

On the extraction of weights from pairwise comparison matrices

Theo K. Dijkstra

Published online: 29 June 2011

© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract We study properties of weight extraction methods for pairwise comparison matrices that minimize suitable measures of inconsistency, ‘average error gravity’ measures, including one that leads to the geometric row means. The measures share essential global properties with the AHP inconsistency measure. By embedding the geometric mean in a larger class of methods we shed light on the choice between it and its traditional AHP competitor, the principal right eigenvector. We also suggest how to assess the extent of inconsistency by developing an alternative to the Random Consistency Index, which is not based on random comparison matrices, but based on judgemental error distributions. We define and discuss natural invariance requirements and show that the minimizers of average error gravity generally satisfy them, except a requirement regarding the order in which matrices and weights are synthesized. Only the geometric row mean satisfies this requirement also. For weight extraction we recommend the geometric mean.

Keywords AHP · Positive reciprocal matrix · Principal right eigenvector · Geometric mean · Functional equations · Invariance · Synthesizing

1 Introduction

In many multi-criteria decision problems people have found it very useful to impose a hierarchy on clusters of the aspects or dimensions of the problem, and to deter-

T. K. Dijkstra (✉)
Economics and Business, University of Groningen,
PO BOX 800, 9700 AV Groningen,
The Netherlands
e-mail: t.k.dijkstra@rug.nl

mine the relative importances clusterwise. Saaty is an outspoken proponent of this approach, which he helped develop in a sequence of publications. See e.g. Saaty (1980, 2005, 2006) for extensive discussions. When combined with a particular method for the determination of weights, the principal right eigenvector of the comparison matrix (more below), the approach is known as the Analytic Hierarchy Process (AHP). The AHP has been applied in an impressively large variety of situations, and has been the subject of numerous publications and dissertations.¹

The heart of the AHP consists of the construction of pairwise comparison matrices. With a pairwise comparison matrix for n items the decision maker indicates how much more important (or how much more suitable, or how much better qualified, or or how much more likely, or whatever the basic comparison mode is) item i is then item j . The decision maker is explicitly required to make more pairwise comparisons than absolutely necessary. Although in principle $n - 1$ comparisons would suffice, $\frac{1}{2}n(n - 1)$ are required: the comparison of item 1 to items 2, 3, ..., n , of item 2 to items 3, 4, ..., n , up to and including item $n - 1$ to item n . The redundancy yields in practice useful information, given the observed difficulty in generating completely consistent value comparisons. The relative importance as expressed in words is translated into numbers, using scales that appear to work well in practice (although other scales have been proposed and applied as well, this is not the topic of the present paper): 1 for equal importance, 3 for moderate importance, 5 for strong, 7 for very strong and 9 for extreme importance; integers in between can be used for refinements, and reciprocals for the inverse judgements. The result is a *positive reciprocal n by n matrix* $A = (a_{ij})$ with ones on the diagonal and $a_{ji} = 1/a_{ij}$. Now if for each i, j , and k it is the case that $a_{ik} \cdot a_{kj} = a_{ij}$ the matrix is called consistent: if i is twice as important as k , which on its turn is three times as important as j , item i is presumably six times as important as j , et cetera.

When there exists a positive n -vector w , the *priority or weight* vector, such that for all indices $a_{ij} = w_i \div w_j$ then the matrix is trivially consistent. (The weight vector is unique apart from a normalization; we will usually impose that the product of its elements equals one). Conversely, consistency entails the existence of a weight vector: if g_i is the geometric mean of the i th-row of A we have

$$g_i^n = a_{i1} \cdot a_{i2} \cdots a_{in} = (a_{ij} \cdot a_{j1}) \cdot (a_{ij} \cdot a_{j2}) \cdots (a_{ij} \cdot a_{jn}) = a_{ij}^n \cdot g_j^n \quad (1)$$

and so $a_{ij} = g_i \div g_j$ and therefore $w = g$ with the right normalization automatically since the product of all elements of a positive reciprocal matrix is trivially 1.

Many methods will yield the weight vector in the consistent case. They include taking simply the row sums, but the principal right eigenvector will do as well. For the practically more relevant case, where A is inconsistent, Saaty proposes to always let the latter represent the weights. In fact, in Saaty (2005, section 2–4), he claims that the principal right eigenvector is the *only plausible candidate for representing*

¹ See e.g. for an overview http://en.wikipedia.org/wiki/Talk:Analytic_Hierarchy_Process#Practical_Examples and for an elaborate example see: http://en.wikipedia.org/wiki/Talk:Analytic_Hierarchy_Process/Example_Leader.

priorities derived from a positive reciprocal near consistent pairwise comparison matrix (bold in the original). It is fair to say that this has been a very controversial position, see e.g. Saaty (1980, 2005, 2006) and its many references, further Lootsma (1999), Barzilai and Golany (1994), Barzilai (1997), et cetera for some lively discussions. The present paper is yet another contribution, attempting to help settle the matter by defining a suitable class of inconsistency measures and imposing reasonable invariance properties.

Saaty’s main argument (Saaty 2005, section 2–4) appears to be the following. For a fixed item i and a fixed integer k the product $a_{i,t_1} \cdot a_{t_1,t_2} \cdots a_{t_{k-1},j}$ where the t ’s are arbitrary choices from $\{1, 2, \dots, n\}$, is a k -step estimate of the dominance of item i over item j . It would equal a_{ij} in case of consistency. The products are summed across all choices of t , yielding the ‘dominance of i over j along paths of length k ’. Summing across $j \in \{1, 2, \dots, n\}$ yields the ‘dominance of i along paths of length k ’. Dividing the result for fixed item i by the sum of the dominances along paths of length k of all items, gives the corresponding relative dominance of item i along paths of length k . The limit of this ratio as $k \rightarrow \infty$ is the i -th component of the principal right eigenvector, normalized to add up to one.

However, to see this result as proof of the ‘plausibility’ of the eigenvector approach requires one to accept the additive averaging of products as the most ‘natural’ operation. If instead one would take the multiplicative mean of the products, Saaty’s process would lead to the geometric row mean of A :

$$\begin{aligned} & \prod_j \left(\prod_{t_1} \prod_{t_2} \cdots \prod_{t_{k-2}} \prod_{t_{k-1}} a_{i,t_1} \cdot a_{t_1,t_2} \cdots a_{t_{k-2},t_{k-1}} \cdot a_{t_{k-1},j} \right) \\ &= \prod_j \left(\prod_{t_1} \prod_{t_2} \cdots \prod_{t_{k-2}} (a_{i,t_1} \cdot a_{t_1,t_2} \cdots a_{t_{k-3},t_{k-2}})^n \cdot g_{t_{k-2}}^n \div g_j^n \right) \\ &= \prod_j \left(\prod_{t_1} \prod_{t_2} \cdots \prod_{t_{k-3}} (a_{i,t_1} \cdot a_{t_1,t_2} \cdots a_{t_{k-5},t_{k-4}})^{n^2} \cdot g_{t_{k-3}}^{n^2} \div g_j^{n^2} \right) \\ &= \cdots = \prod_j \left(g_i^{n^{k-1}} \div g_j^{n^{k-1}} \right) \\ &= g_i^{n^k} \tag{2} \end{aligned}$$

using $\prod_i g_i^n = \prod_{i,j} a_{ij} = 1$ repeatedly and the fact that the geometric column mean is one over the corresponding geometric row mean. So the geometric mean of the daunting product of n^k terms is just g_i . We do not have to introduce ‘relative dominance’ and no limiting process is necessary, since the result is valid for every value of k .

