On grey relation projection model based on projection pursuit

Wang Shuo, Yang Shanlin, Ma Xijun
(Institute of Computer Network Systems, Hefei University of Technology, Hefei 230009, China)

Abstract: Multidimensional grey relation projection value can be synthesized as one-dimensional projection value by using projection pursuit model. The larger the projection value is, the better the model. Thus, according to the projection value, the best one can be chosen from the model aggregation. Because projection pursuit modeling based on accelerating genetic algorithm can simplify the implementation procedure of the projection pursuit technique and overcome its complex calculation as well as the difficulty in implementing its program, a new method can be obtained for choosing the best grey relation projection model based on the projection pursuit technique.

Key words: grey relation projection model; projection pursuit; real coded accelerating genetic algorithm; identification coefficient; objective weight

1 Introduction

In a comprehensive evaluation, a system that evaluates itself can be treated as a grey system. But the factor indexes (objects) in the decision schemes are not separated. In nature, there exists a kind of grey relation between them, which is unclear but does exist[1]. The comprehensive evaluation to systems is in fact a grey multi-objective decision. In Ref. [2] and Ref. [3], the grey system theory is combined with the vector projection principle to research the method of multi-objective decisions and evaluations, and a grey relation projection model is put forward.

The projection pursuit (PP) technique is an exploring data analysis method that is driven by sample data directly, and it is an objective weighted method that can avoid the artificial influence factors. Its basic way of thinking is to turn the data of high dimension into the data of low dimension by combined projection, adopting a projection index function to describe the possibility of system evaluation structure, looking for the best projection value of the projection index function, and analyzing structural characters of the high dimension data by the projection value[4]. Among them the optimization of the projection index function is much more complicated. The calculations required by traditional methods are huge, and the real coded accelerating genetic algorithm is an effective method to solve the problem[5-7].

In this paper, it is proposed that the grey relation projection problem can be converted into the non-linear optimizing problem, which can be solved by using the real coded accelerating genetic algorithm (RAGA). A case study is also conducted in this paper. The index weight generated is indeed objective weight, and a new method to estimate the grey identification coefficient is developed. The results generated from the case study show that grey relation projection method based on projection pursuit is scientific and objective.

2 Projection pursuit to solve the grey relation projection

Assuming an alternative decision set is $A = \{A_1, A_2, \ldots, A_n\}$, and the attribute set is $V = \{V_1, V_2, \ldots, V_m\}$. The value of an alternative A_i on an attribute V_j is Y_{ij} ($i = 1,2,\ldots,n$; $j = 1,2,\ldots,m$). The process of modeling the projection pursuit to solve the grey relation projection problem is as follows.

Step 1: Determining decision matrix. The optimum decision is represented by A_0, and the attribute values of the optimum decision are given by $Y_0 = (Y_{01}, Y_{02}, \ldots, Y_{0m})$. For a benefit attribute, $Y_{0j} = \max[Y_{1j}, Y_{2j}, \ldots, Y_{nj}]$, and for a cost attribute, $Y_{0j} = \min[Y_{1j}, Y_{2j}, \ldots, Y_{nj}]$. Thus, the decision matrix is given by $Y = (Y_{ij})_{(n+1)\times m}$, ($i = 0,1,2,\ldots,n; j = 1, 2,\ldots,m$), representing the evaluations of a set of alternatives A on a set of attributes V, from which the optimum decision can be found.

Step 2: Initializing decision matrix. Initialization is to generate a new decision matrix by dividing each element of a vector of the matrix by the first element of

Received 9 November 2007

the vector, resulting in a decision matrix whose elements have common initial points and no dimension and are all non-negative. The initialized matrix is formally given by $Y' = (Y'_{ij})_{(n+1)\times m}$ with

$$Y'_{ij} = Y_{ij}/Y_{0j}\ ,$$
$$(i = 0,1,2,...,n;\ j = 1,2,...,m) \quad (1)$$

Step 3: Determining grey relation decision matrix. If Y'_{0j} is a principal sequence and Y'_{ij} is a subsequence, the grey relationship between other decisions and the optimum decisions is given by:

$$F_{ij} = \frac{\min_i \min_j |Y'_{0j} - Y'_{ij}| + \rho \max_i \max_j |Y'_{0j} - Y'_{ij}|}{|Y'_{0j} - Y'_{ij}| + \rho \max_i \max_j |Y'_{0j} - Y'_{ij}|} \quad (2)$$

where ρ is an identification coefficient.

