_2 Y_{2i} + \cdots + \hat{\beta}_k Y_{ki}))^2 \end{aligned} \quad (2.3.5)_{12}$$

于是，得到关于待估参数估计值的**正规方程组**：

$$\left\{ \begin{array}{l} \Sigma(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \cdots + \hat{\beta}_k X_{ki}) = \Sigma Y_i \\ \Sigma(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \cdots + \hat{\beta}_k X_{ki}) X_{1i} = \Sigma Y_i X_{1i} \\ \Sigma(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \cdots + \hat{\beta}_k X_{ki}) X_{2i} = \Sigma Y_i X_{2i} \\ \vdots \\ \Sigma(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \cdots + \hat{\beta}_k X_{ki}) X_{ki} = \Sigma Y_i X_{ki} \end{array} \right. \quad (2.3.6)$$

解该（k+1）个方程组成的线性代数方程组，即可得到（k+1）个待估参数的估计值 $\hat{\beta}_j, j = 0, 1, 2, \dots, k$ 。

(2.3.6) 的矩阵形式如下：

$$\begin{pmatrix} n & \sum X_{1i} & \cdots & \sum X_{ki} \\ \sum X_{1i} & \sum X_{1i}^2 & \cdots & \sum X_{1i}X_{ki} \\ \cdots & \cdots & \cdots & \cdots \\ \sum X_{ki} & \sum X_{ki}X_{1i} & \cdots & \sum X_{ki}^2 \end{pmatrix} \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \cdots \\ \hat{\beta}_k \end{pmatrix} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ X_{11} & X_{12} & \cdots & X_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ X_{k1} & X_{k2} & \cdots & X_{kn} \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ \cdots \\ Y_n \end{pmatrix}$$

即：

$$\mathbf{X}'\mathbf{X}\hat{\mathbf{B}} = \mathbf{X}'\mathbf{Y} \quad (2.3.7)$$

由于 $\mathbf{X}'\mathbf{X}$ 满秩，故有

$$\hat{\mathbf{B}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y} \quad (2.3.8)$$

- 估计过程的矩阵表示:

对于模型(2.3.3)式有:

$$\hat{Y} = X\hat{B}$$

被解释变量的观测值与估计值之差的平方和为:

$$\begin{aligned} Q &= \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \\ &= \mathbf{e}'\mathbf{e} = (\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}})'(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}) \end{aligned}$$

其中

$$\mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

根据最小二乘原理，参数估计值应该是下列方程组的解：

$$\frac{\partial}{\partial \hat{\mathbf{B}}} (\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}})'(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}) = 0$$

求解过程如下：

$$\frac{\partial}{\partial \hat{\mathbf{B}}} (\mathbf{Y}' - \hat{\mathbf{B}}'\mathbf{X}')(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}) = 0$$

$$\frac{\partial}{\partial \hat{\mathbf{B}}} (\mathbf{Y}'\mathbf{Y} - \hat{\mathbf{B}}'\mathbf{X}'\mathbf{Y} - \mathbf{Y}'\mathbf{X}\hat{\mathbf{B}} + \hat{\mathbf{B}}'\mathbf{X}'\mathbf{X}\hat{\mathbf{B}}) = 0$$

$$\frac{\partial}{\partial \hat{\mathbf{B}}} (\mathbf{Y}'\mathbf{Y} - 2\mathbf{Y}'\mathbf{X}\hat{\mathbf{B}}' + \hat{\mathbf{B}}'\mathbf{X}'\mathbf{X}\hat{\mathbf{B}}) = 0$$

$$\frac{\partial}{\partial \hat{\mathbf{B}}} (\mathbf{Y}'\mathbf{Y} - 2(\mathbf{X}'\mathbf{Y})'\hat{\mathbf{B}}' + \hat{\mathbf{B}}'\mathbf{X}'\mathbf{X}\hat{\mathbf{B}}) = 0$$

$$-\mathbf{X}'\mathbf{Y} + \mathbf{X}'\mathbf{X}\hat{\mathbf{B}} = 0$$

即得到

$$\mathbf{X}'\mathbf{Y} = \mathbf{X}'\mathbf{X}\hat{\mathbf{B}}$$

于是，参数的最小二乘估计值为：

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

2、最大或然估计

- Y的随机抽取的n组样本观测值的联合概率

$$L(\hat{B}, \sigma_{\mu}^2) = P(y_1, y_2, \dots, y_n)$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}} \sigma_{\mu}^n} e^{-\frac{1}{2\sigma_{\mu}^2} \sum (y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \dots + \hat{\beta}_k x_{ki}))^2}$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}} \sigma_{\mu}^n} e^{-\frac{1}{2\sigma_{\mu}^2} (Y - X\hat{B})'(Y - X\hat{B})}$$

