复杂系统评估指标的量化技术及权重分配

唐伟勤[1,2]

(1.华中科技大学 管理学院,湖北 武汉 430074;2.中南财经政法大学 安全科学与管理学院,湖北 武汉 430073)

摘 要:许多传统的关于指标量化技术提出的公式,在很多情况下,其指标值不能达到固定的值或区间,因此量化后的指标值不能等于1或0。为了克服这一缺陷,提出了一组新的量化公式,这组公式有较广泛的应用面,对于任意一组给定的指标值都能使量化后的至少一个指标值达到1或0。同时,传统的指标权重的确定方法主要有客观赋权法和主观赋权法,客观赋权法确定的权重优势与分目标的实际重要程度相悖,而主观赋权法又存在一定的主观随意性。提出了基于关联度的权重求法和基于满意度的权重求法两种新方法,避免了两类赋权法的缺憾。

关键词:指标量化技术;指标属性;关联度;满意度
中图分类号:F224.0 文献标识码:A 文章编号:1001-7348(2009)03-0116-03

0 引言

在复杂系统多指标评估中,指标是反映对象之间关系的指示量[1],其原始数据一般来源于系统的各种渠道的统计资料,通常有以下 4 个特点:①指标不同,且量纲大都也不同;②不同指标之间,数量级大小相差悬殊;③同一指标的数据大小参差不齐,且大部分是离乱的,有明显的随机性;④多数数据有一定的灰色度,甚至出现短缺、虚假等现象。而不同量纲、不同数量级之间不便于比较,或者难以得到正确的比较结果。因此,为保证建立模型的质量、系统分析的正确结果以及决策的效果,对收集来的各指标的原始数据必须进行数据变换和处理(即量化),使其消除量纲、增强可比性、弱化随机性、初步显示出规律性。

1 复杂系统评估指标的量化

1.1 多指标的复杂系统模型

考虑如下复杂系统多指标评估问题:

$$(\mathrm{MOP})\max f(x)=(f_1(x),f_2(x),\cdots,f_n(x))$$

$$s.t.\ x\in X=\{x_1,x_2,\cdots,x_m\}$$

具体表示为:

$$A=\begin{array}{c} \\ x_1\\ x_2\\ \cdots\\ x_m\end{array}\begin{array}{c}\begin{array}{cccc} f_1 & f_2 & \cdots & f_n\end{array}\\ \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix}\end{array}$$

其中$X=\{x_1,x_2,\cdots,x_m\}$为方案集(对象空间),$D=\{f_1,f_2,\cdots,f_n\}$为指标集(属性空间),$a_{ij}=f_j(x_i)$,$i=1,2,\cdots,m$;$j=1,2,\cdots,n$,为方案 x_i 在指标 f_j 下的属性值,$A=(a_{ij})_{m\times n}$ 为指标值矩阵。

1.2 复杂系统评估指标类型

一般来说,指标的类型(属性)包括效益型、成本型、固定型、区间型、偏离型和偏离区间型 6 种。效益型属性是指标越大越好的属性,成本型属性是指标越小越好的属性,固定型属性是指标值越接近某固定值 α_j 越好的属性,区间型属性是指标值越接近某固定区间 $[q^j_1,q^j_2]$(包括落入该区间)越好的属性,偏离型属性是指标值越偏离某固定值 β_j 越好的属性,偏离区间型属性是指标值越偏离某区间$[p^j_1,p^j_2]$越好的属性。记 $M=\{1,2,\cdots,m\}$为对象集的下标集合,$N=\{1,2,\cdots,n\}$为指标集的下标集合,$N_t(t=1,2,\cdots,6)$分别表示效益型、成本型、固定型、区间型、偏离型和偏离区间型指标的下标集合,则有:

$$N=\bigcup_{t=1}^{6} N_t$$

1.3 复杂系统评估指标量化方法的改进

对于上述 6 种不同属性的指标,其规范化方法(或量化技术)有很多,文献[2]中提出 3 组公式,但实际上,这 3 组公式均存在着不合理性,以第一组公式(公式(1)-(4))为例:

$$z_{ij}=\frac{\max\limits_i|a_{ij}-\alpha_j|-|a_{ij}-\alpha_j|}{\max\limits_i|a_{ij}-\alpha_j|-\min\limits_i|a_{ij}-\alpha_j|},i\in M,j\in N_3 \qquad (1)$$

