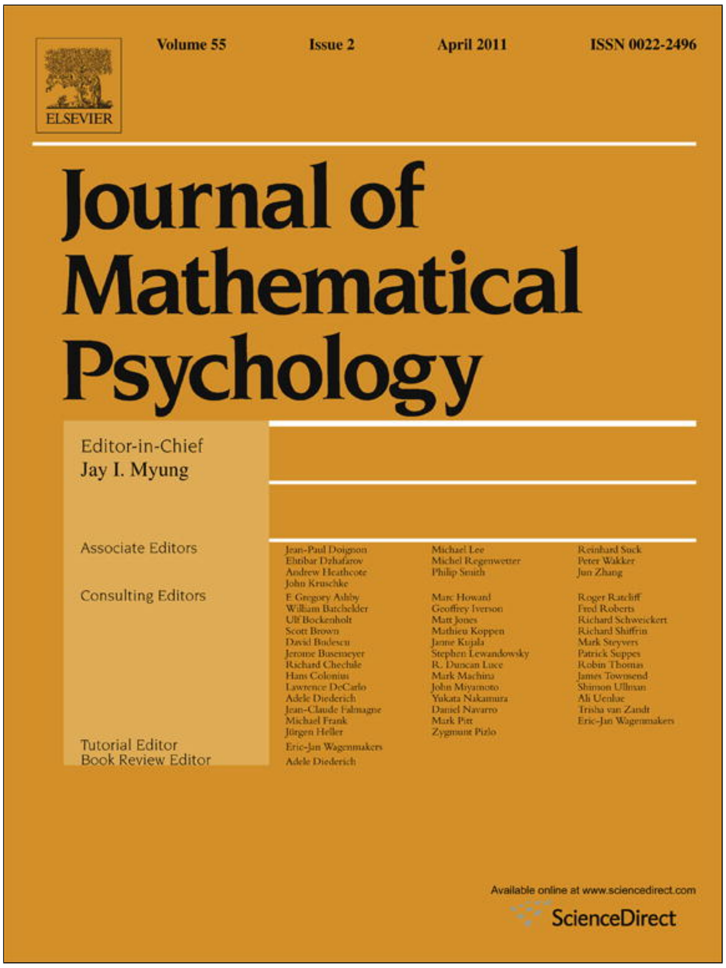
Volume 55
Issue 2
April 2011
ISSN 0022-2496
ELSEVIER
Journal of Mathematical Psychology
Editor-in-Chief
Jay I. Myung
Associate Editors
Consulting Editors
Tutorial Editor
Book Review Editor
Available online at www.sciencedirect.com
ScienceDirect

Journal of Mathematical Psychology 55 (2011) 152–165

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

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

A re-examination of the algebraic properties of the AHP as a ratio-scaling technique

Michele Bernasconi [a], Christine Choirat [b], Raffaello Seri [c,*]

[a] Dipartimento di Scienze Economiche, Università "Ca' Foscari", Cannaregio 873, I-30121 Venezia, Italy
[b] Department of Economics, School of Economics and Business Management, Universidad de Navarra, Edificio de Bibliotecas (Entrada Este), E-31080 Pamplona, Spain
[c] Dipartimento di Economia, Università dell'Insubria, Via Monte Generoso 71, I-21100 Varese, Italy

ARTICLE INFO

Article history:
Received 3 September 2009
Received in revised form
9 November 2010
Available online 11 February 2011

Keywords:
Separable representations
Reciprocally symmetric matrices
Consistency indexes
AHP

ABSTRACT

The Analytic Hierarchy Process (AHP) ratio-scaling approach is re-examined in view of the recent developments in mathematical psychology based on the so-called separable representations. The study highlights the distortions in the estimates based on the maximum eigenvalue method used in the AHP distinguishing the contributions due to random noises from the effects due to the nonlinearity of the subjective weighting function of separable representations. The analysis is based on the second order expansion of the Perron eigenvector and Perron eigenvalue in reciprocally symmetric matrices with perturbations. The asymptotic distributions of the Perron eigenvector and Perron eigenvalue are derived and related to the eigenvalue-based index of cardinal consistency used in the AHP. The results show the limits of using the latter index as a rule to assess the quality of the estimates of a ratio scale. The AHP method to estimate the ratio scales is compared with the classical ratio magnitude approach used in psychophysics.

1. Introduction

The Analytic Hierarchy Process (AHP) (Saaty, 1977, 1980, 1986) is an algebraic technique used in management to derive ratio scales of priorities from positive reciprocally symmetric matrices based on subjective ratio assessments. After an era in which mathematical psychologists considered ratio scaling to be fundamentally unsound (Shepard, 1981) in the past decades progress has been made to comprehend the structural assumptions needed to represent subjective measurement data (Luce & Narens, 2008). Here we study the implications of the recent developments in mathematical psychology for the algebraic techniques used by the AHP. In this introduction we motivate our analysis by reviewing the recent literature on ratio scaling and the AHP and present a summary of the main results of the paper.

1.1. Ratio scale, magnitude estimation and separable representations

As is well known, ratio-scaling procedures have been introduced in the behavioral sciences in the middle of the last century by psychophysicist Stevens (1951, 1957). In the simplest ratio-scaling experiment, also known as ratio magnitude estimation or magnitude estimation with a standard, an individual is asked to compare a set of stimuli $(x_1, \ldots, x_n)$ with a baseline stimulus x_0.[1] Each comparison yields a response value α_{i0}, with $i = 1, \ldots, n$. Stevens assumed that the values $\alpha_{10}, \ldots, \alpha_{n0}$ could directly represent a ratio scale, in the sense that he conjectured the existence of a ratio estimation function of the form $\alpha_{i0} = \left(\frac{x_i}{x_0}\right)^k$, which corresponds to his famous psychophysical law that equal physical ratios produce equal psychological ratios. Stevens' model has always been highly criticized by mathematical psychologists because it lacks normative and descriptive justifications (Michell, 1999, Chapter 4). Recently, however, mainly due to the work of Luce (2002, 2004) and Narens (1996, 2002, 2006), axiomatic developments of subjective measurement approaches have been obtained with stronger theoretical foundations. The new models belong to a class of so-called separable representations, which establish the following relationships between the stimuli and the responses of a ratio

* Corresponding author.
E-mail addresses: bernasconi@unive.it (M. Bernasconi), cchoirat@unav.es (C. Choirat), raffaello.seri@uninsubria.it (R. Seri).

0022-2496/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmp.2010.11.002

[1] The term magnitude estimation was introduced by Stevens and Galanter (1957) (see also Stevens, 1975). A dual scaling procedure, also widely used in psychophysics, is known as ratio production, in which an individual is asked to produce a stimulus s_i which appears p_i times more intense than a reference stimulus s_0 (see Luce, 2004; Steingrimsson & Luce, 2006, for the axiomatic treatments of the different cases of ratio magnitude estimation and ratio production).

estimation exercise:

$$W(\alpha_{i0}) = \frac{\psi(x_i)}{\psi(x_0)}. \tag{1}$$

In Eq. (1), ψ is called the psychophysical function and W the subjective weighting function. The two functions indicate that two independent transformations may occur in a ratio estimation: one of the stimuli intensities (embodied in ψ); and the other of proportions (entailed in W). The latter, in particular, is the key to interpret the number names used by subjects in the experiments, also called numerals, as scientific numbers. Technically, Eq. (1) is based on an assumption referred to by Narens (2006, p. 110) as "commutativity", meaning that, for example, a subjective proportion of 2 multiplied by a subjective proportion of 3 is equivalent to a subjective proportion of 3 multiplied by a subjective proportion of 2, since it is in fact $W(2) \cdot W(3) = W(3) \cdot W(2)$; though neither products of subjective proportions can be treated as equivalent to a subjective proportion of 6, which requires the full force of a property called by Narens (1996) "multiplicativity", implying the linearity of W. Support for separable representations and commutativity has been found in a series of recent experiments, but not for multiplicativity (Bernasconi, Choirat, & Seri, 2008; Ellermeier & Faulhammer, 2000; Steingrimsson & Luce, 2005a,b, 2006, 2007; Zimmer, 2005). It is also remarked that the subjective weighting function in separable representations is closely connected to the possibly most well-known nonlinear probability transformation function in descriptive models for choice under risk and uncertainty (as for example typified in Cumulative Prospect Theory, Tversky & Kahneman, 1992).

While generalizing Stevens' power law model, a characteristic of separable forms is that they maintain the deterministic approach of the former, in the sense that these theories do not involve considerations of errors. This is acknowledged as a limitation (Luce, 1997, p. 81, Narens, 1996, p. 109).

1.2. The AHP

A different approach which emphasizes the presence of inconsistencies in ratio-scaling procedures is the AHP (Saaty, 1980, 1986). In the AHP the vector of stimuli $(x_1, \ldots, x_n)$ corresponds to a set of items relevant in a decision problem. It can be a set of alternatives to be ordered, a set of criteria used to order the alternatives, or a set of outcomes to forecast. The purpose of the AHP is to obtain a vector of subjective priorities measuring on a ratio scale the decision maker's ordering of the items $(x_1, \ldots, x_n)$.[2] Two main differences characterize Saaty's approach from the traditional magnitude estimation. First of all, the AHP allows for errors in the subjective ratio judgments. Secondly, in order to improve the validity of the ratio-scaling procedure, in the AHP the decision maker is asked to compare, two at a time, all the different items $(x_1, \ldots, x_n)$, thus obtaining $n(n-1)/2$ subjective ratio assessments α_{ij}'s for the relative dominances between any pair (x_i, x_j). Subjective ratio judgments are then used to construct an entire $(n \times n)$ matrix $\mathbf{A} = [\alpha_{ij}]$, also known as the AHP response matrix, where it is imposed that $\alpha_{ii} = 1$ and $\alpha_{ji} = 1/\alpha_{ij}$. More formally, for all entries α_{ij} of $\mathbf{A}$ it is assumed that

$$\alpha_{ij} = \frac{u_{0i}}{u_{0j}} \cdot e_{ij} \tag{2}$$

where u_{0i} and u_{0j} are the underlying subjective priority weights belonging to vector $\mathbf{u}_0 = (u_{01}, \ldots, u_{0n})^{\mathsf{T}}$, with $u_{01} > 0, \ldots, u_{0n} > 0$ and by convention $\sum u_{0j} = 1$; and where e_{ij} is a multiplicative error term which, for the way in which matrix $\mathbf{A}$ is constructed, is assumed reciprocally symmetric $e_{ij} = e_{ji}^{-1}$, and with $e_{ii} = 1$ for all i.

The error terms e_{ij} allow for violations of consistency properties often observed in practice in the AHP. The central one is called by Saaty (1977) "cardinal" consistency, requiring that for any three ratio assessments $\alpha_{ij}, \alpha_{ik}, \alpha_{kj}$, the following holds: $\alpha_{ij} = \alpha_{ik} \cdot \alpha_{kj}$. Cardinal consistency is equivalent to multiplicativity in the sense of Narens (1996). A weaker requirement less often violated is "ordinal" consistency, namely that when $a_{ij} > 1$ and $a_{jk} > 1$ then also $a_{ik} > 1$.

The presence of the errors e_{ij} also poses the problem of estimating the priority vector $\mathbf{u}_0$ in an appropriate way. Classical AHP does not make assumptions about the stochastic structure of the error terms but obtains the priority weights using algebraic techniques. It takes as the best approximation for $\mathbf{u}_0$ the right Perron eigenvector of the actual response matrix $\mathbf{A}$, henceforth denoted as $\mathbf{u} = \mathbf{u}(\mathbf{A})$.[3] It is in particular well known by the Perron–Frobenius Theorem that the Perron eigenvector $\mathbf{u}$ is the unique solution of the system of equations

$$\mathbf{A}\mathbf{u} = \boldsymbol{\lambda}\mathbf{u}, \qquad \sum u_i = 1$$

where $\boldsymbol{\lambda}$ denotes the Perron root (maximum eigenvalue) of $\mathbf{A}$. Moreover, when $e_{ij} = 1$ for all $i, j = 1, \ldots, n$ so that any row $(\alpha_{i1}, \ldots, \alpha_{in})$ of $\mathbf{A}$ can be obtained from any other row by the relation $\alpha_{ik} = \alpha_{jk}/\alpha_{ji}$ (this implies that the rank $\mathrm{rk}(\mathbf{A})$ is equal to 1), it is known that the maximum eigenvalue (ME) method delivers the correct priority vector $\mathbf{u}_0$, with the maximum eigenvalue being at its minimum $\lambda = n$. Saaty's argument is to use the same approach even when $\mathbf{A}$ is not perfectly consistent, provided that inconsistencies fall within some bounds (see below).

The AHP has spawned a vast literature and several critiques have been considered against the method (Dyer, 1990). One includes the development of statistical approaches to estimate the priority vector $\mathbf{u}_0$ which pay more attention to the stochastic structure of the data and include the logarithmic least squares method as the most standard alternative (see Crawford & Williams, 1985; de Jong, 1984; Genest & Rivest, 1994).

1.3. AHP in separable form

Another important criticism concerns the fact that no formal justification exists for the ratio-scaling techniques used in the AHP.[4] Separable representations can provide such a justification. In fact, combining classical AHP with the separable form of Eq. (1) in Bernasconi, Choirat, and Seri (2010) we propose the following form for the ratio assessments of an AHP response matrix:

$$\alpha_{ij} = W^{-1}\left(\frac{\psi(x_i)}{\psi(x_j)}\right) \cdot e_{ij} \tag{3}$$

[2] More precisely, in the AHP decision problems are organized in a hierarchical structure and "local" priority weights are derived for the items relevant at every level of the hierarchy. Local priority weights are then combined linearly for associated items through all levels of the hierarchy to obtain "global" priorities. In this paper we do not deal with the function of hierarchical structuring of the AHP (for a recent exposition of all basic functions of the AHP see e.g. Forman & Gass, 2001; for critiques to the method Dyer, 1990).

[3] We follow the habit of calling "principal" or "Perron eigenvalue" the largest eigenvalue of a matrix with positive entries (that is real and unique by the Perron–Frobenius Theorem), and "Perron eigenvector" the eigenvector associated with it.

[4] More precisely, Saaty (1986) provides an axiomatization of the AHP which is concerned with the characterization and the operations which can be performed with the response matrix $\mathbf{A}$, but which does not explicitly deal with the structural assumptions implicit in ratio-scaling techniques for which it is in fact necessary to wait for the literature on separable representations.

where $W^{-1}(\cdot)$ is the inverse of a subjective weighting function from separable representations; $\psi(x_i), \ldots, \psi(x_n)$ are the psychological perceptions of the stimuli intensities corresponding to the priority weights $u_{0i} = \frac{\psi(x_i)}{\sum \psi(x_k)}$, for $i = 1, \ldots, n$; and where the e_{ij}'s are the multiplicative error terms of the AHP.

Eq. (3) represents a generalization of classical AHP in the sense that in Saaty's standard approach W^{-1} is the identity. It can account for systematic distortions in the perception of subjective ratios. The AHP error terms e_{ij}'s make instead explicit the important notion that people are not like robots, but various effects, like lapses of concentration, states of mind, trembling, rounding effects and computational mistakes, imply that no model of human behavior can be assumed to hold deterministically.[5] In our previous paper (Bernasconi et al., 2010) we develop a detailed analysis of representation (3) and propose a statistical method based on a polynomial generalization of the logarithmic least squares, to estimate the priority vector $\mathbf{u}_0$ when $W^{-1}(\cdot)$ is nonlinear and the errors e_{ij}'s are independent with common variance. Moreover, by applying the new estimation method to the data of a ratio estimation experiment, further evidence is provided about the importance of the nonlinearity $W^{-1}(\cdot)$ to generate systematic inconsistencies in subjective measurement data. In experimental contexts in which subjects were asked to conduct pure estimation tasks, we have estimated a concave inverse subjective weighting function, finding a tendency of a vast majority of subjects to underestimate ratios greater than one with underestimation increasing as ratios get increasingly larger.[6] Furthermore, the analysis showed that the inconsistencies in the empirical subjective response matrices due to the nonlinearity of $W^{-1}(\cdot)$ are substantially larger than the ones due to the random noise e_{ij}.

This paper is concerned with a different, but complementary, issue. It studies the effects of the separable forms and of the nonlinearity of the subjective weighting function $W(\cdot)$ for the mathematical behavior of Saaty's maximum eigenvalue method. Even if one does not regard the right Perron eigenvector $\mathbf{u} = \mathbf{u}(\mathbf{A})$ as the best method to estimate the priority vector $\mathbf{u}_0$, looking at its mathematical properties when representation (3) holds is still quite relevant. When there are no distortions, the principal eigenvector method is obviously the most natural method to recover the priority weights. So, it is quite important to know how the natural benchmark behaves when there are inconsistencies in the response data. Such an issue has been studied by Genest and Rivest (1994) for the case when the distortions in the response matrix $\mathbf{A}$ are due to the error terms e_{ij}'s. We extend their analysis to the case in which inconsistencies in the response data may also be due to the subjective weighting function $W^{-1}(\cdot)$.

