

文章编号: 1005—0523(2006)02—0040—04

西部地区交通区域划分问题的研究

李 蓉, 李 宇

(华东交通大学 信息工程学院, 江西 南昌 330013)

摘要:西部地区幅员广阔, 交通与经济的发展表现出很大的地区差异性, 同时交通与社会经济的发展又有着密不可分的关系, 在不同的现实条件下, 交通发展与经济发展之间的适应性严重影响和制约这一个地区的协调与进步. 本文从交通与经济发展相协调的角度出发, 利用主成分分析与层次聚类分析方法对西部地区进行交通区域划分, 研究结果对于了解西部区域间交通发展的差异性及其制定区域交通策略具有一定的意义.

关 键 词:交通区域划分; 西部; 主成分分析; 聚类分析

中图分类号:U11

文献标识码:A

1 引言

西部开发, 交通先行. 我国西部地区幅员辽阔, 各个省份生产力发展水平、产业结构、人口状况、地理环境、民族构成、交通设施的差异很大, 形成各个地区发展不平衡的显著特征. 根据德国科尔(Kohl J. G. 1841)的交通区位理论, 交通发展与人口集中、聚落形成以及经济状况有不可分割的联系, 交通的形成与发展带有一定的区域性. 某一区域的交通建设是该区域社会经济中的一个子系统, 它必须适应地区社会经济发展, 同时又受到地区经济水平的制约. 因此西部地区各省区间的差异性决定了不同地区应采取不同的交通发展策略, 不可一味的照搬照抄中东部地区的模式, 全部“一刀切”, 必须实事求是、具体情况具体对待. 自我国开始逐步实施“贷款修路, 收费还贷”的政策以来, 更是加大了西部地区之间交通发展的不平衡和差别. 在合理选择西部交通的发展模式和研究发展战略问题时, 如何根据区域的差异科学决策, 合理进行投资, 进行区域划分就具有其存在的现实意义.

了解各地区的社会经济发展和交通发展处于

什么水平, 在整个西部范围内属于哪一层次, 可以采取什么样的发展策略和技术水平, 这便是进行交通经济区域划分要解决的问题. 本文首先建立了反映各地区社会经济与交通发展水平的共 11 个评价指标, 然后采用主成分分析方法与聚类分析方法对西部地区的 12 个省市进行了区域划分, 划分出的交通经济区域是在综合了各地的交通发展水平和社会经济发展水平后划分出的一种特定意义的区域, 它可以反映这一地区公路交通发展与社会经济发展的综合水平, 以便在进行交通投资决策与评价时具有较准确的定位, 减少投资的盲目性, 促进区域公路交通发展与区域整个社会经济的协调发展.

2 区域划分方法

本研究采用数理统计中的主成分分析法和聚类分析方法对所收集到的数据进行分析, 给出了西部地区的交通经济区域划分方法和结果.

2.2 主成分分析

主成分分析法通过研究指标体系的内在结构关系, 从而将多个指标体系转化为一个或几个相互独立的综合指标, 并且这些少量的综合指标能够反

收稿日期:2006—02—22

作者简介:李蓉(1979—), 女, 山西长治人, 讲师, 主要从事智能控制方面的研究.

映原来多个指标所反映的绝大部分信息(80%或85%以上).其优点在于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.我们将以上所选取的指标进行主成分分析,步骤如下:

记原数据表为 $X=(x_{ij})_{np}$;其中 n 为样本点个数, p 为变量个数.首先,对数据进行标准化处理,这样可以消除变量的量纲影响,同时将坐标原点移到数据的重心.

* 记 $(x_{ij})'_{np}$ 为标准化后的数据表;
 $x'_{ij}=(x_{ij}-\bar{x}_j)/s_j, i=1,2,\dots,n; j=1,2,\dots,p$
其中: $\bar{x}_j=\frac{1}{n}\sum_{i=1}^n x_{ij}$, 为第 j 个变量的样本均值;
 $s_j=\frac{1}{n}\sum_{i=1}^n (x_{ij}-\bar{x}_j)^2$, 为第 j 个变量的样本标准差.
不失一般性,仍记标准化后的数据表为 X .

* 计算 X 矩阵的相关矩阵 R , 求出 R 的前 m 个特征值 $l_1, l_2, \dots, l_m$; 及对应的特征向量 $m_1, m_2, \dots, m_m$, 它们是标准正交的, m_k 称为主轴.

* 记 y_h 为第 h 主成分, 则 $y_h=X'm_h$.