收稿日期:2007-07-12
作者简介:唐伟勤(1971-),女,湖北武汉人,华中科技大学管理学院博士研究生,中南财经政法大学安全科学与管理学院讲师,研究方向为安全生产管理、应急管理。

$$z_{ij}=\frac{\max\limits_i \max\{a_{ij}-q_1^j,q_2^j-a_{ij}\}-\max\{a_{ij}-q_1^j,q_2^j-a_{ij}\}}{\max\limits_i \max\{a_{ij}-q_1^j,q_2^j-a_{ij}\}-\min\limits_i \max\{a_{ij}-q_1^j,q_2^j-a_{ij}\}}$$

$$\mathrm{i}\in \mathrm{M},\mathrm{j}\in \mathrm{N}_4 \qquad (2)$$

$$z_{ij}=\frac{|a_{ij}-\beta_j|-\min\limits_i |a_{ij}-\beta_j|}{\max\limits_i |a_{ij}-\beta_j|-\min\limits_i |a_{ij}-\beta_j|} \quad i\in M,j\in N_5 \qquad (3)$$

$$z_{ij}=\frac{\max\{p_1^j-a_{ij},a_{ij}-p_2^j\}-\min\limits_i \max\{p_1^j-a_{ij},a_{ij}-p_2^j\}}{\max\limits_i \max\{p_1^j-a_{ij},a_{ij}-p_2^j\}-\min\limits_i \max\{p_1^j-a_{ij},a_{ij}-p_2^j\}} \quad i\in M,j\in N_6 \qquad (4)$$

公式(1)–(4)依次代表固定型、区间型、偏离型、偏离区间型指标的量化方法。对于固定型和区间型指标,只有当指标值 a_{ij} 等于固定值 a_j 且落入固定区间[q'_1,q'_2]时,才被认为是最好的属性值,量化后的指标值 $z_{ij}=1$,这一点在环境质量综合评价中体现得非常明显。而偏离型和偏离区间型指标则刚好相反,$a_{ij}=\beta_j$,$a_{ij}\in[p^j_1,p^j_2]\Rightarrow z_{ij}=0$,但在很多情况下,指标值不能达到固定的值或区间,因此量化后的指标值不能等于 1 或 0。而上述公式(1)–(4),对于任意一组给定的指标值,都能使量化后的至少一个指标值达到 1 或 0,原因在于分子或分母中的最大值或最小值都来源于给定的指标值,当指标值恰好为最大值或最小值时,量化后的指标值就为 1 或 0。因此,为体现一般性,最大值(上限)和最小值(下限)不能产生于给定的指标值内部,而应该由专家或根据经验给出。对于固定型和固定区间型指标,超过上限或下限的指标值量化后均为 0,而对于偏离型和偏离区间型指标, 超过上限或下限的指标值均为 1。另外, 所有的 6 个公式均没有考虑分母可能取零的情形,而实际上,当分母取零时,根据指标属性的含义,它们的取值不仅存在而且是有差异的。设已知上限为 p,下限为 q,下面给出较合理的量化公式:

$$z_{ij}=\begin{cases}0 & a_{ij}<q\\ \dfrac{a_j-a_{ij}}{a_j-q} & a_{ij}\in[a_j,q]\\ 1 & a_{ij}=a_j\\ \dfrac{a_{ij}-a_j}{p-a_j} & a_{ij}\in(a_j,P)\\ 0 & a_{ij}>p\end{cases} \quad i\in M,j\in N_3 \qquad (5)$$

$$z_{ij}=\begin{cases}0 & a_{ij}<q\\ \dfrac{q_1^j-a_{ij}}{a_j-q} & a_{ij}\in[q_j,q_1^j]\\ 1 & a_{ij}\in[q_1^j,q_2^j]\\ \dfrac{a_{ij}-q_2^j}{p-a_j} & a_{ij}\in[q_2^j,p]\\ 0 & a_{ij}>p\end{cases} \quad i\in M,j\in N_4 \qquad (6)$$

$$z_{ij}=\begin{cases}1 & (a_{ij}<q)\\ \dfrac{\beta_j-a_{ij}}{\beta_j-q} & a_{ij}\in[q,\beta_j]\\ 0 & a_{ij}=\beta_j\\ \dfrac{a_{ij}-\beta_j}{p-\beta_j} & a_{ij}\in[\beta_j,p]\\ 1 & a_{ij}>p\end{cases} \quad i\in M,j\in N_5 \qquad (7)$$

$$z_{ij}=\begin{cases}1 & (a_{ij}<q)\\ \dfrac{p_1^j-a_{ij}}{p_1^j-q} & a_{ij}\in[q,p_1^j]\\ 0 & a_{ij}\in[p_1^j,p_2^j]\\ \dfrac{a_{ij}-p_2^j}{p-p_2^j} & a_{ij}\in[p_2^j,p]\\ 1 & a_{ij}>p\end{cases} \quad i\in M,j\in N_6 \qquad (8)$$

