

模糊聚类分析法在房地产企业分类中的应用

□ 文 / 董 昊

提要 在研究多个房地产企业时,常常需要按一定的标准进行分类,运用模糊聚类分析法将各个企业按其核心竞争力的特点进行分类符合实际情况。本文按照房地产企业核心竞争力的相关理论构建指标体系,利用模糊聚类分析法对房地产企业的分类进行研究,并通过实例说明该方法的可行性。

关键词 模糊聚类分析;房地产企业;核心竞争力;编网法

中图分类号 F293.3 文献标识码 A

研究者常常需要对一个地区的多个房地产企业进行分类,传统的方法是根据研究的目的选取少数几个指标对各企业进行人为的分类,其分类结果往往偏离实际情况。本文采用模糊聚类分析法,按房地产企业的核心竞争力设置指标,将核心竞争力特征类似的房地产企业归为一类,将不同特征的企业区分开,其分类结果可以为房地产企业管理层和投资商提供参考,具有一定的理论价值和现实意义。

一、模糊聚类分析

模糊聚类分析是以传统的聚类分析为理论基础,按待辨识对象的属性的亲疏关系进行软划分的一种多元统计方法。它把一个没有类别标记的样本集按某种准则划分成若干个子集(类),使相似的样本尽可能归为一类,而不

相似的样本尽量划分到不同的类中。

1、建立指标体系。本文按照房地产企业的核心竞争力对房地产企业进行分类。核心竞争力最早是由 C.K.Prahalad 和 Gary Hamel 于 1990 年在《企业核心竞争力》一文中提出的,他们指出“核心竞争力是在一组织内部经过整合了的知识和技能,尤其是关于怎样协调多种生产技能的整合了的知识和技能”。房地产企业的核心竞争力是一种合力,最终表现在土地储备能力、技术创新能力、人力资源能力、资源整合和开发能力、企业品牌效应等方面。考虑这些因素,再结合企业的一般特点,本文选取总资产、资产结构(固定/流动)、年销售额、土地储备、管理与科技人才比、高职称人数占员工总数比、品牌度等 7 个指标来建立房地产企业

核心竞争力指标体系。

2、数据标准化。根据模糊矩阵的要求,需要将数据标准化。设 n 为待分类对象的个数, m 为表征对象特性的指标个数,则:

$$x'_{ik} = \frac{x_{ik} - \bar{x}_k}{S_k} \quad (i=1, 2, \dots, n; k=1, 2, \dots, m) \quad (1)$$

$$\text{其中 } \bar{x}_k = \frac{1}{n} \sum_{i=1}^n x_{ik}$$

$$S_k = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{ik} - \bar{x}_k)^2}$$

$$x''_{ik} = \frac{x'_{ik} - \min_{1 \leq i \leq n} \{x'_{ik}\}}{\max_{1 \leq i \leq n} \{x'_{ik}\} - \min_{1 \leq i \leq n} \{x'_{ik}\}} \quad (k=1, 2, \dots, m) \quad (2)$$

表 1 各企业基本情况

指标 企业	总资产 (万元)	资产结构 (固定 / 流动)	年销售额 (万元)	土地储备 (m ²)	管理与科 技人才比	高职称人数占员 工总数比(%)	品牌度 (%)
A	21600	1/4	21000	58000	1/4	24	58
B	18700	1/3	14500	45000	1/4	16	46
C	25300	1/5	26700	69000	1/3	27	65
D	23100	1/5	24300	78000	2/5	23	61
E	19300	1/4	22000	61000	1/4	20	56
F	16500	1/4	17600	51000	1/3	17	50
G	15000	1/3	16500	41000	2/5	16	44
H	20800	1/4	19000	54000	1/3	18	53

于城市的发展速度还是相当缓慢。

(作者单位:重庆大学建设管理与房地产学院)

主要参考文献:

- [1] 秦寿康. TOPSIS 价值函数模型[J]. 系统工程学报, 2003.
- [2] 张颖超, 卢艳. 基于改进的 AHP—TOPSIS 法的和谐社会综合评价研究[J]. 商业现代化, 2008.