In Saaty (2005, section 5–13) a related argument for the principal right eigenvector is offered: if one uses w to weight the judgements in each row of A and then takes the sum, one should recover the weights back proportionately. This is

clearly the defining property of an eigenvector. But again, if instead of adding the $a_{ij} \cdot w_j$'s we multiply them, we find that g is the unique fixed point of the equations $\prod_j a_{ij} \cdot w_j = w_i^n, i = 1, 2, \dots, n$:

$$\prod_j a_{ij} \cdot g_j = g_i^n. \quad (3)$$

The approaches just described generate weights without an explicit minimization of an inconsistency measure. Saaty, see below, derived such a measure by noting that a positive reciprocal matrix is consistent if and only if its largest eigenvalue equals the number of items to be compared. However, choosing weights that minimize Saaty's induced inconsistency measure does not as a rule (re-)produce the eigenvector. We would suggest instead to consistently integrate the choice of weights and the inconsistency measure. It is well-known that the geometric row mean optimizes a suitable inconsistency measure. We will introduce a family of reasonable measures, all sharing essential global properties with Saaty's measure, and use invariance considerations to help select 'the favourite'.

An outline of the paper is as follows. In Sect. 2 we continue to review and discuss important aspects of the AHP. Saaty introduced perturbations (or 'errors' as we will call them) relative to a weight vector, and an inconsistency measure as a function of the perturbations. The average value of Saaty's inconsistency measure for random positive reciprocal comparison matrices is recommended in the AHP approach as a basis for judging whether the observed inconsistency is too large. We argue instead for the use of an appropriate error distribution. We also offer a statistical motivation for the geometric row mean.

In Sect. 3 we will embed Saaty's inconsistency measure in a suitable class of inconsistency measures, defined basically as averages of the errors mapped on an 'error gravity scale'. The geometric row mean is generated by minimization of one of them. We note a few useful consequences related to the assessment of the extent of inconsistency, and how and where to adjust the comparison matrix. We also establish that minimization of Saaty's inconsistency measure yields the principal right eigenvector only when its elementwise reciprocal is the principal left eigenvector. The first non-trivial case, with $n = 3$, is analyzed in some detail: here all weight extraction methods concur.

In Sect. 4 we discuss some invariance and consistency requirements whose naturalness appears hard to argue against. Some have to do with scale-inversion, others with synthesizing judgements and the order in which operations are performed. We characterize reasonable synthesization methods by the requirement that the synthesis of consistent judgements ought to be consistent as well. All methods based on minimization of the average error gravity have the specified invariance properties except the property that the order in which the synthesis of judgements and the extraction of weights is performed is irrelevant. Only the geometric mean has it also. Except for $n = 3$ (and trivially $n = 2$) the principal right eigenvector method shares neither of the invariances. Section 5 concludes.

2 Some preparatory discussions

2.1 Saaty’s AHP: the inconsistency measure

Saaty (1980, p. 180) defines perturbations, or errors, ε_{ij} relative to a weight vector w by $\varepsilon_{ij} := a_{ij} \cdot (w_j \div w_i)$. So a_{ij} , with $a_{ij} = (w_i \div w_j) \cdot \varepsilon_{ij}$, is treated as a multiplicative perturbation of a ratio of weights. We have of course $\varepsilon_{ji} = 1 \div \varepsilon_{ij}$. The errors all equal one if and only if the matrix A is consistent. Saaty noted (on the same page) that *if* the errors are evaluated at the principal right eigenvector, then one can write for the largest eigenvalue λ_{\max} of A :

$$\mu := \frac{\lambda_{\max} - n}{n - 1} = \sum_{i < j} \left[\frac{1}{2} \left(\varepsilon_{ij} + \frac{1}{\varepsilon_{ij}} \right) - 1 \right] \div \left(\frac{1}{2} n(n - 1) \right). \tag{4}$$

Since $f(x) := \frac{1}{2}(x + \frac{1}{x}) - 1$ is strictly convex on the set of positive reals with $f(x) \geq f(1) = 0$, μ is nonnegative and zero only when all errors equal one. Equivalently, the matrix A is consistent if and only if $\lambda_{\max} = n$. Saaty proposed to use μ as *the* measure of inconsistency.

In order to judge when for a given matrix the inconsistency is too large to be tolerated, Saaty suggests to use μ ’s average value for randomly generated positive reciprocal matrices of the same order. The entries of the random matrices are usually independent random drawings from:

$$\{1/9, 1/8, \dots, 1/3, 1/2, 1, 2, 3, \dots, 8, 9\}$$

Note that the average will change with the generating set. A bit of an industry arose in an attempt to estimate the average μ as *accurately* as possible, see e.g. Alonso and Lamata (2006) for an overview. Since the matrices are generated without any concern for consistency, it is suggested (Saaty 2005, section 1–9;1994, section 1–7) as a rule of thumb to accept the inconsistency when the observed μ is smaller than one tenth of the estimated average value (with some variations for $n = 3$ or 4), and to take action otherwise. Incidentally, it can be shown with some straightforward but tedious algebra that when the ε_{ij} ’s in μ are evaluated *not* at the principal right eigenvector but instead at the geometric row mean, μ ’s expected value for random positive reciprocal matrices can be given exactly:

$$E(\mu) = E\left(a^{1-\frac{2}{n}}\right) \cdot \left(E\left(a^{\frac{1}{n}}\right)\right)^{2(n-2)} - 1, \tag{5}$$

where a is a positive random variable² with the property that a and $1/a$ have the same distribution, as e.g. for a random drawing with equal probabilities from $\{1/9, 1/8, \dots, 1/3, 1/2, 1, 2, 3, \dots, 8, 9\}$. For the latter case the numbers are close

² Bold symbols will represent random variables. Ordinary symbols will stand for values, realizations, numbers.

Table 1 Expected value for μ as evaluated at the geometric mean and the principal right eigenvector, respectively

	$n = 3$	$n = 4$	$n = 5$	$n = 6$	$n = 7$	$n = 8$	$n = 9$	$n = 10$
$E(\mu)$ geometric	.52	.87	1.08	1.22	1.32	1.39	1.44	1.48
$E(\mu)$ Saaty	.52	.89	1.11	1.25	1.35	1.40	1.45	1.49

to those obtained (Saaty 2005, section 1–9) by averaging μ for 50,000 random positive reciprocal matrices, evaluated at the principal right eigenvector, as shown in Table 1.

For $n = 3$ the numbers are exactly the same, since the principal right eigenvector is the geometric row mean, see also Sect. 3.2. The table shows that minimization of the expression for μ as a function of w will in general *not* produce the principal right eigenvector. In fact, for $n = 4, \dots, 10$ the geometric row mean will have a better, but not necessarily optimal fit either, in the majority of the cases (this is based on a simulation with 40,000 random positive reciprocal matrices; for $n = 7$ the geometric row mean is better than the principal right eigenvector in 70% of the cases; the larger n the lower this percentage: for $n = 10$ it is 57%).

A result useful for suggesting which entry of the matrix A at hand to change is the derivative of λ_{\max} with respect to a_{ij} , see Saaty (2005, p. 30):

$$\partial\lambda_{\max}/\partial a_{ij} = v_i \cdot w_j - a_{ji}^2 \cdot v_j \cdot w_i \tag{6}$$

where v is the left principal eigenvector, and w the right principal eigenvector; v is normalized such that $v^\top w = 1$. The result is a consequence of $\partial\lambda/\partial A = v \cdot w^\top$ for a non-repeated eigenvalue λ of an arbitrary matrix A with corresponding left and right eigenvector v and w resp., see Neudecker (1967).

2.2 Random errors instead of random matrices

In the previous section we considered Saaty’s suggestion to average the value of a particular inconsistency measure across arbitrary (positive reciprocal) matrices, in order to assess the seriousness of the inconsistency as measured for the matrix at hand. Random matrices generated without any concern for consistency are however rarely ‘acceptable’, in the sense that Saaty’s consistency measure is less than one tenth of the average μ : about one out of five for $n = 3$ and an estimated fourteen (14) out of ten million (10^7) for $n = 7$, see Bozóki and Lewis (2005, p. 257). If the entries of A can be any positive real number, the set of consistent matrices has measure zero. So it is difficult to see what relevance arbitrary matrices have for the matrix at hand. Perhaps it is more relevant to know the typical size of the (judgemental) errors in the actual construction of pairwise comparison matrices. However we do not know of any studies in which error distributions are reported and analyzed. In a sense this is not surprising: errors one is aware of cease to exist, presumably, the moment they are detected. And also, in empirical, statistical studies errors are never observed directly, but they are always defined and constructed relative to a model. Granted this, it could still be a

useful idea to collect systematically error distributions as induced by real world pairwise comparison matrices, whose inconsistency is deemed acceptable by the decision makers who constructed the matrices. The errors could be defined with respect to any scale and any weight extraction method the decision maker is comfortable with.

