Fast Convergence of Distance-based Inconsistency in Pairwise Comparisons

W.W. Koczkodaj · M. Kosiek · J. Szybowski · D. Xu

Received: date / Accepted: date

Abstract This study presents theoretical proof and empirical evidence of the reduction algorithm convergence for the distance-based inconsistency in pairwise comparisons. Our empirical research shows that the convergence very quick. It usually takes less than 10 reductions to bring the inconsistency of the pairwise comparisons matrix below the assumed threshold of 1/3 (sufficient for most applications). We believe that this is the first Monte Carlo study demonstrating such results for the convergence speed of inconsistency reduction in pairwise comparisons.

Keywords pairwise comparisons · distance-based inconsistency · convergence · knowledge management

1 Introduction

Pairwise comparisons (PC) allow us to express preferences more easily and more accurately. These preferences can be highly subjective (e.g., likes/dislikes). Pair-
wise comparisons were most likely used even before numbers were invented. We can easily envision that “weighting” took place during the Stone Age to decide if a fish, in one hand, can be bartered for a bird in another hand. So, it may be one of the oldest scientific methods ever used. Ramon Llull was given credits for discovering the Borda count and Condorcet criterion (Llull winner) in the 13th century after the discovery of his lost manuscripts Ars notandi, Ars eleccionis, and Alia ars eleccionis, in 2001. However, Condorcet published his voting method based on pairwise comparisons in 1785 in [2] and it is generally assumed to be the first documented use of this method. The next formal use of pairwise comparisons is traced to Fechner in 1860 (see [4]).

It is worth stressing the binary nature of pairwise comparisons. Similarly to binary numbers, pairwise comparisons are practically irreducible since comparing one object with itself is not really creative. Empirical software engineering often relies on pairwise comparisons (e.g., the bubble sort) without realizing of their use. In fact, every Ω < condition > then ... else ... construct is a pair of actions to be selected on the basis of the < condition >. So, pairwise comparisons are at the foundations of computer science but using them has been intuitive.

The distance-based inconsistency was introduced in [11]. Its convergence was analyzed in [9]. As a practically “side product”, [9] also provided a proof that limit of the inconsistency reduction is a vector of geometric means. Our results confirm it. This result is of considerable practical importance since inconsistency in pairwise comparisons is undesirable. Reducing it in a systematic sequence of steps is needed for improving the accuracy of data. The distance-based inconsistency was independently analyzed in [1] by practical experimentations. The above publication also stressed that only the distance-based inconsistency localizes the inconsistencies. In [12], a mathematical (existential) proof of convergence has been provided for the distance-based inconsistency. We also need to point that only optimization methods can approximate the given matrix for the assumed norm (e.g., LSM for the Euclidean distance), was recently proposed in [6]) with the new inconsistency indicator. No empirical study has ever been done and this is the first publication showing how rapid this convergence is in practice. Sometimes, the inconsistency indicator is called a “measure” but it is obviously not a measure in the sense of the mathematical theory.

2 The inconsistency reduction

A distance-based adjective has been used by other researchers for the new inconsistency introduced in 1993 in [11]. The distance-based adjective reflects the nature of the inconsistency indicator, which is defined as a minimal distance from the nearest consistent triad in a a pairwise comparisons matrix (PC matrix) \( A \) defined as:

\[
A = \begin{pmatrix}
1 & a_{12} & \cdots & a_{1n} \\
\frac{1}{a_{12}} & 1 & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{1}{a_{1n}} & \frac{1}{a_{2n}} & \cdots & 1
\end{pmatrix}
\]
The inconsistency was studied intensively in the past in [5, 8, 10]. The consistency condition \( a_{ij} = a_{ik} \cdot a_{kj} \) was analyzed in [10] where it was probably introduced for the first time. Assuming that \( a_{ik}, a_{kj}, a_{ij} \) represent ratios of entities \( E_i/E_k, E_k/E_j, E_i/E_j \), the consistency condition simply states that:

\[
\frac{E_i}{E_j} = \frac{E_i}{E_k} \cdot \frac{E_k}{E_j}
\]

According to [14]: “We may assume that when the inconsistency indicator shows the perturbations from consistency are large, the result is unreliable and the information available cannot be used to derive a reliable answer.” Unfortunately, the eigenvalue-based inconsistency definition, provided by Saaty in [14], tolerates an approximation error of any arbitrarily large value. It was evidenced by two counter-examples and mathematical reasoning in [13] where a simple axiomatization for constructing inconsistency indicators, independent of any approximation method, was proposed.

