

大学生综合素质的灰色模糊聚类评价模型

金玲玲¹, 汪文俊¹, 王喜凤²

(1. 海南师范大学 数学与统计学院, 海南 海口 571158;

2. 安徽工业大学 计算机学院, 安徽 马鞍山 243002)

摘要: 学生综合素质评价是高校学生管理人员承担的一项重要而复杂的工作。针对高校教育中大学生综合素质评价相对薄弱的问题, 建立大学生综合素质评价指标体系, 依据灰色系统理论, 确定不同灰数, 通过白化函数生成灰色聚类矩阵获取对象所属灰类的信息, 最后结合模糊聚类方法, 以聚类行向量作为原始数据构造了模糊相似矩阵, 通过聚类分析得到了对象间的关联信息。数据分析表明, 评价结果客观, 合理地反映出学生综合素质的优劣, 方法切实可行。

关键词: 灰色聚类; 模糊; 综合素质; 评价

中图分类号: G515.5

文献标识码: A

文章编号: 1673-629X(2012)05-0109-04

Gray Fuzzy Cluster Model of Evaluating College Students' Comprehensive Quality

JIN Ling-ling¹, WANG Wen-jun¹, WANG Xi-feng²

(1. School of Mathematics and Statistics, Hainan Normal University, Haikou 571158, China;

2. School of Computer, Anhui University of Technology, Ma' anshan 243002, China)

Abstract: Evaluation on college students' comprehensive quality is an important and difficult work that administrative personnel of universities is undertaken, and the quality evaluation is relatively weak in higher education. Based on the problem, evaluation index system is set up. According to gray system theory, different gray digit is given out. The gray information to which the object belongs can get from gray cluster matrix degenerated from whitened function. Finally combined with fuzzy cluster method, construct fuzzy similarity matrix using data of clustering row vectors and get the relevance information among objects. Data analysis shows that assessment result is objective and can reasonably indicate superiority or inferiority of students' comprehensive quality, this method is feasible and effective.

Key words: gray cluster; fuzzy; comprehensive quality; evaluation

0 引言

大学生综合素质是在学生自身的基础上, 经过高等教育和社会环境的影响, 形成的对大学生的发展和提高起决定作用的、基础的、重要的和本质的综合素养和品质^[1,2]。高校学生综合素质评价在教育评价中处于核心地位, 它对大学生就业、高校教学改革和和谐社会发展都具有重要的意义。目前国内对学生综合素质测评方法的研究: 如神经网络评价法^[3]、模糊综合评价法^[4]和 AHP 法^[5]等, 对评价理论和实践操作都具有一定的指导意义。

灰色系统研究的是部分信息明确、部分信息未知的小样本、贫信息、不确定性的系统, 通过对已知信息

的生成去了解认识现实世界^[6,7]。大学生综合素质评价系统就是一个灰色系统。利用灰色系统理论去评价可扩大信息源, 提高评价结果的可信度。文中结合灰色系统和模糊聚类方法, 不仅通过灰色聚类显化各对象间所属灰类的信息, 而且由模糊聚类得出灰色聚类对象间的关系信息, 反映了各样本之间的接近程度。利用这种方法对学生进行聚类分析, 聚类结果具有可比性, 定性和定量分析的结合提高了评判的科学性和合理性。

1 大学生综合素质评价与指标体系

随着素质教育在我国高校的逐步推进, 制定科学、公正、可操作性强的综合素质评价体系, 不仅对大学生的全面素质培养起到导航作用, 也使高校的素质教育工作更具针对性、主动性和实效性。目前国内大部分高校对大学生综合素质评价, 定性分析研究较多, 在定量分析方面较少且过粗, 评价指标体系过于陈旧, 强调

收稿日期: 2011-10-08; 修回日期: 2012-01-15

基金项目: 海南省自然科学基金(610221); 海南师范大学青年教师科研启动资助项目(QN0918)

