ELSEVIER

Information Processing Letters 59 (1996) 197-202

Information Processing Letters

Convergence of inconsistency algorithms for the pairwise comparisons

Włodzimierz Holsztyński [1], Waldemar W. Koczkodaj *

Computer Science Department, Laurentian University, Sudbury, Ontario, Canada P3E 2C6

Received 26 February 1996; revised 5 June 1996
Communicated by D. Gries

Abstract

A formal proof of convergence of a class of algorithms for reducing inconsistency of pairwise comparisons (pc) method is presented. The design of such algorithms is proposed. The convergence of the algorithms justifies making an inference that iterated modifications of the pc matrix made by human experts should also converge. This is instrumental for credibility of practical applications of the pc method.

Keywords: Design of algorithms; Analysis of algorithms; Iterative algorithms; Algorithm convergence; Performance evaluation; Local and global triad inconsistency

1. Basics of the pairwise comparisons (pc) method

The pc method, introduced by Thurstone in 1927 [10] is a powerful inference tool and may be used as a knowledge acquisition technique for knowledge-based systems. This paper is a follow-up to [6] and [2]. Reading [3] may also be helpful for understanding why the proof of convergence of the consistency is important for the pc method. The goal of the pc method is to establish the relative preferences of N stimuli in situations in which it is impractical (or sometimes even impossible) to provide estimations of the stimuli by direct rating (see [7,3]). Experts provide pairwise comparison coefficients $a_{ij} > 0$, $i, j = 1, \ldots, N$, as substitutes for the quotients s_i/s_j of the unknown (or even undefined) absolute values of stimuli $s_i, s_j > 0$. The quotients s_i/s_j are also sometimes called *relative weights*.

The coefficients a_{ij} are expected to satisfy some natural restrictions (e.g., $a_{ii} = 1$, $a_{ij} \cdot a_{ji} = 1$; see below). For the sake of our exposition, we consider the pc $N \times N$ matrices simply as square matrices $A = (a_{ij})$ such that $a_{ij} > 0$ for all i, j.

A pc matrix A is called *reciprocal* if $a_{ij} \cdot a_{ji} = 1$ for all i, j (then automatically $a_{ii} = 1$ for all $i = 1, \ldots, N$). A pc matrix A is *consistent* if $a_{ij} \cdot a_{jk} \cdot a_{ki} = 1$ for every $i, j, k = 1, \ldots, N$. While every consistent matrix is reciprocal, the inverse fails in general. Consistent matrices correspond to the ideal situation in which there are exact values $s_1, \ldots, s_N$ for stimuli. The quotients $a_{ij} = s_i/s_j$ form a consistent matrix. Conversely, the theoretical foundation of the pc inference is Saaty's Theorem (see [8]), which states that for every $N \times N$ consistent matrix $A = (a_{ij})$ there exist positive real numbers $s_1, \ldots, s_N$ such that $a_{ij} = s_i/s_j$ for every

* Corresponding author. Email: icci@nickel.laurentian.ca.
[1] Independent consultant.

PII S0020-0190(96)00113-5

$i, j = 1, \dots, N$. The vector $s = (s_1, \dots, s_N)$ is unique up to a multiplicative constant.

2. Approximation of pairwise comparisons matrices

The challenge posed to the pc method comes from the lack of consistency of the pc matrices that arise in practice since they often reflect subjective judgements expressed by human experts. Such judgments are quite often not only imprecise but inconsistent.

Example. Consider comparing four stimuli A, B, C, and D. Suppose we estimate ratios A/B as 2, B/C as 3, and A/C as 5. Evidently something does not "add up" since (A/B) · (B/C) = 2 · 3 = 6 and obviously is not equal to 5 (that is, A/C). With the inconsistency factor 0.17 the above triad marked by values in frames (see below) is the most inconsistent in the entire matrix (reciprocal values below the main diagonal are not shown). A rush judgment may cause us to believe that A/C should indeed be 6, but we do not have any reason to reject the estimation of B/C as 2.5 or A/B as 5/3. After correcting B/C from 3 to 2.5 (an arbitrary decision usually based on additional knowledge gathering) the next most inconsistent triad is (5, 4, 0.7) with the inconsistency factor 0.13. Adjustment of 0.7 to 0.8 makes this triad fully consistent (5 · 0.8 is 4) but another triad (2.5, 1.9, 0.8) has inconsistency 0.05. By changing 1.9 to 2 the entire table becomes fully consistent. (See Fig. 1.)

