

基于马田系统的区间 Choquet 模糊积分多属性决策方法

常志朋¹, 程龙生², 崔立志¹

(1. 安徽工业大学 商学院, 安徽 马鞍山 243002; 2. 南京理工大学 经济管理学院, 南京 210094)

摘要: 为了扩展马田系统在模糊积分多属性决策领域中的应用, 引入区间样本描述统计量, 将传统的实数型马田系统改进为区间型马田系统, 并在此基础上提出一种基于区间数据的模糊测度计算方法. 为了便于集成区间属性值, 定义区间 Choquet 模糊积分算子. 实例分析表明, 所提方法能够解决属性值为区间数据的模糊积分多属性决策问题, 验证了该方法的可行性.

关键词: 马田系统; 模糊测度; 区间数; Choquet 模糊积分

中图分类号: C934

文献标志码: A

Interval Choquet fuzzy integral multi-attribute decision making method based on Mahalanobis-Taguchi system

CHANG Zhi-peng¹, CHENG Long-sheng², CUI Li-zhi¹

(1. School of Business, Anhui University of Technology, Maanshan 243002, China; 2. School of Economics and Management, Nanjing University of Science & Technology, Nanjing 210094, China. Correspondent: CHANG Zhi-peng, E-mail: changzp@126.com)

Abstract: To expand the application of the Mahalanobis-Taguchi system (MTS) in fuzzy integral multi-attribute decision making, the descriptive statistics of the interval sample is introduced into the MTS, and the interval Mahalanobis-Taguchi system (IMTS) is established. An identification method of fuzzy measures by IMTS is provided. The interval fuzzy Choquet operator is defined to integrate interval data. Finally, an illustrative example is given to demonstrate that the proposed method is feasible and can solve the fuzzy integral multi-attribute decision making problems based on interval data.

Keywords: Mahalanobis-Taguchi system; fuzzy measure; interval number; Choquet fuzzy integral

0 引言

马田系统 (MTS)^[1-2]是在 20 世纪 90 年代初由日本著名质量工程学家田口玄一博士在质量工程学基础上发展起来的一种模式识别技术, 该技术广泛应用于质量检测^[3-4]、疾病诊断^[5]和风险预测^[6]等领域. 为了拓展 MTS 的应用领域, 文献 [7-8] 成功地将 MTS 引入到模糊积分多属性决策领域中, 提出了两种基于 MTS 的模糊测度^[9]计算方法, 并结合 Choquet 模糊积分算子^[10]构建了两种基于 MTS 的模糊积分多属性决策方法. 然而, 文献 [7-8] 提出的两种模糊积分多属性决策方法主要针对单点的实数型数据, 并且模糊测度主要通过借助 ϕ_s 转换函数^[11]间接利用 MTS 来确定, 限制了 MTS 在模糊积分多属性决策领域中的应用.

为了进一步扩展 MTS 在模糊积分多属性决策领

域中的应用, 本文利用区间样本描述统计量对传统的实数型 MTS 进行改进, 定义了区间马氏距离和 3 种区间信噪比, 使其能够处理区间型数据, 并在此基础上提出一种基于区间数据的模糊测度计算方法; 另外, 针对传统 Choquet 模糊积分算子只能处理实数型属性值的问题, 本文定义了区间 Choquet 模糊积分算子, 使其能够处理区间型属性值数据.

1 马田系统和区间数理论

1.1 马田系统理论

MTS 是田口玄一博士通过整合田口方法和马氏距离创立的一种模式识别技术. MTS 的基本理论参见文献 [1-2], 这里不再赘述. 下面详细介绍 MTS 的 3 个关键工具: 马氏距离、信噪比和正交表.

1) 马氏距离 (MD) 是由印度统计学家马哈拉诺

收稿日期: 2014-09-26; 修回日期: 2015-01-24.

基金项目: 国家自然科学基金项目(71271114, 71303004); 教育部人文社会科学青年基金项目(12YJK630005).

作者简介: 常志朋(1978-), 男, 副教授, 博士, 从事模式识别、优化算法、多属性决策、管理综合评价等研究; 程龙生(1964-), 男, 教授, 博士生导师, 从事管理综合评价、多准则决策等研究.

比斯提出的一种协方差距离。在 MTS 中, 如果样本容量为 n , 特征变量数为 p , 则第 k 个样品的马氏距离定义为

$$MD_k = \sqrt{\frac{Z_k R^{-1} Z_k^T}{p}}, \quad k = 1, 2, \dots, n. \quad (1)$$

其中: $Z_k = [z_{k1}, z_{k2}, \dots, z_{kp}]$, R 为相关系数矩阵。

2) 信噪比 (SNR) 在 MTS 中用来评估产品性能。由于产品性能指标的评价标准不同, 田口玄一定义了多种形式的信噪比, 常用的有望小特性信噪比 (LB)、望大特性信噪比 (HB) 和望目特性信噪比 (NB)^[12], 3 种信噪比的具体公式表示如下:

$$\begin{cases} \eta_q^{LB} = -10 \log_{10} \left(\frac{1}{n} \sum_{k=1}^n MD_k^2 \right), \\ \eta_q^{HB} = -10 \log_{10} \left(\frac{1}{n} \sum_{k=1}^n \frac{1}{MD_k^2} \right), \\ \eta_q^{NB} = -10 \log_{10} \left[\frac{1}{n-1} \left(\sum_{k=1}^n MD_k^2 - n(\overline{MD})^2 \right) \right]. \end{cases}$$

其中: n 为异常样品的个数, q 为第 q 次试验, $\overline{MD} = \frac{1}{n} \sum_{k=1}^n MD_k$. 这里: η_q^{LB} 主要针对产品性能指标连续且不为负, 可取 $0 \sim \infty$ 的任何值, 越接近 0 越好, 0 是理想值, 如电厂的污染及零部件的磨损量等可以通过望小特性信噪比衡量; η_q^{HB} 主要针对产品性能指标连续且不为负, 且越大越好, 如材料的强度及弹簧的寿命等可以通过望大特性信噪比衡量; η_q^{NB} 主要针对产品性能指标连续且不为负, 可以取 $0 \sim \infty$ 的任何值, 它的目标值是非零的有限值, 产品性能指标越接近目标值越好, 如电源电路输出电压目标值为 115 V, 可以通过望目特性信噪比衡量。大量实践验证表明, 信噪比通常是越大越好, 这样有利于分析问题^[13]。