1.4. Overview of the results and organization of the paper

We provide several results. First of all, we measure the extent to which the priority weights obtained by the principal eigenvector with the AHP depart from a ratio scale. We approach the problem using the theory of matrix differentials developed in Magnus and Neudecker (1999). We take the second order approximation of the Perron eigenvector $\mathbf{u}$ and of the Perron eigenvalue λ around their ideal values $\mathbf{u}_0$ and $\lambda_0 = n$. Our analysis shows that when the stimuli range is small and the errors e_{ij}'s are independent with common variance, the effects of the deterministic distortions due to $W^{-1}(\cdot)$ and the stochastic terms are comparable; whereas when the stimuli are very different the deterministic distortions are much larger than those due to the stochastic errors. This result provides a justification for a well-known axiom which Saaty (1986) introduced heuristically in the AHP, the so-called "homogeneity" axiom, which requires the stimuli used in the AHP to be in a range of comparability.

Under an appropriate hypothesis on the asymptotic distribution of the errors, we then study the approximate distribution of Saaty's maximum eigenvalue λ and eigenvector $\mathbf{u}$ when Eq. (3) holds. For the case in which W is linear, Genest and Rivest (1994) have shown that λ is upward biased with respect to $\lambda_0 = n$ with an approximate χ^2 distribution; and they have shown how to relate a statistically based measure of consistency with the index of cardinal consistency proposed by Saaty (1977, 1980) for classical AHP and based on the quantity $\mu = (\lambda - n)/(n - 1)$. In particular, a pragmatic procedure used by the AHP is the so-called 10% cut-off rule, which considers the degree of cardinal inconsistency in a response matrix $\mathbf{A}$ too large (and hence rejects the data) when its consistency index μ is larger than one tenth of the average index μ_0, computed from the average of a randomly generated reciprocally symmetric matrix of the same size as $\mathbf{A}$. Genest and Rivest (1994) have shown how the 10% cut-off rule is in fact equivalent to a statistical test of the hypothesis that the variability due to the error terms e_{ij}'s in the response data does not exceed some given threshold. We extend their results and show that when W is allowed to be different from the identity, then λ has approximately a normal distribution or a noncentral χ^2 distribution whose noncentrality parameter depends on the strength of the distortion induced by W. We show how in this case the use of the 10% cut-off rule can lead to severely undersized tests, in particular, when the deterministic perturbations due to W are larger than those caused by the error terms e_{ij}'s. We also derive the asymptotic distribution of $\mathbf{u}$ and show that it is always normal with a bias that depends on the subjective weighting function W and never on the error terms.

Finally, we compare the performance of the AHP with the classical ratio magnitude estimation used in psychophysics. As alluded above, one justification of Saaty to develop the AHP was the intuition that, since the AHP provides all different pairwise comparisons between the stimuli $(x_1, \ldots, x_n)$, it could improve the validity of a procedure like ratio magnitude estimation in which each element of $(x_1, \ldots, x_n)$ is considered only in comparison to a reference stimulus x_0. We show that when the principal eigenvector method is applied to a matrix in which the response data obey Eq. (3), the above intuition is not always valid, but it depends on how the reference stimulus x_0 is chosen and whether the homogeneity axiom is satisfied. We show that when it is respected and all the distortions in the response data are due to random noise, then the AHP is always better than ratio magnitude estimation to obtain an estimate of vector $\mathbf{u}_0$; whereas, when homogeneity is not respected and there are psychophysical distortions in the data, ratio magnitude estimation can be better despite the lower number of comparisons. We give examples when this can be the case. Moreover, since the AHP typically requires more comparisons than a classical ratio magnitude estimation, we also consider the performance of the two procedures when in ratio magnitude estimation each comparison to the reference stimulus is repeated several times with the random errors of all comparisons averaged out. Interestingly, we find that the repeated ratio magnitude estimation would overall require more comparisons than the AHP in order for it to obtain the same standard deviation as the AHP itself.

[5] Obviously, this does not mean that the multiplicative error term e_{ij} of Eq. (2) is the only possible way to take into account unpredictable inconsistencies in the AHP. For example, the experimental economic literature considers different ways of modelling random errors in behavioral data (Loomes, 2005). We will shortly come back to this point in Section 2.

[6] In the context of utility theory for risky gambles the subjective weighting function $W(\cdot)$ on the probability interval [0, 1] is usually estimated with an inverse *s*-shaped form (see e.g., Prelec, 1998; Wu & Gonzalez, 1996, and the references therein).

The paper is organized as follows. We start in Section 2 with a more thorough discussion of a model for the α_{ij}'s based on Eq. (3). In Section 3 we derive three main theorems for the second order expansions of the Perron eigenvector and Perron eigenvalue in reciprocally symmetric matrices with perturbations. In Section 4 these theorems are applied to the AHP to study the algebraic properties of Saaty's method and to derive the asymptotic distributions of interest. In Section 5 we conduct the comparison between the AHP and ratio magnitude estimation. Section 6 concludes.

2. A model for the α_{ij}'s

As indicated in the Introduction, following the recent literature in mathematical psychology, we consider the following model for the α_{ij}'s of an AHP response matrix $\mathbf{A}$:

$$\alpha_{ij} = W^{-1}\left(\frac{\psi(x_i)}{\psi(x_j)}\right) \cdot e_{ij} \tag{3}$$

where $\psi(\cdot)$ and $W(\cdot)$ are, respectively, the psychophysical and subjective weighting functions considered in separable representations and where e_{ij} is a multiplicative, numerical random term, introduced by classical AHP. Without loss of generality henceforth we write $e_{ij} = e^{\nu_{ij}}$.

Form (3) must in particular be considered an empirical realization of a ratio estimation task for which the theoretical model assumed by classical AHP is $\alpha_{0,ij} = \frac{\psi(x_i)}{\psi(x_j)}$. A matrix $\mathbf{A}_0$ filled with these elements is called by Saaty "perfectly consistent" (Saaty, 1977). The right Perron eigenvector and Perron eigenvalue of $\mathbf{A}_0$ are respectively given by:

$$\mathbf{u}_0 = \left[\frac{\psi(x_i)}{\sum_{k=1}^{n}\psi(x_k)}\right],$$

$$\lambda_0 = n.$$

Our objective is to approximate $\mathbf{u}$ and λ, obtained by Saaty's ME method, solutions of the system:

$$\mathbf{A}\cdot\mathbf{u} = \lambda\cdot\mathbf{u}, \tag{4}$$

as functions of the response matrix $\mathbf{A}$ of Eq. (3), which is a perturbation of $\mathbf{A}_0$.

In classical AHP the only perturbations considered are those due to the error term $e_{ij} = e^{\nu_{ij}}$, which by construction of $\mathbf{A}$ imposing the equalities $\alpha_{ii} = 1$ and $\alpha_{ji} = 1/\alpha_{ij}$, must be skew-symmetric $\nu_{ij} = -\nu_{ji}$ and with $\nu_{ii} = 0$. Moreover, perturbations must be small, so that $e^{\nu_{ij}}$ is close to 1. As specified in the introduction, in classical AHP the error term is thought to capture unpredictable violations, mainly of cardinal consistency, arising when $\alpha_{ij} \neq \alpha_{ik}\cdot\alpha_{kj}$. Violations of ordinal consistency, arising when $\alpha_{ij} > 1$ and $\alpha_{jk} > 1$ but $\alpha_{ik} < 1$, are also possible: for example when for a subject $\psi(x_i) > \psi(x_k)$, but she states a ratio $\alpha_{ik} < 1$. They are however less likely.[7]

No assumption is made in classical AHP about the stochastic structure of the ν_{ij}'s. In the first part of our analysis dedicated to various results on general reciprocally symmetric perturbed matrices (Section 3), we also do not make hypotheses about the stochastic structure of the errors.[8] We will introduce the assumption that the errors ν_{ij} are independent with common standard deviation equal to σ in order to study the magnitude of the perturbations in $\mathbf{u}$ and λ (namely, from Section 4 onwards). Moreover, for the analysis of the asymptotic distributions of $\mathbf{u}$ and λ (in Theorem 4.1), we will also require that the errors ν_{ij} are (asymptotically) standard normal random variables.

Both hypotheses are as in Genest and Rivest (1994).[9] More generally, they are coherent with a long psychophysical tradition, possibly going back to Gustav Fechner (1860/1999) and passing through Thurstone (1931, 1959), based on the assumption that people have well-defined systems of values, preferences, judgments, which they apply in actual choice situations with errors. In this respect it is however also important to remark that this standard approach to errors in behavioral data has become recently object of a debate in the experimental economic literature (e.g. Hey, 2005; Loomes, 2005; Myung, Karabatsos, & Iverson, 2005, for discussion and references). An alternative approach which has been proposed is called random preference model by Loomes and Sugden (1995). It suggests to consider people systems of values as inherently stochastic. That is, rather than supposing that each individual has a single true preference function to which a random error is added, noise and random inconsistencies in actual behavior arise because of the imprecision of people to know and therefore to apply the same specification of the theory every time it is used (antecedents of this approach are in Becker, DeGroot, & Marschak, 1963). While we consider the model of mind behind this second approach interesting and stimulating in general, we note that it is intimately incoherent with the AHP, which is indeed based on the idea that people have a well-defined system of priorities entailed in a unique vector $\mathbf{u}_0$, which the AHP techniques aim at recovering. The purpose of this paper is in this sense limited to study the mathematical relationships between the theoretical $\mathbf{u}_0$ and its perturbed counterpart $\mathbf{u}$.

A further remark is also necessary about the assumption that $\alpha_{ii} = 1$ in the main diagonal of $\mathbf{A}$. As noted above, in the AHP this is simply assumed, since no actual comparison is made with a stimulus compared to itself. For example, if one is estimating the distance between pairs of cities, a respondent will be never asked to compare the distance Vilnius–Bratislava to the distance Vilnius–Bratislava. This, however, is different from the case in which two equal stimuli appear among the n items to be compared. In this case random errors occur, as for example when one may be comparing the distance Vilnius–Bratislava to the distance Vilnius–Budapest (even if the two distances are almost identical) or even in a more extreme case when a subject is comparing two distinct line segments which are exactly equal in length.

The second perturbations considered in form (3) are those due to the subjective weighting function $W(\cdot)$. The same arguments about the construction of $\mathbf{A}$ (together with the restrictions on the error terms) require that the subjective weighting function $W(\cdot)$ is a (monotonic) reciprocally symmetric function, namely $W(1/\cdot) = 1/W(\cdot)$, with $W(1) = 1$. Both these conditions are implied by several derivations of separable representations, including a specification developed by Luce (2001, 2002), similar to one which Prelec (1998) proposed in the context of utility theory for risky gambles. A more recent specification proposed by Luce (2004) (and

[7] In particular, the fact that the noise is multiplicative implies that larger errors are associated with larger values of the true ratio, but in such a way that the possibility of reversal of dominance due to multiplication by a random error is reduced for larger ratios. In an empirical test of ordinal consistency conducted with 69 subjects performing three ratio estimation tasks, we found that very few subjects exhibited violations of ordinal consistency (only 5 subjects on one task, 2 subjects on a second task, and 0 on a third task, Bernasconi et al., 2010, p. 707).

[8] In fact, a part of the analysis (including Theorems 3.1 and 3.2) will also cover the more general cases in which neither $\alpha_{ji} = 1/\alpha_{ij}$ nor $\alpha_{ii} = 1$, and hence $\nu_{ij} = -\nu_{ji}$ and $\nu_{ii} = 0$, must necessarily hold.

[9] Weaker hypotheses are possible for the study of asymptotic distributions (see e.g. de Jong, 1984); but since we are mainly interested in deterministic distortions, we will retain the present one.

further discussed by Aczél & Luce, 2007) does not instead impose either reciprocal symmetry or $W(1) = 1$. A part of the analysis below (including Theorems 3.1 and 3.2 of the following section) will also cover this more general case.[10]

For the analysis to be conducted both perturbations caused by W^{-1} and by the error term $e_{ij} = e^{\nu_{ij}}$ must be small, namely asymptotically negligible, so that we can write (the parameters governing the asymptotic behavior are introduced below)[11]:

$$\alpha_{ij} = \alpha_{0,ij} \cdot \exp\left\{\ln\left[\frac{\psi(x_j)}{\psi(x_i)} \cdot W^{-1}\left(\frac{\psi(x_i)}{\psi(x_j)}\right)\right] + \nu_{ij}\right\}$$
$$= \alpha_{0,ij} \cdot e^{\mathrm{d}\varepsilon_{ij}} \tag{5}$$

where we define:

$$\mathrm{d}\varepsilon_{ij} = \ln\left[\frac{\psi(x_j)}{\psi(x_i)} \cdot W^{-1}\left(\frac{\psi(x_i)}{\psi(x_j)}\right)\right] + \nu_{ij}$$

for $j > i$. The reason for the use of the differential symbol d will be clear in the following: we will indeed suppose that α_{ij} is a small perturbation of $\alpha_{0,ij}$ so that $\mathrm{d}\varepsilon_{ij}$ is an infinitesimal quantity. The property of reciprocal symmetry holds when $\mathrm{d}\varepsilon_{ij} = -\mathrm{d}\varepsilon_{ji}$.

Consider the following approximation, used in Bernasconi et al. (2008), in which the function W^{-1} is first log-transformed to w^{-1}:

$$\ln W[\exp(\cdot)] = w(\cdot)$$
$$w^{-1}(\cdot) = \ln W^{-1}[\exp(\cdot)]$$

and w^{-1} is given by a polynomial in its argument (whose degree L can even be infinite):

$$w^{-1}(x) = \sum_{\ell=0}^{L} \phi_\ell \cdot x^\ell$$

with $\phi_0 = 0$, $\phi_1 = 1$ and $\phi_{2n} = 0$ for $n \in \mathbb{N}$. This is generally possible: according to the Weierstrass Approximation Theorem, any continuous function on a compact domain can be approximated to any desired degree of accuracy by a polynomial in its arguments. Therefore:

$$W^{-1}(x) = \exp\left\{\sum_{\ell=0}^{L} \phi_\ell \cdot [\ln(x)]^\ell\right\}$$
$$= x \cdot \exp\left\{\sum_{\ell=2}^{L} \phi_\ell \cdot [\ln(x)]^\ell\right\}$$

so that when $\|\phi_\ell\|_\infty = \max_{2\le\ell\le L} |\phi_\ell| \downarrow 0$, $W^{-1}(x) \to x$.

The infinitesimal error term is:

$$\mathrm{d}\varepsilon_{ij} = \ln\left[\frac{\psi(x_j)}{\psi(x_i)} \cdot W^{-1}\left(\frac{\psi(x_i)}{\psi(x_j)}\right)\right] + \nu_{ij}$$
$$= \sum_{\ell=2}^{L} \phi_\ell \cdot \left[\ln\left(\psi(x_i)/\psi(x_j)\right)\right]^\ell + \nu_{ij}. \tag{6}$$

Under the assumption that the errors ν_{ij}'s are independent with common standard deviation equal to σ, each ν_{ij} converges in probability to 0 when $\sigma \downarrow 0$. Together with $\|\phi_\ell\|_\infty = \max_{2\le\ell\le L} |\phi_\ell| \downarrow 0$, this implies that $\mathrm{d}\varepsilon_{ij}$ is asymptotically negligible.

[10] On the empirical evidence, several direct tests conducted in psychophysical experiments on loudness production, which include Steingrimsson and Luce (2007) and Zimmer (2005), have rejected the behavioral hypothesis underlying the specification with $W(1) = 1$ and have accepted one with $W(1) \neq 1$. Indirect tests based on the inference of separable forms in an experiment measuring the distance ratio between Italian cities and the area ratio between sections of a disk in Bernasconi et al. (2008) have not rejected $W(1) = 1$.

[11] Obviously, the question about the practical relevance of asymptotic results is an empirical one. Below we will shortly report on some evidence available on this issue.

3. Perturbations of reciprocally symmetric matrices with applications to the AHP

Here we show some useful facts about the differentials of reciprocally symmetric matrices, using the theory of matrix differentials in the sense of Magnus and Neudecker (1999). We start from some general results, which are then narrowed down to the case most interesting for the AHP.

Consider a matrix $\mathbf{A}$ and a reciprocally symmetric matrix $\mathbf{A}_0$ such that $\mathbf{A}$ can be considered a perturbation of $\mathbf{A}_0$. We write therefore $\mathbf{A} \simeq \mathbf{A}_0 + \mathrm{d}\mathbf{A} + \frac{1}{2}\mathrm{d}^2\mathbf{A}$ where $\mathrm{d}\mathbf{A}$ and $\mathrm{d}^2\mathbf{A}$ are matrix differentials (therefore asymptotically negligible) in the sense of Magnus and Neudecker (1999).[12] We want to study the behavior of the Perron eigenvalue λ and the right Perron eigenvector $\mathbf{u} = \mathbf{u}(\mathbf{A})$ of $\mathbf{A}$, taken as perturbations of the corresponding quantities λ_0 and $\mathbf{u}_0 = \mathbf{u}(\mathbf{A}_0)$, of the ideal system:

$$\mathbf{A}_0 \cdot \mathbf{u}_0 = \lambda_0 \cdot \mathbf{u}_0. \tag{7}$$

In particular, we want to obtain $\mathrm{d}\mathbf{u}$, $\mathrm{d}^2\mathbf{u}$, $\mathrm{d}\lambda$, $\mathrm{d}^2\lambda$ in the second order approximations of the right Perron eigenvector $\mathbf{u} \simeq \mathbf{u}_0 + \mathrm{d}\mathbf{u} + \frac{1}{2}\mathrm{d}^2\mathbf{u}$ of $\mathbf{A}$ and of the Perron eigenvalue $\lambda \simeq \lambda_0 + \mathrm{d}\lambda + \frac{1}{2}\mathrm{d}^2\lambda$.

