European Journal of Operational Research 252 (2016) 587–600

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Decision Support

Theoretical estimation of the probability of weight rank reversal in pairwise comparisons

Georgia Dede[a,b,*], Thomas Kamalakis[a], Thomas Sphicopoulos[b]

[a] Department of Informatics and Telematics, Harokopio University, 9 Omirou Street, Tavros, Athens, GR17778, Greece
[b] Department of Informatics and Telecommunications, University of Athens, Panepistimiopolis, Ilisia, Athens GR15784, Greece

ARTICLE INFO

Article history:
Received 21 May 2015
Accepted 14 January 2016
Available online 23 February 2016

Keywords:
Decision processes
Multiple criteria analysis
Decision analysis
Pairwise comparisons

ABSTRACT

Pairwise comparison is a key component in multi-criteria decision making. The probability of rank reversal is a useful measure for evaluating the impact of uncertainty on the final outcome. In the context of this paper the type of uncertainty considered is related to the fact that experts have different opinions or that they may perform inconsistent pairwise comparisons. We provide a theoretical model for estimating the probability of the consequent rank reversal using the multivariate normal cumulative distribution function. The model is applied to two alternative weight extraction methods frequently used in the literature: the geometric mean and the eigenvalue method. We introduce a reasonable framework for incorporating uncertainty in the decision making process and calculate the mean value and cross-correlation of the average weights which are required in the application of the model. The theoretical results are compared against numerical simulations and a very good agreement is observed. We further show how our model can be extended in applications of a full multi-criteria decision making analysis, such as the analytic hierarchy process. We also discuss how the theoretical model can be used in practice where the statistical properties of the uncertainty-induced perturbations are unknown and the only information provided by the pairwise comparison matrices of a small group of experts. The methodology presented here can be used to extend the pairwise comparison framework in order to provide some information on the credibility of its outcome.

1. Introduction

Decision making is the process of choosing a specific course of action among several alternatives and is encountered in many areas of human activity (Altuzarra, Moreno-Jimenez, & Salvador, 2007). In situations where complex decisions need to be taken involving high stakes, it is desirable to base these decisions on the combined opinions of several experts in the field rather than simply rely on the skills and intuition of an individual decision maker. In the context of multi-criteria decision making (MCDM) (Triantaphyllou, 2000; Yager, 2004) the first step is to identify a set of criteria on which the decision should be based and then decide on their importance. The pairwise comparison (PWC) method provides a convenient and reliable means to rank both the criteria involved in the decision making process and the alternative courses of action (Saaty, 2008a). In the context of PWC, instead of letting the experts rank the various criteria or alternatives directly, they compare these criteria in combinations of two. This can reduce the influence of subjective point-of-views associated with eliciting the weights directly.

PWC plays a key role in structured decision making systems and especially in MCDM methods, such as the analytic hierarchy process (AHP) and the analytic network process (ANP) (Saaty, 2008b). In this context, PWC has been extensively used on numerous application areas (Chan et al., 2006; Gerdsri & Kocaoglu, 2007; Huang, Chub, & Chiang, 2008; Lee & Kozar, 2006; Liberatore & Nydick, 2008; Zahedi, 1986). PWC has also been applied on a stand-alone basis or on the context of a different decision making framework (Abildtrup et al., 2006; Dede, Varoutas, Kamalakis, Fuentetaja, & Javaudin, 2010; Doumpos & Zopounidis, 2004; Fan & Liu, 2010; Kok & Lootsma, 1985; Traulsen, Pacheco, & Nowak, 2007). Several contributions have attempted to highlight many methodological aspects of the method (Barzilai, 1997; Boenderb, Graan, & Lootsma, 1989; Choo & Wedley, 2004; Deng, 1999; Kwiesielewicz & Van Uden, 2004; Mamat & Daniel, 2007; Marimin, Umano, Hatono, & Tamura, 2002; Mikhailov, 2004).