Alternatively, following statistical practice we could ‘invent’ a distribution, based on pragmatic and ‘reasonable’ considerations only. In this spirit one could impose the condition, using ϵ as a generic symbol for an error in A , that ϵ and $1/\epsilon$ have the same distribution. After all, there is no reason to suppose that errors above the diagonal behave any different from errors below the diagonal. An immediate consequence is that $\text{median}(\epsilon) = 1$. More generally, $\delta := \log(\epsilon)$ must have a distribution symmetrically around zero. There is no lack of choice here, but invoking the maximum entropy principle, we could reduce the choice to two candidates: the Laplace- distribution, or the Gaussian distribution. The latter maximizes the entropy for densities with a finite variance and the real line as support, the former does the same but requires only the existence of the mean (see Kagan et al. 1973, section 13.2). So we could take:

$$p(\delta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{\delta^2}{\sigma^2}\right) \text{ or } p(\delta) = \frac{1}{\sigma\sqrt{2}} \exp\left(-\frac{\sqrt{2}}{\sigma} |\delta|\right) \tag{7}$$

where σ is the standard deviation. One way to get an order of magnitude for σ is to argue that errors out of the range $(1/2, 2)$ are large and therefore exceptional. This is supported partly by anecdotal evidence (the author was employed at a company where more or less routinely comparison matrices are constructed in order to get reasonable weights in certain selection problems). We could translate this into the requirement that σ must be such that $\text{prob}(\epsilon \geq 2) = .025$ or $.05$ say, which again is not unreasonable in the light of anecdotal evidence. If we let the probability be α , then $\sigma = \log\left(\frac{1}{2}\right) \div \Phi^{-1}(\alpha)$ in case we work with the Gaussian distribution, otherwise we have $\sigma = \sqrt{2} \log(2) \div \log(1 \div 2\alpha)$. The numbers are not that different for $\alpha = .025$ or $.05$. We obtain for Gauss $.3536$ and $.4214$ and for Laplace $.3272$ and $.4257$ respectively. Of course, the lower the frequency of large errors, the smaller the standard deviation. The expected value of μ for random errors is for either distribution

$$E(\mu) = E(\epsilon) - 1 = E(\exp(\delta)) - 1 \cong \frac{1}{2}\sigma^2 \tag{8}$$

(for the normal it is exactly $\exp\left(\frac{1}{2}\sigma^2\right) - 1$). So depending on the plausibility or perhaps desirability of errors within the range $(1/2, 2)$ the average μ is roughly $.06$ or $.09$. These correspond with ‘certainty equivalent’ errors of roughly 1.4 and 1.5 resp., where a certainty equivalent error solves $\frac{1}{2}(\epsilon + 1/\epsilon) - 1 = \frac{1}{2}\sigma^2$ for ϵ . The numbers obtained are independent of the choice of the scale for the pairwise comparisons. Clearly, when we require that the observed μ is less than the average(s) just specified, we will be more demanding than Saaty.³ It may be advisable to estimate quantiles of the distribution

³ In Saaty (1980, p. 180–181), Saaty writes $\epsilon_{ij} = 1 + \delta_{ij}$ and argues that for an unbiased judge δ_{ij} would tend to be symmetrically distributed about zero in the interval $(-1, 1)$. This would lead to the approximation that μ is half the variance of δ_{ij} . To my knowledge these ideas were not followed up.

F of μ , perhaps $F^{-1}(.75)$, $F^{-1}(.90)$, and $F^{-1}(.95)$, to help judge the acceptability of its size, but this requires assumptions concerning the joint distribution of the errors, which we sofar have avoided to make (we required until now identical marginal distributions only).

One final statistical note: if we assume that the joint distribution of the δ_{ij} 's is symmetrical in the sense that for their joint density p we have

$$p(\delta_{12}, \delta_{13}, \dots, \delta_{n-1,n}) = p(\pm\delta_{12}, \pm\delta_{13}, \dots, \pm\delta_{n-1,n}) \tag{9}$$

for every sequence of plusses and minusses, then in particular each $\sum_j \delta_{ij}$ is also symmetrically distributed with respect to the origin, and therefore $\prod_j \varepsilon_{ij}$ has a unit median. Now:

$$\log(a_{ij}) = \log(w_i) - \log(w_j) + \log(\varepsilon_{ij}) \tag{10}$$

and so

$$\log\left(\prod_j a_{ij}\right) = \log(w_i^n) + 0 + \log\left(\prod_j \varepsilon_{ij}\right). \tag{11}$$

If we again follow statistical practice and construct errors such that their sample median equals the population median: $\prod_j \varepsilon_{ij} = 1$, our implied estimate for w_i equals g_i . Independence of the errors and identical distributions are not required, a certain symmetry is sufficient (and the condition as specified is sufficient, not necessary).

3 A general class of inconsistency measures

3.1 'Error gravity' functions defined and first consequences

Suppose for a real function f defined on $(0, \infty)$ with $f(1) = 0$ and other properties to be imposed later, we take

$$\Delta := \frac{1}{n(n-1)} \sum_i \sum_j f(\varepsilon_{ij}) \tag{12}$$

as a measure of inconsistency. Using $\varepsilon_{ji} = 1/\varepsilon_{ij}$ we can also write, with $m := \frac{1}{2}n(n-1)$,

$$\Delta = \frac{1}{m} \sum_{i < j} \frac{1}{2} (f(\varepsilon_{ij}) + f(1/\varepsilon_{ij})) \tag{13}$$

It is clear that we may assume that $f(x) = f(1/x)$, because if not, $\tilde{f}(x) := \frac{1}{2}(f(x) + f(1/x))$ does have this property and substituting \tilde{f} for f yields the same

value for Δ . So we will take

$$\Delta := \frac{1}{m} \sum_{i < j} f(\varepsilon_{ij}) \tag{14}$$

with $f(x) = f(1/x)$. The previous section afforded an example with $f(x) = \frac{1}{2}(x + 1/x) - 1$. This function decreases on $(0, 1)$ and increases on $(1, \infty)$. Its absolute minimum, zero, is attained uniquely at $x = 1$. Other examples with the very same properties are $f(x) = \frac{1}{2}(\log(x))^2$ or $f(x) = \frac{1}{4}(1 - x)^2 + \frac{1}{4}(1 - 1/x)^2$, and $f(x) = \frac{1}{2}(x^r + x^{-r}) - 1$ for any real number $r \geq 1$. Any convex combination of the given examples is another example. For all functions specified we have $f(1) = f'(1) = 0$, and $f''(1) = 1$. Now every real-valued f defined on \mathbb{R}_+ with $f(x) = f(1/x)$ can be written as $g(\log(x))$ for a real-valued function g defined on \mathbb{R} with $g(y) = g(-y)$ for every real y . We will take g strictly convex, and at least twice continuously differentiable everywhere, with a unique minimum of zero at $y = 0$. A harmless normalization is $g''(0) = 1$ so that $f''(1) = 1$. The (convex) set of functions f called *error gravity functions* as generated by all possible choices of g will induce our set of inconsistency measures. A bit more explicitly:

Definition A mapping $f : \mathbb{R}_+ \rightarrow [0, \infty)$ is called an error gravity function when $f(\cdot) = g(\log(\cdot))$ for a $g : \mathbb{R} \rightarrow [0, \infty)$ which is strictly convex, at least twice continuously differentiable everywhere, as well as symmetrical with respect to zero. The mapping g is normalized such that $g(0) = 0$ and $g''(0) = 1$.