- 对数或然函数为

$$L^* = Ln(L)$$

$$= -nLn(\sqrt{2\pi}\sigma_{\mu}) - \frac{1}{2\sigma_{\mu}^2} (Y - X\hat{B})' (Y - X\hat{B})$$

- 参数的最大或然估计

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

- 结果与参数的普通最小二乘估计相同

3、矩估计 (Moment Method, MM)

- 用每个解释变量分别乘以模型的两边，并对所有样本点求和，即得到：

$$\left\{ \begin{array}{l} \Sigma y_i = \Sigma(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \mu_i) \\ \Sigma y_i x_{1i} = \Sigma(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \mu_i) x_{1i} \\ \Sigma y_i x_{2i} = \Sigma(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \mu_i) x_{2i} \\ \vdots \\ \Sigma y_i x_{ki} = \Sigma(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \mu_i) x_{ki} \end{array} \right.$$

- 对每个方程的两边求期望，有：

$$\left\{ \begin{array}{l} E(\Sigma y_i) = E(\Sigma(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \mu_i)) \\ E(\Sigma y_i x_{1i}) = E(\Sigma(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \mu_i) x_{1i}) \\ E(\Sigma y_i x_{2i}) = E(\Sigma(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \mu_i) x_{2i}) \\ \vdots \\ E(\Sigma y_i x_{ki}) = E(\Sigma(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \mu_i) x_{ki}) \end{array} \right.$$

- 得到一组矩条件

$$\left\{ \begin{array}{l} \Sigma y_i = \Sigma(\hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \cdots + \hat{\beta}_k x_{ki}) \\ \Sigma y_i x_{1i} = \Sigma(\hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \cdots + \hat{\beta}_k x_{ki}) x_{1i} \\ \Sigma y_i x_{2i} = \Sigma(\hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \cdots + \hat{\beta}_k x_{ki}) x_{2i} \\ \vdots \\ \Sigma y_i x_{ki} = \Sigma(\hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \cdots + \hat{\beta}_k x_{ki}) x_{ki} \end{array} \right.$$

- 求解这组矩条件，即得到参数估计量
- 与**OLS**、**ML**估计量等价

- 矩方法是工具变量方法(**Instrumental Variables, IV**)和广义矩估计方法(**Generalized Moment Method, GMM**)的基础
- 在矩方法中关键是利用了

$$E(\mu_i x_{ji}) = 0, j = 1, 2, \dots, k$$

- 如果某个解释变量与随机项相关，只要能找到1个工具变量，仍然可以构成一组矩条件。这就是**IV**。
- 如果存在 $> k+1$ 个变量与随机项不相关，可以构成一组包含 $> k+1$ 方程的矩条件。这就是**GMM**。

4、多元回归方程及偏回归系数的含义

在经典回归模型的诸假定下，式（2.3.1）两边对 Y 求条件期望得：

$$E(Y_i | X_{1i}, X_{2i}, \dots, X_{ki}) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} \quad (2.3.9)$$

称为多元回归方程（函数）。

多元回归分析（multiple regression analysis）是以多个解释变量的固定值为条件的回归分析，并且所获得的是诸变量X值固定时Y的平均值。诸 β_i 称为**偏回归系数（partial regression coefficients）**。

- 偏回归系数的含义如下：

β_1 度量着在 X_2, X_3, \dots, X_k 保持不变的情况下， X_1 每变化1个单位时， Y 的均值 $E(Y)$ 的变化，或者说 β_1 给出 X_1 的单位变化对 Y 均值的“直接”或“净”（不含其他变量）影响。

