文章编号: 1003-207(2006)03-0051-05

挖掘客户潜在需求确定指标重要程度的证据组合方法

邓明荣[1], 吴红梅[2]

(1. 浙江大学管理学院, 杭州 310058; 2. 浙江大学城市学院, 杭州 310015)

摘 要: 在现实中, 属性(例如产品的各种指标) 的权重常常与特定客户的需求有关, 因此这些客户的偏好信息对权重的确定非常重要。这种信息不一定就直接体现在对属性的重视程度上, 而可能表现在对一组具体产品的偏好关系上。在对产品的指标评价时还会存在一些无知性。针对这些情况, 本文给出一个基于证据理论和目标规划、挖掘客户潜在需求的指标权重估计方法, 并结合一个计算示例, 给出具体的计算过程。

关键词: 权值估计; 证据组合; 目标规划
中图分类号: F224.12 文献标识码: A

1 引言

属性权重的确定是决策分析中的重要问题。例如对产品设计来说, 产品的生命周期正在变得越来越短, 及时发现顾客的潜在需求, 设计相应的产品, 已经成为企业致胜的关键。

已有很多关于权值估计的研究, 典型的是 Keeney 和 Raiffa 提出的一系列两两比较的过程[6]。层次分析法的基本原理也是通过指标间两两比较重要程度, 得到比较一致的互反矩阵后, 计算其特征向量或平均(算术或几何) 向量作为权重估计。但是在实际的操作过程中, 顾客往往往很难给出大量精确、一致的两两比较。Shirland 等为此提出了 3 个一组进行比较的方法[8]。对于比较的结果也常常有不确定性存在, 当这种不确定性表现为模糊性时, 模糊层次分析法是一种有效方法[5, 7]。

另外有一些方法试图采用比专家意见更客观的方法来估计指标的重要程度, 例如根据指标值的分布情况。其中一些方法基于这样一个思想, 即, 如果所有个体在某些指标上有相似的评价, 则这些指标是相对不重要的, 因此可以根据个体各指标值的分布来估计权值[2, 4]。

另外在多属性决策分析的理论中, 常常假定权重的信息是直接给出的。为了表示权重本身的不确定, 多种信息集结了算子中也将权重作为模糊数或语言型变量[11]。Yager 关于权重的研究则更进了一步[9], 在他提出的有序加权平均(OWA) 算子和评估函数(valuation function) 中, 权重不仅与属性有关, 而且还与属性的值有关, 例如悲观主义者将属性值最差的属性权重设为最大。

本文研究的是在以下情况下估计属性权重的问题, 即客户或决策者不能给出关于属性权重的明确信息, 但是却能够给出对一组具体产品或方案的偏好关系。关于每个产品或方案在各个属性上的评价具有某种不确定性, 这种不确定性不是表达的模糊性, 而是来自于无知性。

在[3] 中, 邓明荣等提出了一个从属性评价综合成总体评价的证据组合公式。在本文中, 我们将通过证据组合, 利用效用函数, 将客户偏好关系表示成目标规划模型, 从而根据客户的潜在需求导出权重估计。

2 问题描述

已知一个个体(例如产品、方案) 的集合 $\{a_1, a_2, \cdots, a_M\}$, 对于每个个体有一个属性(指标) 的集合 $\{c_1, c_2, \cdots, c_L\}$ (数量化的或非数量的, 通常表现为一个层次结构)。每个个体的底层属性可以通过主观或者客观的方法得到较直接的评价。

虽然对指标的评价可以采用不同的量度, 为说明方便, 我们将评价的等级统一为一致的集合。例

收稿日期: 2006-02-06; 修订日期: 2006-03-31
作者简介: 邓明荣(1965-), 男(汉族), 浙江大学管理学院, 博士, 副研究员, 浙江大学物流与决策优化研究所副所长, 研究方向: 决策分析、物流与供应链管理等.

