

文章编号: 1000-6788(2003)03-0112-05

一种基于熵的线性组合赋权法

汪泽焱¹, 顾红芳¹, 益晓新², 张申如¹

(1. 解放军理工大学理学院, 江苏 南京 211101; 2. 解放军理工大学通信工程学院, 江苏 南京 210016)

摘要: 给出了一种系统评价指标综合赋权方法——线性组合赋权法。它以优化理论和 Jaynes 最大熵原理为依据, 建立了新的数学模型, 并给出了精确解。应用实例表明了新方法的合理性。

关键词: 多指标评价; 线性组合赋权法; 熵

中图分类号: O 212.6

文献标识码: A

A Method of Determining the Linear Combination Weights Based on Entropy

WANG Ze-yan¹, GU Hong-fang¹, YI Xiao-xin², ZHANG Shen-ru¹

(1. Institute of Sciences, PLAUST, Nanjing 211101, China; 2. Institute of Communication Engineering, PLAUST, Nanjing 210016, China)

Abstract A new method of determining the index's weight, named the linear combination assigning method, is presented. Using optimization theory and Jaynes' maximum entropy principle a new mathematics model is established. The exact solution has been solved and an example demonstrates the validity of this method.

Key words: multiple index evaluating; linear combination assigning method; entropy

1 引言

多指标评价广泛存在于社会经济、工程技术和军事活动中, 是评价人员对多个目标方案进行综合衡量后做出合理的、正确的评价过程。影响多指标评价合理性的一个重要因素是评价指标权系数^[1,2]。目前关于权系数的确定通常有两类方法: 主观赋权法和客观赋权法。前者是由评价人员根据主观上对各指标的重视程度来决定权系数的一类方法, 常见的有专家调查法、循环打分法、二项系数法和 AHP 等。后者则是指利用指标值所反映的客观信息确定权系数的一种方法, 其原始数据由各指标在被评价对象中的实际数据形成, 常见的有均方差法、主成份分析法、离差最大化法、熵值法、代表计数法等。这两类方法各有优缺点: 主观赋权法解释性较强, 但客观性较差; 客观赋权法确定的权系数虽然大多数情况下客观性较强, 但有时会与各指标的实际重要程度相悖, 而且解释性较差, 对所得的结果难以给出明确的解释。

基于上述原因, 近几年来人们提出了综合主、客观赋权法的组合赋权法^[3-5]。文献[3]依据矩阵理论研究了多种权向量的组合赋值方法。文献[4]提出了以 AHP 法、专家调查法和误差逆传播神经网络技术(BP 网)相结合的综合赋权法。文献[5]根据优化理论提出了一种兼顾主观偏好和客观信息的综合赋权法, 最后得出的综合权值是给出的主观权值和客观权值的线性组合。

从数理统计的观点来看, 各指标的真实权系数是一个随机变量, 不同赋值方法得出的权系数只是真实权系数的一个样本值。上述文献中的方法都没有考虑各种赋值法带来的不确定性影响。本文引入 Shannon 熵来描述这种不确定性, 给出一种新的指标综合权系数的计算方法。该方法以优化理论和 Jaynes 最大熵

收稿日期: 2001-11-15

资助项目: 国防科技预研跨行业基金(00J642JB3804); 国防科技重点实验室基金(00JS0441JB3801)

作者简介: 汪泽焱(1972-), 男, 讲师, 主要研究方向为优化理论及应用、网络路由算法等。

原理为依据,建立了确定指标综合权系数的数学模型,并给出了模型的精确解 最后通过实例说明了此方法的合理性 .

