基于投影寻踪的组合评价方法研究

王　硕，杨善林，胡笑旋
(合肥工业大学计算机网络系统研究所，合肥 230009)

[摘要]　以理想解法和灰色关联度分析法为基础进行组合，得到的组合评价方法能够提高评价结果的有效性。在该组合评价方法的研究中，常规的做法是对指标组合权重的运用主观赋值法，组合的偏好系数、灰色关联度的分辨系数采取人为赋值法，没能充分挖掘组合评价方法的优势。针对此问题，提出用投影寻踪法构建处理组合评价法的优化模型，运用实码加速遗传算法处理该非线性优化问题，所得的指标权重为客观权重，同时得到确定组合评价法组合偏好系数和灰色关联度分辨系数的新方法。实例表明，基于投影寻踪的组合评价方法具有科学性和客观性。

[关键词]　组合评价；投影寻踪法；实码加速遗传算法；分辨系数；组合偏好系数
[中图分类号]　O223　[文献标识码]　A　[文章编号]　1009－1742(2008)08－0060－05

1　前言

评价的目的是对客观事物进行综合分析，其结果应能客观公正地反映事物的发展变化，并为评价组织者、决策者所信服和接受。理想解法（technique for order preference by similarity to ideal solution, TOPSIS），是一种有效的多指标综合评价方法[1,2]。灰色关联度分析法是一种在“小样本、贫信息”下处理不确定性系统的评价方法，在大量领域得到广泛应用[3]。文献[4]通过把理想解法和灰色关联度分析法2种评价方法相结合，建立组合评价方法，以提高评价结果的准确性，对提高单一评价方法的有效性进行了有益的研讨。但文献[4]对指标权重的选定运用主观赋值法，组合的偏好系数、灰色关联度的分辨系数采取人为赋值法，对挖掘组合评价方法的优势不够充分。文献[5,6]研究灰色关联度的分辨系数的优化方法，但优化准则的选取不适应组合评价方法。针对文献[4]组合评价模型，提出用投影寻踪法（projection pursuit algorithm, PPA）建立优化模型，运用实码加速遗传算法（real coded accelerating genetic algorithm, RAGA）来处理该优化问题，得到的指标权重为客观权重，同时给出确定组合评价偏好系数和灰色关联度的分辨系数的新方法。实例研究表明，基于投影寻踪的组合评价方法（PPACEM）具有科学性和客观性。

2　基于灰色关联度和理想解法的组合评价方法

记方案集 $A=\{A_1, A_2, \cdots, A_m\}$，指标集 $T=\{T_1, T_2, \cdots, T_n\}$。方案 A_i 在指标 T_j 下的评价值为 $x_i(j)$，$(i=1, 2, \cdots, m; j=1, 2, \cdots, n)$，决策矩阵 $X[x_i(j)]_{m\times n}$，建模过程包括如下步骤：

Step 1　利用向量归一化方法对决策矩阵作标准化处理

用向量归一化法对决策矩阵作标准化处理，得标准化矩阵

$$Y[y_i(j)]_{m\times n} \tag{1}$$

其中

$$y_i(j)=x_i(j)\left\{\sum_{i=1}^{m}[x_i(j)]^2\right\}^{-1/2}。$$

[收稿日期]　2006－12－02；修回日期　2007－03－29
[基金项目]　国家自然科学基金重点项目资助（70631003）；国家自然科学基金项目资助（70471046）；教育部人文社会科学规划项目（07JA790109）；合肥工业大学科学研究发展基金项目（061103F）
[作者简介]　王　硕（1964－），男，安徽合肥市人，合肥工业大学教授，博士后，研究方向为经济预测与评价

Step 2 计算加权标准化判断矩阵

$$U=[u_i(j)]_{m\times n}=[w_j y_i(j)]_{m\times n} \tag{2}$$

其中 w_j为指标 T_j的权重,($i=1,2,\cdots,m$; $j=1,2,\cdots,n$)。

Step 3 确定理想解和负理想解

理想解

$$U_0^+=\{\max_{1\le i\le m} u_i(j)\mid j\in J^+,\ \min_{1\le i\le m} u_i(j)\mid j\in J^-\}=(u_0^+(1),\ u_0^+(2),\ \cdots,\ u_0^+(j),\ \cdots,\ u_0^+(n)) \tag{3}$$