如可以对非数量的指标用基于规划的技术, 对数量化的指标用效用理论, 将它们的评价等级集合转化为统一的集合{“差”,“一般”,“好”,“优”}。记为 $H = \{H_1, H_2, \cdots, H_N\}$。

在现实中, 在对一个体的指标进行评价时, 也常常会有不确定的情况。这里我们主要考虑来自于无知的不确定性。例如对汽车的“熬车性能”, 可以把它评为一个等级“好”; 也可以有 50% 的把握为“好”, 50% 的把握为“优”。特别是当评价数据是综合很多顾客的意见得到、并且有顾客弃权的情形时, 就会综合得到这样的评价。

因此, 对个体 a_1 的指标 c_i, 评价可以表示为,

$$S(c_i(a_1)) = \{H_n, \beta_{n,i}^{(1)}), n = 1, \cdots, N\}, i = 1, \cdots L, l = 1, \cdots, M \tag{1}$$

这里每个 $\beta_{n,i}(a_1) \geqslant 0$, 并且 $\sum_{n=1}^{N} \beta_{n,i}^{(1)} \leqslant 1$

$\beta_{n,i}^{(1)}$ 表示 a_1 在指标 c_i 被评为级别 H_n 的信任程度。

当 $\sum_{n=1}^{N} \beta_{n,i}^{(1)} = 1$ 时, 可以称 a_1 的 c_i 评价是完全的(没有弃权的情况), 特别当有某个 n 使得 $\beta_{n,i}^{(1)} = 1$ 时, 评价是确定的(顾客的意见是一致的)。

对每个个体的综合情况, 当然也可以用同样的方式来评价,

$$S(y(a_1)) = \{(H_n, \beta_n^{(1)}), n = 1, \cdots, N\} \tag{2}$$

从各个指标的评价综合得到总体评价, 是多指标决策分析的一个重要问题。但在现实中, 顾客可能更容易给出对一组个体的偏好关系, 例如:

1) a_i 与 a_j 没什么区别, 记为 $a_i \sim a_j$;

2) a_i 比 a_j 好, 记为 $a_i > a_j$;

3) a_i 比 a_j 好一个倍数, 比如, $L_{ij} a_j$, 即 $a_i > L_{ij} a_j$;

4) a_i 至少比 $m_{ij} a_j$ 好, 但是最多也比不上 $M_{ij} a_j$, 即 $M_{ij} a_j > a_i > m_{ij} a_j$;

我们的问题是, 给出一组个体在底层指标上形似(1) 式的评价, 以及个体集合上一组形似 1), 2), 3) 或 4) 的优先关系, 如何确定每个指标对总体评价的重要程度。

3 证据组合的方法

我们以指标层次为简单两层的情况, 即下层有若干基本指标 $\{c_1, c_2, \cdots, c_L\}$, 上层有一个综合指标 y 为例, 说明如何利用证据组合的方法进行评价的综合。

设 ω_i 是指标 c_i 的权值, 满足, $1 \geqslant \omega_i \geqslant 0$ 且

$$\sum_{i=1}^{L} \omega_i = 1 \tag{3}$$

先用 $N = 4$ 的简单情况来说明证据组合的过程。假设对 a_1 的 c_1, c_2 指标有以下评价:

$$S(c_1(a_1)) = \{(H_1, \beta_{1,1}^1), (H_2, \beta_{2,1}^1), (H_3, \beta_{3,1}^1), (H_4, \beta_{4,1}^1)\}$$

$$S(c_2(a_1)) = \{H_1, \beta_{1,2}^1), (H_2, \beta_{2,2}^1), (H_3, \beta_{3,2}^1), (H_4, \beta_{4,2}^1)\}$$

则根据 $S(c_1(a_1))$ 和 $S(c_2(a_1))$ 可以得到一个评价的组合 $S(c_1(a_1)) \oplus S(c_2(a_1))$, 假设 $S(c_1(a_1))$ 和 $S(c_2(a_1))$ 都是完全的, 并且,

$$m_{n,1} = \omega_1 \beta_{n,1}^1 (n = 1, \cdots, 4),$$

$$m_{H,1} = 1 - \omega_1 \sum_{n=1}^{4} \beta_{n,1}^1 = 1 - \omega_1$$

$$m_{n,2} = \omega_2 \beta_{n,2}^1 (n = 1, \cdots, 4),$$

$$m_{H,2} = 1 - \omega_2 \sum_{n=1}^{4} \beta_{n,2}^1 = 1 - \omega_2$$

这里每个 $m_{n,j}(j = 1, 2)$ 是指标 c_j 评价为 H_n 的基本概率数, $m_{H,j}$ 是关于指标 c_j 评价不确定的概率数。则用 D-S 证据组合的方法, 得到的组合概率数 $m_n(n = 1, \cdots, 4)$ 和 m_H 分别为,

