

文章编号: 1001-0920(2005)05-0533-04

## 一种基于语言评价信息的多指标群聚类方法

于春海, 樊治平, 姜艳萍

(东北大学 工商管理学院, 辽宁 沈阳 110004)

**摘要:** 针对具有语言评价信息的多指标群聚类分析问题, 提出一种基于二元语义信息处理的极大树聚类方法。首先描述了具有语言评价信息的多指标群聚类问题, 并介绍了近年来国际上最新发展的二元语义概念及其算子; 然后基于二元语义信息处理的方法, 将专家给出的语言评价信息进行“量化”集结, 并依据传统的极大树聚类分析方法的基本思路, 给出了解决基于语言评价信息的极大树聚类方法的计算步骤; 最后通过一个算例说明了所提出方法的有效性。

**关键词:** 聚类分析; 语言评价信息; 二元语义; 极大树聚类法; 群聚类

**中图分类号:** TP391.41; N945.25

**文献标识码:** A

## Method for multiple attribute group clustering based on linguistic evaluation information

YU Chun-hai, FAN Zhi-ping, JIANG Yan-ping

(School of Business Administration, Northeastern University, Shenyang 110004, China Correspondent: YU Chun-hai, E-mail: yuchunhai@sina.com)

**Abstract:** The multiple attribute group clustering analysis problem is described, and the two-tuple linguistic concept and its operator developed in recent years are introduced. A maximal tree clustering analysis method is proposed based on two-tuple linguistic information processing. Based on the approach to two-tuple linguistic processing, linguistic assessment information given by experts is quantified and aggregated, according to the basic ideas of traditional maximal tree clustering method, the calculation steps of the maximal tree clustering method with linguistic assessment information are developed. Finally, a numerical example shows the applicability of the proposed method.

**Key words:** clustering analysis; linguistic assessment information; two-tuple linguistic; maximal tree clustering method; group clustering

### 1 引言

聚类分析在经济管理及工程系统具有广泛的实际背景。自Ruspini于1969年提出模糊划分概念以来<sup>[1]</sup>, 有关聚类分析的研究得到了迅猛发展。目前, 比较典型和常用的聚类分析方法有聚合法和分裂法<sup>[2]</sup>, 传递闭包方法<sup>[3]</sup>, 极大树方法<sup>[4]</sup>, C-均值模糊聚类法<sup>[5]</sup>等。需要指出的是, 上述各种方法主要用于数值信息的情形。但在许多实际问题中, 由于被

聚类对象的特征指标的模糊性和不确定性, 经常会遇到专家给出的关于被聚类对象的特征指标评价信息是语言评价短语形式的情形<sup>[6]</sup>。因此, 针对具有语言评价信息的聚类分析问题的研究引起了人们的关注<sup>[7,8]</sup>, 但相关的研究成果甚少。

本文针对多个专家给出的语言评价信息的多指标聚类问题, 给出了一种新的聚类分析方法。在该方法中, 首先采用近年来国际上最新发展的二元语义

收稿日期: 2004-03-08; 修回日期: 2004-06-21

基金项目: 国家自然科学基金项目(70371050, 70301008); 教育部高等学校优秀青年教师教学科研奖励计划项目(教人司[2002]123); 辽宁省自然科学基金项目(20032028)。

作者简介: 于春海(1968—), 男, 吉林长春人, 副教授, 博士, 从事管理决策和管理统计分析等研究; 樊治平(1961—), 男, 江苏镇江人, 教授, 博士生导师, 从事管理决策分析、信息管理 with 知识管理等研究。

信息处理和运算的方法<sup>[9]</sup>,将每个专家给出的语言评价信息进行“量化”集结;然后依据传统的最大树聚类分析方法的基本思路<sup>[4]</sup>,给出一种基于二元语义信息处理的多指标群聚类分析方法的计算步骤

