

误差分析理论在区间数多属性 决策问题中的应用^{*}

樊治平^① 郭亚军

(东北大学工商管理学院,沈阳 110006)

摘 要 基于误差分析理论,针对不确定性区间数多属性决策问题,给出了一种误差分析方法,并将该方法应用于 TOPSIS算法(逼近于理想解的排序方法)。

关键词 多属性决策,区间数,误差分布, TOPSIS算法。

分类号 N 94

多属性决策(MADM: multiple attribute decision making)普遍存在于工程系统和社会经济系统之中,它是多目标决策分析的一个重要内容。值得指出,由于人类对事物本质的认识,常常受到某种局限的影响,所以在实际问题的决策分析中,经常会出现决策信息具有不确定性或模糊性。这样,对于一类具有不确定性区间数的 MADM 问题的研究就有着重要的理论意义和应用背景。目前,对于确定性 MADM 问题的研究,国内外许多学者已经做了大量的工作,但对于不确定性区间数问题的研究还不多见。文献[1,2]分别给出了区间数 MADM 问题的分析方法;文献[3~6]针对具有区间数的判断矩阵权重计算问题给出了分析方法。这些具有区间数的决策问题的处理方法,大多是直接采用区间数的运算法则,如一般限于区间数的加、减、乘、除法的运算,对于复杂的分析算法,有一定的困难。本文则是基于误差分析理论,完整地给出一种不确定性区间数 MADM 问题的分析方法——误差分析法,它是采用误差分布技术测定决策方案综合评价值的复合误差。这种分析方法的最大优点在于遵循了原有确定性 MADM 问题的分析方法,尤其针对 MADM 问题的算法具有复杂性(如非线性)时,是非常有效的。最后,将误差分析应用于解决区间数 MADM 问题的 TOPSIS 算法,并给出了一个计算实例。

1 误差传递公式

考虑随机误差的传递问题。设直接测量值为 x_1, x_2, \dots, x_m , 间接测量值为 y , 它们之间有连续可微的函数关系: $y = f(x_1, x_2, \dots, x_m)$ 。设 x_1, x_2, \dots, x_m 的随机误差分别为 W_1, W_2, \dots, W_m , 相应的均方根差为 $e_{x_1}, e_{x_2}, \dots, e_{x_m}$; y 的随机误差为 W_y , 相应的均方根差为 e_y 。通常,均方根差是表示随机误差分散度的量度。

* 1996-06-11收到。 ① 男, 36, 博士, 副教授。

国家自然科学基金资助项目(编号: 79600006)及辽宁省自然科学基金资助项目(编号: 962163)。

误差传递定理 随机误差传递的一般关系式为

$$e_y^2 = \sum_{i=1}^m \left(\frac{\partial f}{\partial x_i} \right)^2 e_{x_i}^2 + 2 \sum_{\substack{i < j \\ i, j \leq m}} \left(\frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} d_{ij} e_{x_i} e_{x_j} \right) \quad (1)$$

式中, d_{ij} 为相关系数. 若各测量值 x_i 的随机误差 W_i ($i = 1, 2, \dots, m$) 相互独立, 相关系数 $d_{ij} = 0$, 则式 (1) 简化为

$$e_y^2 = \sum_{i=1}^m \left(\frac{\partial f}{\partial x_i} \right)^2 e_{x_i}^2 \quad (2)$$

误差传递定理的详细证明可参见文献 [7], 这里不再详述, 误差传递公式 (1) 是根据均方根差和相关系数的定义并基于台劳公式展开后舍去了高阶项而得到.

值得指出, 在使用误差传递公式 (1) 或 (2) 时, 需要采用均方根差表征随机误差, 但在实际问题中, 用均方根差进行评定是比较困难的. 所以在综合误差的合成时, 可以采用极限测量误差 (即误差的最大估计) 来评定.