作者简介: 金玲玲(1976-), 女, 湖南常德人, 讲师, 硕士, 研究方向为粗糙集与数据挖掘。

知识尺度而忽视道德能力的标准越来越与素质教育价值观和人才培养理念不相适应。

文中在参照大量大学生综合素质指标体系的文献和专家建议的基础上,本着全面科学的评价原则建立了大学生综合素质评价指标体系结构。本指标体系共分三级,其中第一级 A 下设 4 个二级指标和 11 个三级指标,具体是:思想素质 B_1 (政治素质 C_1 、道德素质 C_2 和法纪素质 C_3)、文化素质 B_2 (知识素养 C_4 和文化艺术素养 C_5)、能力素质 B_3 (创新能力 C_6 、实践能力 C_7 、组织能力 C_8 和职业技能 C_9)、身心素质 B_4 (身体素质 C_{10} 和心理素质 C_{11})。

2 聚类评价方法

2.1 灰色聚类分析

聚类分析是采用数学定量手段确定聚类对象间的亲属关系并进行分型化类的一种多元分析方法。灰色聚类是根据灰色关联矩阵或灰数的白化权函数将一些观测指标或观测对象聚集成若干个可定义类别的方法^[8,9]。

步骤 1: 灰色聚类样本的构成。

设聚类对象 $X = \{X_1, X_2, \dots, X_n\}$ 有 n 个, 每个对象有 m 个聚类指标 $Y = \{B_1, B_2, \dots, B_m\}$, 每个指标按评估标准分为 s 个等级, 那么第 i 个对象 ($i = 1 \dots n$) 对第 j 个聚类指标 ($j = 1 \dots m$) 的样本值记为 X_{ij} (即聚类白化数), 将第 i 个对象归入第 k ($k \in \{1, 2, \dots, s\}$) 个灰类中, 称为灰色聚类。

由白化数可以组成矩阵 $M = (X_{ij})_{n \times m}$, 形式如下:

$$M = \begin{pmatrix} X_{11} & X_{12} & X_{13} & \dots & X_{1m} \\ X_{21} & X_{22} & X_{23} & \dots & X_{2m} \\ \dots & \dots & \dots & \dots & \dots \\ X_{n1} & X_{n2} & X_{n3} & \dots & X_{nm} \end{pmatrix}$$

通过问卷调查、专家评定等形式, 采用 1~10 分制打分, 分值越低, 表明其素质就越低。将专家组对学生对象的打分按各类指标取平均值, 得出表 1 的数据, 同时确定素质评价采用 4 个评价灰类, 如表 2 所示。

表 1 聚类白化数 X_{ij}

聚类对象(i)	聚类指标(j)			
	B_1	B_2	B_3	B_4
X_1	9.4	9.2	9.0	7.6
X_2	8.5	9.5	8.5	9.4
X_3	9.6	8.7	7.0	7.8
X_4	7.6	6.8	8.8	9.6
X_5	8.1	7.5	6.5	5.8
X_6	9.8	9.0	7.2	8.0
X_7	6.4	5.5	6.9	6.4
X_8	8.7	5.8	5.2	8.8
X_9	7.2	7.8	7.3	8.2
X_{10}	6.9	7.5	5.6	9.1

表 2 聚类分级标准

聚类指标(j)	灰类(k)			
	优秀	良好	及格	不及格
1	≥ 9.5	8.5 左右	7.5 左右	< 7.0
2	≥ 9.0	8.0 左右	6.5 左右	< 6.0
3	≥ 8.8	7.5 左右	6.5 左右	< 6.0
4	≥ 9.0	8.0 左右	7.0 左右	< 6.0

步骤 2: 确定灰类白化函数。

白化函数是灰色聚类的基础, 它反应了聚类指标与灰色聚类的亲疏关系。将 n 个对象关于指标 j 的取值相应地分为 s 个灰类, 称为 j 指标子类, 其中 j 指标 k 子类的白化函数记为 $f_j^k(\cdot)$ 。一个典型的白化权函数为起点、终点确定的左升、右降连续函数, 如图 1(a) 所示。

$$f(x) = \begin{cases} L(x), & x \in [a_1, b_1) \\ 1, & x \in [b_1, b_2] \\ R(x), & x \in (b_2, a_2] \end{cases}$$

