

# 地区经济发展的现状和潜力分析

吴桂珍, 沈桂红, 姬广林

(吉林大学商学院, 吉林 长春 130012)

**摘要:** 主成分分析法可以避免只依据主观评定法确定指标体系权重的不准确性, 使评估指标体系的权重更加合理、有效。本文利用此方法, 对我国各个地区的经济发展水平和潜力分别进行了比较分析, 并提出把第一主成分作为一个地区综合经济实力的度量, 把第二主成分作为一个地区发展潜力的度量, 并分别给出了各个地区综合经济实力和潜力的排序, 同时也分析了今后东北老工业基地的发展趋势。

**关键词:** 地区经济发展; 潜力分析; 主成分分析

**中图分类号:** F224.0      **文献标识码:** A

## 1 问题的提出

中国的工业化进程、中国国民经济新一轮战略发展目标的实现迫切需要新的经济增长区域的支持。

80年代初、中期以深圳及珠海等经济特区的开辟为标志, 华南地区成为国民经济增长的主要支持区域; 80年代中、后期以苏、锡、常等地乡镇企业振兴和90年代初期上海浦东大规模开发为标志, 华东地区成为国民经济增长的主要支持区域; 90年代初、中期以京津唐高速公路、京石高速公路、第二条“欧亚大陆桥”、京九铁路等重要交通网络的相继开通带动下日益密切的区域经济合作与结构重组, 以山东半岛的振兴与外向发展为目标, 华北地区正在成为国民经济增长的主要支持区域。继此之后, 东北地区理所当然地成为人们的关注焦点和期望所在。

今后, 东北老工业基地将沿着什么样的轨迹加快发展? 东北老工业基地的经济发展趋势如何? 本文试图对全国30个地区2002年统计数据进行分析, 对各个地区的经济发展状况进行分析, 希望能从中挖掘出有用的信息, 为振兴东北老工业基地的有关决策提供参考。

## 2 对综合评价方法的选择与讨论

关于指标体系的综合评价方法有主观权数法、主成分分析法、因子分析法、聚类分析法等。1997年10月起用的新指标体系仍然沿用主观权数法, 即在考虑指标之间相关性强弱和各指标对综合经济效益影响程度的基础上, 采用专家调查法(德尔菲法)确定各指标的权数。该方法的优点在于评价的规范性和可操作性强, 但由于采用的权数是主观的, 即使对指标之间的相关性加以考虑, 也难以准确反映指标体系的内在结构关系; 同时, 使用统一的全国标准值加以对比求值, 而全国标准值本身主观性较强, 而且缺乏足够的理论依据; 更重要的是对于不同地区、不同类型的企业以及不同性质的指标, 均采用统一的全国标准值, 也会在很大程度上影响评价的准确性, 用主成分分析法则能较好地解决上述问题。

随着计算机技术的高度发展, 电子计算机的广泛应用, 主成分分析以其具有综合分析和研究多个指标经济现象的显著特点而越来越受到人们的重视。要综合评价一个城市的综合经济实力, 就不能单独地以某一个指标为依据, 而需要分析这些因素之间的相互关系, 根据它们的不同作用, 按一

定的顺序，进行综合评价。由于影响客观事物发展的各个因素之间都存在着一定的相互关系，故可以对众多的因素进行简化，而简化后的因素中有能综合反映原始因素的信息，它们之间又是互不相关的。主成分分析正是处理这一类问题的有效方法之一。

### 3 方法原理

主成分分析 (Principal Component Analysis, PCA) 是在保持原始数据信息损失最少的前提下，通过线性变换将原始自变量集合由高维空间映射成一个低维空间，从而实现数据的降维。在这个降维空间中，由于变量之间是相互正交的，所以使用这些综合了原始自变量集合最大信息的新综合变量进行回归分析时，将不再会存在使用普通最小二乘法回归的问题。同时，由于主成分分析中产生的新综合变量提取了原始变量最大的信息量，因而不会影响回归分析的结果。也就是说主成分分析是将多个实测变量转换为少数几个不相关的综合指标的多元统计分析方法，在经济管理和科学研究等众多领域中得到了广泛的应用。

