Journal of the Operations Research
Society of Japan
2007, Vol. 50, No. 4, 515-539

MATRIX BALANCING PROBLEM AND BINARY AHP

Koichi Genma
Shizuoka University

Yutaka Kato
Hosei University

Kazuyuki Sekitani
Shizuoka University

(Received October 24, 2006)

Abstract A matrix balancing problem and an eigenvalue problem are transformed into two minimum-norm point problems whose difference is only a norm. The matrix balancing problem is solved by scaling algorithms that are as simple as the power method of the eigenvalue problem. This study gives a proof of global convergence for scaling algorithms and applies the algorithm to Analytic Hierarchy process (AHP), which derives priority weights from pairwise comparison values by the eigenvalue method (EM) traditionally. Scaling algorithms provide the minimum χ square estimate from pairwise comparison values. The estimate has properties of priority weights such as right-left symmetry and robust ranking that are not guaranteed by the EM.

Keywords: AHP, minimum-norm point, matrix scaling, binary AHP, rank reversal

1. Introduction

A variety of biological, statistical and economical science data appear in the form of cross-classified tables of counts, that are often briefly expressed in nonnegative matrices. In practice, such nonnegative matrices have frequently some biased entries and missing data. This practical problem makes the nonnegative matrix being inconsistent with theoretical prior requirements. Then, we must adjust the nonnegative matrix to satisfy the prior consistency requirements.

Adjustment of the entries of the nonnegative matrix is commonly known as matrix balancing problems or matrix scaling problems, that involve both mathematically and statistically well-posed issues of practical interest. A usual way for adjustment of the matrix is to multiply its row and columns by positive constant. A well studied instance of this problem occurring in transportation planning and input-output analysis requires that the matrix be adjusted so that the row and column sums equal fixed positive values. A related problem for a square nonnegative matrix requires that the sum of entries in each of its rows equals the sum of entries in the corresponding column.

This study considers the latter matrix balancing problem, that is applied to adjustment of social accounting matrices, as described in [25]. The matrix balancing problem is equivalent to minimizing the sum of linear ratios over the positive orthant. See Eaves *et al.* [5] for details of the equivalence proof. This paper aims to guarantee a convergence of algorithms for the equivalent fractional minimization problem and to establish an application of the matrix balancing problem to Analytic Hierarchy Process (AHP).

AHP, originally developed by Saaty [22], is a widely applied multiple-criteria decision making technique, where subjective judgments by decision makers are quantified and summarized into a reciprocal matrix under each criterion. An important feature of AHP is that priority weights are given by a principal eigenvector of the reciprocal matrix. In AHP, solving the eigenvalue problem of a reciprocal matrix is called the eigenvalue method (EM). The validity of the EM is discussed in Saaty [23] and Sekitani and Yamaki [26]. The applicability and practicality of AHP can be easily confirmed in [32]. However, it is well known that the EM has shortcomings such as rank reversal phenomena and right-left asymmetry. To overcome this EM shortcomings, this paper applies the minimum χ square method [12, 30] into derivation of the priority weights from the reciprocal matrix. It is expected that the application of the minimum χ square method to AHP will open up a new evaluation method for not only the reciprocal matrix of AHP but also nonnegative matrices corresponding to cross-classified tables, e.g., a supermatrix of Analytic Network Process, the extension of AHP.

This study shows that application of the minimum χ square method to AHP is reduced to solving the matrix balancing problem, whose computation procedure are generally called scaling algorithms [10, 25]. Scaling algorithms are as simple as the power method, which is a well known algorithm for the eigenvalue problem. The existence of the power method contributes AHP practicality. By using equivalence between the matrix balancing problem and the minimum fractional problem [5], this paper also gives a transparent proof for global convergence of scaling algorithms. Therefore, on the computational aspect, the minimum χ square method is compatible to the EM.

For an irreducible nonnegative matrix including a reciprocal matrix, this paper introduces a unified framework, a class of minimum-norm point problems [27], into the matrix balancing problem and the eigenvalue problem. The unified framework shows that these two problems are to find a closest point to the origin under the distinct L_P norms. This fact gives necessary and sufficient conditions of equivalence between the matrix balancing problem and the eigenvalue problem. The equivalence conditions plays a key role of comparison between priority weights of the EM and that of the minimum χ square method in Binary AHP that is simplified from AHP by Takahashi [28] and Jensen [13].

This article is organized as follows: Section 2 shows necessary and sufficient conditions of equivalence between the matrix balancing problem and the eigenvalue problem. Section 3 gives a convergence proof of scaling algorithms for the matrix balancing problem. Section 4 introduces the minimum χ square method of AHP and shows properties of the minimum χ square estimate. Section 5 compares the minimum χ square method with the EM in Binary AHP. The numerical experiments report that the minimum χ square method has the less frequency of rank reversal than the EM. Conclusion and future extensions are documented in Section 6.

2. Matrix Balancing Problem and Eigenvalue Problem

Before discussing priority weights deriving from a pairwise matrix in AHP, we show more general results for nonnegative matrices. For an $n \times n$ nonnegative matrix A, we introduce three definitions "sum-symmetry", "reducible" and "irreducible". The matrix A is called sum-symmetry if

$$\sum_{k=1}^{n} a_{lk} = \sum_{k=1}^{n} a_{kl} \qquad \text{for all } l = 1, \ldots, n. \tag{2.1}$$

The matrix A is said to be reducible, either if A is the 1×1 zero matrix or if $n \geq 2$ and there exists a permutation matrix P such that $PAP^\top = \begin{bmatrix} B & \boldsymbol{O} \\ C & D \end{bmatrix}$ where B and D are square matrices and $\boldsymbol{O}$ is a zero matrix. The matrix A is irreducile if it is not reducible.

Suppose that an $n \times n$ matrix A is irreducible, then a matrix balancing problem for A

is defined as a problem of finding a positive vector $\boldsymbol{x}^*$ such that $\left(a_{ij}\frac{x_j^*}{x_i^*}\right)$ is sum-symmetry. Hereafter, we assume that the matrix A is irreducible. For a positive vector $\boldsymbol{x}$, we define

$$f(\boldsymbol{x}) \equiv \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\frac{x_j}{x_i}. \tag{2.2}$$

The following lemma shows that the matrix balancing problem is equivalent to minimize $f(\boldsymbol{x})$ of (2.2) over the positive orthant.

Lemma 2.1 *There exists an optimal solution of*

$$\min_{\boldsymbol{x}>0} f(\boldsymbol{x}), \tag{2.3}$$

whose optimal solution is unique up to a positive scalar multiple. Furthermore, $\boldsymbol{x}^*$ is an optimal solution of (2.3) if and only if the $n \times n$ matrix $\left(a_{ij}\frac{x_j^*}{x_i^*}\right)$ is sum-symmetry.

Proof: See Theorem 3.4.4 of [2] for the existence of optimal solutions of (2.3). The equivalence between a solution of the matrix balancing problem for A and an optimal solution of (2.3) is shown in Theorem 3 of [5]. □

For a positive vector $\boldsymbol{x}$, we define

$$\boldsymbol{r}(\boldsymbol{x}) \equiv \left(\sum_{j=1}^{n} a_{1j}\frac{x_j}{x_1}, \cdots, \sum_{j=1}^{n} a_{nj}\frac{x_j}{x_n}\right) \quad \text{and} \tag{2.4}$$

$$\boldsymbol{c}(\boldsymbol{x}) \equiv \left(\sum_{i=1}^{n} a_{i1}\frac{x_1}{x_i}, \cdots, \sum_{i=1}^{n} a_{in}\frac{x_n}{x_i}\right). \tag{2.5}$$

By using (2.4) and (2.5), Lemma 2.1 is represented alternatively by the following manner:

Lemma 2.2 *Let $\boldsymbol{e}$ be an n-dimensional vector whose entries are all one, then any positive vector $\boldsymbol{x}$ satisfies*

$$f(\boldsymbol{x}) = \boldsymbol{c}(\boldsymbol{x})\boldsymbol{e} = \boldsymbol{r}(\boldsymbol{x})\boldsymbol{e}. \tag{2.6}$$

A positive vector $\boldsymbol{x}^*$ satisfies

$$\boldsymbol{c}(\boldsymbol{x}^*) = \boldsymbol{r}(\boldsymbol{x}^*) \tag{2.7}$$

if and only if $\boldsymbol{x}^*$ is an optimal solution of (2.3).

Proof: The proof is directly from definitions of $\boldsymbol{r}(\cdot)$ and $\boldsymbol{c}(\cdot)$ and Lemma 2.1. □

A well known theorem of a principal eigenvector for an irreducible matrix, Perron-Frobenius Theorem, is a key way of transforming an eigenvalue problem into an optimization problem:

Lemma 2.3 *Finding the principal eigenvalue of A is to solve*

$$\min_{\boldsymbol{x}>0} \max \left\{\sum_{j=1}^{n} a_{1j}\frac{x_j}{x_1}, \cdots, \sum_{j=1}^{n} a_{nj}\frac{x_j}{x_n}\right\}, \tag{2.8}$$

whose optimal solution is unique up to a positive scalar multiple.

Proof: The proof is directly proved by Perron-Frobenius Theorem. See chapter 4 of [29] for Perron-Frobenius Theorem. □

Two optimization problems (2.3) and (2.8) are reduced to minimum-norm point problems, respectively, whose norms are dual:

Lemma 2.4 *Problem (2.3) and a principal eigenvalue problem are equivalent to*

$$\min_{\boldsymbol{x}>0} \|\boldsymbol{r}(\boldsymbol{x})\|_1 \quad \textit{and} \tag{2.9}$$

$$\min_{\boldsymbol{x}>0} \|\boldsymbol{r}(\boldsymbol{x})\|_\infty , \tag{2.10}$$

respectively, where $\|\boldsymbol{x}\|_1 = \sum_{i=1}^n |x_i|$ *and* $\|\boldsymbol{x}\|_\infty = \max\{|x_1|, \cdots, |x_n|\}$.

Proof: Choose a positive vector $\boldsymbol{x}$. Since $r_i(\boldsymbol{x}) > 0$ for all $i = 1, \dots, n$, $\|\boldsymbol{r}(\boldsymbol{x})\|_1 = \boldsymbol{r}(\boldsymbol{x})^\top \boldsymbol{e}$ and $\|\boldsymbol{r}(\boldsymbol{x})\|_\infty = \max\{r_1(\boldsymbol{x}), \cdots, r_n(\boldsymbol{x})\}$. The proof is complete from lemmas 2.2 and 2.3. □

Kalantari *et al.* [16] also generally consider the matrix balancing problem (2.3) in a framework of L_p-norm minimization problem, however, whose norm is measured for an $n \times n$ vector $(a_{11}, a_{12}x_2/x_1, \cdots, a_{1n}x_n/x_1, \ \cdots, a_{n1}x_1/x_n, \cdots, a_{nn})^\top$. The matrix balancing problem is equivalent to (2.3) and the principal eigenvalue problem is (2.8). From Lemma 2.4, the matrix balancing problem (2.3) and the eigenvalue problem (2.8) implicitly have the dual norms for a common positive vector as their corresponding objective functions. This is summarized into the following relations between the matrix balancing problem and the eigenvalue problem:

Theorem 2.1 *Let* $\boldsymbol{x}^*$ *be an optimal solution of (2.3) and let* $\boldsymbol{x}^\#$ *be a right principal eigenvector of* A*, then we have*

$$\boldsymbol{r}(\boldsymbol{x}^\#) = \lambda \boldsymbol{e} \qquad \textit{and} \tag{2.11}$$

$$\left\|\boldsymbol{r}(\boldsymbol{x}^\#)\right\|_\infty = \frac{1}{n}\left\|\boldsymbol{r}(\boldsymbol{x}^\#)\right\|_1 \geq \frac{1}{n}\left\|\boldsymbol{r}(\boldsymbol{x}^*)\right\|_1 , \tag{2.12}$$

where λ *is the principal eigenvalue of* A *and* $\boldsymbol{e}$ *is an* n*-dimensional row vector whose entries are all 1. Moreover,* $\boldsymbol{x}^\#$ *is equal to a positive scalar multiple of* $\boldsymbol{x}^*$ *if and only if* $\left\|\boldsymbol{r}(\boldsymbol{x}^\#)\right\|_\infty = \frac{1}{n}\left\|\boldsymbol{r}(\boldsymbol{x}^*)\right\|_1$.

Proof: Since $A\boldsymbol{x}^\# = \lambda \boldsymbol{x}^\#$, we have

$$\boldsymbol{r}(\boldsymbol{x}^\#) = \left(\frac{\sum_{j=1}^n a_{1j}x_j^\#}{x_1^\#}, \dots, \frac{\sum_{j=1}^n a_{nj}x_j^\#}{x_n^\#}\right) = \left(\frac{\lambda x_1^\#}{x_1^\#}, \dots, \frac{\lambda x_n^\#}{x_n^\#}\right) = \lambda(1, \dots, 1) = \lambda \boldsymbol{e}, \tag{2.13}$$

implying from Lemma 2.4 that

$$\left\|\boldsymbol{r}(\boldsymbol{x}^\#)\right\|_\infty = \lambda = \frac{1}{n}\left\|\boldsymbol{r}(\boldsymbol{x}^\#)\right\|_1 \geq \frac{1}{n}\min_{\boldsymbol{x}>0}\left\|\boldsymbol{r}(\boldsymbol{x})\right\|_1 = \frac{1}{n}\left\|\boldsymbol{r}(\boldsymbol{x}^*)\right\|_1 . \tag{2.14}$$

Suppose that $\left\|\boldsymbol{r}(\boldsymbol{x}^\#)\right\|_\infty = \frac{1}{n}\left\|\boldsymbol{r}(\boldsymbol{x}^*)\right\|_1$, then it follows from (2.14) that $\left\|\boldsymbol{r}(\boldsymbol{x}^\#)\right\|_1 = \left\|\boldsymbol{r}(\boldsymbol{x}^*)\right\|_1$. Therefore, $\boldsymbol{x}^\#$ is also an optimal solution of (2.9), that is equivalent to (2.3), by Lemma 2.4. Thus, Lemma 2.1 implies that $\boldsymbol{x}^\#$ is equal to a positive scalar multiple of $\boldsymbol{x}^*$.