The following notation will be used throughout the section. For a n-vector $\mathbf{a}$ let $\bar{\mathbf{a}}$ be the n-vector defined by $\bar{\mathbf{a}} = [\bar{a}_i] = \left[a_i^{-1}\right]$. $\mathbf{u}_n$ is a n-vector composed of ones. $\mathbf{I}_n$ is the (n, n)-identity matrix. $\mathbf{U}_n$ is a (n, n)-matrix composed of ones. $\mathbf{e}_i$ is a vector of zeros with a one in the ith position. Let:

$$\mathbf{K}_{nn} \triangleq \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbf{e}_i\mathbf{e}_j^\top \otimes \mathbf{e}_j\mathbf{e}_i^\top,$$

where $\otimes$ denotes the Kronecker product, be the commutation matrix (Magnus & Neudecker, 1999, p. 46). The notations $\overline{\ln}\mathbf{A}$, $\overline{\exp}\mathbf{A}$ and $\mathbf{A}^{\odot\ell}$ denote the elementwise application of natural logarithm, exponential and power function (of degree ℓ) to a matrix $\mathbf{A}$. On the other hand, $\mathbf{A}^\ell$ denotes the ordinary product of the matrix $\mathbf{A}$ by itself, repeated ℓ times. $\mathbf{A}^+$ is the Moore–Penrose inverse of the matrix $\mathbf{A}$.

Our first result yields the Perron eigenvalue and the right Perron eigenvector when $\mathbf{A}_0$ is a general reciprocally symmetric matrix, and the additive perturbation $\mathrm{d}\mathbf{A} + \frac{1}{2}\mathrm{d}^2\mathbf{A}$ does not necessarily yield a reciprocally symmetric matrix $\mathbf{A}$.

Theorem 3.1. *Consider the eigenvalue problems $\mathbf{A}_0 \cdot \mathbf{u}_0 = \lambda_0 \cdot \mathbf{u}_0$ and $\mathbf{A} \cdot \mathbf{u} = \lambda \cdot \mathbf{u}$ where $\mathbf{A} \simeq \mathbf{A}_0 + \mathrm{d}\mathbf{A} + \frac{1}{2}\mathrm{d}^2\mathbf{A}$ and $\mathrm{d}\mathbf{A}$ and $\mathrm{d}^2\mathbf{A}$ are the matrix differentials in the sense of Magnus and Neudecker (1999). λ_0 is a simple eigenvalue with right eigenvector $\mathbf{u}_0$ and left eigenvector $\mathbf{v}_0$. The right eigenvectors $\mathbf{u}_0$ and $\mathbf{u}$ are normalized as $\mathbf{u}_0^\top\mathbf{u}_n = 1$ and $\mathbf{u}^\top\mathbf{u}_n = 1$. The following expansions hold:*

$$\lambda(\mathbf{A}) \simeq n + \mathrm{d}\lambda + \frac{1}{2}\mathrm{d}^2\lambda,$$

$$\mathbf{u}(\mathbf{A}) \simeq \mathbf{u}_0 + \mathrm{d}\mathbf{u} + \frac{1}{2}\mathrm{d}^2\mathbf{u},$$

where $\mathrm{d}\lambda$, $\mathrm{d}^2\lambda$, $\mathrm{d}\mathbf{u}$ and $\mathrm{d}^2\mathbf{u}$ are given in Box I.

Theorem 3.1 is the most general theorem for the second order expansions of $\mathbf{u}$ and λ when $\mathbf{A}_0$ is a reciprocally symmetric matrix. The next theorem holds when matrix $\mathbf{A}_0$ is also perfectly consistent, but $\mathbf{A}$ need not be reciprocally symmetric. $\mathbf{A}_0$ is perfectly consistent if there exists an n-vector $\mathbf{u}$ such that:

$$\mathbf{A}_0 = \mathbf{u}\bar{\mathbf{u}}^\top,$$

[12] We use the notation $\simeq$ in order to intend that a valid Taylor development up to the largest order involved holds (see Magnus & Neudecker, 1999).

$$\mathrm{d}\lambda = \frac{\mathbf{v}_0^{\mathrm{T}}\mathrm{d}\mathbf{A}\mathbf{u}_0}{\mathbf{v}_0^{\mathrm{T}}\mathbf{u}_0}$$

$$\mathrm{d}^2\lambda = \frac{\mathbf{v}_0^{\mathrm{T}}\mathrm{d}^2\mathbf{A}\mathbf{u}_0 + 2\mathbf{v}_0^{\mathrm{T}}\cdot\mathrm{d}\mathbf{A}\cdot\left(\mathbf{I}_n - \frac{\mathbf{u}_0\mathbf{v}_0^{\mathrm{T}}}{\mathbf{v}_0^{\mathrm{T}}\mathbf{u}_0}\right)\cdot(\lambda_0\mathbf{I}_n - \mathbf{A}_0)^+\left(\mathbf{I}_n - \frac{\mathbf{u}_0\mathbf{v}_0^{\mathrm{T}}}{\mathbf{v}_0^{\mathrm{T}}\mathbf{u}_0}\right)\cdot\mathrm{d}\mathbf{A}\cdot\mathbf{u}_0}{\mathbf{v}_0^{\mathrm{T}}\cdot\mathbf{u}_0}$$

$$\mathrm{d}\mathbf{u} = (\mathbf{I}_n - \mathbf{u}_0\mathbf{u}_n^{\mathrm{T}})\cdot(\lambda_0\mathbf{I}_n - \mathbf{A}_0)^+\left(\mathbf{I}_n - \frac{\mathbf{u}_0\mathbf{v}_0^{\mathrm{T}}}{\mathbf{v}_0^{\mathrm{T}}\mathbf{u}_0}\right)\mathrm{d}\mathbf{A}\mathbf{u}_0$$

$$\begin{aligned}\mathrm{d}^2\mathbf{u} &= \left(\mathbf{I}_n - \mathbf{u}_0\mathbf{u}_n^{\mathrm{T}}\right)\cdot(\lambda_0\mathbf{I}_n - \mathbf{A}_0)^+\left(\mathrm{d}^2\mathbf{A} - \mathrm{d}^2\lambda\mathbf{I}_n\right)\mathbf{u}_0 \\ &\quad + 2\left(\mathbf{I}_n - \mathbf{u}_0\mathbf{u}_n^{\mathrm{T}}\right)(\lambda_0\mathbf{I}_n - \mathbf{A}_0)^+(\mathrm{d}\mathbf{A} - \mathrm{d}\lambda\mathbf{I}_n)(\lambda_0\mathbf{I}_n - \mathbf{A}_0)^+\left(\mathbf{I}_n - \frac{\mathbf{u}_0\mathbf{v}_0^{\mathrm{T}}}{\mathbf{v}_0^{\mathrm{T}}\mathbf{u}_0}\right)\mathrm{d}\mathbf{A}\mathbf{u}_0 \\ &\quad - 2\left(\mathbf{I}_n - \mathbf{u}_0\mathbf{u}_n^{\mathrm{T}}\right)(\lambda_0\mathbf{I}_n - \mathbf{A}_0)^+\left(\mathbf{I}_n - \frac{\mathbf{u}_0\mathbf{v}_0^{\mathrm{T}}}{\mathbf{v}_0^{\mathrm{T}}\mathbf{u}_0}\right)\mathrm{d}\mathbf{A}\mathbf{u}_0\cdot\mathbf{u}_n^{\mathrm{T}}(\lambda_0\mathbf{I}_n - \mathbf{A}_0)^+\left(\mathbf{I}_n - \frac{\mathbf{u}_0\mathbf{v}_0^{\mathrm{T}}}{\mathbf{v}_0^{\mathrm{T}}\mathbf{u}_0}\right)\mathrm{d}\mathbf{A}\mathbf{u}_0.\end{aligned}$$

Box I.

with $\mathbf{u}_n^{\mathrm{T}}\cdot\mathbf{u} = 1$. As concerns the right Perron eigenvector $\mathbf{u}_0 = \mathbf{u}(\mathbf{A}_0)$, the solution is given by $\mathbf{u}_0 = \mathbf{u}$ and $\lambda_0 = n$, since:

$$\mathbf{u}\cdot[\bar{\mathbf{u}}^{\mathrm{T}}\mathbf{u}] = \lambda_0\cdot\mathbf{u}$$
$$n\cdot\mathbf{u} = \lambda_0\cdot\mathbf{u}.$$

The fact that λ_0 is the Perron eigenvalue (and $\mathbf{u}_0$ the corresponding Perron eigenvector) can be shown by remarking that $\mathbf{A}_0$ has $(n-1)$ eigenvalues equal to 0 (since it has rank 1) and therefore its only nonnull eigenvalue is equal to its trace, that is n.

Theorem 3.2. *When $\mathbf{A}_0$ is a reciprocally symmetric consistent matrix, the following expansion holds:*

$$\lambda(\mathbf{A}) \simeq n + \mathrm{d}\lambda + \frac{1}{2}\mathrm{d}^2\lambda,$$
$$\mathbf{u}(\mathbf{A}) \simeq \mathbf{u}_0 + \mathrm{d}\mathbf{u} + \frac{1}{2}\mathrm{d}^2\mathbf{u},$$

where:

$$\mathrm{d}\lambda = \frac{1}{n}\bar{\mathbf{u}}_0^{\mathrm{T}}\mathrm{d}\mathbf{A}\mathbf{u}_0$$

$$\mathrm{d}^2\lambda = \frac{1}{n}\bar{\mathbf{u}}_0^{\mathrm{T}}\mathrm{d}^2\mathbf{A}\mathbf{u}_0 + \frac{2}{n^3}\bar{\mathbf{u}}_0^{\mathrm{T}}\mathrm{d}\mathbf{A}\left(n\mathbf{I}_n - \mathbf{u}_0\bar{\mathbf{u}}_0^{\mathrm{T}}\right)\mathrm{d}\mathbf{A}\mathbf{u}_0$$

$$\mathrm{d}\mathbf{u} = \frac{1}{n}(\mathbf{I}_n - \mathbf{u}_0\mathbf{u}_n^{\mathrm{T}})\mathrm{d}\mathbf{A}\mathbf{u}_0$$

$$\begin{aligned}\mathrm{d}^2\mathbf{u} &= \frac{1}{n}\left(\mathbf{I}_n - \mathbf{u}_0\mathbf{u}_n^{\mathrm{T}}\right)\left(\mathbf{I}_n - \frac{\bar{\mathbf{u}}_0\bar{\mathbf{u}}_0^{\mathrm{T}}}{\bar{\mathbf{u}}_0^{\mathrm{T}}\bar{\mathbf{u}}_0}\right)\left(\mathrm{d}^2\mathbf{A} - \mathrm{d}^2\lambda\mathbf{I}_n - \frac{2}{n}\mathrm{d}\lambda\mathrm{d}\mathbf{A}\right)\mathbf{u}_0 \\ &\quad + \frac{2}{n^2}\left(\mathbf{I}_n - \mathbf{u}_0\mathbf{u}_n^{\mathrm{T}}\right)\left(\frac{\mathbf{u}_0\mathbf{u}_0^{\mathrm{T}}}{\mathbf{u}_0^{\mathrm{T}}\mathbf{u}_0}\mathrm{d}\lambda - \frac{\bar{\mathbf{u}}_0\bar{\mathbf{u}}_0^{\mathrm{T}}}{\bar{\mathbf{u}}_0^{\mathrm{T}}\bar{\mathbf{u}}_0}\mathrm{d}\mathbf{A}\right. \\ &\quad \left. + \mathrm{d}\mathbf{A}\left(\mathbf{I}_n - \mathbf{u}_0\mathbf{u}_n^{\mathrm{T}}\right)\right)\mathrm{d}\mathbf{A}\mathbf{u}_0.\end{aligned}$$

As said above, this theorem does not require that the matrix $\mathbf{A}$ be reciprocally symmetric. With reference to the model for the α_{ij}'s discussed in Section 2, Theorem 3.2 may in particular apply when $\mathbf{A}_0$ is filled with elements $\alpha_{0,ij} = \frac{\psi(x_i)}{\psi(x_j)}$, but the standard AHP restrictions $\alpha_{ii} = 1$ and $\alpha_{ji} = 1/\alpha_{ij}$ do not necessarily hold. Such a case could be relevant when an individual is asked to fill the whole AHP response matrix $\mathbf{A}$ (and not only $n(n-1)/2$ elements as customarily done) and either the errors do not satisfy the restrictions $\nu_{ii} = 0$ and $\nu_{ij} = -\nu_{ji}$, or the individual is endowed with a general separable form in which $W(1/\cdot) = 1/W(\cdot)$ and $W(1) = 1$ fail, or when both types of failure occur.

When $\mathbf{A}$ is reciprocally symmetric and the restrictions $\alpha_{ii} = 1$ and $\alpha_{ji} = 1/\alpha_{ij}$ are satisfied, the expansions for the Perron eigenvalue and Perron eigenvector can be further specialized. In particular, in what follows we take $\mathrm{d}\mathbf{A}$ and $\mathrm{d}^2\mathbf{A}$, the perturbations of $\mathbf{A}$, to be determined according to $\alpha_{ij} = \alpha_{0,ij}\cdot e^{\mathrm{d}\varepsilon_{ij}} \simeq \alpha_{0,ij}\cdot\left(1 + \mathrm{d}\varepsilon_{ij} + \frac{1}{2}(\mathrm{d}\varepsilon_{ij})^2\right)$ where $\mathrm{d}\varepsilon_{ij}$'s are the entries of the skew-symmetric matrix:

$$[\mathrm{d}\mathbf{E}]_{ij} = \begin{cases}\mathrm{d}\varepsilon_{ij} & \text{if } j > i \\ -\mathrm{d}\varepsilon_{ij} & \text{if } i > j.\end{cases}$$

We have $\mathbf{A} = \mathbf{A}_0 \odot \overline{\exp}(\mathrm{d}\mathbf{E})$, $\mathrm{d}\mathbf{A} = \mathbf{A}_0 \odot \mathrm{d}\mathbf{E}$ and $\mathrm{d}^2\mathbf{A} = \mathbf{A}_0 \odot \mathrm{d}\mathbf{E} \odot \mathrm{d}\mathbf{E}$, where $\odot$ is the Hadamard or Schur or elementwise product of matrices (see Magnus & Neudecker, 1999, p. 45). In this case the following theorem is obtained.

Theorem 3.3. *Consider the eigenvalue problems $\mathbf{A}_0\cdot\mathbf{u}_0 = \lambda_0\cdot\mathbf{u}_0$ and $\mathbf{A}\cdot\mathbf{u} = \lambda\cdot\mathbf{u}$ where $\mathbf{A} = \mathbf{A}_0 \odot \overline{\exp}(\mathrm{d}\mathbf{E})$. Let $\mathbf{A}_0$ be a reciprocally symmetric and perfectly consistent matrix and $\mathrm{d}\mathbf{E}$ be a skew-symmetric matrix of perturbations. In this case:*

$$\mathrm{d}\lambda = 0$$

$$\mathrm{d}^2\lambda = \frac{1}{n^2}\mathbf{u}_n^{\mathrm{T}}\left(n\left(\mathrm{d}\mathbf{E}\odot\mathrm{d}\mathbf{E}\right) + 2\mathrm{d}\mathbf{E}\mathrm{d}\mathbf{E}\right)\mathbf{u}_n$$

$$\mathrm{d}\mathbf{u} = \frac{1}{n}\left(\mathrm{diag}(\mathbf{u}_0) - \mathbf{u}_0\mathbf{u}_0^{\mathrm{T}}\right)\mathrm{d}\mathbf{E}\mathbf{u}_n$$

$$\begin{aligned}\mathrm{d}^2\mathbf{u} &= \frac{1}{n^2}\left(\mathrm{diag}(\mathbf{u}_0) - \mathbf{u}_0\mathbf{u}_0^{\mathrm{T}}\right) \\ &\quad \times \left(n\left(\mathrm{d}\mathbf{E}\odot\mathrm{d}\mathbf{E}\right) + 2\mathrm{d}\mathbf{E}\left(\mathbf{I}_n - \mathbf{u}_n\mathbf{u}_0^{\mathrm{T}}\right)\mathrm{d}\mathbf{E}\right)\mathbf{u}_n.\end{aligned} \tag{8}$$

Theorem 3.3 is central for analyzing Saaty's eigenvalue method in the AHP whenever there are small distortions in the responses which preserve the property of reciprocal symmetry of the response matrix $\mathbf{A}$, because it gives a measure of the perturbation which depends on the whole matrix and not only on the individual elements α_{ij}'s (as for example in Saaty, 1977, 1980, 1986).