其他参数的含义与之相同。

三、OLS估计量的统计性质

1、线性性

$$\hat{B} = (X'X)^{-1} X'Y \quad (2.3.10)$$

2、无偏性

$$E(\hat{B}) = B \quad (2.3.11)$$

证： $\hat{B} = (X'X)^{-1} X'Y = (X'X)^{-1} X'(XB - N) = B - (X'X)^{-1} X'N$

$$(2.3.12)$$

于是： $E(\hat{B}) = E(B) - E((X'X)^{-1} X'N) = B - (X'X)^{-1} X'E(N) = B$

3、最小方差性

若 B^* 是 B 的任一线性无偏估计量, 则有

$$E[(B^* - B)(B^* - B)'] \geq E[(\hat{B} - B)(\hat{B} - B)'] \quad (2.3.13)$$

证明略。

四、参数估计量的方差-协方差矩阵 和随机误差项 σ^2 方差的估计

1、一个疑问与回答

- 疑问：在无偏性证明中将参数估计量看作随机量，而在正规方程组的推导中又将它看作确定值。如何解释？
- 解释：将一组具体样本资料代入参数估计量的表达式给出的参数估计结果是一个“估计值”，或者“点估计”，是参数估计量的一个具体数值，是确定的；但从另一个角度，仅仅把它看成是参数估计量的一个表达式，那么，则是被解释变量观测值的函数，而被解释变量是随机变量，所以参数估计量也是随机变量，在这个角度上，称之为“估计量”。

2、参数估计量的方差-协方差

- 将参数估计量看作随机量，具有数字特征。
- 参数估计量的方差以及不同参数估计量之间的协方差在模型理论中具有重要性。
- 具体描述如下：

由于矩阵

$$\begin{aligned}
 E((\hat{B} - B)(\hat{B} - B)') &= E \left(\begin{pmatrix} \hat{\beta}_0 - \beta_0 \\ \hat{\beta}_1 - \beta_1 \\ \vdots \\ \hat{\beta}_k - \beta_k \end{pmatrix} \begin{pmatrix} \hat{\beta}_0 - \beta_0 & \hat{\beta}_1 - \beta_1 & \cdots & \hat{\beta}_k - \beta_k \end{pmatrix} \right) \\
 &= \begin{pmatrix} E(\hat{\beta}_0 - \beta_0)^2 & E(\hat{\beta}_0 - \beta_0)(\hat{\beta}_1 - \beta_1) & \cdots & E(\hat{\beta}_0 - \beta_0)(\hat{\beta}_k - \beta_k) \\ E(\hat{\beta}_1 - \beta_1)(\hat{\beta}_0 - \beta_0) & E(\hat{\beta}_1 - \beta_1)^2 & \cdots & E(\hat{\beta}_1 - \beta_1)(\hat{\beta}_k - \beta_k) \\ \cdots & \cdots & \cdots & \cdots \\ E(\hat{\beta}_k - \beta_k)(\hat{\beta}_0 - \beta_0) & E(\hat{\beta}_k - \beta_k)(\hat{\beta}_1 - \beta_1) & \cdots & E(\hat{\beta}_k - \beta_k)^2 \end{pmatrix} \\
 &\quad (3.1.14)
 \end{aligned}$$

主对角线给出了各参数估计 $\hat{\beta}_j$ 的方差，其余部分给出了不同参数估计 $\hat{\beta}_i$ 与 $\hat{\beta}_j$ 的协方差，故称为参数估计向量 \hat{B} 的**方差-协方差矩阵**。

由 (2.3.8) 式可得 $\hat{\mathbf{B}}$ 的方差-协方差矩阵的矩阵符号表达式:

$$\begin{aligned}
 \text{var} - \text{cov}(\hat{\mathbf{B}}) &= E[(\hat{\mathbf{B}} - \mathbf{B})(\hat{\mathbf{B}} - \mathbf{B})'] \\
 &= E\{[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} - \mathbf{B}][(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} - \mathbf{B}']\} \\
 &= E\{[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\mathbf{B} + \mathbf{N}) - \mathbf{B}][(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\mathbf{B} + \mathbf{N}) - \mathbf{B}']\} \\
 &= E\{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{N}[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{N}']\} = E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{N}\mathbf{N}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}] \\
 &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{E}(\mathbf{N}\mathbf{N}')\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \cdot \mathbf{I} \cdot \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \\
 &= \sigma^2(\mathbf{X}'\mathbf{X})^{-1}
 \end{aligned} \tag{2.3.15}$$

