

数量经济学原理 与 经济控制论

张启人 著

(中 册)

湖南财经学院科研处

一九八三年

目 录

· 中 册 ·

第五章 多重回归模型	5—1
------------	-----

§ 5.1 三变量模型	5—1
-------------	-----

§ 5.2 多重回归模型	5—5
--------------	-----

§ 5.3 统计量	5—8
-----------	-----

例题和习题	5—43
-------	------

第六章 计量经济学中若干理论与应用问题	6—1
---------------------	-----

§ 6.1 预 测	6—1
-----------	-----

§ 6.2 时间延迟	6—14
------------	------

§ 6.3 联立方程模型	6—32
--------------	------

§ 6 · 4 计算机仿真简介	6-52
-----------------	------

附 录 经济系统的计算机仿真	附 6-1
----------------	-------

参考文献	附 6-11
------	--------

第七章 经济控制论基本概念	7-1
---------------	-----

§ 7 · 1 经济系统最控制理论的发展	7-1
----------------------	-----

§ 7 · 2 非线性计量经济模型	7-5
-------------------	-----

§ 7 · 3 状态空间模型	7-11
----------------	------

§ 7 · 4 宏观经济最优化和最优控制	7-13
----------------------	------

§ 7 · 5 投入-产出模型与能源动态预测	7-25
------------------------	------

§ 7 · 6 投入产出优化模型	7-28
------------------	------

§ 7 · 7 动态投入产出模型	7-32
------------------	------

§ 7 · 8 动态预测	7-37
--------------	------

第五章 多重回归模型

§5.1 三变量模型

二变量回归模型易于通过多重回归的途径推广到两个以上变量的情形，相应的模型也叫多元回归模型。这里，设 Y 是 $k-1$ 个自变量的线性函数加上误差项：

$$Y = \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \epsilon \quad (5.1.1)$$

注意 X_1 等于1，因 β_1 是常数项。因此 X_1 并非实际变量，自变量数为 $k-1$ 。引进 X_1 只是为了往后写成矩阵形式。

多元回归模型的假定与二元相同，除了必须补充假定有变量多于一个。在这一前提下，必须假定：

⑧ 完全多重共线性 (Perfect Multicollinearity) 或完全共线性在各自变量当中不存在。若一变量是另一个的线性组合，则存在完全多重共线性。例如，若 $X_2 = 5X_3$ ，则 X_2 和 X_3 互为线性组合，在存在完全多重共线性的情况下，不可能估计多元回归方程。如果是 $X_2 = X_3^2$ 呢？由于 X_2 与 X_3 并不互为线性组合，估计才成为可能。完全多重共线性的问题可以通过从模型取消完全共线性变量之一来得到解决。虽然完全多重共线性比较罕见，“不完全多重共线性”或仅为“多重共线性”的问题则往往在多元回归模型中是个棘手的问题。

现在的问题在于估计参数 $(\beta_1, \cdots, \beta_k)$ 使得极小化误差的平方和。计算机发明之前，这个问题实际上是不可解的。由于变量和观测数增加，计算越来越复杂。

下面先讨论三变量模型的代数形式。

三变量模型表成：

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \epsilon \quad (5.1.2)$$

其中 β_1 为常数项； β_2, β_3 为自变量 X_2 和 X_3 的参数。我们想极小化的误差平方和是：

$$\min \sum e^2 = \sum (Y - \beta_1 - \beta_2 X_2 - \beta_3 X_3)^2 \quad (5.1.3)$$

