KYBERNETIKA — VOLUME *39* (2003), NUMBER 5, PAGES 561 – 568

CONSISTENCY–DRIVEN APPROXIMATION OF A PAIRWISE COMPARISON MATRIX

ESTHER DOPAZO AND JACINTO GONZÁLEZ-PACHÓN

The pairwise comparison method is an interesting technique for building a global ranking from binary comparisons. In fact, some web search engines use this method to quantify the importance of a set of web sites.

The purpose of this paper is to search a set of priority weights from the preference information contained in a general pairwise comparison matrix; i.e., a matrix without consistency and reciprocity properties. For this purpose, we consider an approximation methodology within a distance-based framework. In this context, Goal Programming is introduced as a flexible tool for computing priority weights.

Keywords: ranking theory, pairwise comparison, distance-based methods, goal programming

AMS Subject Classification: 90B50, 90C29, 91B08

1. INTRODUCTION

The pairwise comparison method is a powerful inference tool that can be used as a knowledge acquisition technique for knowledge-based systems. It is useful for assessing the relative importance of several objects, when this cannot be done by direct rating. In fact, this perspective has been recently used for measuring the importance of a web site [1]. The problem of interest is to derive priorities from binary comparisons using matrix algebra, [12]. Several approaches (ordinal, cardinal etc..) have been used for this purpose. We focus on a cardinal framework.

Formally, the problem can be formulated as follows. Let us assume that there are a finite set of objects $X = \{x_1, ..., x_n\}$ and an expert comparing these objects in the form of paired comparisons; i.e. assigning the value $m_{ij} > 0$ by answering the question "between object x_i and x_j, which one is more important and by what ratio?" Then, an $n \times n$ pairwise comparison matrix $M = (m_{ij})_{ij}$ is defined. Using this information, the purpose is to assign a set of numerical weights $(w_1, ..., w_n)$ to the n objects reflecting the recorded quantified judgments. In this way, the assigned value m_{ij} (i.e., matrix M) estimates the ratio of weights given by $w_{ij} = w_i/w_j$ (i.e., matrix $W = (w_{ij})_{ij}$).

From this definition, matrix W has the following properties:

- Reciprocity: $w_{ij}w_{ji} = 1$ for all i, j
- Consistency: $w_{ij}w_{jk} = w_{ik}$ for all i, j, k

In practice, noise and/or imperfect judgements lead to non-consistent and/or non-reciprocal pairwise comparison matrices. The preference weights are difficult to obtain from this kind of matrices. The challenge is to get priority weights from non ideal matrices.

In a multicriteria decision making context, where objects $\{x_1, ..., x_n\}$ are criteria, Saaty [10] proposed one possible solution in his Analytical Hierarchical Process (AHP) method. In AHP, a reciprocal matrix M is replaced by a reciprocal and consistent matrix $C = (c_{ij} = s_i/s_j)_{ij}$, where $(s_1, ..., s_n)$ is the eigenvector associated with the largest eigenvalue of M. However, the idea of using the eigenvector to find weights was first formulated by M. G. Kendall [7] and T. H. Wei [12], in the 1950s. Today, this method has acquired considerable currency because of information retrieval applications.

A distance-based point of view may be adopted to solve the above problem. Now, the problem may be stated as follows: how do we find a reciprocal and consistent matrix, B, that is "as close as possible" to M. Priority weights associated with M are obtained from B. Most of the work using this approach has been based on Euclidean distance, see ([8], [3]).

In this paper, we propose a general l^p-distance framework, where the p-parameter has a preference meaning. In this formulation, the least square problem is a particular case ($p = 2$).

On the other hand, when $p \neq 2$, the underlying optimization problems are sometimes very problematic. Goal Programming (GP) provides an interesting tool for solving these cases. Thus, this problem is equivalent to an Archimedean GP model for metric $p = 1$ and is a Chebyshev GP model for $p = \infty$. This equivalence has been used by [4] in the context of ranking aggregation.

The paper is organized as follows. Section 2 states a general distance-based framework for the approximation problem. Within this setting, a GP model for determining priority weights is presented in Section 3. The ideas presented are illustrated with the help of numerical examples in Section 4. The main conclusions derived from this research and its possible application to information retrieval are included in Section 5.