If the attribute weight vector is denoted by $\boldsymbol{W} = [W_1, W_2, ..., W_m]^{\mathrm{T}}$, the grey relation decision matrix is given by

$$\boldsymbol{F} = \begin{bmatrix} W_1 & W_2 & \cdots & W_m \\ W_1 F_{11} & W_2 F_{12} & \cdots & W_m F_{1m} \\ W_1 F_{21} & W_2 F_{22} & \cdots & W_m F_{2m} \\ \vdots & \vdots & & \vdots \\ W_1 F_{n1} & W_2 F_{n2} & \cdots & W_m F_{nm} \end{bmatrix} \quad (3)$$

Step 4: Determining grey relation projection values. The grey relation projection value of the alternative A_i is given by

$$Z_i = \sum_{j=1}^{m} W_j^2 F_{ij} \Big/ \Big[\sum_{j=1}^{m} W_j^2 \Big]^{\frac{1}{2}} \quad (4)$$

Where, W_j is the weight of the attribute V_j, $(i = 1,2,...,n; j = 1,2,...,m)$.

Step 5: Constructing the projection index function. For synthesizing the projection values, the distribution characteristic of the projection value Z_i is required as follows. The local projection points should be as intensive as possible, preferably forming individual point regions, and the point regions should be as widely dispersed as possible in the entire region. So, the scheme index function can be formulated by[4]:

$$Q = S_Z D_Z \quad (5)$$

S_Z is the standard variance of the Z'_is projection value and D_Z is the partial density of the Z'_i s projection value, namely:

$$S_Z = \left[\frac{1}{n-1} \sum_{i=1}^{n} (Z_i - \bar{Z})^2 \right]^{\frac{1}{2}} \quad (6)$$

$$D_Z = \sum_{i=1}^{n} \sum_{j=1}^{n} (R - r_{ij})\, u(R - r_{ij}) \quad (7)$$

In the above equations, $\bar{Z} = \frac{1}{n}\sum_{i=1}^{n} Z_i$ is the average value of Z_i, S_Z is the standard variance of Z_i, and R is the window radius to get the partial density. The selection of R should be not only to ensure that the average number of the projection points in the widow is not too small in order to avoid a large moving average deviation, but also not to make it improve too quickly with the increase of m. The value of R is taken to be $0.1\,S_Z$ [5-7], and the distance to be $r_{ij} = |Z_i - Z_j|$. $u(t)$ is a unit step function, whose value is 0 for $t < 0$, and 1 for $t \geq 0$.

Step 6: Optimizing the projection index function. Through maximizing the projection attribute function, the attribute weights and the identification coefficient can be identified, namely:

$$\max Q = S_Z D_Z$$
$$\text{s.t. } W_j \geq 0,\ \sum_{j=1}^{m} W_j = 1,\ 0 \leq \rho \leq 1 \quad (8)$$

It would be better if decision makers preferences expressed as constraints on weights can be taken into account in the model.

It is a nonlinear optimization problem with the weights $W_1, W_2, ..., W_m$ and the identification coefficient ρ as variables, which is difficult to solve using conventional optimization algorithms due to the strong nonlinearity and possible non-convexity of the objective function. Alternatively, RAGA is simple to use and efficient to find near-optimal solutions for such complicated nonlinear problems. The details about RAGA can be found in Ref. [5-7].

