

基于最小一乘准则的粮食生产模型研究

王桂芝, 赵靖, 朱干江, 江莹 (南京信息工程大学数理学院统计学系, 江苏南京 210044)

摘要 根据影响中国粮食产量诸多因素的分析结果, 运用计量经济学中的CD生产模型对中国粮食生产进行研究, 并采用最小一乘估计建立粮食生产模型, 通过拟合和比较分析得出稳健的中国粮食生产模型。

关键词 粮食生产; CD生产函数; 最小一乘准则; 弹性系数

中图分类号 S11+4 文献标识码 A 文章编号 0517-6611(2007)01-0012-02

Research on the Foodstuff Production Model of China Based on Least-absolute Criteria

WANG Gui-zhi et al (Department of Statistics, College of Mathematics & Physics Nanjing Information Engineering University, Nanjing, Jiangsu 210044)

Abstract According to the analysis results of many elements which influenced the foodstuff production of China, the authors established the foodstuff production model based on Least-absolute criterion. At the same time, some advices were given to solve the problems that existed in the foodstuff production of China.

Key words Foodstuff production; CD production function; Least-absolute criteria; Elasticity

农业是国民经济的基础, 而粮食生产则是农业发展的根本问题。作为农业大国, 粮食生产对我国的经济发展和人民生活水平提高都起着举足轻重的作用, 因此对粮食产量预测及影响产量相关因素的分析就显得尤为重要。

粮食生产是一个复杂的生态过程, 粮食生产能力是指一定地区在一定时期和一定经济技术条件下, 可以稳定地达到一定产量的粮食产出能力。粮食综合生产能力由投入和产出两方面因素构成, 由耕地、资本、劳力、科技、环境等要素的投入能力所决定, 由年度的粮食总产量所表现。

笔者从宏观角度综合考虑影响粮食生产的成灾面积、播种面积、化肥施用量、农业劳动人口、农业机械劳力5大因素, 拟合中国的粮食生产模型, 并对中国粮食生产情况进行综合分析。

1 基于最小二乘准则的粮食生产模型

从计量经济学角度来说, 粮食生产是典型的投入产出行为, 因此选择计量经济学中的CD和CES生产函数来进行相关的拟合和分析。

CD函数的形式为: $Q = AKL$

其中, Q 为产量, L 为劳力投入量, K 为资本投入量; A , α , β 为参数, $0 < \alpha, \beta < 1$ 。 α 和 β 为投入要素的产出弹性系数, 分别表示劳力和资本在生产过程中的相对重要性。

对上式进行形式变换得:

$$\ln Q = \ln(AKK) = \ln A + \alpha \ln K + \beta \ln L;$$

由于将变量转化为单位面积下的资本投入量并不影响参数估计及实证分析结果, 所以笔者使用农业总体数据来拟合模型。令 Q 表示粮食总产量(万吨); X_1 = 播种面积 - 成灾面积, 定义为有效播种面积(万公顷); X_2 表示机械劳动力(万千瓦); X_3 表示化肥施用量(万公斤); X_4 表示农业劳动力(万人)。

根据以上定义的解释变量, 结合处理需要变换函数形式为:

$$\ln Q = a + b \ln X_1 + c \ln X_2 + d \ln X_3 + e \ln X_4;$$

最小二乘法是由著名数学家Legendre和Gauss分别于1805年和1809年提出的。1900年Markov证明了Gauss-Markov定理, 从而奠定了最小二乘法在线性模型参数估计理论的地位。

该文利用1983~2004年粮食生产观测数据(数据来自中国统计年鉴2005), 使用最小二乘法进行参数估计, 使用逐步回归选择变量得表1(由SAS编程计算):

表1		参数估计结果			
变量	df	参数估计值	标准差	t值	Pr > t
截距	1	3.20694	0.85444	3.75	0.0013
lnX1	1	0.54945	0.08119	6.77	<0.0001
lnX3	1	0.31440	0.02175	14.45	<0.0001

得到基于最小二乘估计的中国粮食生产模型1如下:

$$\ln Q = 3.20694 + 0.54945 \ln X_1 + 0.31440 \ln X_3 \quad (1)$$

2 基于最小一乘准则的粮食生产模型的建立与分析

目前, 在模型参数估计中普遍采用的是基于最小二乘准则, 然而理论分析和实际算例都表明, 最小二乘法存在很大的局限: 最小二乘法采用的误差测度是残差的平方和, 这会放大噪声数据对参数估计的干扰; 最小二乘法要求随机误差服从正态分布, 这在实际算例中常常不能完全满足。最小一乘法的出现早于最小二乘法, 由于其误差测度函数的不可微性, 使得算法的实现难度大大增加, 这也是困扰最小一乘法发展的最主要因素。但最小一乘法能克服最小二乘法的局限性, 因此其稳健性和实用性更强^[3-4]。