其中 $L(x)$ 为左增函数, $R(x)$ 为右降函数, a_1 为始点, a_2 为终点, b_1, b_2 为转折点。

根据 a_1, a_2, b_1, b_2 取值的不同, 白化权函数还有三种常见形式, 分别为下限测度白化权函数、适中测度白化权函数和上限测度白化权函数, 记为

$$f_j^k(-, -, b_2, a_2), f_j^k(a_1, b_1, -, a_2), f_j^k(a_1, b_1, -, -),$$

如图 1(b)~(d) 所示。

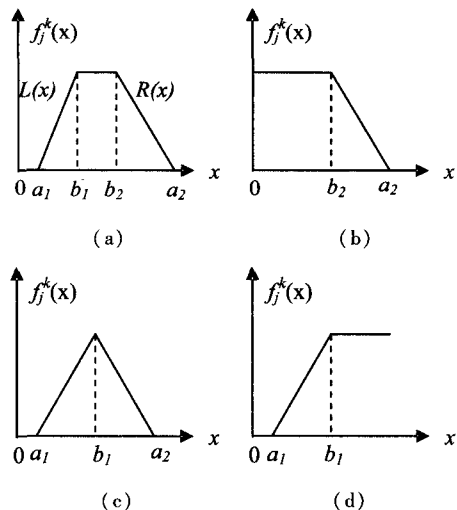


图 1 白权化函数图像

(1) 对于图 1(a) 所示的白化权函数定义为:

$$f_j^k(x) = \begin{cases} 0, & x \notin [a_1, a_2] \\ \frac{x-a_1}{b_1-a_1}, & x \in [a_1, b_1] \\ 1, & x \in [b_1, b_2] \\ \frac{a_2-x}{a_2-b_2}, & x \in [b_2, a_2] \end{cases}$$

(2) 对于图 1(b) 所示的白化权函数定义为:

$$f_j^k(x) = \begin{cases} 0, & x \notin [0, a_2] \\ 1, & x \in [0, b_2] \\ \frac{a_2 - x}{a_2 - b_2}, & x \in [b_2, a_2] \end{cases}$$

(3) 对于图 1(c) 所示的白化权函数定义为:

$$f_j^k(x) = \begin{cases} 0, & x \notin [a_1, a_2] \\ \frac{x - a_1}{b_1 - a_1}, & x \in [a_1, b_1] \\ \frac{a_2 - x}{a_2 - b_1}, & x \in [b_1, a_2] \end{cases}$$

(4) 对于图 1(d) 所示的白化权函数定义为:

$$f_j^k(x) = \begin{cases} 0, & x < a_1 \\ \frac{x - a_1}{b_1 - a_1}, & x \in [a_1, b_1] \\ 1, & x \geq b_1 \end{cases}$$

根据表 2, 思想素质、文化素质、能力素质和身心素质各指标关于各灰类的白化权函数分别设为:

$$\begin{aligned} f_1^1[8.5, 9.5, -, -], & f_1^2[7.5, 8.5, -, 9.5], \\ f_1^3[7.0, 7.5, -, 8.5], & f_1^4[-, -, 7.0, 7.5], \\ f_2^1[8.5, 9.0, -, -], & f_2^2[6.5, 8.0, -, 9.0], \\ f_2^3[6.0, 6.5, -, 8.0], & f_2^4[-, -, 6.0, 6.5], \\ f_3^1[8.0, 8.8, -, -], & f_3^2[6.5, 7.5, -, 8.8], \\ f_3^3[6.0, 6.5, -, 7.5], & f_3^4[-, -, 6.0, 6.5], \\ f_4^1[8.5, 9.0, -, -], & f_4^2[7.0, 8.0, -, 9.0], \\ f_4^3[6.0, 7.0, -, 8.0], & f_4^4[-, -, 6.0, 7.0] \end{aligned}$$