3) 正交表是一套科学安排实验条件的规格化表格, 具有均衡搭配和综合可比的特点, 因此能以少量试验获取较全面的信息。例如, 对于具有 p 个特征变量的试验, 如果进行全面的 2 水平试验, 则需要进行 2^p 次试验, 如果用 $L_q(2^p)$ 正交表筛选, 则只需要进行 q 次试验。由于本文需要测度决策属性集全集的重要程度, 不需要正交表进行筛选属性集的子集。

1.2 区间数理论

定义 1 设 $a = [a^L, a^U]$ 为实数轴上的一个闭区间, 则 a 为一个区间数。其中: $a^L \leq a^U, a^L, a^U \in \mathbf{R}$ 。

定义 2 记 $I(\mathbf{R}) = \{[a^L, a^U] | a^L, a^U \in \mathbf{R}, a^L \leq a^U\}$ 为区间数集。

定义 3 设 $a = [a^L, a^U], b = [b^L, b^U]$ 为任意两个区间数, a 与 b 之间的 Hausdorff 距离^[14]为

$$d_H(a, b) = |m(a) - m(b)| + |s(a) - s(b)|. \quad (2)$$

其中: $m(a) = \frac{a^L + a^U}{2}, s(a) = \frac{a^U - a^L}{2}$ 。

定义 4 设 $a = [a^L, a^U], b = [b^L, b^U]$ 为任意两个区间数, 且记 $l_a = a^U - a^L, l_b = b^U - b^L$, 则称

$$\rho(a \geq b) = \max \left\{ 1 - \max \left(\frac{b^U - a^L}{l_a - l_b}, 0 \right), 0 \right\} \quad (3)$$

为 $a \geq b$ 的可能度^[15]。

2 基于马田系统的区间属性模糊测度确定方法

Sugeno^[16]于 1974 年首次提出了模糊测度的概念。模糊测度作为经典可加测度的拓展, 是一种可以将属性间相关性或交互性列入考虑的一类集函数, 具有很强的表示能力。

定义 5 设 $X = \{x_1, x_2, \dots, x_p\}$ 为有限属性集, $P(X)$ 为 X 的幂集, $g: P(X) \rightarrow [0, 1]$ 为一组函数, 具有如下性质:

$$1) g(\phi) = 0, g(X) = 1;$$

$$2) \forall E, F \in P(X), \text{若 } E \subseteq F, \text{则有 } g(E) \leq g(F).$$

对于 $\forall E, F \in P(X), E \cap F = \phi$, 如果 $g(E \cup F) < g(E) + g(F)$, 则 E 与 F 之间存在消极的交互关系; 如果 $g(E \cup F) > g(E) + g(F)$, 则 E 与 F 之间存在积极的交互关系; 如果 $g(E \cup F) = g(E) + g(F)$, 则 E 与 F 之间是相互独立的^[9]。

MTS 的一个重要功能是可以利用基于马氏距离的信噪比, 测度幂集 $P(X)$ 在分类过程中的重要程度, 即任意子属性集对 MTS 能够正确判断类别的贡献。由于马氏距离是一种协方差距离, 当属性间存在相关性或交互作用时, MTS 能够有效、客观地测度子属性集的重要程度。但是, 传统的 MTS 只能测度单点实数型属性幂集的重要程度, 而在实际应用中, 由于观测误差、信息不完备等原因, 经常需要用区间数据来表示属性值。为此, 本文利用区间样本描述统计量, 对传统的 MTS 进行改进, 使其能够处理区间属性值, 并在此基础上提出一种基于区间数据的模糊测度确定方法, 具体步骤如下。

设 $X = \{x_1, x_2, \dots, x_p\}$ 为 p 个用于分类的属性, $A = \{a_t | t = 1, 2, \dots, m\}, B = \{b_k | k = 1, 2, \dots, n\}$ 分别为正常和异常两类区间样本, 属性 x_j 在样品 a_t 和 b_k 上的区间属性值分别为 $[a_{tj}^L, a_{tj}^U]$ 和 $[b_{kj}^L, b_{kj}^U]$, 它们构成的区间样本矩阵分别为 $[A] = [[a_{tj}^L, a_{tj}^U]]_{m \times p}$ 和 $[B] = [[b_{kj}^L, b_{kj}^U]]_{n \times p}$ 。

Step 1 计算区间样本矩阵 $[A]$ 的均值、标准差和相关系数矩阵。

2000 年, Lynne 等^[17]从均匀分布的角度推导出区间样本均值和标准差的计算公式

$$\bar{X}_j = \frac{1}{2m} \sum_{t=1}^m (a_{tj}^U + a_{tj}^L), \quad (4)$$

$$S_j^2 = \frac{1}{3m} \sum_{t=1}^m [(a_{tj}^U)^2 + a_{tj}^U a_{tj}^L + (a_{tj}^L)^2] -$$

$$\frac{1}{4m^2} \left[\sum_{t=1}^m (a_{tj}^L + a_{tj}^U) \right]^2, \quad (5)$$

$$S_{ij} = \frac{1}{4m} \sum_{t=1}^m (a_{ti}^L + a_{ti}^U)(a_{tj}^L + a_{tj}^U) - \frac{1}{4m^2} \left[\sum_{t=1}^m (a_{ti}^L + a_{ti}^U) \right] \left[\sum_{t=1}^m (a_{tj}^L + a_{tj}^U) \right]. \quad (6)$$

由上述区间样本描述统计量可以得到相关系数

$$r_{ij} = \frac{S_{ij}}{S_i S_j}, \quad i, j = 1, 2, \dots, p, \quad (7)$$

进而得到相关系数矩阵 $R = [r_{ij}]_{p \times p}$.

Step 2 利用式(4)和(5)对区间样本矩阵 $[B]$ 进行标准化

$$[z_{kj}^L, z_{kj}^U] = \left[\frac{b_{kj}^L - \bar{X}_j}{S_j}, \frac{b_{kj}^U - \bar{X}_j}{S_j} \right], \quad (8)$$

可得标准化区间样本矩阵 $[Z] = [[z_{kj}^L, z_{kj}^U]]_{n \times p}$.