这里 β_1, β_2 和 β_3 为回归系数，即总体参数的样本估计值。类似前节补注1，普通微积分知识即能解出：

$$\beta_2 = \frac{(\sum y X_2)(\sum X_3^2) - (\sum y X_3)(\sum X_2 X_3)}{D} \quad (5.1.4)$$

$$\beta_3 = \frac{(\sum y X_3)(\sum X_2^2) - (\sum y X_2)(\sum X_2 X_3)}{D} \quad (5.1.5)$$

$$\beta_1 = \bar{Y} - \beta_2 \bar{X}_2 - \beta_3 \bar{X}_3 \quad (5.1.6)$$

其中分母

$$D = (\sum X_2^2)(\sum X_3^2) - (\sum X_2 X_3)^2 \quad (5.1.7)$$

作为例子，设一名经济学家想要确立一个“技能资本”(Human Model) 模型，其中因变量为收益或收入，教育年限和IQ为自变量。收入用 Y 表示，IQ用 X_2 、教育年限用 X_3 表示。这个模型依据的基本理论是IQ和教育水平都对个人收入有影响。假想数据如表 5.1.1。

计算结果如图 5.1.1 所给。(5.1.4) (5.1.5) 和 (5.1.7) 中的小字母表示偏离平均值的偏差。因此， $\sum y X_2 = 274.500$ ； $\sum y X_3 = 80.500$ ； $\sum X_2^2 = 1.238.5$ ； $\sum X_3^2 = 88.1$ ； $\sum X_2 X_3 = 252.5$ ；和 $D = 45.355.6$ 。因此

$$B_2 = \frac{(274.500)(88.1) - (80.500)(252.5)}{45.355 \cdot 6} = 85.044$$

表 5.1.1 收入—IQ—教育数据

收入 (Y)	IQ (X_2)	教育 (X_3)
5,000	80	9
6,000	95	8
7,000	100	10
8,000	101	10
9,000	103	11
10,000	115	14
11,000	105	15
12,000	116	13
13,000	120	16
14,000	110	17

YX_2	YX_3	X_2^2	X_3^2	X_2X_3
110250	14850	6400.25	10.89	80.85
33250	15050	9025	18.49	40.85
11250	5750	2025	5.29	10.35
5250	3450	1225	5.29	8.05
750	650	225	1.69	1.95
5250	850	11025	2.89	17.85
750	4050	25	7.29	1.35
28750	1750	13225	4.49	8.05
54250	12950	24025	13.69	57.35
24750	21150	3025	22.09	25.85
274500 = SUM, 80500 = SUM, 1238.5 = SUM, 88.1 = SUM, 252.5 = SUM				
45355.6 = D				

图 5.1.1 三变量模型的计算

$$B_3 = \frac{(80,500)(1,238.5) - (274,500)(252.5)}{45,355.6} = 669.992$$

$$B_1 = 9,500 - 85.044(104.5) - 669.992(12.3) = 7,628.13$$

B_2 和 B_3 的标准误差可用下列公式计称。由于常数项 (B_1) 标准误差的计称比较复杂，留待下一小节再交待，即表成矩阵形式。无论如何，计量经济学研究中常数项的显著性检验往往是无关紧要的。

$$SE(B_2) = S \sqrt{\frac{\sum x_3^2}{D}} \quad (5.1.8)$$

$$SE(B_3) = S \sqrt{\frac{\sum x_2^2}{D}} \quad (5.1.9)$$

其中：

$$S = \sqrt{\frac{\sum e^2}{N-3}} \quad (5.1.10)$$

$$\sum e^2 = \sum y^2 - B_2 \sum yx_2 - B_3 \sum yx_3 \quad (5.1.11)$$

用 (5.1.10) 计称“估计值的标准误差”，其平方是回归直线“方差”的估计。在论及预测时估计值的标准误差是很重要的。代入 (5.1.11) 和 (5.1.10)，估计值的标准误差等于 863.63。从 (5.1.8) 和 (5.1.9) 得到的 B_2 与 B_3 的标准误差分别是 38.06 和 142.71。