## 2 问题描述

在多指标群聚类分析问题中,设被聚类对象集为  $X = \{X_1, X_2, \dots, X_n\}$ , 特征指标集为  $Q = \{Q_1, Q_2, \dots, Q_p\}$ , 专家群体集为  $D = \{D_1, D_2, \dots, D_m\}$ , 其中  $D_h$  表示第  $h$  个专家. 假设专家  $D_h$  针对指标集  $Q$  给出具有语言评价信息形式的权重向量为  $w^h = (w_1^h, w_2^h, \dots, w_p^h)^T$ , 其中  $w_k^h$  为专家  $D_h$  从预先定义好的语言(或语言符号)评价集  $L$  中, 选择一个元素作为对指标  $Q_k$  的重要程度描述. 记  $\tilde{x}_{ik}^h$  为专家  $D_h$  针对第  $i$  个被聚类对象  $X_i$  按第  $k$  个指标  $Q_k$  进行测度的结果(即评价值), 其中  $\tilde{x}_{ik}^h$  为从预先定义好的语言(或语言符号)评价集  $S$  中, 选择一个元素作为被聚类对象  $X_i$  对应于  $Q_k$  的评价值(或效用值). 这里,  $S$ (或  $L$ ) 是一个预先定义好的由奇数个元素构成的有序集合, 而本文所考虑的  $S$  和  $L$  均是由 7 个元素(即语言短语)构成的集合, 即  $L = S = \{S_0 = \text{FC}$ (非常差/非常不重要),  $S_1 = \text{HC}$ (很差/很不重要),  $S_2 = \text{C}$ (差/不重要),  $S_3 = \text{YB}$ (一般),  $S_4 = \text{H}$ (好/重要),  $S_5 = \text{HH}$ (很好/很重要),  $S_6 = \text{FH}$ (非常好/非常重要)}. 本文要解决的问题是根据多个专家给出的具有语言评价形式的特征指标评价值  $\tilde{x}_{ik}^h$  ( $i = 1, 2, \dots, n, k = 1, 2, \dots, p, h = 1, 2, \dots, m$ ) 和指标权重  $w_k^h$  ( $k = 1, 2, \dots, p, h = 1, 2, \dots, m$ ), 给出一个将  $X$  合理分类的聚类分析方法

## 3 二元语义及其运算算子

二元语义是基于符号平移概念, 采用二元组  $(s_i, \alpha)$  表示语言评价信息<sup>[9,10]</sup>. 其中  $s_i$  表示事先定义的语言评价集  $S$  中的语言短语,  $\alpha$  表示由计算得到的语言信息与初始语言评价集中最贴近语言短语之间的差别, 并且  $\alpha \in [-0.5, 0.5]$ . 下面对二元语义及其有关的运算算子给予简单介绍<sup>[9]</sup>.

记  $S = \{s_0, s_1, \dots, s_T\}$  表示包含  $T + 1$  个语言短语的集合, 其中  $s_i \in S$  是一个语言短语. 那么,  $s_i$  相应的二元语义  $(s_i, 0)$  可通过如下转换函数  $\theta$  得到:

$$\theta S = S \times [-0.5, 0.5], \quad (1a)$$

$$\theta(s_i) = (s_i, 0), s_i \in S. \quad (1b)$$

设  $\beta$  表示语言短语集结运算的结果, 且  $\beta \in [0, T]$ , 则与  $\beta$  相应的二元语义可由如下函数  $\Delta$  得到:

$$\Delta: [0, T] \rightarrow S \times [-0.5, 0.5]; \quad (2a)$$

$$\Delta(\beta) = (s_i, \alpha) = \begin{cases} s_i, i = \text{Round}(\beta); \\ \alpha = \beta - i, \alpha \in [-0.5, 0.5]. \end{cases} \quad (2b)$$

其中  $\text{Round}$  表示“四舍五入”取整运算

若  $(s_i, \alpha)$  是一个二元语义,  $s_i \in S, \alpha \in [-0.5, 0.5]$ , 则通过如下函数  $\Delta^{-1}$  可将二元语义  $(s_i, \alpha)$  转化为相应的数值  $\beta \in [0, T]$ , 即

$$\Delta^{-1}: S \times [-0.5, 0.5] \rightarrow [0, T], \quad (3a)$$

$$\Delta^{-1}: (s_i, \alpha) = i + \alpha = \beta \quad (3b)$$

设  $(x_1, \alpha_1), (x_2, \alpha_2), \dots, (x_m, \alpha_m)$  是一组要集结的二元语义, 则基于二元语义的算术平均算子  $\Phi$  定义为<sup>[9]</sup>