4. Properties of Saaty's method

In the following we will show how Theorem 3.3 can be used to study the effect of systematic biases and individual variabilities on Saaty's maximum eigenvector $\mathbf{u}$ and maximum eigenvalue λ when form (3) discussed in Section 2 applies. The analysis to be conducted is in this sense a generalization of Genest and Rivest (1994) whose study of the asymptotic distributions of λ and $\mathbf{u}$ is restricted to the case in which W is equal to the identity. Throughout the analysis we assume the following restrictions: $\nu_{ii} = 0$ and $\nu_{ij} = -\nu_{ji}$ for the error terms and $W(1/\cdot) = 1/W(\cdot)$ and $W(1) = 1$ for the subjective weighting function. Moreover, as in Genest and Rivest (1994), we suppose that the errors ν_{ij}'s have common standard deviation σ, but no restriction is made on their distribution unless explicitly stated. Under these hypotheses we rewrite the expressions for $\mathrm{d}\mathbf{u}$, $\mathrm{d}\lambda$, $\mathrm{d}^2\lambda$ from Theorem 3.3 using Eq. (6) of Section 2. In particular, rewriting Eq. (6) as[13]:

$$\mathrm{d}\mathbf{E} = \sum_{\ell=2}^{L} \phi_\ell \cdot [\overline{\ln}\mathbf{A}_0]^{\odot \ell} + \mathbf{N} + o(\|\phi_\ell\|_\infty) + o_{\mathbb{P}}(\sigma)$$

where $\mathbf{N} \triangleq [\nu_{ij}]$ is the matrix of random errors, the relevant terms of Theorem 3.3 become[14]:

$$\begin{aligned}
\mathrm{d}\lambda &= 0,\\
\mathrm{d}^2\lambda &= \left[\sum_{\ell=2}^{L} \phi_\ell \left[\left(\mathbf{I}_{n^2} - \mathbf{K}_{nn}\right) \cdot \mathbf{u}_n \otimes \left(\overline{\ln}\mathbf{u}_0\right)\right]^{\odot \ell} + \mathrm{vec}(\mathbf{N})\right]^{\mathsf{T}}\\
&\quad \times \left\{\frac{2\mathbf{K}_{nn}\left(\mathbf{U}_n \otimes \mathbf{I}_n\right)}{n^2} + \frac{\mathbf{I}_{n^2}}{n}\right\}\\
&\quad \times \left[\sum_{\ell=2}^{L} \phi_\ell \left[\left(\mathbf{I}_{n^2} - \mathbf{K}_{nn}\right) \cdot \mathbf{u}_n \otimes \left(\overline{\ln}\mathbf{u}_0\right)\right]^{\odot \ell} + \mathrm{vec}(\mathbf{N})\right]\\
&\quad + o(\|\phi_\ell\|_\infty^2) + o_{\mathbb{P}}\left(\sigma^2\right),\\
\mathrm{d}\mathbf{u} &= \frac{\sum_{\ell=2}^{L} \phi_\ell}{n} \cdot \left(\mathbf{u}_n^{\mathsf{T}} \otimes \left(\mathrm{diag}\left(\mathbf{u}_0\right) - \mathbf{u}_0 \cdot \mathbf{u}_0^{\mathsf{T}}\right)\right)\\
&\quad \times \left[\left(\mathbf{I}_{n^2} - \mathbf{K}_{nn}\right) \cdot \mathbf{u}_n \otimes \left(\overline{\ln}\mathbf{u}_0\right)\right]^{\odot \ell}\\
&\quad + \frac{1}{n} \cdot \left(\mathbf{u}_n^{\mathsf{T}} \otimes \left(\mathrm{diag}(\mathbf{u}_0) - \mathbf{u}_0 \cdot \mathbf{u}_0^{\mathsf{T}}\right)\right)\\
&\quad \times \mathrm{vec}\left(\mathbf{N}\right) + o\left(\|\phi_\ell\|_\infty\right) + o_{\mathbb{P}}(\sigma).
\end{aligned} \tag{9}$$

4.1. Deterministic and stochastic components in the first order differential of Saaty's priority vector

We remark that the expression for $\mathrm{d}\mathbf{u}$ derived above shows that $\mathbf{u} - \mathbf{u}_0 \simeq \mathrm{d}\mathbf{u}$ can be separated into two additive parts, a deterministic one depending only on the distortions due to W (as expressed through the coefficients $\phi_2, \ldots, \phi_L$) and a stochastic one depending on the random errors (as measured by the standard deviation σ). This is particularly appealing since it allows us to formulate some prescriptive devices concerning the relative contributions of the deterministic and the stochastic components as functions of the values taken by n and $\mathbf{u}_0$.

In order to evaluate the relative contributions of the deterministic and stochastic components of Eq. (9), we measure the first one through the "bias" $\|\mathbb{E}\mathbf{u} - \mathbf{u}_0\| \simeq \|\mathbb{E}\mathrm{d}\mathbf{u}\|$ and the second one through the "standard deviation" $\sqrt{\mathbb{E}\|\mathbf{u} - \mathbb{E}\mathbf{u}\|^2} \simeq \sqrt{\mathbb{E}\|\mathrm{d}\mathbf{u} - \mathbb{E}\mathrm{d}\mathbf{u}\|^2}$. Notice that in general these quantities depend on n and $\mathbf{u}_0$. In particular, when we suppose that $\phi_3 \neq 0$ while $\phi_\ell = 0$ for $\ell \geq 4$, the quantities $\frac{\|\mathbb{E}\mathrm{d}\mathbf{u}\|}{|\phi_3|}$ and $\frac{\sqrt{\mathbb{E}\|\mathrm{d}\mathbf{u} - \mathbb{E}\mathrm{d}\mathbf{u}\|^2}}{\sigma}$ depend only on n and $\mathbf{u}_0$:

$$\begin{aligned}
\frac{\|\mathbb{E}\mathrm{d}\mathbf{u}\|}{|\phi_3|} &= \left\| \frac{1}{n} \cdot \left(\mathbf{u}_n^{\mathsf{T}} \otimes \left(\mathrm{diag}\left(\mathbf{u}_0\right) - \mathbf{u}_0 \cdot \mathbf{u}_0^{\mathsf{T}}\right)\right) \right.\\
&\quad \left. \times \left[\left(\mathbf{I}_{n^2} - \mathbf{K}_{nn}\right) \cdot \mathbf{u}_n \otimes \left(\overline{\ln}\mathbf{u}_0\right)\right]^{\odot 3} \right\|
\end{aligned}$$

and:

$$\frac{\sqrt{\mathbb{E}\left\|\mathrm{d}\mathbf{u} - \mathbb{E}\mathrm{d}\mathbf{u}\right\|^2}}{\sigma} = \frac{1}{n^{1/2}} \cdot \sqrt{\sum_{i=1}^{n} u_{0,i}^2 + \left(\sum_{i=1}^{n} u_{0,i}^2\right)^2 - 2\sum_{i=1}^{n} u_{0,i}^3}.$$

The graphs in Fig. 1 show the two quantities for different choices of n and of $\mathbf{u}_0$ when $\phi_3 \neq 0$. With the names "Constant", "Logarithmic", "Square root", "Linear", "Square" and "Exponential" we denote, in this order, the vectors with $u_{0,j} \propto 1$, $u_{0,j} \propto \ln(1+j)$, $u_{0,j} \propto j^{1/2}$, $u_{0,j} \propto j$, $u_{0,j} \propto j^2$ and $u_{0,j} \propto \exp(j-1)$. These vectors represent situations of increasing dispersion of the real values of the weights. In this sense they relate to the so-called homogeneity axiom of the AHP. In particular, Saaty has always argued that "homogeneity is essential for comparing similar things, as the mind tends to make large errors in comparing widely disparate elements. For example we cannot compare a grain of sand with an orange according to size" (Saaty, 1986, p. 846).

The graphs in the figure clarify the nature of this heuristic argument and the role of the cognitive or deterministic distortions due to the subjective weighting function W in it. The thick black lines in the graphs represent $\frac{\|\mathbb{E}\mathrm{d}\mathbf{u}\|}{|\phi_3|}$ as a function of n while the thick grey lines represent $\frac{\sqrt{\mathbb{E}\|\mathrm{d}\mathbf{u} - \mathbb{E}\mathrm{d}\mathbf{u}\|^2}}{\sigma}$ as a function of n. The index on the x axis starts at 2 since for $n = 1$ both measures are identically 0. The first graphs in the figure are more likely to respect Saaty's homogeneity requirement while the last ones are more prone not to respect it. The graphs show that if $|\phi_3| = \sigma$, the effect of deterministic distortions (multiplied by ϕ_3) is larger than the effect of stochastic ones (multiplied by σ) when the elements of $\mathbf{u}_0$ are very different, while the effects of stochastic distortions are larger when the elements are quite similar.

As a final question, one may wonder how large are comparatively the deterministic and the stochastic part of $\mathrm{d}\mathbf{u}$ (respectively due to the function W and random noise) with respect to each other and how large is $\mathrm{d}^2\mathbf{u}$ with respect to $\mathrm{d}\mathbf{u}$; that is, more generally, an important question concerns the accuracy in practice of the asymptotic analytic results obtained. The answer is obviously an empirical one. A method for the estimation of $\mathbf{u}_0$ and of ϕ_3 and the recovery of the residuals ν_{ij}'s is proposed in Bernasconi et al. (2010). In that source, the method is applied to an experiment in which 69 individuals are required to elicit three comparison matrices respectively concerning 5 probabilities from games of chance, 5 distances of Italian cities from Milan, and the rainfalls in 5 European cities in November 2001. For each individual and each part of the experiment, we obtain estimates $\widehat{\mathbf{u}}_0$ and $\widehat{\phi}_3$ and residuals $\widehat{\nu}_{ij}$'s. Using these values, we compute the norm of the deterministic distortion $\left\|\frac{\phi_3}{n} \cdot \left(\mathbf{u}_n^{\mathsf{T}} \otimes \left(\mathrm{diag}(\mathbf{u}_0) - \mathbf{u}_0 \cdot \mathbf{u}_0^{\mathsf{T}}\right)\right) \cdot \left[\left(\mathbf{I}_{n^2} - \mathbf{K}_{nn}\right) \cdot \mathbf{u}_n \otimes \left(\overline{\ln}\mathbf{u}_0\right)\right]^{\odot 3}\right\|$ and of the stochastic one $\|\frac{1}{n} \cdot \left(\mathbf{u}_n^{\mathsf{T}} \otimes \left(\mathrm{diag}(\mathbf{u}_0) - \mathbf{u}_0 \cdot \mathbf{u}_0^{\mathsf{T}}\right)\right) \cdot \mathrm{vec}(\mathbf{N})\|$, where parameters $\mathbf{u}_0$ and ϕ_3, and errors ν_{ij}'s are replaced by estimates $\widehat{\mathbf{u}}_0$ and $\widehat{\phi}_3$, and residuals $\widehat{\nu}_{ij}$'s. In Table 1, we display the median and the first and third quartiles of the values computed on the 69 individuals for each part of the experiment.

From the figures reported in the table it is quite clear that the deterministic distortion is much larger than the stochastic one and that the numbers are quite similar across experiments. Using the estimates it is also possible to compute the values of $\mathrm{d}\mathbf{u}$ and $\mathrm{d}^2\mathbf{u}$. Therefore we evaluated the ratio of the absolute value of the first order approximation with respect to the second order one for all items and individuals. For all experiments, the median value of this ratio is around 6, precisely 6.196041 while the average value is around 60, precisely 63.85886.[15] This seems

[13] We recall that here and in what follows $o_{\mathbb{P}}(\delta)$ denotes a random variable such that $o_{\mathbb{P}}(\delta)/\delta$ converges in probability to 0 as $\delta \to 0$.

[14] Since, as will be apparent later on, the analytic results depend only marginally on $\mathrm{d}^2\mathbf{u}$, we do not need to rewrite it. Moreover, as we shall argue, the empirical evidence available supports the use of a first order approximation.

[15] Computations available from the authors.

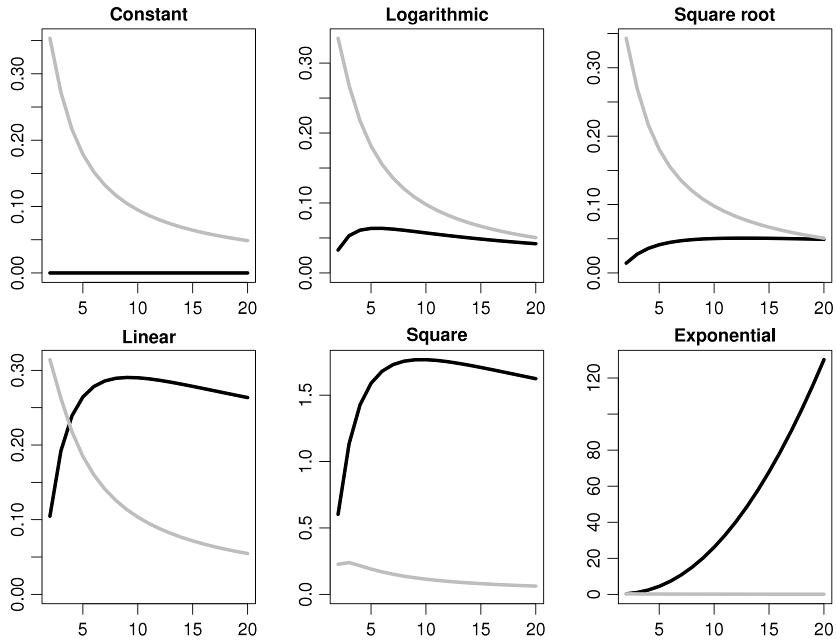

Fig. 1. $\frac{\|\mathbb{E}d\mathbf{u}\|}{|\phi_3|}$ (black) and $\frac{\sqrt{\mathbb{E}\|d\mathbf{u}-\mathbb{E}d\mathbf{u}\|^2}}{\sigma}$ (grey) for different n and $\mathbf{u}_0$.

Table 1
Decomposition of d**u** between deterministic and stochastic components estimated from data in Bernasconi et al. (2010).

		First quartile	Median	Third quartile
Deterministic	Probabilities	0.032002880	0.084831248	0.170602611
	Distances	0.032054720	0.091971090	0.146553681
	Rainfalls	0.060843732	0.086916276	0.134254159
Stochastic	Probabilities	0.0021459689	0.0078615956	0.0213969436
	Distances	0.001922213	0.004797563	0.01155813
	Rainfalls	0.003851987	0.00859049	0.02702352

to justify the approach, even if it may be important to collect further evidence since the accuracy of the approximation may also depend on contexts. Moreover, the computation of the second order approximation may give some guidance on the quality of the first order approximation.

4.2. Asymptotic distributions of Saaty's eigenvector and eigenvalue

Now, we derive the asymptotic distributions of Saaty's eigenvector $\mathbf{u}$ and eigenvalue λ as $\sigma \downarrow 0$ and as $W^{-1}(x) \to x$ (as measured by the coefficients of the polynomial through the condition $\|\phi_\ell\|_\infty \downarrow 0$). In order to conduct the analysis we introduce the assumption that the errors ν_{ij}'s are asymptotically distributed as independent normal random variables with common standard deviation σ. We express the results on the eigenvalue in terms of Saaty's consistency index $\mu = (\lambda - n)/(n-1)$ to facilitate their interpretation and comparison with the previous literature. The results are summarized in the following theorem. It shows that the asymptotic distributions depend on the speed of convergence to 0 of σ and $\|\phi_\ell\|_\infty$.

Theorem 4.1. *Suppose that* $\max\left\{\|\phi_\ell\|_\infty, \sigma\right\} \downarrow 0$. *Consider the errors* ν_{ij}*'s with* $i < j$*: suppose that they are (asymptotically) independent of each other and* $\sigma^{-1}\nu_{ij} \to_{\mathcal{D}} \mathcal{N}(0,1)$ *for every* $i < j$*. Define* $\boldsymbol{\mu}_\phi \triangleq \sum_{\ell=2}^{L} \phi_\ell \cdot \left[\left(\mathbf{I}_{n^2} - \mathbf{K}_{nn}\right) \cdot \mathbf{u}_n \otimes \left(\overline{\ln}\mathbf{u}_0\right)\right]^{\odot \ell}$, $\boldsymbol{\mu} \triangleq \lim_{\|\phi_\ell\|_\infty \downarrow 0} \frac{\boldsymbol{\mu}_\phi}{\|\phi_\ell\|_\infty}$ *and* $\mathbf{U}_0 \triangleq \operatorname{diag}(\mathbf{u}_0) - \mathbf{u}_0 \cdot \mathbf{u}_0^{\mathsf{T}}$.