2. THE OPTIMIZATION PROBLEM IN A DISTANCE–BASED FRAMEWORK

Let $M = (m_{ij})_{ij}$ be the pairwise comparison matrix given by the expert. When M verifies suitable properties (reciprocity and consistency), there exists a set of positive numbers, $\{w_1, \dots, w_n\}$, such that $m_{ij} = w_i/w_j$ for every $i, j = 1, \dots, n$ (see [10]).

However, M does not usually verify these properties because of the existence of noise, imperfect judgements and/or for other psychological reasons. Therefore, the challenge is to search a set of priority weights that synthesize preference information contained in a general pairwise comparison matrix.

The elements of matrix M will be considered as a perturbation of the elements of an ideal matrix W, where reciprocity and consistency are verified. A distance-based framework will be adopted to measure this deviation.

The classical Euclidean distance, used in [8] or [3], is now generalized by an l^p-distance. Thus, the approximation problem can be stated as follows:

$$\min \left[\sum_{i=1}^{n} \sum_{j=1}^{n} |m_{ij} - w_{ij}|^p \right]^{1/p}$$

s.t.

$$\begin{aligned} w_{ij} w_{ji} &= 1 \quad \text{for all } i,j \\ w_{ij} w_{jk} &= w_{ik} \quad \text{for all } i,j,k \\ w_{ij} &> 0 \quad \text{for all } i,j \end{aligned}$$

Note that the first set of constraints is related to conditions of reciprocity of W. Meanwhile, the second set of constraints concerns consistency conditions.

This minimization problem is non-tractable, because there are great number of variables and non-linear constraints involved. We suggest the given value m_{ij} be considered as an estimation of the ratio weights w_i/w_j in the metric p. Because $w_j > 0$ for all j, we assume that

$$w_j m_{ij} - w_i \approx 0.$$

Thus, the problem is stated as the minimization of aggregated residual values $r_{ij} = w_j m_{ij} - w_i$. Note that reciprocity and consistency conditions are implicitly considered in these estimations. This means that we have to deal with only n parameters.

Therefore, the following optimization problem is obtained for metric $p \in [1, \infty)$:

$$\min \left[\sum_{i=1}^{n} \sum_{j=1}^{n} |w_j m_{ij} - w_i|^p \right]^{1/p}$$

s.t.

$$\sum_{i=1}^{n} w_i = 1, \qquad w_i > 0 \quad \text{for all } i$$

Now, the feasible set is defined by normalization and positivity conditions for weights.

For $p = \infty$, the objective function turns into the following expression

$$\min \left[\max_{i,j} |w_j m_{ij} - w_i| \right]$$

In the posed problems, the residual aggregation is affected by the p-parameter. Thus, as $p \in [1, \infty]$ increases, more importance is given to the largest residual values. So, the case $p = 1$ leads to a more robust estimation, whereas the estimation for $p = \infty$ is more sensitive to extreme residual values.

3. GOAL PROGRAMMING FORMULATION

Once the analytical framework has been established, we focus on computing the approximated weights for different metrics p. In this respect, a multi-objective optimization tool, like Goal Programming (GP), provides a flexible and operative tool for managing different p-values.

The optimization problems presented in Section 2 can be reduced to GP formulations considering the relationship between distance function models and mathematical programming (see [9]). Using the change of variable proposed in [2], we introduce the following notation:

$$\begin{aligned} n_{ij} &= \frac{1}{2}\left[|w_i - w_j m_{ij}| + (w_i - w_j m_{ij})\right] \\ p_{ij} &= \frac{1}{2}\left[|w_i - w_j m_{ij}| - (w_i - w_j m_{ij})\right] \end{aligned}$$

Thus, for $p \in [1, \infty)$, the optimization problem is equivalent to the following Archimedean linear GP problem:

$$\min \left[\sum_{i=1}^{n}\sum_{j=1}^{n}(n_{ij} + p_{ij})^p\right]^{1/p} \tag{1}$$

s.t.

$$\begin{aligned} m_{ij}w_j - w_i + n_{ij} - p_{ij} &= 0 \quad \text{for all} \quad i, \\ \sum_{i=1}^{n} w_i = 1, \quad w_i &> 0 \quad \text{for all} \quad i \\ n_{ij}, p_{ij} \geq 0 & \qquad\quad \text{for all} \quad i, j \end{aligned}$$

For $p = \infty$, it can also be shown that the optimization problem is equivalent to the following MINMAX or Chebyshev GP problem ([5]):

$$\min D \tag{2}$$

s.t.

$$\begin{aligned} n_{ij} + p_{ij} &\leq D \quad \text{for all} \quad i, j \\ m_{ij}w_j - w_i + n_{ij} - p_{ij} &= 0 \quad \text{for all} \quad i, j \\ \sum_{i=1}^{n} w_i = 1, \quad w_i &> 0 \quad \text{for all} \quad i \\ n_{ij}, p_{ij} \geq 0 & \qquad\quad \text{for all} \quad i, j \end{aligned}$$

where D is an extra positive variable that quantifies the maximum deviation.