$$\begin{aligned} (\bar{x}, \bar{\alpha}) &= \Phi((x_1, \alpha_1), (x_2, \alpha_2), \dots, (x_m, \alpha_m)) = \\ &= \Delta \left( \frac{1}{m} \sum_{i=1}^m \Delta^{-1}(x_i, \alpha_i) \right), \bar{x} \in S, \bar{\alpha} \in [-0.5, 0.5]. \end{aligned} \quad (4)$$

## 4 基于二元语义信息处理的最大树聚类分析方法

基于上述有关二元语义的概念和运算算子, 可以看到, 通过二元语义信息处理和运算的方法<sup>[9]</sup>, 使原有离散的语言评价信息的评价结果连续化; 通过将每个专家给出的语言评价信息进行“量化”集结, 使“集结”后的结果更加准确, 并最大可能地反映群体的综合评价意见. 此外, 应该说具有语言评价信息的聚类对象之间的边界具有不清晰性, 对这种聚类对象的分类也伴随着模糊性, 而传统的最大树聚类分析方法是基于模糊集理论, 而且是通过建立模糊相似关系后用适当的手段将给定的研究对象进行分类的一种模糊聚类分析方法. 所以本文依据二元语义及其运算算子和传统的最大树聚类分析方法<sup>[4]</sup>的基本思想, 给出解决具有语言评价信息的多指标群聚类问题的聚类分析方法, 其计算步骤如下:

Step 1: 将专家给出的语言评价信息转化为二元语义形式的评价信息. 根据式(1b), 利用转换函数  $\theta$  将语言评价信息  $\tilde{x}_{ik}^h$  和  $w_k^h$  分别转化成  $(\tilde{x}_{ik}^h, 0)$  和  $(w_k^h, 0)$ .

Step 2: 将二元语义形式的评价信息集结为群的综合评价信息. 根据二元语义算术平均算子的计算公式(4), 计算针对每个特征指标的群的特征指标评价值  $(\tilde{x}_{ik}, \alpha_{ik})$  ( $i = 1, 2, \dots, n, k = 1, 2, \dots, p$ ) 和群的特征指标权重  $(\tilde{w}_k, \alpha_k)$ ,  $k = 1, 2, \dots, p$ , 计算公式如下:

$$\begin{aligned} (\tilde{x}_{ik}, \alpha_{ik}) &= \Delta \left( \frac{1}{m} \sum_{h=1}^m \Delta^{-1}(\tilde{x}_{ik}^h, \alpha_{ik}^h) \right), \\ i &= 1, 2, \dots, n, k = 1, 2, \dots, p; \end{aligned} \quad (5)$$

$$\begin{aligned} (\tilde{w}_k, \alpha_k) &= \Delta \left( \frac{1}{m} \sum_{h=1}^m \Delta^{-1}(w_k^h, \alpha_k^h) \right), \\ k &= 1, 2, \dots, p. \end{aligned} \quad (6)$$

式中:  $(\tilde{x}_{ik}, \alpha_k)$  表示聚类对象  $X_i$  针对特征指标  $Q_k$  的群的特征指标评价价值;  $(\tilde{w}_k, \alpha_k)$  表示针对指标  $Q_k$  的群的重要程度描述

Step3: 计算加权特征指标值 对  $(\tilde{x}_{ik}, \alpha_k)$  进行加权计算, 并将计算结果记为  $x_{ik}, i = 1, 2, \dots, n, k = 1, 2, \dots, p$ , 其计算公式为

$$x_{ik} = \frac{\Delta^{-1}(\tilde{x}_{ik}, \alpha_k) \times \Delta^{-1}(\tilde{w}_k, \alpha_k)}{\Delta^{-1}(\tilde{w}_l, \alpha_k)}, \quad i = 1, 2, \dots, n, k = 1, 2, \dots, p. \quad (7)$$

记  $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})$ .