(a) *If* $\frac{\sigma}{\|\phi_\ell\|_\infty} \to +\infty$*:*

(i)

$$\sigma^{-1}(\mathbf{u} - \mathbf{u}_0) \to_{\mathcal{D}} \mathcal{N}\left(\mathbf{0}, \frac{1}{n} \cdot \mathbf{U}_0^2\right);$$

(ii)

$$\sigma^{-2}\mu \to_{\mathcal{D}} \frac{1}{n(n-1)} \cdot \chi^2\left(\frac{(n-1)(n-2)}{2}\right).$$

(b) *If* $\frac{\sigma}{\|\phi_\ell\|_\infty} \to c$*:*

(i)

$$\sigma^{-1}(\mathbf{u} - \mathbf{u}_0) \to_{\mathcal{D}} \mathcal{N}\left(\frac{1}{cn} \cdot \left(\mathbf{u}_n^{\mathsf{T}} \otimes \mathbf{U}_0\right) \cdot \boldsymbol{\mu}, \frac{1}{n} \cdot \mathbf{U}_0^2\right);$$

(ii)

$$\sigma^{-2}\mu \to_{\mathcal{D}} \frac{1}{n(n-1)} \cdot \chi^2\left(\frac{(n-1)(n-2)}{2}; \delta\right)$$

where:

$$\delta \triangleq \frac{1}{c^2} \cdot \boldsymbol{\mu}^{\mathsf{T}} \cdot \left\{\frac{2\mathbf{K}_{nn}\left(\mathbf{U}_n \otimes \mathbf{I}_n\right)}{n^2} + \frac{\mathbf{I}_{n^2}}{n}\right\} \cdot \boldsymbol{\mu}.$$

(c) *If* $\frac{\sigma}{\|\phi_\ell\|_\infty} \to 0$*:*

(i)

$$\lim_{\|\phi_\ell\|_\infty \downarrow 0} \frac{1}{\|\phi_\ell\|_\infty} \cdot (\mathbf{u} - \mathbf{u}_0) = \frac{1}{n} \cdot \left(\mathbf{u}_n^{\mathsf{T}} \otimes \mathbf{U}_0\right) \cdot \boldsymbol{\mu},$$

$$\sigma^{-1}\left(\mathbf{u} - \mathbf{u}_0 - \frac{1}{n} \cdot \left(\mathbf{u}_n^{\mathsf{T}} \otimes \mathbf{U}_0\right) \cdot \boldsymbol{\mu}_\phi\right) \to_{\mathcal{D}} \mathcal{N}\left(0; \frac{1}{n} \cdot \mathbf{U}_0^2\right);$$

(ii)

$$\lim_{\|\phi_\ell\|_\infty \downarrow 0} \frac{1}{\|\phi_\ell\|_\infty^2} \cdot \mu = \frac{1}{n-1} \cdot \boldsymbol{\mu}^\top \cdot \left\{ \frac{2\mathbf{K}_{nn}(\mathbf{U}_n \otimes \mathbf{I}_n)}{n^2} + \frac{\mathbf{I}_{n^2}}{n} \right\} \cdot \boldsymbol{\mu},$$

$$\frac{1}{\sigma \cdot \|\phi_\ell\|_\infty} \left(\mu - \frac{1}{n-1} \cdot \boldsymbol{\mu}_\phi^\top \left\{ \frac{2\mathbf{K}_{nn}(\mathbf{U}_n \otimes \mathbf{I}_n)}{n^2} + \frac{\mathbf{I}_{n^2}}{n} \right\} \cdot \boldsymbol{\mu}_\phi \right) \to_{\mathcal{D}} \mathcal{N}\left(0; \frac{4}{(n-1)^2} \cdot \boldsymbol{\mu}^\top \left\{ \frac{2\mathbf{K}_{nn}(\mathbf{U}_n \otimes \mathbf{I}_n)}{n^3} + \frac{2(\mathbf{U}_n \otimes \mathbf{I}_n)\mathbf{K}_{nn}}{n^3} + \frac{(\mathbf{I}_{n^2} - \mathbf{K}_{nn})}{n^2} - \frac{4\mathbf{U}_{n^2}}{n^4} \right\} \boldsymbol{\mu} \right).$$

Theorem 4.1 distinguishes three cases. In the first case, in which $\frac{\sigma}{\|\phi_\ell\|_\infty} \to +\infty$, the leading term is given by the stochastic perturbation. This case encompasses Genest and Rivest (1994) in which $W(\cdot)$ is the identity and the asymptotic distributions of $\mathbf{u}$ and μ are respectively the normal and the χ^2 distribution (see Genest & Rivest, 1994, p. 490). In particular, suppose that σ_0^2 is such that the equality $\chi^2(\alpha) \approx n(n-1)\mu_0/10\sigma_0^2$ holds approximately, where $\chi^2(\alpha)$ stands for the $100(1-\alpha)\%$ quantile of the χ^2 with $p = (n-1)(n-2)/2$ degrees of freedom and μ_0 is the average value of Saaty's eigenvalue-based consistency index for matrices of size n whose entries are chosen at random within an admissible range. For this case Genest and Rivest (1994, p. 490) have shown that the so-called Saaty's 10% cut-off rule of declaring incoherent a response matrix $\mathbf{A}$ is in fact equivalent to a χ^2-test at significance level α of the hypothesis $H_0 : \sigma^2 \le \sigma_0^2$, that the variance of the background noise in the response data does not exceed a particularly chosen threshold level σ_0^2.

This equivalence, however, does no longer hold for the other two cases considered in Theorem 4.1, where the leading term in the approximations is no longer given by the stochastic perturbation.

The second case covers in particular the situation in which the deterministic and the stochastic perturbation are comparable and shows that the distribution of μ can be different from the χ^2 distribution obtained above. In particular this shows that the interpretation of Saaty's 10% cut-off rule as a test of the hypothesis $\mathrm{H}_0 : \sigma \le \sigma_0^2$ advanced in Genest and Rivest (1994) can lead to distortions when $\delta > 0$. Indeed, even if the interpretation is perfectly legitimate when the stochastic perturbation is the leading one, when a deterministic perturbation is present, this may lead to severely undersized tests.

The third and last case applies when the deterministic distortion induced by W is even larger than the stochastic one.[16] Here most of the variabilities in $\mathbf{u}$ and μ are due to systematic distortions and the asymptotic distribution of μ is so shifted to the right that it behaves as a normal distribution.

5. Ratio magnitude estimation

An analogue of Theorem 3.3, and in particular of Eq. (8) and its application (9) to the AHP, can be obtained when we work on a single row of the AHP response matrix compared with a reference x_0. This is a standard Stevens' ratio magnitude estimation experiment. We ask to compare the stimuli $(x_1, \ldots, x_n)$ with a baseline stimulus x_0 that, we suppose, does not belong to the set $(x_1, \ldots, x_n)$.[17] Each pairwise comparison yields the value $\alpha_{i0} = W^{-1}\left(\frac{\psi(x_i)}{\psi(x_0)}\right) \cdot e^{\nu_{i0}}$; here $\mathrm{d}\varepsilon_{i0} = \ln\left[\frac{\psi(x_0)}{\psi(x_i)} \cdot W^{-1}\left(\frac{\psi(x_i)}{\psi(x_0)}\right)\right] + \nu_{i0}$. Recall that one reason for the AHP to use the pairwise comparison matrix is exactly the fact that increasing the number of comparisons among the n items of a given set increases the amount of information and should generate better estimates.[18]

In this case the following theorem holds.

Theorem 5.1. *Let* $\boldsymbol{\nu} = \begin{bmatrix} \nu_{10} & \nu_{20} & \cdots & \nu_{n0} \end{bmatrix}^\top$. *Under the above-described assumptions:*

$$\mathrm{d}\mathbf{u} = \left(\mathrm{diag}(\mathbf{u}_0) - \mathbf{u}_0 \cdot \mathbf{u}_0^\top\right) \cdot \begin{bmatrix} \mathrm{d}\varepsilon_{10} & \mathrm{d}\varepsilon_{20} & \cdots & \mathrm{d}\varepsilon_{n0} \end{bmatrix}^\top = \left(\mathrm{diag}(\mathbf{u}_0) - \mathbf{u}_0 \cdot \mathbf{u}_0^\top\right) \cdot \left\{ \sum_{\ell=2}^{L} \phi_\ell \cdot \left[\overline{\ln}\mathbf{u}_0 - \ln u_{0,0}\right]^{\odot \ell} + \boldsymbol{\nu} + o\left(\|\phi_\ell\|_\infty\right) + o_{\mathbb{P}}(\sigma) \right\}.$$

In order to compare the vector $\mathbf{u}$ estimated through Saaty's eigenvector method or directly obtained through ratio magnitude estimation, we compute also for the present method the bias $\frac{\|\mathbb{E}\mathrm{d}\mathbf{u}\|}{|\phi_3|}$ and the standard deviation $\frac{\sqrt{\mathbb{E}\|\mathrm{d}\mathbf{u} - \mathbb{E}\mathrm{d}\mathbf{u}\|^2}}{\sigma}$. We have:

$$\frac{\|\mathbb{E}\mathrm{d}\mathbf{u}\|}{|\phi_3|} = \left\| \left(\mathrm{diag}(\mathbf{u}_0) - \mathbf{u}_0 \cdot \mathbf{u}_0^\top\right) \cdot \left[\overline{\ln}\mathbf{u}_0 - \ln u_{0,0}\right]^{\odot 3} \right\|.$$

We notice that in this case the term depends on n, on $\mathbf{u}_0$ but also on the reference point $u_{0,0} = \psi(x_0)$. On the other hand the standard deviation term is easily computable:

$$\frac{\sqrt{\mathbb{E}\,\|\mathrm{d}\mathbf{u} - \mathbb{E}\mathrm{d}\mathbf{u}\|^2}}{\sigma} = \sqrt{\sum_{i=1}^{n} u_{0,i}^2 + \left(\sum_{i=1}^{n} u_{0,i}^2\right)^2 - 2\sum_{i=1}^{n} u_{0,i}^3}.$$

This shows that the AHP reduces the standard deviation of the distortion by a factor $\frac{1}{\sqrt{n}}$. As concerns the bias, on the other hand, the situation is less clear and it is uncertain which method is better. This is in a sense surprising since it would seem that more information would allow for better estimates of $\mathbf{u}$; however, the reason is that when W^{-1} is far from the identity, the principal eigenvector method misuses the additional information to an extent which could produce even more bias than that occurring due to the lower amount of information provided by ratio magnitude estimation.

The graphs in Fig. 2 exemplify the problem. They show the quantities $\frac{\|\mathbb{E}\mathrm{d}\mathbf{u}\|}{|\phi_3|}$ and $\frac{\sqrt{\mathbb{E}\|\mathrm{d}\mathbf{u} - \mathbb{E}\mathrm{d}\mathbf{u}\|^2}}{\sigma}$ for several values of $u_{0,0}$; as in the previous Figure, the names "Constant", "Logarithmic", "Square root", "Linear", "Square" and "Exponential" denote in this order the vectors with $u_{0,j} \propto 1$, $u_{0,j} \propto \ln(1+j)$, $u_{0,j} \propto j^{1/2}$, $u_{0,j} \propto j$, $u_{0,j} \propto j^2$ and $u_{0,j} \propto \exp(j-1)$, while the index n is to be read on the horizontal axis. Here too, n starts from 2. The thin grey line shows $\frac{\sqrt{\mathbb{E}\|\mathrm{d}\mathbf{u} - \mathbb{E}\mathrm{d}\mathbf{u}\|^2}}{\sigma}$ for Saaty's method (the same as in

[16] By the way, it is interesting to notice that the empirical analyses in Table 1 and in Bernasconi et al. (2010) show that the cases in which the stochastic terms are the leading ones in the perturbations are indeed the least likely in practice.

[17] The situation that the reference stimulus belongs to the set of stimuli $(x_1, \ldots, x_n)$ and corresponds, say, to x_1 is simply obtained by setting $\mathrm{d}\varepsilon_{10} = 0$ in the statement of Theorem 5.1.

[18] Obviously, notice that this is different from increasing the number n of the items being compared in a response matrix, which could generate just the opposite effect of increasing the level of inconsistency in the data. The latter observation was anticipated by Saaty (1977) and has been commented extensively by Genest and Rivest (1994, pp. 494–495) on the basis of results equivalent to those referred in the first case of Theorem 4.1 above.

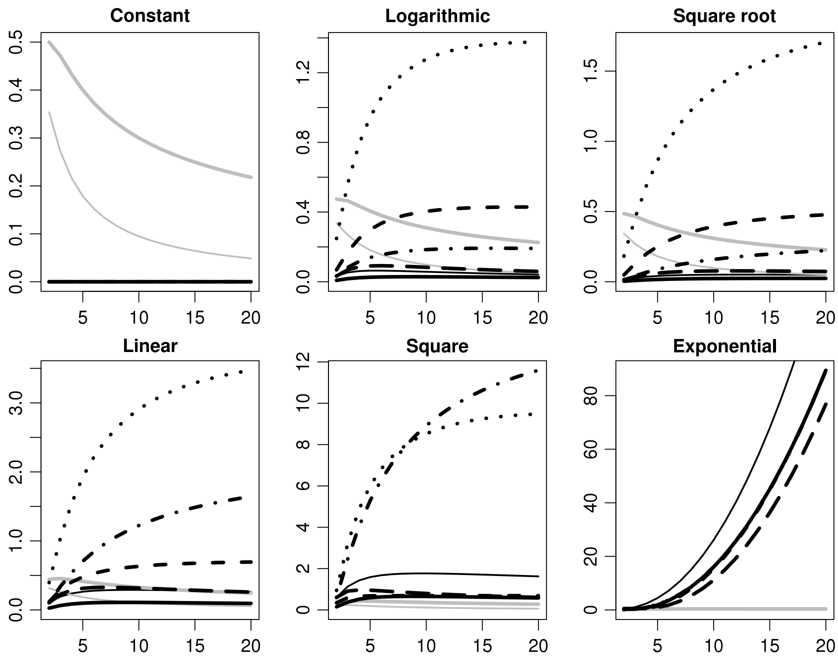

Fig. 2. Comparisons of $\frac{\|\mathbb{E}d\mathbf{u}\|}{|\phi_3|}$ and $\frac{\sqrt{\mathbb{E}\|d\mathbf{u}-\mathbb{E}d\mathbf{u}\|^2}}{\sigma}$ in AHP versus ratio magnitude estimation for different n and $\mathbf{u}_0$.

the previous figure) while the thick grey line shows $\frac{\sqrt{\mathbb{E}\|d\mathbf{u}-\mathbb{E}d\mathbf{u}\|^2}}{\sigma}$ for a ratio magnitude experiment (independent of $u_{0,0}$). The thin black line shows $\frac{\|\mathbb{E}d\mathbf{u}\|}{|\phi_3|}$ for Saaty's method, while $\frac{\|\mathbb{E}d\mathbf{u}\|}{|\phi_3|}$ for a ratio magnitude experiment is displayed in the thick black lines: the dashed line has $u_{0,0} = \left(\prod_{i=1}^n u_{0,i}\right)^{1/2n}$, the solid line has $u_{0,0} = \left(\prod_{i=1}^n u_{0,i}\right)^{1/n}$, the dotted line has $u_{0,0} = \left(\prod_{i=1}^n u_{0,i}\right)^{2/n}$, the dash-dotted line has $u_{0,0} = \min_i u_{0,i}$, the long-dashed line has $u_{0,0} = \max_i u_{0,i}$. It appears that $\frac{\sqrt{\mathbb{E}\|d\mathbf{u}-\mathbb{E}d\mathbf{u}\|^2}}{\sigma}$ is always smaller for Saaty's method, and the ratio increases with n. As concerns $\frac{\|\mathbb{E}d\mathbf{u}\|}{|\phi_3|}$, the situation is more complex. As expected, values of $u_{0,0}$ far away from $\left(\prod_{i=1}^n u_{0,i}\right)^{1/n}$ (or any other measure of centrality of $\mathbf{u}_0$) give larger values. The ratio magnitude estimation experiment using $u_{0,0} = \max_i u_{0,i}$ yields better results than the one with $u_{0,0} = \min_i u_{0,i}$. Moreover Saaty's method yields smaller values of $\frac{\|\mathbb{E}d\mathbf{u}\|}{|\phi_3|}$ for $\mathbf{u}_0$ respecting the homogeneity requirement, while when the values in $\mathbf{u}_0$ are very different Saaty's method is worse than simple ratio magnitude estimation. This has two consequences: first of all, it stresses the relevance of homogeneity. Second, it subverts a widely believed idea: according to common sense and the identification of random noise as the only source of distortion in the AHP, it is usually thought that Saaty's method is better than ratio magnitude estimation since it is based on a larger number of evaluations; this is false whenever homogeneity is not respected and psychophysical distortions are present in the data, while it is always true when all the distortion is due to random noise.

A further interesting issue on the relationships between the AHP and ratio magnitude estimation concerns the different burdens imposed on the respondents due to the different number of comparisons even when starting from the same number of items n: namely $n(n-1)/2$ in the AHP versus n in ratio magnitude estimation. Thus a natural question is whether by repeating K times each comparison in a ratio magnitude experiment the errors can be reduced. We suppose that the comparison between the reference and stimulus i for the kth repetition is:

$$\alpha_{k,i0} = W^{-1}\left(\frac{\psi(x_i)}{\psi(x_0)}\right) \cdot e^{\nu_{k,i0}}, \tag{10}$$

where the random errors $\nu_{k,i0}$ are independent across repetitions. Various methods of aggregations may be considered. In particular, the geometric average with weights $1/K$ yields:

$$\alpha_{i0} = \prod_{k=1}^{K} \alpha_{k,i0}^{1/K} = W^{-1}\left(\frac{\psi(x_i)}{\psi(x_0)}\right) \cdot e^{\frac{1}{K}\sum_{k=1}^{K}\nu_{k,i0}}. \tag{11}$$

We remark that the deterministic term $\|\mathbb{E}d\mathbf{u}\|$ is the same as in the model without repetitions (namely, when $K = 1$), while the standard deviation is reduced by a factor $\frac{1}{\sqrt{K}}$. This has obviously two consequences. The first is that the previous analysis and its implications focused on the systematic part of the error extend without any change even in this case. The second is that, in order for the repeated ratio magnitude estimations task to get the same standard deviation as the AHP, one would need at least n^2 comparisons, that is more than twice those of the AHP. Finally, it is important to remark that the same two implications apply also for other aggregation methods.[19]

6. Conclusions

Recent developments in mathematical psychology, supported by various experimental tests, have shown that ratio-scaling methods in which individuals use number names to express proportions in which they perceive pairs of stimuli cannot be treated as scientific ratios.

The AHP is a ratio-scaling procedure widely used in management decisions. It is a more articulated method than the classical ratio magnitude estimation used in psychophysics. In the AHP a symmetric positive matrix of subjective ratio assessments is constructed and the maximum eigenvalue method is used to extract the single maximum eigenvector, the Perron eigenvector, which is then treated as the ratio scale of interest.