One important aspect of PWC is that the final outcome is undermined by uncertainty, which originates from the fact that

* Corresponding author at: Department of Informatics and Telematics, Harokopio University, 9 Omirou Street, Tavros, Athens, GR17778, Greece. Tel.: +30 2109549416.

E-mail addresses: gdede@hua.gr, gdede@di.uoa.gr (G. Dede), thkam@hua.gr (T. Kamalakis), thomas@di.uoa.gr (T. Sphicopoulos).

http://dx.doi.org/10.1016/j.ejor.2016.01.059
0377-2217/

the experts may produce inconsistent pairwise judgments or that their point-of-views can differ. Uncertainty modeling by itself is of paramount importance and has been the subject of many studies (András, 2007; Aull-Hyde, Erdogan, & Duke, 2006; Carmone, Karab, & Zanakis, 1997; Hahn, 2003; Harker, 1987; Hongyi & Kocaoglu, 2008; Klir & Folger, 1988; Saaty & Vargas, 1987; Stam & Silva, 1997; Shunsuke, Obata, & Daigo, 1998; Wang & Dong, 2009; Wang, Dong, & Yan, 2012; Wang, Yeung, & Tsang, 2001; Yager, 2002; Yuan & Shaw, 1995). The probability of rank reversal P_{RR} (Saaty, 2003; Saaty & Vargas, 1984) is often a useful measure for evaluating the impact of uncertainty. We may define the probability of rank reversal theoretically, by considering the weights W_1 ,..., W_N where N is the number of alternatives, calculated by PWC in the case of an infinite group of experts ($M \rightarrow +\infty$). In practice, since the group size M is finite ($M<+\infty$), the calculated weights w_i will in general turn out different than W_i and hence the ranking may also be different thus introducing an uncertainty in the final outcome. This uncertainty may originate in the difference of opinion among the experts which may bias the result compared to a large (asymptotically infinite) group of experts. The experts may also complete the pairwise comparison in an inconsistent manner (Saaty, 2003). The probability of rank reversal can be formally defined as (Dede, Kamalakis, & Sphicopoulos, 2015):

$$P_{RR} = P\{\text{Ranking of } w_i \text{ is different than that of } W_i\} \tag{1}$$

In our previous work (Dede et al., 2015), we have investigated the convergence properties of the probability of rank reversal in PWC with respect to the size M of the group of experts. The results dictated that there is not much to be gained by increasing the number of experts beyond 15, even if the uncertainty level is large. Moreover, we examined the problem of how the probability of rank reversal can be estimated numerically using Monte Carlo (MC) simulations. Although this is a valid approach, it is often preferable to have a theoretical model for estimating the probability of rank reversal. A theoretical model is often much more straightforward to implement and requires much less computational time than MC simulations. It can also form a solid basis for understanding and extending the PWC application framework. In this paper we discuss how the probability of rank reversal P_{RR} can be estimated theoretically. We show that instead of using MC simulations, the P_{RR} can be estimated through the multivariate normal cumulative distribution function (MVNCDF). This approach is formulated for two alternative weight estimation methods: the eigenvalue (EV) (Saaty & Vargas, 1987) and the geometric mean (GM) method (Crawford, 1987). We show that the MVNCDF yields accurate results compared to MC simulations regardless of the number of criteria and the weight estimation method used. This approach seems to be useful for any potential decision maker willing to investigate the accuracy of the results in real decision problems in practice. We also show how the approach can be extended in the application of a full MCDM analysis, such as the AHP.

The rest of the paper is organized as follows: In Section 2, we briefly summarize the PWC method and discuss how one can incorporate uncertainty in the pairwise comparison matrices. In addition, we discuss the statistical nature of the uncertainty-induced weight perturbations in PWCs and calculate the mean values and the correlation matrices of the average weights required in the MVNCDF estimations. We also elaborate how the P_{RR} can be calculated using the MVNCDF and how the results can be applied in AHP. In Section 3, we validate our theoretical model against MC simulations. Section 4 provides some insight on how the model can be applied in a practical situation where the uncertainty parameters are unknown and can only be inferred by the pairwise comparison matrix elements of a limited number of experts. Some concluding remarks and future outlook are presented in Section 5.