设具有某种关于 0 的对称分布的随机误差 W 的极限误差为 W_m , 对给定的置信概率, 有正实数 t_2^p , 使得:

$$P \left\{ \frac{|W|}{e} \leq \frac{W_m}{e} = t_2^p \right\} = p \quad (3)$$

式中, t_2^p 称为置信系数, 通常可查表得到 [8], 在式 (1) 中, 令

$$e_y = W_{y \lim} / t_y, \quad e_{x_i} = W_{x_i \lim} / t_{x_i}, \quad i = 1, 2, \dots, m \quad (4)$$

则有

$$W_{y \lim}^2 = t_y^2 \left[\sum_{i=1}^m \left(\frac{\partial f}{\partial x_i} \right)^2 \left(\frac{W_{x_i \lim}}{t_{x_i}} \right)^2 + 2 \sum_{\substack{i < j \\ i, j \leq m}} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} d_{ij} \left(\frac{W_{x_i \lim}}{t_{x_i}} \right) \left(\frac{W_{x_j \lim}}{t_{x_j}} \right) \right] \quad (5)$$

如果各单项随机误差服从于正态分布, 则由概率论可知总随机误差必将服从正态分布. 当取同一置信概率 p 时, 可以认为置信系数 $t = t$, 于是在概率意义下, 将式 (5) 可相应地写成:

$$W_{y \lim}^2 = \sum_{i=1}^m \left(\frac{\partial f}{\partial x_i} \right)^2 W_{x_i \lim}^2 + 2 \sum_{\substack{i < j \\ i, j \leq m}} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} d_{ij} W_{x_i \lim} W_{x_j \lim} \quad (6)$$

特别地, 当 x_i ($i = 1, 2, \dots, m$) 相互独立时, 有

$$W_{y \lim}^2 = \sum_{i=1}^m \left(\frac{\partial f}{\partial x_i} \right)^2 W_{x_i \lim}^2 \quad (7)$$

2 决策分析方法

2.1 确定性 MADM 问题的描述

确定性 MADM 问题可由四元组描述, 即 $\langle A, Y, W_A, D \rangle$. 其中, A 为 MADM 问题的指标 (或属性) 体系; Y 为决策 (备选) 方案在指标体系下的属性值; W_A 为决策人对于指标体系的偏好结构; D 为决策方案的综合评价.

关于确定性 MADM 问题的分析方法, 目前比较丰富 [9]. 一般来说, 确定性 MADM 问题的分析过程是根据已有的决策信息 $\langle A, Y, W_A \rangle$, 基于某种分析思路, 采取适应的算法, 可以得到决策方案的综合评价 Y , 最后根据各方案的综合评价, 遵循某种最优决策规则, 进行所有决策方案的排序.

2.2 不确定性区间数 MADM 问题的描述

定义 1 设 R 为实数域, 称闭区间 $[x', x'']$ 为闭区间数, 用 \tilde{x} 表示, 其中, $x', x'' \in R, x' \leq x''$.

定义 2 区间数 $\tilde{x} = [x', x'']$ 也可表示成误差分布形式, 即 $\tilde{x} = x \pm \Delta x$. 其中, $x = (x'' + x') / 2$, $\Delta x = (x'' - x') / 2$ 称 x 和 Δx 分别为区间数 \tilde{x} 的区间中间值和 x 的误差最大估计.

不确定性区间数 MADM 可描述为 $\langle A, Y, W_A, D \rangle$ 或 $\langle A, Y \pm \Delta Y, W_A \pm \Delta W_A, D \pm \Delta D \rangle$. 其中, Y, W_A, D 均为区间数, 其意义相同于确定性 MADM 问题的描述.