Step4: 建立相似系数矩阵  $R (r_{ij})^{[11]}$ . 针对由 Step3 得到的结果, 记  $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})$ , 计算关于  $x_i$  和  $x_j$  的相似系数  $r_{ij}$ , 其计算公式为

$$r_{ij} = \frac{\left| \prod_{k=1}^p (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j) \right|}{\left[ \prod_{k=1}^p (x_{ik} - \bar{x}_i)^2 \prod_{k=1}^p (x_{jk} - \bar{x}_j)^2 \right]^{1/2}}, \quad i, j = 1, 2, \dots, n \quad (8)$$

其中  $\bar{x}_i = \frac{1}{p} \sum_{k=1}^p x_{ik}, \bar{x}_j = \frac{1}{p} \sum_{k=1}^p x_{jk}$

Step5: 画最大树 以  $X_i (i = 1, 2, \dots, n)$  为顶点画出模糊赋权图  $G = (X, r_{ij})$ , 其中相似系数  $r_{ij}$  为模糊图中各边对应的权. 在模糊赋权图  $G = (X, r_{ij})$  中任取一回路, 去掉其中对应权最小的边, 依次类推直至模糊赋权图中没有回路时剩下的各边便构成了最大树

Step6: 分类 首先确定截割水平  $\lambda$  然后根据  $\lambda$  值对最大树进行切割. 具体作法是: 分别比较  $\lambda$  与最大树各边的权 (在此即相似系数  $r_{ij}$ ) 的大小. 当  $\lambda$  大于  $r_{ij}$  时, 将  $r_{ij}$  对应的边截断, 这样剩余的且相互联通的聚类对象便构成了一类, 于是得到关于  $X = \{X_1, X_2, \dots, X_n\}$  的一个分类

Step7: 灵敏度分析 根据聚类分析的需要, 可得出在给定截割水平  $\lambda$  意义下的聚类对象分类的灵敏度分析, 即可选取不同的  $\lambda$  值, 并得出不同  $\lambda$  对应的分类结果. 据此分析  $\lambda$  对分类结果的影响, 并可从中选择适当的分类以满足进一步问题的需要

### 5 算 例

下面采用前面给出的聚类分析方法, 解决一个具有多指标语言评价信息的战略伙伴选择问题<sup>[12]</sup>. 问题的核心是对若干个候选战略伙伴的组织整合能力进行识别并分类. 假设有 7 个候选战略伙伴 (即被聚类对象为  $X_1, X_2, X_3, X_4, X_5, X_6, X_7$ ), 其组织整合

能力由 8 个特征指标 ( $Q_1, Q_2, Q_3, Q_4, Q_5, Q_6, Q_7, Q_8$ ) 来反映, 其涵义是<sup>[13]</sup>:  $Q_1$  表示研究发展能力,  $Q_2$  表示制造能力,  $Q_3$  表示营销能力,  $Q_4$  表示子公司/事业部能力,  $Q_5$  表示功能之间的界面整合能力,  $Q_6$  表示子公司/事业部之间的界面整合能力,  $Q_7$  表示内部管理意识,  $Q_8$  表示核心人才管理能力. 专家群体集为  $D = \{D_1, D_2, D_3, D_4\}$ . 假设 4 个专家针对理想的合作伙伴  $X_0$  和 7 个候选合作伙伴就 8 个特征指标分别给出具有语言评价形式的特征指标评价价值和特征指标权重, 如表 1 和表 2 所示