The above example follows only one path of possible changes while each time there is a continuum (that is the power of the set of all real numbers) of other possibilities since we do not need to achieve a full consistency. This makes the convergency problem mathematically challenging.

Given an inconsistent $N \times N$ matrix A, the theory attempts to provide a consistent $N \times N$ matrix B that differs from A "as little as possible". One solution to this problem was proposed in [8]. Let $s = (s_1, \dots, s_N)$ be the eigenvector of A corresponding to σ, the largest eigenvalue of A. By the theorem of Frobenius σ is unique, positive and simple. Furthermore, s can be chosen to have all coordinates positive.

Let $B = (b_{ij})$ be given by $b_{ij} = s_i/s_j$ for all $i, j = 1$. B is consistent and, moreover, s is an eigenvector of B (corresponding to eigenvalue σ). Sharing this eigenvector with matrix A is an argument in favour of B being a consistent matrix that resembles A the most.

The above procedure of finding a consistent matrix as a replacement for an arbitrary pc matrix takes advantage of the peculiarities of eigenvectors, but the geometric-means solution is simpler to compute [1] and equally good [3]. A standard approach is based on the notion of the best approximation with respect to a distance function. First, introduce a distance function δ in the space of all objects $\mathcal{R}_+^{\mathcal{J}}$, where $\mathcal{J} = \{1, \dots, N\}$ and $\mathcal{R}_+$ is the real half-line of (strictly) positive real numbers. Some of the objects are of special interest. In our application they are the consistent matrices since they form the subspace of "nice objects". The goal is to establish an algorithm that finds a "nice" object B for each object A such that

$$\delta(A, B) \leqslant \delta(A, C)$$

for every "nice" object C. Thus, B minimizes the distance from A to "nice" objects.

At least two different distance functions have been used in the context of pc research. The straight Euclidean distance d in $\mathcal{R}_+^{\mathcal{J}}$ has appeared in the *Least Square Method* (LSM) [9,5]. A different distance δ in $\mathcal{R}_+^{\mathcal{J}}$ seems to provide a deeper insight. Let $\lambda : \mathcal{R}_+^{\mathcal{J}} \to \mathcal{R}^{\mathcal{J}}$ and $\mu : \mathcal{R}^{\mathcal{J}} \to \mathcal{R}_+^{\mathcal{J}}$ be the coordinatewise logarithmic and exponential mappings for $A = (a_{ij})$, $B = (b_{ij})$:

$$\begin{aligned} B = \lambda(A) \ \text{iff} \ b_{ij} = \log(a_{ij}), \\ A = \mu(A) \ \text{iff} \ a_{ij} = \exp(b_{ij}), \end{aligned} \quad \text{for all } i, j.$$

Obviously, the functions λ and μ are inverses. They help to linearize the pc algebra by translating the consistency conditions $a_{ij} \cdot a_{jk} \cdot a_{ki} = 1$ in space $\mathcal{R}_+^{\mathcal{J}}$ into linear conditions $b_{ij} + b_{jk} + b_{ki} = 0$ in linear space $\mathcal{R}^{\mathcal{J}}$. As a result, the image of the (non-linear) subspace of all consistent matrices under mapping λ is a linear subspace of $\mathcal{R}^N$ (and the same is true for the image of the subspace of all reciprocal matrices). Thus it is natural to define distance in $\mathcal{R}_+^{\mathcal{J}}$ as

$$\delta(A', A'') = d(\lambda(A'), \lambda(A'')),$$

where d is the Euclidean distance in $\mathcal{R}^{\mathcal{J}}$.

