

基于时序全局主成分分析的广西经济发展动态描绘*

The Application of Time Series Analysis and All-around PCA in the Economic Dynamic Description of Guangxi Province

郝 寒, 吴群英

YAN Han, WU Qun-ying

(桂林工学院数理系, 广西桂林 541004)

(Department of Mathematics and Physics, Guilin University of Technology, Guilin, Guangxi, 541004, China)

摘要:选取广西 1990~2005 年的地区生产总值、第一产业产值、第二产业产值、第三产业产值、财政收入、社会消费品零售总额、固定资产投资等 7 个指标 16 年的数据,利用时序全局主成分分析方法动态描绘广西 1990~2005 年的经济发展历程。描绘结果显示,1998 年以前广西的经济发展速度较慢,1998 年以后,广西经济发展有所加快,尤其是 2001~2005 年经济发展保持持续快速发展的势头。描绘结果与广西经济发展实际情况基本吻合,时序全局主成分分析方法可以综合评价国民经济的发展状况。

关键词:时序分析 主成分分析 经济发展

中图分类号:O212, F124.1 **文献标识码:**A **文章编号:**1002-7378(2008)03-0168-03

Abstract:In this article, the GDP, the total value of the primary industry, the total value of the secondary industry, the total value of the tertiary industry, financial revenue, the total retail sale of consumer goods and the fixed assets investment from 1990 to 2005 are selected. The economic course of Guangxi from 1990 to 2005 dynamically by means of time series analysis and all-around PCA is described. The result is as follows: the rate of Guangxi economic development is slow before 1998 and speed up after 1998. Particularly, the sustained and rapid development economy is kept in 2001~2005. The result has shown that the dynamic track can accord with the fact. Therefore, time series analysis and all around PCA can be used in comprehensive evaluation of the economic development.

Key words:time series analysis, all-around PCA, economic, development

近年来,如何科学地综合评价国民经济发展已经成为经济统计界十分关注的问题。要全面系统地分析一个地区的经济发展问题,必须考虑众多对经济过程有影响的因素,因为每一个变量都在不同程度上反映了所研究的某些信息,因而使评价系统具

有样本庞大、指标众多及动态变化等特性。而时序全局主成分分析方法正是针对上述研究内容提出的一种解决方法,它为经济决策提供了一种快速、科学的量化分析工具^[1]。主成分分析是一种利用降维的思想,将多指标化为少数几个综合指标的多元统计分析方法^[2],它认为在众多有相关性的因素之间必然存在着起支配作用的共同因素,于是通过对原始变量相关矩阵内部结构关系的研究,找出影响某一经济过程的几个综合指标,这些指标不仅保留了原始的主要信息,又使其彼此之间不相关,比原始变量具有某些更优越的性质,使得人们在研究复杂的经济问题时容易抓住主要矛盾。时序全局主成分分析方

收稿日期:2007-07-25

修回日期:2007-10-31

作者简介:郝 寒(1985-),男,硕士研究生,主要从事概率与统计研究。

* 国家自然科学基金项目(10661006),广西“新世纪十百千人才工程”专项资金(2005214)及广西自然科学基金项目(桂科自 0728212)资助。

法在主成分分析的基础上,以一个综合变量来取代原有的全局变量,再以此为基础描绘出系统的总体水平随时间的变化轨迹^[3]。

从改革开放以后,广西经济得到了很大发展,如何建立一个既能反映广西经济发展又符合客观历史的模型,对于政府决策部门今后做出正确的经济决策至关重要。本文选取广西 1990~2005 年的相关数据,利用时序全局主成分分析方法对广西 1990~2005 年的经济发展进程进行动态描绘,探讨时序全局主成分分析方法在综合评价国民经济发展方面的应用。

1 样本数据选取及数据预处理

1.1 样本数据

选取广西 1990~2005 年的地区生产总值(亿元)(X_1)、第一产业产值(亿元)(X_2)、第二产业产值(亿元)(X_3)、第三产业产值(亿元)(X_4)、财政收入(万元)(X_5)、社会消费品零售总额(万元)(X_6)、固定资产投资(万元)(X_7)等 7 个指标一共 16 年的数据^[4],作为反映经济发展的样本点。各项统计数据见表 1。

表 1 各样本原始指标值^[4]