Conversely, suppose that $\boldsymbol{x}^\#$ is equal to a positive scalar multiple of $\boldsymbol{x}^*$, then $\boldsymbol{x}^*$ is also a right principal eigenvector of A and hence, it follows from (2.13) that

$$\frac{1}{n}\|\boldsymbol{r}(\boldsymbol{x}^*)\|_1 = \frac{1}{n}\|\lambda\boldsymbol{e}\|_1 = \frac{\lambda}{n}\|\boldsymbol{e}\|_1 = \frac{\lambda}{n}n = \lambda = \lambda\|\boldsymbol{e}\|_\infty = \|\lambda\boldsymbol{e}\|_\infty = \left\|\boldsymbol{r}(\boldsymbol{x}^\#)\right\|_\infty . \qquad \square$$

A pair of a left and a right eigenvector of A characterizes an equivalence relation between the matrix balancing problem and the eigenvalue problem as follows:

Corollary 2.1 *Suppose that* $\boldsymbol{x}^*$ *is an optimal solution of (2.3) and that* $\boldsymbol{x}^\#$ *is a right principal eigenvector of* A. *If* $\left(1/x_1^\#, \cdots, 1/x_n^\#\right)$ *is a left principal eigenvector of* A, *then*

$$\left\|\boldsymbol{r}(\boldsymbol{x}^\#)\right\|_\infty = \frac{1}{n}\left\|\boldsymbol{r}(\boldsymbol{x}^*)\right\|_1 \tag{2.15}$$

and vice versa.

Proof: Let λ be the principal eigenvalue of A and let $\boldsymbol{e}$ be an n-dimensional row vector whose entries are all 1. Since $\boldsymbol{x}^\#$ is a right principal eigenvector of A, it follows from (2.11) of Theorem 2.1 that

$$\boldsymbol{r}(\boldsymbol{x}^\#) = \lambda \boldsymbol{e}. \tag{2.16}$$

Suppose that $\left(1/x_1^\#, \cdots, 1/x_n^\#\right)$ is a left principal eigenvector of A, then it follows that

$$\boldsymbol{c}(\boldsymbol{x}^\#) = \left(\sum_{i=1}^n \frac{a_{i1}}{x_i^\#}x_1^\#, \ldots, \sum_{i=1}^n \frac{a_{in}}{x_i^\#}x_n^\#\right) = \left(\frac{\lambda}{x_1^\#}x_1^\#, \ldots, \frac{\lambda}{x_n^\#}x_n^\#\right) = (\lambda, \ldots, \lambda) = \lambda\boldsymbol{e},$$

implying from (2.16) that $\lambda\boldsymbol{e} = \boldsymbol{c}(\boldsymbol{x}^\#) = \boldsymbol{r}(\boldsymbol{x}^\#)$. Therefore, it follows from Lemma 2.1 that $\boldsymbol{x}^\#$ is an optimal solution of (2.3) and $\|\boldsymbol{r}(\boldsymbol{x}^\#)\|_1 = \|\boldsymbol{r}(\boldsymbol{x}^*)\|_1$. Since $\lambda = \left\|\boldsymbol{r}(\boldsymbol{x}^\#)\right\|_\infty = \frac{1}{n}\left\|\boldsymbol{r}(\boldsymbol{x}^\#)\right\|_1$, an optimal solution of (2.3) and a right principal eigenvector of A satisfy (2.15).

Conversely, suppose that an optimal solution $\boldsymbol{x}^*$ of (2.3) and a right principal eigenvector of A satisfy (2.15), then we have $\left\|\boldsymbol{r}(\boldsymbol{x}^\#)\right\|_\infty = \frac{1}{n}\left\|\boldsymbol{r}(\boldsymbol{x}^\#)\right\|_1 = \frac{1}{n}\left\|\boldsymbol{r}(\boldsymbol{x}^*)\right\|_1$ and it follows from Lemma 2.4 that $\boldsymbol{x}^\#$ is also an optimal solution of (2.3). By Lemma 2.1 and (2.16) we have

$$\lambda\boldsymbol{e} = \boldsymbol{r}(\boldsymbol{x}^\#) = \boldsymbol{c}(\boldsymbol{x}^\#) = \left(\sum_{i=1}^n \frac{a_{i1}}{x_i^\#}x_1^\#, \ldots, \sum_{i=1}^n \frac{a_{in}}{x_i^\#}x_n^\#\right),$$

implying that

$$\lambda\left(\frac{1}{x_1^\#}, \ldots, \frac{1}{x_n^\#}\right) = \left(\sum_{i=1}^n \frac{a_{i1}}{x_i^\#}, \ldots, \sum_{i=1}^n \frac{a_{in}}{x_i^\#}\right).$$

Consequently, $\left(1/x_1^\#, \cdots, 1/x_n^\#\right)$ is a left principal eigenvector of A. □

Corollary 2.1 is an evidence that a principal eigenvector of a pairwise comparison matrix has a right-left asymmetry [8, 14]. (Right-left symmetry will be discussed in Theorem 4.2 of section 4)

The following examples illustrate the existence of equivalence between the matrix problem and the eigenvalue problem:

Example 2.1 *Consider a doubly stochastic matrix* $A = \begin{bmatrix} 0.0 & 0.5 & 0.0 & 0.5 & 0.0 \\ 0.8 & 0.0 & 0.0 & 0.2 & 0.0 \\ 0.0 & 0.5 & 0.2 & 0.0 & 0.3 \\ 0.2 & 0.0 & 0.3 & 0.0 & 0.5 \\ 0.0 & 0.0 & 0.5 & 0.3 & 0.2 \end{bmatrix}$, *then* A *has a left principal eigenvector* $[1,1,1,1,1]$ *and a right principal eigenvector* $[1,1,1,1,1]^\top$ *that is also an optimal solution of (2.3).*

Consider a reciprocal matrix $A = \begin{bmatrix} 1 & 1/2 & 1/3 & 1/4 \\ 2 & 1 & 2/3 & 1/2 \\ 3 & 3/2 & 1 & 3/4 \\ 4 & 2 & 4/3 & 1 \end{bmatrix}$ *then,* A *has a right principal eigenvector* $[1,2,3,4]^\top$ *and a left principal eigenvector* $[1, 1/2, 1/3, 1/4]$ *that is also an optimal solution of (2.3).*

3. Scaling Algorithms for the Matrix Balancing Problem

Algorithms for a class of matrix balancing problems are classified into two classes, scaling algorithms and optimization algorithms, which are compared with respect to various viewpoints including quality of solutions and ease of implement and use [25]. Schneider *et al.* [25] conclude that practitioners is easy to implement and use scaling algorithms rather than optimization algorithms. Scaling algorithms for (2.3) is called the diagonal similarity scaling (DSS) algorithms by Schneider *et al.* [25]. Recently, Huang *et al.* [10] proposes one of the DSS algorithms, but they do not complete a proof of global convergence of the algorithm.

The RAS algorithm, developed by [1, 17, 21], is a generalization of the DSS algorithms that can solve (2.3) under constraints of row and column sums being equal to prespecified values. A matrix balancing problem with such equalities constraints is not necessarily feasible under an assumption of irreducibility of a matrix A. A convergence of the RAS algorithms is proved under a stronger assumption of a matrix A than irreducibility that is only our assumption. For instance, see [21] for the stronger assumption.

This section gives a transparent proof of global convergence of the DSS algorithms for solving the matrix balancing problem (2.3) under only an assumption of irreducibility of the matrix. We describe a DSS algorithm consisting of three steps, the second step of which is slightly different from that of existing DSS algorithms by [10] and [25]. (The difference is independent of global convergence of DSS algorithms.)

Algorithm for solving (2.3)

Step1 Choose an initial positive vector $\boldsymbol{x}^0$ such that $\sum_{i=1}^n x_i^0 = 1$ and let $t := 0$.

Step2 Let $\boldsymbol{r} := \boldsymbol{r}(\boldsymbol{x}^t) - (a_{11}, \cdots, a_{nn})$ and $\boldsymbol{c} := \boldsymbol{c}(\boldsymbol{x}^t) - (a_{11}, \cdots, a_{nn})$. Let

$$|r_p - c_p|/x_p := \max_{k=1,\dots,n} |r_k - c_k|/x_k. \tag{3.1}$$

If $|r_p - c_p| \le \epsilon$, then $\boldsymbol{x}$ is an optimal solution of (2.3) and stop.

Step3 Let $\mu := \sqrt{r_p/c_p}$ and

$$x(\mu)_i := \begin{cases} \mu x_i^t & i = p \\ x_i^t & i \ne p \end{cases} \tag{3.2}$$

Set $\boldsymbol{x}^{t+1} := \boldsymbol{x}(\mu)/\sum_i x(\mu)_i$ and $t := t + 1$, and go to Step 2.

A proof of global convergence of the above algorithm consists of four lemmas and one theorem. As a corollary of the theorem, we guarantee global convergence of existing two DSS algorithms. The following lemma corresponds to the termination criterion of Step 2:

Lemma 3.1 *Let $\boldsymbol{r} = \boldsymbol{r}(\boldsymbol{x}) - (a_{11}, \cdots, a_{nn})$ and $\boldsymbol{c} = \boldsymbol{c}(\boldsymbol{x}) - (a_{11}, \cdots, a_{nn})$. A positive vector $\boldsymbol{x}$ is not an optimal solution of (2.3) if and only if $\boldsymbol{r} \ne \boldsymbol{c}$.*

Proof: Since $\boldsymbol{r} \ne \boldsymbol{c}$ is equivalent to $\boldsymbol{r}(\boldsymbol{x}) \ne \boldsymbol{c}(\boldsymbol{x})$, this assertion is directly followed from Lemma 2.2. □

The following lemma indicates that the update process (3.2) of $\boldsymbol{x}^t$ is a certain type of an exact line search of the step size in a descent method of nonlinear programming [33]:

Lemma 3.2 *Choose a positive vector $\boldsymbol{x}$, arbitrarily, and let $\boldsymbol{r} = \boldsymbol{r}(\boldsymbol{x}) - (a_{11}, \dots, a_{nn})$ and $\boldsymbol{c} = \boldsymbol{c}(\boldsymbol{x}) - (a_{11}, \dots, a_{nn})$. Suppose that $c_p \ne r_p$ and let*

$$x(\mu)_i := \begin{cases} \mu x_i & i = p \\ x_i & i \ne p \end{cases}, \tag{3.3}$$

then problem

$$\min_{\mu>0} f(\boldsymbol{x}(\mu)) - f(\boldsymbol{x}) \tag{3.4}$$

has a unique optimal solution $\sqrt{r_p/c_p}$ *and its optimal objective function value is* $-\left(\sqrt{r_p}-\sqrt{c_p}\right)^2$.

Proof: It follows from the definition (3.3) of $\boldsymbol{x}(\mu)$ that

$$\begin{aligned} f(\boldsymbol{x}(\mu)) - f(\boldsymbol{x}) &= \sum_{i\neq p}\sum_{j\neq p} a_{ij}\frac{x_j}{x_i} + \sum_{j\neq p} a_{pj}\frac{x_j}{\mu x_p} + \sum_{i\neq p} a_{ip}\frac{\mu x_p}{x_i} + a_{pp}\frac{\mu x_p}{\mu x_p} \\ &\quad - \left(\sum_{i\neq p}\sum_{j\neq p} a_{ij}\frac{x_j}{x_i} + \sum_{j\neq p} a_{pj}\frac{x_j}{x_p} + \sum_{i\neq p} a_{ip}\frac{x_p}{x_i} + a_{pp}\frac{x_p}{x_p}\right) \\ &= \sum_{i\neq p}\sum_{j\neq p} a_{ij}\frac{x_j}{x_i} + \frac{1}{\mu}r_p + \mu c_p + a_{pp} - \left(\sum_{i\neq p}\sum_{j\neq p} a_{ij}\frac{x_j}{x_i} + r_p + c_p + a_{pp}\right) \\ &= \mu c_p + \frac{1}{\mu}r_p - r_p - c_p. \end{aligned}$$

Let $g(\mu) = \mu c_p + \frac{1}{\mu}r_p - r_p - c_p$, then it follows from $c_p > 0$ and $r_p > 0$ that $g(\mu)$ is a strictly convex function over $\{\mu|\mu > 0\}$. Since $\mu^* = \sqrt{\frac{r_p}{c_p}}$ satisfies $\frac{dg}{d\mu} = c_p - \frac{1}{\mu^2}r_p = 0$, (3.4) has a unique optimal solution $\mu^* = \sqrt{\frac{r_p}{c_p}}$, and hence,

$$\begin{aligned} g(\mu^*) &= \mu^* c_p + \frac{1}{\mu^*}r_p - r_p - c_p = \sqrt{\frac{r_p}{c_p}}c_p + \sqrt{\frac{c_p}{r_p}}r_p - r_p - c_p = \sqrt{r_pc_p} + \sqrt{r_pc_p} - r_p - c_p \\ &= 2\sqrt{r_pc_p} - r_p - c_p = -\left(r_p - 2\sqrt{r_pc_p} + c_p\right) = -(\sqrt{r_p} - \sqrt{c_p})^2. \qquad \square \end{aligned}$$

The update process (3.2) of $\boldsymbol{x}^t$ reduces strictly the objective function value $f(\boldsymbol{x})$ of (2.2). Therefore, the sequence $\{f(\boldsymbol{x}^0), f(\boldsymbol{x}^1), \cdots\}$ generated by the algorithm is strictly decreasing. This is summarized into the following lemma:

Lemma 3.3 *Let p be an index satisfying (3.1), then*

$$-(\sqrt{r_p} - \sqrt{c_p})^2 = f(\boldsymbol{x}^{t+1}) - f(\boldsymbol{x}^t). \tag{3.5}$$

Therefore, $f(\boldsymbol{x}^t) > f(\boldsymbol{x}^{t+1}) \quad t = 0, 1, 2, \ldots.$

Proof: Since $\boldsymbol{x}^{t+1}$ is equal to a positive scalar multiple of $\boldsymbol{x}\left(\sqrt{\frac{r_p}{c_p}}\right)$ in Step 3, it follows from $f(\boldsymbol{x}^{t+1}) = f\left(\boldsymbol{x}\left(\sqrt{\frac{r_p}{c_p}}\right)\right)$ and Lemma 3.2 that

$$0 > -(\sqrt{r_p} - \sqrt{c_p})^2 = \min_{\mu>0} f(\boldsymbol{x}(\mu)) - f(\boldsymbol{x}^t) = f\left(\boldsymbol{x}\left(\sqrt{\frac{r_p}{c_p}}\right)\right) - f(\boldsymbol{x}^t) = f(\boldsymbol{x}^{t+1}) - f(\boldsymbol{x}^t). \square$$