Step 3 根据幂集 $P(X)$ 计算异常样本的区间马氏距离.

定义 6 设 $X = \{x_1, x_2, \dots, x_p\}$ 为有限属性集, 记 $Y = \{Y_1, Y_2, \dots, Y_{2^p}\}$ 为 $P(X)$ 上的 2^p 个子属性集, 则利用子属性集 $Y_q (q = 1, 2, \dots, 2^p)$ 计算的第 k 个异常样本区间马氏距离为

$$[\text{MD}_{kq}^L, \text{MD}_{kq}^U] = \min/\max \left[\sqrt{\frac{Z_{kq}^L R_q^{-1} (Z_{kq}^L)^T}{|Y_q|}}, \sqrt{\frac{Z_{kq}^U R_q^{-1} (Z_{kq}^U)^T}{|Y_q|}} \right].$$

其中: MD_{kq}^L 和 MD_{kq}^U 分别为上式的最小值和最大值; $Z_{kq}^L = [\{z_{kj}^L | j \in Y_q\}]$, $Z_{kq}^U = [\{z_{kj}^U | j \in Y_q\}]$; R_q 为利用属性集 $Y_q (q = 1, 2, \dots, 2^p)$ 计算得到的相关系数矩阵, 当 R_q 奇异时, R_q^{-1} 可用 R_q 的伪逆 R_q^+ 来计算; $|Y_q|$ 为 Y_q 中属性的个数.

例如, 设 $X = \{x_1, x_2, x_3, x_4, x_5\}$, 记 $P(X)$ 上的任意子属性集 $\{x_1, x_3, x_5\}$ 为 Y_{21} , 即 $Y_{21} = \{x_1, x_3, x_5\}$, 利用 Y_{21} 计算的第 k 个区间马氏距离为

$$[\text{MD}_{k,21}^L, \text{MD}_{k,21}^U] = \min/\max \left[\sqrt{\frac{1}{3} \times [z_{k1}^L, z_{k3}^L, z_{k5}^L] R_{21}^{-1} [z_{k1}^L, z_{k3}^L, z_{k5}^L]^T}, \sqrt{\frac{1}{3} \times [z_{k1}^U, z_{k3}^U, z_{k5}^U] R_{21}^{-1} [z_{k1}^U, z_{k3}^U, z_{k5}^U]^T} \right],$$

其中

$$R_{21} = \begin{bmatrix} 1 & r_{13} & r_{15} \\ r_{13} & 1 & r_{35} \\ r_{15} & r_{35} & 1 \end{bmatrix}.$$

因此, 可以根据定义 6 计算得到异常样本的区间马氏距离矩阵 $[\text{MD}] = [[\text{MD}_{kj}^L, \text{MD}_{kj}^U]]_{n \times p}$.

然后, 利用式(9)和(10)对 $[\text{MD}]$ 中区间马氏距离进行规范化

$$d_{kq}^L = \frac{\text{MD}_{kq}^L}{\min_k \min_q (\text{MD}_{kq}^L, \text{MD}_{kq}^U)}, \quad (9)$$

$$d_{kq}^U = \frac{\text{MD}_{kq}^U}{\min_k \min_q (\text{MD}_{kq}^L, \text{MD}_{kq}^U)}, \quad (10)$$

可得规范化异常样本区间马氏距离矩阵

$$[D] = \begin{bmatrix} [d_{11}^L, d_{11}^U] & [d_{12}^L, d_{12}^U] & \cdots & [d_{12^p}^L, d_{12^p}^U] \\ [d_{21}^L, d_{21}^U] & [d_{22}^L, d_{22}^U] & \cdots & [d_{22^p}^L, d_{22^p}^U] \\ \vdots & \vdots & \ddots & \vdots \\ [d_{n1}^L, d_{n1}^U] & [d_{n2}^L, d_{n2}^U] & \cdots & [d_{n2^p}^L, d_{n2^p}^U] \end{bmatrix}.$$

Step 4 计算子属性集 $Y_q (q = 1, 2, \dots, 2^p)$ 的重要程度. 传统信噪比只能处理实数型马氏距离, 本文从均匀分布角度将 3 种特性信噪比改进为区间形式.

定理 1 望小、望大和望目特性区间信噪比为

$$\begin{aligned} \bar{\eta}_{Y_q}^{\text{LB}} &= -10 \log_{10} \left\{ \frac{1}{3n} \sum_{k=1}^n \left[(d_{kq}^U)^2 + d_{kq}^U d_{kq}^L + (d_{kq}^L)^2 \right] \right\}, \\ \bar{\eta}_{Y_q}^{\text{HB}} &= -10 \log_{10} \left(\frac{1}{n} \sum_{k=1}^n \frac{1}{d_{kq}^U d_{kq}^L} \right), \\ \bar{\eta}_{Y_q}^{\text{NB}} &= -10 \log_{10} \left\{ \frac{1}{3(n-1)} \sum_{k=1}^n \left[(d_{kq}^U)^2 + d_{kq}^U d_{kq}^L + (d_{kq}^L)^2 \right] - \frac{1}{4n(n-1)} \left[\sum_{k=1}^n (d_{kq}^U + d_{kq}^L) \right]^2 \right\}. \end{aligned}$$

证明 令 Y_q 为某一特定的随机变量, 且在 $[d_{kq}^L, d_{kq}^U]$ 上服从均匀分布, 即

$$F_{Y_q}(y) = P(Y_q \leq y) = \begin{cases} 0, & y < d_{kq}^L; \\ \frac{y - d_{kq}^L}{d_{kq}^U - d_{kq}^L}, & d_{kq}^L \leq y < d_{kq}^U; \\ 1, & y \geq d_{kq}^U. \end{cases}$$

由此可知, Y_q 的经验分布函数为 n 个服从均匀分布随机变量 $\{Y_{kq} | k = 1, 2, \dots, n\}$ 的综合, 即

$$\begin{aligned} F_{Y_q}(y) &= P(Y_q \leq y) = \frac{1}{n} \{P(Y_{1q} \leq y) + P(Y_{2q} \leq y) + \cdots + P(Y_{nq} \leq y)\} = \\ &= \frac{1}{n} \sum_{k=1}^n F_{Y_{kq}}(y), \end{aligned} \quad (11)$$