Distance δ appeared in the *Logarithmic Least Square Method* (LLSM) [5,9]. LLSM solves the

	A	B	C	D
A	1	2	5	4
B		1	3	1.9
C			1	0.7
D				1

	A	B	C	D
A	1	2	5	4
B		1	2.5	1.9
C			1	0.7
D				1

	A	B	C	D
A	1	2	5	4
B		1	2.5	1.9
C			1	0.8
D				1

	A	B	C	D
A	1	2	5	4
B		1	2.5	2
C			1	0.8
D				1

Fig. 1.

approximation problem elegantly. Given pc matrix A, let $B = \lambda(A)$. Then what exists is a unique best approximation $Z \in \mathcal{R}^{\mathcal{J}}$, such that Z is logarithmically consistent (meaning that $Y = \mu(Z)$ is consistent). Indeed, Z is the orthogonal projection of B onto the linear space of all logarithmically consistent matrices in $\mathcal{R}^{\mathcal{J}}$. This means that Y is the unique best consistent approximation of A (with respect to distance δ).

A different approximation problem has arisen in the context of [6,2] in which measuring the inconsistency of comparisons is based on only three stimuli at a time, rather than the whole comparisons matrix. This allows us to locate the inconsistency and therefore control it.

Let A be an inconsistent pc matrix. In the case of LSM, LLSM, and the eigenvector methods mentioned above, one tries to achieve the best consistent approximation of A, where "best" is meant according to the respective mathematical understanding. In practice, the mathematical meaning of "best" as understood from the application point of view may vary from one application to another. Thus it also seems reasonable to take advantage of expert opinions in the process of arriving at a consistent matrix. This process may be done interactively. After the initial matrix $A_0 = A$ is provided by the experts, an auxiliary algorithm pinpoints the most inconsistent triad (or let us say a triple of indices that contribute most to the inconsistency of matrix A). This information is crucial for making corrections to the coefficients a_{ij}, a_{kj}, a_{ki}. Then the new matrix A_1 is analyzed again and a new triad of a_{ij} elements is presented to the experts, and so on. This interactive process takes advantage of the experts' knowledge to the fullest. It is attractive, but one should worry about its convergence. There is no complete mathematical answer to the formulated non-mathematical problem involving an expert's ability to judge some characteristics of stimuli. Nevertheless, it is important to obtain mathematical indications about the plausibility of the interactive algorithm. The next section provides a basis for such an inference.

Remark. The theorems and algorithms presented above act in the space of all pc matrices. One can easily write these results so that they would act in the more restricted space of reciprocal matrices. Indeed, that is what we would do when implementing these results in the form of computer programs. However, for the conceptual presentation it is simpler, more elegant, and also more general to deal with the full space of pc matrices.

3. The convergence of consistency approximations

Let $\mathbf{L}$ be a non-empty finite family of linear subspaces of $\mathcal{R}^N$. Let $\mathbf{L}^\cap$ be the family of all intersections of subfamilies of family $\mathbf{L}$. The separation $sep(\mathbf{L})$ of the family $\mathbf{L}$ is defined (check Appendix A; appendices to this paper can be found at URL http://www.laurentian.ca.www./math.wkoczkodaj.html) as the minimum of separations $d(V, V')$ taken over all pairs $V, V' \in \mathbf{L}^\cap$ such that $d(V, V') > 0$ – but if $d(V, V') = 0$ for every $V, V' \in \mathbf{L}^\cap$ then we define $sep(\mathbf{L}) = 0$. Note that $sep(\mathbf{L}) = 0$, iff $\mathbf{L}$ is ordered linearly with respect to $\subseteq$. For this reason, our considerations about the consistency approximations will be trivially true when $sep(\mathbf{L}) = 0$. Let $p_V(y)$ be the orthogonal projection of point y onto the linear subspace V. The notion of separation $sep(\mathbf{L})$ will help us to prove the following result:

Theorem 1. *Let $\mathbf{L}$ be a non-empty finite family of the linear subspaces of $\mathcal{R}^N$. Let $W = \bigcap \mathbf{L}$ be the intersection of members of $\mathbf{L}$, let $x \in \mathcal{R}^N$ and $y = p_W(x)$. Furthermore, let $w : \{1,2,\ldots\} \to \mathbf{L}$ be a sequence such that for every $V \in \mathbf{L}$ equality $V = w(n)$*

holds for infinitely many different $n = 1, 2, \ldots$ *Define* $x_0 = x$ *and* $x_n = p_{w(n)}(x_{n-1})$ *for every* $n > 0$. *Then* $\lim_{n=\infty} x_n = y$.