年份	时序	X_1	X_2	X_3	X_4	X_5	X_6	X_7
1990	1	449.06	176.77	118.45	153.84	468305	1754369	685666
1991	2	518.59	195.17	141.02	182.40	559225	2002276	896479
1992	3	646.60	233.03	187.48	226.09	611953	2436189	1410395
1993	4	871.70	250.11	321.10	300.49	959269	3149992	2780754
1994	5	1198.29	333.79	469.81	394.69	622617	3991040	3825871
1995	6	1497.56	453.15	535.86	508.55	794422	4981172	4233742
1996	7	1697.90	534.88	587.37	575.65	905102	5718385	4764200
1997	8	1817.25	582.74	614.07	620.44	991568	6301661	4798023
1998	9	1911.30	586.70	667.29	657.31	1196720	6868810	5717025
1999	10	1971.41	567.72	682.34	721.35	1335647	7404577	6202035
2000	11	2080.04	557.38	732.76	789.90	1470539	8041371	6600146
2001	12	2279.34	576.34	771.18	931.82	1786706	8757053	7312523
2002	13	2523.73	601.99	846.89	1074.85	1867320	9597730	8349852
2003	14	2821.11	658.78	984.08	1178.25	2036578	10768653	9873063
2004	15	3433.50	817.88	1253.70	1361.92	2377721	12222421	12636500
2005	16	4075.75	912.50	1510.70	1652.57	2830359	13970227	17690715

1.2 数据预处理

由于各经济指标具有不同的量纲,在进行主成分分析之前需要对数据进行处理,以消除由于量纲的不同可能带来的不合理影响。对原始数据进行标准化处理,标准化的计算公式为 $x_i = \frac{x_i - \bar{x}}{S}$,其中 x_i 为原始数据, \bar{x} 为数据均值, S 为标准差, x_i 为标准化后的数据值。

为了表述简洁,标准化后的指标仍记为 $X_1 \sim X_7$ 。标准化后的指标值见表 2。

表 2 标准化后的指标值

年份	时序	X_1	X_2	X_3	X_4	X_5	X_6	X_7
1990	1	-1.376	-1.511	-1.394	-1.265	-1.182	-1.356	-1.214
1991	2	-1.308	-1.425	-1.335	-1.200	-1.053	-1.288	-1.166
1992	3	-1.183	-1.250	-1.214	-1.100	-0.978	-1.170	-1.051
1993	4	-0.964	-1.171	-0.864	-0.930	-0.485	-0.977	-0.745
1994	5	-0.646	-0.782	-0.475	-0.715	-0.963	-0.748	-0.511
1995	6	-0.355	-0.229	-0.302	-0.455	-0.719	-0.480	-0.420
1996	7	-0.160	0.151	-0.168	-0.302	-0.562	-0.279	-0.301
1997	8	-0.044	0.373	-0.098	-0.200	-0.439	-0.121	-0.294
1998	9	0.048	0.391	0.041	-0.116	-0.148	0.033	-0.088
1999	10	0.106	0.303	0.081	0.030	0.049	0.178	0.020
2000	11	0.212	0.255	0.213	0.187	0.241	0.351	0.109
2001	12	0.406	0.343	0.313	0.511	0.690	0.545	0.269
2002	13	0.644	0.462	0.511	0.837	0.804	0.774	0.501
2003	14	0.934	0.725	0.870	1.073	1.045	1.092	0.842
2004	15	1.530	1.463	1.575	1.492	1.529	1.486	1.460
2005	16	2.155	1.902	2.247	2.155	2.172	1.961	2.590

2 广西经济发展动态描绘

2.1 主成分选取

用 SPSS13.0 软件对标准化后的数据进行主成分分析得到特征值、载荷矩阵和主成分得分系数分别见表 3、表 4 和表 5。

表 3 标准化后的指标间相关系数矩阵的特征值、方差贡献率和累计方差贡献率

主成分	特征值	方差贡献率 (%)	累计贡献率 (%)
1	6.835	97.650	97.650
2	0.112	1.596	99.246
3	0.040	0.566	99.812
4	0.009	0.127	99.939
5	0.003	4.50E-02	99.984
6	0.001	1.60E-02	100
7	7.86E-13	1.12E-11	100

表 4 载荷矩阵

主成分	X_1	X_2	X_3	X_4	X_5	X_6	X_7
1	0.999	0.996	0.996	0.995	0.987	0.975	0.969
2	0.043	-0.056	0.044	0.003	-0.070	-0.204	0.241
3	-0.006	0.029	-0.072	0.085	-0.142	0.007	0.038
4	-0.016	-0.048	-0.012	-0.034	0.015	0.054	0.044
5	0	0.029	-0.040	-0.013	0.018	-0.007	0.013
6	-0.009	-0.013	-0.010	0.023	0.014	-0.005	0
7	-7.59E-07	3.24E-07	2.83E-07	-4.83E-13	-2.57E-11	3.49E-12	1.59E-07

