

# 泛珠江三角洲区域科技资源配置效率综合评价

## ——基于遗传投影寻踪方法

李浩宾

(华南师范大学 数学科学学院, 广东 广州 510631)

**摘 要:**在系统分析和总结前人相关研究成果的基础上,结合泛珠江三角洲区域的特点,构建了由3个层次指标组成的区域科技资源配置效率评价指标体系,然后剖析了投影寻踪技术的实质,针对评价科技资源配置效率的数据具有高维、非正态特性而不适合用传统统计方法进行分析的实际,提出了应用该指标体系和投影寻踪技术综合评价区域科技资源配置效率的新方法,对泛珠江三角洲区域科技资源配置的现状进行系统的分析和评价。

**关键词:**区域科技;科技资源;配置效率;投影寻踪;遗传算法

中图分类号:F127.65

文献标识码:A

文章编号:1001-7348(2009)06-0133-05

## 0 引言

目前国内对科技资源这一概念比较一致认同的是:科技资源是科技人力资源、科技财力资源、科技物力资源及科技信息资源,科技资源配置是指在一定的经济体制及其运行机制下使科技资源产生正向效果、效率的调配方式<sup>[1]</sup>不同调配方式会造成不同的投入产出,也就是配置效率的不同。而配置效率的高低在一定程度上决定着科技能力的强弱。对于配置效率国内学者有不同的评价方法<sup>[2-4]</sup>,本文将结合泛珠江三角洲区域的特点,构建了由3个层次指标组成的区域科技资源配置效率评价指标体系,然后剖析了投影寻踪技术的实质,针对评价科技资源配置效率的数据具有高维、非正态特性而不适合用传统统计方法进行分析的实际,提出了应用该指标体系和投影寻踪技术综合评价区域科技资源配置效率的新方法,对泛珠江三角洲区域科技资源配置的现状进行系统的分析和评价。

泛珠江三角洲的概念缘起于珠江流域,福建、江西、广西、海南、湖南、四川、云南、贵州和广东9省,位于我国广大的华南、西南地区,直接或间接、或多或少都跟珠江有关。这一区域地域辽阔,相互联系密切,面积199.45万平方千米,人口4.46亿人,占全国面积的20.78%,人口的34.76%。再加上香港、澳门,习惯称为“9+2”。泛珠江三角洲区域包括了东部、中部和西部的省份,该区域的发展有利于我国经济区域协调发展战略的实施,符合我国经济发展的总体战略。科技是第一生产力,泛珠江三角洲区域合理、高效地配置科技资源,提高其配置效率,直接影响该区域经

济的发展,进而影响全国经济的发展。因此正确有效评价该区域的科技资源配置效率是有意义的。

## 1 泛珠江三角洲区域科技资源投入现状分析

20世纪90年代以来,通过实施国家“863”计划、“973”计划、科技攻关计划、自然科学基金资助项目、西部大开放等一系列科技计划以及其它政策措施,政府加大科技投入力度,带动全社会科技投入,有力地促进了经济增长。

(1)区域科技投入总量增加。2005年区域的科技活动经费支出总额为1 074.29亿元,比1995年增长5.242倍。科技活动人员总量由58.46万人升到92.4万人。

(2)R&D经费支出进入高速增长时代。2004、2005年区域R&D经费支出为428.83亿元和515.54亿元,同比增长为14%和20%,远远高于“九五”期间7.3%的年平均增长速度,也大大高于同期主要发达国家的5%左右的增长速度。但是从总量来看,还是比较薄弱的。据2000年R&D资源数据,泛珠三角9个省(区)研究与试验发展人员总数仅占全国的25.4%,R&D经费内部支出仅占全国的24.7%,均比国内生产总值占全国的33.3%要低约8个百分点。区域九省的科学研究和综合技术服务业的R&D人员占全国的20.0%,R&D经费则仅占16.8%。科研机构R&D人员占19.6%,经费占17.1%,高校R&D人员占24.7%,经费占17.3%。区域科技人力投入相对于科技财力投入相对过剩,或者说人均科研经费偏低,科技财力投入相对于人力投入显得更为短缺。