Proof. If $sep(\mathbf{L}) = 0$ then $x_n = y$ for all but a finite number of n, so our theorem holds.

Thus from now on we assume $sep(\mathbf{L}) > 0$. Since $x_{n-1} - x_n$ is orthogonal to $x_n - y$ we obtain $|x_{n-1} - y|^2 = |x_{n-1} - x_n|^2 + |x_n - y|^2$ for every $n \in \{1, 2, \ldots\}$. Thus, by induction

$$|x_0 - y|^2 = \sum_{t=1}^{n} |x_{t-1} - x_t|^2 + |x_n - y|^2$$

for every $n \in 1, 2, \ldots$. It follows that the infinite series $\sum_{t=1}^{\infty} |x_{t-1} - x_t|^2 \leqslant |x_0 - y|^2$ is convergent (summable). Hence $\lim_{t=\infty} |x_{t-1} - x_t| = 0$. For every $\delta > 0$ there exists a positive integer $\Lambda(\delta)$ such that $d(x_{n-1}, x_n) < \delta$ for every $n \geqslant \Lambda(\delta)$. The above equalities also show that $d(x_{n-1}, y) \geqslant d(x_n, y)$ for $n = 1, 2, \ldots$ so $d(x_m, y) \geqslant d(x_n, y)$ whenever $m \leqslant n$.

Let $m \in \{0, 1, \ldots\}$. Then there exists an $n \geqslant m$ such that $\bigcap_{t=m}^{n} w(t) = W$.

Indeed, there exists an $n \geqslant m$ such that every linear space $V \in \mathbf{L}$ occurs among the linear spaces $w(m), \ldots, w(n)$ at least once, i.e., $\{w(m), \ldots, w(n)\} = \mathbf{L}$. Fix arbitrary positive integers $m > 0$ and n such that $\Lambda(\delta) \leqslant m < n$ for some arbitrary $\delta > 0$ and define

$$W_t = \bigcap_{k=m}^{t} w(k) \quad \text{and} \quad y_t = p_{W_t}(x_m)$$

for $t = m, \ldots$.

As in earlier similar considerations, we know that:

$$y_t = p_{W_t}(x_t) \quad \text{for } t = m, \ldots, n$$

therefore we need to estimate the distance $d(x_m, W_n)$. First, we estimate $d(x_m, W_t)$ for every $t = m, \ldots, n$. Let us also note that $d(x_m, W_t) = d(x_m, y_t)$ and $d(x_m, W_m) = 0$.

Furthermore, let $m < t \leqslant n$. Three cases need to be considered:

Case A: $W_{t-1} \subseteq w(t)$. Then $W_t = W_{t-1}$; hence $y_t = y_{t-1}$ and $d(x_m, y_t) = d(x_m, y_{t-1})$.

Case B: $W_{t-1} \supset w(t)$ ("strict" containment). Then $W_t = w(t)$; hence $y_t = x_t$ and

$$\begin{aligned} d(x_m, y_t) &= d(x_m, x_t) \\ &\leqslant d(x_m, y_{t-1}) + d(y_{t-1}, x_{t-1}) \\ &\quad + d(x_{t-1}, x_t) \\ &< 2 \cdot d(x_m, y_{t-1}) + \delta. \end{aligned}$$

Case C: $d(W_{t-1}, w(t)) > 0$. By applying the previous section result (regarding the distance from a point to the intersection of two linear subspaces),

$$\begin{aligned} d(x_m, y_t) &\leqslant d(x_m, y_{t-1}) + d(y_{t-1}, x_{t-1}) \\ &\quad + d(x_{t-1}, y_t) \\ &< 2 \cdot d(x_m, y_{t-1}) + \frac{\delta}{\sin(sep(\mathbf{L}))}, \end{aligned}$$

we combine cases B and C into one case BC.