结合(1)~(4)所定义的公式,可计算不同对象的 $f_j^k(\cdot)$ 值。表3中仅以对象 X_1 为例,列出计算结果。

表3 X_1 对象分指标白化权函数取值

对象	X_1			
指标	f_j^1	f_j^2	f_j^3	f_j^4
B1	0.9	0.1	0	0
B2	1.0	0	0	0
B3	1.0	0	0	0
B4	0	0.6	0.4	0

步骤3:求聚类权。

在白化权函数中,令 λ_j^k 为 j 指标 k 子类的临界值,那么图1(a)中 $\lambda_j^k = (b_1 + b_2)/2$;图1(b)中 $\lambda_j^k = b_2$;图1(c)和(d)中 $\lambda_j^k = b_1$ 。

聚类权 w_j^k 表示各指标对某一灰类的权重:

$$w_j^k = \frac{\lambda_j^k}{\sum_{j=1}^m \lambda_j^k}$$

临界值及聚类权值见表4。

步骤4:求聚类系数。

聚类系数反映了第 i 个聚类对象隶属于第 k 个灰类的程度。表示为:

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

表4 临界值及聚类权值

指标		B ₁	B ₂	B ₃	B ₄
临界值 λ_j^k	λ_j^1	9.5	9.0	8.8	9.0
	λ_j^2	8.5	8.0	7.5	8.0
	λ_j^3	7.5	6.5	6.5	7.0
	λ_j^4	7.0	6.0	6.0	7.0
权 w_j^k	w_j^1	0.2617	0.2479	0.2424	0.2479
	w_j^2	0.2656	0.25	0.2344	0.25
	w_j^3	0.2727	0.2364	0.2364	0.2545
	w_j^4	0.2692	0.2307	0.2307	0.2692

根据公式,当 $i=1$ 时, $\sigma_1^1 = \sum_{j=1}^4 f_j^1(x_{1j}) \cdot w_j^1 = f_1^1(9.4) \times 0.2617 + f_2^1(9.2) \times 0.2479 + f_3^1(9.0) \times 0.2424 + f_4^1(7.6) \times 0.2479 = 0.7258$

同理可得: $\sigma_1^2 = 0.1766, \sigma_1^3 = 0.102, \sigma_1^4 = 0$

步骤5:构造聚类向量,进行聚类。

由 k 个聚类样本构成的聚类系数构成聚类行向量 $\sigma_i = (\sigma_i^1, \sigma_i^2, \dots, \sigma_i^s) (i=1, 2, \dots, n)$,在行向量中聚类系数最大值所对应的灰类为该样本所属类别,即

$\sigma_i^k = \max\{\sigma_i^k\}, (1 \leq k \leq s)$,则对象 i 属于灰类 k' 。

当 $i=1$ 时,向量 $\sigma_1 = (0.7258, 0.1766, 0.102, 0)$,其中 $\max\{\sigma_1^k\} = \sigma_1^1 = 0.7258$,表明该对象属于第一灰类,即学生综合素质评价结果为优秀。同样可得其它对象的聚类结果,如表5所示。

表5 聚类系数及分析结果

对象	系数	σ_i^1	σ_i^2	σ_i^3	σ_i^4	聚类结果
X_1		0.7258	0.1766	0.102	0	优秀
X_2		0.6473	0.3197	0	0	优秀
X_3		0.3609	0.3922	0.1692	0	良好
X_4		0.4903	0.0766	0.4345	0	优秀
X_5		0	0.3260	0.4242	0.2692	及格
X_6		0.5096	0.4141	0.0709	0	优秀
X_7		0	0.0938	0.2436	0.6614	不及格
X_8		0.2010	0.2624	0	0.4614	不及格
X_9		0	0.6042	0.1879	0.1615	良好
X_{10}		0.2479	0.1667	0.0788	0.4999	不及格

