

# 基于灰色定权聚类的我国自然科学基金资助地区的聚类分析<sup>①</sup>

游中胜<sup>1</sup>, 曾波<sup>2</sup>

1. 重庆师范大学 编辑出版中心, 重庆 400047; 2. 重庆工商大学 商务策划学院, 重庆 400067

**摘要:** 由于经济发展水平、高校科研实力及地方科技发展战略的差异, 不同地区的科技竞争力存在较大差异. 以 2009 年度国家自然科学基金项目(青年基金、面上项目)的资助情况为基础, 对我国 30 个省市自治区的科技竞争力进行灰色定权聚类分析, 对其科研发展水平进行分类及评价, 并对原因进行分析, 为我国科技发展战略的制定提供参考依据.

**关键词:** 科技竞争力; 灰色定权聚类; 国家自然科学基金

**中图分类号:** C94

**文献标志码:** A

灰色定权聚类的基本概念如下:

**定义 1**<sup>[1]</sup> 设  $x_{ij}$  ( $i=1, 2, \dots, n; j=1, 2, \dots, m$ ) 为对象  $i$  关于指标  $j$  的观测值,  $f_j^k(\cdot)$  ( $j=1, 2, \dots, m; k=1, 2, \dots, s$ ) 为  $j$  指标  $k$  子类白化权函数. 若  $j$  指标  $k$  子类的权  $\eta_j^k$  ( $j=1, 2, \dots, m; k=1, 2, \dots, s$ ) 与  $k$  无关, 即对任意的  $k_1, k_2 \in \{1, 2, \dots, s\}$ , 总有  $\eta_j^{k_1} = \eta_j^{k_2}$ , 此时可将  $\eta_j^k$  的上标  $k$  略去, 记为  $\eta_j$  ( $j=1, 2, \dots, m$ ), 并称

$$\sigma_i^k = \sum_{j=1}^m f_j^k(x_{ij}) \eta_j$$

为对象  $i$  属于  $k$  灰类的灰色定权聚类系数.

**定义 2**<sup>[1]</sup> 设  $x_{ij}$  ( $i=1, 2, \dots, n; j=1, 2, \dots, m$ ) 为对象  $i$  关于指标  $j$  的观测值,  $f_j^k(\cdot)$  ( $j=1, 2, \dots, m; k=1, 2, \dots, s$ ) 为  $j$  指标  $k$  子类白化权函数. 若对任意的  $j=1, 2, \dots, m$ . 总有  $\eta_j = \frac{1}{m}$ , 则称

$$\sigma_i^k = \sum_{j=1}^m f_j^k(x_{ij}) \cdot \eta_j = \frac{1}{m} \sum_{j=1}^m f_j^k(x_{ij})$$

为对象  $i$  属于  $k$  灰类的灰色等权聚类系数.

**定义 3**<sup>[1]</sup>

- 1) 根据灰色定权聚类系数的值对聚类对象进行归类, 称为灰色定权聚类;
- 2) 根据灰色等权聚类系数的值对聚类对象进行归类, 称为灰色等权聚类.

## 1 灰色定权聚类的步骤

- 1) 给出  $j$  指标  $k$  子类白化权函数  $f_j^k(\cdot)$  ( $j=1, 2, \dots, m; k=1, 2, \dots, s$ ).

① 收稿日期: 2011-03-10

基金项目: 教育部人文社会科学研究一般项目(11YJC630129).

作者简介: 游中胜(1970-), 男, 副教授, 硕士, 主要从事计算机应用、随机经济系统及编辑学的研究.

2) 确定各指标的聚类权  $\eta_j (j=1,2,\dots,m)$ .

3) 从 1) 和 2) 得出的白化权函数  $f_j^k(\cdot) (j=1,2,\dots,m; k=1,2,\dots,s)$ , 聚类权  $\eta_j (j=1,2,\dots,m)$  以及对象  $i$  关于  $j$  指标的观测值  $x_{ij} (i=1,2,\dots,n; j=1,2,\dots,m)$ , 计算出灰色定权聚类系数  $\sigma_i^k = \sum_{j=1}^m f_j^k(x_{ij}) \cdot \eta_j (i=1,2,\dots,n; k=1,2,\dots,s)$ .