$$m_n = k(m_{n,1}, m_{n,2} + m_{H,1} m_{n,2} + m_{n,1} m_{H,2}), (n = 1, \cdots, 4)$$

$$m_H = k(m_{H,1}, m_{H,2})$$

其中,

$$k = \left(1 - \sum_{t=1}^{4} \sum_{\substack{n=1 \\ n \neq t}}^{4} m_{t,1} m_{n,2}\right)^{-1}$$

用同样的方式可以再和第三个指标上的评价进行组合, …, 最后就可以得到一个总的综合评价, 即评为 H_n 的信任度为,

$$\beta_n = \frac{m_n}{1 - m_H}$$

当指标有多个层次时, 可以用同样的方式来进行聚合。更简单的方法是利用以下的聚合计算公式(参见[3])。

定理. 设 $m_{n,j}(n = 1, 2, \cdots, N, j = 1, 2, \cdots, L)$ 是指标 c_j 评价为 H_n 的基本概率数, $m_{H,j}$ 是关于指标 c_j 评价不确定的概率数。对每个 $n = 1, \cdots N$, 设

$$M_n = (m_{n,1} + m_{H,1})(m_{n,2} + m_{H,2}) \cdots (m_{n,L} + m_{H,L}),$$

$M_0 = m_{H,1} m_{H,2} \cdots m_{H,L}$, 则对综合指标 y, 它被评价为每个等级的基本概率数为, $m_n = \left(\sum_{n=1}^{N} M_n - (N-1) M_0\right)^{-1} (M_n - M_0)$, $n = 1, \cdots, N$

$m_H = (\sum_{n=1}^{N} M_n - (N-1)M_0)^{-1}(M_0)$ 为不知如何分配的概率数。

4　权重估计

我们利用效用理论来表示顾客对个体的偏好关系。所谓效用理论, 就是假定存在这样一个效用函数, 使得一个个体比另一个个体优先时, 对应着它的效用函数值比另一个个体的效用函数值大。通常是根据个体的总体评价来确定它的效用, 因此首先要确定评价等级正好为 H_n 的个体效用, μ_n。但是对于顾客来说, 效用并不是一个具体的概念, 很难事先确定一组符合顾客偏好的等级效用值。在我们的方法中, 将它们也当作变量来处理, 并且只作最基本的假设,

$$\mu_1 = 0,\ \mu_N = 1; \tag{4}$$

$$\mu_{n+1} > \mu_n + \delta_2,\ (n = 1, \cdots, N-1); \tag{5}$$

这里, δ_2 是一个很小的正常数, 目的是充分区分开相邻两个等级的效用, 当没有任何其它信息可用时, 可以取作 $1/(2(N-1))$。

这样, 我们可以根据一个个体 a_1 的综合指标评价得到它的期望效用为 $\mu(S_y(a_1)) = \sum_{n=1}^{N} \mu_n \beta_n^{(1)}$

其中, 我们用 $\beta_n^{(1)}$ 来表示 a_1 被评价为 H_n 的综合信任程度, 它是权重向量 $\omega = (\omega_1, \omega_2, \cdots, \omega_L)$ 的非线性函数。

因此顾客给出的 1), 2), 3), 4) 形式的优先关系, 用期望效用可以分别表示为:

$$\sum_{n=1}^{N} \beta_n^{(i)}(\omega)\mu_n = \sum_{n=1}^{N} \beta_n^{(j)}(\omega)\mu_n \tag{6}$$

$$\sum_{n=1}^{N} \beta_n^{(i)}(\omega)\mu_n > \sum_{n=1}^{N} \beta_n^{(j)}(\omega)\mu_n \tag{7}$$

$$\sum_{n=1}^{N} \beta_n^{(i)}(\omega)\mu_n > L_{ij} \sum_{n=1}^{N} \beta_n^{(j)}(\omega)\mu_n \tag{8}$$

$$M_{ij} \sum_{n=1}^{N} \beta_n^{(j)}(\omega)\mu_n > \sum_{n=1}^{N} \beta_n^{(i)}(\omega)\mu_n > m_{ij} \sum_{n=1}^{N} \beta_n^{(j)}(\omega)\mu_n \tag{9}$$