Case BC: W_{t-1} is not contained in $w(t)$. Then $0 < \sin(sep(\mathbf{L})) \leqslant 1$, so

$$d(x_m, y_t) \leqslant 2 \cdot d(x_m, y_{t-1}) + \frac{\delta}{\sin(sep(\mathbf{L}))}.$$

Let $bc(t) = 1$ when case BC holds for index t and let $bc(t) = 0$ otherwise. We obtain by induction

$$d(x_m, y_t) \leqslant \left(2^{\sum_{k=m+1}^{t} bc(k)} - 1\right) \cdot \frac{\delta}{\sin(sep(\mathbf{L}))}.$$

Let $M = \min(N, |\mathbf{L}|) - 1$ where $|\mathbf{L}|$ denotes the number of members of $\mathbf{L}$. Obviously $\sum_{k=m+1}^{n} bc(k) \leqslant M$ hence $d(x_m, y_n) \leqslant (2^M - 1) \cdot \delta / \sin(sep(\mathbf{L}))$.

This essentially proves the required convergence of x_m to y. Indeed, given an arbitrary real $\varepsilon > 0$ let

$$\delta = \frac{\varepsilon \cdot \sin(sep(\mathbf{L}))}{2^M - 1}.$$

Then for every $m > \Lambda(\delta)$ there is an $n > m$ such that $y_n = y$ (by one of the assumptions of the Theorem 1 each $V \in \mathbf{L}$ appears as $w(t)$ for some $t = t_V > m$; thus $y_n = y$ for $n = \max_{V \in \mathbf{L}} t_V$). We can see that: $d(x_m, y) = d(x_m, y_n) < \varepsilon$ for every $m > \Lambda(\delta) = \Lambda(\varepsilon \cdot \sin(sep(bfL))/(2^M - 1))$. $\square$

4. Algorithms for reducing inconsistency

Due to its generality, Theorem 1 assures flexibility in choosing designs of algorithms that approximate the limit point y. We will call such algorithms "correct". To be more specific, by a *correct algorithm* we mean a constructive choice of a sequence $w : \{1,2,\ldots\} \to \mathbf{L}$ such that $\lim_{n\to\infty} x_n = y$, where $x_0 = x$ is the initial point and $x_n = p_{w(n)}(x_{n-1})$ for every $n > 0$ (see Theorem 1).

The easiest way to achieve correctness is by defining sequence w as a periodic sequence that exhausts $\mathbf{L}$. First, we let $w(1), \ldots, w(|\mathbf{L}|)$ enumerate all members of $\mathbf{L}$. We can extend the sequence w indefinitely by the recursion: $w(n) = w(n - |\mathbf{L}|)$ for every $n > |\mathbf{L}|$. Such a sequence w satisfies the assumption of Theorem 1 about each member of $\mathbf{L}$ appearing infinitely many times in $w(1), w(2), \ldots$. By Theorem 1 we have obtained a correct algorithm. In practice, other choices of sequences w might be more efficient, meaning that the corresponding sequence (x_n) may converge faster. Thus let us provide another option, the greedy algorithm.

Assume that $w(1), \ldots, w(n-1) \in \mathbf{L}$ are already prescribed ("computed"), so that $x_0, \ldots, x_{n-1}$ are also given: $x_0 = x$ and $x_k = p_{w(k)}(x_{k-1})$ for $k = 1, \ldots, n-1$. Then $w(n) \in \mathbf{L}$ is a *greedy choice* if:

$$d(x_{n-1}, w(n)) = \max_{V\in\mathbf{L}} d(x_{n-1}, V)$$

and, obviously, we let $x_n = p_{w(n)}$. The (whole) algorithm is *greedy* when all choices are greedy, for $n = 1, 2, \ldots$.

