

文章编号: 1674-9057(2014)04-0785-06

doi:10.3969/j.issn.1674-9057.2014.04.032

基于循环修正组合优化模型的 广西各城市综合经济实力评价

杨飞, 邓光明, 刘艳萍

(桂林理工大学 理学院, 广西 桂林 541004)

摘要: 选用2011年广西各市12项有关经济发展的指标数据,采用4种定性和定量的综合评价方法对广西各市的经济实力进行了评价。为了避免评价结果的不同,采用了循环修正的组合评价模型对广西各市的经济实力进行了组合评价,得到更加科学合理的广西各城市综合经济实力排名,为广西各市的经济发展规划与决策提供理论依据。

关键词: 组合优化模型; 组合评价模型; 综合经济; 广西

中图分类号: O157; F224

文献标志码: A

在对我国城市综合经济实力的评价研究中,不同学者从不同角度采用了不同的评价方法,如经济分析法、专家评价法、主成分分析法^[1]、因子分析法、TOPSIS法和熵值法^[2]等。这些评价方法大致分为两类,即主观赋权法和客观赋权法:主观赋权法无法避免人为因素带来的偏差;客观赋权法由于忽略了指标本身的重要程度,有时确定的指标权数与预期不一致,另外对于同样的指标体系,在不同的样本中客观赋权法对于同一指标确定的权数也可能不同。如果将单一模型进行组合,就能有效避开这两类模型各自的局限性,使结果更加全面、科学、合理。国际上对于组合评价的研究比较少,主要集中在单一评价方法^[3]及其结论不一致的原因的分析^[4]。国内研究有郭显光关于“平均值法”、“Borda法”以及“Copeland法”3种组合评价方法的论述^[5];唐俊等也对3种组合评价方法的有效性进行了分析^[6];由于3种组合评价方法结果也存在一些差异性,郭显光提出一种新的组合评价方法:模糊Borda法,对3种组合模型的结果进行组合^[7];刘艳萍也提出一种新的

组合评价方法,即组合优化模型^[8]。

笔者选用2011年广西壮族自治区关于经济发展的12项指标数据,分别采用综合指数法、加权平均法、因子分析法和熵值法,对广西壮族自治区14个市的综合经济实力进行了综合评价和排位;利用3种组合评价方法对4种单一评价方法重新进行综合评价和排位;最后使用循环修正的组合优化法,使14个市的排名达到一致,结果更加科学、合理,为广西各市的经济发展规划与决策提供理论依据。

1 模型的构建

1.1 指标体系的构建

城市综合经济实力是指城市所拥有的全部实力、潜力及其在国内外经济社会中的地位和影响力。由于城市的发展重心不同,选取指标时应该侧重共有的基础指标,基础指标的大小直接反映出一个城市的经济实力强弱。依据城市综合经济实力的要求,遵循科学性、系统性、可比性、典型

收稿日期: 2014-04-01

基金项目: 国家社会科学基金项目(13BTJ009); 广西自然科学基金项目(2014GXNSFBA118006); 广西高校人才小高地建设创新团队计划项目(桂教人[2011]47)

作者简介: 杨飞(1989—),男,硕士研究生,研究方向:应用统计,245565622@qq.com。

通讯作者: 邓光明,教授, dgm@glut.edu.cn。

引文格式: 杨飞, 邓光明, 刘艳萍. 基于循环修正组合优化模型的广西各城市综合经济实力评价 [J]. 桂林理工大学学报, 2014, 34(4): 785-790.

性与可操作性等原则,综合分析现有文献对城市综合发展指标体系的讨论和应用。本文从经济、人口、居住环境、交通4个方面入手,选取了反映城市综合发展水平的12个指标作为变量,分别为: X_1 ,非农业人口数; X_2 ,工业总产值; X_3 ,货运总量; X_4 ,批发零售住宿餐饮业从业人数; X_5 ,地方财政一般预算内收入; X_6 ,城乡居民储蓄年末余额; X_7 ,在岗职工平均人数; X_8 ,在岗职工工资总额; X_9 ,人均居住面积(市辖区); X_{10} ,每万人拥有公共汽车数; X_{11} ,人均城市道路面积; X_{12} ,人均绿地面积^[9]。