2.2 模糊聚类分析

上述的聚类分析,显化了各个聚类对象所属灰类的信息,但各个聚类对象间的相互关系信息并未显化,因此还有必要对分组信息进一步挖掘。模糊聚类分析^[10-12]利用模糊的方法来处理聚类问题,最终得到样本属于各个类别的可能性程度,进而通过确定一个适当的标准 λ 进行具体分类。

步骤1 数据矩阵及规范化处理。

由灰色聚类办法得到的若干聚类行向量组成原始数据矩阵 σ ,其中 σ 的行指标为聚类对象,列指标为灰类指标,矩阵形式如下:

$$\sigma = \begin{pmatrix} \sigma_1^1 & \sigma_1^2 & \sigma_1^3 & \sigma_1^4 \\ \sigma_2^1 & \sigma_2^2 & \sigma_2^3 & \sigma_2^4 \\ \dots & \dots & \dots & \dots \\ \sigma_{10}^1 & \sigma_{10}^2 & \sigma_{10}^3 & \sigma_{10}^4 \end{pmatrix}$$

对矩阵按行进行归一化处理后得:

$$\sigma = \begin{pmatrix} 0.723 & 0.176 & 0.102 & 0 \\ 0.669 & 0.331 & 0 & 0 \\ 0.391 & 0.425 & 0.184 & 0 \\ 0.490 & 0.077 & 0.433 & 0 \\ 0 & 0.320 & 0.416 & 0.264 \\ 0.552 & 0.448 & 0.077 & 0.0 \\ 0 & 0.094 & 0.244 & 0.662 \\ 0.217 & 0.284 & 0 & 0.499 \\ 0 & 0.633 & 0.197 & 0.169 \\ 0.25 & 0.168 & 0.079 & 0.503 \end{pmatrix}$$

步骤2 建立模糊相似矩阵。

采用最大最小法确定相似系数 $r_{ij} = R(\sigma_i, \sigma_j)$, 计算公式如下:

$$r_{ij} = \frac{\sum_{k=1}^m (\sigma_i^k \wedge \sigma_j^k)}{\sum_{k=1}^m (\sigma_i^k \vee \sigma_j^k)}$$

建立模糊相似矩阵 $R = (r_{ij})_{10 \times 10}$:

$$R = \begin{pmatrix} 1.0 & 0.7 & 0.5 & 0.5 & 0.2 & 0.6 & 0.1 & 0.2 & 0.2 & 0.3 \\ 0.7 & 1.0 & 0.8 & 0.4 & 0.2 & 0.7 & 0.1 & 0.3 & 0.2 & 0.3 \\ 0.5 & 0.8 & 1.0 & 0.5 & 0.3 & 0.8 & 0.2 & 0.4 & 0.4 & 0.3 \\ 0.5 & 0.4 & 0.5 & 1.0 & 0.3 & 0.5 & 0.2 & 0.2 & 0.2 & 0.3 \\ 0.2 & 0.2 & 0.3 & 0.3 & 1.0 & 0.2 & 0.4 & 0.4 & 0.5 & 0.3 \\ 0.6 & 0.7 & 0.8 & 0.5 & 0.2 & 1.0 & 0.1 & 0.3 & 0.3 & 0.3 \\ 0.1 & 0.1 & 0.2 & 0.2 & 0.4 & 0.1 & 1.0 & 0.4 & 0.3 & 0.5 \\ 0.2 & 0.3 & 0.4 & 0.2 & 0.4 & 0.3 & 0.4 & 1.0 & 0.3 & 0.8 \\ 0.2 & 0.2 & 0.4 & 0.2 & 0.5 & 0.3 & 0.3 & 0.3 & 1.0 & 0.3 \\ 0.3 & 0.3 & 0.3 & 0.3 & 0.3 & 0.3 & 0.5 & 0.8 & 0.3 & 1.0 \end{pmatrix}$$