2 线性组合赋权法

2.1 符号定义

设待评价的方案为 n 个,记为 $P = \{P_1, P_2, \dots, P_n\}$. 评价指标为 m 个,记为 $I = \{I_1, I_2, \dots, I_m\}$. 方案 P_i 对第 j 个指标 I_j 的评价值用 a_{ij} 表示 ($i = 1, 2, \dots, n; j = 1, 2, \dots, m$), 则 n 个方案的 $m n$ 个评价值构成矩阵 $A = [a_{ij}]_{n \times m}$, 称为方案集对指标集的评价矩阵

设评价分析人员利用主、客观赋权法,得到 l 种评价指标权向量 W^1, \dots, W^l . 设第 k 种权值向量为 $W^k = (w_1^k, w_2^k, \dots, w_m^k)^T$, 满足 $\sum_{j=1}^m w_j^k = 1, w_j^k \geq 0 (j = 1, 2, \dots, m; k = 1, 2, \dots, l)$.

2.2 评价指标规范化

由于在实际评价问题中各评价指标数量级和量纲会存在不同,而且属性也不一样,有的指标要求越小越好,有的指标要求越大越好,有的指标则要求稳定在某一确定值——理想值 因此为了消除各指标间的不可公度性和统一各指标的趋势要求,在进行评价前要对评价矩阵进行规范化处理

目前常用的指标类型有效益型、成本型和固定型等 效益型指标是指评价值越大越好的指标,成本型指标是指评价值越小越好的指标,固定型指标则是指评价值越接近某一固定值越好的指标 文献[6]对各种指标属性进行了比较,并综述了现有的指标类型的规范化方法,本文采用其中一种规范化方法

设 O_1, O_2, O_3 分别为成本型指标、效益型指标、固定型指标的下标集,则 $O_1 \cup O_2 \cup O_3 = \{1, 2, \dots, m\}$, 且 $O_s \cap O_t = \Phi (s \neq t, s, t = 1, 2, 3)$.

对成本型指标 I_j , 令

$$r_{ij} = \frac{a_j^{max} - a_{ij}}{a_j^{max} - a_j^{min}}, i = 1, 2, \dots, n, j \in O_1 \tag{1}$$

对效益型指标 I_j , 令

$$r_{ij} = \frac{a_{ij} - a_j^{min}}{a_j^{max} - a_j^{min}}, i = 1, 2, \dots, n, j \in O_2 \tag{2}$$

对固定型指标 I_j , 令

$$r_{ij} = \frac{b_j^{max} - b_{ij}}{b_j^{max} - b_j^{min}}, i = 1, 2, \dots, n, j \in O_3 \tag{3}$$

其中 $a_j^{min} = \min\{a_{ij} | i = 1, 2, \dots, n\}$, $a_j^{max} = \max\{a_{ij} | i = 1, 2, \dots, n\}$, 分别表示评价指标 I_j 的最小值和最大值 ($j \in O_1 \cup O_2$). $b_{ij} = |a_{ij} - \alpha_j|$, α_j 是固定型指标 I_j 的理想值, $b_j^{min} = \min\{b_{ij} | i = 1, 2, \dots, n\}$, $b_j^{max} = \max\{b_{ij} | i = 1, 2, \dots, n\}$ ($j \in O_3$). 记规范化后的评价矩阵为 $R = [r_{ij}]_{n \times m}$. 注意矩阵 R 中的每个元素满足: $0 \leq r_{ij} \leq 1$, 且每列中至少有一个元素等于 1. 从每列中取一个元素 1, 构成一个新的评价方案,称之为理想方案,可以理解为当前最好的方案 理想方案的评价值为 m 维全 1 向量

2.3 确定指标的综合权值

文献[5]中给出的综合权值为主观、客观权值的线性组合,即设 $\alpha = (\alpha_1, \dots, \alpha_n)^T$ 为主观权值, $\beta = (\beta_1, \dots, \beta_n)^T$ 为客观权值,则最后的综合权值为 $x\alpha + (1-x)\beta$, x 为 $[0, 1]$ 间的任一数,代表偏好系数,需要预先给定 本文在此结论基础上,给出一种确定组合权向量的方法