设  $X = (x_1, x_2, \dots, x_p)'$  为  $p$  维随机变量，构造

$$Z_i = L_i' X \quad i = 1, \dots, p$$

其中  $L_i$  为  $p$  维正交化向量 ( $L_i \times L_i' = 1$ )， $Z_i$  之间互不相关且按照方差由大到小排列，则称  $Z_i$  为  $X$  的第  $i$  主成分。设  $X$  的协方差矩阵为  $\Sigma$ ，则  $\Sigma$  必为半正定对称矩阵，求特征值  $\lambda_i$  (按从大到小排序) 及其特征向量，可以证明， $\lambda_i$  所对应的正交化特征向量，即为第  $i$  主成分  $Z_i$  所对应的系数向量  $L_i$ ，

而  $Z_i$  的方差贡献率定义为  $\lambda_i / \sum_j \lambda_j$ ，通常要求提取的主成分的数目  $k$  满足  $\sum_{i=1}^k \lambda_i / \sum_{j=1}^p \lambda_j > 0.80$ 。

在确定了选择  $k$  个主成分之后，关键的一步是要对主成分做经济解释，即要对每个主成分赋予新的意义，给出合理的解释，这个解释应该根据主成分的计算结果结合定性分析进行。主成分是原始数据的线性组合。在这个线性组合中各变量的系数有大有小，有正有负。一般而言，线性组合中各变量系数的绝对值大的指标表明其对主成分的属性做出了较大的贡献，若几个变量系数相当时，则应认为这一主成分是一些数据指标的综合。

最后，由回归法估计出主成分得分，以各主成分的方差贡献率占其总方差贡献率的比重作为权重进行加权汇总，得出综合得分  $E$ ，即：

$$E_k = \sum_{i=1}^k (F_i Z_i) \quad (F_i \text{ 为主成分的方差贡献率})$$

PCA 的优点在于它确定的权数是基于由数据分析而得出的指标之间的内在的结构关系，不受主观因素的影响，有较好的客观性，而且得出的综合指标 (主成分) 之间相互独立，减少信息的交叉，这对分析评价极为有利。

PCA 方法通过主成分的提取，消除了评价指标之间共线性的影响。这种共线性意味着评价指标所代表的信息原来是有重叠的。PCA 方法能够按照主成分的方差贡献率，客观地确定各个主成分的权重。

### 4 确定指标体系

实施地区协调发展战略，首先要对各地区经济社会发展水平的现状进行准确的描述；其次为了消除地区发展的不均衡，还要对各地区经济社会发展的潜力进行大体上的估计，必须构建地区经济

社会发展评估指标体系。在具体指标的设置上，既应该考虑各地区的整体发展水平，还要兼顾各地区的潜力发展水平，因此要注意以下几个方面：

(1) 指标的全面性：指标体系既要反映各地区发展实力，其中包括经济发展水平及社会发展等各方面的状况，又要反映各地区的发展潜力，也就是可持续发展水平。

(2) 指标的代表性：指标本身有着强烈的经济和社会意义。

(3) 指标的可得性：有些指标很有意义，但因数据不可得只得放弃，或以相近的指标代替，例如交通密度指标，它反映地区交通基本状况，但资料不可得，只好以人均货运总量来代替。

(4) 指标的整合性：必须注重数量指标和质量指标、规模指标和效率指标的结合运用。

(5) 指标的简洁性：所选指标要尽量精炼使用，因为主成分分析要求样本容量大于变量个数。

因此，考虑选择以下 12 项指标：

X1: 国内生产总值 (GDP)	X2: 人均 GDP
X3: 人均出口额	X4: 地方财政收入占 GDP 比重
X5: 第三产业产值占 GDP 比重	X6: 人均外商直接投资
X7: 人均固定资产投资	X8: 人均货运总量
X9: 每万人普通高校在校学生数	X10: 人均社会消费品零售额
X11: 每万人拥有卫生技术人员数	X12: 人均邮电业务总量

上述指标分别从经济发展及其环境发育程度（资金供应能力）、医疗、教育、交通、通信等方面描述地区的社会经济综合发展状况和发展潜力。

## 5 实证分析

现在我们将对全国 30 个地区（将重庆市的数据合并到了四川省）上述 12 项指标的数据进行主成分分析，为消除量纲，我们首先将各项指标进行了标准化。结果如下：

### (1) 特征值和特征向量

由表 1 知： $\text{Var}(Z_1)=7.11517$ ， $\text{Var}(Z_2)=2.57735$ ， $Z_1$  的方差贡献率为 59.29308%， $Z_2$  的方差贡献率为 21.47791%，前两个主成分已反映原变量 80.77098% 的信息，而且由于第三主成分  $Z_3$  的特征根小于 1，只反映原变量 7.38715% 的信息，按照特征根选取主成分的原则，在这里我们只选取前面两个主成分来做进一步的分析。