进而可以得到 Y_q 的经验分布函数

$$F_{Y_q}(y) = P(Y_q \leq y) = \frac{1}{n} \left\{ \sum_{k=1}^n \left(\frac{y - d_{kq}^L}{d_{kq}^U - d_{kq}^L} \right) + |\{k | y \geq d_{kq}^U\}| \right\}.$$

因此, Y_q 的经验密度函数为

$$f_{Y_q}(y) = F'_{Y_q}(y) = \frac{1}{n} \sum_{k=1}^n \frac{1}{d_{kq}^U - d_{kq}^L}. \quad (12)$$

利用式 (12) 可以推导出如下 3 种区间信噪比:

1) 望小特性区间信噪比为

$$\begin{aligned} \bar{\eta}_{Y_q}^{LB} = & -10\log_{10}\left(\int_{d_{kq}^L}^{d_{kq}^U} y^2 f_{Y_q}(y) dy\right) = \\ & -10\log_{10}\left[\frac{1}{n} \sum_{k=1}^n \left(\int_{d_{kq}^L}^{d_{kq}^U} \frac{y^2}{d_{kq}^U - d_{kq}^L} dy\right)\right] = \\ & -10\log_{10}\left\{\frac{1}{3n} \sum_{k=1}^n \left[\frac{(d_{kq}^U)^3 - (d_{kq}^L)^3}{d_{kq}^U - d_{kq}^L}\right]\right\} = \\ & -10\log_{10}\left\{\frac{1}{3n} \sum_{k=1}^n \left[(d_{kq}^U)^2 + d_{kq}^U d_{kq}^L + (d_{kq}^L)^2\right]\right\}; \end{aligned}$$

2) 望大特性区间信噪比为

$$\begin{aligned} \bar{\eta}_{Y_q}^{HB} = & -10\log_{10}\left(\int_{d_{kq}^L}^{d_{kq}^U} \frac{1}{y^2} f_{Y_q}(y) dy\right) = \\ & -10\log_{10}\left[\frac{1}{n} \sum_{k=1}^n \left(\int_{d_{kq}^L}^{d_{kq}^U} \frac{1}{y^2} \times \frac{1}{d_{kq}^U - d_{kq}^L} dy\right)\right] = \\ & -10\log_{10}\left[\frac{1}{n} \sum_{k=1}^n \left(\frac{d_{kq}^U - d_{kq}^L}{d_{kq}^U d_{kq}^L} \times \frac{1}{d_{kq}^U - d_{kq}^L}\right)\right] = \\ & -10\log_{10}\left(\frac{1}{n} \sum_{k=1}^n \frac{1}{d_{kq}^U d_{kq}^L}\right); \end{aligned}$$

3) 望目特性区间信噪比为

$$\begin{aligned} \bar{\eta}_{Y_q}^{NB} = & -10\log_{10}\left\{\frac{n}{n-1} \left[\frac{1}{3n} \sum_{k=1}^n \left[(d_{kq}^U)^2 + \right. \right. \right. \\ & \left. \left. d_{kq}^U d_{kq}^L + (d_{kq}^L)^2\right] - \left[\frac{1}{2n} \sum_{k=1}^n (d_{kq}^U + d_{kq}^L)\right]^2\right\} = \\ & -10\log_{10}\left\{\frac{1}{3(n-1)} \sum_{k=1}^n \left[(d_{kq}^U)^2 + d_{kq}^U d_{kq}^L + \right. \right. \\ & \left. \left. (d_{kq}^L)^2\right] - \frac{1}{4n(n-1)} \left[\sum_{k=1}^n (d_{kq}^U + d_{kq}^L)\right]^2\right\}. \quad \square \end{aligned}$$

由于在测度属性集 Y_q ($q = 1, 2, \dots, 2^p$) 的重要程度时, 每个属性集在决策过程中发挥的作用越大越好, 本文采用望大特性区间信噪比, 即

$$\eta_{Y_q} = -10\log_{10}\left[\frac{1}{n} \sum_{k=1}^n \frac{1}{d_{kq}^U d_{kq}^L}\right]. \quad (13)$$

Step 5 计算属性集 Y_q ($q = 1, 2, \dots, 2^p$) 的模糊测度.

定理 2 g_{Y_q} 为属性集 Y_q 的模糊测度, 且有

$$g_{Y_q} = \frac{\eta_{Y_q}}{\eta_X}, Y_q \subseteq X, q = 1, 2, \dots, 2^p, \quad (14)$$

其中 η_{Y_q} 和 η_X 分别为利用属性集 Y_q 和 X 计算的望大特性信噪比.

证明 由定义 5 可知, 模糊测度需要满足单调性

和有界性, 下面分别证明 g_{Y_q} 满足单调性和有界性.

1) 单调性. 规范化区间马氏距离的下限和上限

$$d_{kq}^L = \sqrt{Z_{kq}^L R_q^{-1} (Z_{kq}^L)^T}, d_{kq}^U = \sqrt{Z_{kq}^U R_q^{-1} (Z_{kq}^U)^T}$$

为实二次型, 故对于 $\forall V = (v_1, v_2, \dots, v_p) \neq 0$ 和 $\forall U = (u_1, u_2, \dots, u_p) \neq 0$, 都可以通过正交变换 $Z_{kq}^L = CV_{kq}$ 和 $Z_{kq}^U = CU_{kq}$ 将其变换为标准型, 即

$$d_{kq}^L = \sqrt{\sum_{j \in Y_q} \lambda_j v_{kj}^2}, d_{kq}^U = \sqrt{\sum_{j \in Y_q} \lambda_j u_{kj}^2}. \quad (15)$$

其中: $V_{kq} = \{v_{kj} | j \in Y_q\}$, $U_{kq} = \{u_{kj} | j \in Y_q\}$, $\lambda_1, \lambda_2, \dots, \lambda_p$ 为 R_q^{-1} 的特征值.