负理想解

$$U_0^-=\{\min_{1\le i\le m} u_i(j)\mid j\in J^+,\ \max_{1\le i\le m} u_i(j)\mid j\in J^-\}=(u_0^-(1),\ u_0^-(2),\ \cdots,\ u_0^-(j),\ \cdots,\ u_0^-(n)) \tag{4}$$

式中,J^+为效益型指标集合;J^-为成本型指标集合。

Step 4 计算各方案到理想解和负理想解之间的距离:

方案 A_i到理想解距离

$$L_i^+=\left(\sum_{j=1}^{n}[u_i(j)-u_0^+(j)]^2\right)^{1/2},\ (i=1,2,\cdots,m) \tag{5}$$

方案 A_i到负理想解距离

$$L_i^-=\left(\sum_{j=1}^{n}[u_i(j)-u_0^-(j)]^2\right)^{1/2},\ (i=1,2,\cdots,m) \tag{6}$$

Step 5 计算各方案到理想解和负理想解之间的灰色关联度:

方案 A_i与理想方案关于指标 T_j的灰色关联系数

$$e_{ij}^+=[\min_i\min_j|u_0^+(j)-u_i(j)|+\xi\max_i\max_j|u_0^+(j)-u_i(j)|]/[|u_0^+(j)-u_i(j)|+\xi\max_i\max_j|u_0^+(j)-u_i(j)|] \tag{7}$$

其中,ξ为分辨系数,$0\le\xi\le1$。

则方案 A_i与理想方案的灰色关联度为

$$E_i^+=\frac{1}{n}\sum_{j=1}^{n}e_{ij}^+,\ (i=1,2,\cdots,m) \tag{8}$$

方案 A_i与负理想方案关于指标 T_j的灰色关联系数

$$e_{ij}^-=[\min_i\min_j|u_0^-(j)-u_i(j)|+\xi\max_i\max_j|u_0^-(j)-u_i(j)|]/[|u_0^-(j)-u_i(j)|+\xi\max_i\max_j|u_0^-(j)-u_i(j)|] \tag{9}$$

其中,ξ为分辨系数,$0\le\xi\le1$。

方案 A_i与负理想方案的灰色关联度为

$$E_i^-=\frac{1}{n}\sum_{j=1}^{n}e_{ij}^-,\ (i=1,2,\cdots,m) \tag{10}$$

Step 6 归一化各方案到理想解和负理想解距离,归一化各方案到理想解和负理想解灰色关联度。P_i分别代表 L_i^+, L_i^-, E_i^+, E_i^-,其归一化为 $P_{new}=\frac{P_i}{\max\limits_{1\le i\le m}P_i}$ 仍记为 L_i^+, L_i^-, E_i^+, E_i^-。

Step 7 归一化后的各方案到理想解距离和灰色关联度组合、归一化后的各方案到负理想解距离和灰色关联度组合分别为

$$V_i^+=\alpha_1L_i^-+\alpha_2E_i^+,\quad V_i^-=\beta_1L_i^++\beta_2E_i^-,\ (i=1,2,\cdots,m) \tag{11}$$

其中 $\alpha_1+\alpha_2=1$, $\beta_1+\beta_2=1$, $\alpha_1\ge0$, $\alpha_2\ge0$, $\beta_1\ge0$, $\beta_2\ge0$。

Step 8 计算各方案的相对贴近度

方案 A_i的相对贴近度

$$Z_i=V_i^+/(V_i^++V_i^-),\ (i=1,2,\cdots,m) \tag{12}$$

把方案 A_i的相对贴近度 Z_i作为方案 A_i综合评价值。Z_i越大,则方案 A_i越优。

Step 9 构造投影指标函数。在综合投影值时,要求投影值 Z_i的散布特征应为:局部投影点尽可能密集,最好凝聚成若干个点团,而在整体上投影点团之间尽可能散开。投影指标函数可构造为[7,8]

$$Q=S_ZD_Z \tag{13}$$

式中,S_Z为投影值 Z_i的标准差,D_Z为投影值 Z_i的局部密度,即

$$S_Z=\left(\frac{1}{m-1}\sum_{i=1}^{m}(Z_i-\bar{Z})^2\right)^{1/2} \tag{14}$$

$$D_Z=\sum_{i=1}^{m}\sum_{j=1}^{m}(R-r_{ij})u(R-r_{ij}) \tag{15}$$

其中,$\bar{Z}=\frac{1}{m}\sum_{i=1}^{m}Z_i$为 Z_i的均值,S_Z为 Z_i的标准差;R为求局部密度的窗口半径,它的选取既要使包含在窗口内的投影点的平均个数不太少,避免滑动平均偏差太大,又不能使它随着 m的增大而增加太快,R取值可通过试验选取,文献[8,9]的研究表明,R可取为 $0.1S_Z$;距离 $r_{ij}=|Z_i-Z_j|$;$u(t)$为单位阶跃函数,当 $t<0$时其函数值为0,当 $t\ge0$时其函数值为1。