In data and knowledge processing, the importance of inconsistency analysis is expressed by the popular adage GIGO (garbage in – garbage out). GIGO summarizes well what has been known for a long time: processing “dirty data” cannot guarantee meaningful results. The distance-based inconsistency allows us to localize the most inconsistent triad (or triads). It is the maximum of all triads \((a_{ik}, a_{ij}, a_{kj})\) of elements of \(A\) (say, with all indexes \(i, j, k\) distinct) of their inconsistency indicators, which in turn are defined as:

\[
ii = \min(|1 - \frac{a_{ij}}{a_{ik}a_{kj}}|, |1 - \frac{a_{ik}a_{kj}}{a_{ij}}|) . \tag{1}
\]

It has been recently simplified to:

\[
ii = 1 - \min(\frac{a_{ij}}{a_{ik}a_{kj}}, \frac{a_{ik}a_{kj}}{a_{ij}}) . \tag{2}
\]

For a triad \((x, y, z)\), the pseudocode is even simpler:

\[
ii = 1 - \min(xy/y, y/x/z) . \tag{3}
\]

where \(x > 0, y > 0, z\) and \(ii = 0\) for \(y = xz\).

When a triad \((x, y, z)\) with the maximal inconsistency localized, we modify values of \(x, y, \text{or } z\) to make the replaced triad consistent. This method was first described in [11], extended in [3], analyzed [12] and finally simplified in [13]. The inconsistency axiomatization is proposed in [13].

2.1 The convergence of the triad reduction method

To show the method more explicitly, let us consider instead of each matrix:

\[
A = \begin{pmatrix}
1 & a_{12} & \cdots & a_{1n} \\
\frac{1}{a_{12}} & 1 & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{1}{a_{1n}} & \frac{1}{a_{2n}} & \cdots & 1
\end{pmatrix}
\]
the skew symmetric matrix

\[
\log A = \begin{pmatrix}
0 & \log a_{12} & \cdots & \log a_{1n} \\
-\log a_{12} & 0 & \cdots & \log a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
-\log a_{1n} & -\log a_{2n} & \cdots & 0
\end{pmatrix}
\]

Denote by \( M(a_{ik}, a_{ij}, a_{kj}) \) the set of logarithms of all matrices which are consistent with respect to a given triad \((a_{ik}, a_{ij}, a_{kj})\). It follows from the consistency condition that \( M(a_{ik}, a_{ij}, a_{kj}) \) is a linear subspace of the space of all skew symmetric matrices. Moreover, the intersection of all such subspaces is equal to the subspace of all skew symmetric matrices which are logarithms of consistent matrices.

Consider a step of the algorithm which is, for a given triad \((a_{ik}, a_{ij}, a_{kj})\), an orthogonal projection of \( \log A \) onto \( M(a_{ik}, a_{ij}, a_{kj}) \) with respect to some inner product.

Assume \( a_{ik}a_{kj} \neq a_{ij} \). We are looking for a value \( v \) such that a new triad

\[
\tilde{a}_{ik} = (1 + v)a_{ik}, \quad \tilde{a}_{ij} = (1 + v)^{-1}a_{ij}, \quad \tilde{a}_{kj} = (1 + v)a_{kj},
\]

is consistent. It gives us:

\[
(1 + v)^3 = \frac{a_{ij}}{a_{ik}a_{kj}}.
\]

hence

\[
1 + v = a_{ik}^{-1/3}a_{kj}^{-1/3}a_{ij}^{1/3}
\]

The above is equivalent to:

\[
\tilde{a}_{ik} = a_{ik}^{2/3}a_{kj}^{-1/3}a_{ij}^{1/3},
\]

\[
\tilde{a}_{ij} = a_{ik}^{1/3}a_{kj}^{1/3}a_{ij}^{2/3},
\]

\[
\tilde{a}_{kj} = a_{ik}^{-1/3}a_{kj}^{2/3}a_{ij}^{-1/3},
\]

which is the same result as by the orthogonal projections. Such algorithm applied sequentially to a sequence of triads is convergent, as stated in [9] where the proof was incomplete. A complete proof was finally provided in [12].