Typically, f will be quasi-convex (:all level sets are convex) but the examples given except $\frac{1}{2}(\log(x))^2$ are strictly convex also. For x close to 1 all error gravity functions with a third derivative will be close to one another: repeatedly differentiating $f(x) = f(1/x)$ yields $f'''(1) = -3 \cdot f''(1) = -3$ (higher order derivatives if they exist are free) and so

$$f(x) = \frac{1}{2}(x - 1)^2 - \frac{1}{2}(x - 1)^3 + O((x - 1)^4). \tag{15}$$

In particular this entails that $Ef(\epsilon) = Ef(\exp(\delta)) = \frac{1}{2}\sigma^2 + O(\sigma^4)$, so all sufficiently smooth inconsistency measures can plausibly be judged on the same basis.

We propose to extract weights from A by minimizing the *average error gravity* Δ as a function of w , subject to a convenient normalization like $\prod_j w_j = 1$. Differentiating the Lagrangian $L := \Delta(w) - \lambda \cdot c(w)$, where $c(w)$ is the normalizing constraint, with respect to w_j yields $\frac{\partial L}{\partial w_j} = \frac{\partial \Delta}{\partial w_j} - \lambda \cdot \frac{\partial c}{\partial w_j} = 0$. Since Δ is such that multiplication of w by any non-zero scalar does not change it, the optimal λ must be zero.⁴ So the optimal weights make the first-order derivatives of Δ with respect to the weights zero. With some algebra, using the implication of $f(x) = f(1/x)$

⁴ Euler tells us that the property of Δ just referred to, homogeneity of degree zero, implies that $\sum_i w_i \cdot \partial \Delta / \partial w_i = 0$, from which $\lambda = 0$ easily follows.

that $x \cdot f'(x) = -(1/x) \cdot f'(1/x)$ we obtain that the optimal weights satisfy:

$$\sum_j f'(\varepsilon_{kj}) \cdot \varepsilon_{kj} = 0 \quad \text{for } k = 1, 2, \dots, n. \tag{16}$$

or equivalently,

$$\sum_i f'(\varepsilon_{ik}) \cdot \varepsilon_{ik} = 0 \quad \text{for } k = 1, 2, \dots, n. \tag{17}$$

For the special case of Saaty’s error gravity function $f(x) = \frac{1}{2}(x + 1/x) - 1$ this entails that the k th column sum, $\sum_j \varepsilon_{jk}$ must equal the k th row sum $\sum_j \varepsilon_{kj}$. To spell out: we must have

$$\left(\sum_j a_{kj} \cdot w_j \right) \cdot (1/w_k) = \left(\sum_j a_{jk} \cdot (1/w_j) \right) \cdot w_k. \tag{18}$$

So when w is the principal right eigenvector the left hand side equals λ_{\max} and we can write:

$$\sum_j a_{jk} \cdot (1/w_j) = \lambda_{\max} \cdot (1/w_k) \tag{19}$$

Consequently, the principal right eigenvector minimizes Saaty’s inconsistency measure *only* when $(1/w^T) := (1/w_1, 1/w_2, \dots, 1/w_n)$ is the principal *left* eigenvector. We will see below that this is the case for *any* positive reciprocal 3×3 matrix A , but the table in Sect. 2.1 already showed that for $n \geq 4$ and inconsistent A this will happen only exceptionally. In fact Saaty (1980, p. 192) conjectured that it will never happen when A with $n \geq 4$ is inconsistent. But (see Sect. 4.1) there are positive reciprocal matrices structured in a very special way where it *does* happen. For these matrices, as well as for all 3×3 matrices, the principal right eigenvector and the geometric row mean are identical.

For $f(x) = \frac{1}{2}(\log(x))^2$ we get the well-known result that $\sum_j \log(\varepsilon_{ij}) = 0$ must be zero, or $\prod_j \varepsilon_{ij} = 1$ and $w_{\text{optimal}} = g$. Clearly, the expected value of the inconsistency measure for random errors equals $\frac{1}{2}E(\delta^2) = \frac{1}{2}\sigma^2$, and certainty equivalent errors which solve $\frac{1}{2}(\log(\varepsilon))^2 = \frac{1}{2}\sigma^2$ are simply $\exp(\sigma)$. If we are willing to assume that the errors are not only identically distributed, but also Gaussian as well as independent, then quantiles of the inconsistency measure can be derived via a $\chi^2(m)$ -distribution.

For random positive reciprocal matrices, where the upper diagonal entries are independent copies of \mathbf{a} , a positive random variable with the property that \mathbf{a} and $1/\mathbf{a}$ have the same distribution, the average inconsistency measure can be shown to equal:

$$(1 - 2/n) \cdot E\left(\frac{1}{2}(\log(\mathbf{a}))^2\right) \tag{20}$$

With \mathbf{a} uniformly distributed on $\{1/9, 1/8, \dots, 1/3, 1/2, 1, 2, 3, \dots, 8, 9\}$ the expected value of $\frac{1}{2} (\log(\mathbf{a}))^2$ equals 1.3791. If one feels that random matrices are relevant for judging the size of the observed inconsistency, as in the AHP-approach, one could take one tenth of the average inconsistency measure as a critical boundary.

We close this section with the fact that at the *optimal* w

$$\partial\Delta/\partial a_{ij} = \frac{1}{m} f'(\varepsilon_{ij}) \cdot \varepsilon_{ij} \div a_{ij}. \tag{21}$$

This is useful for suggesting where to adjust the matrix when one wants to decrease the inconsistency. For $f(x) = \frac{1}{2}(x + 1/x) - 1$ this yields: $\partial\Delta/\partial a_{ij} = \frac{1}{2m} (\varepsilon_{ij} - \varepsilon_{ji}) \div a_{ij}$ and for $f(x) = \frac{1}{2} (\log(x))^2$ we get $\partial\Delta/\partial a_{ij} = \frac{1}{m} \log(\varepsilon_{ij}) \div a_{ij}$.

3.2 A special case: $n = 3$. All methods concur

Consider an arbitrary positive reciprocal 3×3 matrix A :

$$A := \begin{bmatrix} 1 & a & b \\ 1/a & 1 & c \\ 1/b & 1/c & 1 \end{bmatrix} \tag{22}$$

where a, b and c are any positive real numbers. A is consistent if and only if $a \cdot c = b$. Take any error gravity function f and define

$$\Delta := \frac{1}{3} \cdot f\left(a \frac{w_2}{w_1}\right) + \frac{1}{3} \cdot f\left(b \frac{w_3}{w_1}\right) + \frac{1}{3} \cdot f\left(c \frac{w_3}{w_2}\right). \tag{23}$$

Since $f(x) = g(\log(x))$ with g strictly convex and symmetric about zero we can write:

$$\Delta = \frac{1}{3} \cdot g\left(\log\left(a \frac{w_2}{w_1}\right)\right) + \frac{1}{3} \cdot g\left(\log\left(1/\left(b \frac{w_3}{w_1}\right)\right)\right) + \frac{1}{3} \cdot g\left(\log\left(c \frac{w_3}{w_2}\right)\right) \tag{24}$$

which is at least as large as

$$\begin{aligned} &g\left(\frac{1}{3} \log\left(a \frac{w_2}{w_1}\right) + \frac{1}{3} \log\left(1/\left(b \frac{w_3}{w_1}\right)\right) + \frac{1}{3} \log\left(c \frac{w_3}{w_2}\right)\right) \\ &= g\left(\frac{1}{3} \log\left(\frac{ac}{b}\right)\right) = f\left(\left(\frac{ac}{b}\right)^{\frac{1}{3}}\right), \end{aligned} \tag{25}$$

with equality only when:

$$a \frac{w_2}{w_1} = 1/\left(b \frac{w_3}{w_1}\right) = c \frac{w_3}{w_2}. \tag{26}$$