2.2 层次聚类分析方法

聚类分析(Cluster Analysis)是统计学所研究的“物以类聚”问题的一种方法,能够将一批样本数据按照他们在性质上的亲疏程度在没有先验知识的情况下自动进行分类.常用的聚类方法有:C—均值聚类法、K—均值聚类法、层次聚类分析法等,本文采用层析聚类方法.

层次聚类分析(Hierarchical Cluster Analysis)方法的基本思想是:在聚类分析的开始,每个样本自成一类;然后,按照某种方法度量所有样本之间的亲疏程度,并把其中最亲密或最相似的样本首先聚成一小类;接下来,度量剩余的样本和小类之间的亲疏程度,并将当前最亲密的样本或小类再聚成一类;如此反复,直到所有的样本分别聚成一类为止.由此可见,层次聚类方法中,度量数据之间的亲疏

程度是极为关键的.
假设有 n 个对象,描述第 i 个对象的 m 个属性值分别对应于连续变量值 $x_{i1}, x_{i2}, \dots, x_{im}$, 描述第 j 个对象的 m 个属性值分别对应于连续变量值 $x_{j1}, x_{j2}, \dots, x_{jm}$, 那么对象 i 与 j 之间的差异度 $d(i, j)$ 的计算方法主要有如下3种:

- * 欧几里德距离(Euclidean distance)
$$d(i, j)=\sqrt{(x_{i1}-x_{j1})^2+(x_{i2}-x_{j2})^2+\dots+(x_{im}-x_{jm})^2}, \quad i \in \{1, 2, \dots, n\}; j \in \{1, 2, \dots, n\}$$
- * 绝对值距离(Manhattan distance)
$$d(i, j)=|x_{i1}-x_{j1}|+|x_{i2}-x_{j2}|+\dots+|x_{im}-x_{jm}|, \quad i \in \{1, 2, \dots, n\}; j \in \{1, 2, \dots, n\}$$
- * 明考斯基距离(Minkowski distance)
$$d(i, j)=(|x_{i1}-x_{j1}|^p+|x_{i2}-x_{j2}|^p+\dots+|x_{im}-x_{jm}|^p)^{1/p}, \quad i \in \{1, 2, \dots, n\}; j \in \{1, 2, \dots, n\}$$

距离越近,表明对象 i 与 j 之间越相似,差异越小;距离越远,表明对象 i 与 j 之间越不相似,差异越大.样本数据与小类、小类与小类间亲疏程度的度量方法主要有平均值距离、平均距离、最大距离、最小距离等.

3 交通区域划分

为了能综合反映各地区经济与交通发展水平,本研究共选取反映区域经济和交通发展水平两个方面共11个主要指标进行评价和划分.同时为减少运算和分析问题带来的不必要的麻烦,抓住事物的主要矛盾,首先采用主成分分析方法对原始指标进行处理,然后用层次聚类分析方法,进行区域分类.

3.1 评价指标的确定

我们本着系统、客观的原则共选取了反映各地区公路交通发展水平和社会经济发展水平的11个指标来完成区域划分.各指标及各地区的取值如表1所示.

表1 西部各地区各项评价指标值

指标	人均 GDP (x_1)	人口 (x_2)	消费价格 指数(x_3)	民用汽车 拥有量(x_4)	旅客周 转量(x_5)	货运周 转量(x_6)	公路面积 密度(x_7)	公路人口 密度(x_8)	公路投资额 增长率(x_9)	高速公路 面积密度(x_{10})	高速公路 人口密度(x_{11})
地区	元/人	万人	上年=100	万辆	亿人公里	亿吨公里	公里/百平 方公里	公里/万人	百分比	公里/百平方 公里	公里/万人
内蒙古	8 975	2 380	102.2	49.85	204.8	1 160.3	6.7	31.2	31.9	0.03	0.138
广西	5 969	4 857	101.1	43.44	463.4	863.4	25.4	12	65.5	0.44	0.208
重庆	7 209	3 130	100.6	34.25	235.7	367.7	38.3	10	61	0.707	0.185
四川	6 418	8 700	101.7	113.87	547.9	768.3	23.1	12.9	41.8	0.308	0.173

贵州	3 603	3 870	101.2	32.46	248.6	547	26.6	11.7	32.5	0.19	0.083
云南	5 662	4 376	101.2	77.09	229.5	612.2	42.2	38	9	0.27	0.243
西藏	6 871	270	100.9	6.06	5.5	27.1	3.4	152.9	108.9	0	0
陕西	6 480	3 690	101.7	54.01	382.6	849.1	24.4	13.6	−5.2	0.412	0.229
甘肃	5 022	2 603	101.1	27.39	243.6	738.7	9	15.5	39.1	0.076	0.131
青海	7 277	534	102	10.47	38	124.2	3.4	45.7	74	0.016	0.221
宁夏	6 691	580	101.7	12.39	51	244.5	18.1	20.5	21.3	0.797	0.906
新疆	9 700	1 934	100.4	46.76	249.8	636.6	5.227	43.245	193.4	0.027	0.223