Theorem 2. *The greedy algorithm is correct.*

Proof. Let $x \in \mathcal{R}^n$ and let $\mathbf{L}$ be a finite family of linear subspaces of $\mathcal{R}^n$. Let $w : \{1,2,\ldots\} \to \mathbf{L}$ be a greedy sequence. Let $\mathbf{L}_0$ be the family of all $V \in \mathbf{L}$ that occur in sequence $w : \{1,2,\ldots\}$ infinitely many times. Let α be the first index such that $w(n) \in \mathbf{L}_0$ for every $n > \alpha$. We may then apply Theorem 1 to $x' = x_\alpha$, $\mathbf{L}_0$ and to sequence $v : \{1,2,\ldots\} \to \mathbf{L}_0$ such that $v(n) = w(n+\alpha)$ for every $n = 1, 2, \ldots$. By Theorem 1, $\lim_{n\to\infty} x'_n = y' \in W_0$, where $W_0 = \bigcap \mathbf{L}_0$ and $x'_n = x_{n+\alpha}$ for $n = 1, 2, \ldots$ and $y' = p_{W_0}(x')$. Let $V \in \mathbf{L}$ be arbitrary. We will show that $y' \in V$. Otherwise $d(y', V) > 0$ and there exists an n such that $d(x'_n, y') < d(x'_n, V)$. But $d(x'_n, V') \leqslant d(x'_n, W_0) = d(x'_n, y')$ for every $V' \in \mathbf{L}_0$, so

$$d(x_{n+\alpha}, V) = d(x'_n, V) > \max_{V'\in\mathbf{L}_0} d(x'_n, V') = \max_{V'\in\mathbf{L}_0} d(x_{n+\alpha}, V')$$

which is a contradiction.

Thus we have shown that $y' \in V$ for every $V \in \mathbf{L}$, i.e., $y' \in W = \bigcap \mathbf{L}$. Since $W \subseteq W_0$ we see that $y' = y = p_W(x)$ proving the correctness of the greedy algorithm. $\square$

5. Convergence to consistency

Let $\mathcal{R}_+ = \{x \in \mathcal{R} : x > 0\}$. Let $J = \{1, \ldots, N\}^2$ for a positive integer N. Then $\mathcal{R}_+^{\mathcal{J}}$ is the set of all $N \times N$ pc matrices. Let $\lambda : \mathcal{R}_+^{\mathcal{J}} \to \mathcal{R}^{\mathcal{J}}$ be a mapping such that for $A = (a_{ij}) \in \mathcal{R}_+^{\mathcal{J}}$ and $B = (b_{ij}) \in \mathcal{R}^{\mathcal{J}}$ we have $B = \lambda(A)$ iff $b_{ij} = \log(a_{ij})$ for every $i, j = 1, \ldots, N$. Obviously λ is a bijection (a 1-1 onto mapping) between $\mathcal{R}_+^{\mathcal{J}}$ and $\mathcal{R}^{\mathcal{J}}$. Thus let $\mu : \mathcal{R}^{\mathcal{J}} \to \mathcal{R}_+^{\mathcal{J}}$ be the inverse of λ.

Let $L_{ijk} = \{B \in \mathcal{R}^{\mathcal{J}} \mid b_{ij} + b_{jk} + b_{ki} = 0\}$ for $i, j, k = 1, \ldots, N$, and let

$$\mathbf{L} = \{L_{ijk} \mid i, j, k, = 1, \ldots, N\} \quad \text{and} \quad \Lambda = \bigcap \mathbf{L}.$$

A pc matrix $A \in \mathcal{R}_+^{\mathcal{J}}$ is consistent iff $\lambda(A) \in \Lambda$. From now on we call matrices $B \in \Lambda$ *logarithmically consistent*. Thus A is consistent iff $\lambda(A)$ is logarithmically consistent. The following results are immediate consequences of Theorems 1 and 2.

Theorem 3. *Let $A \in \mathcal{R}_+^{\mathcal{J}}$ and $B = B_0 = \lambda(A)$. Let $w : \{1,2,\ldots\} \to \{1,\ldots,N\}^3$ be such that each triple (i,j,k) appears infinitely many times in $w(1), w(2), \ldots$. Finally, let $Y = p_\Lambda(B)$ and let $B_n = p_{L_{w(n)}}(B_{n-1})$ for $n = 1, 2, \ldots$. Then sequence $A_n = \mu(B_n)$ of pc matrices is convergent to the consistent matrix $X = \mu(Y)$, and* $\lim_{n\to\infty} A_n = X$.