表 1 专家给出的具有语言评价信息形式的特征指标评价

| $D_h$ | $X_i$ | $Q_1$ | $Q_2$ | $Q_3$ | $Q_4$ | $Q_5$ | $Q_6$ | $Q_7$ | $Q_8$ |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $D_1$ | $X_1$ | FC    | YB    | FC    | HC    | C     | FC    | C     | HC    |
|       | $X_2$ | H     | YB    | HC    | FH    | FH    | HC    | H     | HC    |
|       | $X_3$ | FH    | FH    | H     | HC    | FH    | C     | YB    | HH    |
|       | $X_4$ | YB    | FH    | HH    | C     | YB    | YB    | HC    | HC    |
|       | $X_5$ | FH    | H     | HH    | H     | HH    | HH    | HH    | HH    |
|       | $X_6$ | C     | H     | YB    | HC    | HC    | FC    | H     | HC    |
|       | $X_7$ | H     | YB    | HH    | YB    | C     | H     | YB    | C     |
| $D_2$ | $X_1$ | HC    | YB    | FC    | C     | C     | FC    | C     | HC    |
|       | $X_2$ | H     | YB    | C     | FH    | FH    | C     | H     | HC    |
|       | $X_3$ | HH    | FH    | HH    | HC    | H     | HC    | H     | HH    |
|       | $X_4$ | C     | FH    | H     | C     | YB    | YB    | HC    | C     |
|       | $X_5$ | FH    | H     | HH    | HH    | H     | H     | HH    | H     |
|       | $X_6$ | HC    | H     | H     | HC    | HC    | HC    | H     | HC    |
|       | $X_7$ | H     | YB    | H     | H     | YB    | H     | YB    | HC    |
| $D_3$ | $X_1$ | FC    | YB    | FC    | HC    | C     | FC    | C     | HC    |
|       | $X_2$ | HH    | YB    | HC    | FH    | FH    | C     | H     | C     |
|       | $X_3$ | FH    | HH    | H     | HC    | FH    | C     | YB    | HH    |
|       | $X_4$ | H     | FH    | H     | HC    | YB    | YB    | HC    | HC    |
|       | $X_5$ | FH    | YB    | HH    | H     | H     | H     | H     | H     |
|       | $X_6$ | YB    | HH    | YB    | HC    | HC    | FC    | HH    | HC    |
|       | $X_7$ | H     | YB    | HH    | YB    | YB    | H     | C     | C     |
| $D_4$ | $X_1$ | HC    | C     | FC    | C     | HC    | FC    | HC    | HC    |
|       | $X_2$ | H     | YB    | HC    | FH    | FH    | HC    | H     | FC    |
|       | $X_3$ | FH    | FH    | H     | HC    | FH    | C     | YB    | H     |
|       | $X_4$ | YB    | HH    | FH    | C     | H     | YB    | HC    | HC    |
|       | $X_5$ | FH    | HH    | HH    | H     | FH    | HH    | H     | HH    |
|       | $X_6$ | C     | H     | C     | HC    | C     | FC    | HH    | HC    |
|       | $X_7$ | H     | YB    | HH    | YB    | HC    | H     | YB    | C     |

表 2 专家给出的具有语言评价信息形式的特征指标权重

| $D_h$ | $w_k^h$ | $Q_1$ | $Q_2$ | $Q_3$ | $Q_4$ | $Q_5$ | $Q_6$ | $Q_7$ | $Q_8$ |
|-------|---------|-------|-------|-------|-------|-------|-------|-------|-------|
| $D_1$ | $w_k^1$ | YB    | HH    | C     | YB    | YB    | FH    | H     | YB    |
| $D_2$ | $w_k^2$ | H     | HH    | HC    | C     | H     | FH    | YB    | H     |
| $D_3$ | $w_k^3$ | YB    | FH    | C     | HC    | H     | HH    | C     | H     |
| $D_4$ | $w_k^4$ | C     | HH    | FC    | C     | C     | FH    | YB    | HH    |

下面根据前面给出的聚类分析方法的计算步骤, 对被聚类对象 ( $X_0, X_1, X_2, X_3, X_4, X_5, X_6, X_7$ ) 进行分类. 为节省篇幅, 这里仅给出其中的一些计算结果. 根据 Step3, 计算得到对应于群的特征指标评价价值  $\tilde{x}_{ik}$  和特征指标权重的二元语义表示如表 3 所示