Step 10 优化投影指标函数

通过最大化投影指标函数,确定指标权重组合编号系数、灰色关联度分辨系数,即

$$\max Q=S_ZD_Z$$

$$s.t\sum_{j=1}^{n} w_j=1,\quad \alpha_1+\alpha_2=1,\quad \beta_1+\beta_2=1,$$
$$w_1\geqslant 0,\quad w_2\geqslant 0,\quad \cdots,\quad w_n\geqslant 0,\quad \alpha_1\geqslant 0,\quad \alpha_2\geqslant 0,$$
$$\beta_1\geqslant 0,\quad \beta_2\geqslant 0,\quad 0\leqslant\xi\leqslant 1 \tag{16}$$

这是一个以 $w_1, w_2, \cdots, w_n, \alpha_1, \alpha_2, \beta_1, \beta_2, \xi$ 为优化变量的复杂非线性优化问题,常规优化方法较难处理。RAGA作为一种通用的全局优化方法,用它来求解该问题则十分简便而有效。RAGA算法包括实数编码、RAGA初始化、适应度评价、选择、交叉、变异、演化迭代过程、加速循环等 8个步骤,通过如此加速循环,优秀个体的变化区间将逐步调整和收缩,与最优点的距离将越来越近,直至最优个体的目标函数值小于某一设定值或算法运行达到预定加速(循环)次数,结束整个算法的运行,并把当前群体中最佳个体或优秀个体的平均值指定为 RAGA的结果。在文献[7,9]的研究基础上,结合实际研究问题,RAGA算法中控制参数的配置为:RAGA种群规模为 300,优秀个体数为 10,以个体选择概率值进行交叉变异。RAGA的具体算法参见文献[7,9]。文献[9]通过算法测试表明:RAGA具有处理复杂优化问题的能力,即使存在许多局部最优点也无妨;RAGA调整、压缩搜索区间性能稳健,不易产生早熟收敛;RAGA所需迭代次数相对较少;RAGA能得到高精度的解。

Step 11 方案排序

把由 Step 10得到的 $w_1, w_2, \cdots, w_n, \alpha_1, \alpha_2, \beta_1, \beta_2, \xi$ 代入方案 A_i 的综合评价模型式(12),按照 Z_i 值从大到小的排列次序,得到相应方案的优劣次序。

3 算例

供应商选择的评价指标包括:产品价格 T_1、售后服务 T_2、地理位置 T_3、产品合格率 T_4、新产品开发率 T_5、供应能力 T_6、净资产收益率 T_7、准时交货率 T_8、市场占有率 T_9。有 6种方案可供选择:供应商 $A_1, A_2, \cdots, A_6$。各评价指标的定量化表示以及供应商指标的原始数据如表 1所示[4]。

表 1 供应商选择的评价指标

Table 1 The evaluation indexes on provider choice

	T_1/元	T_2/h	T_3/km	T_4	T_5	T_6	T_7	T_8	T_9
A_1	335	3.2	15	0.80	0.12	230	0.10	0.82	0.13
A_2	268	1.4	37	0.92	0.25	130	0.08	0.96	0.15
A_3	304	1.9	22	0.99	0.09	220	0.14	0.99	0.20
A_4	270	2.0	16	0.98	0.35	180	0.12	0.96	0.21
A_5	310	0.8	26	0.86	0.20	150	0.15	0.80	0.12
A_6	303	2.7	10	0.95	0.19	170	0.16	0.91	0.19

