

国内工业生产总产值的时间序列模型研究*

The Time Series Model of the Gross Industrial Output Value in China

梁 鑫, 谢佳利, 荣小军

LIANG Xin, XIE Jia-li, RONG Xiao-jun

(广西师范大学数学科学学院, 广西桂林 541004)

(College of Mathematics, Guangxi Normal University, Guilin, Guangxi, 541004, China)

摘要: 选择非季节模型、季节模型、乘积季节模型, 在 SPSS 系统下, 采集 2000 年 7 月到 2006 年 12 月国内工业生产总产值的原始数据进行分析找出其特征, 从模型识别、参数估计、适应性检验和实际拟合 4 个方面来确定最符合国内工业生产总产值发展规律的时间序列模型。分析结果表明, 季节模型、乘积季节模型的 Q 值都小于相应的 $i_{0.05}(M)$ 值, 而非季节模型相反; 乘积季节模型对 2006 年各个月份国内工业生产总产值的预测值与实际值的平均相对误差低于季节性模型, 预测值与实际值拟合最好。乘积季节模型最符合国内工业生产总产值的发展规律, 对各大中型企业和国家相关部门统一规划有较大的参考价值。

关键词: 时间序列模型 SPSS 系统 工业生产总产值

中图分类号: O212 文献标识码: A 文章编号: 1005-9164(2008)01-0035-03

Abstract Under the SPSS system, this paper carries out the characteristic analysis by adopting the primary data of the gross industrial output value in China from July 2000 to December 2006, and chooses the nonseasonal model, the seasonal model and the multiplicative seasonal model, to confirm the time series model that mostly fit in with the development rule of the gross industrial output value in China from four aspects: model identification, parameter estimation, diagnostic checking and actual fitting. The research presents the new result that the Q value of the seasonal model and the multiplicative seasonal model is less than the relevant value, while the nonseasonal model is just on the contrary. To the gross industrial output value in China, compared with the actual value, the average relative error of the forecast value of the multiplicative seasonal model is less than the seasonal model's. That is, the forecast value of the multiplicative seasonal model closely fits the actual value. Compare with the other typical statistic forecast models, the multiplicative seasonal model mostly fits the development rules of the gross industrial output value in China, and it can make a valuable reference for the unified programming of large-sized or mid-sized corporation and the related departments of the nation.

Key words time series model, SPSS system, the gross industrial output value

工业经济是国民经济的一个重要组成部分, 现阶段新型工业经济更是带动国民经济发展的主要动力, 是国民经济的主要支柱之一。准确分析和预测国内工业生产总产值发展规律, 对各大中型企业和国家相关部

门的统一规划具有较大的参考价值。很多学者对国内工业生产总产值进行了研究: 文献 [1] 建立了工业生产总产值的模型; 文献 [2] 运用 Box-Jenkins 识别方法来定阶模型; 文献 [3] 运用移动平均法建立自回归模型; 文献 [4] 建立多个工业生产总产值模型并进行比较, 最后进行预测, 但上述文献都没有对工业生产总产值的季节性、周期性进行分析。文献 [5] 也运用 Box-Jenkins 识别方法, 而且考虑了季节性、周期性, 但是该文献是基于 Eniew's 3.1 系统下对数据进行处理, 并且对模型的

收稿日期: 2007-08-28

修回日期: 2007-11-14

作者简介: 梁 鑫 (1978-), 男, 讲师, 主要从事概率统计研究。

* 广西自然科学基金项目 (0728091) 和广西师范大学青年科学基金项目 (2007) 资助。

适应性检验是基于 ACF(自相关系数)和 PACF(偏自相关系数)图形进行直观判断.本文在上述研究成果的基础上,基于 SPSS 系统^[6]运用非季节性模型、季节模型、乘积季节模型及非参数统计方法对国内工业生产总产值的发展规律进行了研究,得到较好的预测效果.

1 预测模型^[7]选择

(1)非季节性 ARIMA (p, d, q) 模型,其一般形式为 $H(B)\nabla^d x_t = \Theta(B)e_t$.

(2)季节性 ARIMA $(P, D, Q)_s$ 模型,其一般形式为 $U(B^s)\nabla_s^D x_t = V(B^s)X_t$.

(3)乘积季节 ARIMA $(p, d, q) \times (P, D, Q)_s$ 模型,其一般形式为 $H(B)U(B^s)\nabla^d \nabla_s^D x_t = V(B^s)\Theta(B)a_t$.

