

Cobb-Douglass函数在经济预测模型中的应用及其参数估计^{*}

张明礼

(西安武警技术学院 710086)

摘要 本文讨论了 C-D函数在建立工农业生产总量预测模型中的应用及其参数估计方法。

关键词 Cobb-Douglass函数 误差项选择 参数估计

1 引言

C-D函数在建立经济预测模型,以及衡量技术进步在经济增长中的作用等方面有着广泛的应用。象《预测》杂志仅1996年第3期就刊登了三篇有关使用C-D函数的论文。在文献[5],[6],[7]中,分别建立了农业生产预测模型,工业生产函数模型和玉米螟产卵期预测模型,并得出较好的结果。但是,建立C-D函数模型的关键,一是要模型形式准确,变量选择合理;二是要误差项设定正确,有比较好的参数估计方法。否则,所得的模型或者与实际背景矛盾,或者有较大偏差,就不能有效地进行预测。

文献[5]使用主成份回归方法建立了1978~1992年江苏省农业生产预测模型为:

$$\begin{cases} \hat{y} = 0.0008x_1^{0.0104}x_2^{0.0196}x_3^{2.4711}x_4^{-0.7625} \\ x_2 = -53.9325 + 6.4265x_1 \end{cases} \quad (1)$$

其中, x_1 为耕地面积, x_2 为农业劳动力, x_3 为农业机械总动力, x_4 为农业化肥实物施用量, $\hat{y}$ 为农业总产值。诚然,使用主成份回归是消除自变量复共线性的有效方法,但作者以为,模型(1)是不适宜的。我们知道,C-D函数 $y = Ax_1^{U_1}x_2^{U_2}\cdots x_m^{U_m}$ 中, U_i 为投入要素 x_i 的产出弹性。化肥投入量的弹性 $U_4 = -0.7625$ 违背了弹性系数为正的基本假设。显然,不能认为化肥施用量增加1%,产出反而下降0.76%,这也不符合实际情况(当然化肥施用量是有一个限度的)。另外, U_3 比 U_1 高出236倍,比 U_2 高出125倍,也不合常情。因为由

此将得出 y 的增长主要由农业机械总动力 x_3 所决定,而土地 x_1 ,劳动力 x_2 的作用甚微。这原因我以为是对解释变量选择不当,模型形式也尚可改进。文献[6][7]中使用的参数估计方法均是取对数后再求线性回归参数。当然,可以用其它方法去估计,以使误差变小,增加预测精度。

C-D函数的误差项 X 可有两种假定^[4],即

$$y = A_0x_1^{U_1}x_2^{U_2}\cdots x_m^{U_m}e^X \quad (2)$$

与

$$y = A_0x_1^{U_1}x_2^{U_2}\cdots x_m^{U_m} + X \quad (3)$$

(2)式和(3)式分别假定具有乘法误差和加法误差。一般地,对具有乘法误差项的(2)式可取对数线性化再求参数,而对具有加法误差项的(3)式则不可。正如D. Gujarati在文献[4]中(P455)“设定偏倚”一节所指出的,设定偏倚还可能是由于“不正确地设定误差项”,例如在Cobb-Douglass函数中,用方程15.2.1式(即(2)式)而不用15.3.1式(即(3)式)中设定的误差项”。因此,有必要对建立C-D函数模型的几个关键问题加以讨论。

2 C-D函数模型形式的选择

C-D函数的最初形式为(只写出确定性部分):

$$y = Ax_1^{U_1}x_2^{U_2}\cdots x_m^{U_m} \quad (4)$$

其中, y 为产出量, U_i 为投入量 x_i 的产出弹性系数。通常可假定有乘法误差和加法误差项。由于C-D函数在经济增长分析中起到了很好作用,引起不少学者的讨论与研究。有人又将技术水平 A 用一个随时间变化的量 $A_t = A_0e^{rt}$ 代替,即有:

$$y = A_0e^{rt}x_1^{U_1}x_2^{U_2}\cdots x_m^{U_m} \quad (5)$$

其中, A_0 是基年的技术水平, r 是技术进步系数。根据文献[1],[2]中的结果与作者多年从事生产函数测算的经验,对于工业生产函数,一般用

* 收稿日期: 1996-12-30

$$y = A_0 e^{rt} K^T L^U \quad (6)$$

测算,其中 T 和 U 分别为投入资金 K 和劳动 L 的产出弹性。有时,为使技术进步贡献从投入要素对总产出中分离出来,还可采用规模收益不变的假定,即令 $T + U = 1$ 而对农业生产函数模型,一般可用