定义 1 称经多个权向量线性组合得到的向量为线性组合权向量

令 $W^* = (w_1^*, w_2^*, \dots, w_m^*)^T$ 为上述 l 种权向量 W^1, \dots, W^l 的线性组合权向量,用向量形式表示为

$$W^* = \sum_{k=1}^l x_k W^k \tag{4}$$

其中 x_k 表示线性组合系数,满足

$$\sum_{k=1}^l x_k = 1, x_k \geq 0 \tag{5}$$

令 $x = (x_1, x_2, \dots, x_l)^T$ 表示组合系数向量 很容易验证 (4) 式确定的 w^* 是权向量, 即满足

$$\sum_{j=1}^m w_j^* = \sum_{j=1}^m \sum_{k=1}^l x_k w_j^k = 1 \quad (6)$$

设在组合权向量 w^* 下第 i 种方案 P_i 的综合评价值为 V_i , 则

$$V_i = \sum_{j=1}^m w_j^* r_{ij} = \sum_{j=1}^m \sum_{k=1}^l x_k w_j^k r_{ij} = 1, \quad i = 1, 2, \dots, n \quad (7)$$

这时方案 P_i 与理想方案的广义距离为

$$d_i = \sum_{j=1}^m \sum_{k=1}^l x_k w_j^k (1 - r_{ij}), \quad i = 1, 2, \dots, n \quad (8)$$

从数理统计的观点来看, 在现实系统中, 各评价指标的真实权系数是一个随机变量, 它们构成的真实权向量是一个随机向量, 可将权向量 w_k 理解为真实权向量的第 k 个样本值, (4) 式中 w^k 的线性组合系数 x_k 在满足 (5) 式的条件下可以理解为真实权向量取样本值 w^k 的概率 因此组合系数向量 x 具有不确定性, 这种不确定性可以用 Shannon 信息熵^[7]来表示:

$$H = - \sum_{k=1}^l x_k \ln x_k \quad (9)$$

求线性组合权向量的目的, 就是要确定合适的组合系数向量 x , 一方面使所有方案与理想方案的加权广义距离和为最小, 即

$$\min \sum_{i=1}^n d_i = \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^l x_k w_j^k (1 - r_{ij}) \quad (10)$$

$$\text{s.t.} \quad \sum_{k=1}^l x_k = 1, x_k \geq 0 \quad (11)$$

而另一方面, 应尽量消除组合系数向量 x 的不确定性 根据 Jaynes 最大熵原理, 确定的指标综合权系数应使 Shannon 熵取极大, 即

$$\max H = - \sum_{k=1}^l x_k \ln x_k \quad (12)$$

$$\text{s.t.} \quad \sum_{k=1}^l x_k = 1, x_k \geq 0 \quad (13)$$

因此, 求线性组合权向量是一个多目标优化问题 为解此问题, 构造如下的单目标优化 (SP):

$$\min \mu \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^l x_k w_j^k (1 - r_{ij}) + (1 - \mu) \sum_{k=1}^l x_k \ln x_k \quad (14)$$

$$\text{s.t.} \quad \sum_{k=1}^l x_k = 1, x_k \geq 0 \quad (15)$$

其中正参数 $0 < \mu < 1$ 为参数, 用来表示两个目标之间的平衡系数, 可根据实际问题预先给出

定理 单目标优化问题 (SP) 有唯一解, 其解为

$$x = \left(\frac{s_1}{\sum_{k=1}^l s_k}, \frac{s_2}{\sum_{k=1}^l s_k}, \dots, \frac{s_l}{\sum_{k=1}^l s_k} \right) \quad (16)$$

其中 $s_k = \exp \left\{ - \left[1 + \mu \sum_{i=1}^n \sum_{j=1}^m w_j^k (1 - r_{ij}) / (1 - \mu) \right] \right\}, k = 1, 2, \dots, l$

证明 构造 Lagrange 函数

$$L(x, \lambda) = \mu \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^l x_k w_j^k (1 - r_{ij}) + (1 - \mu) \sum_{k=1}^l x_k \ln x_k - \lambda \left(\sum_{k=1}^l x_k - 1 \right)$$