(3)专利数量、论文数量及水平近年有显著增长。2006

年,泛珠三角区域“9+2”省区国内专利申请及授权分别为139 482件和70 770件,同比增长22.2%和24%,分别占国内专利申请总量及授权总量的29.7%和31.6%;其中,泛珠三角区域内发明专利申请及授权分别为33 537件和5 078件,同比增长48.5%和21.2%,分别占国内发明专利申请总量及授权总量的27.4%和20.2%。

在“9+2”省区中,广东以90 886件专利申请位居全国首位,四川、福建、湖南分别以13 109件、10 351件和10 249件分列全国第11、13位以及第14位。据了解,从1985~2006年,泛珠三角区域“9+2”省区共申请国内专利793 919件,授权专利462 189件,分别占全国的29.1%和31%。

(4)分析9个省份投入状况,比较突出的是广东与四川,2000年R&D人员占区域总量的56.1%,R&D经费占68.8%,是区域科技投入大省。贵州、海南2000年R&D人员只占区域总量的5.21%,R&D经费占2.26%,人均R&D经费只有区域平均水平的60%,贵州、海南科技投入薄弱,科技能力弱。

## 2 区域科技资源配置研究常用的传统统计分析方法技巧不足

目前科技资源配置研究使用的统计分析方法通常有回归分析、聚类分析、主成分分析等。但是这些传统的多元统计分析方法是建立在总体服从某种分布,比如正态分布这个假定的基础上的,这就是采用所谓的“假定—模拟—检验”的数据分析法(Confirmatory Data Analysis,简称CDA)。如果实际中数据偏离正态分布,结果就有可能发生畸变,需要用稳健的或非参数方法来解决。遗憾的是当数据的维数较高时,遇到很大的困难。

(1)随着数据维数的增高,数据的计算量迅速增大,传统的多元统计分析方法很难处理大量的数据。

(2)对于高维数据,存在着高维空间中点稀疏的“维数祸根”,传统的多元统计分析方法也很难处理。

(3)传统的多元统计分析方法用于低维数据稳健性较好,但用于高维数据时,稳健性较差。

投影寻踪方法作为一种新的多元统计分析方法,基本思想就是把高维数据投影到低维子空间,使得投影子空间中得到的构形最能反映原高维数据的结构和特征,是一种稳健的适应于高维、非正态数据的探索性数据分析方法。国内外在工业、农业、水利、医学及遥感等领域已有成功应用的例子<sup>[6-9]</sup>。本文区域科技资源配置效率研究流程图如图1所示。

## 3 基于遗传算法的投影寻踪模型

本文采用动态多指标投影寻踪模型,直接利用具有省份、时间和指标的三维数据矩阵,通过某种组合投影到低维空间上,利用全局收敛的加速遗传算法优化投影方向,在低维空间上得到多维数据的综合评价。投影寻踪模型的建模过程包括如下6个步骤<sup>[6]</sup>:

步骤1:建立决策矩阵。

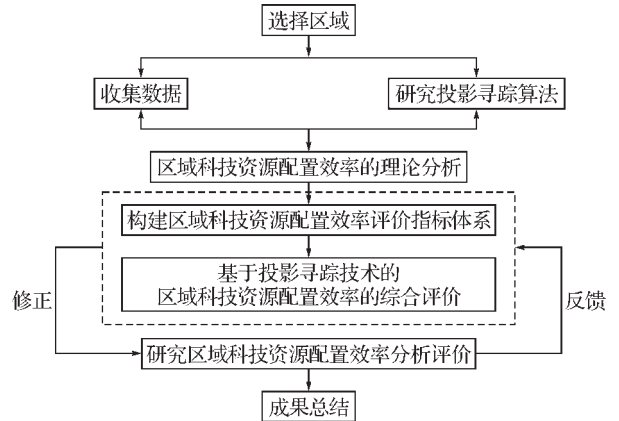


图1 区域科技资源配置效率研究流程

设区域科技资源配置问题的时段为 $T_i$ ,指标为 $P_j$ ,拟定的省份为 $S_k$ ,并假定 $S_k$ 均为有效解,省份 $S_k$ 对应于时段 $T_i$ 、指标 $P_j$ 的属性值为 $a_{ijk}$ (效益型属性值取正值,成本型属性值取负值), $i=1, \dots, n_i, j=1, \dots, n_j, k=1, \dots, n_k$ 。其中分别为 $n_i, n_j$ ,和 $n_k$ 为时段、指标和省份的数目,对应第 $k$ 个省份方案的矩阵就是 $(a_{ijk})_{n_i \times n_j \times n_k}$ 。

步骤2:数据矩阵的规范化。

为消除各指标的量纲效应,使建模具有通用性,可用下式对数据矩阵 $(a_{ijk})_{n_i \times n_j \times n_k}$ 进行规范化值处理<sup>[10]</sup>:

$$b_{ijk} = \frac{a_{ijk}}{\left[ \sum_{k=1}^{n_k} \sum_{i=1}^{n_i} a_{ijk}^2 / n_k \right]^{0.5}} \quad (1)$$

其中, $i=1, \dots, n_i, j=1, \dots, n_j, k=1, \dots, n_k$ 。

另外,文献[11]根据“效益型”、“成本型”、“固定型”和“区间型”的不同类型指标,分别给出了相应的规范化方法,应用时可参考。

步骤3:确定理想决策矩阵和负理想决策矩阵。

由规范化数据矩阵 $(a_{ijk})_{n_i \times n_j \times n_k}$ 可得相应的理想数据矩阵 $B^+$ 和理想数据矩阵 $B^-$ :

$$B^+ = (b_{ij}^+)_{n_i \times n_j}$$

$$B^- = (b_{ij}^-)_{n_i \times n_j}$$

其中

$$b_{ij}^+ = \max_k \{b_{ijk}\}$$

$$b_{ij}^- = \min_k \{b_{ijk}\}$$

步骤4:构造投影指标函数。

PP方法就是把三维数据

$$\{b_{ijk} | i=1, \dots, n_i, j=1, \dots, n_j, k=1, \dots, n_k\}$$

综合成以 $a=(w_1, w_2, \dots, w_n, a_1, a_2, \dots, a_n)$ 为投影方向的一维投影值 $z(k)$

$$z(k) = \frac{\left( \sum_{i=1}^{n_i} a_i \sum_{j=1}^{n_j} w_j (b_{ijk} - b_{ij}^-)^2 \right)^{0.5}}{\left( \sum_{i=1}^{n_i} a_i \sum_{j=1}^{n_j} w_j (b_{ijk} - b_{ij}^+)^2 \right)^{0.5} + \left( \sum_{i=1}^{n_i} a_i \sum_{j=1}^{n_j} w_j (b_{ijk} - b_{ij}^-)^2 \right)^{0.5}} \quad (2)$$

其中,  $w_j$  为第  $j$  个指标的权重,  $a_i$  为第  $i$  个时段的权重,  $z(k)$  为省份  $S_k$  距离理想数据方案  $B^+$  和负理想数据方案  $B^-$  的相对接近度。投影时希望投影值尽可能散开, 据此投影指标函数可构造为

$$Q(a) = S_z \tag{3}$$

式中,  $S_z$  为投影值  $z(k)$  的标准差

$$S_z = \sum_{k=1}^{n_k} [(z(k) - E_z)^2 / n_k]^{0.5} \tag{4}$$

式中,  $E_z$  为序列  $\{z(k) | k=1, \dots, n_k\}$  的均值。

步骤 5: 优化投影目标函数, 确定最佳投影方向。

通过求解投影指标函数最大化来估计最佳投影方向, 最佳投影方向可以最大可能地暴露高维数据特征结构的投影方向。即:

$$\begin{aligned} \max Q(a) = S_z, \text{ 满足} \\ w(j) > 0, \sum_{j=1}^p w_j = 1, a_i > 0, \sum_{i=1}^p a_i = 1 \end{aligned} \tag{5}$$