Solving the latter equation yields $w_1 = (ab)^{\frac{1}{3}}$, $w_2 = (c/a)^{\frac{1}{3}}$ and $w_3 = (bc)^{-\frac{1}{3}}$. In other words, the geometric row mean g is the optimal w for every suitable function f . In fact more is true: g is also the principal right eigenvector, and $1./g$ is the principal left eigenvector.⁵ The remainder of this section contains a simple proof.⁶

First, it is easily verified that $A \cdot g = \lambda_1 \cdot g$ and $(1./g^\top) \cdot A = \lambda_1 \cdot (1./g^\top)$ where $\lambda_1 := 1 + t + (1/t)$ with $t := (\frac{ac}{b})^{\frac{1}{3}}$. It remains to show that $\lambda_1 = \lambda_{\max}$. Observe that $\lambda_1 \geq 3$, and $\lambda_1 = 3$ if and only if $t = 1$ (or A is consistent). The characteristic polynomial of A equals

$$c(\lambda) := -\lambda^3 + 3\lambda^2 + |A| \quad \text{with } |A| := t^3 + t^{-3} - 2. \quad (27)$$

The determinant $|A|$ is nonnegative and zero only when $t = 1$. In the latter case $c(\lambda)$ has one root equal to 3 and two roots equal to 0. For $|A| > 0$ we can factor out $(\lambda - \lambda_1)$ leaving the quadratic polynomial $-\lambda^2 + (3 - \lambda_1)\lambda + (3 - \lambda_1)\lambda_1$. Since now $\lambda_1 > 3$ this polynomial has two complex roots with the same modulus equal to $\sqrt{\frac{1}{2}\lambda_1(\lambda_1 - 3)}$, which is definitely smaller than λ_1 . And so we do indeed have $\lambda_1 = \lambda_{\max}$.

4 Some consistency requirements and invariance properties

4.1 Independence-of-scale-inversion

The decision maker is asked to specify for each entry (i, j) of A how much more important, or qualified or suitable or whatever, the first item i is than the second item j . To the decision maker, the information supplied for each entry (i, j) is exactly equivalent to the statement expressing *conversely* how much more important item j is than item i . The information contained in A is equivalent to the information contained in A^\top . The preferences or judgements are the same, only the encoding is different. As Barzilai (1997, p. 1228) argues, this means that the output should be equivalent. So for w as a function of A one should have $w(1./A) = 1./w(A)$ where $1./x$ takes the reciprocal of a matrix x elementwise. In other words, if the best approximation of A is $w \cdot (1./w^\top)$ then the transpose of the latter is the best approximation of the transpose of A . Barzilai (1997, p. 1228) call this the *independence-of-scale-inversion* property. He states that the geometric mean does have this property, as opposed to the principal right eigenvector. We have just seen that for $n = 3$ both methods yield identical results, and there are larger special matrices, see below, for which this is also true, but generally ‘one over the principal right eigenvector’ is not the principal left eigenvector. So Saaty’s approach does not satisfy the independence-of-scale-inversion property.

⁵ Saaty (1980, p. 191) showed by other means than we will employ that one over the principal right eigenvector is the principal left eigenvector for $n = 3$. Apparently, but we do not know that for sure, the link with the geometric means was not observed.

⁶ A reviewer kindly pointed out that Tuma (1987) has a proof as well.

We would like to point out that the optimizers of the inconsistency measures we defined *all* have the desired property: $f(x) = f(1/x)$ entails that

$$f\left(\left(A^\top\right)_{ij} \cdot w_j/w_i\right) = f(1/(a_{ji} \cdot w_j/w_i)) = f(a_{ij} \cdot (1/w_j)/(1/w_i)), \tag{28}$$

so when w minimizes Δ for A , then $1./w$ minimizes Δ for $1./A$. An immediate corollary is that there is no Δ whose minimizer is the principal right eigenvector for *all* positive reciprocal matrices. This is not to say that Saaty’s weight vector does not minimize any metric. Fichtner (1984) developed a metric $\delta(A, B)$, where A and B are positive reciprocal matrices, with just that property: a (conditional) term involving the largest eigenvalues of A and B (which is zero if and only if the matrices are both consistent or equal) plus the Euclidean distance between their principal right eigenvectors. Minimization of this metric for consistent matrices B yields Saaty’s weight vector. However, since the principal right eigenvectors in this metric can be replaced by weight vectors (of the same type) that will reproduce A and B in case of consistency, the approach is not a justification of the eigenvector method per se.

An example of a 5×5 matrix where all methods discussed sofar yield the same result, is based on a matrix used by Bozóki (2008, p. 356) for unrelated purposes. We generalize slightly, by making the matrix A dependent on two positive real valued parameters a and b :

$$A = \begin{bmatrix} 1 & a & b & 1/b & 1/a \\ 1/a & 1 & a & b & 1/b \\ 1/b & 1/a & 1 & a & b \\ b & 1/b & 1/a & 1 & a \\ a & b & 1/b & 1/a & 1 \end{bmatrix} \tag{29}$$

Bozóki used $a = 6$ and $b = 1$. Clearly, the matrix is purposely constructed for numerical testing, it is rather unusual as a pairwise comparison matrix. If it were a legitimate comparison matrix, one would suspect the decision maker who created it, to be in a ‘mental tie’. We observe that all row products equal one. Moreover, the row sums and the column sums are all equal to $1 + a + 1/a + b + 1/b$. So this is an eigenvalue, the largest one in fact, which exceeds 5 always unless $a = b = 1$. Therefore $(1, 1, 1, 1, 1)^\top$ is the principal right eigenvector *and* the principal left eigenvector, the latter being trivially equal to one over the former. We can prove⁷ that optimizing any inconsistency measure Δ yields the same $w_{opt} = (1, 1, 1, 1, 1)^\top$. This result seems eminently reasonable given the presumed confused state the decision maker is in. Larger matrices with similar properties are easily constructed, but we have no systematic way of characterizing them. We suspect that they have measure zero in the space of all positive reciprocal matrices.

⁷ Write Δ as two averages, one containing a and weights only, the other b and weights only. Use $f(x) = g(\log(x))$ with $g(\cdot)$ strictly convex and symmetrical, to deduce that the averages always exceed $f(a)$ and $f(b)$ respectively, unless the errors in A corresponding with a are equal and similarly for b . With some algebra this yields that all optimal weights are 1.

It is not clear yet to what extent the results concerning the equality of the AHP method and our alternatives can be generalized. It seems that whenever all Δ 's yield the same result, the AHP concurs and the principal left and right eigenvectors are inversely related, and conversely, when the latter is true, all Δ 's concur. This is something we hope to sort out in the future.

We close this section with a small but interesting detour. [Bozóki \(2008\)](#) designed a special, global optimization method, that transforms the least squares problem of minimizing the Frobenius norm of the difference between A and $w \cdot (1./w^\top)$, into finding the roots of multivariate polynomials. It can handle arbitrary matrices up to size 8×8 but computing time (as of 2008⁸) can be an issue for $n = 7$ or 8 . He finds for $a = 6$ and $b = 1$ rather startling results: the problem has 5 local as well as 5 global minimizers. The local minima are the same, as are the global minima of course. The latter are a mere 0.001% smaller than the former. Nevertheless, the sets of minimizers are rather different. Each set of five can be obtained by rotating clockwise any one of the solutions. And although $(1, 1, 1, 1, 1)^\top$ is a critical point, it is not a (local) minimizer but a saddlepoint⁹ with a function value about 1.6% higher than the global minimum. Given the rotational symmetry of the solutions we necessarily have that $(1, 1, 1, 1, 1)^\top$ is the geometric mean of the sets of minimizers separately and jointly, if normalized to have a product equal to one. Also, $(1/5, 1/5, 1/5, 1/5, 1/5)$ is the additive mean for the same sets normalized to add up to one. So, our solution is the barycenter of rather conflicting solutions to the least squares problem. We refer to [Bozóki \(2003, 2008\)](#); [Bozóki and Lewis \(2005\)](#) for an in-depth analysis of the numerical issues surrounding the minimization of the Frobenius norm of the difference between A and $w \cdot (1./w^\top)$, and for the design of appropriate algorithms.