2.3 不确定性区间数 MADM 问题的分析方法

不确定性区间数 MADM 问题的决策分析思想是: 首先将不确定性区间数 MADM 问题 $\langle A, Y, W_A, D \rangle$ 分解成确定性 MADM 问题 $\langle A, Y, W_A, D \rangle$ 和基于误差分布的 MADM 问题 $\langle A, \Delta Y, \Delta W_A, \Delta D \rangle$. 然后对于 $\langle A, Y, W_A, D \rangle$, 根据确定性 MADM 问题的某种分析方法 (即算法), 最终将得出决策方案的综合评价值 D , 即为不确定性区间数 MADM 问题的决策方案评价值的区间中点; 对于 $\langle A, \Delta Y, \Delta W_A, \Delta D \rangle$ 仍根据确定性 MADM 问题的分析方法 (即函数关系), 利用前述随机误差传递公式, 最终将得出决策方案综合评价值的误差分布 ΔD .

在区间数 MADM 问题中, ΔY 和 ΔW_A 是对应于区间中点值 Y 和 W_A 的误差值最大估计, 即为极限误差. 在大多数实际问题中, Y 和 W_A 的误差可以认为其服从正态分布. 当 Y 和 W_A 的数量有限且误差范围不是很大时, D 的误差 (即总的合成误差) 也将服从正态分布. 另外, 所有的 Y 和 W_A 的误差针对实际问题而言, 可以认为是相互独立的 (即使不独立, 也认为是弱相关的). 因此, 在基于误差分布的 MADM 问题 $\langle A, \Delta Y, \Delta W_A, \Delta D \rangle$ 的分析中, 可以有效地使用误差传递公式 (7).

区间数 MADM 问题的决策方案排序方法, 取决于计算出来的决策方案综合评价值 $D = D \pm \Delta D$ 或 $D = [D - \Delta D, D + \Delta D]$. 如果考虑到方案的综合评价值是愈大, 相应方案愈优, 下面给出关于决策规则的定义.

定义 3 若 $\tilde{D}_i = \tilde{D}_k$, 即 $D_i \pm \Delta D_i = D_k \pm \Delta D_k$, 则 $P_i \sim P_k$, 即决策方案 P_i 等价于决策方案 P_k .

定义 4 若 $\tilde{D}_i \cap \tilde{D}_k = \emptyset$ 且 $(D_i - \Delta D_i) \geq (D_k + \Delta D_k)$, 则 $P_i \succ P_k$, 即决策方案 P_i 优于决策方案 P_k .

由定义 3 和定义 4 可知:

若 $\tilde{D}_i = \tilde{D}_k, \tilde{D}_k = \tilde{D}_j$, 则 $P_i \sim P_k, P_k \sim P_j$, 故有 $P_i \sim P_j$.

若 $(D_i - \Delta D_i) \geq (D_k + \Delta D_k), (D_k - \Delta D_k) \geq (D_j + \Delta D_j)$, 则 $P_i \succ P_k, P_k \succ P_j$. 故有 $P_i \succ P_j$.

对于 $\tilde{D}_i \cap \tilde{D}_k = \emptyset$ 即两个方案评价值的区间数存在相互交叉部分, 这种情况比较复杂, 严格来讲, 在方案 P_i 和方案 P_k 之间是无法选优的. 如果遇到这种情况, 那么主要取决于决策者的偏好与风险. 实际上, 这是另一个需要深入研究的新问题.

3 应用于 TOPSIS 算法

TOPSIS 算法是一种接近于简单加权法的排序方法, 它是根据不同方案接近决策者所期望理想解的程度来进行方案排序的方法, 也是求解 MADM 问题的一种非常有效的方法.

对于求解确定性 MADM 问题, TOPSIS 算法的步骤可归纳如下^[9].

第一步 设有一多属性决策问题, 其决策矩阵 $Y =$

	属性 1	属性 2	...	属性 m
方案 1	y_{11}	y_{12}	...	y_{1m}
方案 2	y_{21}	y_{22}	...	y_{2m}
\vdots	\vdots	\vdots	\vdots	\vdots
方案 n	y_{n1}	y_{n2}	...	y_{nm}

将矩阵 Y 构

成规范化决策矩阵 Z , 其中 Z 的元素 z_{ij} 为

$$z_{ij} = y_{ij} \sqrt{\frac{1}{\sum_{k=1}^n y_{kj}^2}} \quad i = 1, 2, \dots, n; j = 1, 2, \dots, m \quad (8)$$

第二步 构成加权的规范化决策矩阵 X , 其中 X 的元素 x_{ij} 为

$$x_{ij} = w_j z_{ij} \quad i = 1, 2, \dots, n; j = 1, 2, \dots, m \quad (9)$$

式中, $w_j (j = 1, 2, \dots, m)$ 是第 j 个属性的权重.