The following lemma guarantees that there exists a compact set in the positive orthant such that the compact set contains $\{\boldsymbol{x}^0, \boldsymbol{x}^1, \cdots\}$:

Lemma 3.4 *Let $\boldsymbol{e}$ be an n-dimensional row vector whose entries are all 1. Let* $\Omega = \{\boldsymbol{x} \mid \boldsymbol{e}\boldsymbol{x} = 1 \text{ and } \boldsymbol{x} > 0\}$ *and* $\delta = \min\{a_{ij} \mid a_{ij} > 0\}$. *Choose $\boldsymbol{x}^0 \in \Omega$ arbitrarily, then, the iterate $\boldsymbol{x}^t$ of the algorithm satisfies*

$$\boldsymbol{x}^t \in \left\{ \boldsymbol{x} \;\middle|\; \boldsymbol{e}\boldsymbol{x} = 1 \text{ and } \boldsymbol{x} \geq \frac{\delta^n}{nf(\boldsymbol{x}^0)^n}\boldsymbol{e}^\top \right\} \tag{3.6}$$

for all $t = 0, 1, \ldots.$

Proof: Since $\boldsymbol{x}^t \in \Omega$, there exists an index pair (k, l) such that $x_k^t = \min_{i=1,\dots,n} x_i^t$, $x_l^t = \max_{i=1,\dots,n} x_i^t$ and $x_l^t \geq 1/n$. Since the matrix A is irreducible, there exist distinct indices $i_0, i_1, \cdots, i_q$ such that $i_0 = k, i_q = l$ and $a_{i_j i_{j+1}} > 0$ for all $j = 0, \dots, q-1$. By Lemma 3.3 and the arithmetic mean–the geometric mean inequality, we have

$$
\begin{aligned}
f(\boldsymbol{x}^0) &\geq f(\boldsymbol{x}^t) \geq \frac{1}{q} f(\boldsymbol{x}^t) \\
&\geq \frac{1}{q} \sum_{i=1}^{n} \sum_{j=1} a_{ij} \frac{x_j^t}{x_i^t} \geq \frac{1}{q} \sum_{j=0}^{q-1} a_{i_j i_{j+1}} \frac{x_{i_{j+1}}^t}{x_{i_j}^t} \geq \left(\prod_{j=0}^{q-1} a_{i_j i_{j+1}} \frac{x_{i_{j+1}}^t}{x_{i_j}^t} \right)^{1/q} \\
&\geq \left(\prod_{j=0}^{q-1} \delta \frac{x_{i_{j+1}}^t}{x_{i_j}^t} \right)^{1/q} = \delta \left(\prod_{j=0}^{q-1} \frac{x_{i_{j+1}}^t}{x_{i_j}^t} \right)^{1/q} = \delta \left(\frac{x_{i_1}^t}{x_{i_0}^t} \cdot \frac{x_{i_2}^t}{x_{i_1}^t} \cdots \frac{x_{i_{q-1}}^t}{x_{i_{q-2}}^t} \cdot \frac{x_{i_q}^t}{x_{i_{q-1}}^t} \right)^{1/q} \\
&= \delta \left(\frac{x_{i_q}^t}{x_{i_0}^t} \right)^{1/q} = \delta \left(\frac{x_l^t}{x_k^t} \right)^{1/q} \geq \delta \left(\frac{x_l^t}{x_k^t} \right)^{1/n} \geq \delta \left(\frac{1}{n x_k^t} \right)^{1/n}.
\end{aligned}
$$

Therefore, it follows from the definition of the index k that

$$
\frac{1}{n} \left(\frac{\delta}{f(\boldsymbol{x}^0)} \right)^n \leq x_k^t \leq x_i^t \qquad \text{for all } i = 1, 2, \dots, n. \qquad \square
$$

The above proof is based on (a) of Lemma 2 of Eaves *et al.* [5]. Lemma 3.4 means that the set $\left\{ \boldsymbol{x} \,\middle|\, \boldsymbol{e}\boldsymbol{x} = 1 \text{ and } \boldsymbol{x} \geq n^{-1} \left(\delta / f(\boldsymbol{x}^0) \right)^n \boldsymbol{e}^\top \right\}$ of (3.6) is compact and it is a proper subset of Ω. Therefore, accumulation points of $\{\boldsymbol{x}^0, \boldsymbol{x}^1, \cdots\}$ exists and all the functions $f(\cdot)$, $\boldsymbol{r}(\cdot)$ and $\boldsymbol{c}(\cdot)$ are well defined at any accumulation points. We can show global convergence of the algorithm as follows:

Theorem 3.1 *Let $\epsilon = 0$ in Step 2 and suppose that the algorithm provides an infinite sequence $\{\boldsymbol{x}^0, \boldsymbol{x}^1, \cdots\}$, then any accumulation point of $\{\boldsymbol{x}^0, \boldsymbol{x}^1, \cdots\}$ is an optimal solution of (2.3). Moreover, the algorithm converges globally.*

Proof: Since $\{f(\boldsymbol{x}^0), f(\boldsymbol{x}^1), \cdots\}$ is monotonically decreasing and $f(\boldsymbol{x}^t) > 0$ for all $t = 0, 1, 2, \dots$, the sequence converges to a finite value, say $\bar{f}$.

Let $\delta = \min\{a_{ij} \,|\, a_{ij} > 0\}$ and let

$$
\Omega_\delta = \left\{ \boldsymbol{x} \,\middle|\, \boldsymbol{e}^\top \boldsymbol{x} = 1 \text{ and } \boldsymbol{x} \geq \frac{1}{n} \left(\frac{\delta}{f(\boldsymbol{x}^0)} \right)^n \boldsymbol{e} \right\},
$$

where $\boldsymbol{e}$ is an n-dimensional row vector whose entries are all 1. By Lemma 3.4, the compact set Ω_δ contains $\{\boldsymbol{x}^0, \boldsymbol{x}^1, \cdots\}$, which has an accumulation point $\bar{\boldsymbol{x}} \in \Omega_\delta$. Therefore, all functions $f()$, $\boldsymbol{c}()$ and $\boldsymbol{r}()$ are well defined at $\bar{\boldsymbol{x}}$.

Let $\{\boldsymbol{x}^t \,|\, t \in T\}$ be a subsequence converging to $\bar{\boldsymbol{x}}$. Let $\bar{\boldsymbol{r}}(\boldsymbol{x}) = \boldsymbol{r}(\boldsymbol{x}) - (a_{11}, \cdots, a_{nn})$ and $\bar{\boldsymbol{c}}(\boldsymbol{x}) = \boldsymbol{c}(\boldsymbol{x}) - (a_{11}, \cdots, a_{nn})$, then the function $\max_{k=1,\dots,n} |\bar{r}_k(\boldsymbol{x}) - \bar{c}_k(\boldsymbol{x})| / x_k$ is continuous over $\{\boldsymbol{x} \,|\, \boldsymbol{e}\boldsymbol{x} = 1 \text{ and } \boldsymbol{x} > 0\}$ and the right-hand of (3.1) is $\max_{k=1,\dots,n} |\bar{r}_k(\boldsymbol{x}^t) - \bar{c}_k(\boldsymbol{x}^t)| / x_k^t$.

Let p^t be an index satisfying (3.1) at the tth iteration, then it follows from the continuity of $\max_{k=1,\dots,n} |\bar{r}_k(\boldsymbol{x}) - \bar{c}_k(\boldsymbol{x})| / x_k$ that there exists an index $\bar{p}$ satisfying (3.1) at $\bar{\boldsymbol{x}}$ such that $\{\boldsymbol{x}^t | \, t \in T \text{and } p^t = \bar{p}\}$ is an infinite sequence. Let $\bar{T} = \{t \in T \,|\, p^t = \bar{p}\}$, then it follows from $\lim_{t \to \infty} f(\boldsymbol{x}^t) = \bar{f}$ that

$$
\lim_{\substack{t \in \bar{T} \\ t \to \infty}} f(\boldsymbol{x}^t) = \lim_{\substack{t \in \bar{T} \\ t \to \infty}} f(\boldsymbol{x}^{t+1}) = \bar{f}.
$$

Therefore, it follows from (3.5) that

$$\begin{aligned} 0 &= \lim_{\substack{t\in\bar{T}\\ t\to\infty}} f(\boldsymbol{x}^t) - f(\boldsymbol{x}^{t+1}) = \lim_{\substack{t\in\bar{T}\\ t\to\infty}} (\sqrt{r_{p^t}} - \sqrt{c_{p^t}})^2 \\ &= \lim_{\substack{t\in\bar{T}\\ t\to\infty}} \left(\sqrt{r_{\bar{p}}(\boldsymbol{x}^t) - a_{\bar{p}\bar{p}}} - \sqrt{c_{\bar{p}}(\boldsymbol{x}^t) - a_{\bar{p}\bar{p}}}\right)^2 = \left(\sqrt{r_{\bar{p}}(\bar{\boldsymbol{x}}) - a_{\bar{p}\bar{p}}} - \sqrt{c_{\bar{p}}(\bar{\boldsymbol{x}}) - a_{\bar{p}\bar{p}}}\right)^2. \end{aligned}$$

Thus, we have $r_{\bar{p}}(\bar{\boldsymbol{x}}) = c_{\bar{p}}(\bar{\boldsymbol{x}})$ and it follows from the definition of $\bar{p}$ that $\boldsymbol{r}(\bar{\boldsymbol{x}}) = \boldsymbol{c}(\bar{\boldsymbol{x}})$. Lemma 2.2 implies that $\bar{\boldsymbol{x}}$ is an optimal solution $\boldsymbol{x}^*$ of (2.3), and hence, $\bar{f} = f(\bar{\boldsymbol{x}}) = f(\boldsymbol{x}^*)$.

Since an optimal solution $\boldsymbol{x}^*$ of (2.3) is unique in $\{\, \boldsymbol{x} \mid \boldsymbol{e}\boldsymbol{x} = 1 \text{ and } \boldsymbol{x} > 0\}$, any convergent subsequence of $\{\boldsymbol{x}^0, \boldsymbol{x}^1, \cdots\}$ has the same limit point $\boldsymbol{x}^*$. Therefore, the algorithm converges globally. □

For any index p satisfying $r_p(\boldsymbol{x}^t) \neq c_p(\boldsymbol{x}^t)$, even if it is not maximal of (3.1), all lemmas in this section holds. The p selection (3.1) of the second step corresponds to a function $\max_{1\le k\le n} |r_k(\boldsymbol{x}) - c_k(\boldsymbol{x})|/x_k$ at a positive vector $\boldsymbol{x}$. As stated in the proof of Theorem 3.1, the continuity of the corresponding function over the positive orthant plays a key role of global convergence of the algorithm. Therefore, we may have a globally convergent algorithm by replacing (3.1) with an alternative way of the p selection. In fact, existing two DSS algorithms adopt p selections other than (3.1). This is summarized into the following corollary:

Corollary 3.1 *Consider the algorithm replacing (3.1) with*

$$|r_p - c_p| \;:=\; \max_{k=1,\dots,n} |r_k - c_k| \qquad \textit{or} \tag{3.7}$$

$$|\sqrt{r_p} - \sqrt{c_p}| \;:=\; \max_{k=1,\dots,n} |\sqrt{r_k} - \sqrt{c_k}| \tag{3.8}$$

and let $\{\boldsymbol{x}^t \mid t = 0, 1, 2, \dots\}$ *be an infinite sequence of the modified algorithm, then modified algorithm has global convergence and its limit point of* $\{\boldsymbol{x}^t \mid t = 0, 1, 2, \dots\}$ *is an optimal solution of (2.3).*

Proof: Let $\Omega = \{\, \boldsymbol{x} \in R^n \mid \sum_{i=1}^n x_i = 1 \text{ and } \boldsymbol{x} > 0\}$. Let $\bar{\boldsymbol{r}}(\boldsymbol{x}) = \boldsymbol{r}(\boldsymbol{x}) - (a_{11}, \cdots, a_{nn})$ and $\bar{\boldsymbol{c}}(\boldsymbol{x}) = \boldsymbol{c}(\boldsymbol{x}) - (a_{11}, \cdots, a_{nn})$, then the function $\max_{k=1,\dots,n} |\bar{r}_k(\boldsymbol{x}) - \bar{c}_k(\boldsymbol{x})|$ is continuous over Ω. The right-hand of (3.7) is $\max_{k=1,\dots,n} |\bar{r}_k(\boldsymbol{x}^t) - \bar{c}_k(\boldsymbol{x}^t)|$. Therefore, we can prove global convergence of the modified algorithm replacing (3.1) with (3.7) in the same manner as the proof of Theorem 3.1.

In the similar way to the above argument, we can also prove global convergence of the modified algorithm replacing (3.1) with (3.8). □

Huang *et al.* adopt (3.7) as the p selection criteria. The selection criteria of (3.8) is in [25]. The Zangwill global convergence theory [33] may also show another proof of global convergence of these DSS algorithms.

Optimization algorithms for (2.3) are recently studied [15, 16]. Kalantari *et al.* [16] shows the polynomial-time solvalitity of (2.3) to any prescribed accuracy by using interior-point Newton methods. Johnson *et al.* [15] develop the DomEig algorithm by reducing (2.3) to the minimum dominant eigenvalue problem.

4. A χ-square Estimation Model for AHP and Binary AHP

AHP, first proposed by Saaty [22], is a decision making approach in which a decision maker's preferences are elicited through pairwise comparison of $n(\geq 2)$ alternatives on a ratio scale.

The comparison of alternative i with alternative j is quantified into pairwise comparison value a_{ij}, whose meaning is that alternative i is a_{ij} times as important as alternative j. Let $a_{ii} = 1$ for all $i = 1, \ldots, n$, then all pairwise comparison values are stored in a reciprocal matrix $A = [a_{ij}]$, that is, $a_{ij} = 1/a_{ji}$ and $a_{ij} > 0$ for all $i, j = 1, \ldots, n$. We often assume that a priority vector $\boldsymbol{x}$, whose ith entry is an ideal importance x_i of alternative i, satisfies

$$a_{ij} \approx \frac{x_i}{x_j} \qquad i, j = 1, \ldots, n. \tag{4.1}$$

Saaty recommends eigenvalue method (EM) to deriving a priority vector from a reciprocal matrix A. Therefore, a priority vector of Saaty's AHP is a principal eigenvector of A, that is, an optimal solution of (2.3). However, the EM is criticized both from prioritization and consistency points of view and several new techniques are developed on the statistical model (4.1), e.g., the geometric means of the row entries of A and an optimal solution of

$$\min.\ Q(\boldsymbol{x}) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{(a_{ij} - x_i/x_j)^2}{x_i/x_j} \tag{4.2}$$

that is called the minimum χ square method by Jensen [12]. (Strictly speaking, $Q(\boldsymbol{x})$ is not an index of Pearson's χ^2 goodness-of-fit because a_{ij} is not an observed frequency and x_i/x_j is not an expected (theoretical) frequency.)