Let us notice that if \( v \) above is a small value (taking place for a relatively small \( ii \)) then

\[
(1 + v)^{-1} \approx 1 - v.
\]

Thus, we may replace \( \tilde{a}_{ij} \) with \( \bar{a}_{ij} \) given by formulas

\[
\bar{a}_{ik} = (1 + c)a_{ik}, \quad \bar{a}_{ij} = (1 - c)a_{ij}, \quad \bar{a}_{kj} = (1 + c)a_{kj}.
\]

Then, consistency of this new triad leads us to a quadratic equation for \( c \). The particular formulas will be presented in the next subsection.
On the other hand, from the consistency of this triad we get
\[
\frac{(1 + c)^2}{1 - c} = \frac{a_{ij}}{a_{ik}a_{kj}}.
\] (9)

Comparing (5) and (9), we obtain the equality
\[
(1 + v)^3 = \frac{(1 + c)^2}{1 - c}.
\] (10)

Thus,
\[
(1 + v)^3 - \frac{(1 + c)^3}{1 - c^2} > (1 + c)^3.
\]

If \(v > 0\), then
\[
0 < c < v.
\] (11)

If \(-1 < v < 0\), then from (10) we have
\[
(1 + v)^3 \leq (1 + c)^2.
\]

From the Bernoulli inequality,
\[
1 + c \geq (1 + v)^\frac{3}{2} \geq 1 + \frac{3}{2}v,
\]

which implies that
\[
0 > c \geq \frac{3}{2}v.
\] (12)

From (11) and (12) we get
\[
|c| \leq \frac{3}{2}|v|,
\]

which is precise form of the assertion (7). Hence, if we have two sequences \(v_n\), \(c_n\) connected by the equality (10), then the condition \(v_n \to 0\) implies \(c_n \to 0\).

By Theorem 1 of [9], the method of orthogonal projections is convergent which implies \(v_n \to 0\). Consequently, \(c_n \to 0\) which means that the method of quadratic equation is also convergent.

This theoretical proof of convergence of possibly infinite number of the triad reduction is supported by the presented empirical experimentation.

2.2 The distance-based inconsistency reduction algorithm

Inconsistency analysis allows us to locate the most inconsistent triad. In practice, we change only one value in a triad. Depending on the application, it may take days or even weeks to call an expert panel, gather data, analyze it, and make a decision about which value should be altered. In our experimentation, we modify all three values: \(a_{ik}\), \(a_{kj}\), and \(a_{ij}\). This is done by splitting the total modification to three elements of a triad by minimizing the affect of the modification on the initial PC matrix. For it, let us assume that the most inconsistent triad is \((a_{ik}, a_{ij}, a_{kj})\).

According to equation (1):
\[
ii = 1 - \min\left(\frac{a_{ik}a_{kj}}{a_{ij}}, \frac{a_{ij}}{a_{ik}a_{kj}}, \frac{a_{ij}}{a_{ik}a_{kj}}\right)
\]
To make this triad consistent ($ii = 0$), three variables (let us say $\Delta ik, \Delta kj, \Delta ij$) are added to each entry in this triad. The following equations can be obtained when $\Delta ik, \Delta kj, \Delta ij$ meet above requirements:

If $ai_ikak_j < aij$ then
\[(aik + \Delta ik)(akj + \Delta kj) = (aij - \Delta ij),\] (13)
where values:
$\Delta ik, \Delta kj, \Delta ij$
are positive.

If $ai_ikak_j > aij$ then
\[(ai_k - \Delta ik)(ak_j - \Delta kj) = (aij + \Delta ij),\] (14)
where values:
$\Delta ik, \Delta kj, \Delta ij$
are positive.

By assigning three values to $\Delta ik, \Delta kj$ and $\Delta ij$ respectively, the triad will be fully consistent. We assign values to $\Delta ik, \Delta kj$ and $\Delta ij$ according to the formulas
Hence, we can come to the following:
\[\Delta ik = ai_kc, \Delta kj = ak_jc, \Delta ij = aijc\] (15)
where $c$ is a positive constant.

By combining equations (13), (14) and (15), we can get the following quadratic polynomials:
\[ai_ikak_jc^2 + (aij + 2ai_ikak_j)c + ai_ikak_j - aij = 0,\] (16)
for $ai_ikak_j < aij$, and
\[ai_ikak_jc^2 - (aij + 2ai_ikak_j)c + ai_ikak_j - aij = 0,\] (17)
for $ai_ikak_j > aij$.

By solving equations (16) and (17), $c$ can be obtained and then all $\Delta ik, \Delta kj, \Delta ij$ can be determined. The discriminant for both equations is the same and is equal to:
\[(aij + 2ai_ikak_j)^2 - 4(ai_ikak_j - aij)ai_ikak_j =\]
\[= a_i^2 + 8ai_ikak_j > 0,\]
and it implies that each of the both equations has exactly two solutions $c_1$ and $c_2$.

Furthermore, when $ai_ikak_j < aij$, from the Vieta’s formulas, the roots of equation (16) satisfy
\[c_1c_2 = \frac{ai_ikak_j - aij}{ai_ikak_j} < 0,\]
which implies that only one of them is positive. In this case, we take the positive root as the solution.