4.2 Synthesizing judgements

Suppose there are two experts whose pairwise comparison matrices for the same problem are A and B respectively, and it is desired to synthesize their judgements. We want a matrix C , say, that can reasonably be said to be a proper compromise between A and B . It is simplest and hardly unreasonable when we synthesize elementwise with the same function:

$$C = \Sigma(A, B) := (\sigma(a_{ij}, b_{ij})). \quad (30)$$

Here the synthesizing function $\sigma : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ will have to be continuous (in at least one point), and we impose $\sigma(a, a) = a$ for all real, positive a . The first condition is an extremely mild regularity condition, ensuring uniqueness of the type of solutions of a certain functional equation to be specified shortly. The second condition says simply that when the experts agree about every judgement, so $A = B$, then their

⁸ The computing time is inversely proportional to the capacity of processors, so in 2011 it is a third to a quarter of the time required in 2008; personal communication by Dr. S. Bozóki, 10-5-2011.

⁹ Personal communication by Dr. S. Bozóki, 9-11-2009. I am grateful for his willingness to test $(1, 1, 1, 1, 1)$ using his software.

compromise shares their opinions: $C = A = B$. Thirdly, if the experts disagree we will want

$$c_{ij} \in [\min(a_{ij}, b_{ij}), \max(a_{ij}, b_{ij})]. \tag{31}$$

As a fourth and final condition we impose that C is consistent whenever A and B are. This is perhaps relatively contentious when compared with the other demands. It is offered with some confidence though because arguments to the contrary seem hard to find. So we want

$$\sigma(a_{ik}, b_{ik}) \cdot \sigma(a_{kj}, b_{kj}) = \sigma(a_{ij}, b_{ij}) \tag{32}$$

whenever $a_{ik} \cdot a_{kj} = a_{ij}$ and $b_{ik} \cdot b_{kj} = b_{ij}$. We can write this for positive values of x_i and y_i as the functional equation

$$\sigma(x_1, y_1) \cdot \sigma(x_2, y_2) = \sigma(x_1 \cdot x_2, y_1 \cdot y_2). \tag{33}$$

With $\tilde{\sigma} := \log(\sigma)$ and $\tilde{x}_i := \log(x_i)$, $\tilde{y}_i := \log(y_i)$ we have the well-known (multi-place) functional *Cauchy equation*:

$$\tilde{\sigma}(\tilde{x}_1, \tilde{y}_1) + \tilde{\sigma}(\tilde{x}_2, \tilde{y}_2) = \tilde{\sigma}(\tilde{x}_1 + \tilde{x}_2, \tilde{y}_1 + \tilde{y}_2) \tag{34}$$

whose solutions are all of the form $\tilde{\sigma}(\tilde{x}, \tilde{y}) = \theta_1 \cdot \tilde{x} + \theta_2 \cdot \tilde{y}$ for real numbers θ_1 and θ_2 , see [Aczél and Dhombres \(1991, section 4.1\)](#). In terms of the original variables and function we have $\sigma(a, b) = a^{\theta_1} \cdot b^{\theta_2}$, in fact:

$$\sigma(a, b) = a^\theta \cdot b^{1-\theta} \tag{35}$$

with $\theta \in [0, 1]$ because $\sigma(a, a) = a$ and because of the compromise nature of C : $\sigma(a, b) \in [\min(a, b), \max(a, b)]$.

We conclude that the ‘only reasonable’ way to synthesize judgements, such that synthesized values are consistent whenever the underlying judgements are, is by using (weighted) geometric means.

The parameter θ could be just $\frac{1}{2}$ or it could reflect competence, experience, political clout or otherwise. Since σ is necessarily concave in the judgements, highly diverse opinions will tend to be assigned a compromise value in the neighborhood of 1.

Suppose now that we have two weight vectors v and w for the same problem, and it is desired to define a suitable compromise between the two. It seems reasonable to demand for the synthesizer of the weights that $\sigma(w_3, v_3) / \sigma(w_2, v_2) = \sigma(w_2, v_2) / \sigma(w_1, v_1)$ whenever $w_3/w_2 = w_2/w_1$ and $v_3/v_2 = v_2/v_1$. In other words, the synthesized values ought to honour the proportionalities. We propose in fact the following functional equation:

$$\sigma(\lambda x, \mu y) = \sigma(x, y) \cdot R(\lambda, \mu) \tag{36}$$

where both σ and R are functions from $\mathbb{R}_+ \times \mathbb{R}_+$ in \mathbb{R}_+ with $\sigma(x, x) = x$ for all x . Of course, it is basically the same functional equation as before: with $x = y = 1$ we get $\sigma(\lambda, \mu) = \sigma(1, 1) \cdot R(\lambda, \mu)$ and so we necessarily have $R \equiv \sigma$. All solutions continuous in at least one point are the weighted geometric means.

It is somewhat remarkable that the AHP-approach advocates the same device, for pairwise comparisons as well as for weights. See e.g. Saaty (1980, 2005, 2006). The motivation is different though: Aczél and Saaty (1983), for the geometric mean, and Aczél and Alsina (1986) for the weighted version, basically adopt the axioms that lead to the quasiarithmetic means, and then add homogeneity and a ‘reciprocal property’. The latter means that $\sigma(1/a, 1/b) = 1/\sigma(a, b)$ and it is clearly related to the independence-of-scale-inversion. Our motivation in terms of consistency is linked in a direct way to the context in which the synthesizing is to take place. But whatever the motivation, the outcome is the same: the AHP also effectively requires consistent inputs to be transformed into consistent output.

4.3 Synthesizing ratios or weights, the order of things

Barzilai and Golany (1994) observed, as no doubt did many others, that the geometric row mean of $A^\theta \circ B^{1-\theta} := (a_{ij}^\theta \cdot b_{ij}^{1-\theta})$, the synthesis of a pair of conformable positive reciprocal matrices A and B is just $g_a^\theta \circ g_b^{1-\theta} := (g_{a,i}^\theta \cdot g_{b,i}^{1-\theta})$, the synthesis of the corresponding geometric row means. This entails that whether we synthesize the ratios first and then extract the weightvector, or extract the weightvectors from both matrices and then synthesize them, we obtain exactly the same result. The question arises whether this invariance property applies to more methods (it does not apply to the principal right eigenvector). For the class of inconsistency measures we introduced in Sect. 3.1 we have that:

$$\begin{aligned} \Delta(A^\theta \circ B^{1-\theta}, w_{optimal}) &\leq \Delta(A^\theta \circ B^{1-\theta}, w_a^\theta \circ w_b^{1-\theta}) \\ &\leq \theta \cdot \Delta(A, w_a) + (1 - \theta) \cdot \Delta(B, w_b) \end{aligned} \tag{37}$$

where w_a and w_b minimize the inconsistency measures for A and B respectively. The first inequality is by definition, the second is a consequence of our assumption that $f(x) = g(\log(x))$ with $g(\cdot)$ strictly convex, or equivalently, $f(\exp(x))$ is strictly convex in x and

$$\begin{aligned} &\Delta(A^\theta \circ B^{1-\theta}, w_a^\theta \circ w_b^{1-\theta}) \\ &= \frac{1}{m} \sum_{i < j} f(\exp[\theta \cdot \log(a_{ij} \cdot (w_{a,j}/w_{a,i})) + (1 - \theta) \cdot \log(b_{ij} \cdot (w_{b,j}/w_{b,i}))]) \end{aligned}$$

And the conclusion follows. So synthesizing w_a and w_b is for general f a step in the right direction towards the best weights for the synthesis of A and B . For $f(x) = \frac{1}{2}(\log(x))^2$ the stepsize is precisely right. We will show that this characterizes the geometric mean by proving the following proposition:

Proposition Let $w_a := \arg \min_w \left(\frac{1}{m} \sum_{i < j} f(\varepsilon_{ij}(A, w)) \right)$ and similarly for w_b for two conformable positive reciprocal matrices A and B , and the function f belongs to the class of error gravity functions.