The minimum χ square method for the matrix A is to minimize the sum of linear ratios over the positive orthant. This minimization fractional problem is reduced into a matrix balancing problem (2.3) for $\left[a_{ij}^2 + 1\right]$ by the following two lemmas:

Lemma 4.1 *For every matrix A and every positive vector $\boldsymbol{x}$,*

$$Q(\boldsymbol{x}) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{x_j}{x_i}\left(a_{ij} - \frac{x_i}{x_j}\right)^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_{ij}^2 + 1\right)\frac{x_j}{x_i} - 2\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}. \tag{4.3}$$

Proof: Since $\sum_{i=1}^{n}\sum_{j=1}^{n} x_i/x_j = \sum_{i=1}^{n}\sum_{j=1}^{n} x_j/x_i$, it follows that

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{x_j}{x_i}\left(a_{ij} - \frac{x_i}{x_j}\right)^2 &= \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{x_j}{x_i}\left(a_{ij}^2 - 2a_{ij}\frac{x_i}{x_j} + \left(\frac{x_i}{x_j}\right)^2\right) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_{ij}^2\frac{x_j}{x_i} - 2a_{ij} + \frac{x_i}{x_j}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_{ij}^2 + 1\right)\frac{x_j}{x_i} - 2\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}. \quad \square\end{aligned}$$

Lemma 4.2 *Suppose that A is an $n \times n$ matrix, then the optimization problem (4.2) has a unique optimal solution up to a positive scalar multiple, that is equal to an optimal solution of $\min_{\boldsymbol{x}>0}\sum_{j=1}^{n}\sum_{i=1}^{n}\left(a_{ij}^2 + 1\right)\frac{x_j}{x_i}$. If the optimal value of (4.2) is 0, then $a_{ij} = x_i/x_j$ for all $i, j = 1, \ldots, n$.*

Proof: Lemma 4.1 implies that (4.2) is equivalent to $\min_{\boldsymbol{x}>0}\sum_{j=1}^{n}\sum_{i=1}^{n}\left(a_{ij}^2 + 1\right)\frac{x_j}{x_i}$.

Let $\boldsymbol{x}$ be a positive vector, then we have $(x_j/x_i)(a_{ij} - x_i/x_j)^2 \geq 0$, where the equality is valid if and only if $a_{ij} = x_i/x_j$. Therefore, $Q(\boldsymbol{x}) = 0$ is equivalent to $a_{ij} = x_i/x_j$ for all $i, j = 1, \ldots, n$. $\square$

Lemma 4.2 implies that the minimum χ square estimate for the reciprocal matrix A is unique up to a scalar multiple and it is completely solved by applying the scaling algorithms to the matrix $\left[a_{ij}^2 + 1\right]$. Furthermore, the minimum χ square estimate derived from (4.2) guarantees following two properties that are desirable from the prioritization point of view:

Theorem 4.1 (Row dominance) *Suppose that A is a nonnegative matrix and that there is a pair of (l,k) such that $a_{lj} \geq a_{kj}$ and $a_{jl} \leq a_{jk}$ for all $j = 1,\ldots,n$, then an optimal solution $\boldsymbol{x}^*$ of (4.2) satisfies $x_l^* \geq x_k^*$. Moreover, if there is an index h such that $a_{lh} > a_{kh}$ or $a_{hl} < a_{hk}$, then $x_l^* > x_k^*$.*

Proof: Let $b_{ij} = a_{ij}^2 + 1$ for all $i,j = 1,\ldots,n$. Without loss of generality, suppose that $a_{1j} \geq a_{2j}$ for all $j = 1,\ldots,n$ and $a_{i1} \leq a_{i2}$ for all $i = 1,\ldots,n$, then we have

$$b_{1j} \geq b_{2j} \quad \text{for all } j = 1,\ldots,n \ \text{ and } \ b_{i1} \leq b_{i2} \quad \text{for all } i = 1,\ldots,n. \tag{4.4}$$

Let $\boldsymbol{x}^*$ be an optimal solution of (4.2), then it follows from Lemma 4.2 and Lemma 2.1 that

$$\sum_{j=1}^n b_{1j}\frac{x_j^*}{x_1^*} = \sum_{i=1}^n b_{i1}\frac{x_1^*}{x_i^*} \text{ and } \sum_{j=1}^n b_{2j}\frac{x_j^*}{x_2^*} = \sum_{i=1}^n b_{i2}\frac{x_2^*}{x_i^*}. \tag{4.5}$$

It follows from (4.5), (4.4) and $\boldsymbol{x}^* > 0$ that

$$\frac{(x_1^*)^2}{(x_2^*)^2} = \frac{\left(\dfrac{\sum_{j=1}^n b_{1j}x_j^*}{\sum_{i=1}^n b_{i1}x_i^*}\right)}{\left(\dfrac{\sum_{j=1}^n b_{2j}x_j^*}{\sum_{i=1}^n b_{i2}x_i^*}\right)} = \frac{\left(\sum_{j=1}^n b_{1j}x_j^*\right)\cdot\left(\sum_{i=1}^n b_{i2}x_i^*\right)}{\left(\sum_{j=1}^n b_{2j}x_j^*\right)\cdot\left(\sum_{i=1}^n b_{i1}x_i^*\right)} \geq 1\cdot 1 = 1, \tag{4.6}$$

implying that $x_1^* \geq x_2^*$.

If $a_{1h} > a_{2h}$ then, we have $b_{1h} > b_{2h}$ and hence, $\sum_{j=1}^n b_{1j}x_j^* > \sum_{j=1}^n b_{2j}x_j^*$. This means from (4.6) that $x_1^* > x_2^*$. □

Both the minimum χ square method and the EM satisfy row dominance, (this result is also proved in [12] by using the fact that A is reciprocal), however the EM does not satisfy the right-left symmetry that is also desirable on psychological grounds [8, 14]. The minimum χ square method satisfies right-left symmetry as follows:

Theorem 4.2 (Right-Left Symmetry) *Suppose that A is an $n \times n$ matrix. Let $\boldsymbol{x}^*$ be an optimal solution of (4.2) and let $\boldsymbol{y}^*$ be an optimal solution of*

$$\min_{\boldsymbol{y}>0} \sum_{i=1}^n \sum_{j=1}^n \left(a_{ji} - \frac{y_i}{y_j}\right)^2 \frac{y_j}{y_i}, \tag{4.7}$$

then there exists a constant $c > 0$ such that $x_i^* = c/y_i^*$ for all $i = 1,2,\ldots,n$.

Proof: It follows from Lemma 4.2 that (4.7) is equivalent to

$$\min_{\boldsymbol{y}>0} \sum_{i=1}^n \sum_{j=1}^n \left(a_{ji}^2 + 1\right) \frac{y_j}{y_i}. \tag{4.8}$$

Let $b_{ij} = a_{ij}^2 + 1$ and $x_i = 1/y_i$, then we have

$$\begin{aligned} \min_{\boldsymbol{y}>0} \sum_{i=1}^n \sum_{j=1}^n \left(a_{ji}^2 + 1\right) \frac{y_j}{y_i} &= \min_{\boldsymbol{y}>0} \sum_{i=1}^n \sum_{j=1}^n b_{ji}\frac{y_j}{y_i} = \min_{\boldsymbol{y}>0} \sum_{j=1}^n \sum_{i=1}^n b_{ji}\frac{y_j}{y_i} = \min_{\boldsymbol{y}>0} \sum_{i=1}^n \sum_{j=1}^n b_{ij}\frac{y_i}{y_j} \\ &= \min_{\boldsymbol{x}>0} \sum_{i=1}^n \sum_{j=1}^n b_{ij}\frac{x_j}{x_i} = \min_{\boldsymbol{x}>0} \sum_{i=1}^n \sum_{j=1}^n \left(a_{ij}^2 + 1\right) \frac{x_j}{x_i}. \end{aligned}$$

Therefore, it follows from Lemma 4.2 that there exists a positive number c such that $x_i^* = c/y_i^*$ for all $i = 1,\ldots,n$. □

Positivity of all entries of the matrix B is a key premise for both two proofs of Theorem 4.1 and Theorem 4.2. In fact, Theorem 4.1 holds for any positive matrix B satisfying (4.4) and $b_{1h} > b_{2h}$ for some $h \in \{1, \ldots, n\}$. Theorem 4.2 holds for any positive matrix B. Therefore, both theorems do not request an assumption of reciprocity of the given matrix A.

Saaty recommends that pairwise comparison value a_{ij} is chosen among $\{1/9, 1/7, 1/5, 1/3, 1, 3, 5, 7, 9\}$, that is called the scale 1-9. Takahashi [28] and Jensen [13] independently simplify the scale 1-9 into $\{1/\alpha, \alpha\}$. Here, α is a coding parameter which is greater than 1. The simplest AHP, say Binary AHP, has any pairwise comparison value $a_{ij} \in \{1/\alpha, \alpha\}$, whose simplification leads to an analytical priority weights [7, 28]. See Nishizawa [19, 20] for recent extensions of Binary AHP.

For the given coding parameter α, the pairwise comparison matrix of Binary AHP is denoted by $A(\alpha)$, whose off-diagonal element is α or $1/\alpha$. The matrix $A(\alpha)$ is also reciprocal. In AHP with the 1-9 scale, the minimum χ square estimate for the pairwise comparison matrix A is a solution of the matrix balancing problem for the matrix $\left[a_{ij}^2 + 1\right] (\neq A)$. However, the minimum χ square estimate for $A(\alpha)$ in Binary AHP is also a solution of the matrix balancing problem for $A(\alpha)$ as follows:

Theorem 4.3 *Choose any $\alpha \geq 1$ and let $A = A(\alpha)$, then an optimal solution $\boldsymbol{x}^*$ of (4.2) coincides with that of (2.3) and $Q(\boldsymbol{x}^*) = (\alpha + 1/\alpha)\left(f(\boldsymbol{x}^*) - n^2\right)$.*

Proof: Let $K^+ = \{(i,j) \,|\, a_{ij} = \alpha\}$ and $K^- = \{(i,j) \,|\, a_{ij} = 1/\alpha\}$, then $(i,j) \in K^+$ if and only if $(j,i) \in K^-$. Therefore,

$$\begin{aligned}
\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\frac{x_j}{x_i} &= \sum_{(i,j)\in K^+} a_{ij}\frac{x_j}{x_i} + \sum_{(i,j)\in K^-} a_{ij}\frac{x_j}{x_i} + \sum_{i=1}^{n} a_{ii}\frac{x_i}{x_i} = \sum_{(i,j)\in K^+} \alpha\frac{x_j}{x_i} + \sum_{(i,j)\in K^-} \frac{1}{\alpha}\frac{x_j}{x_i} + n \\
&= \sum_{(i,j)\in K^+} \alpha\frac{x_j}{x_i} + \sum_{(i,j)\in K^+} \frac{1}{\alpha}\frac{x_i}{x_j} + n \\
&= \sum_{(i,j)\in K^+} \left(\alpha\frac{x_j}{x_i} + \frac{1}{\alpha}\frac{x_i}{x_j}\right) + n = \sum_{(i,j)\in K^+} \frac{\alpha^2 x_j^2 + x_i^2}{\alpha x_i x_j} + n. \qquad (4.9)
\end{aligned}$$

In the same fashion, we have

$$\begin{aligned}
\sum_{i=1}^{n}\sum_{j=1}^{n} \left(a_{ij}^2 + 1\right)\frac{x_j}{x_i} &= \sum_{(i,j)\in K^+} \left(\alpha^2 + 1\right)\frac{x_j}{x_i} + \sum_{(i,j)\in K^-} \left(\frac{1}{\alpha^2} + 1\right)\frac{x_j}{x_i} + 2n \\
&= \sum_{(i,j)\in K^+} \left(\alpha^2 + 1\right)\frac{x_j}{x_i} + \sum_{(i,j)\in K^+} \left(\frac{\alpha^2 + 1}{\alpha^2}\right)\frac{x_i}{x_j} + 2n \\
&= \left(\alpha^2 + 1\right)\sum_{(i,j)\in K^+} \left(\frac{x_j}{x_i} + \frac{1}{\alpha^2}\frac{x_i}{x_j}\right) + 2n \\
&= \left(\alpha^2 + 1\right)\sum_{(i,j)\in K^+} \frac{\alpha^2 x_j^2 + x_i^2}{\alpha^2 x_i x_j} + 2n \\
&= \frac{(\alpha^2 + 1)}{\alpha}\sum_{(i,j)\in K^+} \frac{\alpha^2 x_j^2 + x_i^2}{\alpha x_i x_j} + 2n. \qquad (4.10)
\end{aligned}$$

It follows from Lemma 4.2, (4.9) and (4.10) that both optimization problems (4.2) and (2.3) are reduced to $\min_{\boldsymbol{x}>0} \sum_{(i,j)\in K^+} \left(\alpha^2 x_j^2 + x_i^2\right) / (\alpha x_i x_j)$. Since $\sum_{(i,j)\in K^+} \left(\alpha^2 x_j^2 + x_i^2\right) / (\alpha x_i x_j)$ $= \sum_{i=1}^{n} \sum_{j\neq i} a_{ij} x_j / x_i$, it follows from Lemma 2.1 that both optimization problems (4.2) and