$$y = A_0 e^{rt} L^{U_1} D^{U_2} K^{U_3} P^{U_4} \quad (7)$$

测算,其中 L 为劳力, D 为播种面积, K 为投入资金, P 为虚变量(这里使用单位耕地面积的产值),它表示了当年的天气、水份、土地质量等因素的影响。

如果所研究的模型中没有时间变量,象文献 [7] 中玉米螟产卵期模型,使用 (4) 式即可。

3 C-D函数误差项的设定

若 C-D函数误差项的设定不同,则所用的测算方法与结果就有差异。如何确定所给的数据是服从乘法误差还是加法误差?一般来说,若随着投入量 x 的变化, y 的变化摆动增大,便可用乘法误差假定;若随着 x 的变化, y 的变化摆动无明显增大,则用加法误差假定。但因经济时间序列样本数 N 较小,故上述方法不易于判定。实际应用中,我们可以在某种假定下先进行估计,若得到的参数在统计上不显著,或者有错误符号,可再用另一种假定误差模型去估计,看看存在问题能否解决,误差方差的估计是否减小,最后分析,比较确定出一个较为合理的假定误差模型。因此,模型的建立往往是一个反复尝试的过程。文献 [7] 中给出的预测模型假定有乘法误差项,所得到的结果是:

$$y = 281311985.5x_1^{-3.6966}x_2^{-0.7528}x_3^{-0.1884} \quad (8)$$

其误差方差 $Var = 3.6617$;若假定为加法误差项,用非线性回归方法去做,则为

$$y = 176872000x_1^{-3.3063}x_2^{-0.9746}x_3^{-0.1826} \quad (9)$$

其方差 $Var = 3.2874$ 从参数看, A 变化较大,但其余参数符号相同,变化不大。但假定乘法误差时方差比假定有加法误差项时方差高 11.4%;当然,对这组数据,两种假定都可以接受,文献 [6] 中模型也有类似情形。但有些数据则差异较大,这将在最后举例讨论。

4 C-D函数的参数估计方法

C-D函数由于选择的形式与误差设定不同,可有多种估计参数的方法,这里介绍五种。

4.1 线性回归方法

若设定为乘法误差,且形如 (5) 式,则可取对数化为线性模型,即得

$$\ln y = \ln A_0 + rt + U_1 \ln x_1 + U_2 \ln x_2 + \dots + U_m \ln x_m \quad (10)$$

令 $a = \ln y, a = \ln A_0, x_0 = t, x_1 = \ln x_1, \dots, x_m = \ln x_m$,

$$\text{则有 } a = a + rx_0 + U_1 x_1 + \dots + U_m x_m \quad (11)$$

只须求一组参数 $(a, r, U_1, \dots, U_m)$, 使

$$\begin{aligned} Q_0 &= \sum_{i=1}^N [a_i - (a + rx_{0i} + U_1 x_{1i} + \dots + U_m x_{mi})]^2 \\ &= \sum_{i=1}^N [a_i - \hat{a}_i]^2 \end{aligned} \quad (12)$$

最小即可。这是最常见的多元线性回归方法。

4.2 非线性回归方法

若假定为加法误差项,且形如 (5) 式,则不能线性化。这时须用非线性回归方法估计。即求一组参数 $(A_0, r, U_1, \dots, U_m)$ 使

$$Q_1 = \sum_{i=1}^N [y_i - A_0 e^{rt} x_{1i}^{U_1} \dots x_{mi}^{U_m}]^2 = \sum_{i=1}^N [y_i - \hat{y}_i]^2 \quad (13)$$

最小;求解方法是将 $f_i(\theta) = A_0 e^{rt} x_{1i}^{U_1} \dots x_{mi}^{U_m}$ 在 θ_i 附近展开为泰勒级数,仅取前二项,经推导可得迭代公式

$$\theta_{i+1} = \theta_i + (J^T(\theta_i)J(\theta_i))^{-1}J^T(\theta_i)(y_i - f(\theta_i)) \quad (14)$$

其中 $J(\theta_i)$ 为 Jacobi 矩阵, θ 表示等估参数向量。这就是高斯—牛顿法。此法的不足是公式复杂,收敛解与初始值关系密切,故使用起来麻烦。但这是求非线性回归模型参数的主要方法 [3]。

4.3 回归分析加权估计

因为用高斯—牛顿法求非线性参数具有复杂性,我们可采用线性回归分析加权估计 [8]。经推导,可得 (13) 式与 (12) 式右端关系为:

$$y_i - \hat{y}_i \approx y_i (\hat{a}_i - \hat{a}_i)$$

从而有

$$Q_2 = \sum_{i=1}^N (y_i - \hat{y}_i)^2 \approx \sum_{i=1}^N y_i^2 (\hat{a}_i - \hat{a}_i)^2 \quad (15)$$