由相关系数矩阵 R_q 为半正定矩阵可知, R_q^{-1} 也为半正定矩阵, 故 $\lambda_1 \geq 0, \lambda_2 \geq 0, \dots, \lambda_p \geq 0$. 进而由式 (15) 可知, 当 $Y_q \subseteq Y_h \subseteq X$ 时, 有 $d_{kq}^L \leq d_{kh}^L, d_{kq}^U \leq d_{kh}^U, q, h = 1, 2, \dots, 2^p$, 故有 $\eta_{Y_q} \leq \eta_{Y_h}$, 即 $\frac{\eta_{Y_q}}{\eta_X} \leq \frac{\eta_{Y_h}}{\eta_X}$, 由此可得 $g_{Y_q} \leq g_{Y_h}$ 满足单调性.

2) 有界性. 由规范化的区间马氏距离 $d_{kq}^L \geq 1, d_{kq}^U \geq 1$ 可得

$$\eta_{Y_q} = -10\log_{10}\left[\frac{1}{n} \sum_{k=1}^n \frac{1}{d_{kq}^U d_{kq}^L}\right] \geq 0. \quad (16)$$

由式 (14) 可知 $Y_1 = \phi$, 故

$$g_{Y_1} = \frac{\eta_\phi}{\eta_X} = 0.$$

由于 $Y_{2^p} = X$, 故 $g_{Y_{2^p}} = \frac{\eta_X}{\eta_X} = 1$, 即模糊测度满足有界性. \square

3 区间 Choquet 模糊积分

Choquet 模糊积分^[18]算子只能对单点实数型属性值进行集成, 而在实际应用中, 由于数据的不确定性、测量误差、专家思维的模糊性等原因, 属性值往往通过区间值表示, 为了使 Choquet 模糊积分算子能够对区间型属性值进行集成, 本文定义了区间 Choquet 模糊积分.

定义 7 设 $X = \{x_1, x_2, \dots, x_p\}$ 为有限属性集, $P(X)$ 为 X 的幂集, g 为定义在 $(X, P(X))$ 上的模糊测度, $\bar{f}: X \rightarrow I(\mathbf{R})$ 关于 g 的区间 Choquet 模糊积分为

$$\int \bar{f} dg = \sum_{j=1}^p \bar{f}(x_{(j)}) (g(X_{(j)}) - g(X_{(j+1)})). \quad (17)$$

其中: (j) 为按照 $\bar{f}(x_{(1)}) \leq \bar{f}(x_{(2)}) \leq \dots \leq \bar{f}(x_{(p)})$ 进行排序后的下标, $X_{(p+1)} = \phi, X_{(j)} = \{x_{(j)}, x_{(j+1)}, \dots, x_{(p)}\}$.

定理 3 设区间值 $\bar{f}(x_j) = [f^L(x_j), f^U(x_j)]$ ($j = 1, 2, \dots, p$) 满足递增序列 $\bar{f}(x_{(1)}) \leq \bar{f}(x_{(2)}) \leq \dots \leq \bar{f}(x_{(p)})$, 则区间 Choquet 模糊积分为

$$\int \bar{f} dg = \left[\int f^L dg, \int f^U dg \right]. \quad (18)$$

证明

$$\begin{aligned} \int \bar{f} dg &= \sum_{j=1}^p \bar{f}(x_{(j)})(g(X_{(j)}) - g(X_{(j+1)})) = \\ &= \sum_{j=1}^p [f^L(x_j), f^U(x_j)](g(X_{(j)}) - g(X_{(j+1)})) = \\ &= \sum_{j=1}^p [f^L(x_j)(g(X_{(j)}) - g(X_{(j+1)})), \\ &= f^U(x_j)(g(X_{(j)}) - g(X_{(j+1)}))] = \\ &= \left[\sum_{j=1}^p f^L(x_j)(g(X_{(j)}) - g(X_{(j+1)})), \right. \\ &= \left. \sum_{j=1}^p f^U(x_j)(g(X_{(j)}) - g(X_{(j+1)})) \right] = \\ &= \left[\int f^L dg, \int f^U dg \right]. \quad \square \end{aligned}$$

4 实例分析

有 4 种投资方案分别为 I_1 、 I_2 、 I_3 和 I_4 ，假设一个投资者在这 4 种投资方案中考虑 5 个决策属性，分别是房屋面积 (x_1)、设施水平 (x_2)、小区环境 (x_3)、房屋价格 (x_4) 和小区与工作单位的距离 (x_5)。在这 5 个决策属性中， x_1 、 x_2 和 x_3 为效益型， x_4 和 x_5 为成本型，这 4 种投资方案的区间数决策矩阵为

$$[A] = \begin{bmatrix} [90, 120] & [7, 11] & [6, 10] & [2\ 900, 3\ 300] & [9, 12] \\ [70, 100] & [3, 8] & [7, 11] & [3\ 050, 3\ 200] & [6, 10] \\ [65, 70] & [5, 13] & [11, 12] & [3\ 150, 3\ 250] & [7, 11] \\ [60, 80] & [4, 6] & [8, 9] & [3\ 050, 3\ 250] & [8, 14] \end{bmatrix}.$$

Step 1 利用下式构造异常样品区间矩阵:

1) 当属性 x_j 为效益型时, 有

$$\begin{cases} [b_{1j}^L, b_{1j}^U] = [\max_k a_{kj}^L, \max_k a_{kj}^U], \\ [b_{2j}^L, b_{2j}^U] = [\min_k a_{kj}^L, \min_k a_{kj}^U]; \end{cases} \quad (19)$$

2) 当属性 x_j 为成本型时, 有

$$\begin{cases} [b_{1j}^L, b_{1j}^U] = [\min_k a_{kj}^L, \min_k a_{kj}^U], \\ [b_{2j}^L, b_{2j}^U] = [\max_k a_{kj}^L, \max_k a_{kj}^U]. \end{cases} \quad (20)$$

由此可得异常样品区间矩阵

$$[B] = \begin{bmatrix} [90, 120] & [7, 13] & [11, 12] & [2\ 900, 3\ 200] & [6, 10] \\ [60, 70] & [3, 6] & [6, 9] & [3\ 150, 3\ 300] & [9, 14] \end{bmatrix}.$$

Step 2 计算决策矩阵 $[A]$ 的均值、标准差和相关系数矩阵, 并利用式 (8) 对矩阵 $[B]$ 进行标准化, 可得

$$\begin{aligned} Z_1 &= ([0.097\ 3, 0.456\ 9], [-0.017\ 4, 0.786\ 9], [0.187\ 8, \\ &= 0.295\ 1], [-0.077\ 5, 0.017\ 9], [-0.373\ 5, 0.038\ 1]), \\ Z_2 &= ([-0.262\ 2, -0.142\ 4], [-0.553\ 6, -0.151\ 5], \\ &= [-0.348\ 7, -0.026\ 8], [0.002\ 0, 0.049\ 7], \\ &= [-0.064\ 8, 0.449\ 6]). \end{aligned}$$

Step 3 根据属性集 $Y_q (q = 1, 2, \dots, 32)$ 计算 Z_1 和 Z_2 的区间马氏距离, 如表 1 所示.