- [3] 陈鸿彬. 城乡统筹发展定量评价指标体系的构建[J]. 地域研究与开发, 2007.
- [4] 李岳云, 陈勇, 孙林. 城乡统筹及其评价方法[J]. 农业技术经济, 2004. 1.
- [5] 顾益康, 许勇军. 城乡一体化评估指标体系研究[J]. 浙江社会科学, 2004. 6.
- [6] Hwang C L, Yoon K. Multiple attribute decision making methods and applications[M]. 1981.
- [7] 郭翔宇, 颜华. 城乡统筹发展—理论、

- 机制、对策[M]. 中国农业出版社, 2007.
- [8] 杨建涛. 河南省城乡统筹发展测度研究[D]. 河南大学硕士学位论文, 2008.
- [9] 王丽娟. 城乡统筹发展的系统思考[D]. 西南交通大学硕士学位论文, 2005.
- [10] 张翠霞. 统筹城乡发展问题探讨[D]. 南京, 南京林业大学硕士学位论文, 2007.
- [11] 刘洪彬. 基于集群理论的统筹城乡发展研究[D]. 哈尔滨, 东北林业大学博士论文, 2006.

3. 建立模糊相似矩阵。相似系数表示两个对象之间的相似程度, 计算相似系数较简便的方法是最大最小法, 即:

$$r_{ij} = \frac{\sum_{k=1}^m (x_{ik} \wedge x_{jk})}{\sum_{k=1}^m (x_{ik} \vee x_{jk})} \quad (3)$$

于是便可利用相似系数建立模糊相似矩阵 $R=(r_{ij})_{n \times n}$ 。

4. 聚类。目前, 常用的聚类方法有传递闭包法、最大树法和编网法。其中, 编网法是将建立的模糊相似矩阵改造为编网图, 具体做法是: 由对称性取 R 的下三角矩阵, 首先把对角线上的“1”改写成对应元素的符号; 对给定的水平 λ , 当矩阵 R 的元素 $r_{ij} \geq \lambda$ 时, 用“*”号取代元素 r_{ij} , 而当 $r_{ij} < \lambda$ 时, 则抹去 r_{ij} , 它的位置变为空格。把画“*”的位置称为结点, 由结点向矩阵的主对角线引竖直的经线和水平的纬线, 经过同一结点的经线和纬线看作被捆扎在一起了。一捆经纬线的末端所系的对角线上的元素属于同一类。

5. 确定最佳阈值。在聚类分析中, 不同的 λ 可能对应不同的分类, 这时可用 F -统计量确定 λ 的最佳值。设 r 为对应于 λ 值的类数, n_i 为第 i 类元素的个数, \bar{x}_k 表示第 i 类元素的第 k 个特征的平均值, \bar{x}_k 表示全体样本第 k 个特征的平均值, 则 F -统计量为:

$$F = \frac{\sum_{i=1}^r \sum_{k=1}^m (\bar{x}_k - \bar{x}_k)^2 / (r-1)}{\sum_{i=1}^r \sum_{j=1}^n \sum_{k=1}^m (x_{ik} - \bar{x}_k)^2 / (n-r)} \quad (4)$$

查 F 临界值得 F_{α} (α 常取 0.05)。在满足 $F > F_{\alpha}$ 的所有情形中, 差值 $F - F_{\alpha}$ 的最大者所对应的分类即为最佳分类。

二、实例分析

选取某市 8 家有代表性的房地产开发公司为研究对象, 用前文所建立的指标体系来表征其核心竞争力, 各企业的具体情况, 如表 1 所示。(表 1)

用公式(1)、(2)得标准化矩阵 R' :

$$R' = \begin{pmatrix} 0.64 & 0.38 & 0.53 & 0.46 & 0 & 0.73 & 0.67 \\ 0.36 & 1 & 0 & 0.11 & 0 & 0 & 0.1 \\ 1 & 0 & 1 & 0.76 & 0.56 & 1 & 1 \\ 0.79 & 0 & 0.80 & 1 & 1 & 0.64 & 0.81 \\ 0.42 & 0.38 & 0.61 & 0.54 & 0 & 0.36 & 0.57 \\ 0.15 & 0.38 & 0.25 & 0.27 & 0.56 & 0.09 & 0.29 \\ 0 & 1 & 0.16 & 0 & 1 & 0 & 0 \\ 0.56 & 0.38 & 0.37 & 0.35 & 0.56 & 0.18 & 0.43 \end{pmatrix}$$

用公式(3)算出相似系数, 得到对称的模糊相似矩阵 R :

$$R = \begin{pmatrix} 1 & & & & & & \\ 0.23 & 1 & & & & & \\ 0.53 & 0.09 & 1 & & & & \\ 0.53 & 0.09 & 0.72 & 1 & & & \\ 0.76 & 0.27 & 0.44 & 0.46 & 1 & & \\ 0.36 & 0.26 & 0.28 & 0.30 & 0.41 & 1 & \\ 0.11 & 0.37 & 0.11 & 0.19 & 0.12 & 0.36 & 1 \\ 0.57 & 0.27 & 0.43 & 0.45 & 0.59 & 0.70 & 0.28 & 1 \end{pmatrix}$$