(2.3) have the same optimal solution as $\min_{\boldsymbol{x}>0} \sum_{i=1}^n \sum_{j\neq i} a_{ij}x_j/x_i$. Let $\boldsymbol{x}^*$ be an optimal solution of (4.2), then it follows from $\sum_{i=1}^n \sum_{j=1}^n a_{ij} = (\alpha + 1/\alpha)(n-1)n/2 + n$ that

$$\begin{aligned} Q(\boldsymbol{x}^*) &= (\alpha+1/\alpha)(f(\boldsymbol{x}^*)-n) + 2n - 2\sum_{i=1}^n\sum_{j=1}^n a_{ij} \\ &= (\alpha+1/\alpha)(f(\boldsymbol{x}^*)-n) + 2n - 2\{(\alpha+1/\alpha)(n-1)n/2+n\} \\ &= (\alpha+1/\alpha)(f(\boldsymbol{x}^*)-n) - (\alpha+1/\alpha)(n-1)n = (\alpha+1/\alpha)\left(f(\boldsymbol{x}^*)-n^2\right). \end{aligned}$$ □

Sekitani and Yamaki [26] has an interpretation of the EM in AHP with the scale 1-9 by using the ith entry of $\boldsymbol{r}(\boldsymbol{x})$, i.e.,

$$\sum_{j=1}^n a_{ij}\frac{x_j}{x_i} = \sum_{j\neq i} a_{ij}\frac{x_j}{x_i} + a_{ii} = \sum_{j\neq i} a_{ij}\frac{x_j}{x_i} + 1. \tag{4.11}$$

In (4.11), x_i is called the self-evaluation value of alternative i and $\sum_{j\neq i}^n a_{ij}x_j$ is called the external evaluation value of alternative i. Then, $\sum_{j\neq i} a_{ij}x_j/x_i$ means the gap between the self-evaluation value and the external evaluation value with respect to alternative i, which is called the over-estimation ratio of alternative i. Therefore, each entry of $\boldsymbol{r}(\boldsymbol{x})$ is the corresponding over-estimation ratio plus 1. From Lemma 2.4, the EM is to solve $\min_{\boldsymbol{x}>0} \|\boldsymbol{r}(\boldsymbol{x})\|_\infty$. Let $\boldsymbol{e}$ be all one vector, then each entry of $\boldsymbol{r}(\boldsymbol{x}) - \boldsymbol{e}$ is the corresponding over-estimation ratio. Since $\min_{\boldsymbol{x}>0} \|\boldsymbol{r}(\boldsymbol{x}) - \boldsymbol{e}\|_\infty$ is equivalent to $\min_{\boldsymbol{x}>0} \|\boldsymbol{r}(\boldsymbol{x})\|_\infty$, the EM is to minimize the largest over-estimation ratio [26]. Theorem 4.3 implies that the minimum χ square method is equivalent to $\min_{\boldsymbol{x}>0} \|\boldsymbol{r}(\boldsymbol{x})\|_1$. Since $\min_{\boldsymbol{x}>0} \|\boldsymbol{r}(\boldsymbol{x})\|_1$ is equivalent to $\min_{\boldsymbol{x}>0} \|\boldsymbol{r}(\boldsymbol{x}) - \boldsymbol{e}\|_1$, the minimum χ square method for $A(\alpha)$ in Binary AHP is to minimize the sum of over-estimation ratio.

Binary AHP has the following relation between the EM and the minimum χ square method:

Corollary 4.1 *Choose any $\alpha \geq 1$. Let $A = A(\alpha)$ and let $\lambda_{\max}$ be the principal eigenvalue of $A(\alpha)$, then $n(\alpha+1/\alpha)(\lambda_{\max} - n) \geq Q(\boldsymbol{x}^*)$, where $\boldsymbol{x}^*$ is an optimal solution of (4.2).*

Proof: Let $\boldsymbol{x}^*$ be an optimal solution of (4.2), then it follows from Theorem 4.3 and (2.12) that

$$Q(\boldsymbol{x}^*) = (\alpha+1/\alpha)\left(\|\boldsymbol{r}(\boldsymbol{x}^*)\|_1 - n^2\right) \leq n(\alpha+1/\alpha)(\lambda_{\max} - n).$$ □

5. Comparisons with Eigenvalue Method on Binary AHP

Takahashi [28] and Genest *et al.* [7] independently show surprisingly elegant properties of priority weights of a pairwise matrix $A(\alpha)$ by the EM. By introducing such their properties of priority weights, this section shows sufficient conditions of equivalence between the minimum χ square method and the EM.

Genest *et al.* [7] define two classes of $A(\alpha)$ as follows: $A(\alpha)$ is called maximally intransitive if $\sum_{j=1}^n a_{1j} = \cdots = \sum_{j=1}^n a_{nj}$. $A(\alpha)$ is called ordinally transitive if $a_{ij} > 1$ and $a_{jk} > 1$ imply $a_{ik} > 1$. A pairwise matrix $A(\alpha)$ is ordinally transitive if and only if there is a permutation matrix P such that any upper off-diagonal entry of $P^\top A(\alpha)P$ is α (see, e.g., [28]). Hereafter, without loss of generality, we assume that an ordinally transitive matrix $A(\alpha)$ satisfies $a_{ij} = \alpha$ for all $i < j$.

The following lemmas shows that a pairwise comparison matrix $A(\alpha)$ has a left principal eigenvector and a right one that are equivalent up to a positive scalar multiple if $A(\alpha)$ is ordinally transitive or maximally intransitive.

Lemma 5.1 *Suppose that $A(\alpha)$ is ordinally transitive, then $A(\alpha)$ has a right principal eigenvector $\left(\alpha^{-1/n}, \alpha^{-3/n}, \cdots, \alpha^{-(2n-1)/n}\right)^\top$ and a left one $\left(\alpha^{1/n}, \alpha^{3/n}, \cdots, \alpha^{(2n-1)/n}\right)$.*
Proof: See Theorem 4 of [28] and Proposition 4.3 of [7]. □

Lemma 5.2 *Suppose that $A(\alpha)$ is maximally intransitive, then $A(\alpha)$ has a right principal eigenvector $(1, 1, \cdots, 1)^\top$ and a left principal eigenvector $(1, 1, \cdots, 1)$.*

Proof: Suppose that $A(\alpha)$ is maximally intransitive, then A has a right principal eigenvector $(1, 1, \cdots, 1)^\top$. Since $A(\alpha)$ is reciprocal, we have $\{j \mid a_{ij} = \alpha\} = \{j \mid a_{ji} = 1/\alpha\}$ and $\{j \mid a_{ij} = 1/\alpha\} = \{j \mid a_{ji} = \alpha\}$. This means that $\sum_{i=1}^n a_{i1} = \cdots = \sum_{i=1}^n a_{in}$ and hence, A has a left principal eigenvector $(1, 1, \cdots, 1)$. □

The equivalence between a left and a right eigenvectors implies that the minimum χ square method is equivalent to the EM as follows:
Lemma 5.3 *Suppose that $A(\alpha)$ is ordinally transitive, then an optimal solution of (4.2) for $A = A(\alpha)$ is equal to a positive scalar multiple of $\left(\alpha^{-1/n}, \alpha^{-3/n}, \cdots, \alpha^{-(2n-1)/n}\right)^\top$.*
Proof: Let $\boldsymbol{x}^\#$ be a right principal eigenvector of $A(\alpha)$ that is ordinally transitive, then Lemma 5.1 and Corollary 2.1 imply that $\boldsymbol{x}^\#$ is an optimal solution of (4.2). Therefore, it follows from Theorem 4.3 that $\boldsymbol{x}^\#$ is also an optimal solution of (2.3). □

Lemma 5.4 *Suppose that $A(\alpha)$ is maximally intransitive, then an optimal solution of (4.2) is equal to a positive scalar multiple of $(1, 1, \cdots, 1)^\top$.*

Proof: In the same way as the proof of Lemma 5.3, we can prove it. □

When a pairwise comparison matrix $A(\alpha)$ is of size $n \leq 3$, the minimum χ square method coincides with the EM as follows:
Lemma 5.5 *Suppose that a pairwise comparison matrix $A(\alpha)$ is of size $n \leq 3$, then a right principal eigenvector of $A(\alpha)$ and an optimal solution of (4.2) coincide for all $\alpha \geq 1$.*

Proof: If $A(\alpha)$ is of size $n = 2$, then it suffices to consider $A(\alpha) = \begin{bmatrix} 1 & \alpha \\ 1/\alpha & 1 \end{bmatrix}$. A left principal eigenvector of $A(\alpha)$ is equal to a positive scalar multiple of $[1/\alpha, 1]$ and a right principal eigenvector of $A(\alpha)$ is $[\alpha, 1]^\top$. Therefore, a right principal eigenvector of $A(\alpha)$ and an optimal solution of (4.2) coincide for all $\alpha \geq 1$.

The matrix $A(\alpha)$ with size $n = 3$ is ordinally transitive or maximally intransitive. Hence, it follows from Lemma 5.3 and Lemma 5.4 that a right principal eigenvector of $A(\alpha)$ with size $n = 3$ and an optimal solution of (4.2) coincide for all $\alpha \geq 1$. □

For a general size n of $A(\alpha)$, we have sufficient conditions of equivalence between the minimum χ square method and the EM as follows:
Theorem 5.1 *A principal right eigenvector of $A(\alpha)$ and an optimal solution of (4.2) coincide for all $\alpha \geq 1$ whenever*
1. *$A(\alpha)$ is ordinally transitive;*
2. *$A(\alpha)$ is maximally intransitive; or*
3. *$A(\alpha)$ is of size $n \leq 3$.*

Proof: This immediately follows from Lemma 5.3, Lemma 5.4 and Lemma 5.5. □

Hereafter, we compare the ranking from priority weights by the minimum χ square method with the ranking from priority weights by the EM. All priority weights $(x_1, \cdots, w_n)$

is ordered according each value. That is, the rank of i is higher than that of j if $x_i > x_j$, and i and j have the same rank if $x_i = x_j$.

Corollary 5.1 *Two rankings deriving from a principal right eigenvector of $A(\alpha)$ and an optimal solution of (4.2) coincide for all $\alpha > 1$ whenever*

1. *$A(\alpha)$ is ordinally transitive;*
2. *$A(\alpha)$ is maximally intransitive; or*
3. *$A(\alpha)$ is of size $n \le 4$ and there is no permutation matrix P such that $P^\top A(\alpha)P =$*

$$\begin{bmatrix} 1 & \alpha & \alpha & 1/\alpha \\ 1/\alpha & 1 & \alpha & \alpha \\ 1/\alpha & 1/\alpha & 1 & \alpha \\ \alpha & 1/\alpha & 1/\alpha & 1 \end{bmatrix}. \tag{5.1}$$

Proof: Let $\boldsymbol{x}^*$ be priority weights by the minimum χ square method and let $\boldsymbol{x}^\#$ be priority weights by the EM. If $A(\alpha)$ is of size $n \le 3$, ordinally transitive, or maximally intransitive, then it follows from Theorem 5.1 that the ranking from $\boldsymbol{x}^*$ is equal to that from $\boldsymbol{x}^\#$.

For any 4×4 matrix $A(\alpha)$n there is a permutation matrix P such that $P^\top A(\alpha)P$ is one among the following four different matrices:

$$A^0 = \begin{bmatrix} 1 & \alpha & \alpha & \alpha \\ 1/\alpha & 1 & \alpha & \alpha \\ 1/\alpha & 1/\alpha & 1 & \alpha \\ 1/\alpha & 1/\alpha & 1/\alpha & 1 \end{bmatrix}, \qquad A^1 = \begin{bmatrix} 1 & \alpha & \alpha & \alpha \\ 1/\alpha & 1 & \alpha & 1/\alpha \\ 1/\alpha & 1/\alpha & 1 & \alpha \\ 1/\alpha & \alpha & 1/\alpha & 1 \end{bmatrix},$$

$$A^2 = \begin{bmatrix} 1 & 1/\alpha & 1/\alpha & 1/\alpha \\ \alpha & 1 & 1/\alpha & \alpha \\ \alpha & \alpha & 1 & 1/\alpha \\ \alpha & 1/\alpha & \alpha & 1 \end{bmatrix} \text{ and } A^3 = \begin{bmatrix} 1 & \alpha & \alpha & 1/\alpha \\ 1/\alpha & 1 & \alpha & \alpha \\ 1/\alpha & 1/\alpha & 1 & \alpha \\ \alpha & 1/\alpha & 1/\alpha & 1 \end{bmatrix}.$$

Note that every permutation matrix P satisfies $P^\top A^i P \ne A^k$ for all $k \ne i$. Since A^3 is the matrix (5.1), the proof for $A(\alpha)$ of size $n = 4$ is sufficient to discuss rankings for three matrices A^0, A^1 and A^2.

Consider the case where $A(\alpha)$ is reduced to A^0. The matrix A^0 is ordinally transitive and it follows from Theorem 5.1 that $\boldsymbol{x}^\#$ and $\boldsymbol{x}^*$ provide the same ranking.

Consider that $A(\alpha)$ is reduced to A^1. Genest [7] shows that the matrix A^1 has a right principal eigenvector $\boldsymbol{x}^\# = (3\alpha\beta, 1, 1, 1)^\top$, where $\beta = \left[(\alpha + 1/\alpha)/2 + \sqrt{(\alpha + 1/\alpha)^2/4 + 3}\right]^{-1}$. Since $3\alpha\beta > 1$ for all $\alpha > 1$, the ranking from the right principal eigenvector $\boldsymbol{x}^\#$ is given as

$$x_1^\# > 1 = x_2^\# = x_3^\# = x_4^\#. \tag{5.2}$$

The matrix balancing problem (2.3) for A^1 has an optimal solution $\boldsymbol{x}^* = (\alpha, 1, 1, 1)^\top$. In fact, the optimal solution $\boldsymbol{x}^* = (\alpha, 1, 1, 1)^\top$ satisfies

$$\begin{bmatrix} \frac{1}{x_1^*} & 0 & 0 & 0 \\ 0 & \frac{1}{x_2^*} & 0 & 0 \\ 0 & 0 & \frac{1}{x_3^*} & 0 \\ 0 & 0 & 0 & \frac{1}{x_4^*} \end{bmatrix} \begin{bmatrix} 1 & \alpha & \alpha & \alpha \\ \frac{1}{\alpha} & 1 & \alpha & \frac{1}{\alpha} \\ \frac{1}{\alpha} & \frac{1}{\alpha} & 1 & \alpha \\ \frac{1}{\alpha} & \alpha & \frac{1}{\alpha} & 1 \end{bmatrix} \begin{bmatrix} x_1^* & 0 & 0 & 0 \\ 0 & x_2^* & 0 & 0 \\ 0 & 0 & x_3^* & 0 \\ 0 & 0 & 0 & x_4^* \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & \alpha & \frac{1}{\alpha} \\ 1 & \frac{1}{\alpha} & 1 & \alpha \\ 1 & \alpha & \frac{1}{\alpha} & 1 \end{bmatrix}, \tag{5.3}$$

that is sum-symmetry. It follows from Theorem 4.3 that $\boldsymbol{x}^*$ is also an optimal solution of (4.2) for A^1. Therefore, the ranking from $\boldsymbol{x}^*$ is given as

$$x_1^* = \alpha > 1 = x_2^* = x_3^* = x_4^*, \tag{5.4}$$

for all $\alpha > 1$. The ranking (5.4) from the optimal solution $\boldsymbol{x}^*$ of (4.2) coincides with (5.2) from the right principal eigenvector $\boldsymbol{x}^\#$ of A^1.