We should note that properties like reciprocity or consistency do not explicitly appear as constraints in these formulations. From a computational point of view, the advantage of these GP formulations is the facility of solving this kind of problems. In fact, these formulations reduce these problems to linear problems that can be solved using the simplex method.

4. A NUMERICAL EXAMPLE

Let $A = \{a_1, a_2, a_3, a_4\}$ be a set of objects in a pairwise comparison problem. Let be cardinal pairwise comparisons over A represented by the matrix M:

$$M = \begin{pmatrix} 1.20 & 2.00 & 0.50 & 3.00 \\ 0.40 & 0.90 & 0.25 & 1.50 \\ 1.50 & 3.00 & 1.00 & 5.00 \\ 0.25 & 0.50 & 0.33 & 1.10 \end{pmatrix}$$

Note that M is not reciprocal and there are elements other than 1 in the main diagonal. Preference weights and their associated ratio-matrix are computed to illustrate the proposed methods. In this example, we focus on the most frequently used metrics, $p = 1$, $p = 2$ and $p = \infty$. The numerical results are compared with the ones obtained by Saaty's eigenvector method.

For $p = 1$, the following goal programming problem is obtained:

Achievement Function:

$$\min \left[\sum_{i=1}^{4} \sum_{j=1}^{4} (n_{ij} + p_{ij}) \right]$$

Goals and Constraints:

$$\begin{array}{ll}
1.20\, w_1 - w_1 + n_{11} - p_{11} = 0 & 0.40\, w_1 - w_2 + n_{21} - p_{21} = 0 \\
2.00\, w_2 - w_1 + n_{12} - p_{12} = 0 & 0.90\, w_2 - w_2 + n_{22} - p_{22} = 0 \\
0.50\, w_3 - w_1 + n_{13} - p_{13} = 0 & 0.25\, w_3 - w_2 + n_{23} - p_{23} = 0 \\
3.00\, w_4 - w_1 + n_{14} - p_{14} = 0 & 1.50\, w_4 - w_2 + n_{24} - p_{24} = 0
\end{array}$$

$$\begin{array}{ll}
1.50\, w_1 - w_3 + n_{31} - p_{31} = 0 & 0.25\, w_1 - w_4 + n_{41} - p_{41} = 0 \\
3.00\, w_2 - w_3 + n_{32} - p_{32} = 0 & 0.50\, w_2 - w_4 + n_{42} - p_{42} = 0 \\
1.00\, w_3 - w_3 + n_{33} - p_{33} = 0 & 0.33\, w_3 - w_4 + n_{43} - p_{43} = 0 \\
5.00\, w_4 - w_3 + n_{34} - p_{34} = 0 & 1.10\, w_4 - w_4 + n_{44} - p_{44} = 0
\end{array}$$

$$w_i > 0 \quad \text{for all } i = 1, 2, 3, 4$$

$$w_1 + w_2 + w_3 + w_4 = 1$$

For $p = \infty$ the following MINMAX or Chebyshev GP model is obtained:

Achievement Function:

$$\min D$$

Goals and Constraints of above model plus:

$$\begin{array}{llll}
n_{11} + p_{11} \le D & n_{21} + p_{21} \le D & n_{31} + p_{31} \le D & n_{41} + p_{41} \le D \\
n_{12} + p_{12} \le D & n_{22} + p_{22} \le D & n_{32} + p_{32} \le D & n_{42} + p_{42} \le D \\
n_{13} + p_{13} \le D & n_{23} + p_{23} \le D & n_{33} + p_{33} \le D & n_{43} + p_{43} \le D \\
n_{14} + p_{14} \le D & n_{24} + p_{24} \le D & n_{34} + p_{34} \le D & n_{44} + p_{44} \le D
\end{array}$$

The numerical results of the above optimization problems are shown in the following table. The solutions of the least square procedure ($p = 2$) and the eigenvector method (EM) are also included.