注:数据资料来源于“中国统计年鉴 2004”。高速公路里程不包括城市主干线。

3.2 主成分分析

本文借助于 SPSS 统计分析软件包完成数据处理及分析。SPSS (Statistical Product and Service Solutions) 是一种为统计产品与服务提供解决方案的软件包, 为了实现对西部地区的划分, 我们分别利用了 SPSS 中的“Analyze”模块中的“Data Reduction”与“Classify”功能来实现主成分分析与聚类过程。通过对西部 12 个省份地区的各项数据作标准化处理以及相关系数矩阵的计算, 我们取前 4 个主成分(累积贡献率达到了 86.691%)进行下一步的计算, 各指标因子在主成分中的载荷如表 2 所示。

表 2 旋转前后各指标因子在主成分中的载荷									
变量	旋转前				旋转后				
	主成分 1	主成分 2	主成分 3	主成分 4	主成分 1	主成分 2	主成分 3	主成分 4	
1	−0.377	0.327	0.268	0.748	−0.014	0.027	0.937	−0.051	
2	0.893	0.312	−0.127	−0.045	0.902	−0.002	−0.313	−0.049	
3	0.184	−0.211	0.901	−0.187	0.048	0.0300	0.013	0.960	
4	0.793	0.405	0.029	0.136	0.900	−0.017	−0.047	0.020	
5	0.861	0.381	−0.095	0.136	0.944	0.032	−0.124	−0.081	
6	0.660	0.421	0.382	0.164	0.808	−0.085	0.153	0.321	
7	0.706	−0.344	−0.478	−0.074	0.423	0.471	−0.622	−0.251	
8	−0.755	0.235	−0.125	−0.251	−0.600	−0.524	0.144	−0.220	
9	−0.589	0.520	−0.351	0.445	−0.188	−0.294	0.645	−0.633	
10	0.413	−0.749	−0.265	0.351	0.078	0.919	−0.255	−0.103	
11	0.010	−0.782	0.187	0.487	−0.244	0.863	0.150	0.237	

我们可以利用西部各地区第一、第二、第三、第四主成分得分, 进行中国交通经济区域的划分。

$$y_1 = -0.014x_1 + 0.902x_2 + 0.048x_3 + 0.900x_4 + 0.944x_5 + 0.808x_6 + 0.423x_7 - 0.600x_8 - 0.188x_9 + 0.078x_{10} - 0.244x_{11}$$
$$y_2 = 0.027x_1 - 0.002x_2 + 0.030x_3 - 0.017x_4 + 0.032x_5 - 0.085x_6 + 0.471x_7 - 0.524x_8 - 0.294x_9 + 0.919x_{10} + 0.863x_{11}$$
$$y_3 = 0.937x_1 - 0.323x_2 + 0.013x_3 - 0.047x_4 - 0.124x_5 + 0.153x_6 - 0.622x_7 + 0.144x_8 + 0.645x_9 - 0.255x_{10} + 0.150x_{11}$$

$$y_4 = -0.051x_1 - 0.049x_2 + 0.960x_3 + 0.020x_4 - 0.081x_5 + 0.321x_6 - 0.251x_7 - 0.220x_8 - 0.633x_9 - 0.103x_{10} + 0.237x_{11}$$

再利用这 4 个主成分各自的方差贡献率, 加权合成构造西部地区交通经济发展综合评价模型:

$$Y = 39.708y_1 + 21.384y_2 + 13.891y_3 + 11.708y_4$$

这样我们可计算出各因子的得分, 从而得到各区域的因子得分情况(见表 3 所示), 这样我们就可以对我国的区域进行分类了。

表 3 西部各地区各项主成分因子的得分				
地区名称	主成分 1 得分	主成分 2 得分	主成分 3 得分	主成分 4 得分
内蒙古	0.477	−0.616	1.295	1.779
广西	0.822	0.285	−0.164	−0.477
重庆	−0.004	1.144	−0.351	−1.465
四川	1.855	−0.185	−0.078	0.179
贵州	−0.091	−0.532	−1.591	−0.086
云南	0.388	0.113	−1.060	−0.220
西藏	−1.703	−1.512	−0.124	−0.839
陕西	0.591	0.280	−0.286	0.772
甘肃	−0.107	−0.678	−0.507	0.261
青海	−1.232	−0.513	0.365	0.947
宁夏	−1.280	2.403	0.275	0.747
新疆	0.284	−0.189	2.225	−1.597