Consider the case of A^2, then $(A^1)^\top = A^2$ and a right principal eigenvector of A^2 is a positive scalar multiple of $\boldsymbol{x}^\# = (3\beta/\alpha, 1, 1, 1)^\top$, where $\beta = \left[(\alpha + 1/\alpha)/2 + \sqrt{(\alpha + 1/\alpha)^2/4 + 3}\right]^{-1}$. Genest *et al.* [7] shows that the principal eigenvalue $\lambda_{\max}$ of A^2 is $\beta^{-1}+1$. Since $Q(\boldsymbol{x}) \geq 0$ for any positive vector $\boldsymbol{x}$, it follows from Corollary 4.1 that $4 = n \leq \lambda_{\max} = \beta^{-1} + 1$, implying that $3\beta \leq 1$ and $3\beta/\alpha < 1$ for all $\alpha > 1$. The ranking from $\boldsymbol{x}^\#$ is given as

$$x_1^\# < 1 = x_2^\# = x_3^\# = x_4^\#. \tag{5.5}$$

Let $\boldsymbol{x}^*$ be an optimal solution of (4.2) for A^2, then it follows from $(A^1)^\top = A^2$ and Theorem 4.2 that $\boldsymbol{x}^* = (1/\alpha, 1, 1, 1)^\top$ for all $\alpha > 1$. The ranking from $\boldsymbol{x}^*$ is

$$x_1^* < 1 = x_2^* = x_3^* = x_4^*, \tag{5.6}$$

which is the same order as that from (5.5). □

As stated above, by using a permutation of a pairwise comparison matrix $A(\alpha)$, $A(\alpha)$ with $n = 4$ is classified into four equivalent classes. Whenever $A(\alpha)$ with $n = 4$ is not equivalent to (5.1), the minimum χ square method provides the same ranking as the EM. For the equivalent class to (5.1), the ranking by the minimum χ square method is different from the ranking by the EM as follows:

Corollary 5.2 *Suppose that $A(\alpha)$ is (5.1). Let $\boldsymbol{x}^\#$ be a right principal eigenvector of $A(\alpha)$ and let $\boldsymbol{x}^*$ be an optimal solution of (4.2). For any $\alpha > 1$ we have*

$$x_1^\# > x_2^\# > x_4^\# > x_3^\# \text{ and} \tag{5.7}$$

$$x_2^* > x_1^* > x_4^* > x_3^*. \tag{5.8}$$

Proof: See Proposition 4.4 of [8] for the proof of (5.7). The proof of (5.8) is given in Appendix. □

Consider a 5×5 matrix similar to (5.1) as follows[7]:

$$\begin{bmatrix} 1 & \alpha & \alpha & \alpha & 1/\alpha \\ 1/\alpha & 1 & \alpha & \alpha & \alpha \\ 1/\alpha & 1/\alpha & 1 & \alpha & \alpha \\ 1/\alpha & 1/\alpha & 1/\alpha & 1 & \alpha \\ \alpha & 1/\alpha & 1/\alpha & 1/\alpha & 1 \end{bmatrix}. \tag{5.9}$$

Genest *et al.* [7] report from some numerical experiments that a ranking by the EM for (5.9) depends on choice of coding parameter of α. For a small values of $\alpha > 1$, a right principal eigenvector $\boldsymbol{x}^\#$ of (5.9) satisfies

$$x_1^\# > x_2^\# > x_3^\# > x_5^\# > x_4^\#. \tag{5.10}$$

For any $\alpha > 3.7$, a right principal eigenvector of (5.9) usually satisfies

$$x_1^\# > x_2^\# > x_5^\# > x_3^\# > x_4^\#. \tag{5.11}$$

Checking rankings from the minimum χ estimate of the matrix (5.9) with $\alpha = 2, 3, \ldots, 100$, we confirm numerically that an optimal solution $\boldsymbol{x}^*$ of (4.2) satisfies

$$x_2^* > x_1^* > x_3^* > x_5^* > x_4^* \tag{5.12}$$

Table 1: Comparison of the EM with minimum χ square method on Rank reversal

n	Number of equivalent classes	Eigenvector method: number of rank reversals	Eigenvector method: percentage of rank reversals	χ square method: number of rank reversals	χ square method: percentage of rank reversals
1	1	0	0%	0	0%
2	1	0	0%	0	0%
3	2	0	0%	0	0%
4	4	0	0%	0	0%
5	12	1	8.3%	0	0%
6	56	10	17.9%	2	3.8%
7	456	113	24.8%	20	4.4%
8	6880	2425	35.2%	524	7.6%
9	191536	85805	44.8%	21218	11.1%

for all $\alpha = 2, 3, \ldots, 100$ and hence, the ranking from $\boldsymbol{x}^*$ of (5.9) may be independent of the choice of coding parameter of α. Genest *et al.* [7] document that such rank reversibility by the EM increases dramatically with size $n = 6, 7, 8, 9$ of $A(\alpha)$.

Comparing rank reversibility by the EM with that by the minimum χ square method, we carry out numerical experiments to measure the rank reversals frequency generated by the two methods. For $\alpha = 2, 3, 5, 7, 9$, let $\boldsymbol{x}^*(\alpha)$ be an optimal solution of (4.2) and let $\boldsymbol{x}^{\#}(\alpha)$ be an right principal eigenvector of $A(\alpha)$. For a sufficiently small positive number $\epsilon > 0$, we define

$$K^*_\epsilon = \left\{(i,j) \;\Big|\; \left|x^*_i(2) - x^*_j(2)\right| \geq \epsilon\right\}. \tag{5.13}$$

We say that the rank reversal occurs by the minimum χ square method, if and only if there exists an $\alpha \in \{2, 3, 5, 7, 9\}$ such that an optimal solution $\boldsymbol{x}(\alpha)^*$ of (4.2) violates at least one of

$$\begin{aligned} \left|x^*_i(\alpha) - x^*_j(\alpha)\right| \geq \epsilon \text{ and } \left(x^*_i(2) - x^*_j(2)\right)\left(x^*_i(\alpha) - x^*_j(\alpha)\right) &> 0 \quad \text{for all } (i,j) \in K^*_\epsilon, \\ \left|x^*_i(\alpha) - x^*_j(\alpha)\right| &< \epsilon \quad \text{for all } (i,j) \notin K^*_\epsilon. \end{aligned} \tag{5.14}$$

In the similar way to (5.13) of the minimum χ square estimate, we define $K^{\#}_\epsilon$ for a right principal eigenvector of $\boldsymbol{x}^{\#}(2)$ and consider two conditions for $\boldsymbol{x}^{\#}(\alpha)$ that are similar to (5.14). Then we say that the rank reversal occurs by the EM, if and only if there exists an $\alpha \in \{2, 3, 5, 7, 9\}$ such that $\boldsymbol{x}^{\#}(\alpha)$ violates at least one of two conditions for $\boldsymbol{x}^{\#}(\alpha)$.

As stated in the proof of Corollary 5.1, a set of $A(\alpha)$ with size $n = 4$ is classified into four equivalent classes. The size n of $A(\alpha)$ and the number of the equivalent classes on the whole set of $A(\alpha)$ with the corresponding size n are listed in the first and second column of Table 1, respectively.

Let $\epsilon = 10^{-7}$ of K^*_ϵ and $K^{\#}_\epsilon$ and choose one matrix $A(\alpha)$ from each equivalent class, we check rank reversals occurrence for the minimum χ square method and the EM. The third column of Table 1 indicates the number of the equivalent classes where rank reversals occurs by the EM. The forth column is given the proportion of the equivalence classes including rank reversals by the EM. In the same manner, the resulting rank reversals occurrence by the minimum χ square method are listed in the fifth and sixth columns of Table 1.

For the EM and the minimum χ square method, the two proportions of the equivalent classes including the rank reversals increase with size n. Especially, the rank reversals with size $n = 9$ occurs by the EM at the almost 50% probability. The rank reversals occurrence by the EM is at least 4 times as much as that by the minimum χ square method.

It is well known in regression analysis [9] that the L_1-norm estimation is robust for outliers on the comparison with L_2-norm and L_∞-norm. The minimum χ square method is the minimum-norm point problem (2.9) whose norm is L_1 norm ($\|\cdot\|_1$) and the EM is the minimum-norm point problem (2.10) whose norm is L_∞-norm ($\|\cdot\|_\infty$). By analogy with the regression analysis, it is natural that rank preservation of the minimum χ square method is superior to that of the EM.

6. Conclusion

This research shows necessary and sufficient conditions (Theorem 2.1) of equivalence between the eigenvalue problem and the matrix balancing problem for a nonnegative matrix, by using a unified framework, the minimum-norm point problem, for the two problems. Furthermore, Theorem 3.1 shows global convergence of scaling algorithms for solving the matrix balancing problem. A variety of nonnegative matrix analyses appear frequently in social sciences, e.g., economics, management science and operations research, where Theorem 2.1 and Theorem 3.1 may be useful for modeling and simulation.

This study also illustrates a contribution of Theorem 2.1 and Theorem 3.1 by applying the minimum χ square method to AHP. Lemma 4.2 shows that the minimum χ square method (4.2) is to solve the matrix balancing problem. The minimum χ square method has the desirable properties of Theorem 4.1 and Theorem4.2, that is not satisfied with the EM. Furthermore, Theorem 4.1 and Theorem4.2 do not request a usual assumption of AHP that is reciprocal of a pairwise comparison matrix.

The simplest AHP, Binary AHP, can compare directly the EM with the minimum χ square method, by virtue of Theorem 4.3. Consequently, Theorem 5.1, Corollary 5.1, and Corollary 5.2 show how pairwise comparison matrix $A(\alpha)$ of Binary AHP is when the priority weights of the EM coincide with ones of the minimum χ square method. In the other case, i.e., when the priority weights of the EM and the minimum χ square method are not equivalent , dual norms incorporated into each method make a remarkable gap between their frequencies of rank reversals occurrence. The numerical experiments indicate that the rank reversals frequency of the EM is at least 4 times as much as that of the the minimum χ square method.

The minimum χ square method of AHP has a potential merit that is to incorporate a data reliability $\beta_{ij}(=\beta_{ji})$ for a pairwise comparison value a_{ij} into (4.2) as follows:

$$\min_{\boldsymbol{x}>0} \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} \frac{x_j}{x_i} \left(a_{ij} - \frac{x_i}{x_j} \right)^2 . \tag{6.1}$$

This optimization problem (6.1) is reduced to

$$\min_{\boldsymbol{x}>0} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(a_{ij}^2 + 1 \right) \beta_{ij} \frac{x_j}{x_i}, \tag{6.2}$$

which is equivalent to the matrix balancing problem for $\left[\left(a_{ij}^2 + 1 \right) \beta_{ij} \right]$. By using β_{ij}, the minimum χ square method is directly available to incomplete information case, where all pairwise comparisons are not carried out. Let $C = \{(i,j) \,|\, (i,j) \text{ are not compared}\}$ and $D = \{(i,j) \,|\, (i,j) \text{ are compared}\}$ and set $\beta_{ij} = 1$ for all $(i,j) \in D$ and $\beta_{ij} = 0$ for all

$(i, j) \in C$, then the minimum χ square method under the incomplete information is to solve the matrix balancing problem as follows:

$$\min_{\boldsymbol{x}>0} \sum_{(i,j)\in D} \left(a_{ij}^2 + 1\right) \frac{x_j}{x_i}.$$

AHP including multiple decision makers, group AHP, can be applied by the minimum χ square method. Suppose that L decision makers individually provide L pairwise comparison matrices $[a_{ij}^l]$ $(l = 1, \ldots, L)$, then the minimum χ square method (4.2) is straightly modified as follows:

$$\min_{\boldsymbol{x}>0} \sum_{l=1}^{L} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{x_j}{x_i} \left(a_{ij}^l - \frac{x_i}{x_j}\right)^2. \tag{6.3}$$

The modified problem (6.3) is equivalent to

$$\min_{\boldsymbol{x}>0} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\sum_{l=1}^{L} \left(a_{ij}^l\right)^2 + L\right) \frac{x_j}{x_i}. \tag{6.4}$$

Therefore, the minimum χ square method for group AHP is to solve the matrix balancing problem for $\left[\sum_{l=1}^{L} \left(a_{ij}^l\right)^2 + L\right]$.

One of future investigations in this study is to apply the minimum χ square method to real-world decision making problems such as merit-based personnel systems [31] and strategic decision of supply chain management [18]. Another of future investigations is to apply the minimum χ square method into parameter estimation of stochastic models, e.g., Bradley-Terry model [3], Huff model [11] and contingency analysis [6]. Their models often use maximum likelihood estimation, whose validity of the model is checked by a likelihood ratio test. The minimum χ square estimate gives the minimum χ square value that is immediately used in Pearson's χ square test of goodness of fit.

Acknowledgments

The third author is partly supported by *Grant-in-Aid for the Japan Society for the Promotion of Science*, No. 18510121. The author wishes to thank Prof. Yoshitsugu Yamamoto, Prof. Akiko Yoshise and Prof. Tomomi Matsui for stimulating discussions and constructive comments.

References

[1] A. Bachem and B. Korte: On the RAS-algorithm. *Computing*, **23** (1979), 189–198.

[2] R.B. Bapatn and T.E.S. Raghaven: *Nonnegative Matrices and Applications* (Cambridge University Press, Cambridge, 1997).

[3] R.A. Bradley and M.E. Terry: Rank analysis of incomplete block designs: the method of paired comparisons. *Biometrika*, **39** (1952), 324–345.

[4] L.M. Bregman: Proof of the convergence of Sheleikhovskii's method for a problem with transportation constraints. *USSR Computational Mathematics and Mathematical Physics*, **7** (1967), 191–204.