根据极值存在的必要条件, 有

$$\frac{\partial L}{\partial x_k} = \mu \sum_{i=1}^n \sum_{j=1}^m w_j^k (1 - r_{ij}) + (1 - \mu) (\ln x_k + 1) - \lambda = 0, k = 1, 2, \dots, l \quad (17)$$

$$\frac{\partial}{\partial \lambda} = \sum_{k=1}^l x_k - 1 = 0$$

(18)

由(17)式得

$$x_k = \exp \left[\left(\lambda - \mu \sum_{i=1}^n \sum_{j=1}^m w_j^k (1 - r_{ij}) \right) / (1 - \mu) - 1 \right], \quad k = 1, 2, \dots, l$$

(19)

代入(18)式得

$$\exp [\lambda / (1 - \mu)] = 1 / \sum_{k=1}^l \exp \left\{ - \left[1 + \mu \sum_{i=1}^n \sum_{j=1}^m w_j^k (1 - r_{ij}) / (1 - \mu) \right] \right\}$$

(20)

将(20)式代回(19)式, 得到

$$x_k = \frac{\exp \left\{ - \left[1 + \mu \sum_{i=1}^n \sum_{j=1}^m w_j^k (1 - r_{ij}) / (1 - \mu) \right] \right\}}{\sum_{k=1}^l \exp \left\{ - \left[1 + \mu \sum_{i=1}^n \sum_{j=1}^m w_j^k (1 - r_{ij}) / (1 - \mu) \right] \right\}}$$

$k=1, 2, \dots, l$ 从而定理得证

3 实例分析

短波跳频通信以其较强的抗干扰能力和较大的组网容量, 成为现代军事通信的一种重要手段^[8]. 为了较好地评价短波跳频通信系统的抗干扰能力, 可以依据如下八个指标来做出合理的抉择, 即同步概率(I_1)、同步时间(I_2)、同步保持时间(I_3)、同步频率更换时间(I_4)、迟入网成功概率(I_5)、迟入网时间(I_6)、跳速(I_7)和跳频图案(I_8). 现有四种短波跳频通信设备(即四种待评方案 P_1 、 P_2 、 P_3 和 P_4), 某部邀请专家对它们进行性能测试, 从中挑选性能最好的设备作为部队建设的重点. 专家给出的八种指标综合评价结果列于表 1.

表 1 专家评价价值表

指标 方案	I_1	I_2	I_3	I_4	I_5	I_6	I_7	I_8
P_1	46.5	56.0	42.7	29.8	35.6	27.6	33.4	41.5
P_2	36.8	51.0	46.5	45.6	49.5	70.4	35.4	35.7
P_3	62.1	48.7	52.6	46.8	42.9	54.7	42.9	26.4
P_4	50.0	33.5	29.8	26.7	18.6	57.1	38.2	52.6

首先计算规范化评价矩阵 R . 这里属于成本型指标的是 I_2 、 I_4 和 I_6 , 属于效益型指标的是 I_1 、 I_3 、 I_5 、 I_7 和 I_8 . 根据(1)式和(2)式, 得到规范化评价矩阵为

$$R = \begin{bmatrix} 0.3834 & 0.0000 & 0.5658 & 0.8458 & 0.5502 & 1.0000 & 0.0000 & 0.5763 \\ 0.0000 & 0.2222 & 0.7325 & 0.0597 & 1.0000 & 0.0000 & 0.2105 & 0.3550 \\ 1.0000 & 0.3244 & 1.0000 & 0.0000 & 0.7864 & 0.3668 & 1.0000 & 0.0000 \\ 0.5217 & 1.0000 & 0.0000 & 1.0000 & 0.0000 & 0.3107 & 0.5053 & 1.0000 \end{bmatrix}$$