表3 对应于群的特征指标评价 $\tilde{x}_{ik}$ 和特征指标权重的二元语义表示

| $X_i$         | $Q_1$        | $Q_2$        | $Q_3$        | $Q_4$       | $Q_5$        | $Q_6$        | $Q_7$        | $Q_8$        |
|---------------|--------------|--------------|--------------|-------------|--------------|--------------|--------------|--------------|
| $X_0$         | (FH, 0)      | (FH, 0)      | (FH, 0)      | (FH, 0)     | (FH, 0)      | (FH, 0)      | (FH, 0)      | (FH, 0)      |
| $X_1$         | (HC, - 0.5)  | (YB, - 0.25) | (FC, 0)      | (C, - 0.25) | (C, - 0.25)  | (FC, 0)      | (C, - 0.25)  | (HC, 0)      |
| $X_2$         | (H, 0.25)    | (YB, 0)      | (C, - 0.5)   | (FH, 0)     | (FH, 0)      | (C, - 0.5)   | (H, 0)       | (HC, 0)      |
| $X_3$         | (FH, - 0.25) | (FH, - 0.25) | (H, 0.25)    | (HC, 0)     | (FH, - 0.5)  | (C, - 0.25)  | (YB, 0.25)   | (HH, - 0.25) |
| $X_4$         | (YB, 0)      | (FH, - 0.25) | (HH, - 0.25) | (C, - 0.25) | (YB, 0.25)   | (YB, 0)      | (HC, 0)      | (HC, 0.25)   |
| $X_5$         | (FH, 0)      | (H, 0)       | (HH, 0)      | (H, 0.25)   | (HH, - 0.25) | (HH, - 0.5)  | (HH, - 0.5)  | (HH, - 0.5)  |
| $X_6$         | (C, 0)       | (H, 0.25)    | (YB, 0)      | (HC, 0)     | (HC, 0.25)   | (FC, 0.25)   | (HH, - 0.5)  | (HC, 0)      |
| $X_7$         | (H, 0)       | (YB, 0)      | (HH, - 0.25) | (YB, 0.25)  | (C, 0.25)    | (H, 0)       | (YB, - 0.25) | (C, - 0.25)  |
| $\tilde{x}_k$ | (YB, 0)      | (HH, 0.25)   | (HC, 0.25)   | (C, 0)      | (YB, 0.25)   | (FH, - 0.25) | (YB, 0)      | (H, 0)       |

根据式(8), 可得相似系数矩阵R( $r_{ij}$ )的计算结果

$$R(r_{ij}) = \begin{bmatrix} 1 & 0.39 & 0.24 & 0.61 & 0.74 & 0.96 & 0.33 & 0.83 \\ 0.39 & 1 & 0.56 & 0.75 & 0.65 & 0.22 & 0.87 & 0.07 \\ 0.24 & 0.56 & 1 & 0.42 & 0.38 & 0.24 & 0.38 & 0.13 \\ 0.61 & 0.75 & 0.42 & 1 & 0.72 & 0.60 & 0.67 & 0.29 \\ 0.74 & 0.65 & 0.38 & 0.72 & 1 & 0.65 & 0.61 & 0.70 \\ 0.96 & 0.22 & 0.24 & 0.60 & 0.65 & 1 & 0.21 & 0.85 \\ 0.33 & 0.87 & 0.38 & 0.67 & 0.61 & 0.21 & 1 & 0.16 \\ 0.83 & 0.07 & 0.13 & 0.29 & 0.70 & 0.85 & 0.16 & 1 \end{bmatrix}$$

由此依据 Step5 得到最大树如图 1 所示

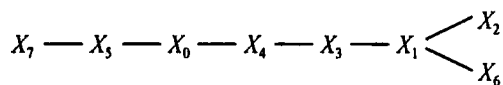


图1 最大树

根据给定的不同的截割水平 $\lambda$  聚类结果的灵敏度分析如表4所示 根据表4的聚类结果, 再结合问题的实际情况, 决策者可选取不同的截割水平, 得到适用的分类结果 以 $\lambda$ 取值为0.72  $\lambda < 0.75$ 为例, 此时与理想合作伙伴 $X_0$ 分在一类的有 $X_5$ 和 $X_7$ . 这说明选择合作伙伴首先考虑的应是 $X_5$ 和 $X_7$ , 至于选择 $X_5$ 还是 $X_7$ , 则需再根据其他情况由决策者取舍. 类似地,  $\lambda$ 取值为0.87  $\lambda < 0.96$ 时, 最佳合作伙伴为 $X_5$  其他情况依次类推