[5] B.C. Eaves, A.J. Hoffman, U.G. Rothblum and H. Schneider: Line-sum-symmetric scalings of square nonnegative matrices. *Mathematical Programming Study*, **25** (1985), 124–141.

[6] D. Friedlander: A technique for estimating a contingency table, given the marginal totals and some supplementary data. *Journal of the Royal Statistical Society:Series A*, **124** (1961), 412–420.
[7] C. Genest, F. Lapointe and S.W. Drury: On a proposal of Jensen for the analysis of ordinal pairwise preferences using Saaty's eigenvector scaling method. *Journal of Mathematical Psychology*, **37** (1993), 575–610.
[8] C. Genest and C.-É. M'lan: Deriving priorities from the Bradley-Terry model. *Mathematical and Computer Modelling*, **29** (1999), 87–102.
[9] R. Gonin and A.H. Money: *Nonlinear* L_p*-Norm Estimation* (Marcel Dekker, New York, 1989).
[10] T. Huang, W. Li and W. Sun: Estimates for bounds of some numerical characters of matrices. *Linear Algebra and its Applications*, **319** (2000), 137–145.
[11] D.L. Huff: A probabilistic analysis of shopping center trade area. *Land Economics*, **39** (1963), 81–90.
[12] R.E. Jensen: Comparisons of eigenvector, least squares, chi square and logarithmic least squares methods of scaling a reciprocal matrix. *Trinity University, USA, Working Paper*, **127** (1984).
[13] R.E. Jensen: Comparison of consensus methods for priority ranking problems. *Decision Sciences*, **17** (1986), 195–211.
[14] C.R. Johnson, W.B. Beine and T.J. Wang: Right-left asymmetry in an eigenvector ranking procedure. *Journal of Mathematical Psychology*, **19** (1979), 61–64.
[15] C.R. Johnson, J. Pitkin and D.P. Stanford: Line-sum symmetry via the DomEig algorithm. *Computational optimization and applications*, **17** (2000), 5–10.
[16] B. Kalantari, L. Khachiyan and A. Shokoufandeh: On the complexity of matrix balancing. *SIAM Journal on Matrix Analysis and Applications*, **18** (1997), 450–463.
[17] J. Kruithof: Telefoonverkeersrekening. *De Ingenieur*, **52** (1937) 15-25.
[18] T. Nakagawa and K. Sekitani: A use of analytic network process for supply chain management. *Asia Pacific Management Review*, **9** (2004), 783–800.
[19] K. Nishizawa: A consistency improving method in Binary AHP. *Journal of the Operations Research Society of Japan*, **35** (1995), 21–33.
[20] K. Nishizawa: A method to estimate results of unknown comparisons in Binary AHP. *Journal of the Operations Research Society of Japan*, **40** (1997), 105–121.
[21] B.N. Parlett and T.L. Landis: Methods for scaling to doubly stochastic form. *Linear Algebra and its Applications*, **48** (1982), 53–79.
[22] T.L. Saaty: *Analytic Hierarchy Process* (McGraw-Hill, New York, 1980).
[23] T.L. Saaty: Decision-making with the AHP: Why is the principal eigenvector necessary. *European Journal of Operational Research*, **145** (2003), 85–91.
[24] T.L. Saaty and L.G. Vargas: Inconsistency and rank preservation. *Journal of Mathematical Psychology*, **20** (1984), 347–354.
[25] M.H. Schneider and S.A. Zenios: A comparative study of algorithms for matrix balancing. *Operations Research*, **38** (1990), 439–455.
[26] K. Sekitani and N. Yamaki: A logical interpretation for the eigenvalue method in AHP. *Journal of the Operations Research Society of Japan* **42** (1999), 219–232.
[27] K. Sekitani and Y. Yamamoto: A recursive algorithm for finding the minimum norm point in a polytope and a pair of closest points in two polytopes. *Mathematical Programming*, **61** (1993), 233–249.

[28] I. Takahashi: AHP applied to binary and ternary comparisons. *Journal of the Operations Research Society of Japan*, **33** (1990), 199–206.

[29] A. Takayama: *Mathematical Economics* (Cambridge University Press, Cambridge, 1985).

[30] Q.H. Vuong and W. Wang: Minimum chi-square estimation and tests for model selection. *Journal of Econometrics*, **56** (1993), 141–168.

[31] N. Yamaki and K. Sekitani: A large-scale AHP including incomplete information and multiple evaluators and its applications to personnel administration evaluation system. *Journal of the Operations Research Society of Japan*, **42** (1999), 405–421.(In Japanese)

[32] F. Zahedi: The Analytic Hierarchy Process – a survey of the method and its applications. *Interfaces*, **16** (1986), 96–108.

[33] W.I. Zangwill: *Nonlinear Programming: A Unified Approach* (Prentice-Hall, New Jersey, 1969).

Appendix

Proof of Corollary 5.2

Since $\boldsymbol{x}^*$ is an optimal solution of (4.2), it follows from Theorem 4.3 that $\boldsymbol{x}^*$ is also an optimal solution of the problem (2.3) whose A is replaced with the matrix (5.1). Hence, it follows from Lemma 2.1 that the matrix

$$\begin{bmatrix} 1 & \alpha\frac{x_2^*}{x_1^*} & \alpha\frac{x_3^*}{x_1^*} & (1/\alpha)\frac{x_4^*}{x_1^*} \\ (1/\alpha)\frac{x_1^*}{x_2^*} & 1 & \alpha\frac{x_3^*}{x_2^*} & \alpha\frac{x_4^*}{x_2^*} \\ (1/\alpha)\frac{x_1^*}{x_3^*} & (1/\alpha)\frac{x_2^*}{x_3^*} & 1 & \alpha\frac{x_4^*}{x_3^*} \\ \alpha\frac{x_1^*}{x_4^*} & (1/\alpha)\frac{x_2^*}{x_4^*} & (1/\alpha)\frac{x_3^*}{x_4^*} & 1 \end{bmatrix} \tag{6.5}$$

is sum-symmetry. The sum-symmetry of (6.5) is equivalent to the following four equations:

$$(x_1^*)^2 = \frac{\alpha x_2^* + \alpha x_3^* + x_4^*/\alpha}{1/(\alpha x_2^*) + 1/(\alpha x_3^*) + \alpha/x_4^*}, \tag{6.6}$$

$$(x_2^*)^2 = \frac{x_1^*/\alpha + \alpha x_3^* + \alpha x_4^*}{\alpha/x_1^* + 1/(\alpha x_3^*) + 1/(\alpha x_4^*)}, \tag{6.7}$$

$$(x_3^*)^2 = \frac{x_1^*/\alpha + x_2^*/\alpha + \alpha x_4^*}{\alpha/x_1^* + \alpha/x_2^* + 1/(\alpha x_4^*)}, \tag{6.8}$$

$$(x_4^*)^2 = \frac{\alpha x_1^* + x_2^*/\alpha + x_3^*/\alpha}{1/(\alpha x_1^*) + \alpha/x_2^* + \alpha/x_3^*}. \tag{6.9}$$

We will show $x_1^* > x_4^*$. The proof is by contradiction. Assume $x_4^* \geq x_1^*$, then we have

$$1 \leq \left(\frac{x_4^*}{x_1^*}\right)^2 = \frac{x_1^*/x_2^* + 2/\alpha^2 + x_3^*/(\alpha^2 x_2^*) + x_1^*/x_3^* + x_2^*/(\alpha^2 x_3^*) + +\alpha^2 x_1^*/x_4^* + x_2^*/x_4^* + x_3^*/x_4^*}{x_2^*/x_1^* + x_3^*/x_1^* + x_4^*/(\alpha^2 x_1^*) + \alpha^2 + \alpha^2 x_3^*/x_2^* + x_4^*/x_2^* + \alpha^2 x_2^*/x_3^* + \alpha^2 + x_4^*/x_3^*}.$$

This implies from $x_4^*/x_1^* \geq 1$ and $\alpha > 1$ that

$$\begin{aligned} 0 \geq\ & x_2^*/x_1^* + x_3^*/x_1^* + x_4^*/(\alpha^2 x_1^*) + \alpha^2 + \alpha^2 x_3^*/x_2^* + x_4^*/x_2^* + \alpha^2 x_2^*/x_3^* + \alpha^2 + x_4^*/x_3^* \\ & - \left(x_1^*/x_2^* + 1/\alpha^2 + x_3^*/(\alpha^2 x_2^*) + x_1^*/x_3^* + x_2^*/(\alpha^2 x_3^*) + 1/\alpha^2 + \alpha^2 x_1^*/x_4^* + x_2^*/x_4^* + x_3^*/x_4^*\right) \\ =\ & (x_2^* + x_3^*)\,(1/x_1^* - 1/x_4^*) + x_4^*/(\alpha^2 x_1^*) + 1/\alpha^2 x_4^*/x_1^* - \alpha^2 x_1^*/x_4^* + 2\alpha^2 - 2/\alpha^2 \end{aligned}$$

$$\begin{aligned}
& +\alpha^2 \left(x_3^*/x_2^* + x_2^*/x_3^*\right) - 1/\alpha^2 \left(x_3^*/x_2^* + x_2^*/x_3^*\right) + x_4^* \left(1/x_2^* + 1/x_3^*\right) - x_1^* \left(1/x_2^* + 1/x_3^*\right) \\
= \; & \left(x_2^* + x_3^*\right)\left(1/x_1^* - 1/x_4^*\right) + \left(x_4^* - x_1^*\right)\left(1/x_2^* + 1/x_3^*\right) + x_4^*/(\alpha^2 x_1^*) - \alpha^2 x_1^*/x_4^* \\
& +2\left(\alpha^2 - 1/\alpha^2\right) + \left(\alpha^2 - 1/\alpha^2\right)\left(x_3^*/x_2^* + x_2^*/x_3^*\right) \\
\geq \; & \left(x_2^* + x_3^*\right)\left(1/x_1^* - 1/x_4^*\right) + \left(x_4^* - x_1^*\right)\left(1/x_2^* + 1/x_3^*\right) + 1/\alpha^2 - \alpha^2 + 2\left(\alpha^2 - 1/\alpha^2\right) \\
& + \left(\alpha^2 - 1/\alpha^2\right)\left(x_3^*/x_2^* + x_2^*/x_3^*\right) \\
= \; & \left(x_2^* + x_3^*\right)\left(1/x_1^* - 1/x_4^*\right) + \left(x_4^* - x_1^*\right)\left(1/x_2^* + 1/x_3^*\right) + \alpha^2 - 1/\alpha^2 \\
& + \left(\alpha^2 - 1/\alpha^2\right)\left(x_3^*/x_2^* + x_2^*/x_3^*\right) \\
> \; & 0,
\end{aligned}$$

which is contradiction. Hence, we have $x_1^* > x_4^*$.

Suppose that $x_3^*/x_1^* \geq 1$, then we have

$$\begin{aligned}
1 \;\geq\; & \left(\frac{x_1^*}{x_3^*}\right)^2 = \frac{(\alpha x_2^* + \alpha x_3^* + x_4^*/\alpha) \cdot (\alpha/x_1^* + \alpha/x_2^* + 1/(\alpha x_4^*))}{(1/(\alpha x_2^*) + 1/(\alpha x_3^*) + \alpha/x_4^*) \cdot (x_1^*/\alpha + x_2^*/\alpha + \alpha x_4^*)} \\
= \; & \frac{\alpha^2 x_2^*/x_1^* + \alpha^2 x_3^*/x_1^* + x_4^*/x_1^* + \alpha^2 + \alpha^2 x_3^*/x_2^* + x_4^*/x_2^* + x_2^*/x_4^* + x_3^*/x_4^* + 1/\alpha^2}{x_1^*/(\alpha^2 x_2^*) + + x_4^*/x_2^* + 1/\alpha^2 + x_1^*/(\alpha^2 x_3^*) + x_2^*/(\alpha^2 x_3^*) + x_4^*/x_3^* + x_1^*/x_4^* + x_2^*/x_4^* + \alpha^2.}
\end{aligned}$$

This implies that

$$\begin{aligned}
0 \;\geq\; & \alpha^2 x_2^*/x_1^* + \alpha^2 x_3^*/x_1^* + x_4^*/x_1^* + \alpha^2 + \alpha^2 x_3^*/x_2^* + x_4^*/x_2^* + x_2^*/x_4^* + x_3^*/x_4^* + 1/\alpha^2 \\
& - \left(x_1^*/(\alpha^2 x_3^*) + x_2^*/(\alpha^2 x_3^*) + x_4^*/x_3^* + \alpha^2 + x_1^*/(\alpha^2 x_2^*) + x_4^*/x_2^* + x_1^*/x_4^* + x_2^*/x_4^* + 1/\alpha^2\right) \\
= \; & \alpha^2 x_2^*/x_1^* - x_2^*/(\alpha^2 x_1^*) + \alpha^2 x_3^*/x_1^* - x_1^*/(\alpha^2 x_3^*) + x_4^*/x_1^* - x_4^*/x_3^* + \alpha^2 x_3^*/x_2^* \\
& - x_1^*/(\alpha^2 x_2^*) + x_3^*/x_4^* - x_1^*/x_4^* \\
> \; & (1 - 1/1) x_2^*/x_1^* + x_3^*/x_1^* - x_1^*/x_3^* + x_4^*(1/x_1^* - 1/x_3^*) + \left(1 \cdot x_3^* - x_1^*/1\right)/x_2^* + (x_3^* - x_1^*)/x_4^* \\
= \; & x_3^*/x_1^* - x_1^*/x_3^* + x_4^*(1/x_1^* - 1/x_3^*) + \left(x_3^* - x_1^*\right)/x_2^* + (x_3^* - x_1^*)/x_4^* \\
\geq \; & 0,
\end{aligned}$$

which is contradiction. Hence, we have $x_1^* > x_3^*$.