记 c_{ij} 为矩阵 $(X'X)^{-1}$ 中第 i 行第 j 列元素，比较 (3.1.14) 与 (3.1.15) 式，知第 i 个回归参数估计量 $\hat{\beta}_i$ ($i=0, 1, 2, \dots, k$) 的方差、标准差、协方差为：

$$\text{var}(\hat{\beta}_i) = \sigma^2 c_{ii} \quad (2.3.16)$$

$$\sqrt{\text{var}(\hat{\beta}_i)} = \sigma \sqrt{c_{ii}} \quad (2.3.17)$$

$$\text{cov}(\hat{\beta}_i, \hat{\beta}_j) = \sigma^2 c_{ij} \quad (2.3.18)$$

3、随机误差项 σ^2 方差的估计

可以证明，随机误差项方差 σ^2 的无偏估计为：

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{n - (k + 1)} = \frac{\mathbf{e}'\mathbf{e}}{n - k - 1} = \frac{\mathbf{Y}'\mathbf{Y} - \hat{\mathbf{B}}'\mathbf{X}'\mathbf{Y}}{n - k - 1} \quad (2.3.19)$$

于是，参数估计量 $\hat{\beta}_i$ 的方差、标准差、协方差可分别用其样本方差 $s^2(\hat{\beta}_i)$ 、标准差 $s(\hat{\beta}_i)$ 、协方差 $ss(\hat{\beta}_i, \hat{\beta}_j)$ 加以估计：

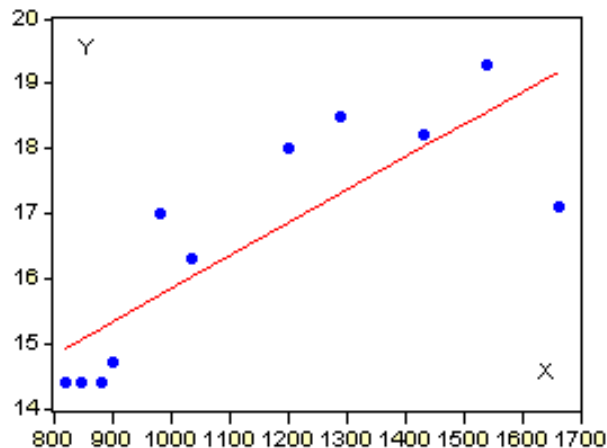
$$s^2(\hat{\beta}_i) = \hat{\sigma}^2 c_{ii} = \frac{\sum e_i^2}{n - k - 1} c_{ii} \quad (2.3.20)$$

$$s(\hat{\beta}_i) = \sqrt{\hat{\sigma}^2 c_{ii}} = \sqrt{\frac{\sum e_i^2}{n - k - 1} c_{ii}} \quad (2.3.21)$$

$$ss(\hat{\beta}_i, \hat{\beta}_j) = \hat{\sigma}^2 c_{ij} = \frac{\sum e_i^2}{n - k - 1} c_{ij} \quad i \neq j \quad (2.3.22)$$

例 人均鲜蛋需求量Y与人均可支配收入X关系

(file: li-2-1)



Dependent Variable: Y
Method: Least Squares
Date: 02/12/07 Time: 08:46
Sample: 1988 1998
Included observations: 11

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	10.76616	1.396736	7.708087	0.0000
X	0.005069	0.001183	4.283328	0.0020
R-squared	0.670895	Mean dependent var	16.57273	
Adjusted R-squared	0.634328	S.D. dependent var	1.845042	
S.E. of regression	1.115713	Akaike info criterion	3.219829	
Sum squared resid	11.20333	Schwarz criterion	3.292174	
Log likelihood	-15.70906	F-statistic	18.34690	
Durbin-Watson stat	1.320391	Prob(F-statistic)	0.002040	