表 1 Z_1 和 Z_2 的区间马氏距离

q	Y_q	$[MD_{1q}^L, MD_{1q}^U]$	$[MD_{2q}^L, MD_{2q}^U]$
1	{ ϕ }	[0.000 0, 0.000 0]	[0.000 0, 0.000 0]
2	{1}	[0.001 9, 0.041 8]	[0.004 1, 0.013 7]
3	{2}	[0.000 1, 0.123 8]	[0.004 6, 0.061 3]
4	{3}	[0.007 1, 0.017 4]	[0.000 1, 0.024 3]
5	{4}	[0.000 1, 0.001 2]	[0.000 1, 0.000 4]
6	{5}	[0.000 2, 0.027 9]	[0.000 8, 0.040 4]
7	{1, 2}	[0.002 0, 0.163 5]	[0.008 5, 0.074 2]
8	{1, 3}	[0.009 1, 0.060 1]	[0.004 2, 0.038 7]
9	{1, 4}	[0.003 1, 0.041 8]	[0.004 5, 0.013 8]
10	{1, 5}	[0.029 8, 0.042 0]	[0.014 6, 0.044 5]
11	{2, 3}	[0.007 1, 0.140 0]	[0.004 7, 0.084 6]
12	{2, 4}	[0.001 3, 0.123 9]	[0.005 1, 0.061 3]
13	{2, 5}	[0.028 0, 0.124 1]	[0.045 0, 0.062 1]
14	{3, 4}	[0.008 3, 0.017 5]	[0.000 6, 0.024 3]
15	{3, 5}	[0.017 7, 0.034 7]	[0.025 2, 0.040 5]
16	{4, 5}	[0.000 3, 0.029 1]	[0.000 8, 0.040 9]
17	{1, 2, 3}	[0.009 2, 0.180 6]	[0.008 7, 0.098 1]
18	{1, 2, 4}	[0.003 2, 0.163 6]	[0.009 0, 0.074 2]
19	{1, 2, 5}	[0.029 9, 0.163 8]	[0.049 0, 0.075 0]
20	{1, 3, 4}	[0.010 3, 0.060 2]	[0.004 7, 0.038 7]
21	{1, 3, 5}	[0.036 8, 0.060 4]	[0.039 6, 0.044 7]
22	{1, 4, 5}	[0.031 0, 0.042 1]	[0.014 6, 0.045 0]
23	{2, 3, 4}	[0.008 3, 0.140 1]	[0.005 2, 0.084 6]
24	{2, 3, 5}	[0.034 8, 0.140 3]	[0.085 5, 0.045 1]
25	{2, 4, 5}	[0.029 2, 0.124 2]	[0.045 5, 0.062 1]
26	{3, 4, 5}	[0.017 8, 0.035 9]	[0.025 2, 0.041 0]
27	{1, 2, 3, 4}	[0.010 4, 0.180 6]	[0.009 2, 0.098 1]
28	{1, 2, 3, 5}	[0.036 9, 0.180 9]	[0.049 1, 0.099 0]
29	{1, 2, 4, 5}	[0.031 1, 0.163 9]	[0.049 5, 0.075 0]
30	{1, 3, 4, 5}	[0.038 0, 0.060 5]	[0.039 6, 0.045 2]
31	{2, 3, 4, 5}	[0.036 0, 0.140 4]	[0.045 6, 0.085 5]
32	{1, 2, 3, 4, 5}	[0.038 1, 0.180 9]	[0.049 6, 0.099 0]

Step 4 对表 1 的区间马氏距离进行规范化, 利用式 (13) 计算属性集 $Y_q (q = 1, 2, \dots, 32)$ 的重要程度, 利用式 (14) 计算属性集 $Y_q (q = 1, 2, \dots, 32)$ 的模糊测度, 如表 2 所示.

表 2 属性集的模糊测度

q	Y_q	g_{Y_q}	q	Y_q	g_{Y_q}
1	{ ϕ }	0	17	{1, 2, 3}	0.877 3
2	{1}	0.663 1	18	{1, 2, 4}	0.828 1
3	{2}	0.586 1	19	{1, 2, 5}	0.976 5
4	{3}	0.465 1	20	{1, 3, 4}	0.772 6
5	{4}	0.135 1	21	{1, 3, 5}	0.919 4
6	{5}	0.517 2	22	{1, 4, 5}	0.858 1
7	{1, 2}	0.804 7	23	{2, 3, 4}	0.834 4
8	{1, 3}	0.763 9	24	{2, 3, 5}	0.978 4
9	{1, 4}	0.681 4	25	{2, 4, 5}	0.955 4
10	{1, 5}	0.856 5	26	{3, 4, 5}	0.850 5
11	{2, 3}	0.825 6	27	{1, 2, 3, 4}	0.883 2
12	{2, 4}	0.751 5	28	{1, 2, 3, 5}	0.998 5
13	{2, 5}	0.953 5	29	{1, 2, 4, 5}	0.978 2
14	{3, 4}	0.594 5	30	{1, 3, 4, 5}	0.921 0
15	{3, 5}	0.848 3	31	{2, 3, 4, 5}	0.980 1
16	{4, 5}	0.545 2	32	{1, 2, 3, 4, 5}	1.000 0

Step 5 计算各决策方案区间 Choquet 模糊积分综合评价价值.