Step 7: Alternative ranking. ρ and W_j generated in step 5 are used in the grey relation projection model for calculating an index for alternative A_i as follows:

$$Z_i = \sum_{j=1}^{m} W_j^2 F_{ij} \Big/ \Big[\sum_{j=1}^{m} W_j^2 \Big]^{\frac{1}{2}}$$
$$(i = 1,2,...,n) \quad (9)$$

The alternatives are ranked on the basis of the magnitude of the index values. An alternative with a large index value is ranked higher than another alternative with a lower index value.

3 A numerical example

Example[3]: The selection of locations for a pumped-storage power station. Table 1 shows the attribute values of the three planning alternative locations: the Qingshi Mountain Ridge, the Buyun Mountain and the Pushi River in Liaoning Province, China.

Table 1 The attribute values of the 3 planning locations for the pumped-storage power station in Liaoning Province

Index	A_1	A_2	A_3
V_1	1 236	1 244	1 132
V_2	432	405	482
V_3	427.3	287.3	337.0

cont.

Index	A_1	A_2	A_3
V_4	12.2	7.7	6.6
V_5	176 ×2/1 064.3	100 ×2/1 064.3	194 ×2/1 064.3
V_6	21.0	21.6	20.7
V_7	6.75	5.25	6.75
V_8	62/176	34/100	80/194
V_9	682.64/176	468.98/100	717.27/194

There are three alternative planning locations to choose from: A_1—Qingshi Mountain Ridge, A_2—Buyun Mountain, and A_3—the Pushi River. The evaluation attributes are defined as follows: V_1 is the power station infrastructure investment in yuan·(kW)$^{-1}$, V_2 is the supporting transmission investment in yuan·(kW)$^{-1}$, V_3 is the largest water head in meter, V_4 is the landscape index, V_5 is the flood control coefficient, V_6 is the system total cost in 10^9 yuan, V_7 is the project period in year, V_8 is the amount of coal saved in ton per year, and V_9 is the amount of construction required in m^3·(kW)$^{-1}$. Only three attributes V_3, V_5 and V_8 are benefit attributes, whilst the rest are all cost attributes.

So the attribute values of the ideal location A_0 are given by

Y_0 = [1 132, 405, 427.3, 6.6, 388, 20.7, 5.25, 0.412, 3.697],

The decision matrix is given by

$$Y = \begin{bmatrix} 1\,132 & 405 & 427.3 & 6.6 & 388 & 20.7 & 5.25 & 0.412 & 3.697 \\ 1\,236 & 432 & 427.3 & 12.2 & 352 & 21.0 & 6.75 & 0.352 & 3.879 \\ 1\,244 & 405 & 287.3 & 7.7 & 200 & 21.6 & 5.25 & 0.34 & 4.689\,8 \\ 1\,132 & 482 & 337.0 & 6.6 & 388 & 20.7 & 6.75 & 0.412 & 3.697 \end{bmatrix}$$

The initialization decision matrix is given by

$$Y' = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0.915\,9 & 0.937\,5 & 1 & 0.541\,0 & 0.907\,2 & 0.985\,7 & 0.777\,8 & 0.854\,4 & 0.953\,1 \\ 0.910\,0 & 1 & 0.672\,4 & 0.857\,1 & 0.515\,5 & 0.958\,3 & 1 & 0.825\,2 & 0.788\,3 \\ 1 & 0.840\,2 & 0.788\,7 & 1 & 1 & 1 & 0.777\,8 & 1 & 1 \end{bmatrix}$$

The projection attribute function is gained. By using RAGA to optimize the optimization problem in Eq.(8), the maximum projection attribute value is gained as 0.120 675, the objective weight vector is $\boldsymbol{W}$ = [0.027 741, 0.000 027, 0.922 951, 0.005 47, 0.023 337, 0.003 908, 0.014 949, 0.001 531, 0.000 087]$^{\mathrm{T}}$, the identification coefficient is ρ = 0.000 001; and the grey relation projection value vector is $\boldsymbol{Z}$ = [0.851 047, 0.000 003 74, 0.000 037 3]$^{\mathrm{T}}$. So the ranking order of the planning locations should be $A_1 > A_3 > A_2$.