第三步 确定理想解 x^+ 和负理想解 x^- :

$$x^+ = \left\{ (\max_i x_{ij} | j \in J), (\min_i x_{ij} | j \in J') | i = 1, 2, \dots, n \right\} = \{x_1^+, x_2^+, \dots, x_m^+\} \quad (10)$$

$$x^- = \left\{ (\min_i x_{ij} | j \in J), (\max_i x_{ij} | j \in J') | i = 1, 2, \dots, n \right\} = \{x_1^-, x_2^-, \dots, x_m^-\} \quad (11)$$

式中, J 是效益型指标的集; J' 是成本型指标的集.

第四步 计算距离. 每个解到理想解和负理想解的距离分别是

$$s_i^+ = \sqrt{\sum_{j=1}^m (x_{ij} - x_j^+)^2} \quad i = 1, 2, \dots, n \quad (12)$$

$$s_i^- = \sqrt{\sum_{j=1}^m (x_{ij} - x_j^-)^2} \quad i = 1, 2, \dots, n \quad (13)$$

第五步 计算每个解对理想解的相对接近度:

$$c_i^+ = s_i^- / (s_i^- + s_i^+), \quad 0 \leq c_i^+ \leq 1, \quad i = 1, 2, \dots, m \quad (14)$$

第六步 排列方案的优先次序. 按 c_i^+ ($i = 1, 2, \dots, m$) 由大到小的顺序排列, 排在前面的方案应优先采用.

对于不确定性区间数 MADM 问题 $\langle A, Y, W_A, D \rangle$, 将分解成 $\langle A, Y, W_A, D \rangle$ 和 $\langle A, \Delta Y, \Delta W_A, \Delta D \rangle$. 求解 $\langle A, Y, W_A, D \rangle$, 可直接采用上述 TOPSIS 算法, 即式 (8) 至式 (14); 求解 $\langle A, \Delta Y, \Delta W_A, \Delta D \rangle$, 将根据 TOPSIS 算法的函数映射关系式 (8), (9), (12), (13), (14), 使用误差传递公式 (7), 得出每步的误差分布算法, 其步骤可归纳如下.

第一步 计算规范化决策矩阵 Z 的误差分布 ΔZ , 其中 ΔZ 的元素 Δz_{ij} 为

$$\Delta z_{ij}^2 = \frac{1}{\left(\sum_{k=1}^n y_{kj}^2\right)^3} \left[\Delta y_{ij}^2 \left(\sum_{\substack{k=1 \\ k \neq i}}^n y_{kj}^2\right)^2 + \frac{1}{4} y_{ij}^2 \left(\sum_{\substack{k=1 \\ k \neq i}}^n y_{kj}^2\right) \right] \quad (15)$$

$$i = 1, 2, \dots, n; j = 1, 2, \dots, m$$

第二步 计算加权规范化决策矩阵 X 的误差分布 ΔX , 其中 ΔX 的元素 Δx_{ij} 为

$$\Delta x_{ij}^2 = z_{ij}^2 \Delta w_j^2 + w_j^2 \Delta z_{ij}^2, \quad i = 1, 2, \dots, n; j = 1, 2, \dots, m \quad (16)$$

第三步 确定理想解 x^+ 和负理想解 x^- 的误差分布 Δx^+ 和 Δx^- .