效益型指标集合 J^+ 包括 $T_4, T_5, T_6, T_7, T_8, T_9$,成本型指标集合 J^- 包括 T_1, T_2, T_3。

对原始决策矩阵进行归一化处理得标准化矩阵

$$Y=\begin{bmatrix}
0.457\,0 & 0.607\,6 & 0.269\,0 & 0.355\,3 & 0.225\,3 & 0.511\,7 & 0.318\,6 & 0.368\,0 & 0.311\,7\\
0.365\,6 & 0.265\,8 & 0.663\,5 & 0.408\,6 & 0.469\,4 & 0.289\,2 & 0.254\,9 & 0.430\,9 & 0.359\,6\\
0.414\,7 & 0.360\,7 & 0.394\,5 & 0.439\,7 & 0.169\,0 & 0.489\,5 & 0.446\,1 & 0.444\,4 & 0.479\,5\\
0.368\,4 & 0.379\,7 & 0.286\,9 & 0.435\,3 & 0.657\,2 & 0.400\,5 & 0.382\,4 & 0.430\,9 & 0.503\,4\\
0.422\,9 & 0.151\,9 & 0.466\,2 & 0.382\,0 & 0.375\,6 & 0.333\,7 & 0.477\,9 & 0.359\,1 & 0.287\,7\\
0.413\,4 & 0.512\,6 & 0.179\,3 & 0.422\,0 & 0.356\,8 & 0.378\,2 & 0.509\,8 & 0.408\,4 & 0.455\,5
\end{bmatrix},$$

得到投影指标函数后,再用 RAGA优化由式(16)所确定的优化问题,得最大投影指标函数值为 2.487 234 9,客观权重向量为

$W=$[0.108 395 0.103 259 0.112 678 0.109 994 0.111 738 0.109 713 0.112 958 0.115 328 0.115 950]^T。

同时求得方案到理想解距离和灰色关联度组合的偏好系数分别为 $\alpha_1=0.116\,959$, $\alpha_2=0.883\,041$, $\beta_1=0.079\,555$, $\beta_2=0.920\,445$,分辨系数为 $\xi=0.002\,080$。于是有

$$U=\begin{bmatrix} 0.049540 & 0.062737 & 0.030308 & 0.039084 & 0.025178 & 0.056145 & 0.036000 & 0.042447 & 0.036123 \\ 0.039632 & 0.027447 & 0.074759 & 0.044947 & 0.052455 & 0.031734 & 0.028800 & 0.049694 & 0.041681 \\ 0.044956 & 0.037250 & 0.044451 & 0.048366 & 0.018884 & 0.053704 & 0.050400 & 0.051246 & 0.055574 \\ 0.039928 & 0.039211 & 0.032328 & 0.047878 & 0.073437 & 0.043940 & 0.043200 & 0.049694 & 0.055574 \\ 0.045843 & 0.015684 & 0.052533 & 0.042015 & 0.041964 & 0.036616 & 0.054000 & 0.041411 & 0.033345 \\ 0.044808 & 0.052934 & 0.020205 & 0.046412 & 0.039866 & 0.041498 & 0.057600 & 0.047105 & 0.052796 \end{bmatrix},$$

$$U_0^{+}=[0.039632\ 0.015684\ 0.020205\ 0.048366\ 0.073437\ 0.056145\ 0.057600\ 0.051246\ 0.058353]^T,$$

$$U_0^{-}=[0.049540\ 0.062737\ 0.074759\ 0.039084\ 0.018884\ 0.031734\ 0.028800\ 0.041411\ 0.033345]^T。$$

归一化后的各方案到理想解距离与灰色关联度组合为

$$V^{+}=[0.440429\quad 0.464198\quad 0.815953\quad 1.000000\quad 0.467430\quad 0.837170]^T。$$

归一化后的各方案到负理想解距离和灰色关联度组合为

$$V^{-}=[1.000000\quad 0.962681\quad 0.376911\quad 0.053787\quad 0.669099\quad 0.083498]^T。$$

由式(12),得方案的综合评价值向量为

$$Z=[0.305762\quad 0.325324\quad 0.684029\quad 0.948958\quad 0.411279\quad 0.909307]^T,$$

各方案排列顺序为

$$A_4>A_6>A_3>A_5>A_2>A_1。$$

在文献[4]中,指标权重向量为

$W=[0.11\quad 0.16\quad 0.15\quad 0.07\quad 0.14\quad 0.12\quad 0.10\quad 0.09\quad 0.06]^T$,属于主观赋权。用 PPACEM所得的指标权重是基于样本数据驱动所得到的客观权重,它由数据本身的特征所决定,没有人为因素影响,具有科学性和客观性。