专家依据不同赋权方法给出的三种评价指标权向量为

$$W^1 = (0.155, 0.138, 0.143, 0.105, 0.119, 0.096, 0.112, 0.132)^T$$

$$W^2 = (0.102, 0.163, 0.129, 0.101, 0.122, 0.095, 0.108, 0.180)^T$$

$$W^3 = (0.096, 0.127, 0.146, 0.112, 0.128, 0.181, 0.082, 0.128)^T$$

应用这三种权向量得到四种待评方案的综合评价值及结果如表 2 所示

从表 2 可以看出, 不同的赋权方法得出的结果不尽相同, 从而不能给决策者提供一个很好的决策依据

现以本文提出的线性组合赋权法来给出评价, 取平衡系数 $\mu=0.8$



依 (16) 式求出线性组合系数向量为 $x = (0.3403, 0.3258, 0.3339)^T$, 代入 (4) 式得组合权向量为 $w^* = (0.1180, 0.1425, 0.1394, 0.1060, 0.1230, 0.1241, 0.1007, 0.1463)^T$ 最后代入 (7) 式得四种方案的综合评价值为 $(0.4899, 0.3362, 0.5466, 0.5458)^T$. 运算结果示于表 2 最后一列

从表 2 还可以看出, 应用组合权向量 w^* 来评价四种方案, 虽然得到的排序结果为 $P_3 > P_4 > P_1 > P_2$, 但方案 P_3 与 P_4 的综合评价价值相差无几, 但都比余下两种方案的综合评价价值要好, 因此可以认为 P_3 与 P_4 性能基本相当, 都可以作为候选方案. 因此决策者可以依据不同的环境选择 P_3 与 P_4 中的任一种. 如果单凭一种指标赋权法是得不出这种结论的.

表 2 利用不同权向量得到的评价结果表

权向量 方案	w^1	w^2	w^3	w^*
P_1	0.4667	0.4634	0.5393	0.4899
P_2	0.3311	0.3454	0.3326	0.3362
P_3	0.5836	0.5227	0.5322	0.5466
P_4	0.5423	0.5813	0.5148	0.5458
最终排序结果	$P_3 > P_4 > P_1 > P_2$	$P_4 > P_3 > P_1 > P_2$	$P_1 > P_3 > P_4 > P_2$	$P_3 > P_4 > P_1 > P_2$

注: 表 2 中各权向量对应列的数据为方案的综合评价价值

4 结束语

评价指标权系数的选择是多指标系统评价中的关键问题. 本文利用优化理论和 Jaynes 最大熵原理, 给出了一种计算指标权系数的方法——线性组合赋值法. 从文中的数学模型可以看出, 本文中提出的方法既考虑了待评方案与当前理想方案的差距为最小, 又考虑到了权系数的不确定性. 最后的应用实例说明了此方法的合理性. 由此可见, 文中提出的模型和方法对于系统评价的指标权系数的综合确定具有一定的实际意义.

参考文献:

[1] 陈珏. 决策分析[M]. 北京: 科学出版社, 1987.
[2] 胡毓达. 实用多目标最优化[M]. 上海: 上海科学技术出版社, 1990.
[3] 魏巍贤, 冯佳. 多目标权系数的组合赋值方法研究[J]. 系统工程与电子技术, 1998, 20(2): 14- 16.
[4] 梁杰, 侯志伟. AHP 法专家调查法与神经网络相结合的综合定权方法[J]. 系统工程理论与实践, 2001, 21(3): 59- 63.
[5] 陶菊春, 吴建民. 综合加权评分法的综合权值确定新探[J]. 系统工程理论与实践, 2001, 21(8): 43- 48.
[6] 刘树林, 邱菀华. 多属性决策基础理论研究[J]. 系统工程理论与实践, 1998, 18(1): 38- 43.
[7] 朱雪龙. 应用信息论基础[M]. 北京: 清华大学出版社, 2001.
[8] 汪泽炎, 张申如, 益晓新, 等. 短波跳频通信系统抗干扰性能的一个评价算法[J]. 军事通信技术, 2001, 22(4): 6- 11.