1.2 相关的数据分析及说明

本文选取广西14个城市2011年12个相关指标数据^[10]作为样本,所有结果都由SPSS 19.0计算完成。

原始数据有1项缺失值(百色市每万人拥有公共汽车数),为了不影响分析,使用均值代替缺失值。

由于12项原始指标的量纲不同,所以要对指标进行无量纲化处理,即第*i*个城市的第*j*项指标 A_{ij} 的无量纲化处理:

$$X_{ij} = \frac{A_{ij} - A_{\min(j)}}{A_{\max(j)} - A_{\min(j)}}, \quad (1)$$

其中: A_{ij} 为原始指标; $A_{\max(j)} = \max_i A_{ij}$; $A_{\min(j)} = \min_i A_{ij}$; $i = 1, 2, \dots, 14$; $j = 1, 2, \dots, 12$ 。

无量纲化后的数据 X_{ij} 在 $[0, 1]$ 区间内。本文的分析均使用无量纲化后的数据。

2 单项综合评价模型的建立

2.1 综合指数法

综合指数法是主观赋权法,即根据人们对各项经济指标的重要程度的理解,对各项经济指标赋予相应的权数。具体步骤为

①以各项指标的平均值为基数,求各地区每项指标的指数值

$$X_{ij}' = X_{ij} / \bar{X}_j, \quad i = 1, 2, \dots, 14; j = 1, 2, \dots, 12. \quad (2)$$

其中, \bar{X}_j 为第*j*个指标的样本平均值。

②主观的给出各项指标的经验权数 f_j , $j = 1, 2, \dots, 12$, 本文给定的权数为1,3,2,1,2,3,1,2,2,1,1,1。

③求各项指标指数值的加权平均数,得各地区综合经济指数 y_i 并排序。

$$y_i = \sum_{j=1}^{12} X_{ij}' f_j / \sum_{j=1}^{12} f_j, \quad i = 1, 2, \dots, 14. \quad (3)$$

2.2 加权算术平均法

加权算术平均法是一种客观赋权法,是对每个指标 X_i 给定权重 f_i ,然后用加权的算术平均数来综合计算。本文采用标准差系数作为相应的权重,其具体步骤如下:

①求原始指标的标准差系数

$$cv_j = S_j / \bar{x}_j, \quad j = 1, 2, \dots, 12. \quad (4)$$

其中: cv_j 是第*j*个指标的标准差系数; S_j 是第*j*个指标的样本标准差; \bar{x}_j 代表第*j*个指标的样本均值。

②用标准差系数确定的权重计算算术平均数

$$y_i = \sum_{j=1}^{12} x_{ij} f_j, \quad i = 1, 2, \dots, 14. \quad (5)$$

其中, $f_j = cv_j / \sum_{j=1}^{12} cv_j$, $j = 1, 2, \dots, 12$ 。

③求得各样本的分值并排序。

2.3 因子分析法

因子分析法^[11]的思想是根据相关性大小把原始变量分组,使得同组内的变量之间相关性较高,而不同组的变量间的相关性较低。每组变量代表一个基本结构,并用一个不可观测的综合变量表示,这个基本结构就称为公共因子。其步骤如下:

①求数据的标准化矩阵 X' 及其协方差矩阵,以及原始矩阵的相关系数矩阵 R 。

②求出相关系数矩阵 R 的特征值、贡献率和累积贡献率(表1),确定因子个数。

③提取前2个主成分 F_1 和 F_2 ,求出因子载荷矩阵(表2),为了方便解释各因子含义,对原始因子载荷矩阵进行正交旋转(表3)。

从旋转后的因子载荷矩阵可以看出,第1因子 F_1 在非农业人口数、货运总量、批发零售住宿餐饮业从业人数、地方财政一般预算内收入、城乡居民储蓄年末余额、在岗职工平均人数、在岗职工工资总额和人均绿地面积8个指标中有较大的载荷,称之为城市规模及经济发展水平;第2因子 F_2 在工业总产值、人均居住面积、每万人拥有公共汽车数、人均城市道路面积4个指标中有较大的载荷,称之为城市的基础设施水平和居民住房条件因子。