4) 若  $\max_{1 \leq k \leq s} \{\sigma_i^k\} = \sigma_i^{k^*}$ , 则断定对象  $i$  属于灰类  $k^*$ .

## 2 我国自然科学基金资助地区的灰色定权聚类

2009 年度我国国家自然科学基金项目的资助情况, 如表 1 所示<sup>[2]</sup>. 现按照科技竞争实力强(I 类地区)、科技竞争实力较强(II 类地区)、科技竞争实力一般(III 类地区)、科技竞争实力较弱(IV 类地区), 对我国 30 个省市自治区的科研实力进行分类.

表 1 2009 年国家自然科学基金项目资助情况(按地区统计)

序号	地区	青年项目		面上项目	
		资助项数 / 项	资助金额 / 万元	资助项数 / 项	资助金额 / 万元
1	北京	1 235	24 431.4	2 297	78 238
2	天津	163	3 261.2	294	9 589.7
3	河北	42	819	120	3 839.2
4	湖北	413	8 119.6	683	22 322.9
5	山西	37	752	90	2 863
6	内蒙古	9	166	7	262
7	辽宁	214	4 165.7	410	13 003.5
8	吉林	127	2 540	257	8 562
9	黑龙江	184	3 664.5	242	7 705
10	上海	676	13 345.6	975	31 371.5
11	江苏	569	11 159	827	27 396.2
12	浙江	330	6 452	475	14 900
13	安徽	157	3 153.2	258	8 961
14	福建	91	1 783.7	211	6 886.5
15	江西	40	785.5	29	971.7
16	山东	252	4 959.5	416	13 769.6
17	河南	88	1 699	152	4 948.1
18	湖南	197	3 860.2	330	10 452.2
19	湖北	413	8 119.6	683	22 322.9
20	广东	370	7 314.4	681	21 659.8
21	广西	13	268	23	696
22	海南	2	41	1	26
23	重庆	160	3 206.5	265	8 185
24	贵州	17	337	28	1 229
25	四川	213	4 268.3	331	10 761.8
26	云南	51	1 023	59	1 903
27	陕西	307	6 191.7	459	14 940.3
28	甘肃	83	1 753	108	3 960
29	青海	5	103	11	347
30	宁夏	3	53	1	8

1) 将指标和灰类编号, 将  $j$  指标  $k$  子类白化权函数  $f_j^k(\cdot) (j=1,2,3,4; k=1,2,3,4)$  分别取为

$$\begin{matrix}
 f_1^1[500, 1000, -, -] & f_1^2[100, 250, -, 500] & f_1^3[20, 50, -, 100] & f_1^4[-, -, 20, 80] \\
 f_2^1[10000, 20000, -, -] & f_2^2[5000, 7500, -, 10000] & f_2^3[500, -, 2000, 5000] & f_2^4[-, -, 40, 500] \\
 f_3^1[600, 1200, -, -] & f_3^2[100, 250, -, 600] & f_3^3[20, 50, -, 100] & f_3^4[-, -, 20, 80] \\
 f_4^1[30000, 60000, -, -] & f_4^2[6000, 8000, -, 30000] & f_4^3[500, -, 2000, 5000] & f_4^4[-, -, 40, 500]
 \end{matrix}$$

2) 取青年项目资助项数、青年项目资助金额、面上项目资助项数、面上项目资助金额的权重分别为： $\eta_1 = 0.25, \eta_2 = 0.15, \eta_3 = 0.35, \eta_4 = 0.25$ 。

3) 将表 1、白化权函数信息、指标权重信息按照一定的格式分别输入 Excel 文件，通过灰色建模软件<sup>[1]</sup>进行灰色定权聚类的计算，可得到如下图 1 所示的界面。