其中 B 为后移算子, ∇ 为差分符号, $\nabla = 1 - B$, ∇_s 为周期是 s 的差分, $\nabla_s = 1 - B^s$, d, D 均为差分阶数, $H(B) = 1 - h_1B - h_2B^2 - \dots - h_pB^p$, $U(B^s) = 1 - u_1B^s - u_2B^{2s} - \dots - u_pB^{ps}$, $V(B^s) = 1 - v_1B^s - v_2B^{2s} - \dots - v_qB^{qs}$, $\Theta(B) = 1 - \theta_1B - \theta_2B^2 - \dots - \theta_qB^q$, X_t 为扰动序列, a_t 为白噪声序列.

2 模型比较与分析

在 SPSS 系统下,采集 2000 年 7 月到 2006 年 12 月期间国内工业生产总产值的原始数据^[8]进行分析,找出其特征,看是否满足 3 种预测模型的要求,从模型识别、参数估计、适应性检验和实际拟合 4 个方面来确定最符合国内工业生产总产值发展规律的模型.

2.1 数据分析

数据分析结果如图 所示.

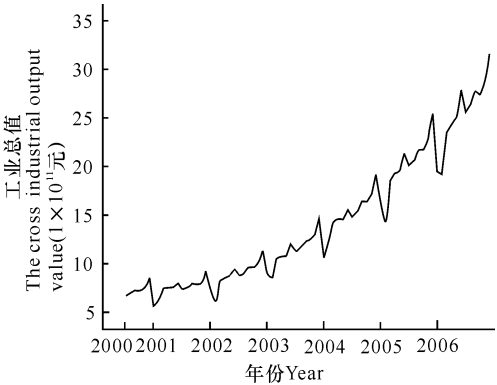


图 1 国内工业生产总产值时间序列

Fig. 1 The time series of the gross industrial output value in China

从图 1 中可以看出该序列是非平稳的,并且国内工业生产总产值发展具有如下基本趋势:(1)有波动现象,而且保持了长达数年的高速增长趋势,这与许多

其它的重要宏观经济指标(尤其是 GDP)相似.(2)以一年为周期变化,不同年份中同一月份呈现出较强的相似性,同一年份中不同月份也具有一定的相关性.

2.2 模型识别

对预测模型阶数的各种可能排列进行重复拟合,对明显不合适或提出不适合警告的排列形式予以删除,并采用最佳准则函数定阶法中的 AIC 准则与 BIC 准则^[7]判定模型阶数.结果如表 1 所示.

表 1 预测模型的阶数及对应的 AIC 与 BIC

Table 1 The order and relevant AIC value and BIC value of the forecast model

模型 Model	阶数 Order	AIC 值 AIC value	BIC 值 BIC value
ARIMA (p, d, q)	(2, 2, 1)	1320.948	1330.271
ARIMA $(P, D, Q)_s$	(2, 2, 0) ₁₂	855.874	861.841
ARIMA $(p, d, q) \times (P, D, Q)_s$	$(0, 1, 1) \times (2, 2, 0)_{12}$	848.330	856.211

经比较在收敛标准最大值为 10,参数变化为 0.001%,平方和变化为 0.001%的情况下,3 个模型分别取表 1 中阶数对 AIC 值和 BIC 值达到最小.

2.3 参数估计

取表 1 中各模型阶数作为暂定值,得到相应的各模型参数估计的结果如表 2 所示.

表 2 预测模型参数估计

Table 2 Parameter estimation in the forecast model

模型 Model	AR1	AR2	MA1	MA2	SAR1	SAR2	CONS-TANT
ARIMA (2, 2, 1)	-0.33	-0.473	0.993				9.224
ARIMA (2, 2, 0) ₁₂					-0.922	-0.634	1096.620
ARIMA (0, 1, 1) \times (2, 2, 0) ₁₂			0.999		-0.902	-0.624	-5.017

2.4 适应性检验

通过对序列原始数据与拟合数据的误差(常称为残差)序列进行模型适应性(优劣性)检验,若残差序列为白噪声,则意味着所建立的模型已包含了原始序列的所有趋势,从而模型应用于预测是合适的;否则,模型需要改进.对预测模型残差序列白噪声检验结果如表 3 所示.

表 3 残差序列自相关函数

Table 3 Auto correlation function of the residual sequence

模型 Model	M	Q 值	$i_{0.05}(M)$
ARIMA (2, 2, 1)	8	30.464	15.51
ARIMA (2, 2, 0) ₁₂	7	1.617	14.067
ARIMA (0, 1, 1) \times (2, 2, 0) ₁₂	7	1.891	14.067

表 4 2006年国内工业生产总产值比较

Table 4 The comparison about the gross industrial output value in China in 2006

月份 Month	实际值 Actual value (亿元)	预测值 Forecast value(亿元)	
		ARIMA(2, 2, 0) ₁₂	ARIMA(0, 1, 1)× (2, 2, 0) ₁₂
1	19449.00	20924.58	21026.8
2	19253.50	19952.21	19882.02
3	23530.99	24199.10	24131.42
4	24392.08	25030.88	24937.75
5	25127.53	25097.13	24979.00
6	27840.80	27243.00	27130.04
7	25715.29	25898.76	25791.32
8	26367.85	26497.00	26369.18
9	27838.21	27752.59	27608.02
10	27480.28	27943.98	27795.06
11	28617.47	29095.95	28930.26
12	31661.15	32165.80	31986.04