可由下列公式导出“多重可决系数”：

$$R^2 = \frac{B_2 \sum yx_2 + B_3 \sum yx_3}{\sum y^2} \quad (5.1.12)$$

注意 $\sum y^2$ 已在 § 4.3 称过，且等于 82,500,000。因此

$$R^2 = \frac{(85.04)(274,500) + (669.99)(80,500)}{82,500,000} = .937$$

§5.2 多重回归模型

按前节所述收入—IQ—教育的“技能资本”模型，写成矩阵形式是：

$$Y = XB + E \quad (5.2.1)$$

式中

$$Y = [5000 \ 6000 \ 7000 \ 8000 \ 9000 \ 10000 \ 11000 \ 12000 \ 13000 \ 14000]^T$$

$$X = \begin{bmatrix} 1 & 80 & 9 \\ 1 & 95 & 8 \\ 1 & 100 & 10 \\ 1 & 101 & 10 \\ 1 & 103 & 11 \\ 1 & 115 & 14 \\ 1 & 105 & 15 \\ 1 & 116 & 13 \\ 1 & 120 & 16 \\ 1 & 110 & 17 \end{bmatrix}$$

回归系数阵 $B = [B_1 \ B_2 \ B_3]^T$ ，可以证明：

$$B = (X^T X)^{-1} X^T Y \quad (5.2.2)$$

其中在本例中：

$$X^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 80 & 95 & 100 & 101 & 103 & 115 & 105 & 116 & 120 & 110 \\ 9 & 8 & 10 & 10 & 11 & 14 & 15 & 13 & 16 & 17 \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 10 & 1045 & 123 \\ 1045 & 110441 & 13106 \\ 123 & 13106 & 1601 \end{bmatrix}$$

$$(\Sigma X)^{-1} = \begin{bmatrix} 11.13160227 & -0.13450820 & 0.24589466 \\ -0.13450820 & 0.00194243 & -0.00556712 \\ 0.24589466 & -0.00556712 & 0.02730644 \end{bmatrix}$$

$$X^T Y = \begin{bmatrix} 95000 \\ 10202000 \\ 1249000 \end{bmatrix}$$

用 $X^T Y$ 乘 $(X^T X)^{-1}$, 得 B 向量:

$$B = \begin{bmatrix} -7628.13 \\ 85.04 \\ 669.99 \end{bmatrix}$$

在 §4.3 中, 已称出来包括教育在内的同一数值例的 R^2 值, 手续仍相同. 首先称出 $Y_c = -7,628.13 + 85.04 X_2 + 669.99 X_3$, 然后称出 Σe^2 再用 Σy^2 除之. 最后用 1 减去后一结果即得 R^2 . 计算结果如表 5.2.1.

表 5.2.1 R^2 和误差阵 (E) 的计算

Y	Y_c	$(Y - Y_c) = E$	$(Y - Y_c)^2$
\$ 5,000	5,205.45	-205.45	42,209.70
6,000	5,811.11	188.89	35,679.43
7,000	7,576.31	-576.31	332,133.21
8,000	7,661.36	338.64	114,677.04
9,000	8,501.44	498.56	248,562.07
10,000	11,531.94	-1,531.94	2,346,844.10
11,000	11,351.51	-351.51	123,559.28
12,000	10,947.00	1,053.00	1,108,809.00
13,000	13,297.15	-297.15	88,298.12
14,000	13,116.72	883.28	780,183.55
			5,220,951.30 = Σe^2

$$R^2 = 1 - \frac{\Sigma e^2}{\Sigma y^2} = 1 - \frac{5,220,951.3}{82,500,000} = 0.937$$

至此，可以直接用(5.1.10) (不用(5.1.11)式)计算估计值的标准误差：

$$S = \sqrt{\frac{\sum e^2}{N-3}} = \sqrt{\frac{5,220,951.3}{7}} = 863.63$$

用 $(X^T X)^{-1}$ 对应主对角元素的平方根乘估计值的标准误差可得回归系数的标准误差

$$SE(B_1) = 863.63 \sqrt{11.1316} = 2,881.42$$

$$SE(B_2) = 863.63 \sqrt{.0019} = 38.06$$

$$SE(B_3) = 863.63 \sqrt{.0273} = 142.71$$

推广到任意的变量和观测数，令 N 为观测数， K 为变量数（自变量加因变量），由(5.1.1)确定的多重回归模型可写成矩阵形式：

$$Y = X\beta + \epsilon \quad (5.2-3)$$

其中 $Y (N \times 1)$ ； $X (N \times K)$ ， $\beta (K \times 1)$ ， $\epsilon (N \times 1)$ ：