19 For example, if instead of the geometric we use the arithmetic mean, it is possible to show that the first order results are the same, while minor differences affect the higher order error terms.

In this paper we have used recent developments in mathematical psychology based on the so-called separable forms, to study the types of distortions which can arise in the AHP when the maximum eigenvalue method is used. The analysis has emphasized the difference between the distortions due to random noise from the systematic or cognitive distortions embodied in the separable representations. The cognitive distortions highlight the importance of the so-called homogeneity axiom of the AHP to keep under control the bias arising in the estimate of the ratio scale.

We have also studied the asymptotic distributions of the maximum eigenvalue and maximum eigenvector under separable representations and have shown the limits of using the eigenvalue-based index of cardinal consistency of classical AHP as a rule to assess the quality of the estimate of the ratio scale.

The analysis has also shown that in some cases, when the cognitive distortions in the data are larger than those due to random noise and homogeneity is not fully respected, the classical ratio magnitude estimation used in psychophysics can be a better ratio-scaling procedure than the AHP, despite the greater amount of comparisons used by the latter method.

Acknowledgment

The first and third authors works were supported by a grant from MIUR (Ministero dell'Istruzione, dell'Università e della Ricerca). We are particularly grateful to two anonymous referees and the editor for several helpful comments on an earlier version of the paper.

Appendix. Proofs

Before passing to the mathematical details, we remark that the proofs of the previous theorems yield more than declared in the statements. Indeed, they prove the existence of a real function λ and of a vector function $\mathbf{u}$ defined for all $\mathbf{A}$ in some neighborhood $N(\mathbf{A}_0) \subset \mathbb{R}^{n\times n}$ of $\mathbf{A}_0$, such that

$$\lambda(\mathbf{A}_0) = \lambda_0$$
$$\mathbf{u}(\mathbf{A}_0) = \mathbf{u}_0$$
$$\mathbf{A}\cdot\mathbf{u} = \lambda\cdot\mathbf{u}$$
$$\mathbf{u}^{\mathrm{T}}\mathbf{u}_n = 1$$

for $\mathbf{A} \in N(\mathbf{A}_0)$. Moreover, the functions λ and $\mathbf{u}$ are ∞ times differentiable on $N(\mathbf{A}_0)$, and the differentials at $\mathbf{A}_0$ have the form provided in the statements.

Proof of Theorem 3.1. First of all we prove the analyticity of λ and $\mathbf{u}$. Our proof follows the scheme of Theorem 7 on p. 158 in Magnus and Neudecker (1999, in the following MN) but is more complicated because of the nonstandard normalization of the eigenvector and of the nonsymmetry of $(\lambda_0\mathbf{I}_n - \mathbf{A}_0)$. Consider the vector function $f : \mathbb{R}^{n+1}\times\mathbb{R}^{n\times n} \to \mathbb{R}^{n+1}$ defined by:

$$f(\mathbf{u},\lambda;\mathbf{A}) = \begin{bmatrix} (\lambda\mathbf{I}_n - \mathbf{A})\cdot\mathbf{u} \\ \mathbf{u}_n^{\mathrm{T}}\cdot\mathbf{u} - 1 \end{bmatrix} \tag{12}$$

(remark the difference in the second line with respect to the proof in MN). f is ∞ times differentiable on $\mathbb{R}^{n+1}\times\mathbb{R}^{n\times n}$ and $f(\mathbf{u}_0,\lambda_0;\mathbf{A}_0) = \mathbf{0}$. The matrix $(\lambda_0\mathbf{I}_n - \mathbf{A}_0)$ has reduced rank $n-1$ since λ_0 is a simple eigenvalue and we can apply Theorem 4 on p. 43 and Theorem 3 on p. 41 in MN:

$$\begin{vmatrix} \lambda_0\mathbf{I}_n - \mathbf{A}_0 & \mathbf{u}_0 \\ \mathbf{u}_n^{\mathrm{T}} & 0 \end{vmatrix} = -\mathbf{u}_n^{\mathrm{T}}(\lambda_0\mathbf{I}_n - \mathbf{A}_0)^{\sharp}\mathbf{u}_0 \tag{13}$$

$$= -\mathbf{u}_n^{\mathrm{T}}\left[\mu(\lambda_0\mathbf{I}_n - \mathbf{A}_0)\frac{\mathbf{u}_0\mathbf{v}_0^{\mathrm{T}}}{\mathbf{v}_0^{\mathrm{T}}\mathbf{u}_0}\right]\mathbf{u}_0 \tag{14}$$

$$= -\mu(\lambda_0\mathbf{I}_n - \mathbf{A}_0) \neq 0 \tag{15}$$

where $\mathbf{B}^{\sharp}$ is the adjoint matrix of $\mathbf{B}$ defined on p. 40 of MN, and $\mu(\mathbf{B})$ is the product of the nonzero eigenvalues of $\mathbf{B}$. This implies that the conditions of the Implicit Function Theorem (Theorem A.3 in the Appendix of Chapter 7 in MN) are satisfied and there exist a neighborhood $N(\mathbf{A}_0) \subset \mathbb{R}^{n\times n}$ of $\mathbf{A}_0$, a unique real-valued function $\lambda : N(\mathbf{A}_0) \to \mathbb{R}$ and a unique (up to the sign) vector function $\mathbf{u} : N(\mathbf{A}_0) \to \mathbb{R}^n$ such that:

1. λ and $\mathbf{u}$ are ∞ times differentiable on $N(\mathbf{A}_0)$;
2. $\lambda(\mathbf{A}_0) = \lambda_0$ and $\mathbf{u}(\mathbf{A}_0) = \mathbf{u}_0$;
3. $\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$, $\mathbf{u}_n^{\mathrm{T}}\mathbf{u} = 1$ for every $\mathbf{A} \in N(\mathbf{A}_0)$.

This completes the existence proof.

Differentiating $\mathbf{A}\cdot\mathbf{u} = \lambda\cdot\mathbf{u}$ around the point $(\mathbf{A},\lambda,\mathbf{u}) = (\mathbf{A}_0,\lambda_0,\mathbf{u}_0)$ we get

$$\mathrm{d}\mathbf{A}\cdot\mathbf{u}_0 + \mathbf{A}_0\cdot\mathrm{d}\mathbf{u} = \mathrm{d}\lambda\cdot\mathbf{u}_0 + \lambda_0\cdot\mathrm{d}\mathbf{u} \tag{16}$$

and premultiplying this by $\mathbf{v}_0^{\mathrm{T}}$ we get:

$$\mathbf{v}_0^{\mathrm{T}}\cdot\mathrm{d}\mathbf{A}\cdot\mathbf{u}_0 + \mathbf{v}_0^{\mathrm{T}}\cdot\mathbf{A}_0\cdot\mathrm{d}\mathbf{u} = \mathrm{d}\lambda\cdot\mathbf{v}_0^{\mathrm{T}}\mathbf{u}_0 + \lambda_0\cdot\mathbf{v}_0^{\mathrm{T}}\cdot\mathrm{d}\mathbf{u}$$
$$\mathbf{v}_0^{\mathrm{T}}\cdot\mathrm{d}\mathbf{A}\cdot\mathbf{u}_0 = \mathrm{d}\lambda\cdot\mathbf{v}_0^{\mathrm{T}}\mathbf{u}_0$$
$$\mathrm{d}\lambda = \frac{\mathbf{v}_0^{\mathrm{T}}\cdot\mathrm{d}\mathbf{A}\cdot\mathbf{u}_0}{\mathbf{v}_0^{\mathrm{T}}\mathbf{u}_0}. \tag{17}$$

Now, we take the second differential of $\mathbf{A}\cdot\mathbf{u} = \lambda\cdot\mathbf{u}$ around the point $(\mathbf{A},\lambda,\mathbf{u}) = (\mathbf{A}_0,\lambda_0,\mathbf{u}_0)$:

$$\mathrm{d}^2\mathbf{A}\cdot\mathbf{u}_0 + 2\mathrm{d}\mathbf{A}\cdot\mathrm{d}\mathbf{u} + \mathbf{A}_0\cdot\mathrm{d}^2\mathbf{u} = 2\mathrm{d}\lambda\cdot\mathrm{d}\mathbf{u} + \mathrm{d}^2\lambda\cdot\mathbf{u}_0 + \lambda_0\cdot\mathrm{d}^2\mathbf{u} \tag{18}$$

and premultiplying it by $\mathbf{v}_0^{\mathrm{T}}$ we get:

$$\mathbf{v}_0^{\mathrm{T}}\cdot\mathrm{d}^2\mathbf{A}\cdot\mathbf{u}_0 + 2\mathbf{v}_0^{\mathrm{T}}\cdot\mathrm{d}\mathbf{A}\cdot\mathrm{d}\mathbf{u} + \mathbf{v}_0^{\mathrm{T}}\cdot\mathbf{A}_0\cdot\mathrm{d}^2\mathbf{u} = 2\mathrm{d}\lambda\cdot\mathbf{v}_0^{\mathrm{T}}\cdot\mathrm{d}\mathbf{u} + \mathrm{d}^2\lambda\cdot\mathbf{v}_0^{\mathrm{T}}\cdot\mathbf{u}_0 + \lambda_0\cdot\mathbf{v}_0^{\mathrm{T}}\cdot\mathrm{d}^2\mathbf{u}$$
$$\mathbf{v}_0^{\mathrm{T}}\cdot\mathrm{d}^2\mathbf{A}\cdot\mathbf{u}_0 + 2\mathbf{v}_0^{\mathrm{T}}\cdot\mathrm{d}\mathbf{A}\cdot\mathrm{d}\mathbf{u} = 2\mathrm{d}\lambda\cdot\mathbf{v}_0^{\mathrm{T}}\cdot\mathrm{d}\mathbf{u} + \mathrm{d}^2\lambda\cdot\mathbf{v}_0^{\mathrm{T}}\cdot\mathbf{u}_0$$

and:

$$\mathrm{d}^2\lambda = \frac{\mathbf{v}_0^{\mathrm{T}}\mathrm{d}^2\mathbf{A}\mathbf{u}_0 + 2\mathbf{v}_0^{\mathrm{T}}\mathrm{d}\mathbf{A}\mathrm{d}\mathbf{u} - 2\mathrm{d}\lambda\mathbf{v}_0^{\mathrm{T}}\mathrm{d}\mathbf{u}}{\mathbf{v}_0^{\mathrm{T}}\cdot\mathbf{u}_0} = \frac{\mathbf{v}_0^{\mathrm{T}}\mathbf{u}_0\cdot\mathbf{v}_0^{\mathrm{T}}\mathrm{d}^2\mathbf{A}\mathbf{u}_0 + 2\mathbf{v}_0^{\mathrm{T}}\mathbf{u}_0\cdot\mathbf{v}_0^{\mathrm{T}}\mathrm{d}\mathbf{A}\mathrm{d}\mathbf{u} - 2\mathbf{v}_0^{\mathrm{T}}\mathrm{d}\mathbf{A}\mathbf{u}_0\cdot\mathbf{v}_0^{\mathrm{T}}\mathrm{d}\mathbf{u}}{\left(\mathbf{v}_0^{\mathrm{T}}\cdot\mathbf{u}_0\right)^2} \tag{19}$$

where $\mathrm{d}\mathbf{u}$ will be obtained in the following.

We start from $\mathbf{A}_0\cdot\mathbf{u}_0 = \lambda_0\cdot\mathbf{u}_0$ and we define by $\mathbf{u}_0$ the vector normalized as $\mathbf{u}_0^{\mathrm{T}}\mathbf{u}_n = 1$ and by $\widetilde{\mathbf{u}}_0$ the vector normalized as $\widetilde{\mathbf{u}}_0^{\mathrm{T}}\widetilde{\mathbf{u}}_0 = 1$. We have $\mathbf{u}_0 = \frac{\widetilde{\mathbf{u}}_0}{\widetilde{\mathbf{u}}_0^{\mathrm{T}}\mathbf{u}_n}$. In the same way, we consider $\mathbf{A}\cdot\mathbf{u} = \lambda\cdot\mathbf{u}$ and we define by $\mathbf{u}$ the vector normalized by $\mathbf{u}^{\mathrm{T}}\mathbf{u}_n = 1$ and by $\widetilde{\mathbf{u}}$ the vector normalized by $\widetilde{\mathbf{u}}^{\mathrm{T}}\widetilde{\mathbf{u}}_0 = 1$. We have $\mathbf{u} = \frac{\widetilde{\mathbf{u}}}{\widetilde{\mathbf{u}}^{\mathrm{T}}\mathbf{u}_n}$. On the other hand $\widetilde{\mathbf{u}} \simeq \widetilde{\mathbf{u}}_0 + \mathrm{d}\widetilde{\mathbf{u}} + \frac{1}{2}\mathrm{d}^2\widetilde{\mathbf{u}}$ and:

$$\mathbf{u} = \frac{\widetilde{\mathbf{u}}}{\widetilde{\mathbf{u}}^{\mathrm{T}}\mathbf{u}_n} \simeq \frac{\widetilde{\mathbf{u}}_0}{\widetilde{\mathbf{u}}_0^{\mathrm{T}}\mathbf{u}_n} + \left(\mathbf{I}_n - \frac{\widetilde{\mathbf{u}}_0\mathbf{u}_n^{\mathrm{T}}}{\widetilde{\mathbf{u}}_0^{\mathrm{T}}\mathbf{u}_n}\right)\cdot\frac{\mathrm{d}\widetilde{\mathbf{u}}}{\widetilde{\mathbf{u}}_0^{\mathrm{T}}\mathbf{u}_n} + \frac{1}{2}\cdot\left(\mathbf{I}_n - \frac{\widetilde{\mathbf{u}}_0\mathbf{u}_n^{\mathrm{T}}}{\widetilde{\mathbf{u}}_0^{\mathrm{T}}\mathbf{u}_n}\right)\cdot\frac{\mathrm{d}^2\widetilde{\mathbf{u}}}{\widetilde{\mathbf{u}}_0^{\mathrm{T}}\mathbf{u}_n} - \left(\mathbf{I}_n - \frac{\widetilde{\mathbf{u}}_0\mathbf{u}_n^{\mathrm{T}}}{\widetilde{\mathbf{u}}_0^{\mathrm{T}}\mathbf{u}_n}\right)\cdot\frac{\mathbf{u}_n^{\mathrm{T}}\mathrm{d}\widetilde{\mathbf{u}}\cdot\mathrm{d}\widetilde{\mathbf{u}}}{\left(\widetilde{\mathbf{u}}_0^{\mathrm{T}}\mathbf{u}_n\right)^2}.$$

Since $\mathbf{u} = \frac{\widetilde{\mathbf{u}}_0}{\widetilde{\mathbf{u}}_0^{\mathrm{T}}\mathbf{u}_n}$, this means that the equalities:

$$\mathrm{d}\mathbf{u} = \left(\mathbf{I}_n - \frac{\widetilde{\mathbf{u}}_0\mathbf{u}_n^{\mathrm{T}}}{\widetilde{\mathbf{u}}_0^{\mathrm{T}}\mathbf{u}_n}\right)\cdot\frac{\mathrm{d}\widetilde{\mathbf{u}}}{\widetilde{\mathbf{u}}_0^{\mathrm{T}}\mathbf{u}_n} \tag{20}$$

$$\mathrm{d}^2\mathbf{u} = \left(\mathbf{I}_n - \frac{\widetilde{\mathbf{u}}_0\mathbf{u}_n^{\mathrm{T}}}{\widetilde{\mathbf{u}}_0^{\mathrm{T}}\mathbf{u}_n}\right)\cdot\frac{\mathrm{d}^2\widetilde{\mathbf{u}}}{\widetilde{\mathbf{u}}_0^{\mathrm{T}}\mathbf{u}_n} - 2\left(\mathbf{I}_n - \frac{\widetilde{\mathbf{u}}_0\mathbf{u}_n^{\mathrm{T}}}{\widetilde{\mathbf{u}}_0^{\mathrm{T}}\mathbf{u}_n}\right)\cdot\frac{\mathbf{u}_n^{\mathrm{T}}\mathrm{d}\widetilde{\mathbf{u}}\cdot\mathrm{d}\widetilde{\mathbf{u}}}{\left(\widetilde{\mathbf{u}}_0^{\mathrm{T}}\mathbf{u}_n\right)^2} \tag{21}$$

hold. Theorem 8 on p. 161 in MN states that:

$$\mathrm{d}\widetilde{\mathbf{u}} = (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ \left(\mathbf{I}_n - \frac{\widetilde{\mathbf{u}}_0 \widetilde{\mathbf{v}}_0^\mathsf{T}}{\widetilde{\mathbf{v}}_0^\mathsf{T} \widetilde{\mathbf{u}}_0}\right) \cdot \mathrm{d}\mathbf{A} \cdot \widetilde{\mathbf{u}}_0. \tag{22}$$