由表 3 结果可以看出 ARIMA(2, 2, 0)₁₂, ARIMA(0, 1, 1)×(2, 2, 0)₁₂ 的 Q 值都小于相应的 $i_{0.05}^2(M)$ 值, 可以认为这两个预测模型的残差序列属于白噪声序列, 即这两个预测模型有效, 可以较好地模拟工业生产总产值的时间序列, 而 ARIMA(2, 2, 1) 的 Q 值大于相应的 $i_{0.05}^2(M)$ 值, 说明该预测模型的残差序列蕴含着相关信息, 不是白噪声序列, 这个模型的预测效果不好.

2.5 实际拟合

预测模型 ARIMA(2, 2, 0)₁₂和 ARIMA(0, 1, 1)×(2, 2, 0)₁₂与实际值进行拟合的结果如图 2 所示.

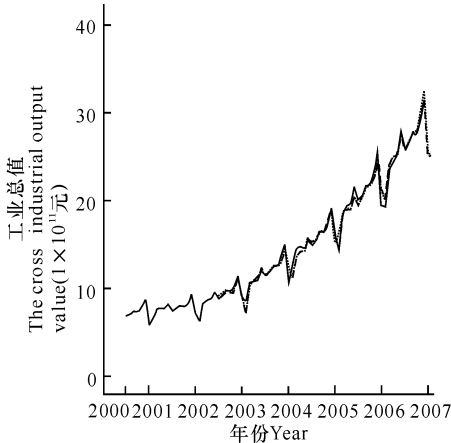


图 2 拟合值

Fig. 2 The sequence chart about fitting value of each year and predicted value in 2007

— : 实际值; - - - : ARIMA(2, 2, 0)₁₂预测值;
 ·····: ARIMA(0, 1, 1)×(2, 2, 0)₁₂预测值.
 — : Actual value; - - - : Forecast value of
 ARIMA(2, 2, 0)₁₂; ·····: Forecast value of ARIMA(0, 1, 1)×
 (2, 2, 0)₁₂.

由图 2 结果可以看出 ARIMA(0, 1, 1)×(2, 2, 0)₁₂ 模型的图形拟合效果最好, 我国的工业生产总产值保持高速平稳的增长趋势, 而且每年的 12 月份是一个最高点. 将 2006 年各个月份的工业生产总产值的实际值与预测值进行比较的结果(表 4), 计算 ARIMA(2, 2, 0)₁₂模型的平均相对误差为 0.0212, ARIMA(0, 1, 1)×(2, 2, 0)₁₂模型的平均相对误差为 0.0197. ARIMA(0, 1, 1)×(2, 2, 0)₁₂的拟合效果优于 ARIMA(2, 2, 0)₁₂.

3 结束语

基于 SPSS 系统所建立的乘积季节模型 ARIMA(0, 1, 1)×(2, 2, 0)₁₂很好地反映了国内工业经济的发展规律, 对各大中型企业和国家相关部门的统一规划具有较大的参考价值. 我们可以利用基于 SPSS 系统建立乘积季节模型的方法对现实生活中具有明显趋势和季节性特征的数据进行建模, 研究其内在规律, 并利用得出的结果对未来进行规划.

参考文献:

[1] 张德铭, 谭元发. 我国国内工业总产值动态分析 [J]. 矿冶工程, 2005, 25(4): 89-92.
 [2] 王志平, 张海帆. 中国工业时间序列分析及预测 [J]. 统计与决策, 2004, 12(8): 17-18.
 [3] 周国利, 易可. 贵州省工业生产总产值的趋势分析和预测 [J]. 贵州工业大学学报, 2001, 30(3): 1-4.
 [4] 陈玉娟, 查奇芬. 江苏省工业生产总产值的预测 [J]. 统计与决策, 2003, 23(11): 63-64.
 [5] 蓝斌, 李晓青. 厦门市工业总产值时间序列分析研究 [J]. 厦门理工学院学报, 2006(1): 11-14.
 [6] 杨善朝, 张军舰. SPSS 统计软件应用基础 [M]. 桂林: 广西师范大学出版社, 2001.
 [7] 王振龙. 时间序列分析 [M]. 北京: 中国统计出版社, 2000.
 [8] 中国国家统计局. 中国统计数据 [R/OL]. [2007-01-29]. <http://www.china.com.cn/ch-company>.

(责任编辑: 尹 闯 邓大玉)