④计算各因子综合得分并进行排序。各因子得分由SPSS软件可以直接获得,以各个因子的信息贡献率为权数,计算各地区的综合得分:

表1 总方差解释
Table 1 Total variance explained

成分	初始特征值			提取载荷平方和			旋转载荷平方和		
	合计	方差贡献率/%	累积方差贡献率/%	合计	方差贡献率/%	累积方差贡献率/%	合计	方差贡献率/%	累积方差贡献率/%
1	9.052	75.431	75.431	9.052	75.431	75.431	6.706	55.884	55.884
2	1.335	11.122	86.553	1.335	11.122	86.553	3.680	30.669	86.553
3	0.645	5.377	91.931						
4	0.462	3.848	95.779						
5	0.269	2.245	98.024						
6	0.150	1.252	99.275						
7	0.066	0.550	99.825						
8	0.011	0.094	99.920						
9	0.005	0.043	99.963						
10	0.004	0.029	99.992						
11	0.001	0.005	99.997						
12	0.000	0.003	100.000						

表2 因子载荷矩阵
Table 2 Component matrix

指 标	F ₁	F ₂
Zscore:非农业人口数	0.980	-0.165
Zscore:工业总产值(当年价,万元)	0.769	0.288
Zscore:货运总量(万t)	0.626	-0.395
Zscore:批发零售住宿餐饮业从业人数	0.967	-0.222
Zscore:地方财政一般预算内收入(万元)	0.981	-0.103
Zscore:城乡居民储蓄年末余额(万元)	0.955	-0.190
Zscore:在岗职工平均人数	0.976	-0.179
Zscore:在岗职工工资总额(万元)	0.980	-0.174
Zscore:人均居住面积(市辖区)	0.657	0.622
Zscore:每万人拥有公共汽车数	0.796	0.287
Zscore:人均城市道路面积(m ²)	0.654	0.663
Zscore:人均绿地面积	0.945	-0.042

表3 旋转后的因子载荷矩阵
Table 3 Rotated Component matrix^a

指 标	F ₁	F ₂
Zscore:非农业人口数	0.909	0.402
Zscore:工业总产值(当年价,万元)	0.483	0.664
Zscore:货运总量(万t)	0.740	0.015
Zscore:批发零售住宿餐饮业从业人数	0.930	0.348
Zscore:地方财政一般预算内收入(万元)	0.876	0.455
Zscore:城乡居民储蓄年末余额(万元)	0.902	0.368
Zscore:在岗职工平均人数	0.912	0.389
Zscore:在岗职工工资总额(万元)	0.913	0.395
Zscore:人均居住面积(市辖区)	0.205	0.881
Zscore:每万人拥有公共汽车数	0.505	0.678
Zscore:人均城市道路面积(m ²)	0.180	0.913
Zscore:人均绿地面积	0.812	0.486

$$F = (55.884F_1 + 30.669F_2)/86.553. \quad (6)$$

2.4 熵值法

熵值法^[7]的思想是对于某项指标而言,如果该项指标值之间的差距越大,则该指标在综合评价中所起的作用越大;如果某项指标值全部相等,则该指标在综合评价中不起作用(表4)。

熵值法的步骤是:

①将各项指标同度量化,计算第j项指标下第i个城市指标值的比重 p_{ij}:

$$p_{ij} = x_{ij} / \sum_{i=1}^{14} x_{ij}, j = 1, 2, \dots, 12. \quad (7)$$

其中, x_{ij} 是第j项指标下第i个城市的指标值。

②计算第j项指标的熵值 e_j:

$$e_j = -k \sum_{i=1}^{14} p_{ij} \ln p_{ij}, j = 1, 2, \dots, 12. \quad (8)$$

其中, k > 0, e_j ≥ 0, 一般情况下, k = 1/ln n, n 为方案的个数,在本文中, k = 1/ln 14。

③计算权重 f_j:

$$f_j = \frac{1 - e_j}{\sum_{i=1}^{14} (1 - e_j)}, j = 1, 2, \dots, 12. \quad (9)$$