根据基于最小二乘建立的模型1进行误差分析, 可以看出2004年粮食数据为一噪声数据, 产生这一噪声的原因是: 2004年, 我国许多省市遭受了罕见的虫灾, 而该文所使用的《中国统计年鉴2005》所统计的成灾面积数据仅包括水灾和旱灾。在这种噪声数据存在的情况下, 由于最小二乘估计使用残差的平方和为误差度量测度, 所以会放大噪声数据对建模的干扰; 同时最小二乘法要求随机误差服从正态分布, 根据检验该例中的随机误差对正态分布的服从度不高, 因此使用最小二乘估计建模存在很大局限。使用最小一乘法进行参数估计, 其误差累积测度函数为残差的绝对值和, 并且不要求随机误差服从正态分布, 因此最小一乘估计比最小二乘

估计的稳健性更强,统计性能更优。但最小一乘的误差累积函数不可微,所以算法设计难度大。使用最小一乘法建立粮食生产模型包括以下几点:

首先,将最小一乘问题转化为无约束规划问题,然后迭代计算可得基于最小一乘的参数估计。

其次,保留噪声数据,利用最小一乘估计,得基于最小一

乘的中国粮食生产模型2:

$$\ln Q = 2.118624 + 0.658155 \ln X_1 + 0.328258 \ln X_3 \quad (2)$$

再次,剔除噪声数据,利用最小一乘估计,得基于最小一乘的中国粮食生产模型3:

$$\ln Q = 2.418887 + 0.613290 \ln X_1 + 0.341455 \ln X_3 \quad (3)$$

然后,分别计算3个模型的误差并进行误差分析(表2):

表2 3个模型的误差分析结果

年份	模型1 误差	模型2 误差	模型3 误差	年份	模型1 误差	模型2 误差	模型3 误差
1983	-0.0223	-0.03561	-0.02148	1994	0.0295	0.032443	0.026781
1984	0.0145	0.000837	0.014246	1995	-0.0193	-0.03152	-0.03244
1985	0.0051	0.004707	0.012232	1996	0.0174	-3E-06	0.000003
1986	0.003803	0.000825	0.007836	1997	0.0394	0.032385	0.027376
1987	-0.00052	-0.00837	-0.00001	1998	0.0289	0.013847	0.011648
1988	-0.0153	-0.01833	-0.01326	1999	0.0317	0.019304	0.01586
1989	-0.0218	-0.02815	-0.02353	2000	0.0194	0.023653	0.013226
1990	-0.00808	-0.02509	-0.01785	2001	-0.0109	-0.00735	-0.018
1991	0.0101	0.005469	0.006098	2002	-0.0241	-0.02411	-0.03366
1992	0.0119	0.006367	0.006528	2003	-0.0145	-6E-06	-0.01583

注:2004年数据为噪声,故不将其纳入误差分析范围。模型1平均绝对误差为0.016720,模型2平均绝对误差为0.015507,模型3平均绝对误差为0.015456。

由表2可知,利用最小二乘估计所得模型1平均绝对误差为0.016720,为3个模型中平均绝对误差最大,并且稳健性最弱;保留噪声数据,基于最小一乘估计的模型2的平均绝对误差为0.015507,介于模型1和模型3之间;剔除噪声数据,基于最小一乘估计的模型3。其平均绝对误差最小为0.015456,拟合效果最好,且稳健性最强。

综上,选择稳健的中国粮食生产模型3作为基于最小一乘的中国粮食生产模型。

根据CD生产函数的经济意义和实际模型可知,有效播种面积的弹性系数为 $0.613290 < 1$,说明有效播种面积的投入为规模报酬递减的,即有效播种面积单位投入增量所带来的粮食产量增量是越来越小的。化肥施用量的弹性系数为 0.341455 也小于1,同样表明化肥施用量的投入也是规模报酬递减的。可以看到,对中国粮食产量影响最大的两个因素都是属于资本类因素,表明中国农业属于资本密集型行业;另外从 $0.613290 + 0.341455 = 0.954745 < 1$,即两因素弹性系

数之和小于1,可以看出,我国粮食生产主要依靠资本投入,资本投入增加,粮食产量也随之增加,但增加的速率递减。

3 小结

根据对影响中国粮食产量诸多因素分析的结果,分别建立了基于最小二乘的粮食生产模型和基于最小一乘的粮食生产模型。通过拟合和比较分析发现,基于最小一乘的粮食生产模型拟合精度更高,稳健性更强,更适合用于中国的粮食产量分析。

参考文献

- [1] 何忠伟. 中国粮食供求模型及其预测研究[J]. 北京电子科技学院学报,2005,13(1):22-25.
- [2] 雷钦礼. 中国粮食生产的价格作用机制分析[J]. 统计研究,2005(3):24-28.
- [3] 陈希孺. 最小一乘线性回归上[J]. 数理统计与管理,1989(5):48-55.
- [4] 陈希孺. 最小一乘线性回归下[J]. 数理统计与管理,1989(6):48-56.
- [5] MICHAEL D, G BODIN, CHENG H. 经济计量模型、技术与应用[M]. 北京: 中国社会科学出版社,2004.