3.3 聚类分析

我们利用这 4 主成分因子用层次聚类分析方法进行分类, 可以把西部 12 个地区划分为 4 个交通经济发展水平层次, 划分结果见表 4 所示。

表 4 西部地区交通经济区域划分			
交通经济水平层次	主成分综合得分	交通经济水平评价	区域范围
第一层次	≥30	发达	四川、广西、陕西、内蒙古
第二层次	10~30	中等发达	宁夏、新疆
第三层次	0~10	初级发达	云南、重庆
第四层次	≤0	不发达	贵州、青海、甘肃、西藏

从表 4 的层次划分可见, 我国西部地区各省经济发展水平和交通发展水平存在明显的差异, 其总

体格局是以四川、广西、陕西、内蒙古构成了西部地区的经济发达地区,中等发达地区为宁夏、新疆;初级发达地区集中为云南、重庆;而贵州、青海、甘肃、西藏构成了西部地区中最不发达的区域。

4 西部交通经济区域差异分析及对策研究

划分结果反映了各类型地区的相似性及其亲疏程度,在对我国西部各地区进行交通经济区域划分时,高于发展平均水平、评价值为正值地区中有8个,它们是四川、广西、陕西、内蒙古、宁夏、新疆、云南、重庆;其余各省份地区均处于平均发展水平之下,评价值最高的为四川(70.69616),评价值最低的为西藏(-111.503),二者相差达182.1989。可见西部地区的整体交通发展与经济发展水平存在较大的不平衡性,对西部地区的交通规划应该在充分考虑这种不平衡性的基础上因地制宜,取得最佳的经济效益。

第一层次:包括四川、广西、陕西、内蒙古,综合评价值在30分以上,为西部地区经济与交通环境发达的地区。这些西部相对发达的地区,人口多、生活水平高、外部投资环境与交通发展环境优越、竞争力强,发展势头好,是带动整个西部发展的核心地区。对于这类区域的公路收费系统的建设,宜本着适度超前的原则,交通系统规划与建设可参照东部发达地区的作法。

第二层次:包括宁夏、新疆,综合评价值在10~30分之间,为西部地区经济较为发达、具有较好发展潜力的地区,但交通运输事业的发展相对薄弱。

对于这类区域的交通建设,应该做好系统规划,根据地区发展情况作出适当的调整,不宜过于超前,也不宜太过落后,否则会对区域发展造成阻碍。

第三层次:包括云南、重庆,综合评价值在0~10分之间,综合评分处于平均水平之上,具有一定经济发展潜力、交通条件落后,收费公路密度均处于平均水平之下,是整个西部初级发达的区域。区域交通须根据其外部影响因素做出合理的选择。

第四层次:包括贵州、青海、甘肃、西藏,综合评价值在0分以下,处于平均水平之下,是西部地区最不发达的区域。该区域交通建设应采取“适当落后”的原则,达到节约系统运营成本和养护成本的目的。

由以上的分析,西部要发展,公路建设必须满足经济增长的需要。区域划分结果不打乱行政区划,而是若干省、自治区的联合。我们并不要求建立什么行政组织,而是强调在进行系统决策时要有区域的概念,要谋求各种资源最有效的利用,促进区域经济、交通的发展。

参考文献:

[1] 陈大道.中国区域发展的理论与实践[M].北京:科学出版社,2003.
[2] 陈栋生.中国区域经济学[M].北京:经济科学出版社,2005.
[3] 徐文洁,杨德刚,张小雷.西部地区投资环境的区域差异[J].干旱区资源与环境,2004,(18),2.
[4] 王 斌,杨德刚,徐文洁.西部十二省(市)经济发展水平的区域差异研究[J].干旱区资源与环境,(18),1.

The Study for Regional Clusters of Traffic in the Western Region

LI Rong, LI Yu

(School of Information Engineering, East China Communication University, Nanchang 330013, China)

Abstract: The Western Region is with a vast territory, traffic development and social-economic development in a society has not only put up obvious regional otherness, but also existed intrinsic relationship. In the different condition, the adaptability of Traffic development, social-economic development and regional progress have been deeply effected and restricted regional progress. Grounded on the equilibrium of traffic development and social-economic development in the Western Region, this study introduce cluster analysis and principal component analysis method to do regional clusters of traffic in the western region, the significance of this study lies in its contribution to regional traffic policy.

Key words: regional clusters of traffic; western region; cluster analysis; principal component analysis