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix} \quad X = \begin{bmatrix} 1 & X_{21} & \cdots & X_{K1} \\ 1 & X_{22} & \cdots & X_{K2} \\ \vdots & \vdots & & \vdots \\ 1 & X_{2N} & \cdots & X_{KN} \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_K \end{bmatrix} \quad \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_N \end{bmatrix}$$

β 和 ϵ 的最小二乘估计是：

$$B = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_K \end{bmatrix} = (X^T X)^{-1} X^T Y$$

$$E = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{bmatrix} = Y - XB$$

$$S = \sqrt{\frac{\sum e^2}{N-K}}$$

§5.3 统计量

1. 偏回归系数

图 5.3.1 表示了计算机打印输出, 包括回归系数、估计值的标准误差、每一回归系数的标准误差, 以及它们的 t 值、 R^2 、修正 R^2 、整个回归的 F 值、自由度、Durbin-Watson 统计量和相关矩阵。首先讨论回归系数。

REGRESSION COEFFICIENTS

-7628.13

85.044

669.992

STD. ERROR OF ESTIMATE = 863.63

SE(B₁) = 2881.42

T₁ = -2.64735

SE(B₂) = 38.0628

T₂ = 2.23431

SE(B₃) = 142.712

T₃ = 4.69471

R SQUARED = .936715

ADJUSTED R SQUARED = .918634

F = 51.8055

DEGREES OF FREEDOM = 7

DW = 2.36762

CORRELATION MATRIX

1 .858751 .944237

.858751 1 .764408

.944237 .764408 1

图 5.3.1 收入在 IQ 和教育上的回归

怎样解释多重回归方程 $Y = -7.628 \cdot 13 + 85.04 X_2 + 669.99 X_3$?

假定有接受相同教育的两个具体的人 (X_3 保持常数), 平均地说 IQ 中一点之差将对应于收入中 85.04 元的差别。类似地, 教育相差一年, 保持 IQ (X_2) 不变, 对应于收入中 669.99 元之差 (平均而言)。因此, X_2 上的回归系数代表当 X_3 保持或控制得不变化时可以预料收入中的变化; X_3 上的系数表明当 X_2 保持常数时收入中的变化。这些系数叫偏回归系数。多重回归分析的长处亦在于此: 控制一变量为常数时解释另一变量所具有的相对重要性。

二元模型中, IQ 上的系数是 221.64 (表 4.3.) 控制住教育后, 数值仅为 85.04。这就是说, 在阐明收入时 IQ 的作用在保持教育水平后要小得多。

2. 假设检验

为了检验回归系数的显著性, 可以应用自由度 $N-K$ 的 t 检验。

$$T(B_1) = \frac{B_1}{SE(B_1)} = \frac{7.628 \cdot 13}{2.881 \cdot 42} = 2.647$$

$$T(B_2) = \frac{B_2}{SE(B_2)} = \frac{85.04}{38.06} = 2.234$$

$$T(B_3) = \frac{B_3}{SE(B_3)} = \frac{669.99}{142.71} = 4.695$$

虽然计算机可能未打印 t 的绝对值 (图 5.3.1), 检验零假设时必须只应用绝对值。表 5.3.1 表示在已知自由度 = 7 时各种概率水平上的临界 t 值。因 $T(B_1)$ 大于 2.365, 常数项在 .05 水平上是显著的。类似地, 2.234 大于 1.895 (但非 2.365) 并因此 IQ 在 .1 水平上是显著的, 但在 .05 水平上则否。因 4.695 大于 3.499, 在 .01 水平上教育是显著的。