If $w_a^\theta \circ w_b^{1-\theta} = \arg \min_w \left(\frac{1}{m} \sum_{i < j} f(\varepsilon_{ij}(A^\theta \circ B^{1-\theta}, w)) \right)$

then $f(x) = \frac{1}{2} (\log(x))^2$.

Proof In terms of the first order conditions, $\sum_j f'(\varepsilon_{a,kj}) \cdot \varepsilon_{a,kj} = 0$ and $\sum_j f'(\varepsilon_{b,kj}) \cdot \varepsilon_{b,kj} = 0$ entail that

$$\sum_j f'(\varepsilon_{a,kj}^\theta \cdot \varepsilon_{b,kj}^{1-\theta}) \cdot \varepsilon_{a,kj}^\theta \cdot \varepsilon_{b,kj}^{1-\theta} = 0. \tag{38}$$

Or in a more simple notation, for positive x and y values: $\sum_j f'(x_j) \cdot x_j = 0$ and $\sum_j f'(y_j) \cdot y_j = 0$ entail that

$$\sum_j f'(x_j^\theta \cdot y_j^{1-\theta}) \cdot x_j^\theta \cdot y_j^{1-\theta} = 0. \tag{39}$$

Now f is an error gravity function so $f(\exp(u))$ is strictly convex in u . Its first derivative $h(u) := f'(\exp(u)) \cdot \exp(u)$ has itself therefore a first derivative that is strictly positive everywhere and so $h(\cdot)$ is invertible. We also have $h(0) = 0$ and $h'(0) = 1$. With the substitutions $x_j = \exp(u_j)$ and $y_j = \exp(v_j)$ our implication reads:

$$\sum_j h(u_j) = 0 \ \& \ \sum_j h(v_j) = 0 \Rightarrow \sum_j h(\theta \cdot u_j + (1 - \theta) \cdot v_j) = 0 \tag{40}$$

Define $H(u) := \sum_j h(u_j)$. By inserting $v = 0$ in the implication we find that $H(\theta \cdot u) = 0$ whenever $H(u) = 0$. Differentiating $H(\theta \cdot u)$ with respect to θ yields $H(\theta \cdot u) = H(0) + \theta \cdot \nabla H(0) \cdot u + o(\theta)$, where $H(0) = 0$ and the gradient of H evaluated at zero, $\nabla H(0)$, equals $(1, 1, \dots, 1)$. So $0 = \nabla H(0) \cdot u + o(\theta) / \theta$ where the last term tends to zero when θ does. Therefore $\nabla H(0) \cdot u = 0$ for all u satisfying $H(u) = 0$. Since $h(\cdot)$ is invertible with inverse $k(\cdot)$, say, this functional equation is equivalent to $\sum_j k(u_j) = 0$ for $\sum_j u_j = 0$. Any solution of this well-known equation, see [Aczél \(2006, p. 47\)](#), is linear and therefore $h(\cdot)$ is linear. Retracing our steps leads in terms of x to: $f'(x) = c \cdot \log(x) \div x$ for a constant c , which is solved uniquely by $f(x) = \frac{1}{2} (\log(x))^2$ (taking the normalization into account). \square

4.4 Consistent extraction of weights in a hierarchy

It happens frequently in practice that many more aspects are relevant for the choice or the ranking of objects than can be reliably compared pairwise. As Saaty has pointed out repeatedly, decision makers can handle 7 ± 2 aspects relatively confidently, but higher numbers can pose a real challenge. In many cases the aspects can be arranged more or less naturally in a hierarchy, as layers of a pyramid. The more tangible aspects are lined up at the lowest layer, the base. The concepts in each layer except the base represent effectively a partition of the aspects or concepts of the layer directly below it.

The number of concepts at the various layers is kept small to facilitate pairwise comparisons. So instead of one large matrix many smaller matrices are constructed, each allowing the extraction of relative weights. The absolute weights at each layer are then defined by the product of the relative weights along the direct path from the aspects or concepts to the pinnacle of the pyramid. Basically, the pairwise comparisons between aspects of different ‘blocks’ at the same layer are dealt with indirectly by the *inter*-block comparisons performed on a higher layer. The final result is a complete ‘super-matrix’ of pairwise comparisons for all (basic) aspects. It would present a somewhat problematic choice when the weights extracted from this super-matrix induce a (substantially) different super-matrix. This cannot happen though for any of the Δ -minimization methods. We will demonstrate that for a simple two-layer pyramid, extension to more blocks and more layers is not difficult.

So consider a positive reciprocal $n_1 \times n_1$ -matrix A with a vector of geometric row means g_a and a positive reciprocal $n_2 \times n_2$ -matrix B with a vector of geometric row means g_b . Assume that at the next higher layer or level the ‘A-concept’ has been compared with the ‘B-concept’, with ensuing weights (geometric row means as well) θ_a and θ_b respectively. The standard normalization is imposed, so the product of weights per group equals one, in particular $\theta_a \cdot \theta_b = 1$. The implicit comparison between aspect i of A with aspect j of B is $\theta_a \cdot g_{a,i} \div (\theta_b \cdot g_{b,j})$. We can collect these in a matrix C , say, and get the following super-matrix:

$$\begin{bmatrix} A & C \\ 1./C^\top & B \end{bmatrix}. \tag{41}$$

Now if we take the geometric mean of the i th-row of $[A, C]$ we get:

$$\begin{aligned} & \left\{ \prod_{j=1}^{j=n_1} a_{ij} \cdot \prod_{k=1}^{k=n_2} (\theta_a \cdot g_{a,i}) \div (\theta_b \cdot g_{b,k}) \right\}^{1/(n_1+n_2)} \\ &= \left\{ g_{a,i}^{n_1+n_2} \cdot \theta_a^{n_2} \div \theta_b^{n_2} \right\}^{1/(n_1+n_2)} \\ &= g_{a,i} \cdot \theta_a^{2n_2/(n_1+n_2)} \end{aligned} \tag{42}$$

So the best consistent approximation of A is the same as before. And similarly for B since we get $g_{b,j} \cdot \theta_b^{2n_1/(n_1+n_2)}$, when we take the geometric mean of the j th-row of $[1./C^\top, B]$. The *explicit* comparison between aspect i of A with aspect j of B is now:

$$g_{a,i} \cdot \theta_a^{2n_2/(n_1+n_2)} \div \left(g_{b,j} \cdot \theta_b^{2n_1/(n_1+n_2)} \right) = \theta_a \cdot g_{a,i} \div (\theta_b \cdot g_{b,j}) \tag{43}$$

as determined *implicitly* before. In other words, the weights extracted from the super-matrix are fully consistent with the weights used in its construction. We note that since C is of rank 1 *all extraction methods based on optimization of a suitable Δ will share the invariance with the geometric row mean*: we can minimize Δ by minimization of the terms related to A and to B separately, which yields w_a and w_b , and then rescale them in order to reproduce C exactly.

If the same invariance were to be valid for the principal right eigenvector, then $[\phi(\theta_a) \cdot e_a^\top, \psi(\theta_b) \cdot e_b^\top]$ where e_a and e_b are the relevant eigenvectors of A and B respectively, would have to be the transpose of the principal right eigenvector of the super-matrix.¹⁰ Here the real-valued functions $\phi(\cdot)$ and $\psi(\cdot)$ defined on positive reals are such that

$$\phi(\theta_a) \div \psi(\theta_b) = \theta_a \div \theta_b. \quad (44)$$

Postmultiplication by the candidate eigenvector yields the transpose of

$$\left[(\lambda_{a,\max} + n_2) \cdot \phi(\theta_a) \cdot e_a^\top, (\lambda_{b,\max} + n_1) \cdot \psi(\theta_b) \cdot e_b^\top \right] \quad (45)$$

in an obvious notation. Clearly, we have invariance only if $\lambda_{a,\max}$ differs from n_1 by exactly the same positive amount that $\lambda_{b,\max}$ differs from n_2 . This is only accidentally so, e.g. when A and B are both consistent.