$$\tilde{x}_{ij} = [x'_{ij}, x''_{ij}] = [x_{ij} - \Delta x_{ij}, x_{ij} + \Delta x_{ij}], \quad i = 1, 2, \dots, n; j = 1, 2, \dots, m \quad (17)$$

$$x^{\prime 4} = \left\{ (\max_i x'_{ij} | j \in J), (\min_i x'_{ij} | j \in J') | i = 1, 2, \dots, n \right\} = \{x_1^{\prime 4}, x_2^{\prime 4}, \dots, x_m^{\prime 4}\} \quad (18)$$

$$x^{\prime \prime 4} = \left\{ (\max_i x''_{ij} | j \in J), (\min_i x''_{ij} | j \in J') | i = 1, 2, \dots, n \right\} = \{x_1^{\prime \prime 4}, x_2^{\prime \prime 4}, \dots, x_m^{\prime \prime 4}\} \quad (19)$$

$$x^{\prime -} = \left\{ (\min_i x'_{ij} | j \in J), (\max_i x'_{ij} | j \in J') | i = 1, 2, \dots, n \right\} = \{x_1^{\prime -}, x_2^{\prime -}, \dots, x_m^{\prime -}\} \quad (20)$$

$$x^{\prime \prime -} = \left\{ (\min_i x''_{ij} | j \in J), (\max_i x''_{ij} | j \in J') | i = 1, 2, \dots, n \right\} = \{x_1^{\prime \prime -}, x_2^{\prime \prime -}, \dots, x_m^{\prime \prime -}\} \quad (21)$$

$$\Delta x^+ = (x^h - x^l) / 2 = \{\Delta x_1^+, \Delta x_2^+, \cdots, \Delta x_m^+\}$$

(22)

$$\Delta x^- = (x^{h-} - x^{l-}) / 2 = \{\Delta x_1^-, \Delta x_2^-, \cdots, \Delta x_m^-\}$$

(23)

第四步 计算距离 s_i^+ 和 s_i^- 的误差分布 Δs_i^+ 和 Δs_i^- .

$$(\Delta s_i^+)^2 = \frac{1}{\sum_{j=1}^m (x_{ij} - x_j^+)^2} \left[\sum_{j=1}^m (x_{ij} - x_j^+)^2 (\Delta x_{ij}^2 + (\Delta x_j^+)^2) \right], \quad i = 1, 2, \cdots, m$$

(24)

$$(\Delta s_i^-)^2 = \frac{1}{\sum_{j=1}^m (x_{ij} - x_j^-)^2} \left[\sum_{j=1}^m (x_{ij} - x_j^-)^2 (\Delta x_{ij}^2 + (\Delta x_j^-)^2) \right], \quad i = 1, 2, \cdots, m$$

(25)

第五步 计算相对接近度 c_i^+ 的误差分布 Δc_i^+ .

$$(\Delta c_i^+)^2 = \frac{1}{(s_i^- + s_i^+)^2} \left[(s_i^+)^2 (\Delta s_i^-)^2 + (s_i^-)^2 (\Delta s_i^+)^2 \right], \quad i = 1, 2, \cdots, m$$

(26)

第六步 略.

4 算 例

考虑一个购买战斗机问题. 决策者根据战斗机的性能和费用, 考虑了 6 项评价指标, 决策矩阵如表 1 所示. 在 6 个指标中, 第 4 个指标 (购买费用) 为成本型指标, 其余 5 个指标均为效益型指标. 第 6 个指标 (灵敏度) 的单位为 10 点制, 即从 1 (最差) 到 10 (最好). 假设决策者对 6 个指标已经有如下优先权重: $w = (0.2 \pm 0.05, 0.1 \pm 0.05, 0.1 \pm 0.05, 0.1 \pm 0.05, 0.2 \pm 0.05, 0.3 \pm 0.05)$.