通过极小化 Q_2 求出的参数是与 Q_1 极小化所得参数较接近的,即这是求非线性模型参数的近似方法。由此法求出解称为线性回归加权估计。还可使用迭代法,显然

$$y_i - \hat{y}_i = \frac{y_i - \hat{y}_i}{\hat{a}_i - \hat{a}_i} (\hat{a}_i - \hat{a}_i)$$

$$(t = 1, 2, \dots, N)$$

可得

$$Q_3 = \sum_{i=1}^N (y_i - \hat{y}_i)^2 = \sum_{i=1}^N \left[\frac{y_i - \hat{y}_i}{\hat{a}_i - \hat{a}_i} \right]^2 (\hat{a}_i - \hat{a}_i)^2 \quad (16)$$

这里, $\hat{a}_0$ 与 $\hat{y}_0$ 为由 (15) 式求出的参数估计与 y 的估计,由 (16) 式求出 $\hat{a}_i$ 后,反复迭代可求得几乎接近真值的参数来。此法称回归分析加权迭代估计。这种方法,可编制一个程序,一次计算线性回归,加权估计和加权迭代解,见文献 [8]

4.4 主成份估计

若用以上方法求出参数与符号与实际不符,或误差方差较大,或自变量复共线性严重,可考虑使用主成份估计。文献 [5], [9]中已有详尽叙述,在此不再赘述。

4.5 在约束条件下的复形调优法

对 C-D函数 (5),一般假定 $A_0, r, U_i (i = 1, 2, \dots, m)$ 均大于零。若使用以上各法均不能达到这一要求,或使用主成份估计后,虽参数为正,但误差方差较大,须优化参数,便可使用约束条件下的复形调优法。即在 $(M, R, b$ 为常数)

$$0 < A_0 < M, 0 < r < R, 0 < U_i < b_i \\ (i = 1, 2, \dots, m)$$

条件下,使用随机投点的统计方法求一组参数使 (13) 式中 Q_i 最小。方法的原理和程序见文献 [10]

综上所述可知,对于给定的数据和模型,先用回归分析加权估计一次求出线性回归解和加权迭代估计,即可得 C-D函数在具有两种误差项时的参数,再经比较分析选择一种即可;否则,继续用主成份估计或复形调优法,直至满意。

5 算例

5.1 陕西省工业生产函数预测模型

这里使用陕西 1952~ 1985年工业总产值、资金

和劳动力的时间序列数据,分别建立 $y = AK^{U_1}L^{U_2}$ 模型 (见表 1) 与 $y = A_0e^{rt}K^{U_1}L^{U_2}, (U_1 + U_2 = 1)$ 模型, (见表 2),由表 1可见,设定加法误差项,即用非线性估计的参数比设定乘法误差项,即用线性化方法估计的参数要优良得多,不仅误差方差少 85.45,而且参数 U_1 与 U_2 的比例比较接近实际,不象线性化方法所得参数 U_2 是 U_1 的 6.88倍。

表 2的结果说明,线性化方法不可取,主成份估计虽消除了复共线性,但使用非线性回归方法结果似乎更好。这里由于使用前几种参数估计方法即得出较好参数估计,故无需使用复形法。

对照表 1,表 2结果及文献 [2]研究,使用 (6)式估计工业生产函数模型为好。

5.2 渭南地区农业生产函数预测模型

取 1977~ 1988年渭南地区的农业统计数据。首先,使用 (4) 式估计,得不到合理参数。最后使用 (7) 式,选用劳力、播种面积、资金、虚变量、产值的时间数据,测算结果见表 3 从表 3知,使用线性回归、加权迭代、非线性回归方法均无效,而主成份方法较好。如果使用复形调优法,可使误差方差降低到主成份的 38%,且参数也合理。

表 1 工业生产函数 $y = AK^{U_1}L^{U_2}$ 测算结果表

参 数	$\hat{A}$	$\hat{U}_1$	$\hat{U}_2$	残差平方和 Q	残差方差 Var	相关指数 r
线性回归	0.2592	0.1607	1.1057	4177.3	134.75	0.9694
回归加权	0.4273	0.5863	0.5724	1614.9	52.09	0.9883
加权迭代	0.3176	0.5541	0.6690	1537.9	49.61	0.9886
非线性回归	0.2762	0.5418	0.7113	1528.47	49.30	0.9889
主成份	0.5418	0.4001	0.6918	3924.02	126.58	0.9713