表4 截割水平 $\lambda$ 不同取值及其分类结果

| $\lambda$               | 聚类结果   |
|-------------------------|--|
| $0.56 < \lambda < 0.72$ | $\{X_2\}, \{X_0, X_1, X_3, X_4, X_5, X_6, X_7\}$                     |
| $0.72 < \lambda < 0.75$ | $\{X_2\}, \{X_3\}, \{X_4\}, \{X_1, X_6\}, \{X_0, X_5, X_7\}$         |
| $0.75 < \lambda < 0.87$ | $\{X_2\}, \{X_3\}, \{X_4\}, \{X_1, X_6\}, \{X_0, X_5\}, \{X_7\}$     |
| $0.87 < \lambda < 0.96$ | $\{X_2\}, \{X_1\}, \{X_3\}, \{X_6\}, \{X_4\}, \{X_7\}, \{X_0, X_5\}$ |

## 6 结 语

针对具有语言评价信息的多指标群聚类问题, 给出了一种聚类分析方法 该方法通过采用近年来国际上最新发展的二元语义及其集结运算算子对语言评价信息进行“量化”集结处理, 将传统最大树聚类法扩展到具有语言评价信息的群聚类情形 这种

“量化”集结处理能充分反映语言评价信息的模糊性, 而且具有使语言评价信息运算概念清晰, 计算简单的特点

## 参考文献 (References)

- [1] Ruspini E H. A new approach to clustering [J] *Information and Control*, 1969, 15(1): 22-32
- [2] Tamura S, Higuchi S, Tanaka K. Pattern classification based on fuzzy relations[J] *IEEE Trans on Systems, Man and Cybernetics*, 1971, 1(1): 217-242
- [3] Zkin L. Fuzzy relation compositions and pattern recognition [J] *Information Sciences*, 1996, 89 (1-2): 107-130
- [4] Wu Z, Leathy R. An optimal graph theoretic to data clustering: Theory and its application to image segmentation [J] *IEEE Trans on Pattern Analysis Machine Intelligence*, 1993, 15(11): 1101-1113
- [5] Yang M S, Ko C H. On a class of fuzzy C-numbers clustering problems for fuzzy data[J] *Fuzzy Sets and Systems*, 1996, 84(1): 49-60
- [6] Delgado M, Verdegay J L, Vila M A. On aggregation operations of linguistic label[J] *Int J of Intelligent Systems*, 1993, 8(3): 351-370
- [7] Sonbaty Y E, Ismail M A. Fuzzy clustering for symbolic data [J] *IEEE Trans on Fuzzy Systems*, 1998, 6(2): 195-201
- [8] Kalyani M, Sushmita M. Clustering and its validation in a symbolic framework [J] *Pattern Recognition Letters*, 2003, 24(14): 2367-2376
- [9] Herrera F, Martinez L. A 2-tuple fuzzy linguistic representation model for computing with words [J] *IEEE Trans on Fuzzy Systems*, 2000, 8(6): 746-752
- [10] Herrera F, Martinez L. A model based on linguistic 2-tuples for dealing with multigranularity hierarchical linguistic contexts in multiexpert decision making [J] *IEEE Trans on Systems, Man and Cybernetics-Part B: Cybernetics*, 2001, 31(2): 227-234
- [11] 杨论标, 高英仪. 模糊数学原理及应用[M] 广州: 华南理工大学出版社, 2002: 115-125

(下转第540页)

