IFAC/IFORS/IIASA/TIMS

The International Federation of Automatic Control
The International Federation of Operational Research Societies
The International Institute for Applied Systems Analysis
The Institute of Management Sciences

SUPPORT SYSTEMS FOR DECISION AND NEGOTIATION PROCESSES

Preprints of the IFAC/IFORS/IIASA/TIMS Workshop

Warsaw, Poland

June 24-26, 1992

Editors:

Roman Kulikowski
Zbigniew Nahorski
Jan W. Owsiński
Andrzej Straszak

Systems Research Institute
Polish Academy of Sciences
Warsaw, Poland

VOLUME 1:

Names of first authors: **A-K**

SYSTEMS RESEARCH INSTITUTE, POLISH ACADEMY OF SCIENCES

SUPPORT SYSTEMS FOR DECISION AND NEGOTIATION PROCESSES
Preprints, IFAC/IFORS/IIASA/TIMS Workshop, June 24-26, 1992, Warsaw, Poland

SOME PROBABILISTIC PROPERTIES OF THE NEAREST ADJOINING ORDER METHOD FOR THE CASE WHEN COMPARISONS ARE NOT INDEPENDENT

Leszek KLUKOWSKI

Systems Research Institute, Polish Academy of Sciences
Newelska 6, 01-447 Warszawa, Poland

ABSTRACT

In the paper some probabilistic properties of the nearest adjoining order method are presented for the case when comparisons of objects may be not independent (in stochastic sense) and probabilities of comparison errors may be not known exactly, i.e. that basic assumptions commonly used in literature are not applied. The results presented contain: evaluation of the probability that the errorless solution is obtained and some asymptotic properties of that errorless solution in the case, when comparisons of pairs of different objects (i.e. pairs (x_i,x_j) and (x_k,x_l) for $k\neq i,j$ and $l\neq i,j$) are uncorrelated.

Keywords: pairwise comparisons, nearest adjoining order method, weak preference relation estimation

1. Introduction

The nearest adjoining order method (NAO) was presented firstly in Slater (1961) and developed by many other authors e.g. Remage and Thomson (1966), Flueck and Korsh (1975); more references are given in David (1988).

The main idea of this method consists in determining such the partition, of the given set of objects, for which the number of inconsistencies with regard to results of comparisons is minimal (it is said that comparison is not consistent with regard to a given partition if the direction of preference in the pair compared is not the same as in this partition).

Properties of the NAO method have been obtained under two basic assumptions: (i) results of comparisons are independent and (ii) probability of the error for each comparison is known - in the case of one comparison for each pair - or alternatively the number of independent comparisons for each pair is greater than one.

In practice, the assumptions (i) and (ii) may be not fullfiled; for example if comparisons result from a statistical test, then they are usually not independent and the probabilities of comparison errors are known very often only approximately. The same may be true when comparisons are made by experts.

The purpose of this paper is to investigate the basic probabilistic properties of the NAO method in the case when both assumptions mentioned are not satisfied. These properties comprise: the evaluation of the probability that the NAO solution is errorless and some asymptotic characteristics of this solution (as number of objects tends to infinity), in the case, when comparisons of different objects are uncorrelated.

2. Formulation of the problem.

The general formulation of the ordering problem can be stated as follows.

Given a finite set of elements $X = \{x_1,\dots,x_m\}$, $m\ge 3$. It is assumed that there exists (but is unknown) a complete, reflexive and transitive preference relation R on X of the form:

$$R = I \cup P \tag{1}$$

where: I - the eqivalence relation and P - the strict preference relation.

The preference relation R generates partition $\chi_1^*,\dots,\chi_n^*$ $(n\le m)$, in which each element $x_i \subseteq \chi_k^*$, is preferred to the element $x_j \subseteq \chi_l^*$, $k<l$ and each of the subsets χ_ν^*, $(1\le\nu\le n)$ includes equivalent elements only.