In Ref. [3], the attribute weight vector is subjectively given by $\boldsymbol{W}$ = [0.150, 0.150, 0.043, 0.202, 0.050, 0.077, 0.097, 0.202, 0,029]$^{\mathrm{T}}$. The weights shown above are generated from the sample data, which are decided by the feature of the data itself, so it is scientific and objective without artificial factor. However, attribute weights should be the representation of the decision makes preferences. In nature, they should be subjective rather than objective.

In Ref. [3], the grey relation identification coefficient was assumed to be 0.5 using a traditional method. The identification coefficient value of ρ is always a variable which is difficult to choose. ρ = 0.000 001 was estimated according to the feature of the data itself, which is accurately generated using grey relation projection method based on projection pursuit.

In Ref. [3], the evaluation values vector is [0.232 2, 0.273 3, 0.346 4]$^{\mathrm{T}}$, the alternative locations are in the range of (0.232 2, 0.346 4), rather narrow for alternative ranking. The range is (0.000 003 74, 0.851 047) in the paper, which is sufficiently wide to support alternative ranking.

According to the objective weights generated in the paper, the degrees of contributions from the attributes to projection values can be analyzed. In general, attributes with larger degrees of contribution have more influence on the comprehensive evaluation of alternatives. In this example, the contribution degree of the 9 attributes can be calculated by 2.774 1 %, 0.002 7 %, 92.295 1 %, 0.547 %, 2.333 7 %, 0.390 8 %, 1.494 9 %, 0.153 1 %, 0.008 7 %.

4 Conclusions

By using the projection pursuit technique to solve grey relation projection models, objective weight can be obtained and the grey identification coefficient can be determined with scientific basis. The deviation between alternatives is enlarged, which is useful to support alternative ranking, meanwhile the contribution rate of evaluation attributes to projection values can be analyzed according to objective weights generated. This study shows that it is not only feasible but simple that projection pursuit model driven directly by the sample data is applied to grey relation projection model. The result is reliable and objective because the projection value has dispersivity, applicability and operability, which is easy to support decision making and avoid subjectivity.

References

[1] Liu Sifeng, Dang Yaoguo, Fang Zhigeng. Grey System Theory and Its Application (3rd Education) [M]. Beijing: Science Press, 2004:5-8. (in Chinese)
[2] Lu Feng, Cui Xiaohui. Multi-criteria decision grey relation projection method and its application[J]. System Engineering Theory and Practice, 2002, 22(1):103-107.
[3] Fu Qiang. Data Process Method and Its Application to Agriculture [M]. Beijing: Science Press, 2006:27-31.
[4] Friedman J H, Turkey J W. A projection pursuit algorithm for exploratory data analysis[J]. IEEE Trans on Computer, 1974, 23 (9):881-890.
[5] Fu Qiang, Zhao Xiaoyong. Projection Pursuit Model principle and Its Application [M]. Beijing: Science Press, 2006:18-28.
[6] Jin Juliang, Ding Jing. Water Resource System Engineer[M]. Chengdu: Cichuan Science and Technology Press, 2002:162-167.
[7] Wang Shuo, Zhang Libing, Jin Juliang. System Forecasting and Comprehensive Evaluation Method[M]. Hefei: Hefei University of Technology Press, 2006:15-20.

Authors

Wang Shuo, male, born in 1964, now is a professor, PhD, post doctor. Mr. Wang can be reached by E-mail: wangshuo@ah.edu.cn or by phone:0551-2901501(O), 13665516068(m).

Yang Shanlin, male, born in 1948, now is a professor, doctoral tutor. Mr. Yang can be reached by E-mail: slyang@mail.hf.ah.cn or by phone: 0551-2901501(O).

Ma Xijun, female, born in 1958, now is a professor. Ms. Ma can be reached by E-mail: mxinjun@mail.hf.ah.cn or by phone: 0551-2901501(O).

Foundation item: The Key Project of NSFC (No. 70631003), the Liberal Arts and Social Science Programming Project of Chinese Ministry of Education (No. 07JA790109).