④计算各城市综合得分 y_i 并排序

$$y_i = \sum_{j=1}^{12} f_j p_{ij}, i = 1, 2, \dots, 14. \quad (10)$$

从4种单一模型的结果来看,大部分城市的排序是一致的,比如南宁、柳州、桂林、防城港、贵港、玉林、百色、贺州、来宾和崇左;但是还有一部分城市排序不一致,比如梧州、北海、钦州和河池。为了消除差异,使用组合评价模型。

表4 4种单一评价方法的结果
Table 4 Results of four single comprehensive models

城市	综合指数法		加权算术平均		因子分析		熵值法	
	得分	排序	得分	排序	得分	排序	得分	排序
南宁	0.898 249	1	0.935 291	1	2.140 713	1	0.940 586	1
柳州	0.602 771	2	0.552 350	2	0.809 756	2	0.543 524	2
桂林	0.423 597	3	0.416 390	3	0.480 650	3	0.414 094	3
梧州	0.224 945	7	0.203 664	7	-0.191 060	6	0.194 098	7
北海	0.241 730	6	0.208 732	6	-0.247 820	8	0.202 423	6
防城港	0.134 556	10	0.105 176	10	-0.409 070	10	0.092 407	10
钦州	0.182 814	8	0.137 885	9	-0.243 660	7	0.125 825	9
贵港	0.112 270	11	0.087 369	11	-0.419 140	11	0.084 373	11
玉林	0.309 013	4	0.277 258	4	0.109 103	4	0.273 958	4
百色	0.297 726	5	0.257 410	5	-0.044 730	5	0.247 722	5
贺州	0.015 410	14	0.014 015	14	-0.669 030	14	0.013 885	14
河池	0.152 450	9	0.147 005	8	-0.270 500	9	0.143 230	8
来宾	0.055 090	13	0.047 958	13	-0.545 620	13	0.044 383	13
崇左	0.073 993	12	0.069 921	12	-0.499 590	12	0.067 566	12

2.5 4种单一评价方法的比较

4种不同的方法从不同的角度得出评价结果,其结果可能会存在差异,但是对于同一样本,评价结果不应该有过大的差异。笔者使用 Spearman 等级相关系数法^[7]来检验不同评价方法评价结果的相关程度,其公式为

$$r = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}, \quad (11)$$

其中, d_i 为 2 种排序的等级差。

由式(11)可以计算出 4 种评价方法的等级相关系数矩阵(表 5)。

表5 4种模型的相关系数矩阵

Table 5 Correlation coefficient matrix of four models

	综合指数	加权算术平均	因子分析	熵值法
综合指数	1			
加权算术平均	0.995 6	1		
因子分析	0.986 8	0.978 0	1	
熵值法	0.995 6	1.000 0	0.978 0	1

给定显著性水平 $\alpha = 0.05$, $c_\alpha = 0.464$, $r_{ij} > c_\alpha$, 可以认为 4 种评价方法具有一致性。从表 5 可以看出, 4 种方法最终的结果关系相当密切, 其中加权算术平均法和熵值法的相关系数为 1; 而且 4 种单一评价方法结果具有一致性, 可以进行组合评价。

3 3种组合评价模型

3.1 平均值法^[5]

设 r_{ik} 为 x_i 方案在第 k 种方法下所排的位次 ($i = 1, 2, \dots, n; k = 1, 2, \dots, p$), 先用排序打分法将每种方法排序的名次转换成分数: $R_{ik} = n - r_{ik} + 1$, 即第 1 名得 n 分, 第 n 名得 1 分, 第 k 名得 $n - k + 1$ 分, 其中如果有相同的名次, 则取这几个位置的平均分; 然后计算不同方法得分的平均值 $\bar{R}_i = (\sum_{k=1}^p R_{ik})/p$, 按平均值重新排列, 若有 2 个方案得分相同, $\bar{R}_i = \bar{R}_j$, 则计算在不同方法下得分的方差 $\sigma_i^2 = \sum (R_{ik} - \bar{R}_i)^2/p$, 方差小者为优。