估计结果:

$$\hat{\sigma} = 1.1157, \hat{\sigma}^2 = 1.1157^2 = 1.2448$$

$$\hat{Var}(\hat{\beta}_0) = 1.3967^2 = 1.9508$$

$$\hat{Var}(\hat{\beta}_1) = 0.0012^2 = 0.00000144$$

关于 $\mathbf{e}'\mathbf{e} = \mathbf{Y}'\mathbf{Y} - \hat{\mathbf{B}}'\mathbf{X}'\mathbf{Y}$ 的证明:

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

$$\mathbf{e}'\mathbf{e} = (\mathbf{M}\mathbf{Y})'(\mathbf{M}\mathbf{Y}) = \mathbf{Y}'\mathbf{M}'\mathbf{M}\mathbf{Y}$$

可以证明, $\mathbf{M} = (\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')$ 为对称等幂矩阵,

即 $\mathbf{M}' = \mathbf{M}$, $\mathbf{M}^2 = \mathbf{M}$, 于是:

$$\begin{aligned}\mathbf{e}'\mathbf{e} &= \mathbf{Y}'\mathbf{M}\mathbf{Y} = \mathbf{Y}'[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{Y} \\ &= \mathbf{Y}'\mathbf{Y} - \mathbf{Y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \mathbf{Y}'\mathbf{Y} - \mathbf{Y}'\mathbf{X}\hat{\mathbf{B}}\end{aligned}$$

由于 $\mathbf{Y}'\mathbf{X}\hat{\mathbf{B}}$ 为一数量, 故 $\mathbf{Y}'\mathbf{X}\hat{\mathbf{B}} = (\mathbf{Y}'\mathbf{X}\hat{\mathbf{B}})' = \hat{\mathbf{B}}'\mathbf{X}'\mathbf{Y}$,

于是:

$$\mathbf{e}'\mathbf{e} = \mathbf{Y}'\mathbf{Y} - \hat{\mathbf{B}}'\mathbf{X}'\mathbf{Y}$$

五、样本容量问题

1. 最小样本容量

- 所谓“最小样本容量”，即从最小二乘原理和最大或然原理出发，欲得到参数估计量，不管其质量如何，所要求的样本容量的下限。
- 样本最小容量必须不少于模型中解释变量的数目（包括常数项）。

2、满足基本要求的样本容量

- 从参数估计角度： $>3 \times$ 解释变量数目
- 从检验的有效性角度： >30

3、模型的良好性质只有在大样本下才能得到理论上的证明

六、多元线性回归模型实例

中国消费函数模型

- 根据消费模型的一般形式，选择消费总额为被解释变量，国内生产总值和前一年的消费总额为解释变量，变量之间关系为简单线性关系，选取1981年至1996年统计数据为样本观测值。

• 中国消费数据表 单位：亿元

年 份	消费总额	国内生产总值	前一年消费额	年 份	消费总额	国内生产总值	前一年消费额
1981	3309	4901	2976	1989	10556	16466	9360
1982	3638	5489	3309	1990	11362	18320	10556
1983	4021	6076	3638	1991	13146	21280	11362
1984	4694	7164	4021	1992	15952	25864	13146
1985	5773	8792	4694	1993	20182	34501	15952
1986	6542	10133	5773	1994	27216	47111	20182
1987	7451	11784	6542	1995	34529	59405	27216
1988	9360	14704	7451	1996	40172	68498	34529

• 模型估计结果

Dependent Variable: CONS

Method: Least Squares

Date: 02/25/03 Time: 17:47

Sample: 1981 1996

Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	540.5286	84.30153	6.411848	0.0000
GDP	0.480948	0.021861	22.00035	0.0000
CONS1	0.198545	0.047409	4.187969	0.0011
R-squared	0.999773	Mean dependent var	13618.94	
Adjusted R-squared	0.999739	S.D. dependent var	11360.47	
S.E. of regression	183.6831	Akaike info criterion	13.43166	
Sum squared resid	438613.2	Schwarz criterion	13.57652	
Log likelihood	-104.4533	F-statistic	28682.51	
Durbin-Watson stat	1.450101	Prob(F-statistic)	0.000000	

- 拟合效果