5 Conclusion

In this paper we introduced and analyzed a class of inconsistency measures for positive reciprocal matrices. Our purpose was to contribute to the discussion concerning the choice of proper methods for the extraction of weights, a discussion which is largely dominated by two classical contenders, the AHP with the principal right eigenvector and the geometric row mean. Using global properties of the function used by Saaty for measuring inconsistency we developed a class of functions, all suitable for the same purpose. We offered some suggestions as to how to assess the size of the inconsistency. More importantly we showed that optimizers of all members of the class of error gravity measures share relatively desirable and reasonable invariance properties, like invariance-of-scale-inversion and the consistent extraction of weights in hierarchies. We also showed that when matrices are synthesized in a natural way, in general the geometric mean is uniquely invariant with respect to the order in which operations are performed. For the first non-trivial situation, with three aspects or items, we demonstrated that all methods proposed including the AHP, yield the same results. But for more aspects or items the AHP does not possess the invariances discussed, except accidentally, and when it does (for the examples analyzed), all other contenders yield the same result.

We refer to [Lootsma \(1999\)](#), [Barzilai \(1997\)](#), [Barzilai and Golany \(1994\)](#) e.g. for additional invariance properties of the geometric row mean approach, emphasizing that the order in which operations can be carried out is not relevant for the outcome when the order has no intrinsic value or logic. Barzilai in particular is an outspoken proponent of the geometric row mean approach. In fact, in [Barzilai \(1997\)](#) he introduced two requirements for weight extraction methods, called ‘axioms’, that ought to settle the matter once and for all: the first axiom stipulates that when applied to a

¹⁰ We do not need new symbols for θ_a and θ_b , because for only two aspects the geometric row mean and the principal right eigenvector are identical.

consistent matrix, the ensuing weight vector ought to be the true one, and the second axiom demands essentially that each individual weight is a function of the entries in the corresponding row of the pairwise comparison matrix *only*. Barzilai shows that the only extraction method satisfying these axioms is the geometric row mean. Although sympathetic to the conclusion, we are somewhat uncomfortable with its premise. The second axiom effectively forbids the use of any method that requires a renormalization of its output (to ensure that the product of the weights equals one), even when the ensuing consistent approximation is perfect. Since the product of all elements of a positive reciprocal matrix is trivially 1 and therefore the product of the geometric row means is that also, the second axiom appears to capitalize rather heavily on a normalization for which no substantive arguments (not that we know of anyway) are supplied. Although we do grant that in a *hierarchy* of matrices the normalization is rather sensible, as the analysis in the previous section substantiates.

Additional, interesting results are obtained by [Blanquero et al. \(2006\)](#), adopting an approach rather different from ours. Whereas we define a one-dimensional inconsistency measure, to be minimized, they look at the $n^2 - n$ dimensional vector of discrepancies $(|a_{ij} - (w_i/w_j)|)_{i \neq j}$ for a given choice of w , and ask whether it is possible to replace w by a vector whose discrepancies are not larger, and strictly smaller for at least one index. It turns out that when w is the geometric row mean this is never possible: a diminishing of the discrepancies somewhere is inevitably paid for by deteriorations elsewhere. In the parlance of multicriteria optimization the geometric row mean is *efficient*. They show that the principal right eigenvector does not share this property, although the example given is certainly not dramatic.

We close this section with some ‘personal comments’. It is inevitable that value judgements and esthetic considerations enter discussions about the choice of ‘proper’ weight extraction methods for *inconsistent* comparison matrices. Words like ‘plausible’, ‘natural’, ‘invariant’, ‘suitable’ et cetera all have rather positive connotations and rhetorical value and are freely employed by all involved in the discussion. We did not refrain from their use either. Our personal bias will be clear: we think that the geometric row mean is a robust and dependable workhorse for use in decision problems with multiplicative input. Invariance with respect to the order of operations, where the particular order is not carried by substantive considerations, the naturalness (here we go again) of the approach from a statistical point of view, as well as its easy calculations, are all strong arguments in its favor. This is absolutely *not* meant to deny that other methods have merit, or that research in e.g. the really baffling intricacies of the least squares method is very valuable. But from a practitioner’s perspective it appears that the geometric row mean, possibly supplemented with a robustness analysis of its outcome, will do an acceptable job.

Finally, we iterate our call for a deeper analyses of ‘real life measurement errors’ as measured in substantial and realistic decision problems. It would be very valuable to know to what extent the ‘errors’ *materially* depend on the context, the scale (the numerical code for relative value statements, such as Saaty’s, or Lootsma’s power scale¹¹ or any other alternative), or the extraction method. We may or may not find that

¹¹ [Lootsma \(1999\)](#) chapter 4 e.g.

a scale fixed for *all* applications is too rigid, meaning that occasionally inconsistencies can be induced which are unnecessarily ‘large’. This will be very useful information too. In short, we will welcome anything that helps define a standard distribution of ‘acceptable’ errors on a solid empirical basis.

Acknowledgments The constructive comments and questions by the reviewers are greatly appreciated. References to Aczél and Saaty (1983), Tuma (1987) and Blanquero et al. (2006) are gratefully acknowledged.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Aczél J (2006) Lectures on functional equations and their applications. Dover Publishing Company. (unabridged republication of the 1966 edition by Academic Press; errata list and corrigenda added)
- Aczél J, Saaty TL (1983) Procedures for synthesizing ratio judgments. *J Math Psychol* 27:93–102
- Aczél J, Alsina C (1986) On synthesis of judgments. *Socioecon Plann Sci* 20:333–339
- Aczél J, Dhombres J (1991) Functional equations in several variables. Cambridge University Press, Cambridge
- Alonso JA, Lamata MT (2006) Consistency in the analytic hierarchy process: a new approach. *Int J Uncertain, Fuzziness Knowl-Based Syst* 14(4):445–459
- Barzilai J, Golany B (1994) AHP rank reversal, normalization and aggregation rules. *INFOR* 32:57–64
- Barzilai J (1997) Deriving weights from pairwise comparison matrices. *J Oper Res Soc* 48:1226–1232
- Blanquero R, Carrizosa E, Condo E (2006) Inferring inefficient weights from pairwise comparison matrices. *Math Meth Oper Res* 64:271–284
- Bozóki S (2003) A method for solving LSM problems of small size in the AHP. *Cent Eur J Oper Res* 11:17–33
- Bozóki S, Lewis RH (2005) Solving the least squares method problem in the AHP for 3×3 and 4×4 matrices. *Cent Eur J Oper Res* 13:255–270
- Bozóki S (2008) Solution of the least squares method problem of pairwise comparison matrices. *Cent Eur J Oper Res* 16:345–358
- Fichtner J (1984) Some thoughts about the mathematics of the analytic hierarchy process. Report 8403, Institut für Angewandte Systemforschung und Operations Research, Universität der Bundeswehr, München
- Kagan AM, Linnik Yu V, Rao CR (1973) Characterization problems in mathematical statistics. Wiley, New York
- Lootsma FA (1999) Multi-criteria decision analysis via ratio and difference judgement. Kluwer, Dordrecht
- Neudecker H (1967) On matrix procedures for optimising differentiable scalar functions of matrices. *Statistica Neerlandica* 21:101–107
- Saaty TL (1980) The analytic hierarchy process. McGraw-Hill, New York
- Saaty TL (2005) Theory and applications of the analytic network process. RWS Publications, Pittsburgh
- Saaty TL (2006) Fundamentals of decision making and priority theory. Vol. VI of the AHP Series. RWS Publications, Pittsburgh
- Tuma JJ (1987) Engineering mathematics handbook, 3rd edn. McGraw-Hill, New York