首先, 对区间决策矩阵 $[A]$ 进行规范化处理, 具体处理方法如下:

1) 当属性 x_j 为效益型时, 有

$$h_{tj}^L = \frac{a_{tj}^L}{\sum_{t=1}^m a_{tj}^U}, \quad h_{tj}^U = \frac{a_{tj}^U}{\sum_{t=1}^m a_{tj}^L};$$

2) 当属性 x_j 为成本型时, 有

$$h_{tj}^L = \frac{1}{\sum_{t=1}^m \frac{1}{a_{tj}^U}}, \quad h_{tj}^U = \frac{1}{\sum_{t=1}^m \frac{1}{a_{tj}^L}}.$$

然后, 得到规范化的区间决策矩阵

$$[H] = \begin{bmatrix} [0.243\ 2, 0.421\ 1] & [0.184\ 2, 0.578\ 9] & [0.142\ 9, \\ [0.189\ 2, 0.350\ 9] & [0.078\ 9, 0.421\ 1] & [0.166\ 7, \\ [0.175\ 7, 0.245\ 6] & [0.131\ 6, 0.684\ 2] & [0.261\ 9, \\ [0.162\ 1, 0.280\ 7] & [0.105\ 3, 0.315\ 8] & [0.190\ 5, \\ & 0.312\ 5] & [0.229\ 9, 0.280\ 1] & [0.152\ 7, 0.321\ 4] \\ & 0.343\ 8] & [0.231\ 7, 0.266\ 3] & [0.183\ 3, 0.482\ 2] \\ & 0.375\ 0] & [0.233\ 5, 0.257\ 9] & [0.166\ 6, 0.413\ 3] \\ & 0.281\ 2] & [0.233\ 5, 0.266\ 3] & [0.130\ 9, 0.361\ 6] \end{bmatrix}.$$

下面以方案 I_1 为例, 说明区间 Choquet 模糊积分综合评价值的计算过程.

首先, 利用式 (3) 确定方案 I_1 的区间属性值排序

$$\bar{f}(x_3) \leq \bar{f}(x_5) \leq \bar{f}(x_4) \leq \bar{f}(x_1) \leq \bar{f}(x_2).$$

利用式 (18) 计算区间 Choquet 模糊积分的上限和下限

$$\begin{aligned} \int f_1^L dg &= f^L(x_3)(1 - g_{5\ 412}) + f^L(x_5)(g_{5\ 412} - g_{412}) + \\ &f^L(x_4)(g_{412} - g_{12}) + f^L(x_1)(g_{12} - g_2) + f^L(x_2)g_2 = \\ &0.193\ 4, \\ \int f_1^U dg &= f^U(x_3)(1 - g_{5\ 412}) + f^U(x_5)(g_{5\ 412} - g_{412}) + \\ &f^U(x_4)(g_{412} - g_{12}) + f^U(x_1)(g_{12} - g_2) + f^U(x_2)g_2 = \\ &0.498\ 8. \end{aligned}$$

然后, 得到方案 I_1 的区间 Choquet 模糊积分综合评价价值 $\int \bar{f}_1 dg = [0.193\ 4, 0.498\ 8]$, 同理可分别得到方案 I_2 、 I_3 和 I_4 的区间 Choquet 模糊积分综合评价价值

$$\begin{aligned} \int \bar{f}_2 dg &= [0.171\ 6, 0.423\ 8], \\ \int \bar{f}_3 dg &= [0.198\ 6, 0.489\ 5], \\ \int \bar{f}_4 dg &= [0.163\ 1, 0.314\ 9]. \end{aligned}$$

最后, 利用式 (3) 进行排序, 可得 $I_1 \succ I_3 \succ I_2 \succ I_4$, 故方案 I_1 为最优方案.

为了验证本文方法的有效性, 下面分别与不同的决策方法进行比较分析.

1) 与文献 [19] 方法进行比较分析. 由于该方法是一种基于可加测度 (权重) 的决策方法, 为了便于比较, 本文利用下式^[15]计算各属性的客观权重:

$$w_j = \frac{\sum_{t=1}^m \sum_{k=1}^m d_H(h_{tj}, h_{kj})}{\sum_{j=1}^p \sum_{t=1}^m \sum_{k=1}^m d_H(h_{tj}, h_{kj})}, \quad j = 1, 2, \dots, p.$$

通过计算可以得到各属性的权重分别为: $w_1 = 0.204\ 9$, $w_2 = 0.433\ 7$, $w_3 = 0.154\ 4$, $w_4 = 0.023\ 5$, $w_5 = 0.183\ 5$.

本文方法和文献 [19] 方法的可能度排序值如表 3 所示.

表 3 可能度排序值

方法	I_1	I_2	I_3	I_4
文献 [19]	0.273 9	0.242 9	0.277 4	0.205 8
本文方法	0.276 8	0.248 9	0.276 3	0.198 0

从表 3 可以看出, 两种方法的最优方案与次优方案排序值比较接近, 但是本文方法得到的方案排序为 $I_1 \succ I_3 \succ I_2 \succ I_4$, 文献 [19] 方法得到的方案排序为 $I_3 \succ I_1 \succ I_2 \succ I_4$. 分析其原因, 主要因为本文方法一方面采用了非可加测度 (模糊测度) 描述属性间客观存在的交互性或关联性, 另一方面采用非线性 Choquet 模糊积分算子集成属性值信息, 从而使本文方法更符合实际决策问题.

2) 与文献 [20] 方法进行比较分析. 文献 [20] 方法是一种基于属性权重和属性间交互度的模糊测度确定方法, 通过与该方法进行比较来验证本文提出的模糊测度确定方法的合理性和有效性. 基于上述权重, 首先利用文献 [20] 中 ϕ_s 转换函数, 分别计算交互度为 -0.99 、 -0.50 、 1 、 10 、 50 和 100 时的模糊测度; 然后计算各方案的区间 Choquet 模糊积分综合评价价值; 最后利用可能度公式计算各方案的排序值, 所得排序值如表 4 所示.