Using the proportionality between $\widetilde{\mathbf{u}}_0$ and $\mathbf{u}_0$ on the one hand, and $\widetilde{\mathbf{v}}_0$ and $\mathbf{v}_0$ on the other hand, (20) becomes:

$$\begin{aligned} \mathrm{d}\mathbf{u} &= \frac{1}{\widetilde{\mathbf{u}}_0^\mathsf{T} \mathbf{u}_n} \cdot \left(\mathbf{I}_n - \mathbf{u}_0 \mathbf{u}_n^\mathsf{T}\right) \cdot (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ \left(\mathbf{I}_n - \frac{\widetilde{\mathbf{u}}_0 \widetilde{\mathbf{v}}_0^\mathsf{T}}{\widetilde{\mathbf{v}}_0^\mathsf{T} \widetilde{\mathbf{u}}_0}\right) \cdot \mathrm{d}\mathbf{A} \cdot \widetilde{\mathbf{u}}_0 \\ &= \left(\mathbf{I}_n - \mathbf{u}_0 \mathbf{u}_n^\mathsf{T}\right) \cdot (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ \left(\mathbf{I}_n - \frac{\mathbf{u}_0 \mathbf{v}_0^\mathsf{T}}{\mathbf{v}_0^\mathsf{T} \mathbf{u}_0}\right) \cdot \mathrm{d}\mathbf{A} \cdot \mathbf{u}_0. \end{aligned}$$

From this, we get $\mathrm{d}^2\lambda$ given in Box II. Expanding the product $\mathbf{v}_0^\mathsf{T} \left(\mathbf{u}_0 \mathbf{v}_0^\mathsf{T} \mathrm{d}\mathbf{A} - \mathrm{d}\mathbf{A}\mathbf{u}_0 \cdot \mathbf{v}_0^\mathsf{T}\right) \left(\mathbf{I}_n - \mathbf{u}_0 \mathbf{u}_n^\mathsf{T}\right)$ and using the fact that $\mathbf{v}_0^\mathsf{T} \mathbf{u}_0$ is a scalar, it can be seen that

$$\begin{aligned} &\mathbf{v}_0^\mathsf{T} \left(\mathbf{u}_0 \mathbf{v}_0^\mathsf{T} \mathrm{d}\mathbf{A} - \mathrm{d}\mathbf{A}\mathbf{u}_0 \cdot \mathbf{v}_0^\mathsf{T}\right) \left(\mathbf{I}_n - \mathbf{u}_0 \mathbf{u}_n^\mathsf{T}\right) \\ &\quad = \mathbf{v}_0^\mathsf{T} \mathbf{u}_0 \cdot \mathbf{v}_0^\mathsf{T} \mathrm{d}\mathbf{A} - \mathbf{v}_0^\mathsf{T} \mathrm{d}\mathbf{A}\mathbf{u}_0 \cdot \mathbf{v}_0^\mathsf{T} \\ &\quad = \mathbf{v}_0^\mathsf{T} \mathbf{u}_0 \cdot \mathbf{v}_0^\mathsf{T} \mathrm{d}\mathbf{A} \cdot \left(\mathbf{I}_n - \frac{\mathbf{u}_0 \mathbf{v}_0^\mathsf{T}}{\mathbf{v}_0^\mathsf{T} \mathbf{u}_0}\right). \end{aligned}$$

As concerns the second differential of the eigenvector, we differentiate twice $\mathbf{A} \cdot \widetilde{\mathbf{u}} = \lambda \cdot \widetilde{\mathbf{u}}$ around the point $(\mathbf{A}, \lambda, \widetilde{\mathbf{u}}) = (\mathbf{A}_0, \lambda_0, \widetilde{\mathbf{u}}_0)$ and we get:

$$\mathrm{d}^2\mathbf{A} \cdot \widetilde{\mathbf{u}}_0 + 2\mathrm{d}\mathbf{A} \cdot \mathrm{d}\widetilde{\mathbf{u}} + \mathbf{A}_0 \cdot \mathrm{d}^2\widetilde{\mathbf{u}} = 2\mathrm{d}\lambda \cdot \mathrm{d}\widetilde{\mathbf{u}} + \mathrm{d}^2\lambda \cdot \widetilde{\mathbf{u}}_0 + \lambda_0 \cdot \mathrm{d}^2\widetilde{\mathbf{u}}$$

from which:

$$(\lambda_0 \mathbf{I}_n - \mathbf{A}_0) \cdot \mathrm{d}^2\widetilde{\mathbf{u}} = \left(\mathrm{d}^2\mathbf{A} - \mathrm{d}^2\lambda \mathbf{I}_n\right) \cdot \widetilde{\mathbf{u}}_0 + 2\left(\mathrm{d}\mathbf{A} - \mathrm{d}\lambda \mathbf{I}_n\right) \cdot \mathrm{d}\widetilde{\mathbf{u}}.$$

We apply the arguments on p. 160 in the proof of Theorem 7 on p. 158 of MN. Indeed, consider the matrix $\mathbf{C}_0 = (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ (\lambda_0 \mathbf{I}_n - \mathbf{A}_0) + \widetilde{\mathbf{u}}_0 \widetilde{\mathbf{u}}_0^\mathsf{T}$. By the properties of the Moore–Penrose inverse and using the fact that $\widetilde{\mathbf{u}}_0$ is an eigenvector of $\mathbf{A}_0$ with eigenvalue λ_0, both $(\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)$ and $\mathbf{C}_0$ are symmetric and idempotent and:

$$\begin{aligned} \mathrm{rank}(\mathbf{C}_0) &= \mathrm{tr}\,(\mathbf{C}_0) \\ &= \mathrm{tr}((\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)) + 1 \\ &= \mathrm{rank}(\lambda_0 \mathbf{I}_n - \mathbf{A}_0) + 1 = n. \end{aligned} \tag{23}$$

Therefore $\mathbf{C}_0 = \mathbf{I}_n$ and, premultiplying the equality with $(\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+$, we get:

$$\begin{aligned} \left(\mathbf{I}_n - \widetilde{\mathbf{u}}_0 \widetilde{\mathbf{u}}_0^\mathsf{T}\right) \cdot \mathrm{d}^2\widetilde{\mathbf{u}} &= (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ \left(\mathrm{d}^2\mathbf{A} - \mathrm{d}^2\lambda \mathbf{I}_n\right) \\ &\quad \cdot \widetilde{\mathbf{u}}_0 + 2(\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ \left(\mathrm{d}\mathbf{A} - \mathrm{d}\lambda \mathbf{I}_n\right) \cdot \mathrm{d}\widetilde{\mathbf{u}}. \end{aligned}$$

However, from $\widetilde{\mathbf{u}}^\mathsf{T} \widetilde{\mathbf{u}}_0 = 1$ it is clear that $\widetilde{\mathbf{u}}_0^\mathsf{T} \cdot \mathrm{d}^2\widetilde{\mathbf{u}} = 0$:

$$\begin{aligned} \mathrm{d}^2\widetilde{\mathbf{u}} &= (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ \left(\mathrm{d}^2\mathbf{A} - \mathrm{d}^2\lambda \mathbf{I}_n\right) \cdot \widetilde{\mathbf{u}}_0 + 2\left(\lambda_0 \mathbf{I}_n - \mathbf{A}_0\right)^+ \\ &\quad \times \left(\mathrm{d}\mathbf{A} - \mathrm{d}\lambda \mathbf{I}_n\right) \cdot \mathrm{d}\widetilde{\mathbf{u}}. \end{aligned}$$

Now, we replace this term in (21):

$$\begin{aligned} \mathrm{d}^2\mathbf{u} &= \left(\mathbf{I}_n - \frac{\mathbf{u}_0 \mathbf{u}_n^\mathsf{T}}{\mathbf{u}_0^\mathsf{T} \mathbf{u}_n}\right) \cdot (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ \left(\mathrm{d}^2\mathbf{A} - \mathrm{d}^2\lambda \mathbf{I}_n\right) \mathbf{u}_0 \\ &\quad + 2\left(\mathbf{I}_n - \frac{\mathbf{u}_0 \mathbf{u}_n^\mathsf{T}}{\mathbf{u}_0^\mathsf{T} \mathbf{u}_n}\right) (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ \left(\mathrm{d}\mathbf{A} - \mathrm{d}\lambda \mathbf{I}_n\right) (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ \\ &\quad \times \left(\mathbf{I}_n - \frac{\mathbf{u}_0 \mathbf{v}_0^\mathsf{T}}{\mathbf{v}_0^\mathsf{T} \mathbf{u}_0}\right) \mathrm{d}\mathbf{A}\mathbf{u}_0 - 2\left(\mathbf{I}_n - \frac{\mathbf{u}_0 \mathbf{u}_n^\mathsf{T}}{\mathbf{u}_0^\mathsf{T} \mathbf{u}_n}\right) (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ \\ &\quad \times \left(\mathbf{I}_n - \frac{\mathbf{u}_0 \mathbf{v}_0^\mathsf{T}}{\mathbf{v}_0^\mathsf{T} \mathbf{u}_0}\right) \mathrm{d}\mathbf{A}\mathbf{u}_0 \cdot \mathbf{u}_n^\mathsf{T} (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ \\ &\quad \times \left(\mathbf{I}_n - \frac{\mathbf{u}_0 \mathbf{v}_0^\mathsf{T}}{\mathbf{v}_0^\mathsf{T} \mathbf{u}_0}\right) \mathrm{d}\mathbf{A}\mathbf{u}_0. \quad \square \end{aligned}$$

Proof of Theorem 3.2. We replace in the previous formulas $\mathbf{v}_0$ with $\bar{\mathbf{u}}_0$ and we use the fact that $\left(n\mathbf{I}_n - \mathbf{u}_0 \bar{\mathbf{u}}_0^\mathsf{T}\right)^+ = \frac{1}{n}\mathbf{I}_n - \frac{\mathbf{u}_0 \mathbf{u}_0^\mathsf{T}}{n \cdot \mathbf{u}_0^\mathsf{T} \mathbf{u}_0} - \frac{\bar{\mathbf{u}}_0 \bar{\mathbf{u}}_0^\mathsf{T}}{n \cdot \bar{\mathbf{u}}_0^\mathsf{T} \bar{\mathbf{u}}_0} + \frac{\mathbf{u}_0 \bar{\mathbf{u}}_0^\mathsf{T}}{\mathbf{u}_0^\mathsf{T} \mathbf{u}_0 \cdot \bar{\mathbf{u}}_0^\mathsf{T} \bar{\mathbf{u}}_0}$ (Meyer, 1973). □

Proof of Theorem 3.3. First of all, applying an elementwise expansion to $\mathbf{A} = \mathbf{A}_0 \odot \overline{\exp}\,(\mathrm{d}\mathbf{E})$ remark that $\mathbf{A} \simeq \mathbf{A}_0 + \mathrm{d}\mathbf{A} + \frac{1}{2}\mathrm{d}^2\mathbf{A}$ with $\mathrm{d}\mathbf{A} = \mathbf{A}_0 \odot \mathrm{d}\mathbf{E}$ and $\mathrm{d}^2\mathbf{A} = \mathbf{A}_0 \odot \mathrm{d}\mathbf{E} \odot \mathrm{d}\mathbf{E}$. The differentials $\mathrm{d}\mathbf{A}$ and $\mathrm{d}^2\mathbf{A}$ can then be written as:

$$\begin{aligned} \mathrm{d}\mathbf{A} &= \mathrm{diag}(\mathbf{u}_0) \cdot \mathrm{d}\mathbf{E} \cdot \mathrm{diag}(\bar{\mathbf{u}}_0) \\ \mathrm{d}^2\mathbf{A} &= \mathrm{diag}(\mathbf{u}_0) \cdot (\mathrm{d}\mathbf{E} \odot \mathrm{d}\mathbf{E}) \cdot \mathrm{diag}(\bar{\mathbf{u}}_0). \end{aligned}$$

This gives the result in the statement. □

Proof of Theorem 4.1. First of all, we remark that $\mathbf{u} = \mathbf{u}_0 + \mathrm{d}\mathbf{u} + o\left(\|\phi_\ell\|_\infty\right) + o_\mathbb{P}\left(\sigma\right)$ and $\lambda = n + \frac{1}{2}\mathrm{d}^2\lambda + o\left(\|\phi_\ell\|_\infty^2\right) + o_\mathbb{P}\left(\sigma^2\right)$. In terms of Saaty's consistency index (i.e. $\mu = \frac{\lambda - n}{n-1}$), we have $\mu = \frac{\mathrm{d}^2\lambda}{2(n-1)} + o\left(\|\phi_\ell\|_\infty^2\right) + o_\mathbb{P}\left(\sigma^2\right)$. Therefore, using the commutation matrix, we derive some alternative formulation of the results of Theorem 3.3:

$$\begin{aligned} \mathrm{d}^2\lambda &= \frac{2 \cdot \mathbf{u}_n^\mathsf{T} \left(\mathrm{d}\mathbf{E} \cdot \mathrm{d}\mathbf{E}\right) \mathbf{u}_n}{n^2} + \frac{\mathbf{u}_n^\mathsf{T} \left(\mathrm{d}\mathbf{E} \odot \mathrm{d}\mathbf{E}\right) \mathbf{u}_n}{n} \\ &= \frac{2\,\mathrm{vec}\left(\mathrm{d}\mathbf{E}^\mathsf{T} \cdot \mathbf{u}_n\right)^\mathsf{T} \cdot \mathrm{vec}\left(\mathrm{d}\mathbf{E} \cdot \mathbf{u}_n\right)}{n^2} + \frac{\mathrm{tr}\left(\mathrm{d}\mathbf{E} \cdot \mathrm{d}\mathbf{E}^\mathsf{T}\right)}{n} \\ &= \frac{2\left(\left(\mathbf{u}_n^\mathsf{T} \otimes \mathbf{I}_n\right) \cdot \mathrm{vec}\left(\mathrm{d}\mathbf{E}^\mathsf{T}\right)\right)^\mathsf{T} \cdot \left(\mathbf{u}_n^\mathsf{T} \otimes \mathbf{I}_n\right) \cdot \mathrm{vec}\left(\mathrm{d}\mathbf{E}\right)}{n^2} \\ &\quad + \frac{\mathrm{vec}\left(\mathrm{d}\mathbf{E}\right)^\mathsf{T} \cdot \mathrm{vec}\left(\mathrm{d}\mathbf{E}\right)}{n} \\ &= \mathrm{vec}\left(\mathrm{d}\mathbf{E}\right)^\mathsf{T} \cdot \left\{\frac{2\mathbf{K}_{nn}\left(\mathbf{U}_n \otimes \mathbf{I}_n\right)}{n^2} + \frac{\mathbf{I}_{n^2}}{n}\right\} \cdot \mathrm{vec}\left(\mathrm{d}\mathbf{E}\right) \\ \mathrm{d}\mathbf{u} &= \frac{1}{n} \cdot \left(\mathbf{u}_n^\mathsf{T} \otimes \left(\mathrm{diag}(\mathbf{u}_0) - \mathbf{u}_0 \cdot \mathbf{u}_0^\mathsf{T}\right)\right) \cdot \mathrm{vec}\left(\mathrm{d}\mathbf{E}\right). \end{aligned}$$

We recall that $\mathrm{d}\mathbf{E} = \sum_{\ell=2}^{L} \phi_\ell \cdot \left[\overline{\ln}\mathbf{A}_0\right]^{\odot \ell} + \mathbf{N} + o\left(\|\phi_\ell\|_\infty\right) + o_\mathbb{P}(\sigma)$. Moreover:

$$\overline{\ln}\mathbf{A}_0 = \overline{\ln}\left(\mathbf{u}_0 \bar{\mathbf{u}}_0^\mathsf{T}\right) = \left(\overline{\ln}\mathbf{u}_0\right) \cdot \mathbf{u}_n^\mathsf{T} - \mathbf{u}_n \cdot \left(\overline{\ln}\mathbf{u}_0\right)^\mathsf{T}$$

and taking the vec's on both sides:

$$\begin{aligned} \mathrm{vec}\left(\overline{\ln}\mathbf{A}_0\right) &= \mathrm{vec}\left[\left(\overline{\ln}\mathbf{u}_0\right) \cdot \mathbf{u}_n^\mathsf{T} - \mathbf{u}_n \cdot \left(\overline{\ln}\mathbf{u}_0\right)^\mathsf{T}\right] \\ &= \mathrm{vec}\left[\left(\overline{\ln}\mathbf{u}_0\right) \cdot \mathbf{u}_n^\mathsf{T}\right] - \mathrm{vec}\left[\mathbf{u}_n \cdot \left(\overline{\ln}\mathbf{u}_0\right)^\mathsf{T}\right] \\ &= \mathbf{u}_n \otimes \left(\overline{\ln}\mathbf{u}_0\right) - \left(\overline{\ln}\mathbf{u}_0\right) \otimes \mathbf{u}_n \\ &= \left(\mathbf{I}_{n^2} - \mathbf{K}_{nn}\right) \cdot \mathbf{u}_n \otimes \left(\overline{\ln}\mathbf{u}_0\right). \end{aligned}$$

Therefore:

$$\begin{aligned} \mathrm{vec}\left(\mathrm{d}\mathbf{E}\right) &= \sum_{\ell=2}^{L} \phi_\ell \cdot \left[\mathrm{vec}\left(\overline{\ln}\mathbf{A}_0\right)\right]^{\odot \ell} + \mathrm{vec}(\mathbf{N}) \\ &\quad + o\left(\|\phi_\ell\|_\infty\right) + o_\mathbb{P}\left(\sigma\right) \\ &= \sum_{\ell=2}^{L} \phi_\ell \cdot \left[\left(\mathbf{I}_{n^2} - \mathbf{K}_{nn}\right) \cdot \mathbf{u}_n \otimes \left(\overline{\ln}\mathbf{u}_0\right)\right]^{\odot \ell} + \mathrm{vec}\left(\mathbf{N}\right) \\ &\quad + o\left(\|\phi_\ell\|_\infty\right) + o_\mathbb{P}(\sigma) \\ &= \boldsymbol{\mu}_\phi + \mathrm{vec}(\mathbf{N}) + o\left(\|\phi_\ell\|_\infty\right) + o_\mathbb{P}(\sigma). \end{aligned}$$