Theorem 4. *Let $A \in \mathcal{R}_+^{\mathcal{J}}$ and $B = B_0 = \lambda(A)$. Let $w : \{1,2,\ldots\} \to \{1,\ldots,N\}^3$ be a sequence of "greedy choices", meaning that w and the sequence of matrices B_n such that $B_n = p_{L_{w(n)}}(B_{n-1})$ for every $n = 1, 2, \ldots$ satisfy the condition $d(B_{n-1}, L_{w(n)}) = \max_{V\in\mathbf{L}} d(B_{n-1}, V)$ for every $n = 1, 2, \ldots$. Furthermore, let $Y = p_\Lambda(B)$, $X = \mu(Y)$ and $A_n = \mu(B_n)$ for $n = 0, 1, \ldots$. Then* $X = \lim_{n\to\infty} A_n$ *is a consistent matrix.*

Remark. With any of the above algorithms, when we pass from B_{n-1} to B_n only three entries of the matrix are modified, corresponding to some pairs of indices $(i,j),(j,k),(k,i)$, these indices i,j,k do not have to be different (even the case $i = j = k$ is allowed). In the restricted domain of the reciprocal matrices, each modification would involve different indices $i \neq j \neq k \neq i$. Furthermore, together with the matrix terms corresponding to the pairs $(i,j),(j,k),(k,i)$, we would also modify terms corresponding to $(i,k),(k,j),(j,i)$, so that the matrices are reciprocal for every step.

6. Conclusions

We are able to find weights for a given pc matrix by various methods. The accuracy of such weights does not depend strongly on the method (e.g., [3]), but there are still some intriguing questions: *Does the solution make any sense? Is it of any use?* Frequently, both questions must be answered *no* because all methods for finding weights mentioned in this paper require and silently assume a certain degree of consistency of judgments, which are often subjective and imprecise. The inconsistency of judgements must be small enough for the weights to have any practical meaning; therefore controlling inconsistency is instrumental for the pc method.

It is encouraging to know that triad-based algorithms for improving consistency are convergent. It also means that theoretical grounds for trust in the pc method have been established. In the past, one could only hope or believe that the results for non-consistent pc matrices were close enough to real values of stimuli. It is often impossible to verify the results since the pc method is used when stimuli cannot be measured.

The refinement of human expert judgments by the interactive process of decreasing inconsistency has been observed during lengthy practical experiments. There was, however, a danger of a cyclic degradation: improving one triad could cause another to deteriorate, which when improved could badly influence the former one. This paper provides mathematical evidence that the process of gradual improvement of local consistency converges. It is quite acceptable from the common sense point of view: global improvement is usually achieved by a number of local improvements. We believe that the theory presented has a much broader application to other situations where local/global issues occur. This is only the tip of the iceberg that we hope will, in time, contribute to moving attention from improving inexact data by themselves to a more constructive approach of stepwise improvement by consistency-driven procedures.

The triad inconsistency definition provides an opportunity to design an algorithm for reducing the inconsistency of the experts' judgments. It should be seen as a technique for data validation in the knowledge acquisition process. The inconsistency measure of a comparisons matrix is the measure of the quality of knowledge.

References

[1] G. Crawford, The geometric mean procedure for estimating the scale of judgement matrix, *Math. Model.* **9** (1987) 327-334.

[2] Z. Duszak and W.W. Koczkodaj, The generalization of a new definition of consistency for pairwise comparisons, *Inform. Process. Lett.* **52** (1994) 273-276.

[3] M. Herman and W.W. Koczkodaj, Monte Carlo study of pairwise comparisons, *Inform. Process. Lett.* **57** (1996) 25-29.

[4] F.R. Gantmacher, *Applications of the Theory of Matrices* (Interscience, New York, 1959).

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

[6] W.W. Koczkodaj, A new definition of consistency of pairwise comparisons, *Math. Comput. Modelling* **18** (1993) 79-84.

[7] W.W. Koczkodaj, Statistically accurate evidence of improved error rate by pairwise comparisons, *Perceptual and Motor Skills* **82** (1996) 43-48.

[8] T.L. Saaty, A scaling methods for priorities in hierarchical structure, *J. Math. Psychology* **15** (1977) 234-281.

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

[10] L.L. Thurstone, A law of comparative judgments, *Psychological Rev.* **34** (1927) 273-286.