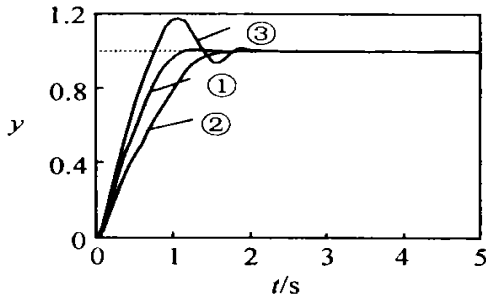


图2 线性系统的阶跃响应

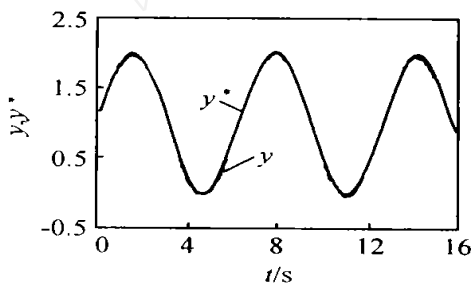
例2(非线性系统被控对象) 采样周期为0.01 s, 控制对象为一个非线性系统

$$y(k+1) =$$

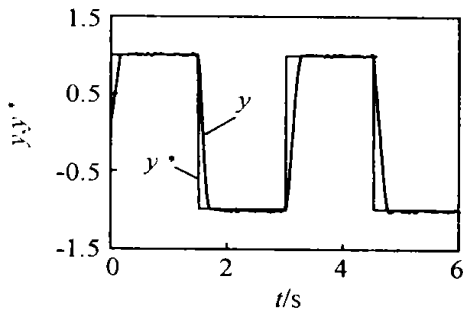
$$\frac{y(k)y(k-1)(0.4y(k)+1)}{1+y^2(k)+y^2(k-1)} + 0.4y(k) +$$

$$0.2y(k-1) + 0.8u(k) + 0.2u(k-1).$$

图3为跟踪响应仿真结果, 其中图3(a)为正弦函数跟踪响应, 参考输入  $y^* = 1 + \sin t$ , 系统输出为  $y$ ;



(a) 正弦函数跟踪响应



(b) 正弦函数跟踪响应

图3 非线性系统的跟踪响应仿真

图3(b)为方波函数跟踪响应, 参考输入  $y^*$  为方波函数, 系统输出为  $y$ 。

## 6 结论

支持向量机比神经网络具有更强的理论依据和泛化能力, 它克服了模糊神经网络依赖经验设计结构、过学习、泛化能力不强等缺陷。本文的模糊支持向量机控制器, 结合了模糊控制和支持向量机的优点, 既具有模糊逻辑不依赖被控对象模型、鲁棒性强的特点, 又具有支持向量机小样本学习、泛化能力强等优点。对于线性、非线性系统, 所设计的控制器控制性能均较优。

## 参考文献(References)

- [1] 王耀南. 智能控制系统[M]. 长沙: 湖南大学出版社, 1996.
- [2] Nello Cristianini, John Shawe-Taylor. *An introduction to support vector machines and other kernel-based learning methods*[M]. Cambridge: Cambridge University Press, 2000.
- [3] Vapnik V N. An overview of statistical learning theory [J]. *IEEE Trans on Neural Networks*, 1999, 10(5): 988-999.
- [4] Christopher J C Burges. A tutorial on support vector machines for pattern recognition [J]. *Data Mining and Knowledge Discovery*, 1998, 2(2): 121-167.
- [5] Chan W C, Chan C W, Cheung K C, et al. On the modelling of nonlinear dynamic system using support vector neural networks [J]. *Engineering Applications of Artificial Intelligence*, 2001, 14(2): 105-113.
- [6] Suykens J A K, Vandewalle J, Moor B De. Optimal control by least squares support machines [J]. *Neural Networks*, 2001, 14(1): 23-35.
- [7] Platt J. Sequential minimal optimization: A fast algorithm for training support vector machine [R]. Washington: Microsoft Research, 1998.
- [8] LeCun Y A, Jackel L D, Bottou L, et al. Learning algorithms for classification: A comparison on handwritten digit recognition [A]. *Neural Networks: The Statistical Mechanics Perspective* [C]. Singapore, 1995: 261-276.

(上接第536页)

- [12] 冯蔚东, 陈剑. 虚拟企业中核心能力的定性定量识别 [J]. *系统工程理论与实践*, 2002, 16(5): 48-54.  
(Fen W D, Chen J. Identifying the core competence for virtual enterprises: A qualitative framework and a quantitative method [J]. *Systems Engineering — Theory and Practice*, 2002, 22(5): 48-54.)

- [13] 王毅, 陈劲. 企业核心能力高标定位研究 [J]. *管理工程学报*, 2002, 16(4): 22-27.  
(Wang Y, Chen J. Benchmarking core competence in China's firms [J]. *J of Industrial Engineering/Engineering Management*, 2002, 16(4): 22-27.)