Method	Weight vector	Ratio-matrix
$p=1$	$\begin{bmatrix} 0.3030 \\ 0.1515 \\ 0.4545 \\ 0.0909 \end{bmatrix}$	$\begin{bmatrix} 1.0000 & 2.0000 & 0.6666 & 3.3333 \\ 0.5000 & 1.0000 & 0.3333 & 1.6666 \\ 1.5000 & 3.0000 & 1.0000 & 5.0000 \\ 0.3000 & 0.6000 & 0.2000 & 1.0000 \end{bmatrix}$
$p=2$	$\begin{bmatrix} 0.2919 \\ 0.1485 \\ 0.4644 \\ 0.0951 \end{bmatrix}$	$\begin{bmatrix} 1.0000 & 1.9657 & 0.6287 & 3.0681 \\ 0.5087 & 1.0000 & 0.3198 & 1.5608 \\ 1.5906 & 3.1266 & 1.0000 & 4.8800 \\ 0.3259 & 0.6406 & 0.2049 & 1.0000 \end{bmatrix}$
$p=\infty$	$\begin{bmatrix} 0.2717 \\ 0.1630 \\ 0.4619 \\ 0.1032 \end{bmatrix}$	$\begin{bmatrix} 1.0000 & 1.6666 & 0.5882 & 2.6315 \\ 0.6000 & 1.0000 & 0.3529 & 1.5789 \\ 1.7000 & 2.8333 & 1.0000 & 4.4736 \\ 0.3800 & 0.6333 & 0.2235 & 1.0000 \end{bmatrix}$
EM	$\begin{bmatrix} 0.2979 \\ 0.1303 \\ 0.4672 \\ 0.1045 \end{bmatrix}$	$\begin{bmatrix} 1.0000 & 2.2849 & 0.6376 & 2.8507 \\ 0.4376 & 1.0000 & 0.2790 & 1.2477 \\ 1.5683 & 3.5834 & 1.0000 & 4.4709 \\ 0.3507 & 0.8015 & 0.2236 & 1.0000 \end{bmatrix}$

In this example, the solutions obtained by different methods (l^1-norm, l^∞-norm, l^2-norm and EM) are closed. The reason is that the matrix M is "very consistent" in the sense of Saaty's consistency index (−0.0293).

To illustrate the effect of the p parameter(l^p-norm) on the solution, let us see what happens if an incorrect value is entered in the above matrix M; for instance, the value of m_{14} changes to 30.0 instead of 3.00. In this case, the following results are obtained:

Method	Weight vector	Ratio-matrix
$p=1$	$\begin{bmatrix} 0.3296 \\ 0.1648 \\ 0.4945 \\ 0.0109 \end{bmatrix}$	$\begin{bmatrix} 1.0000 & 2.0000 & 0.6666 & 30.0245 \\ 0.4999 & 1.0000 & 0.3333 & 15.0118 \\ 1.4999 & 3.0000 & 1.0000 & 45.0364 \\ 0.0333 & 0.0666 & 0.0222 & 1.0000 \end{bmatrix}$
$p=2$	$\begin{bmatrix} 0.3632 \\ 0.1619 \\ 0.4597 \\ 0.0149 \end{bmatrix}$	$\begin{bmatrix} 1.0000 & 2.2418 & 0.7903 & 24.1351 \\ 0.4460 & 1.0000 & 0.3525 & 10.7657 \\ 1.2651 & 2.8363 & 1.0000 & 30.5356 \\ 0.0414 & 0.0928 & 0.0327 & 1.0000 \end{bmatrix}$
$p=\infty$	$\begin{bmatrix} 0.4100 \\ 0.2050 \\ 0.3627 \\ 0.0220 \end{bmatrix}$	$\begin{bmatrix} 1.0000 & 2.0000 & 1.1304 & 18.5729 \\ 0.4999 & 1.0000 & 0.5652 & 9.2862 \\ 0.8846 & 1.7692 & 1.0000 & 16.4298 \\ 0.0538 & 0.1076 & 0.0608 & 1.0000 \end{bmatrix}$
EM	$\begin{bmatrix} 0.5364 \\ 0.0860 \\ 0.3128 \\ 0.0647 \end{bmatrix}$	$\begin{bmatrix} 1.0000 & 6.2361 & 1.7148 & 8.2884 \\ 0.1603 & 1.0000 & 0.2749 & 1.3291 \\ 0.5831 & 3.6364 & 1.0000 & 4.8332 \\ 0.1206 & 0.7523 & 0.2068 & 1.0000 \end{bmatrix}$

We should note that the priority weights obtained from the different procedures are not similar. Accordingly, we would like to point out that the consistency index is -0.4772, lower than in the first case. We find that the incorrect entry has had more influence on the l^∞-solution and on the EM-solution than on the l^1-solution. They even rank objects differently. In fact, the influence of the perturbation on l^p-solutions is greater, the higher p is. These differences in the results may be useful for the analyst to test the reliability of input data.