The relation R can be characterized by the function $T : X \times X \to D$, $D=\{0,\mp1,\dots,\mp(n-1)\}$, defined as follows:

$$\left.\begin{array}{l} T(x_i,x_j) = d \Leftrightarrow x_i \subseteq \chi_k^*,\ \ x_j^* \subseteq \chi_l^*,\ \ d=k-l \\ T(x_i,x_j) = -T(x_i,x_j) \text{ for } T(x_i,x_j) \ne 0 . \end{array}\right\} \tag{2}$$

The preference relation R is to be determined on the basis of pairwise comparisons made by an expert, under following assumptions:

A1. The comparisons are made for each pair $x_i,x_j \subseteq X$ and each comparison points out an element which is preferred or indicates that they are equivalent. Thus the result of a comparison can be described by the function $g : X \times X \to \{-1,0,1\}$, defined as follows:

$$g(x_i,x_j) = \begin{cases} -1 & \text{if the expert preferes } x_i \text{ to } x_j \\ 0 & \text{ - - considers } x_i \text{ and } x_j \text{ as equivalent} \\ 1 & \text{ - - preferes } x_j \text{ to } x_i . \end{cases} \qquad (3)$$

The result of comparison is correct if one of the following conditions is satisfied :

$$\left.\begin{array}{l} \operatorname{sgn} g(x_i,x_j) = \operatorname{sgn} T(x_i,x_j) \\ g(x_i,x_j) = T(x_i,x_j) = 0 . \end{array}\right\} \qquad (4)$$

In opposite case, the result is incorrect.

A2. Incorrect results (errors) are of random nature; for the joint distribution of all comparisons (i.e. for all pairs $(x_i,x_j) \in X$) the following inequalities:

$$\Pr\{[\operatorname{sgn} g(x_i,x_j)=\operatorname{sgn} T(x_i,x_j);T(.)\neq 0] \text{ or } [g(x_i,x_j)=T(x_i,x_j)=0]\} \geq$$
$$\geq 1 - \delta \qquad (5)$$

where: $\delta \in (0, \frac{1}{2})$,

are satisfied .

A3. Results of comparisons $g(x_i,x_j)$ and $g(x_k,x_l)$ are uncorrelated for $k\neq i,j$; $l\neq i,j$:

$$\operatorname{cov}[g(x_i,x_j),g(x_k,x_l)] = 0 . \qquad (6)$$

The essential role in above problem formulation is played by the assumptions A1 and A2, which determine basic features of the probabilistic structure of comparisons. The assumption A3 is in general not neccessary, but it provides valuable properties of the NAO solution. It is also significantly weaker than the assumption on the independency among all the comparisons, commonly used in literature.

3. Definitions and notions.

For a given partition $\chi_1,\ldots,\chi_r$, $(r\leq m)$, being a feasible solution to the formulated problem, i.e. such that the relations:

$$\bigcup_{i=1}^{r} \chi_i = X , \quad \chi_i \cap \chi_j = \{0\}, (j\neq i), \quad \chi_i \neq \{0\}, (i=1,\ldots,r), \qquad (7)$$

are satisfied, the following notation is used:

S_χ - the set of all feasible solutions;

R_χ - the set of all the pairs of indices $\langle i,j \rangle$ for which the conditions :

$$1 \leq i,j \leq m , \quad j > i \qquad (8)$$

hold;

$I(\chi_1,\dots,\chi_r)$, $P_1(\chi_1,\dots,\chi_r)$, $P_2(\chi_1,\dots,\chi_r)$ - subsets of the set R_χ such that:

$$I(\chi_1,\dots,\chi_r) = \{\langle i,j\rangle \mid \exists q, (1\le q\le r) : x_i, x_j \in \chi_q\} ; \quad (9)$$

$$P_1(\chi_1,\dots,\chi_r) = \{\langle i,j\rangle \mid x_i\in \chi_k, x_j\in \chi_l ; k - l < 0\} ; \quad (10)$$

$$P_2(\chi_1,\dots,\chi_r) = \{\langle i,j\rangle \mid x_i\in \chi_k, x_j\in \chi_l ; k - l > 0\} . \quad (11)$$