由于在顾客给出的优先关系中, 可能存在某种不一致性, 我们在每个关系中加入偏差变量 d。另外引进一个小的正常数 δ_1 来区分开两个有优先关系的个体, 这样(6), (7), (8), (9)变为,

$$\sum_{n=1}^{N} u_n \beta_n^{(i)}(\omega) - \sum_{n=1}^{N} u_n \beta_n^{(j)}(\omega) - d_{ij}^{+} + d_{ij}^{-} = 0 \tag{10}$$

$$\sum_{n=1}^{N} u_n \beta_n^{(i)}(\omega) - \sum_{n=1}^{N} u_n \beta_n^{(j)}(\omega) - d_{ij}^{+} + d_{ij}^{-} \geqslant \delta_1 \tag{11}$$

$$M_{ij} \sum_{n=1}^{N} u_n \beta_n^{(j)}(\omega) - \sum_{n=1}^{N} u_n \beta_n^{(i)}(\omega) - d_{ij}^{+} + d_{ij}^{-} \geqslant \delta_1 \tag{12}$$

$$\sum_{n=1}^{N} u_n \beta_n^{(i)}(\omega) - m_{ij} \sum_{n=1}^{N} u_n \beta_n^{(j)}(\omega) - d_{ij}^{+} + d_{ij}^{-} \geqslant \delta_1 \tag{13}$$

$$d_{ij}^{+} \times d_{ij}^{-} = 0 \tag{14}$$

这里 d_{ij}^{+} 和 d_{ij}^{-} 是满足互补条件的非负偏差变量。

首先我们要确定顾客给出的优先关系是否是一致的, 这可以通过计算是否存在零偏差的可行解来判断。即解数学规划:

$$\min\left\{ \sum_{i,j \in I^1} (d_{ij}^{+} + d_{ij}^{-}) + \sum^{i,j \in I^2} d_{ij}^{-} \right\}$$
$$\text{s.t.}\ (\omega, u, d) \in \Omega \tag{15}$$

这里 Ω 表示由(3–5)和(10–14)式表示的可行区域; 目标函数中 I^1 表示所有如(10)等式约束的个体指标对, I^2 表示有(11)、(12)、(13)式不等式约束的指标对。

如果对于得到的最优解$(\hat{\omega}, \hat{u}, \hat{d}) \in \Omega$, 其最优值为 0, 即对所有$(ij) \in I^1$ 有 $\hat{d}_{ij}^{+} = 0$, $\hat{d}_{ij}^{-} = 0$; 对所有$(ij) \in I^2$ 有 $\hat{d}_{ij}^{+} \geqslant 0$, $\hat{d}_{ij}^{-} = 0$ 时, 说明顾客给出的优先关系是一致的, 否则说明顾客给出的优先关系中存在不一致, 这时需要顾客修正优先关系。例如, 放松某些关系, 放弃某一个不肯定的关系, 等等。

当给出的优先关系中不存在矛盾时, 规划(15)的最优解中的 ω 也不一定就是我们所要求的指标权值, 因为可能存在着无数多的满足条件的权值。这时, 我们可以通过求解以下两个数学规划来得到每个指标权值的最小值与最大值 $\overline{\omega}_i$, $\underline{\omega}_i$,

$$\begin{array}{ll} \min\ \omega_i & \max\ \omega_i \\ \text{s.t.}\ (\omega, u, \hat{d}) \in \Omega & \text{s.t.}\ (\omega, u, \hat{d}) \in \Omega \end{array} \tag{17}$$

得到指标 c_i 的一个估计区间$[\overline{\omega}_i, \omega_i]$。当 $\overline{\omega}_i > \underline{\omega}_j$ 时, 我们可以肯定指标 c_i 比 c_j 更重要。但是, 当两个区间有重叠部分时, 情况变得复杂了。这时, 为了进一步确定它们之间的关系, 可以用两种方法, 一是简单地通过计算平均权值得到权值的一个估计, 在区间长度较短时, 可以适用。当区间长度较大时, 就有必要采用第二种途径, 即扩充优先关系的集合, 这等价于顾客提供更多的信息。