3.2 Borda 法^[5]

若评价认为 x_i 优于 x_j 的个数大于认为 x_j 优于 x_i 的方法个数, 记为 $x_i S_{x_j}$ 。定义 Borda 矩阵 $B = \{b_{ij}\}_{n \times n}$, 其中:

$$b_{ij} = \begin{cases} 1, & x_i S_{x_j}; \\ 0, & \text{其他。} \end{cases} \quad (12)$$

地区 i 的得分 $b_i = \sum_{j=1}^n b_{ij}$, 按得分 b_i 重新排序, 若有 $b_i = b_j$, 如 3.1 节方法中方差小者为优。

3.3 Copeland 法^[5]

Copeland 法与 Borda 法不同之处是在计算“优”次数的同时还要计算“劣”的次数, 即定义

$$c_{ij} = \begin{cases} 1, & x_i S_{x_j}; \\ 0, & \text{其他}; \\ -1, & x_j S_{x_i}. \end{cases} \quad (13)$$

再定义方案 x_i 的得分为 $c_i = \sum_{j=1}^n c_{ij}$, 根据 c_i 的大小对 x_i 重新排序, 若有 $c_i = c_j$, 如 3.1 节方法中方差小者为优。

3.4 计算组合评价价值

通过 3 种组合评价模型计算得分并重新排序后, 计算结果如表 6 所示。从结果可以明显看出, 3 种组合模型最终排名基本一致。但是钦州和河池排名仍略有差异: 在平均值法中, 钦州排名高于河池; 在 Borda 法和 Copeland 法中, 2 个城市的得分是一样的。除此之外, 其他城市的排名完全一致。

表 6 3 种组合评价模型的结果

Table 6 Results of three combination evaluation models

城市	平均值法		Borda 法		Copeland 法		方差
	得分	排序	得分	排序	得分	排序	
南宁	14	1	13	1	13	1	0
柳州	13	2	12	2	11	2	0
桂林	12	3	11	3	9	3	0
梧州	8.25	7	7	7	1	7	0
北海	8.5	6	8	6	3	6	0
防城港	5	10	4	10	-5	10	0
钦州	6.75	8	5	9	-2	9	0.67
贵港	4	11	3	11	-7	11	0
玉林	11	4	10	4	7	4	0
百色	10	5	9	5	5	5	0
贺州	1	14	0	14	-13	14	0
河池	6.5	9	5	8	-2	8	0.67
来宾	2	13	1	13	-11	13	0
崇左	3	12	2	12	-9	12	0

进行组合评价, 可使组合后结果的方差持续减小; 重复使用上步, 直至方差趋近于 0, 即几种评价结果完全相同。

组合优化模型的步骤是:

①验证组合评价模型结果是否有差异, 若有差异, 对几种评价组合值重新组合;

②重复步骤①, 至评价模型结果方差趋近于 0 或者几种组合评价模型结果一致时结束。

通过表 6 发现, 3 种组合评价方法的结果不一致, 钦州和河池的排序方差不为 0, 因此以 3 种组合评价方法的得分值为依据, 继续使用 3 种组合评价方法进行组合评价结果见表 7。

可见利用组合评价方法重新组合评价后, 可以看到 3 种方法的结果一致, 方差为 0, 循环结束。

表 7 重新组合后的评价模型结果

Table 7 Re-evaluation results of the combined model

城市	平均值法		Borda 法		Copeland 法		方差
	得分	排序	得分	排序	得分	排序	
南宁	0.33	1	13	1	13	1	0
柳州	0.67	2	12	2	11	2	0
桂林	1.00	3	11	3	9	3	0
梧州	2.33	7	7	7	1	7	0
北海	2.00	6	8	6	3	6	0
防城港	3.33	10	4	10	-5	10	0
钦州	3.00	8	6	8	-1	8	0
贵港	3.67	11	3	11	-7	11	0
玉林	1.33	4	10	4	7	4	0
百色	1.67	5	9	5	5	5	0
贺州	4.67	14	0	14	-13	14	0
河池	2.67	9	5	9	-3	9	0
来宾	4.33	13	1	13	-11	13	0
崇左	4.00	12	2	12	-9	12	0