### (2) 主成分的表达式及其含义

由表 1，我们可以发现：第一主成分（ $Z_1$ ）在各变量上的系数都为正，而且数值上相差不大，因而可以认为  $Z_1$  代表地区综合发展水平， $Z_1$  越高，表明地区综合发展水平实力越强；第二主成分（ $Z_2$ ）在变量前系数有正有负，在表示经济水平的变量 X1、X3、资金供应能力变量 X6、X7、教育水平变量 X9、医疗水平变量 X11 上的系数为正；而在市场化水平 X4、通信能力变量 X12 等这些变量上的系数为负。除国内生产总值（GDP）外，教育水平变量 X9、医疗水平变量 X11 这两个代替潜力的指标的系数最为显著，因而可以认为  $Z_2$  是地区经济发展水平与地区发展潜力的比较，它作为一个形状因子，为正值时，表示发展潜力超过其经济发展水平；接近于零时，表明地区经济发展水平与其潜力较为均衡；为负值时，表明经济发展水平超过其潜力。

表 1 主成分分析结果

		第 1 主成分 (Z <sub>1</sub> )	第 2 主成分 (Z <sub>2</sub> )	第 3 主成分 (Z <sub>3</sub> )	第 4 主成分 (Z <sub>4</sub> )
特 征 向 量	X1	0.12884	0.53244	-0.10393	-0.19677
	X2	0.36094	-0.01861	-0.15852	0.18220
	X3	0.34023	0.00013	-0.33680	-0.18752
	X4	0.29581	-0.12833	0.44353	0.03280
	X5	0.24530	-0.12678	0.67844	-0.18739
	X6	0.33480	0.02120	-0.34706	-0.21487
	X7	0.35244	-0.12180	0.02318	-0.03236
	X8	0.27753	-0.08066	-0.09500	0.88200
	X9	0.12477	0.55541	0.15073	0.06746
	X10	0.36480	-0.02650	-0.04639	-0.00899
	X11	0.05846	0.58826	0.19534	0.18060
	X12	0.35339	-0.08120	0.02439	-0.16498
特征值		7.11517	2.57735	0.88646	0.51746
贡献率(%)		59.29308	21.47791	7.38715	4.31217
累积贡献率(%)		59.29308	80.77098	88.5813	92.47031

### (3) 主成分得分

由此我们得到的两个主成分所代表的经济意义分别为：Z<sub>1</sub>—地区综合发展水平，Z<sub>2</sub>—地区发展水平与地区发展潜力的比较。

计算两个主成分的值，并根据主成分得分公式：

$$E=0.5929308Z_1+0.2147791Z_2$$

计算各个地区的主成分得分，结果如表 2。由表 2 可知，按照各个地区主成分得分排序，我们不难发现，总的发展水平（包括综合发展水平和地区发展潜力）在前 10 名的依次为上海、北京、广东、天津、浙江、江苏、辽宁、山东、福建、四川这 10 个地区，从这 10 个地区的分布来看，东部占 9 个，上海、北京、天津、福建这四个地区，经济发展水平很高，但是其潜力为负值，表明经济发展水平超过其潜力。而广东、浙江、江苏、辽宁、山东、四川这六个地区，发展潜力超过其经济发展水平，尤其是广东、江苏、山东和四川资源配置合理，可待发展潜力很大；位于中间 10 名的依次为河北、湖北、黑龙江、山西、河南、安徽、陕西、湖南、吉林和新疆，从这 10 个地区的分布来看，中部占 7 个，河南发展潜力最大；位于后 10 名的依次为内蒙古、海南、云南、江西、广西、青海、贵州、宁夏、甘肃和西藏，除海南和江西外，都是西部地区。由此说明中国各地区综合水平很不平衡，东部较发达，中部次之，而西部最为落后。