顾客越能提供丰富的信息, 就越能确定指标的重要程度。

5　计算示例

考虑一个六种汽车的评价问题。假设顾客最关

心的七个指标是,加速性能(acceleration),熬车性能(braking),马力(horsepower),操作性能(handling),行驶质量(ride quality),动力(powertrain)和耗油(fuel economy),其中四个指标是定量指标,三个是定性指标。

假设已通过其他方法将所有指标的评价等级统一为{W, P, A, G, E, T},分别表示“最差”(worst),“差”(poor),“一般”(average),“好”(good),“优秀”(excellent)和“顶级”(top)。然后得到的六种汽车的各个指标的评价数据如下:

表 1 六种汽车的性能评价

性能	Car 1	Car 2	Car 3	Car 4	Car 5	Car 6
耗油	{(P, 0.2), (A, 0.8)}	{(G, 0.5), (E, 0.5)}	{(E, 0.75), (T, 0.25)}	{(A, 0.4), (G, 0.6)}	{(G, 0.5), (E, 0.5)}	{(G, 0.25), (E, 0.75)}
加速	{(G, 1.0)}	{(E, 0.333), (T, 0.667)}	{(G, 0.5), (E, 0.5)}	{(P, 0.75), (A, 0.25)}	{(P, 1.0)}	{(E, 1.0)}
操作	{(A, 0.4), (G, 0.6)}	{(E, 0.6), (T, 0.4)}	{(A, 0.4), (G, 0.6)}	{(A, 1.0)}	{(G, 1.0)}	{(E, 0.6), (T, 0.4)}
马力	{(E, 0.333), (T0.667)}	{(P, 0.533), (A, 0.467)}	{(G, 0.462), (E, 0.538)}	{(G, 0.385), (E, 0.615)}	{(W, 0.467), (P, 0.533)}	{(A, 0.267), (G, 0.733)}
熬车	{(G, 0.6), (E, 0.4)}	{(A, 1.0)}	{(A, 0.4), (G, 0.6)}	{(G, 1.0)}	{(G, 1.0)}	{(G, 0.6), (E, 0.4)}
动力	{(A, 0.4), (G, 0.6)}	{(G, 1.0)}	{(E, 0.6), (T, 0.4)}	{(A, 0.4), (G, 0.6)}	{(G, 0.6), (E, 0.4)}	{(E, 0.6), (T, 0.4)}
行驶	{(G, 1.0)}	{(G, 1.0)}	{(E, 1.0)}	{(G, 1.0)}	{(A, 1.0)}	{(G, 1.0)}

为了估计每个汽车指标在顾客心目中的重要程度,需要顾客给出对六种汽车的偏好关系。

首先假设顾客给出以下优先关系,

Car 5> Car 4> Car 1> Car 2;

Car6> Car 3> Car 2.

通过求解目标规划(15),其最优值为零,说明顾客给出的优先关系是一致的。再解相应的数学规划(17),可以得到每个指标的权值估计区间,

$\omega_1= 0.026\sim 0.490$, $\omega_2= 0\sim 0.223$, $\omega_3= 0\sim 0.273$,

$\omega_4= 0\sim 0.302$, $\omega_5= 0\sim 0.741$, $\omega_6= 0\sim 0.607$,

$\omega_7= 0\sim 0.264$.

到这里,我们还不能看出各个指标的重要程度,也就是说,顾客给出的优先关系还不能本质地反映各指标的重要程度。

为了得到更确定的指标权值,要求顾客进一步提供信息。

比如,顾客又想起来,Car 5> Car 3.

因此我们得到新的数学规划模型,求解后,除第一个指标基本不变外,其它指标得到了更精确的估计,

$\omega_2= 0.038\sim 0.490$, $\omega_3= 0\sim 0.273$, $\omega_4= 0\sim 0.195$,

$\omega_5= 0.283\sim 0.741$, $\omega_6= 0\sim 0.558$, $\omega_7= 0\sim 0.222$。

现在我们大致上可以知道,第五个指标最重要,但是对于其它各个指标还不能有确切的结论。

现在假设顾客进一步给出了一个优先关系,例如,Car4> Car3.