4 组合优化模型

由于 3 种组合模型结果略有差异, 无法准确地对各城市作出排序, 所以下面使用组合优化模型, 对结果进行优化, 使各城市最终排序趋于一致。

组合优化模型的思想^[7]: 单一评价模型对同一问题采用不同的方法进行综合评价, 其结果应该具有一致性。通过对单一模型的有效组合, 可以减小评价的方差, 这就是组合评价模型。当结果有差异时, 可以使用常见的组合评价方法继续

5 模型和结果分析

(1) 从单一的综合评价模型来看, 对于同一问题, 使用不同的综合评价模型, 其结果应该具有一致性。

(2) 主观赋权法和客观赋权法都有自身的局限性, 使用组合评价模型进行综合, 能够得到更加科学、合理的评价结果。

(3) 不同的组合评价模型最后得到的结果也可能不一致, 组合优化模型可以很好的综合各个组合评价的结果, 使最后结果一致。

(4) 从结果来看: 南宁、柳州和桂林稳居广西城市经济综合实力的前3名; 最后3名是崇左、来宾和贺州, 这些城市应该根据各自经济的构成特点, 加快促进自身的经济发展, 缩小和其他城市的差距; 钦州和河池最后的排名也可以确定, 钦州排名第8、河池第9; 组合优化模型最后的结果与人们直观排序是基本一致的, 说明了组合优化模型的有效性、实用性、科学性。

参考文献:

[1] 刘艳春, 高立. 组合评价模型在辽宁省主要地区综合经济实力评价中的构建和应用 [J]. 数学的实践与认识, 2005, 35 (3): 44 - 50.

[2] 郭丽婷, 姚凤桐. 内蒙古自治区各盟市综合经济实力组合评价研究 [J]. 内蒙古科技与经济, 2012, 11 (6): 49 - 50.

[3] Warnick G R, Albers J J. A comprehensive evaluation of the heparin-manganese precipitation procedure for estimating high

density lipoprotein cholesterol [J]. Journal of Lipid Research, 1978, 19 (1): 65 - 76.

[4] Clemen R T, Fischer G W, Winkler R L. Assessing dependence: Some experimental results [J]. Management Science, 2000, 46 (8): 1100 - 1115.

[5] 郭显光. 一种新的综合评价方法——组合评价法 [J]. 统计研究, 1995 (5): 56 - 59.

[6] 唐俊, 张明清. 三种组合评价方法的有效性分析 [J]. 计算机工程与设计, 2009, 30 (15): 3568 - 3572.

[7] 郭显光. 熵值法及其在综合评价中的应用 [J]. 财贸研究, 1994 (6): 56 - 60.

[8] 刘艳春. 一种循环修正的组合评价方法 [J]. 数学的实践与认识, 2007, 37 (2): 88 - 94.

[9] 蔡龙, 章波, 黄贤金, 等. 我国城市基础设施现代化水平综合评价研究 [J]. 城市基础建设, 2004 (4): 50 - 54.

[10] 广西壮族自治区统计局. 广西统计年鉴——2012 [M]. 北京: 中国统计出版社, 2012.

[11] 何晓群. 多元统计分析 [M]. 北京: 中国人民大学出版社, 2008: 57 - 95, 192 - 206.

Evaluation of comprehensive economic strength for Guangxi cities based on cycling modification of combination optimization model

YANG Fei, DENG Guang-ming, LIU Yan-ping

(College of Science, Guilin University of Technology, Guilin 541006, China)

Abstract: The combination optimization model can be useful for reducing the variance and optimizing the results of different combination of evaluation models. 12 indicators are selected which are related to the city economic development of Guangxi in 2011, by 4 qualitative and quantitative methods to evaluate the economic strength of cities in Guangxi. The cycling modification of combination optimization model is applied in order to unify the evaluation results. More reasonable rank of the economic strength of cities in Guangxi is got, with a theoretical basis for the economic development program and decision-making.

Key words: combination optimization model; combination evaluation model; comprehensive economic; Guangxi