From (7) - (11) it follows that :

$$I(.) \cup P_1(.) \cup P_2(.) = R_\chi , \quad (12)$$

$$I(.)\cap P_1(.) = \{0\}, \quad I(.)\cap P_2(.) = \{0\}, \quad P_1(.)\cap P_2(.) = \{0\} , \quad (13)$$

$$\mathrm{card}(R_\chi) = M = \tfrac{1}{2}m(m - 1) . \quad (14)$$

The problem of determining the NAO solution can be formulated, using the introduced notation, as the minimization problem of the form:

$$\min\{ \sum_{I(\chi_1,\dots,\chi_r)} |g(x_i,x_j)| + \sum_{P_1(\chi_1,\dots,\chi_r)} h^{(1)}(x_i,x_j) + \\ + \sum_{P_2(\chi_1,\dots,\chi_r)} h^{(2)}(x_i,x_j) \} , \quad (15)$$

where:

$$h^{(1)}(x_i,x_j) = \begin{cases} 0 & \text{if } g(x_i,x_j) = -1 \\ 1 & \text{otherwise} , \end{cases}$$

$$h^{(2)}(x_i,x_j) = \begin{cases} 0 & \text{if } g(x_i,x_j) = 1 \\ 1 & \text{otherwise} , \end{cases}$$

under condition : $\chi_1,\dots,\chi_r \in S_\chi$.

The optimal solution (solutions) of the above problem will be denoted by $\hat\chi_1,\dots,\hat\chi_{\hat n}$ ($\hat\chi_1^{(i)},\dots,\hat\chi_{n_i}^{(i)}$, $i=1,\dots,\nu$; ν - the number of solutions with the same minimal value of the function (15)).

Let us define the random variable $W(\chi_1,\dots,\chi_r)$ for any partition $\chi_1,\dots,\chi_r$ from S_χ, of the form :

$$W(\chi_1,\dots,\chi_r) = \sum_{I(.)} U_{ij}(\chi_1,\dots,\chi_r) + \sum_{P_1(.)} V_{ij}(\chi_1,\dots,\chi_r) + \\ + \sum_{P_2(\,)} Z_{ij}(\chi_1,\dots,\chi_r) \quad (16)$$

where:

$$U_{ij}(.) = \begin{cases} 0 & \text{if } g(x_i,x_j) = 0 \text{ for } \langle i,j\rangle \in I(.) \\ 1 & \text{if } g(x_i,x_j) \neq 0 \text{ for } \langle i,j\rangle \in I(.) \end{cases} \tag{17}$$

$$V_{ij}(.) = \begin{cases} 0 & \text{if } g(x_i,x_j) = -1 \text{ for } \langle i,j\rangle \in P_1(.) \\ 1 & \text{if } g(x_i\ x_j) \geq\ 0 \text{ for } \langle i,j\rangle \in P_1(.) \end{cases} \tag{18}$$

$$Z_{ij}(.) = \begin{cases} 0 & \text{if } g(x_i,x_j) = 1 \text{ for } \langle i,j\rangle \in P_2(.) \\ 1 & \text{if } g(x_i,x_j) \leq 0 \text{ for } \langle i,j\rangle \in P_2(.) . \end{cases} \tag{19}$$

The family of the random variables W(.), generated by the set S_χ will be denoted by $\mathbb{W}$.

It follows from (15) and (16) that the optimal solution of the problem (15) generates random variable $W(\hat{\chi}_1,...\hat{\chi}_n)$, which assumes the minimal value in the family $\mathbb{W}$.

To simplify the notation the variables corresponding to the errorless solution $\chi_1^*,...,\chi_n^*$ will be marked with asterisks, e.g. I^*, P_1^*, P_2^*, U_{ij}^*,V_{ij}^*,Z_{ij}^*,W^* ,while those corresponding to any other solution $\tilde{\chi}_1$,..., $\tilde{\chi}_r$ will be denoted as follows: $\tilde{I}$, $\tilde{P}_1$, $\tilde{P}_2$, etc.

4. Basic theorems.

Theorem 1.

If the assumptions A1 and A2 are satisfied, then for any random variable $\tilde{W}$ from $\mathbb{W}$ the following inequalities hold true:

$$E(W^* - \tilde{W}) < 0 \tag{20}$$

$$Pr(W^* < \tilde{W}) \geq 1 - 2\delta . \tag{21}$$

Proof of this theorem is given in Klukowski and Wagner (1989).

The inequality (20) shows that the variable W^*, corresponding to the solution $\chi_1^*,...,\chi_n^*$, assumes minimal expected value in the family $\mathbb{W}$, while the inequality (21) provides some evaluation (based on Tchebysheff inequality) of the probability (or frequency in large number of trials) of the event $\{W^* < \tilde{W}\}$. In other words the inequality (21) evaluates the frequency of the event that the NAO solution $\hat{\chi}_1,...,\hat{\chi}_n$ is equivalent to the errorless solution; this evaluation is close to one if the value of δ is close to zero.