这样,根据数学规划我们又得到每个指标的权值估计如下,

$\omega_1= 0.084\sim 0.459$, $\omega_2= 0\sim 0.148$, $\omega_3= 0\sim 0.170$,

$\omega_4= 0\sim 0.189$, $\omega_2= 0.422\sim 0.741$, $\omega_2= 0\sim 0.384$,

$\omega_7= 0\sim 0.215$。

现在可以看出,第一辆汽车和第三辆汽车之间的优先关系对我们确定各指标的重要性起着关键的作用。假设面对第一辆汽车和第三辆汽车时,顾客选择了第三辆汽车。通过同样的过程,我们得到,

$\omega_1= 0.116\sim 0.459$, $\omega_2= 0\sim 0.139$, $\omega_3= 0\sim 0.210$,

$\omega_4= 0\sim 0.185$, $\omega_5= 0.425\sim 0.646$, $\omega_6= 0\sim 0.358$,

$\omega_7= 0\sim 0.210$。

这样,指标之间的权重关系一点一点地显露出来了。

6 结语

本文中,我们提出了一个基于证据组合和目标规划的属性权值估计方法。通过挖掘客户对一组产品的偏好程度,结合这些产品的指标状况,通过不断与客户交互,就可以分析客户的潜在需求,确定产品的各个指标权值,最终指导产品的设计。

与事先确定权重的方法比较,这种方法更基于事实和行为的结果分析。

参考文献:

[1] Barzilai J. Deriving weight frojm pairwise comparison matri-

ces[J]. Journal of the Operational Research Society, 1997, 48: 1226- 1232.

[2] Deng H, Yeh CH And Willis R. Inter- company comparison using modified TOPSIS with objective weights[J]. Computers & Operations Research, 2000, 27: 963- 973.

[3] Deng M, Xu W and Yang J- B. Estimating the attribute weights through evidential reasoning and programming[J]. International Journal of Information Technology & Decision Making, 2000, (3): 3: 419- 428.

[4] Diakoulaki D, Mavrotas G and Papayannakis L. Determining objective weights in multiple criteria problems: the CRITIC method[J]. Computers & Operations Research, 1995, 22: 963- 790.

[5] Kahraman C., Cebeci U. and Ruan D. Multi- attribute comparison of catering service companies using fuzzy AHP: the case of Turkey[J]. International Journal of Production Economics, 2004, 87: 171- 184.

[6] Keeney R anmd Raiffa H. Decision with Multiple Objective, Preferences and Value Trade Offs[M]. Wiley, New Yrok, 1976.

[7] Mihailov L. Fuzzy analytical approach to partnership selection information of virtual enterprises[J]. Omega, 2002, 30: 393- 401.

[8] Shirland L. E., Jesse R. R., Thompson R. L. et al. Determining attribute weights using mathematical programming [J]. Omega, 2003, 31: 423- 437.

[9] Yager R R. On the determination of strength of belief for decision support under uncertainty - part I: generating strengths of belief[J]. Fuzzy Sets and Systems, 2004, 142: 117- 128.

[10] Yang JB. Rule and utility based evidential reasoning approach for multiattribute decision analysis under uncertainities[J]. European Journal of Operational Research, 2001, 131: 31- 61.

[11] 徐泽水. 不确定多属性决策方法及应用[M]. 清华大学出版社, 2004.

An Approach for Customer Requirement Mining and Attribute Weight Estimating Based on Evidence Combination

DENG Ming-rong[1], WU Hong-mei[2]

(1. School of Management, Zhejiang University, Hangzhou 310058, China;
2. City College, Zhejiang University, Hangzhou 310015, China)

Abstract: In reality, the weights of attributes(e. g. indices of a product) often depend on the requirement of the customer. Hency the customer preference information is important for weight identifyiing. Such information may be not direct expressed as importance attached to attributes, but implied in their preference on a set of given alternatives. Along with the attribute evaluation, there also maybe exists uncertainty of ignorance. For this kind of circumstance, an attribute weight estimating method based on mining the latent customer requirement is proposed. The theory of evidence combination and goal programming is employed in the procedure. An illustration example is given.

Key words: weight estimation; evidence combination; goal programming