表 5.3.1 自由度 = 7 时 t 的临界值

概率	.2	.1	.05	.02	.01	0.001
t 值	1.415	1.895	2.365	2.998	3.499	5.405

3. 修正 R^2

R^2 是被自变量解释的因变量中变化部分的质量。不过，为了解释因变量而应用 R^2 时有两个主要问题。首先，我们不能肯定模型是否已正确设定，因而用 R^2 比较各种模型时能引出一个模型优于另一模型的错误结论。其次， R^2 是模型中自变量数的函数，通常随自变量增加而上升。为了说明后者，在图 5.3.2 中已在计算机上用随机数函数生成了第三自变量。这个随机变量统计上是不显著的 ($T = .68$)，但 R^2 ($.941$) 大于三变量模型的 R^2 ($.937$)。

REGRESSION COEFFICIENTS

-8028.44

88.8863

709.574

-902.453

STO. ERROR OF ESTIMATE = 898.388

SE (B1) = 3053.87

T 1 = -2.62894

SE (B2) = 39.9901

T 2 = 2.22271

SE (B3) = 159.312

T 3 = 4.454

SE (B4) = 1318.01

T 4 = -.684711

R SQUARED = .941302

ADJUSTED R SQUARED = .911953

F = 32.0725

DEGREES OF FREEDOM = 6

DW = 2.61276

CORRELATION MATRIX

1 .858751 .944237 .543762

.858751 1 .764408 .536025

.944237 .764408 1 .607221

.543762 .536025 .607221 1

图 5.3.2 包含随机变量的多元回归模型

为了修正当更多变量加进多元回归模型时 R^2 增大的问题，可以找出把模型中变量数考虑进来的一种“修正” R^2 ，用 $\bar{R}^2$ 表示：

$$\bar{R}^2 = 1 - (1 - R^2) \frac{N - 1}{N - K} \quad (5.3.1)$$

图 5.3.2 中修正 $\bar{R}^2$ 为 .912，比图 5.3.1 中原始模型的修正 $\bar{R}^2$ 低。在任意二模型间比较修正 $\bar{R}^2$ 仅适用于因变量相同的情况。还应当指出修正 $\bar{R}^2$ 可以为负。例如，如果 $R^2 = .18$ ， $K = 3$ 和 $N = 10$ ，则 $\bar{R}^2 = -.05$ ，这是个无意义的值。

3. F 检验和方差分析

由回归程序打印出来的 F 统计量能帮助我们检验 $\beta_2 = \beta_3 = \dots = \beta_K = 0$ 的零假设。若零假设成立，则在解释因变量时整个模型并不显著，而 R^2 与零并无显著差别。分子的自由度为 $K - 1$ ，分母为 $N - K$ ：

$$F_{K-1, N-K} = \frac{R^2(N-K)}{(1-R^2)(K-1)} \quad (5.3.2)$$

将图 5.3.1 中的回归代入 (5.3.2) 式：

$$F = \frac{.9367(7)}{.0633(2)} = 51.81$$

自由度为 (2, 7)、显著性水平 .01 时 F 的临界值是 9.55。因为 51.81 大于 9.55，零假设被拒绝。

通常计算机回归程序在个别回归系数 T 统计量位置上或另外加上打印出来的 F 统计量。F 与 T 统计量的关系是 $F = T^2$ 。表 5.3.2 说明三变量模型中 β_2 与 β_3 的 F 统计量是如何计算的。这有时也叫方差分析检验。 Σ_2 (没有 Σ_3) 指回归仅以两个变量

—— Y 和 Σ_2 运行。本例中解释的平方和 (ESS) 等于 60,839,927。然后回归以方程中的 Σ_2 和 Σ_3 二者运行，而 ESS 增加 16,439,122。类似地，把 Σ_2 的 3,723,430 加进 ESS 以与仅有 Σ_3 的 Y 回归模

型比较。残差平方和 (RSS) 是 $\sum e^2$, F 的公式是

$$F(B_k) = \frac{(\text{加进 } x_k \text{ 在 ESS 的增加量})(N-K)}{\text{RSS}} \quad (5.3.3)$$

F 值可与具有 (1, 7) 自由度的下表比较。注意 T 检验中有 $N-K$ 个自由度, 而 F 检验的分子自由度为 1, 分母为 $N-K$ 。表 5.3.3 表示了 (1, 7) 自由度的临界 F 值: 针对回归系数显著性的结论与从 T 检验所得完全相同。