这是一个以  $a=(w_1, w_2, \dots, w_n, a_1, a_2, \dots, a_n)$  为优化变量的复杂非线性优化问题, 用常规优化方法处理较困难。模拟生物优胜劣汰规则与群体内部染色体信息交换机制的加速遗传算法(RAGA)是一种通用的全局优化方法, 通过多元替代和叠加运算, 加强了计算结果的稳定性和精确性。

步骤 6: 确定投影值, 综合评价分析。

把最佳投影方向  $a^*$  代入(2)式得各样本的投影值  $z^*(k)$ , 投影值是各评价指标的最佳投影方向与评价标准值的加权, 按照  $z^*(k)$  值从大到小的排列顺序就可得到相应的省份科技资源配置效率的优劣次序, 反映了各评价指标的综合评价结果。

### 4 科技资源配置效率评价指标体系的构建

关于科技投入与产出的内容指标, 常见于科技实力、科技创新能力和科技竞争力等衡量国家和区域的科技发展水平的概念之中。对此, 许多学者从投入产出及其相关因素关系的角度来对科技系统进行分析。他们都主要强调了科技竞争力的三大构成要素: 投入要素、产出要素、环境要素。投入主要是指人力和财力两个方面。本文采用泛珠江三角洲区域九省 2000~2004 年科技投入与产出相关指标。

表 1 泛珠江三角洲区域九省科技投入与产出指标

一级指标	二级指标	三级指标
科技投入水平	人力投入	R&D 人员总数 万人 R&D 科学家和工程师
	财力投入	R&D 经费支出总额 地方财政科技拨款
科技产出水平	专利及技术贸易	发明专利申请受理量 技术市场成交额
	科技论文	科技论文
	高技术产业	高技术产业增加值 高新技术产品出口额

这些指标均为效益型指标, 时间跨度为 5 年, 相关数据见表 2。

### 5 实证分析

在 Matlab 编程实现 RAGA, 输入数据, 得到最大投影值为 0.2907, 最佳投影方向为

$$a = (0.0176, 0.0659, 0.4556, 0.0026, 0.2167, 0.1170, 0.0498, 0.0265, 0.0483, 0.3560, 0.1876, 0.1050, 0.0161, 0.3353)$$

分别表示 9 个指标 R&D 人员, 万人 R&D 科学家和工程师, R&D 经费, 地方财政科技拨款, 高技术产业规模以上企业增加值, 高技术产品出口额, 专利申请受理量, 国内期刊论文数, 技术市场成交合同金额和 5 个时间段。

将  $a$  代入(2)中, 得到各省的最佳投影值

$$Z(k) = (0.0115, 0.0300, 0.0529, 0.0753, 0.0835, 0.1721, 0.1945, 0.2750, 1.0000)$$

最佳投影值越大, 说明该省的科技资源配置效率越高。

结果分析:

(1) 从各省的最佳投影值, 投影值分布在 0.0115~1 之间, 离散性强, 能够很好区分差别, 我们得到区域省份科技资源配置效率排名:

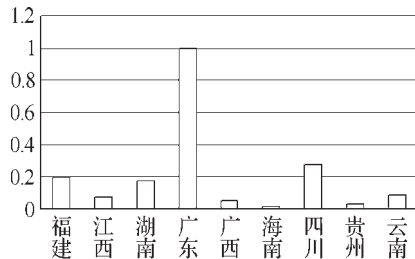


图 2 区域科技资源配置效率

对上述结果进行分类, 可以分为以下 4 组, 解释如下:

第一组: 广东。

广东为全国的经济强省, 全省 GDP、人均 GDP 均位居区域首位, 对科技的投入规模和投入强度较大, 广东 2005 年科技活动人员占区域总量的 34.68%, 科技经费支出额占 41.69%, R&D 经费占 47.28%。广东一省约占了三分之一以上的区域科技资源, 而且重视科技成果的转化和应用, 科技产出的 4 项指标(专利申请受理量、国内期刊论文数、技术市场成交合同金额、高技术产品增加值)均居区域首位。呈现高投入、高产出, 资源配置效率遥遥领先, 居区域首位。

第二组: 四川、福建、湖南。

四川、福建、湖南 3 省 2005 年科技活动人员占区域总量的 42.33%, 科技经费支出额占 41.86%, R&D 经费占 37.77%。四川是区域西部地区的科技大省, 投入规模位居区域第二, 产出方面除了高技术产业规模以上企业增加值和技术市场成交合同金额略比福建少, 专利申请受理量和国内中文期刊科技论文数要胜于福建, 属于中等偏上投

表2 泛珠江三角洲区域九省 2000-2004 年指标相关数据

省份 年份	R&D 人员 (万人)	万人 R&D 科工*	R&D 经费 (亿元)	地方财政科技 拨款(亿元)	高技术产业规模以上 企业增加值(亿元)	高技术产品出 口额(亿美元)	专利申请 受理量(件)	国内中文期刊科 技论文数(千篇)	技术市场成交 合同金额(亿元)
福建									
2000	2.20	4.898	21.20	10.84	127.82	33.24	4 211	642	17.26
2001	2.48	5.552	22.60	8.48	124.94	35.01	4 971	4 140	13.69
2002	2.24	5.136	24.40	8.90	180.32	60.99	6 522	4 511	12.90
2003	2.66	5.820	37.50	10.18	226.63	84.52	7 236	4 826	16.68
2004	3.18	6.836	45.89	11.60	289.60	115.03	7 498	6 246	14.14
江西									
2000	1.80	2.657	8.20	2.20	37.54	0.49	1 557	89	6.93
2001	1.51	2.293	7.80	3.25	37.24	1.24	1 778	2 092	6.27
2002	1.53	2.676	11.72	3.30	47.07	0.97	2 037	2 456	6.29
2003	1.70	2.844	16.98	3.63	41.56	1.93	2 434	2 730	8.33
2004	1.92	3.408	21.53	4.30	51.60	2.05	2 685	3 811	9.37
湖南									
2000	2.90	3.106	19.20	8.58	29.61	2.72	4 117	1 025	28.68
2001	2.87	3.366	24.00	9.33	31.46	2.28	4 292	7 550	29.39
2002	2.92	3.515	26.21	11.30	38.63	2.46	4 859	9 146	32.34
2003	2.70	3.272	30.09	9.14	41.94	1.94	6 054	10 362	36.93
2004	3.13	3.882	37.04	9.40	61.70	2.70	7 693	13 489	40.83
广东									
2000	7.10	6.711	107.10	39.42	677.77	353.35	21 123	1 472	48.21
2001	7.91	8.377	137.40	40.77	759.79	444.79	27 596	14 677	53.97