Table 1
Nine level scale.

$P_{ij}^{(m)}$	Explanation
1	C_i and C_j are equally important
3	C_i is slightly more important than C_j
5	C_i is strongly more important than C_j
7	C_i is very strongly more important than C_j
9	C_i is absolutely more important than C_j
2,4,6,8	Intermediate values
Reciprocals of above	Used in analogous manner when C_j is more important than C_i

2. Methodology and model description

2.1. The pairwise comparison method

The judgment of the mth expert concerning any two criteria is stored in a square pairwise comparison matrix $\mathbf{P}^{(m)}$ and each matrix element $P_{ij}^{(m)}$ reflects the relative importance of criterion C_i compared to C_j. In order to complete the pairwise comparison matrices, two alternative methods widely used in the literature can be considered. In the first approach (referred to hereafter as method "A"), each expert is required to grade the relative importance between any two criteria C_i and C_j by assigning values between 0 and 100 (Gerdsri & Kocaoglu, 2007). Assuming that $i<j$, the mth expert assigns a value between 0 and 100 in the element $A_{ij}^{(m)}$ of a matrix $\mathbf{A}^{(m)}$ comparing the importance of criterion C_i to criterion C_j. For example, if an expert assigns $A_{ij}^{(m)}=60$, then this implies that the weight of C_i is 60 percent compared to the total weight of both criteria, while that of C_j is 40 percent. Using the upper triangular elements of $\mathbf{A}^{(m)}$, the elements of the pairwise comparison matrix are calculated setting $P_{ij}^{(m)}=A_{ij}^{(m)}/(100-A_{ij}^{(m)})$ for the upper diagonal elements ($i<j$), $P_{ij}^{(m)}=1/P_{ij}^{(m)}$ for the lower diagonal elements ($i>j$) while $P_{ii}^{(m)}$ are set equal to 1. According to the alternative approach (referred to hereafter as method "B"), proposed by Saaty (2008b), the nine-level scale shown in Table 1 is used to carry out the comparisons. Again, one needs to complete only the upper triangular elements ($i<j$), since $P_{ij}^{(m)}=1/P_{ji}^{(m)}$ and $P_{ii}^{(m)}=1$. This type of pairwise matrix completion will be referred as method "B" in the following analysis. In both cases, each experts is expected to carry out a total of $N(N-1)/2$ comparisons.

After the experts have completed the pairwise matrices, one can calculate the weight $w_k^{(m)}$ of criterion C_k according to the mth expert. For the extraction of the weights from a PWC matrix, several methods have been proposed. Two of the most popular ones are the EV and the GM methods. In the EV method the eigenvalues of $\mathbf{P}^{(m)}$ are calculated and the eigenvector $\mathbf{x}_1^{(m)}=[x_{1k}^{(m)}]$ associated with the largest eigenvalue $\lambda_{max}^{(m)}$ is determined (Saaty, 2003). The weight $w_k^{(m)}$ are obtained normalizing the sum of the elements of $\mathbf{x}_1^{(m)}$ to unity,

$$w_k^{(m)} = x_{1k}^{(m)} \left[\sum_{l=1}^{N} x_{1l}^{(m)} \right]^{-1} \tag{2}$$

The PWC matrix is said to be perfectly consistent if all its elements are of the form $P_{ij}^{(m)}=q_i^{(m)}/q_j^{(m)}$, where $q_i^{(m)}$ are positive real numbers. If one assumes that all $q_i^{(m)}$ are normalized so that $\Sigma_i q_i^{(m)}=1$ then it follows that $w_k^{(m)}=q_k^{(m)}$. In practice however, the PWC matrices $\mathbf{P}^{(m)}$ will generally be inconsistent since the experts perform their comparisons without having to conform to any such restrictions. The average weight w_k for each criterion C_k is calculated by averaging out the weights $w_k^{(m)}$ calculated by all