表 2 主成分值和主成分得分

按得分排序	按 $Z_1$ 值排序	按 $Z_2$ 值排序	地区	$Z_1$ 值	$Z_2$ 值	得分
1	1	25	上海	8.6248	-1.8011	4.7270
2	2	24	北京	7.1138	-1.7588	3.8402
3	4	4	广东	3.5306	2.3559	2.5994
4	3	26	天津	4.0343	-1.8027	2.0049
5	5	9	浙江	2.1503	0.8086	1.4487
6	6	2	江苏	1.2141	3.1470	1.3958
7	7	8	辽宁	1.1099	1.0724	0.8884
8	9	1	山东	0.1189	3.3178	0.7831
9	8	16	福建	0.5517	-0.2762	0.2678
10	11	3	四川	-0.7225	2.4915	0.1067
11	13	7	河北	-0.8911	1.6970	-0.1639
12	16	6	湖北	-1.1056	1.8716	-0.2536
13	14	11	黑龙江	-0.9053	0.4710	-0.4356
14	10	18	山西	-0.6991	-0.5576	-0.5343
15	27	5	河南	-1.8119	2.2850	-0.5836
16	21	12	安徽	-1.4073	0.4256	-0.7430
17	20	13	陕西	-1.3339	0.1874	-0.7507
18	23	10	湖南	-1.5088	0.5674	-0.7727
19	19	15	吉林	-1.2709	-0.2193	-0.8006
20	15	23	新疆	-0.9377	-1.1690	-0.8071
21	17	21	内蒙古	-1.1406	-1.0187	-0.8951
22	12	27	海南	-0.8562	-2.0070	-0.9388
23	22	19	云南	-1.4180	-0.7936	-1.0112
24	25	14	江西	-1.6795	-0.1618	-1.0306
25	28	17	广西	-1.8334	-0.4434	-1.1823
26	18	29	青海	-1.2130	-2.2880	-1.2106
27	29	22	贵州	-1.9000	-1.1460	-1.3727
28	24	28	宁夏	-1.5849	-2.0409	-1.3781
29	30	20	甘肃	-2.0313	-0.9903	-1.4171
30	26	30	西藏	-1.6910	-2.3400	-1.5052

众所周知，长江地带上的上海与环渤海地带的北京都是国际化的大都市，一个是中国的经济门户，一个是中国的政治门户，上海最特有的是其拥有曾经作为亚洲最大的金融、贸易和制造中心的“知识积累”或“历史记忆”。而作为首都的北京伴随申奥成功拥有商机无限，绿色北京和绿色奥运对资本的需求大增，日益完美的生活环境也使得消费市场潜力巨大，更加开放的北京对投资者的吸引力在迅速增强。与其强大的经济实力相比，其发展潜力相对较小，但是这并不阻碍它们发展。

## 6 图解分析和聚类分析

我们采用图解样本的方法，分别以  $F_1$ 、 $F_2$  的值为横坐标和纵坐标，绘制其散点图（见图 1）；同样，我们用聚类分析方法绘出谱系图（见图 2），这样就能非常直观和清楚地表示出各地区的发展水平及地区的发展潜力。