2002	8.69	9.161	156.45	57.70	998.68	647.81	34 352	17 175	68.45
2003	9.38	9.706	179.84	56.58	1 552.11	950.27	43 186	20 003	80.57
2004	9.31	9.501	211.21	65.40	1 880.00	1 303.92	52 201	28 296	57.27
广西									
2000	1.30	2.450	8.40	4.06	19.11	0.91	1 762	120	1.77
2001	0.95	1.754	8.00	4.64	22.33	0.88	1 838	2 121	3.78
2002	1.21	2.157	9.05	6.00	24.70	0.91	1 927	2 580	4.44
2003	1.32	2.203	11.24	8.87	28.53	1.19	2 250	3 218	4.18
2004	1.48	2.618	11.87	6.80	34.60	1.85	2 202	4 163	9.10
海南									
2000	0.10	1.144	0.80	0.62	6.03	1.02	502	14	0.00
2001	0.09	1.005	0.80	0.61	7.42	4.94	390	415	8.40
2002	0.08	0.996	1.22	0.70	9.02	5.60	546	524	0.91
2003	0.10	1.110	1.21	0.98	9.21	2.17	445	511	1.20
2004	0.14	1.589	2.09	1.00	10.50	7.88	375	589	0.19
四川									
2000	6.00	4.923	44.90	7.29	112.98	6.78	4 496	1 261	10.42
2001	4.82	4.086	57.50	8.59	111.43	12.49	5 039	7 716	12.63
2002	6.13	5.189	61.92	8.90	126.64	15.62	5 997	9 390	7.75
2003	5.79	5.057	79.42	10.07	133.85	18.87	7 443	11 127	12.87
2004	6.01	5.307	78.01	10.80	143.50	17.45	7 260	14 798	16.56
贵州									
2000	0.81	1.674	4.20	3.09	26.40	0.18	986	72	0.06
2001	0.95	1.658	5.30	3.56	28.49	28.49	950	1 687	0.06
2002	0.90	1.590	6.07	3.90	33.40	0.99	1 260	1 624	1.35
2003	0.86	1.680	7.89	4.31	39.96	2.38	1 242	1 911	1.79
2004	0.78	1.588	8.68	5.10	43.70	3.76	1 486	2 614	1.35
云南									
2000	1.10	2.052	6.80	6.65	13.31	1.09	1 710	322	18.77
2001	1.17	2.263	7.70	7.84	15.56	2.79	1 793	2 686	25.53
2002	1.39	2.539	9.79	9.00	21.30	2.55	1 780	3 114	17.95
2003	1.29	2.217	11.01	9.17	20.24	1.65	1 966	3 408	22.87
2004	1.47	3.796	12.51	8.40	22.20	1.95	2 132	4 038	21.56