When $a_{ik}a_{kj} > a_{ij}$ the product of roots of equation (17) is given by the same formula, so it is positive. At the same time, again from the Vieta’s formulas,

$$c_1 + c_2 = \frac{a_{ij} + 2a_{ik}a_{kj}}{a_{ik}a_{kj}} > 0,$$

the above inequality and the positivity of the product imply that both of them are positive. In this case we take the smaller value as the answer to this triad. If we took the bigger one, $a_{ik} - \Delta_{ik}$ and $a_{kj} - \Delta_{kj}$ would be negative.

Our computations clearly indicate that the convergence takes place in fewer steps than we have anticipated it. As such, it is “good enough” (also known as “satisfying”) for practical applications of pairwise comparisons. As opposed to optimal decisions, satisfying, a portmanteau of satisfy and suffice, is a decision-making strategy that attempts to meet an acceptability threshold. In our case, this threshold for inconsistency has been assumed (as a heuristic) to be 1/3. It is worth noticing that “a satisfying strategy” may often be (near) optimal if the costs of the decision-making process itself are considered as a part of the objective function. By “costs”, we understand not the financial problem but “other aspects” related to solving our decision problem. It may vary from obtaining the complete information (usually, an impossible task) to assessing the impact of our decision on “public safety” or “public acceptance”.

2.3 An example of the quadratic inconsistency reduction

Let us assume we have a triad (4.2, 1.8, 0.7). This triad is inconsistent since $4.2 \times 0.7 = 2.94 > 1.8$, but $a_{ij} = 2.94$ makes this triad consistent. However, this may cause bigger changes in other triads. After all, the change of $a_{ij}$ from 1.8 to 2.94 is a relatively significant change (nearly doubled). An improvement is expected by finding the corresponding variables $\Delta_{ik}, \Delta_{kj}, \Delta_{ij}$ from this equation: $(4.2 - \Delta_{ik})(0.7 - \Delta_{kj}) = (1.8 + \Delta_{ij})$

Let us verify that values of $\Delta_{ik}, \Delta_{kj}, \Delta_{ij}$ are not significantly affecting another triad (or triads). According to equation (14), we get:

$$\Delta_{ik} = 4.2c \ \Delta_{kj} = 0.7c \ \Delta_{ij} = 1.8c$$

By solving this equation:

$$(4.2 - 4.2c) \cdot (0.7 - 0.7c) = (1.8 + 1.8c)$$

we get $c_1 \simeq 2.454, c_2 \simeq 0.158$.

According to the earlier discussion, we take $c_2$ as our solution. Therefore, we have:

$$\Delta_{ik} = 0.66$$
$$\Delta_{kj} = 0.11$$
$$\Delta_{ij} = 0.28$$

hence:
\[ a_{ik} = 4.2 - 0.66 = 3.54 \]
\[ a_{kj} = 0.7 - 0.11 = 0.59 \]
\[ a_{ij} = 1.8 + 0.28 = 2.08 \]

and the new triad is \((3.54, 2.08, 0.59)\).

2.4 Not-so-inconsistent matrices

Using completely PC matrices for testing has very little scientific merits since they are just random numbers and as such defy any principles of learning (machine or natural). Common sense dictates the use of somehow inconsistent matrices but not just a proverbial “bunch of random numbers”. We will call such a PC matrix ”not-so-inconsistent” (NSI) PC matrix. NSI matrix was defined in [7] as follows. We obtain NSI PC matrix \(M\) from a random vector \(v\) with positive coordinates by:

\[ M = \left[ \frac{v_i}{v_j} \right] \]

where \(i, j = 1, 2, \ldots, n\). We deviate \(M\) randomly by random multipliers \(m_{ij} := m_{ij} \cdot \text{rand}()\).

Our computing results demonstrate that the quadratic inconsistency reduction algorithm can efficiently reduce the global inconsistency of a “Not-so-inconsistent” (NSI) PC matrix to a certain threshold value (1/3 is usually considered as the acceptable inconsistent level for most applications). NSI PC matrices are not totally random. Totally random matrices have nothing in common with PC matrices. NSI PC matrices are slightly deviated from what we call PC matrices. The initial PC matrix is not expected to be fully consistent. Solving real-life problems usually involves inconsistent assessments. However, a matrix with large inconsistency is undesirable according to “garbage in, garbage out (GIGO)” principle. Inconsistencies often reflect assessing “every criterion being more important than another”.