步骤3 聚类。

按照直接聚类法,取不同 λ 求相似类,并逐层进行相似类的合并,最后获得 λ 不同水平上的等价分类,其动态聚类图如图2所示。

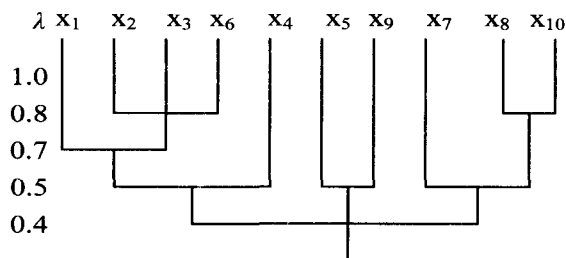


图2 详细动态聚类图

3 结果讨论

依据表5,从行方向分析:聚类对象 X_1, X_2, X_4, X_6

均属于第1灰类,即素质评价为优秀。 X_3 和 X_9 属于第2灰类,评价为良好, X_5 属于第3灰类,评价为及格,其它为第4灰类。

从列方向分析:在第1列中 $\sigma_1^1 > \sigma_2^1 > \sigma_6^1 > \sigma_4^1 > \sigma_3^1 > \sigma_{10}^1 > \sigma_8^1 > \sigma_5^1 = \sigma_7^1 = \sigma_9^1$,说明从接近第1灰类的程度看, X_1 对象的聚类系数高于其它对象,说明其综合素质状况最好。同理第4列, X_7 对象的 σ_7^4 值高于其它对象,说明其综合素质状况最差。

由图2模糊聚类的结果来看,当 $\lambda = 0.8$ 时,所有对象的综合素质可划分为7个类别, $\{X_2, X_3, X_6\}, \{X_8, X_{10}\}, \{X_1\}, \{X_4\}, \{X_5\}, \{X_7\}, \{X_9\}$; 当 $\lambda = 0.5$ 时,对象划分为3个类别, $\{X_1, X_2, X_3, X_4, X_6\}, \{X_5, X_9\}$ 和 $\{X_7, X_8, X_{10}\}$ 。

4 结束语

针对大学生综合素质评价体系指标内涵的不确定性,采用灰色和模糊聚类方法建立评价模型,并结合实例予以说明。该模型计算过程简单,充分利用了各种灰类程度的评价信息,评价结果可靠。从总体上看,评价结果能较客观地反应实际情况,方法科学合理。

参考文献:

- [1] 赵锦兰,李健希.浅谈当代大学生综合素质的培养[J].中国职工教育,2007(5):48-49.
- [2] 甘泉,胡俊英.高校学生综合素质评价的内容体系研究[J].广西教育学院学报,2011(1):84-87.
- [3] 李平,莫春立.基于神经网络方法的大学生素质综合评价[J].沈阳工业学院学报,2003(2):43-45.
- [4] 王鹏,季海鹏,王柏娜.基于模糊数学的大学生综合素质评价体系研究[J].现代服务,2005(8):173-174.
- [5] 支敏,卢云辉.基于AHP的大学生综合素质评估[J].贵州民族学院学报(哲学社会科学版),2006(4):168-170.
- [6] 吴顺祥.灰色粗糙集模型及其应用[M].北京:科学出版社,2009.
- [7] Deng Julong. Control Problems of Grey Systems[J]. Systems & Control Letters, 1982, 1(5): 288-294.
- [8] 张国辉,吴艳,张蜜.基于灰色聚类分析的企业应急管理能力评价[J].经济数学,2011,28(1):94-99.
- [9] 颜雪丽,杨胜发,吴琼.灰色聚类法在水库水环境评价中的应用[J].重庆交通大学学报(自然科学版),2010,29(5):816-819.
- [10] 谢季坚,刘承平.模糊数学方法及其应用[M].第3版.武汉:华中科技大学出版社,2006.
- [11] Zadh L A. Fuzzy sets[J]. InfoCont, 1965(8):338-353.
- [12] Hall L O, Bensaid A M, Clarke L P, et al. A comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain[J]. IEEE Transactions on Neural Networks, 1992, 3(5): 672-682.