5. CONCLUSIONS

This paper generalizes the classical least square consistent approximation to a given pairwise comparison matrix, by considering a general l^p-norm. This framework provides an alternative method to the Saaty's classical eigenvector method. Additionally, we propose transforming the associated nonlinear optimization problem into a linear one by introducing a GP formulation. This leads to more flexible tools for computing priority weights from pairwise comparison information.

Other advantages of the new framework are as follows.

The value of p (l^p-norm) can be chosen by considering the final objective of the analysis and the structure of information data. In this way, each l^p-solution has a precise preference meaning. The l^1-norm is useful for obtaining robust approximations that rule out gross data errors or inaccuracies. Meanwhile, the l^∞-norm tries to retain the original DM information by minimizing the maximum data deviation from the model. Middle p-values represent different preservation degrees of the original DM preferences. In this respect, l^1 and l^2 estimates are more robust than those yielded by Saaty's eigenvector method.

The proposed framework is able to handle generalized pairwise comparison matrices without a reciprocity property. This includes the possibility of values different to one in the main diagonal. This provides the user (analyst) with more flexible data acquisition.

Finally, the l^p-closest consistent matrix may be useful for obtaining a technique for data validation in the knowledge acquisition process. The relative distance of the derived ratio-matrix to the input matrix provides a useful validation criterion for the DM judgements. If the relative distance is less than a fixed tolerance level then the priority weights define an acceptable synthesis of the DM preferences; i.e., no revision of judgements/information is necessary. Otherwise, a revision is justified.

ACKNOWLEDGEMENTS

The authors would like to thank Prof. Carlos Romero for his comments. Suggestions raised by reviewers have greatly improved the presentation and accuracy of the paper. Thanks are due to Mrs. Rachel Elliot for reviewing the English.

(Received April 5, 2002.)

REFERENCES

[1] S. Brin and L. Page: The anatomy of a large-scale hypertextual web search engine. Computer Networks and ISDN Systems *30* (1998), 107–117.

[2] A. Charnes and W. W. Cooper: Goal programming and multiple objective optimization: Part 1. European J. Oper. Res. *1* (1977), 39–54.

[3] M. T. Chu: On the optimal consistent approximation to pairwise comparison matrices. Linear Algebra Appl. *272* (1998), 155–168.

[4] J. González-Pachón and C. Romero: Distance-based consensus methods: a goal programming approach. Omega *27* (1999), 341–347.

[5] J. P. Ignizio and T. M. Cavalier: Linear Programming. Prentice-Hall, Englewood Cliffs, N.J. 1994.

[6] R. E. Jensen: An alternative scaling method for priorities in hierarchical structures. J. Math. Psychology *28* (1984), 317–332.

[7] M. G. Kendall: Further contributions to the theory of paired comparisons. Biometrics *11* (1955), 43–62.

[8] W. Koczkodaj and M. Orlowski: Computing a consistent approximation to a generalized pairwise comparisons matrix. Comput. Math. Appl. *37* (1999), 79–85.

[9] C. Romero: Handbook of Critical Issues in Goal Programming. Pergamon Press, London 1991.

[10] T. L. Saaty: The Analytic Hierarchy Process. McGraw-Hill, New York 1980.

[11] T. L. Saaty and L. G. Vargas: Comparison of eigenvalues, logarithmic least square and least square methods in estimating ratios. Math. Model. *5* (1984), 309–324.

[12] T. H. Wei: The Algebraic Foundations of Ranking Theory. Cambridge Univ. Press, Cambridge 1952.

Dr. Esther Dopazo, Facultad de Informática, Universidad Politécnica de Madrid, Campus de Montegancedo s/n, 28660-Boadilla del Monte, Madrid. Spain.
e-mail: edopazo@fi.upm.es

Dr. Jacinto González-Pachón, Facultad de Informática, Universidad Politécnica de Madrid, Campus de Montegancedo s/n, 28660-Boadilla del Monte, Madrid. Spain.
e-mail: jgpachon@fi.upm.es