表 4 不同交互度下各决策方案的可能度排序值

λ	Y_1	Y_2	Y_3	Y_4	排序
-0.99	0.2824	0.2430	0.2761	0.1984	$I_1 \succ I_3 \succ I_2 \succ I_4$
-0.50	0.2821	0.2392	0.2785	0.2002	$I_1 \succ I_3 \succ I_2 \succ I_4$
0	0.2840	0.2400	0.2715	0.2046	$I_1 \succ I_3 \succ I_2 \succ I_4$
1	0.2883	0.2415	0.2614	0.2088	$I_1 \succ I_3 \succ I_2 \succ I_4$
10	0.2933	0.2436	0.2536	0.2095	$I_1 \succ I_3 \succ I_2 \succ I_4$
100	0.2953	0.2448	0.2461	0.2137	$I_1 \succ I_3 \succ I_2 \succ I_4$

从表 4 可以看出, 当交互度 λ 在 $-0.99 \sim 100$ 之间时, 4 个决策方案的排序一直稳定为 $I_1 \succ I_3 \succ I_2 \succ I_4$.

I_4 , 与本文的排序结果完全一致, 进一步验证了本文方法的有效性.

5 结 论

为了将广泛应用于质量工程学领域中的马田系统引入到区间模糊积分多属性决策领域中, 本文主要做了以下几个方面工作: 1) 利用区间样本描述统计量, 构建了区间马氏距离及其标准化方法; 2) 从均匀分布角度推导出区间信噪比公式; 3) 证明了区间信噪比所计算的属性幂集重要程度为模糊测度; 4) 定义了区间 Choquet 模糊积分算子, 拓展了 Choquet 模糊积分算子的应用领域. 从实例的推演过程可以看出, 本文方法易于操作、原理简单, 但是本方法是一种客观决策方法, 决策结果可能会偏离决策者的主观偏好. 下一步研究的重点将会引入决策者的主观偏好, 使决策结果更加科学、合理.

参考文献(References)

- [1] Taguchi G, Jugulum R. The Mahalanobis-Taguchi strategy: A pattern technology system[M]. New York: John Wiley & Sons, 2002: 19-57.
- [2] Taguchi G, Chowdhury S, Wu Y. The Mahalanobis-Taguchi system[M]. New York: McGraw-Hill, 2001: 23-60.
- [3] Edgar R, Luis A. Binary ant colony optimization applied to variable screening in the Mahalanobis-Taguchi system[J]. Expert Systems with Applications, 2013, 40(2): 634-637.
- [4] Soylemezoglu A, Jagannathan S. Mahalanobis-Taguchi system as a multi-sensor based decision making prognostics tool for centrifugal pump failures[J]. IEEE Trans on Reliability, 2011, 60(4): 864-878.
- [5] Wang Z P, Liu C. Fault diagnosis and health assessment for bearings using the Mahalanobis-Taguchi system based on EMD-SVD[J]. Trans of the Institute of Measurement and Control, 2013, 35(6): 798-807.
- [6] Lee Y C, Teng H L. Predicting the financial crisis by Mahalanobis-Taguchi system—Examples of Taiwan's electronic sector[J]. Expert Systems with Applications, 2009, 36(4): 7469-7478.
- [7] 常志朋, 程龙生. 基于马田系统和 ϕ_s 转换的模糊积分多属性决策方法[J]. 系统工程与电子技术, 2013, 35(8): 1702-1710.
(Chang Z P, Cheng L S. Fuzzy integral multi-attribute decision making method based on Mahalanobis-Taguchi system and ϕ_s tranformation[J]. Systems Engineering and Electronic, 2013, 35(8): 1702-1710.)
- [8] 常志朋, 程龙生. 基于施密特正交马田系统和 ϕ_s 转换的灰模糊积分关联度决策模型[J]. 控制与决策, 2014, 29(7): 1257-1261.
(Chang Z P, Cheng L S. Grey fuzzy integral correlation degree decision model based on Mahalanobis-Taguchi Gram-Schmidt and ϕ_s tranformation[J]. Control and Decision, 2014, 29(7): 1257-1261.)
- [9] Sugeno M. Fuzzy measure and fuzzy integrals, a survey, fuzzy automata and decision processes[M]. New York: North-Holland, 1977: 89-102.
- [10] Ishii K, Sugeno M. A model of humanevaluationprocess using fuzzy measure[J]. Int J of Man-Machine Studies, 1985, 22(1): 19-38.
- [11] Tukamoto Y. A measure theoretic approach to evaluation of fuzzy set defined on probability space[J]. J of Fuzzy Mathematics, 1982, 2(3): 89-98.
- [12] 王万中. 试验的设计与分析[M]. 北京: 高等教育出版社, 2004: 192-194.
(Wang W Z. The design of the test and analysis[M]. Beijing: Higher Education Press, 2004: 192-194.)
- [13] 陈魁. 试验设计与分析[M]. 北京: 清华大学出版社, 1996: 203-208.
(Chen K. The design of the test and analysis[M]. Beijing: Tsinghua University Press, 1996: 203-208.)
- [14] Francisco A T, Renata S, Maric C, et al. Adaptive Hausdorff distances and dynamic clustering of symbolic interval data[J]. Pattern Recognition Letters, 2006, 27(3): 167-179.
- [15] 徐泽水, 孙在东. 一类不确定型多属性决策问题的排序方法[J]. 管理科学学报, 2002, 5(3): 35-39.
(Xu Z S, Sun Z D. Priority method for a kind of multi-attribute decision-making problems[J]. J of Management Sciences in China, 2002, 5(3): 35-39.)
- [16] Sugeno M. Theory of fuzzy integral and its applications[D]. Tokyo: Tokyo Institute of Technology, 1974.
- [17] Lynne B, Edwin D. Analysis of symbolic data[M]. New York: Springer-Verlag, 2000: 78-80.
- [18] Murofushi T, Sugeno M. An interpretation of fuzzy measure and the Choquet integral as an integral with respect to a fuzzy measure[J]. Fuzzy Sets and Systems, 1989, 29(2): 201-227.
- [19] 徐泽水, 达庆利. 区间数排序的可能度法及其应用[J]. 系统工程学报, 2003, 18(1): 67-70.
(Xu Z S, Da Q L. Possibility degree method for ranking interval numbers and its application[J]. J of System Engineering, 2003, 18(1): 67-70.)
- [20] Takahagi E. A fuzzy measure identification method by diamond pairwise comparisons and ϕ_s transformation[J]. Fuzzy Optimization and Decision Making, 2008, 7(3): 219-232.