Theorem 2.

If the assumptions A1 - A3 are satisfied then:

$$\operatorname{var}(W^{*}) \le \frac{1}{2}\, m(m - 1)(2m - 3)\delta(1 - \delta) . \tag{22}$$

Proof of this theorem is given in Klukowski (1990).

Two conclusions important for further considerations can be derived from the inequality (22).

Conclusion 1. Under the assumptions A1 - A3 the variance of the random variable W^{*}/M converges to zero as $m \to \infty$, i.e.

$$\lim_{m \to \infty} \operatorname{var}(W^{*}/M) = 0 . \tag{23}$$

Conclusion 2. Under the assumptions A1 - A2 the variance of any variable $\tilde{W}$ from the family $\mathcal{W}$ satisfies the conditions :

$$\operatorname{var}(\tilde{W}) < \frac{1}{8}\, m(m - 1)(2m - 3) , \tag{24}$$

$$\lim_{m \to \infty} \operatorname{var}(\tilde{W}/M) = 0 . \tag{25}$$

Proofs of the relations (23) - (25) are given in Klukowski (1990).

From (23) and (25) it follows that if the random variable $\tilde{W}$ satisfies the inequality :

$$\lim_{m \to \infty} E\, [\tfrac{1}{M}\, (W^{*} - \tilde{W})] < 0 , \tag{26}$$

then the variable W^{*}/M converges (in stochastic sense) to a limit lower than that corresponding to the variable $\tilde{W}/M$. In this case the probability of the event $\{W^{*} < \tilde{W}\}$ converges to one as $m \to \infty$. It can be shown that the inequality (26) holds for each partition $\tilde{\chi}_1,\ldots,\tilde{\chi}_r$, which satisfies at least one of the following conditions

$$\left.\begin{array}{l} \lim\, [\operatorname{card}(\tilde{I} - I^{*})]/M > 0 , \\ \lim\, [\operatorname{card}(\tilde{P}_1 - P_1^{*})]/M > 0 , \\ \lim\, [\operatorname{card}(\tilde{P}_2 - P_2^{*})]/M > 0 . \end{array}\right\} \tag{27}$$

This fact indicates that for large m the application of the NAO method makes it possible to eliminate systematic errors.

In the case, when all the probabilities of comparisons errors are close to δ, the evaluation (24) can be also used to construct the rough test (based on the Tchebysheff inequality for variance) for the hypothesis that the partition $\hat{\chi}_1,\ldots,\hat{\chi}_{\hat{n}}$ is errorless against the alternative that it cannot be accepted as the solution to the formulated problem.

5. Concluding remarks.

Results, presented in this paper, show that the NAO method provides a reasonable solution of the ranking problem under much weaker assumptions than those commonly used in literature. Especially, it should be emphasized that one can evaluate the probability of determining the errorless solution even if results of comparisons are not independent and probabilities of errors are not known. Moreover if the covariation structure of comparisons satisfies the assumption A3, then the NAO method makes it possible to eliminate systematic errors.

The approach used in this paper can be also applied in the case of $N > 1$ independent comparisons for each pair. The first results on this subject are presented in Klukowski (1990).

Literature.

David H.A. (1988) The Method of Paired Comparisons (sec. ed.), Ch. Griffin & Comp. LTD, London.

Flueck J.A.,Korsh J.F. (1975) A generalized approach to maximum likelihood paired comparisons ranking. Biometrika 61, 621-26.

Klukowski L. (1990) Ranking of alternatives on the basis of pairwise comparisons with random errors (in Polish). In Badania Systemowe t. 3 Podstawy metodologiczne i budowa systemów komputerowych, edited by R. Kulikowski and J. Kacprzyk. Omnitech Press Warszawa, 212-268.

Klukowski L, Wagner D. (1989) Uncertainty in the analytic hierarchy process. Proc. of the 3-rd Polish-finnish Symp. : Methodology and application of decision support systems, edited by R. Kulikowski, IBS PAN Warszawa, 106-120.

Remage R. Jr, Thompson W.A. Jr (1966) Maximum - likelihood paired comparison rankings. Biometrika 53, 143-149.

Slater P. (1961) Inconsistences in a schedule of paired comparisons. Biometrika 48, 303-312.