表5 第一主成分得分系数

原始变量	第一主成分
X_1	0.1461
X_2	0.1417
X_3	0.1456
X_4	0.1458
X_5	0.1426
X_6	0.1456
X_7	0.1444

由表3结果可知第一主成分的累积方差贡献率已达97.65%，基本上保留了原来指标的信息。另外由表4结果可知第一主成分地区生产总值与各原始变量的因子负荷量都为正值且都在0.969以上，说明第一主成分基本反映了广西各年份的经济发展水平的信息。

2.2 构造综合评价模型

确定主成分后，用一个新变量 F 来代替原来的7个指标，即可用 F 为综合评价指标来进行综合评价。由表5可知 F 的线性组合为 $F = 0.1461X_1 + 0.1417X_2 + 0.1456X_3 + 0.1458X_4 + 0.1426X_5 + 0.1456X_6 + 0.1444X_7$ ，代入数据即得出16个综合评价值 F (见表6)。

表6 各样本综合评价值

年份	时序	F
1990	1	-1.3438
1991	2	-1.2687
1992	3	-1.1489
1993	4	-0.8872
1994	5	-0.6989
1995	6	-0.4277
1996	7	-0.2349
1997	8	-0.1197
1998	9	0.0225
1999	10	0.1106
2000	11	0.2266
2001	12	0.4446
2002	13	0.6557
2003	14	0.9517
2004	15	1.5231
2005	16	2.1953

2.3 时序分析

根据每年的综合评价值 F 绘制的时间序列变化结果如图1所示。

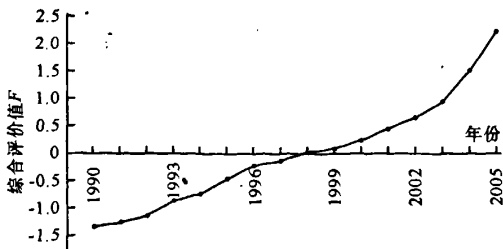


图1 广西经济时序变化曲线

图1动态轨迹结果表明，1998年以前广西的经济发展速度较慢；1998年以后，广西经济发展有所加快，尤其是2001~2005年经济发展保持持续快速发展的势头。

3 结束语

广西1998年以前经济发展速度较慢，工业基础薄弱。改革开放以来，全国经济迅速发展，但是广西的经济发展起步较晚，1990~1998年地区生产总值只增加了1462.24亿元，其中，第一产业409.93亿元，增长5.8%；第二产业548.84亿元，增长13.9%；第三产业503.47亿元，增长6.4%。1998年以后，广西实施了“三大战略，六大突破”，为经济发展打下了良好的基础。同时，西部大开发战略的实施，给处于西部地区的广西经济发展提供强大的支撑力。特别是中国-东盟自由贸易区的构建以及中国-东盟博览会永久落户广西南宁，为推动广西经济的迅速发展提供了极好的发展机遇。1998年至2005年地区生产总值增加了2164.45亿元，其中，第一产业325.80亿元；第二产业843.41亿元；第三产业995.26亿元^[4]。

随着近年来东部地区土地、劳动力等成本优势的逐渐丧失，产业结构进入调整与升级阶段，广西抓住这一机遇，从2001年开始大规模的承接东部产业转移，通过充分发掘比较优势，积极承接东部产业转移，加快经济发展，使经济发展保持了持续快速发展的势头。

由以上的分析可以看出，时序全局主成分分析模型所绘出的曲线对广西经济发展的描述与广西实际情况基本吻合，说明时序全局主成分分析方法对广西经济发展的动态描述是合理的，其分析方法可以用于综合评价国民经济的发展状况，其评价结果可以为有关部门的决策提供参考依据。

参考文献:

- [1] 乔峰,姚俭.主成分分析方法在经济发展动态描绘中的应用[J].数理统计与管理,2003,22(2):1-5.
- [2] 何晓群.多元统计分析[M].北京:中国人民大学出版社,2004.
- [3] 任若恩,王惠文.多元统计数据分析——理论、方法、实例[M].北京:国防工业出版社,1997.
- [4] 广西壮族自治区统计局.广西统计年鉴2006[M].北京:中国统计出版社,2006.

(责任编辑:韦廷宗)