以 R 为基础用编网法进行聚类, 做出编网图:

①当 $0.76 < \lambda \leq 1$ 时, 有:

$$R_{\lambda} = \begin{pmatrix} A & & & & & & \\ & B & & & & & \\ & & C & & & & \\ & & & D & & & \\ & & & & E & & \\ & & & & & F & \\ & & & & & & G \\ & & & & & & & H \end{pmatrix}$$

故 X 分为 8 类: $\{A\}$, $\{B\}$, $\{C\}$, $\{D\}$, $\{E\}$, $\{F\}$, $\{G\}$, $\{H\}$ 。

②当 $0.72 < \lambda \leq 0.76$ 时, 有:

$$R_{\lambda} = \begin{pmatrix} A & & & & & & \\ | & B & & & & & \\ | & & C & & & & \\ | & & & D & & & \\ * & - & - & - & E & & \\ & & & & & F & \\ & & & & & & G \\ & & & & & & & H \end{pmatrix}$$

故 X 分为 7 类: $\{A, E\}$, $\{B\}$, $\{C\}$, $\{D\}$, $\{F\}$, $\{G\}$, $\{H\}$ 。

同理, 当 $0.70 < \lambda \leq 0.72$ 时, X 分为 6 类: $\{A, E\}$, $\{C, D\}$, $\{B\}$, $\{F\}$, $\{G\}$, $\{H\}$; 当 $0.59 < \lambda \leq 0.70$ 时, X 分为 5 类: $\{A, E\}$, $\{C, D\}$, $\{F, H\}$, $\{B\}$, $\{G\}$; 当 $0.53 < \lambda \leq 0.59$ 时, X 分为 4 类: $\{A, E, F, H\}$, $\{C, D\}$, $\{B\}$, $\{G\}$; 当 $0.37 < \lambda \leq 0.53$ 时, X 分为 3 类: $\{A, C, D, E, F, H\}$, $\{B\}$, $\{G\}$; 当 $0.36 < \lambda \leq 0.37$ 时, X 分为 2 类: $\{A, C, D, E, F, H\}$, $\{B, G\}$; 当 $0 \leq \lambda \leq 0.36$ 时, X 分为 1 类: $\{A, B, C, D, E, F, G, H\}$ 。

将表 1 中的数据代入公式(4)计算各方案的 F 值, 查出各方案的临界值 $F_{0.05}$, 结果如表 2 所示。(表 2)

三、结论

上述分类结果中, 8 家房地产开发公司各自成类和全部并为一类没有实际应用价值, 不予考虑。其他的 6 个分类方案中, 只有当 $0.59 < \lambda \leq 0.70$ 时, F 值大于 $F_{0.05}(r-1, n-r)$, 分类比较合理, 因此按照所选取的房地产企业核心竞争力 7 个指标的相近程度, 将这 8 家公司分为 5 类为最佳分类, 用集合表示为: $\{A, E\}$, $\{C, D\}$, $\{F, H\}$, $\{B\}$, $\{G\}$ 。

按核心竞争力对房地产企业进行模糊聚类分析, 不仅将具有相同特征的企业并到了一块, 而且将不同特征的企业区分开来, 其分类结果充分反映了各企业在核心竞争力方面的特点。进行聚类后, 处于同一类的企业在认识到自身特点的同时, 可以在制定发展战略、提升自身核心竞争力方面相互借鉴; 另外, 对于投资商而言, 将欲投资的企业聚类后, 可以在分类结果中根据投资的目的和企业的特点做进一步的比较和分析, 为最终的投资决策提供依据。

(作者单位: 西安建筑科技大学管理学院)

主要参考文献:

- [1] 高新波. 模糊聚类分析及其应用[M]. 西安: 西安电子科技大学出版社, 2004.
- [2] C. K. Prahalad and Gary Hamel. The Core Competence of the Corporation[J]. Harvard Business Review, 1990.
- [3] 梁保松, 曹殿立. 模糊数学及其应用[M]. 北京: 科学出版社, 2007.
- [4] 刘林. 应用模糊数学(第二版)[M]. 西安: 陕西科学技术出版社, 2008.

表 2 各 λ 所对应的 F -统计量

λ 值	1	0.76	0.72	0.70	0.59	0.53	0.37	0.36
r 类	8	7	6	5	4	3	2	1
F	0	11.75	6.54	9.17	5.86	2.57	4.56	0
$F_{0.05}(r-1, n-r)$	—	234.0	19.30	9.12	6.59	5.79	5.99	—