This leads us to the following expressions that will be used extensively in the following:

$$\begin{aligned} \mathbf{u} &= \mathbf{u}_0 + \frac{1}{n} \cdot \left(\mathbf{u}_n^\mathsf{T} \otimes \left(\mathrm{diag}(\mathbf{u}_0) - \mathbf{u}_0 \cdot \mathbf{u}_0^\mathsf{T}\right)\right) \\ &\quad \cdot \left[\boldsymbol{\mu}_\phi + \mathrm{vec}(\mathbf{N})\right] + o\left(\|\phi_\ell\|_\infty\right) + o_\mathbb{P}(\sigma) \end{aligned}$$

$$d^2\lambda = \frac{\mathbf{v}_0^\mathsf{T}\mathbf{u}_0 \cdot \mathbf{v}_0^\mathsf{T} d^2\mathbf{A}\mathbf{u}_0}{\left(\mathbf{v}_0^\mathsf{T}\cdot\mathbf{u}_0\right)^2} + \frac{2\mathbf{v}_0^\mathsf{T}\left(\mathbf{u}_0\mathbf{v}_0^\mathsf{T} d\mathbf{A} - d\mathbf{A}\mathbf{u}_0\cdot\mathbf{v}_0^\mathsf{T}\right)\left(\mathbf{I}_n - \mathbf{u}_0\mathbf{u}_n^\mathsf{T}\right)\cdot\left(\lambda_0\mathbf{I}_n - \mathbf{A}_0\right)^+\left(\mathbf{I}_n - \frac{\mathbf{u}_0\mathbf{v}_0^\mathsf{T}}{\mathbf{v}_0^\mathsf{T}\mathbf{u}_0}\right)(d\mathbf{A})\,\mathbf{u}_0}{\left(\mathbf{v}_0^\mathsf{T}\cdot\mathbf{u}_0\right)^2}$$

Box II.

$$\mu = \frac{1}{2(n-1)}\cdot\left[\boldsymbol{\mu}_\phi + \mathrm{vec}(\mathbf{N})\right]^\mathsf{T}\cdot\left\{\frac{2\mathbf{K}_{nn}(\mathbf{U}_n\otimes\mathbf{I}_n)}{n^2} + \frac{\mathbf{I}_{n^2}}{n}\right\} \cdot\left[\boldsymbol{\mu}_\phi + \mathrm{vec}(\mathbf{N})\right] + o_\mathbb{P}\left(\|\phi_\ell\|_\infty\cdot\sigma\right).$$

We need expressions for the variances $\mathbb{V}[\mathrm{vec}(\mathbf{N})]$ and $\mathbb{V}(\frac{1}{n}\cdot(\mathbf{u}_n^\mathsf{T}\otimes(\mathrm{diag}[\mathbf{u}_0]-\mathbf{u}_0\cdot\mathbf{u}_0^\mathsf{T}))\cdot\mathrm{vec}(\mathbf{N}))$. It can be shown that:

$$\mathbb{V}(\mathrm{vec}(\mathbf{N})) = \sigma^2\cdot\left(\mathbf{I}_{n^2} - \mathbf{K}_{nn}\right)$$

and:

$$\mathbb{V}\left(\frac{1}{n}\cdot\left(\mathbf{u}_n^\mathsf{T}\otimes\left(\mathrm{diag}(\mathbf{u}_0)-\mathbf{u}_0\cdot\mathbf{u}_0^\mathsf{T}\right)\right)\cdot\mathrm{vec}(\mathbf{N})\right) = \frac{\sigma^2}{n}\cdot\left(\mathrm{diag}(\mathbf{u}_0)-\mathbf{u}_0\cdot\mathbf{u}_0^\mathsf{T}\right)^2.$$

Now we pass to the proof of the statements.

(a) The proofs in this case are exactly the same as those of point (b) when $c = +\infty$.

(b) We start from $\sigma^{-1}\cdot(\boldsymbol{\mu}_\phi + \mathrm{vec}(\mathbf{N}))$ and we remark that it tends to the following normal distribution:

$$\sigma^{-1}\cdot\left(\boldsymbol{\mu}_\phi + \mathrm{vec}(\mathbf{N})\right) = \frac{\|\phi_\ell\|_\infty}{\sigma}\cdot\frac{\boldsymbol{\mu}_\phi}{\|\phi_\ell\|_\infty} + \sigma^{-1}\cdot\mathrm{vec}(\mathbf{N}) \to_\mathcal{D} \frac{1}{c}\cdot\boldsymbol{\mu} + \mathcal{N}\left(\mathbf{0},\left(\mathbf{I}_{n^2}-\mathbf{K}_{nn}\right)\right).$$

From this the asymptotic result for $\mathbf{u}$ follows. As concerns the result for μ, we use Theorem 3.1 in Tan (1977), identifying his μ with $\frac{1}{c}\cdot\lim_{\|\phi_\ell\|_\infty\downarrow 0}\frac{\boldsymbol{\mu}_\phi}{\|\phi_\ell\|_\infty}$, his A with $\left\{\frac{2\mathbf{K}_{nn}(\mathbf{U}_n\otimes\mathbf{I}_n)}{n^2}+\frac{\mathbf{I}_{n^2}}{n}\right\}$ and his V with $\left(\mathbf{I}_{n^2}-\mathbf{K}_{nn}\right)$. The eigenvalues of $V\cdot A = \left(\mathbf{I}_{n^2}-\mathbf{K}_{nn}\right)\cdot\left\{\frac{2\mathbf{K}_{nn}(\mathbf{U}_n\otimes\mathbf{I}_n)}{n^2}+\frac{\mathbf{I}_{n^2}}{n}\right\}$ can be found reasoning as follows. We want to show that $\mathbf{A}_1 \triangleq \frac{n}{2}\cdot(\mathbf{I}_{n^2}-\mathbf{K}_{nn})\cdot\left\{\frac{2\mathbf{K}_{nn}(\mathbf{U}_n\otimes\mathbf{I}_n)}{n^2}+\frac{\mathbf{I}_{n^2}}{n}\right\}$ has $\frac{(n-1)(n-2)}{2}$ eigenvalues equal to 1. We use the equality $\left(\mathbf{I}_{n^2}-\mathbf{K}_{nn}\right)\cdot\left(\mathbf{I}_{n^2}-\mathbf{K}_{nn}\right) = 2(\mathbf{I}_{n^2}-\mathbf{K}_{nn})$ and we remark that the eigenvalues of $\mathbf{A}_1$ are the same as the ones of the symmetric matrix $\mathbf{A}_2 \triangleq \frac{n}{4}\cdot\left(\mathbf{I}_{n^2}-\mathbf{K}_{nn}\right)\cdot\left\{\frac{2\mathbf{K}_{nn}(\mathbf{U}_n\otimes\mathbf{I}_n)}{n^2}+\frac{\mathbf{I}_{n^2}}{n}\right\}\cdot(\mathbf{I}_{n^2}-\mathbf{K}_{nn})$. Now, by exploiting the relations $\left(\mathbf{I}_{n^2}-\mathbf{K}_{nn}\right)\cdot(\mathbf{I}_{n^2}-\mathbf{K}_{nn}) = 2\left(\mathbf{I}_{n^2}-\mathbf{K}_{nn}\right)$, $\left(\mathbf{I}_{n^2}-\mathbf{K}_{nn}\right)\cdot\mathbf{K}_{nn} = \left(\mathbf{K}_{nn}-\mathbf{I}_{n^2}\right)$ and $\mathbf{K}_{nn}\mathbf{U}_{n^2} = \mathbf{U}_{n^2}$, it is possible to show through some extremely tedious algebra, available from the authors upon request, that $\mathbf{A}_2$ is idempotent. The number of nonzero eigenvalues is therefore given by the trace of matrix $\mathbf{A}_1$, that can be shown to be equal to $\frac{(n-2)(n-1)}{2}$. Since all χ^2 random variables appearing in the linear combination of Theorem 3.1 in Tan (1977) have the same weight, the asymptotic distribution is a noncentral chi-squared distribution whose noncentrality parameter can be computed as:

$$\delta \triangleq \frac{1}{c^2}\cdot\left[\lim_{\|\phi_\ell\|_\infty\downarrow 0}\frac{\boldsymbol{\mu}_\phi}{\|\phi_\ell\|_\infty}\right]^\mathsf{T}\cdot\left\{\frac{2\mathbf{K}_{nn}(\mathbf{U}_n\otimes\mathbf{I}_n)}{n^2}+\frac{\mathbf{I}_{n^2}}{n}\right\} \cdot\left[\lim_{\|\phi_\ell\|_\infty\downarrow 0}\frac{\boldsymbol{\mu}_\phi}{\|\phi_\ell\|_\infty}\right].$$

(c) Statement (i) is simple. Statement (ii) is obtained opening the square $\left[\boldsymbol{\mu}_\phi+\mathrm{vec}(\mathbf{N})\right]^\mathsf{T}\cdot\left\{\frac{2\mathbf{K}_{nn}(\mathbf{U}_n\otimes\mathbf{I}_n)}{n^2}+\frac{\mathbf{I}_{n^2}}{n}\right\}\cdot\left[\boldsymbol{\mu}_\phi+\mathrm{vec}(\mathbf{N})\right]$, remarking that the leading term under $\frac{\sigma}{\|\phi_\ell\|_\infty}\to 0$ is $\boldsymbol{\mu}_\phi^\mathsf{T}\cdot\left\{\frac{2\mathbf{K}_{nn}(\mathbf{U}_n\otimes\mathbf{I}_n)}{n^2}+\frac{\mathbf{I}_{n^2}}{n}\right\}\cdot\boldsymbol{\mu}_\phi$ and that the second order term (the one that determines the asymptotic distribution) is $2\boldsymbol{\mu}_\phi^\mathsf{T}\cdot\left\{\frac{2\mathbf{K}_{nn}(\mathbf{U}_n\otimes\mathbf{I}_n)}{n^2}+\frac{\mathbf{I}_{n^2}}{n}\right\}\cdot\mathrm{vec}(\mathbf{N}) = O_\mathbb{P}\left(\|\phi_\ell\|_\infty\cdot\sigma\right)$ while $\mathrm{vec}(\mathbf{N})^\mathsf{T}\cdot\left\{\frac{2\mathbf{K}_{nn}(\mathbf{U}_n\otimes\mathbf{I}_n)}{n^2}+\frac{\mathbf{I}_{n^2}}{n}\right\}\cdot\mathrm{vec}(\mathbf{N})$ is asymptotically negligible. □

References

Aczél, J., & Luce, R. D. (2007). A behavioral condition for Prelec's weighting function on the positive line without assuming $W(1) = 1$. *Journal of Mathematical Psychology*, *51*, 126–129.

Becker, G., DeGroot, M., & Marschak, J. (1963). Stochastic models of choice behavior. *Behavioral Science*, *8*, 41–55.

Bernasconi, M., Choirat, C., & Seri, R. (2008). Measurement by subjective estimation: testing for separable representations. *Journal of Mathematical Psychology*, *52*, 184–201.

Bernasconi, M., Choirat, C., & Seri, R. (2010). The analytic hierarchy process and the theory of measurement. *Management Science*, *56*, 699–711.

Crawford, G. B., & Williams, C. (1985). A note on the analysis of subjective judgment matrices. The geometric mean procedure for estimating the scale of a judgment matrix. *Journal of Mathematical Psychology*, *29*, 387–405.

de Jong, P. (1984). A statistical approach to Saaty's scaling method for priorities. *Journal of Mathematical Psychology*, *28*, 467–478.

Dyer, J. S. (1990). Remarks on the analytic hierarchy process. *Management Science*, *36*, 249–258.

Ellermeier, W., & Faulhammer, G. (2000). Empirical evaluation of axioms fundamental to Stevens' ratio-scaling approach. I. Loudness production. *Perception & Psychophysics*, *62*, 1505–1511.

Fechner, G. T. (1860). *Elemente der Psychophysik*. Leipzig: Breitkopf und Härtel, (Reprinted, Bristol: Thoemmes Press, 1999).

Forman, E. H., & Gass, S. I. (2001). The analytic hierarchy process—an exposition. *Operations Research*, *49*, 469–486.

Genest, C., & Rivest, L.-P. (1994). A statistical look at Saaty's method of estimating pairwise preferences expressed on a ratio scale. *Journal of Mathematical Psychology*, *38*, 477–496.

Hey, J. D. (2005). Why we should not be silent about noise. *Experimental Economics*, *8*, 325–345.

Loomes, G. (2005). Modelling the stochastic component of behaviour in experiments: some issues for the interpretation of data. *Experimental Economics*, *8*, 301–323.

Loomes, G., & Sugden, R. (1995). Incorporating a stochastic element into decision theories. *European Economic Review*, *39*, 641–648.

Luce, R. D. (1997). Several unresolved problems of mathematical psychology. *Journal of Mathematical Psychology*, *41*, 79–97.

Luce, R. D. (2001). Reduction invariance and Prelec's weighting functions. *Journal of Mathematical Psychology*, *45*, 167–179.

Luce, R. D. (2002). A psychophysical theory of intensity proportions, joint presentations, and matches. *Psychological Review*, *109*, 520–532.

Luce, R. D. (2004). Symmetric and asymmetric matching of joint presentations. *Psychological Review*, *111*, 446–454.

Luce, R. D., & Narens, L. (2008). Measurement, theory of. In S. N. Durlauf, & L. E. Blume (Eds.), *The new palgrave dictionary of economics* (second ed.) Basingstoke: Palgrave Macmillan.

Magnus, J. R., & Neudecker, H. (1999). *Wiley series in probability and statistics*, *Matrix differential calculus with applications in statistics and econometrics*. Chichester: John Wiley and Sons Limited.

Meyer, C. D. (1973). Generalized inversion of modified matrices. *SIAM Journal on Applied Mathematics*, *24*, 315–323.

Michell, J. (1999). *Measurement in psychology: a critical history of a methodological concept*. Cambridge, UK: Cambridge University Press.

Myung, J. I., Karabatsos, G., & Iverson, G. I. (2005). A Bayesian approach to testing decision making axioms. *Journal of Mathematical Psychology*, *49*, 205–225.

Narens, L. (1996). A theory of ratio magnitude estimation. *Journal of Mathematical Psychology*, *40*, 109–788.

Narens, L. (2002). The irony of measurement by subjective estimations. *Journal of Mathematical Psychology*, *46*, 769–788.

Narens, L. (2006). Symmetry, direct measurement, and Torgerson's conjecture. *Journal of Mathematical Psychology*, *50*, 290–301.

Prelec, D. (1998). The probability weighting function. *Econometrica*, *66*, 497–527.

Saaty, T. L. (1977). A scaling method for priorities in a hierarchichal structure. *Journal of Mathematical Psychology*, *15*, 234–281.

Saaty, T. L. (1980). *The analytic hierarchy process*. New York: McGraw-Hill.
Saaty, T. L. (1986). Axiomatic foundation of the analytic hierarchy process. *Management Science*, *32*, 841–855.
Shepard, R. N. (1981). On the status of "Direct" psychological measurement. In C. W. Savage (Ed.), *Minnesota studies in the philosophy of science*: *Vol. 9*. Minneapolis: University of Minnesota Press.
Steingrimsson, R., & Luce, R. D. (2005a). Evaluating a model of global psychophysical judgments—I: behavioral properties of summations and productions. *Journal of Mathematical Psychology*, *49*, 290–307.
Steingrimsson, R., & Luce, R. D. (2005b). Evaluating a model of global psychophysical judgments—II: behavioral properties linking summations and productions. *Journal of Mathematical Psychology*, *49*, 308–319.
Steingrimsson, R., & Luce, R. D. (2006). Empirical evaluation of a model of global psychophysical judgments: III. A form for the psychophysical function and perceptual filtering. *Journal of Mathematical Psychology*, *50*, 15–29.
Steingrimsson, R., & Luce, R. D. (2007). Empirical evaluation of a model of global psychophysical judgments: IV. Forms for the weighting function. *Journal of Mathematical Psychology*, *51*, 29–44.
Stevens, S. S. (1951). *Handbook of experimental psychology*. New York: Wiley.
Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, *64*, 153–181.
Stevens, S. S. (1975). *Psychophysics: introduction to its perceptual, neural, and social prospects*. New York: Wiley.
Stevens, S. S., & Galanter, E. (1957). Ratio scales and category scales for a dozen perceptual continua. *Journal of Experimental Psychology*, *54*, 377–411.
Tan, W. Y. (1977). On the distribution of quadratic forms in normal random variables. *Canadian Journal of Statistics*, *5*, 241–250.
Thurstone, L. L. (1931). The indifference function. *The Journal of Social Psychology*, *2*, 139–167.
Thurstone, L. L. (1959). *The measurement of values*. Chicago: The University of Chicago Press.
Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323.
Wu, G., & Gonzalez, R. (1996). Curvature of the probability weighting function. *Management Science*, *42*, 1676–1690.
Zimmer, K. (2005). Examining the validity of numerical ratios in loudness fractionation. *Perception & Psychophysics*, *67*, 569–579.