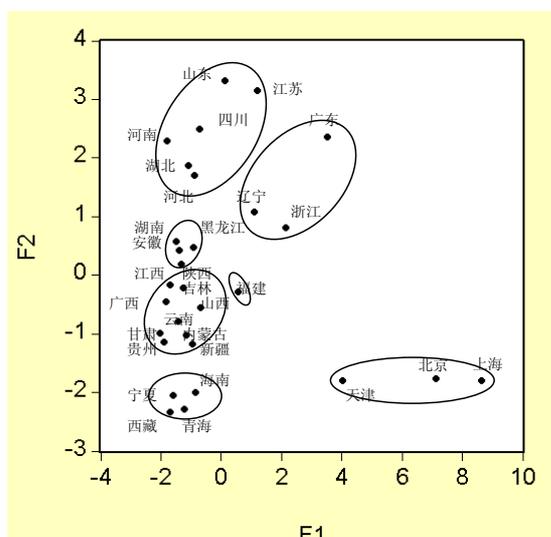


图 1 图解样本散点图

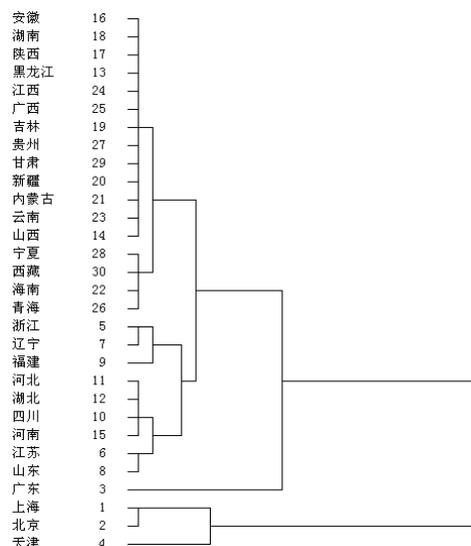


图 2 聚类分析谱系图

根据图 1 和图 2，我们把全国 30 个地区进行如下分类：高经济发展水平，潜力相对较小的地区：天津、上海、北京；较高经济发展水平，潜力较好的地区：辽宁、浙江、广东；经济发展水平较好，潜力一般的地区：福建；经济发展水平一般，高潜力的地区：河南、湖北、河北、四川、山东、江苏；经济发展水平一般，潜力也一般的地区：湖南、安徽、陕西、黑龙江；经济发展水平落后，潜力较差的地区：甘肃、贵州、广西、云南、内蒙古、新疆、江西、山西、吉林；经济发展水平落后，潜力最差的地区：西藏、宁夏、青海、海南。

## 7 结论

总的来看，综合经济发展水平较高的城市基本处于东部沿海一带，特别是改革开放较早的沿海经济特区，相比其它城市有较好的发展机遇，全方位优惠的政策和经济自主权，使它们率先成为中国充满经济活力的城市而集聚了较多资本。开放的市场环境使得人力资源充裕，能够满足各个层次的需求，但伴随城市经济发展水平的提高而提升的劳动力成本会对增量资本产生一定的排斥作用，所以更多的企业会谋求到中部城市发展。

位于中部的城市，近年来由于基础设施的着力改善，加之招商引资政策优惠力度的加大，伴随企业改制释放出大量劳动力质量好、价格低，整体环境得到较快发展。但一方面由于长期体制观念的羁绊，投资软环境的建设赶不上硬环境的建设速度；另一方面，城市的经济腹地效应和产业聚

合力还比较弱，所以位于中部地区的城市的综合经济发展水平都是偏低的。

排行靠后的城市基本位于西部地区，尽管政府出台了許多鼓励政策和优惠措施，但由于地理条件的局限，与经济基础相联系的物的因素与人的因素均不尽如人意，整体综合经济发展水平相对处于劣势，短时间内要形成规模引资的能力是十分有限的。

但是，东北地区相对于较为落后的经济发展水平而言，其经济发展潜力相对较大（从表 2 中  $F_2$  的排序可看出，辽宁、黑龙江、吉林的发展潜力分别位于第 8、第 11 和第 15 位）。东北地区具有对外开放的良好区位条件；东北地区拥有极为丰富的各种资源；东北地区拥有规模最大、数量最多、分布最为密集的国有大、中型企业；东北地区拥有大机器长期训练下素质较高的产业工人队伍；东北地区具有高度城市化和科技人才优势。再加上党的十一届六中全会做出了振兴东北地区等老工业基地的战略部署。对于如何振兴东北老工业基地，中央领导已经明确提出“东北地区等老工业基地具有重要的战略位置，要把老工业基地调整、改造和振兴提到更加突出的位置，用新思路、新机制、新方式，走出加快老工业基地的新路子。”有政府的支持，再加上东北等地区的自身优势，东北等地区的经济将在新世纪面临新的机遇。

#### 参考文献

- [1] Alvin C. Rencher. Methods of Multivariate Analysis [M]. 1995.
- [2] 国家统计局. 中国统计年鉴2003 [M]. 北京: 中国统计出版社, 2003.
- [3] 何晓群. 现代统计方法与应用 [M]. 北京: 中国人民大学出版社, 1998.
- [4] 于秀林, 任雪松. 多元统计分析 [M]. 北京: 中国统计出版社, 1999.
- [5] 王玲玲, 周纪芾. 常用统计方法 [M]. 上海: 华东师范大学出版社, 1993.

## Analysis of Actuality and Potential on Region Economics Development

Wu Guizhen, Shen Guihong, Ji Guanglin

(Business School of Jilin University , Changchun 130012, China)

**Abstract:** The principal component analysis can avoid the inaccuracy of the decision of weight due to subjective judge means, and reflects properly and effectively the weight of evaluation system. The paper compares and analyses comprehensive economic power and potential in China, suggests that the first principal component should be used as the measure of an area's synthetic economic power, and put in order for every area of China. At the same time the paper also analyses the development trend in the future.

**Keywords:** Region Economics Development; potential analysis; principal component analysis

收稿日期: 2004-8-10

基金项目: 吉林大学人文社会科学基金项目“地区经济差距的定量研究”;

**作者简介：**吴桂珍（1964- ），女，吉林省通化人，吉林大学商学院副教授，博士研究生。