表 5.3.2 方差分析 (F 检验)

Y 中的变化, 源于:

x_2 (没有 x_3)

平方和
60,839.927

加上 x_3

16,439.122

残差或误差

5,220.951

总 计

82,500,000

Y 中的变化, 源于:

x_3 (没有 x_2)

73,555,619

加上 x_2

3,723.430

残差或误差

5,220.951

总 计

82,500,000

$$F(B_2) = \frac{(3,723.430)(7)}{5,220.951} = 4.99 \quad T^2 = (2.234)^2 = 4.99$$

$$F(B_3) = \frac{(16,439.122)(7)}{5,220.951} = 22.04 \quad T^2 = (4.695)^2 = 22.04$$

表 5.3.3 (1, 7) 自由度的临界 F 值

0.05 水平 = 5.59

0.01 水平 = 12.25

4 相关矩阵

回归程序打印的相关矩阵对于每一二之四归来讲正如§4.3定义的, 包含R("简单"相关)。例如, 矩阵第一行和第一列(1,1)是因变量本身的相类。虽然这必须等于1, 其他表示每个自变量本身相类的对角元素(2,2)和(3,3)也一样。第一行和第二列(1,2), 包括因变量Y与 X_2 的相类(0.858751), 而(1,3)元素是Y与 X_3 的相类(0.944237), (2,3)元素是 X_2 与 X_3 的相类(0.764408), 这里应等于(3,2)元素, 即 X_3 与 X_2 的相类。类似地, 矩阵的其他元素也是重复出现的, 比方(1,2)=(2,1)。

5. 多重共线性

多重回归模型唯一的补充假定是不存在完全多重共线性。设我们把第三自变量加进模型: 教育的月数(X_4), 虽然 X_4 等于12倍 X_3 , 后者是教育的年数(表5.3.4)

表5.3.4 解释完全多重共线性的数据

教育年龄(X_3)	9	8	10	10	11	14	15	13	16	17
教育月数(X_4)	108	96	120	120	132	168	180	156	192	204

CANT INVERT MATRIX AT LINE 390

REGRESSION COEFFICIENTS

-0.179605E 9

- 148229

7916.67

0.25331E 8

STD. ERROR OF ESTIMATE = 0.221239E 10

SE(B1) = 0.105537E 9

T1 = -1.70182

SE(B2) = 0.97507E 8

T2 = -1.52019E -2

```

DIVISION BY 0 AT LINE 660
SE(B 3)=0      T3 = 0
DIVISION BY 0 AT LINE 660
SE(B 4)=0      T4 = 0
RSQUARED=-.355977E 12
ADJUSTED R SQUARED=-.533965E 12
F = -2
DEGREES OF FREEDOM=6
DW=.231045E -1
CORRELATION MATRIX
      1  .858751  .944237  -.944237
      .858751  1  .764408  -.764408
      .944237  .764408  1  1
      .944237  .764408  1  1

```

图 5.3.3 有完全多重共线性的回归

图 5.3.3 表示了计算机的一个回归运行记录，收入为因变量 Y 、教育年数和教育月数为自变量。计算机打印结果表明矩阵 $(X'X)$ 不能求逆。因此，所有的输出均无意义，除了自由度和相关矩阵，其中用 1 表示 X_3 与 X_4 的相关。因此，当两个或更多自变量完全相关或共线，不能用 OLS 法。

实际上，完全多重共线性仅表现为模型规范中的一种误差。这种误差易于通过考查相关矩阵予以检出，因而变量之一与另一变量完全共线时可以从模型剔除。更一般地说，如果两个或更多自变量相互间高度相关，就成为不完全多重共线性，简称多重共线性。那样一个变量，就会跟着把另一个也取消了，使得通过保持另一变量不变时很难解释回归系数，为阐明多重共线性的后果，考虑 Y 为收入、 X_2 为 IQ、 X_3 为教育使之与 X_2 高度相关 (98)