显然,这样一组量化公式克服了文献[3]中公式的局限,适用于任意给定的指标值。

2 权重的分配

权重向量 W 是复杂系统评估中不能回避的另一个重要参数,对各指标赋权的合理与否直接关系到决策结果的可靠性。目前,指标权重的确定方法主要有主观赋权法和客观赋权法[3]。主观赋权法是根据各指标的主观重视程度进行赋权的一类方法,如专家调查法、二项系数法、层次分析法(AHP)等;客观赋权法是各个指标根据一定的规则进行自动赋权的一类方法,如主成分分析法、熵技术法、多目标规划法等。在上面的权重分配法中,各分目标的轻重程度主要是由人的主观价值观来决定的。因此,客观赋权法确定的权重优势与分目标的实际重要程度相悖;而主观赋权法又存在一定的主观随意性,专家组对各分目标之间的比较、评分也存在较大工作量,最后确定的权重也是折中、调和的产物,很难人人满意。下面给出两种能避免两类赋权法缺憾的方法。

2.1 基于关联度的权重求法

设待选的方案为 m 个, 影响各方案综合评价值的分目标为 n 个。用 x_{ij} 表示方案 i 的第 j 个分目标值,则m个方案的n个分目标值构成矩阵 $X=(x_{ij})_{m\times n}$。

(1)利用前面的量化技术得到消除量纲后的新矩阵 $X'=(x'_{ij})_{m\times n}$,从矩阵X'各列中选出最大值,得到最优对象:

$$S_0^+=(x_{0^+1},x_{0^+2}Lx_{0^+n})$$

(2)从其各列中选出最小值,得到最劣对象:

$$S_0^-=(x_{0^-1},x_{0^-2}Lx_{0^-n})$$

(3)计算各方案与最优对象、最劣对象在各分目标的关联系数,得:

$$\zeta_{0^+i}(j)=\frac{\min\limits_i\min\limits_j \Delta_{0^+ij}+0.5\max\limits_i\max\limits_j \Delta_{0^+ij}}{\Delta_{0^-ij}+0.5\max\limits_i\max\limits_j \Delta_{0^+ij}}$$

$$\zeta(j)=\frac{\min\limits_i\min\limits_j \Delta_{0^-ij}+0.5\max\limits_i\max\limits_j \Delta_{0^-ij}}{\Delta_{0^-ij}+0.5\max\limits_i\max\limits_j \Delta_{0^-ij}}$$

其中 $\Delta_{0^+ij}=|x'_{0^+j}-x'_{ij}|$ 　$\Delta_{0^-ij}=|x'_{0^-j}-x'_{ij}|$ $(i=1,2,\cdots,m,j=1,2,\cdots,n)$

若指标为收益性指标或成本性指标(其它的 4 类指标较为复杂,在此不一一讨论),则有$\min\limits_i\min\limits_j \Delta_{0^+ij}=\min\limits_i\min\limits_j \Delta_{0^-ij}=0$,$\max\limits_i\max\limits_j \Delta_{0^+ij}=\max\limits_i\max\limits_j \Delta_{0^-ij}=1$。

故此时：

$$\zeta_{0^+i}(j)=\frac{0.5}{\Delta_{0^+ij}+0.5}=\frac{1}{2\Delta_{0^+ij}+1}$$

$$\zeta(j)=\frac{0.5}{\Delta_{0^-ij}+0.5}=\frac{1}{2\Delta_{0^-ij}+1}$$

(4)模型的建立与分析。$\zeta_{0^+i}(j)$愈大，表示待评方案与最优方案在分目标j上的关联系数愈大，$\zeta_{0^-i}(j)$则相反，其值愈小，表示待评方案与最优方案在分目标j上的关联系数愈大。

设n个分目标的权重向量为$w=(w_1,w_2,\cdots,w_n)$（待定），则：$f(w)=\sum_{j=1}^{n}w_j^2[1-\zeta_{0^+i}(j)]^2+\sum_{j=1}^{n}w_j^2\cdot\zeta_{0^-i}^2(j)$。对应着方案$i$离最优方案与最劣方案的加权距离的平方和。从距离的意义讲，$f_i(w)$应越小越好。由此建立下列多目标规划模型：

$$\min f(w)=(f_1(w),f_2(w),\cdots,f_m(w))^T$$

$$s.t.\begin{cases}\sum_{i=1}^{n}w_j=1\\ w_j\geq0,j=1,2,\cdots,n\end{cases}\quad s.t.\begin{cases}min\sum_{i=1}^{m}f_i(\omega)\\ \sum_{i=1}^{n}w_j=1\\ w_j\geq0,j=1,2,\cdots,n\end{cases}$$

由于$f_i(w)\geq0(i=1,2,\cdots,m)$，故上述多目标规划可归结为单目标规划：

构造拉格朗日函数：

$$F(w,\lambda)=\sum_{i=1}^{m}\sum_{j=1}^{n}w_j^2\{[1-\zeta_{0^+i}(j)]^2+\zeta_{0^-i}^2(j)\}-\lambda(\sum_{j=1}^{n}w_j-1)$$