\*注：“科工”为科学家和工程师。



表3 区域省份科技资源配置放弃排名

省份	福建	江西	湖南	广东	广西	海南	四川	贵州	云南
投影值	0.1945	0.0753	0.1721	1.0000	0.0529	0.0115	0.2750	0.0300	0.0835
排名	3	6	4	1	7	9	2	8	5

入,中等偏上产出。这说明四川在科技成果产业化方面显得较广东为薄弱,未能充分发挥科技对经济的促进作用。福建为区域东部沿海经济较发达省份,地缘优势突出,拥有众多的港口,交通网络四通八达,与国际经济联系密切,有利于吸引外资和促进进出口贸易的发展,人均GDP位居区域第二,仅次于广东,投入和产出各方面相对均衡。虽然投入、产出总量比四川要少,但是人均投入和产出却胜于四川,属于中等投入和中等偏上产出。湖南为中等投入,中等产出。

第3组:云南、江西、广西。

云南、江西、广西3省2005年R&D人员占区域总量的18.67%,R&D经费占12.50%。云南、江西、广西属于投入较低,而产出相对较高。该组总体上科技配置效率中等。

第4组:贵州、海南。

贵州、海南2005年R&D人员只占区域总量的4.33%,R&D经费占2.45%,属于低科技投入,但更低科技产出,总体科技配置效率低。

(2)从最佳投影方向 $a$ ,我们可以得到各个指标和年份在配置效率中的权重。

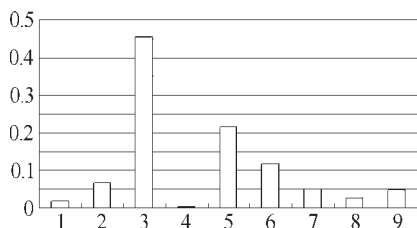


图3 科技资源效率评价指标

在科技投入指标中,R&D经费(亿元)(0.4556)对配置效率的影响程度远大于科技人力投入R&D人员(万人),万人R&D科工(0.0176,0.0659)。我们认为区域科技人力投入相对于科技财力投入相对过剩,或者说人均科研经费偏低。可见区域科技财力投入与区域科技人力投入不甚协调,结构不合理,科技财力投入相对于人力投入显得更为短缺。这也可从区域科技资源配置效率排名分组中验证。4组当中,除了第一组广东,科技活动人员占区域总量的比例小于科技经费支出额和R&D经费占区域总量的比例,其余3组科技活动人员占区域总量的比例均大于科技经费支出额和R&D经费占区域总量的比例。

在科技产出指标中,高技术产业规模以上企业增加值和高技术产品出口额(亿美元)对配置效率的影响程度最大,分别是0.2167和0.1170,其次为专利申请受理量和技术市场成交合同金额。这说明科技成果产业化指标高的省份科技资源配置效率高,与区域省份科技资源配置效率排名结果一致,区域应该进一步加快科技成果产业化进程。

在5个年份中,2000年影响最大,最小为2003年,其余年份相差不大。说明综合排序决策时主要看2000年各省份的R&D经费投入、万人R&D科学家和工程师、高技术产业规模以上企业增加值、高技术产品出口额,说明在5个年份中2000年这4个指标含有的决策信息量较大。

## 6 结论

通过对泛珠江三角洲区域科技资源配置效率的定性和定量分析,得到几点结论:

(1)经济发达的科技大省排名靠前,如广东。科技资源投入偏高的省份,科技效率偏高,科技投入少的省份效率比较低。

(2)区域科技财力投入与区域科技人力投入不甚协调,结构不合理,科技财力投入相对于人力投入显得更为短缺。

(3)R&D经费投入、万人R&D科学家和工程师、高技术产业规模以上企业增加值、高技术产品出口额这4个指标对配置效率影响比较大。在科技资源有限的情况下,提高科技资源配置水平和科技实力,重要的途径就是合理、高效配置科技资源,实现科技链与产业链有机衔接,重视和加快科技成果产业化进程,实现科技与经济的有机结合,从而推动区域经济发展。

### 参考文献:

- [1] 周寄中.科技资源论[M].西安:陕西人民出版社,1999:43-115.
- [2] 雷睿勇,罗敏,邹吉鸿.对我国科技资源配置效率评价方法的述评[J].山地农业生物学报,2004,23(5):448-453.
- [3] 徐建国.我国区域科技资源配置能力分析[J].中国软科学,2002(9):98-99.
- [4] 李石柱,李钢.我国地区科技资源配置现状分析[J].北京理工大学学报(社会科学版),2002(4):107-109.
- [5] 鲁勇兵,梁婉君,李双成.河北省区域科技资源配置效率的因子分析[J].科技进步与对策,2006(7):73-75.
- [6] 金菊良,汪淑娟,魏一鸣.动态多指标决策问题的投影跟踪模型[J].中国管理科学,2004(1):64-67.
- [7] 王顺久,张欣莉,丁晶,等.投影跟踪聚类模型及其应用[J].长江科学院院报,2002,19(6):53-55,61.
- [8] 王顺久,张欣莉,侯玉,等.洪水灾情投影跟踪评估模型[J].水文,2002,22(4):1-4.
- [9] 麻素挺,汤洁,林年丰.基于GIS与RS多空间信息的吉林西部生态环境综合评价[J].资源科学,2004,26(4):140-145.
- [10] 樊治平,肖四汉.有时序多指标决策的理想矩阵法[J].系统工程,1996,11(1):64-65.
- [11] 樊治平,肖四汉.一类动态多指标决策问题的关联分析法[J].系统工程,1995,13(1):23-27.

(责任编辑:陈晓峰)