表 1 决策矩阵

方案 P_i	最大速度 马赫	飞行范围 10 ³ 公里	最大负荷 10 ⁴ 磅	购买费用 10 ⁶ 美元	可靠性 %	灵敏度
P_1	2.0 ± 0.2	1.5 ± 0.3	2.0 ± 0.2	5.5 ± 0.1	85 ± 5	9 ± 1
P_2	2.5 ± 0.2	2.7 ± 0.3	1.8 ± 0.2	6.5 ± 0.1	75 ± 5	5 ± 1
P_3	1.8 ± 2.2	2.0 ± 0.3	2.1 ± 0.2	4.5 ± 0.1	90 ± 5	7 ± 1
P_4	2.2 ± 0.2	1.8 ± 0.3	2.0 ± 0.2	5.0 ± 0.1	85 ± 5	5 ± 1

根据前述算法, 即式 (8)~ (26), 最终得到每个决策方案对理想方案的相对接近度如表 2 所示.

表 2 计算结果

决策方案	区间中点值 C_i^+	误差分布 ΔC_i^+	区间值 $[C_i^+, C_i^-]$
P_1	0.700 086	0.229 030	[0.471 056, 0.929 116]
P_2	0.319 834	0.213 792	[0.106 042, 0.533 626]
P_3	0.480 073	0.262 552	[0.217 521, 0.742 625]
P_4	0.227 866	0.239 235	[0.011 369, 0.467 101]

由表 2 可以看出, 由于 $c_1 > c_4$, 所以决策方案的最优排序应该是 $[(P_1, P_3, P_2), P_4]$, 由于方案 P_1, P_3, P_2 的区间值有相互交叉, 所以在它们之间进行排序是困难的. 到底哪一个方案为最优方案, 可由决策者来决定. 这完全取决于决策者的偏好与风险. 这个例子使人感到有兴趣的

是由于决策矩阵及指标权重上的小小波动(或误差),将会对决策分析产生很大的影响.

5 结 束 语

本文在介绍了误差传递公式后,给出了一种解决不确定性区间数 MADM 问题的误差分析方法.这种误差分析法遵循了原有确定性 MADM 问题的算法,当其算法具有复杂性(如非线性)时,是非常有效的.同时,也应该指出,由于误差传递公式的得出是基于台劳公式展开并线性化,所以在使用误差传递公式计算总的合成误差时,要存在一定的偏差,即使这样,也可以认为它是处理不确定性区间数决策问题(包括群体决策问题)的一种可行方法.本文的研究内容实际上是一个新的课题,应该寻求解决区间数决策问题更有效的新方法.

参考文献

- 1 徐玖平.指标为区间的多指标决策.见:首届中国青年运筹学者大会论文集.北京:冶金工业出版社,1994. 301~ 305
- 2 张吉军.指标为区间数的多指标决策方法研究.见:管理科学与系统科学进展——全国青年管理科学与系统科学论文集(第 3 卷).上海:上海交通大学出版社,1995. 121~ 124
- 3 魏毅强,刘进生,王绪柱.不确定型 AHP 中判断矩阵的一致性概念及权重.系统工程理论与实践,1994, 14(4): 16~ 22
- 4 刘进生,魏毅强,王绪柱.区间数判断矩阵的建立及其权重计算.系统工程,1993, 11(3): 42~ 46
- 5 许树柏.实用决策方法——层次分析法原理.天津:天津大学出版社,1988. 149~ 162
- 6 Yoon K. The propagation of errors in multiple-attribute decision analysis: A practical approach. Journal of the Operational Research Society, 1989, 40(7): 681~ 686
- 7 梁晋文,陈林才,何贡.误差理论与数据处理.北京:中国计量出版社,1989. 78~ 81
- 8 杨志超.误差理论.长沙:中南工业大学出版社,1987. 93~ 103
- 9 陈挺.决策分析.北京:科学出版社,1987. 184~ 186

An Application of Error Analysis Theory to the Multiple Attribute Decision Making with Interval Numbers

Fan Zhiping, Guo Yajun

ABSTRACT Based on the error analysis theory, an analysis method for the multiple attribute decision making with uncertain interval numbers is proposed, and this method is applied to the TOPSIS(Technique for Order Preference by Similarity to Ideal Solution) algorithm.

KEY WORDS multiple attribute decision making, interval number, error propagation, TOPSIS algorithm.

(Received June 11, 1996)