令

$$\begin{cases}\frac{\partial F}{\partial w_j}=2w_j\sum_{i=1}^{m}\left\{[1-\zeta_{0^+i}(j)]^2+\zeta_{0^-i}^2(j)\right\}-\lambda=0\\ \frac{\partial F}{\partial\lambda}=\sum_{j=1}^{n}w_j-1=0\end{cases}$$

此线性方程组的系数行列式为：

$$\mathbf{d}=\begin{vmatrix}2u_1&0&\cdots&0&-1\\0&2u_2&\cdots&0&-1\\\cdots&\cdots&\cdots&\cdots&\cdots\\0&0&\cdots&2u_n&-1\\1&1&\cdots&1&0\end{vmatrix}=\prod_{i=1}^{n}u_j\sum_{i=1}^{n}\frac{1}{u_i}$$

式中$u_i=\sum_{i=1}^{m}\left\{[1-\zeta_{0^+i}(j)]^2+\zeta_{0^-i}^2(j)\right\}$

d≠0，故原方程组一定有解。

求解方程组得$\omega_j=\frac{1}{u_j\sum_{i=1}^{n}\frac{1}{u_i}}$，由此可得各指标的权重。

2.2 基于满意度的权重求法

下面所考虑的无信息多指标决策模型同1.1，其指标加权向量为：$W=(w_1,w_2,\cdots,w_m)^T,W^TW=1,W\geq0$。W为未知，仍记规范化后的决策矩阵为$z=(z_{ij})_{m\times n}$，方案$x_i$的综合评价值为：$d_i=\sum_{j=1}^{m}w_jz_{ij},i=1,2,\cdots,m$。

综合评价理想值和负理想值分别为d_i^+、d_i^-，用单目标优化模型定义为：

$$(M_1)\begin{cases}\max d_i^+=\sum_{j=1}^{n}w_jz_{ij}\\ s.t.\sum_{j=1}^{n}w_jz_{ij}=1\\ w_i\geq0\end{cases}\qquad(M_2)\begin{cases}\min d_i^-=\sum_{j=1}^{n}w_jz_{ij}\\ s.t.\sum_{j=1}^{n}w_jz_{ij}=1\\ w_j\geq0\end{cases}$$

μ_i为方案x_i的满意度，若

$$\mu_i=\frac{d_i}{d_i^+-d_i^-}\quad i=1,2,\cdots,m$$

显然满意度μ_i越大越好，为综合评价方案$x_1,x_2,\cdots,x_m$，建立如下多目标优化模型：

$$(M_3)\begin{cases}\max\mu=(\mu_1,\mu_2,\cdots,\mu_m)\\ s.tW^TW=1\\ W\geq0\end{cases}$$

(M_3)显然是非线性多目标规划模型，目前这方面的算法很多，如STEM法、SEMOP法、Geffrion法等[4]，以下探索其解析解法。由于各决策方案之间是公平竞争的，不存在任何偏好关系，可采用等权的线性权和法将其转化为下列单目标最优化模型：

$$(M_4)\begin{cases}\max\sum_{i=1}^{m}\mu_i\\ s.tW^TW=1\\ W\geq0\end{cases}$$

又$\sum_{i=1}^{m}\mu_i=\sum_{i=1}^{m}\frac{d_i}{d_i^+-d_i^-}=\sum_{i=1}^{m}\frac{d_i}{d_i^+-d_i^-}=\sum_{i=1}^{m}\frac{1}{d_i^+-d_i^-}d_i=\sum_{i=1}^{m}k_i\sum_{j=1}^{m}w_jz_{ij}=\sum_{i=1}^{m}\sum_{j=1}^{n}k_iw_iz_{ij}$。

其中$k_i=\frac{1}{d_i^+-d_i^-}$为常数，将目标函数写成向量形式：

$$\sum_{i=1}^{m}\sum_{j=1}^{n}k_iw_jz_{ij}=\sum_{i=1}^{m}k_id_i=K^TD(W)$$

其中

$$K^T=(k_1,k_2,\cdots,k_m)$$

$$D(W)=(d_1,d_2,\cdots,d_m)^T=ZW$$

为求解该单目标模型，作Lagrange函数：

$$L(W,\lambda)=K^TZW=\lambda(W^TW-1)$$

令$\frac{\partial L}{\partial W}=0$，得$Z^TK+2\lambda W=0$。

联立$W^TW=1$，且$W\geq0$解得：

$$W=\frac{Z^TK}{\sqrt{K^TZZ^TK}}\qquad\lambda=-\frac{1}{2}\sqrt{K^TZZ^TK}$$

如果要得到归一化权向量，则有$\overline{W}=\frac{W}{E_{m1}^TW}$，其中$E_{m1}^T=(1,1,\cdots,1)$。

由此得各指标的权重。

3 结论

指标的量化技术及权重的分配是复杂系统评估过程