表 2 工业生产函数 $y = A_0e^{rt}K^{U_1}L^{U_2}$ 测算结果表

参 数	A_0	r	U_1	U_2	残差平方和 Q	残差方差 Var	相关指数 r
线性回归	0.3351	0.0406	-0.1288	1.1288	1549.6	51.49	0.9888
回归加权	0.5126	0.0221	0.2791	0.7209	1473.3	49.28	0.9892
加权迭代	0.4709	0.0252	0.2321	0.7679	1432.7	47.79	0.9896
非线性回归	0.4492	0.0269	0.2073	0.7927	1426.9	47.57	0.9898
主成份	0.5608	0.0160	0.2598	0.7402	2605.41	86.85	0.9810

(下转 65 页)

表 2 运价表

$i \backslash j$	1	2	3	u_i
1	0	0.107		0
2	0.119	0.012	0.227	-0.095
3	0.411	0.304	0.065	-0.257
v_j	0	0.107	0.322	

由于 $1 - nQ = 1 - 3 \times 0.1 = 0.7$, 按方程组 (12) 得到运输问题的平衡表, 用西北角法求得方程组 (12) 的初始基本可行解 (y_{ij}) , 见表 1 $m = 2$ 时 (14) 式中只能取 $k = 0$, 于是

$$c_{ij} = \left| \frac{x_{ij} - 0.1}{x_{2j} - 0.1} - \frac{x_{0i}}{x_{1i}} \right| \quad (17)$$

由于 $j = 1, 2, 3$ 时, $(x_{1j} - 0.1) / (x_{2j} - 0.1)$ 分别为 1.2, 1.093, 0.854; x_{0i} / x_{1i} 分别为 1.2, 1.081, 0.789, 按 (17) 式作运价表 (c_{ij}) , 并求出位势 u_i, v_j , 见表 2

由于非基变量 y_{ij} 对应的检验数 $\gamma_{ij} = c_{ij} - u_i - v_j$ 都是正的, 故上述初始基本可行解即运输问题 (D) 的

最优解. 按 (16) 式计算得:

$$(q_{ij}) = \begin{pmatrix} 0.5 & 0.2 & 0 \\ 0 & 0.532 & 0.168 \\ 0 & 0 & 0.7 \end{pmatrix}$$

按 (15) 式计算得:

$$(p_{ij}) = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.1 & 0.632 & 0.268 \\ 0.1 & 0.1 & 0.8 \end{pmatrix}$$

于是预测第 4 个月的市场占有分布:

$$x(3) = x(2)(p_{ij}) = (0.2125, 0.3296, 0.4579)$$

参 考 文 献

[1] 暴奉贤, 陈宏立. 经济预测与决策方法. 暨南大学出版社, 1991
[2] 董云庭, 胡胜鹏. 管理数值分析. 电子工业出版社, 1987
[3] J. P. 伊格尼西奥. 目标规划及其应用. 哈尔滨工业大学出版社, 1988

(上接 52 页)

表 3 农业生产函数 $y = A_0 e^{rL^{U_1} D^{U_2} K^{U_3} P^{U_4}}$ 测算结果表

参数	A_0	r	U_1	U_2	U_3	U_4	残差方差 Var	相关指数 r
线性回归	690.5	-0.0091	0.0978	-0.0319	0.01622	0.9877	0.00124	0.9998
回归加权	710.3	-0.0073	0.0607	-0.0066	0.0092	0.9910	0.00106	0.9998
加权迭代	640.4	-0.0084	0.0903	-0.0159	0.0140	0.9877	0.00105	0.9999
非线性	611.1	-0.0083	0.934	-0.0098	0.0119	0.9893	0.00104	0.9999
主成份	0.0648	0.02127	0.7104	0.2745	0.3244	0.2746	0.86810	0.9907
复形调优	0.0988	0.0021	0.8431	0.3931	0.1680	0.6064	0.32730	0.9966

参 考 文 献

[1] 史清琪. 技术进步与经济增长. 科技文献出版社, 1985
[2] 乌家培. 经济模型在国民经济管理中的作用. 经济科学出版社, 1987
[3] D. A. RAtkowsky, 洪再吉译. 非线性回归模型. 南京大学出版社, 1986
[4] D. Gujarati, 庞皓译. 基础经济计量学. 科技文献出版社, 1986
[5] 高春津等. 主成份回归与经济预测模型. 预测, 1996 (3)

[6] 孙敬水. 测算广义技术进步增长率的新方法. 预测, 1996 (3)
[7] 宫锡鸿. 用 C-D 函数模型预测玉米螟产卵期. 预测, 1996 (3)
[8] 张明礼. C-D 函数弹性系数的回归加权估计. 武警技术学院学报, 1994 (2)
[9] 张明礼. 技术进步测算的主成分估计法. 武警技术学院学报, 1987 (4)
[10] 徐士良. FORTRAN 常用算法程序集. 清华大学出版社, 1992
[11] 葛宏立. 部分简单非线性方程线性化后参数求解问题探讨. 数学的实践与认识, 1989 (3)