图 1 基于灰色建模软件的我国自然科学基金资助地区的灰色定权聚类

### 3 定权聚类结果分析

从图 1 可以得出表 2 所示的聚类结果。

表 2 我国自然科学基金资助地区的灰色定权聚类结果

类别	省份
科技竞争实力强(I类地区)	北京, 上海, 江苏
科技竞争实力较强(II类地区)	天津, 湖北, 辽宁, 黑龙江, 浙江, 安徽, 山东, 湖南, 湖北, 广东, 重庆, 四川, 陕西
科技竞争实力一般(III类地区)	吉林, 福建, 河南, 甘肃
科技竞争实力较弱(IV类地区)	河北, 山西, 内蒙古, 江西, 广西, 海南, 贵州, 云南, 青海, 宁夏

从表 2 不难看出, 北京、上海和江苏是我国科技竞争实力最强的三个区域, 这些区域集中了全国较大比例的 985 及 211 高校, 分别处于我国的长三角及环渤海经济圈, 经济发达, 初步形成了科技实力与经济

发展良性互动的局面;而以天津为代表的科技竞争实力较强的13个地区,这些地区的经济发展、高校数量及实力总体上与I类地区存在一定的差异,但在全国尚处于中上水平;对于III类地区,科技实力对经济发展的支撑作用有限;对于IV类地区,其科技竞争实力相对较弱。

## 4 结束语

本文以2009年度国家自然科学基金项目的资助情况为统计数据,对我国30个省市自治区的科技竞争力进行灰色定权聚类分析,对其科研发展水平进行分类及评价,并对原因进行了分析。评价一个地区的科技竞争力,会受到很多因素的影响。本文对国家自然科学基金项目的资助情况进行了分析,能在一定程度上客观反应该地区的科技竞争力,可以为我国科技发展战略的制定提供参考依据。

### 参考文献:

- [1] 刘思峰,党耀国,方志耕. 灰色系统理论及应用 [M]. 北京:科学出版社,2007:104-107.
- [2] 国家自然科学基金管理委员会. 国家自然科学基金资助项目统计资料 [EB/OL]. [2010-09-10](2010-05-05). [http://www.nsf.gov.cn/nsfc/cen/xmtj/psd/2008\\_table.pdf](http://www.nsf.gov.cn/nsfc/cen/xmtj/psd/2008_table.pdf).
- [3] 徐小峰,赵金楼,刘丙泉. 基于改进灰色定权的区域造船能力比较研究 [J]. 统计与信息论坛,2008,23(12):5-8.
- [4] 杨皎平,赵宏霞. 灰色聚类的软件研发项目团队有效性评价 [J]. 计算机工程与应用,2010,46(9):231-233.
- [5] 董里. 基于灰色聚类评价的省区增设研究 [J]. 西南交通大学学报:社会科学版,2009,10(2):125-128.
- [6] 张恩祥,郁广健. 基于旅游流空间下的信息流灰色系统综合评价方法 [J]. 统计与信息论坛,2006,21(4):57-59.

# Clustering Analysis of Assistant Regions of Natural Science Foundation of China Based on Grey Fixed Weight Clustering

YOU Zhong-sheng<sup>1</sup>, ZENG Bo<sup>2</sup>

1. Editing and Publishing Center, Chongqing Normal University, Chongqing 400047, China;

2. Strategic Planning College, Chongqing Technology and Business University, Chongqing 400067, China

**Abstract:** Considerable differences exist among different regions in their technological competitiveness because of their different economic development level, universities' research strength and development strategies. Based on the data of the research projects financially supported by the National Natural Science Foundation of China in 2009, this paper analyses the technological competitiveness of 30 provinces and autonomous regions of the country with the method of grey fixed weight clustering, sorts and evaluates the level of their research and development, and explores the underlying causes for the differences, so as to provide reference for the formulation of S & T development strategy.

**Key words:** technological competitiveness; grey fixed weight clustering; National Natural Science Foundation of China

责任编辑 张 枸