The concept of an NSI PC matrix was introduced in [7] by the first author of this study. Monte Carlo experiments in [7] demonstrated (on the basis of 1,000,000 cases) no statistical difference between the geometric means and eigenvalue methods of computing weights. A randomly selected deviation was applied to elements of a fully consistent matrix rendering it inconsistent. The same method is also used in this study. For an inconsistency to occur, a minimum size of 3 for PC matrix is required since at least one triad needs to exits. Needless to say that for two comparisons, inaccuracy (not inconsistency) takes place. We use \(n = 7\) as the maximal PC matrix size. For a matrix with \(n\) elements, there are \(n(n - 1)/2\) comparisons. It gives us 21 comparisons for \(n = 7\) and it is a psychological limit for most respondents to cooperate (we wonder who would agree to compare 100 objects giving 4950 pair combinations?).

3 The relationship of deviation and maximal inconsistency

We produce not-so-inconsistent (NSI) PC matrices by using a random deviation \(\Delta > 0\). For \(\Delta = 0\), the PC matrix, generated from a random vector with positive coordinates, is fully consistent. By increasing \(\Delta\), the inconsistency of the PC matrix is also expected to increase. In order to examine the relationship between \(\Delta\)
and maximal inconsistency, we follow this item:

1. Generate random PC matrices,
2. Adjust the deviation of each matrix from 0 to 0.5 with increasing 0.0005 each iteration,
3. Record the maximal inconsistency of 1,000 matrices for each deviation,
4. Compute the average maximal inconsistency of 1,000 matrices for each deviation.

The NSI PC matrix is obtained by:

1. Randomize a vector (say $v$ with coordinates $v_i$)
2. Generate the fully consistent PC matrix (say $A$) by $a_{ij} = v_i / v_j$.

Fig. 1 shows the histogram of inconsistency in a NSI PC matrix generated by adding to each element of a consistent matrix a deviation randomly generated from $[0, 0.5]$. It looks like a normal distribution.

Fig. 1 Histogram of Inconsistency
Fig. 2 shows the result of the relation between deviation and inconsistency.

As we can see from Fig. 2, the maximal inconsistency increases with the deviation. It is nearly linear (but not quite) dependency for 1,000 generated NSI PC matrices. The maximal inconsistency is still below 0.7, since the deviation was not significantly high (between 0 and 0.5)

4 The quadratic inconsistency reduction

In order to test the convergence of quadratic inconsistency reduction method, we:

1. Generate random 7 by 7 NSI PC matrices as described above.
2. Add a random deviation for each entry in the upper PC matrix triangle.
3. Record the maximal inconsistency of each PC matrix.
4. Count the number of triads with the inconsistency larger than 1/3.
5. Count iterations needed to reduce the maximal inconsistency to a maximal value of 1/3.
Fig. 3 shows the histogram of numbers of iterations needed to reduce the inconsistency to equal or less than 1/3. To analyze the data in details, we compute the average number of iterations for maximal inconsistency for each NSI PC matrix.

The relationship between maximal inconsistency and average number of iterations needed to bring the inconsistency under the required level of acceptance 1/3 is shown by Fig. 4. It is encouraging to see that not more than seven iterations are needed.

Fig. 5 shows the the number of iterations needed to bring the inconsistency under the required level of acceptance 1/3 for the given number of inconsistent triads.

The statistical evidence shows that the number of iterations actually depends more on the number of triads with an inconsistency larger than 1/3.

5 Conclusions

We have generated 1,000 NSI PC matrices with ranks ranging from 4 by 4 to 7 by 7. The convergence rate was rapid. Bringing matrices to an inconsistency below 1/3 takes place usually in no more than 10 iterations, for the worst randomly generated case. The inconsistency reduction problem in pairwise comparisons is one of the most fundamental problem. Simply, it is unreasonable to expect an accurate output from an inaccurate input as the adage illustrates by: “garbage-in,
garbage-out" (GIGO). The inconsistency should be reduced whenever we are able to do so and fortunately, it can be done in not so many steps. For this reason, results of this study may be considered as essential for the PC research.

Acknowledgements Authors would like to thank R. Szwarc, Wroclaw University, Poland for invaluable help in the crucial theoretical aspects of this study.

The authors would also like to express appreciation to: T. Kakiashvli, MD, Amanda Dion-Groleau (Laurentian University, Psychology), and Grant O. Duncan (a part-time graduate student at Laurentian University; Team Lead, Business Intelligence, Integration and Development, Health Sciences North, Sudbury, Ontario, Canada) for the editorial improvements.

References

Fig. 5 The relationship between the number of triads with inconsistency greater than 1/3 and the average number of iterations.