在文献[4]中,各方案到理想解距离和灰色关联度组合的偏好系数皆人为取定为 0.5,缺乏理论依据。基于样本数据驱动求得方案到负理想解距离和方案与理想方案的灰色关联度组合的偏好系数为 $\alpha_1=0.116959$, $\alpha_2=0.883041$,方案到理想解距离和方案与负理想方案的灰色关联度组合的偏好系数为 $\beta_1=0.079555$, $\beta_2=0.920445$,具有科学性和客观性。

在文献[4]中,依据一般的传统做法,灰关联分析的分辨系数取 $\xi=0.5$。ξ值一直是灰色系统中较难优化的变量,文献[5]以实际值与模拟值之间的误差最小为优化目标给出一种分辨系数优化的方法,文献[6]利用差异信息熵理论给出分辨系数优化方法,文献[4~6]确定分辨系数的方法不适应所研究的灰色关联度和理想解法的组合评价方法。依据数据本身的特征,用 PPACEM所得到 $\xi=0.002080$,这在所研究的灰色关联度和理想解法的组合评价方法中具有精确性,同时也建立了关于优化灰分辨系数新的优化目标和求解方法。

文献[4]中方案评价值在(0.4245, 0.6206)内,没有拉开方案档次,不便于方案排序。用 PPACEM所得到方案的评价值在(0.305762, 0.948958)内,方案档次相对拉开较大,有利于优劣排序。

根据 PPACEM所得到的客观权重可以分析评价指标对投影值的贡献率,贡献率越大的指标,对方案的综合评价值影响越大。各指标的贡献率分别为 10.8395%, 10.3259%, 11.2678%, 10.9994%, 11.1738%, 10.9713%, 11.2958%, 11.1738%, 11.5328%, 11.5950%。

4 结语

基于投影寻踪的组合评价方法所得的权重为客观权重,能科学地确定组合的偏好系数和灰关联分析的分辨系数。评价值档次较大拉开,有利于进行选优排序。根据客观权重,可以分析方案指标对方案评价值的贡献率。

参考文献

[1] Wang Shuo, Xu Ruomei, Xie Zhong. A research on internal mechanism of evaluation system[J]. Engineering Sciences, 2005, 3(1): 64-68

[2] 徐泽水. 不确定多属性决策方法及应用[M]. 北京: 清华大学出版社, 2004: 76-79

[3] 刘思峰, 党耀国, 方志耕. 灰色系统理论及其应用[M]. 第三版. 北京: 科学出版社, 2004: 55-61

[4] 孙晓东, 焦玥, 胡劲松. 基于灰色关联度和理想解法的决策方法研究[J]. 中国管理科学, 2005, 13(4): 63-68

[5] 温坤礼, 何明镜, 吴汉雄. 气体绝缘放电预测中的辨识系数之最佳值研究[C]. 灰色系统研究新进展, 1996: 32-36

[6] 张岐山. 基于遗传算法的分辨系数的优化方法[J]. 中国管理

科学，2003，11(5)：76－79
[7] Yang Shanlin, Wang Shuo, Gong Daning. Approach to weighted geometric evaluation based on projection pursuit[J]. Engineering Sciences, 2006, 4(1)：85－88
[8] Friedman JH, Turkey JW. A projection pursuit algorithm for exploratory data analysis[J]. IEEE Trans on Computer, 1974, 23(9)：881－890
[9] 王 硕，张礼兵，金菊良. 系统预测与综合评价方法[M]. 合肥：合肥工业大学出版社，2006, 30－40, 136－145

Combinational evaluation method based on projection pursuit algorithm

Wang Shuo, Yang Shanlin, Hu Xiaoxuan
(Computer Network System Institute of Hefei University of Technology, Hefei 230009, China)

[Abstract] The combinational evaluation method which combines TOPSIS with gray correlation degree analysis can improve validity of the evaluation result. In such method, a usual way to assign combinational weight of index is subjective weighting and to assign combinational preference coefficient and identification coefficient is manual weighting. The advantages of combinational evaluation method are not fully mined. Aim at this problem, a method uses projection pursuit algorithm to establish combinational evaluation model is presented. The real coded accelerating genetic algorithm is used to dispose those non-linear problems. The gained index weights are objective. In the same time, a new method of confirms combinational preference coefficients and identification coefficients of gray correlation degree is presented. The example shows the combinational evaluation method based on projection pursuit algorithm is scientific and objective.

[Key words] combinational evaluation; projection pursuit algorithm; real coded accelerating genetic algorithm; identification coefficient; combinational preference coefficient