Assume that $x_2^* \leq x_3^*$, then we have

$$\begin{aligned}
1 \;\geq\; & \left(\frac{x_2^*}{x_3^*}\right)^2 = \frac{(x_1^*/\alpha + \alpha x_3^* + \alpha x_4^*) \cdot (\alpha/x_1^* + \alpha/x_2^* + 1/(\alpha x_4^*))}{(\alpha/x_1^* + 1/(\alpha x_3^*) + 1/(\alpha x_4^*)) \cdot (x_1^*/\alpha + x_2^*/\alpha + \alpha x_4^*)} \\
= \; & \frac{1 + \alpha^2 x_3^*/x_1^* + \alpha^2 x_4^*/x_1^* + x_1^*/x_2^* + \alpha^2 x_3^*/x_2^* + \alpha^2 x_4^*/x_2^* + x_1^*/(\alpha^2 x_4^*) + x_3^*/x_4^* + 1}{1 + x_2^*/x_1^* + \alpha^2 x_4^*/x_1 + x_1^*/(\alpha^2 x_3^*) + x_2^*/(\alpha^2 x_3^*) + x_4^*/x_3^* + x_1^*/(\alpha^2 x_4^*) + x_2^*/(\alpha^2 x_4^*) + 1}.
\end{aligned}$$

This means that

$$\begin{aligned}
0 \;\geq\; & (x_2^*)^2 - (x_3^*)^2 \\
= \; & 1 + \alpha^2 x_3^*/x_1^* + \alpha^2 x_4^*/x_1^* + x_1^*/x_2^* + \alpha^2 x_3^*/x_2^* + \alpha^2 x_4^*/x_2^* + x_1^*/(\alpha^2 x_4^*) + x_3^*/x_4^* + 1 \\
& - \left\{1 + x_2^*/x_1^* + \alpha^2 x_4^*/x_1 + x_1^*/(\alpha^2 x_3^*) + x_2^*/(\alpha^2 x_3^*) + x_4^*/x_3^* + x_1^*/(\alpha^2 x_4^*) + x_2^*/(\alpha^2 x_4^*) + 1\right\} \\
= \; & \alpha^2 x_3^*/x_1^* - x_2^*/x_1^* + x_3^*/x_4^* - x_2^*/(\alpha^2 x_4^*) + \alpha^2 x_3^*/x_2^* + \alpha^2 x_4^*/x_2^* - x_1^*/(\alpha^2 x_3^*) - x_2^*/(\alpha^2 x_3^*) \\
& + x_1^*/x_2^* - x_4^*/x_3^* \\
= \; & \alpha^2(x_3^* - x_2^*)/x_1^* + (x_3^* - x_2^*/\alpha^2)/x_4^* + (1/x_2^* - 1/(\alpha^2 x_3^*))x_1^* + (\alpha^2/x_2^* - 1/x_3^*)x_4^* \\
& + \alpha^2 x_3^*/x_2^* - x_2^*/(\alpha^2 x_3^*) \\
> \; & 0,
\end{aligned}$$

which is contradiction. Hence, we have $x_2^* > x_3^*$.

Suppose that $x_2^* \leq x_4^*$, then we have

$$\begin{aligned} 1 \;\geq\; \left(\frac{x_2^*}{x_4^*}\right)^2 &= \frac{(x_1^*/\alpha + \alpha x_3^* + \alpha x_4^*)\cdot(1/(\alpha x_1^*) + x_2^*/\alpha + \alpha/x_3^*)}{(\alpha/x_1^* + 1/(\alpha x_3^*) + 1/(\alpha x_4^*))\cdot(\alpha x_1^* + x_2^*/\alpha + x_3^*/\alpha)} \\ &= \frac{1/\alpha^2 + x_3^*/x_1^* + x_4^*/x_1^* + x_1^*/x_2^* + \alpha^2 x_3^*/x_2^*\alpha^2 x_4^*/x_2^* + x_1^*/x_3^* + \alpha^2 x_4^*/x_3^* + \alpha^2}{\alpha^2 + x_2^*/x_1^* + x_3^*/x_1^* + x_1^*/x_3^* + x_2^*/(\alpha^2 x_3^*) + 1/\alpha^2 + x_1^*/x_4^* + x_2^*/(\alpha^2 x_4^*) + x_3^*/(\alpha^2 x_4^*).} \end{aligned}$$

This means that

$$\begin{aligned} 0 \;\geq\;& (x_2^*)^2 - (x_4^*)^2 \\ =\;& 1/\alpha^2 + x_3^*/x_1^* + x_4^*/x_1^* + x_1^*/x_2^* + \alpha^2 x_3^*/x_2^* + \alpha^2 x_4^*/x_2^* + x_1^*/x_3^* + \alpha^2 x_4^*/x_3^* + \alpha^2 \\ & - \left\{\alpha^2 + x_2^*/x_1^* + x_3^*/x_1^* + x_1^*/x_3^* + x_2^*/(\alpha^2 x_3^*) + 1/\alpha^2 + x_1^*/x_4^* + x_2^*/(\alpha^2 x_4^*) + x_3/(\alpha^2 x_4^*)\right\} \\ =\;& (x_4^* - x_2^*)/x_1^* + \left(\alpha^2 x_4^* - x_2^*/\alpha^2\right) x_3^* + x_1^*\left(1/x_2^* - 1/x_4^*\right) + x_3^*\left(\alpha^2/x_2^* - 1/(\alpha^2 x_4^*)\right) \\ & + \alpha^2 x_4^*/x_2^* - x_2^*/(\alpha^2 x_4^*) \\ >\;& 0, \end{aligned}$$

which is contradiction. Hence, we have $x_2^* > x_4^*$.

Since $x_1^* > x_3^*$, $x_1^* > x_4^*$, $x_2^* > x_3^*$ and $x_2^* > x_4^*$, the arithmetic mean and the harmonic one of $\{\alpha x_2^*, \alpha x_3^*, x_4^*/\alpha\}$ satisfy

$$\max\{\alpha x_2^*, \alpha x_3^*, x_4^*/\alpha\} > \frac{1}{3}\left(\alpha x_2^* + \alpha x_3^* + \frac{x_4^*}{\alpha}\right) > \min\{\alpha x_2^*, \alpha x_3^*, x_4^*/\alpha\} \tag{6.10}$$

$$\max\{\alpha x_2^*, \alpha x_3^*, x_4^*/\alpha\} > 3\left(\frac{1}{\alpha x_2^*} + \frac{1}{\alpha x_3^*} + \frac{\alpha}{x_4^*}\right)^{-1} > \min\{\alpha x_2^*, \alpha x_3^*, x_4^*/\alpha\}, \tag{6.11}$$

respectively. It follows from (6.6), (6.10) and (6.11) that

$$\begin{aligned} \max\{\alpha x_2^*, \alpha x_3^*, x_4^*/\alpha\} &= \sqrt{(\max\{\alpha x_2^*, \alpha x_3^*, x_4^*/\alpha\})^2} \\ &> \sqrt{\frac{1}{3}\left(\alpha x_2^* + \alpha x_3^* + \frac{x_4^*}{\alpha}\right) 3\left(\frac{1}{\alpha x_2^*} + \frac{1}{\alpha x_3^*} + \frac{\alpha}{x_4^*}\right)^{-1}} = x_1^* \qquad (6.12) \\ &> \sqrt{(\min\{\alpha x_2^*, \alpha x_3^*, x_4^*/\alpha\})^2} = \min\{\alpha x_2^*, \alpha x_3^*, x_4^*/\alpha\}. \end{aligned}$$

In the same manner as above, we have

$$\max\left\{\frac{x_1^*}{\alpha}, \alpha x_3^*, \alpha x_4^*\right\} > \quad x_2^* \quad > \min\left\{\frac{x_1^*}{\alpha}, \alpha x_3^*, \alpha x_4^*\right\}, \tag{6.13}$$

$$\max\left\{\frac{x_1^*}{\alpha}, \frac{x_2^*}{\alpha}, \alpha x_4^*\right\} > \quad x_3^* \quad > \min\left\{\frac{x_1^*}{\alpha}, \frac{x_2^*}{\alpha}, \alpha x_4^*\right\}, \tag{6.14}$$

$$\max\left\{\alpha x_1^*, \frac{x_2^*}{\alpha}, \frac{x_3^*}{\alpha}\right\} > \quad x_4^* \quad > \min\left\{\alpha x_1^*, \frac{x_2^*}{\alpha}, \frac{x_3^*}{\alpha}\right\}. \tag{6.15}$$

Without loss of generality, we assume $x_1^* = \alpha$, then it follows from $x_1^* > x_4^*$, $x_2^* > x_3^*$, $\alpha > 1$ and (6.12) that $x_4^* < x_1^* = \alpha < \max\{\alpha x_2^*, \alpha x_3^*, x_4^*/\alpha\} = \alpha x_2^*$. This implies that

$$1 < x_2^*. \tag{6.16}$$

Furthermore, it follows from (6.14) and (6.13) that

$$\begin{aligned} x_3^* &> \min\left\{\frac{x_1^*}{\alpha}, \frac{x_2^*}{\alpha}, \alpha x_4^*\right\} \geq \min\left\{1, \frac{1}{\alpha}\min\left\{\frac{x_1^*}{\alpha}, \alpha x_3^*, \alpha x_4^*\right\}, \alpha x_4^*\right\} \\ &= \min\left\{1, \frac{1}{\alpha}\min\left\{1, \alpha x_3^*, \alpha x_4^*\right\}, \alpha x_4^*\right\} = \min\left\{1, \frac{1}{\alpha}, x_3^*, x_4^*, \alpha x_4^*\right\} \\ &= \min\left\{\frac{1}{\alpha}, x_4^*\right\}. \end{aligned}$$

Assume $x_3^* \leq 1/\alpha$, then we have $1/\alpha \geq x_3^* > x_4^*$. Therefore, it follows from (6.13) that $\max\{x_1^*/\alpha, \alpha x_3^*, \alpha x_4^*\} = 1 > x_2^*$. This is contradiction for (6.16). Hence, we have

$$\alpha x_3^* > 1. \tag{6.17}$$

Assume that $x_1^* \geq x_2^*$, then we have

$$\begin{aligned} 1 &\geq \left(\frac{x_2^*}{x_1^*}\right)^2 = \frac{x_1^*/\alpha + \alpha x_3^* + \alpha x_4^*}{\alpha/x_1^* + 1/(\alpha x_3^*) + 1/(\alpha x_4^*)} \cdot \frac{1/(\alpha x_2^*) + 1/(\alpha x_3^*) + \alpha/x_4^*}{\alpha x_2^* + \alpha x_3^* + x_4^*/\alpha} \\ &= \frac{x_1^*/(\alpha^2 x_2^*) + x_3^*/x_2^* + x_4^*/x_2^* + x_1^*/(\alpha^2 x_3^*) + 1 + x_4^*/x_3^* + x_1^*/x_4^* + (\alpha^2 x_3^*)/x_4^* + \alpha^2}{\alpha^2 x_2^*/x_1^* + \alpha x_3^*/x_1^* + x_4^*/x_1^* + x_2^*/x_3^* + x_4^*/(\alpha^2 x_3^*) + 1 + x_2^*/x_4^* + x_3^*/x_4^* + 1/\alpha^2}. \end{aligned}$$

Therefore, it follows that

$$\begin{aligned} 0 &\leq (x_1^*)^2 - (x_2^*)^2 = 1/\alpha^2 - \alpha^2 + \alpha^2 x_2^*/x_1^* - x_1^*/(\alpha^2 x_2^*) + \alpha^2 x_3^*/x_1^* - x_1^*/(\alpha^2 x_3^*) + x_4^*/(\alpha^2 x_3^*) \\ &\quad -\alpha^2 x_3^*/x_4^* + x_4^*/x_1^* - x_1^*/x_4^* + x_2^*/x_3^* - x_3^*/x_2^* + x_2^*/x_4^* - x_4^*/x_2^* + x_3^*/x_4^* - x_4^*/x_3^* \\ &= \left(\frac{1}{\alpha^2} + \alpha^2\frac{x_2^*}{x_1^*}\right) - \left(\alpha^2 + \frac{x_1^*}{\alpha^2 x_2^*}\right) + \left(\alpha^2\frac{x_3^*}{x_1^*} + \frac{x_3^*}{x_4^*}\right) - \left(\frac{x_3^*}{x_2^*} + \alpha^2\frac{x_3^*}{x_4^*}\right) \\ &\quad + \left(\frac{x_2^*}{x_3^*} + \frac{x_4^*}{\alpha^2 x_3^*}\right) - \left(\frac{x_1^*}{\alpha^2 x_3^*} + \frac{x_4^*}{x_3^*}\right) + \left(\frac{x_4^*}{x_1^*} + \frac{x_2^*}{x_4^*}\right) - \left(\frac{x_4^*}{x_2^*} + \frac{x_1^*}{x_4^*}\right) \\ &\leq \left(\frac{1}{\alpha^2} + \alpha^2\right) - \left(\alpha^2 + \frac{1}{\alpha^2}\right) + \left(\alpha^2\frac{x_3^*}{x_1^*} + \frac{x_3^*}{x_4^*}\right) - \left(\frac{x_3^*}{x_2^*} + \alpha^2\frac{x_3^*}{x_4^*}\right) \\ &\quad + \left(\frac{x_2^*}{x_3^*} + \frac{x_4^*}{\alpha^2 x_3^*}\right) - \left(\frac{x_1^*}{\alpha^2 x_3^*} + \frac{x_4^*}{x_3^*}\right) + \left(\frac{x_4^*}{x_1^*} + \frac{x_1^*}{x_4^*}\right) - \left(\frac{x_4^*}{x_2^*} + \frac{x_1^*}{x_4^*}\right) \\ &= \left(\alpha^2\frac{x_3^*}{x_1^*} + \frac{x_3^*}{x_4^*}\right) - \left(\frac{x_3^*}{x_2^*} + \alpha^2\frac{x_3^*}{x_4^*}\right) + \left(\frac{x_2^*}{x_3^*} + \frac{x_4^*}{\alpha^2 x_3^*}\right) - \left(\frac{x_1^*}{\alpha^2 x_3^*} + \frac{x_4^*}{x_3^*}\right) + x_4^*\left(\frac{1}{x_1^*} - \frac{1}{x_2^*}\right) \\ &\leq \left(\alpha^2\frac{x_3^*}{x_1^*} + \frac{x_3^*}{x_4^*}\right) - \left(\frac{x_3^*}{x_2^*} + \alpha^2\frac{x_3^*}{x_4^*}\right) + \left(\frac{x_2^*}{x_3^*} + \frac{x_4^*}{\alpha^2 x_3^*}\right) - \left(\frac{x_1^*}{\alpha^2 x_3^*} + \frac{x_4^*}{x_3^*}\right) \\ &= \alpha^2 x_3^*\left(\frac{1}{x_1^*} - \frac{1}{x_4^*}\right) + \frac{1}{\alpha^2}\left(\frac{x_4^*}{x_3^*} - \frac{x_1^*}{x_3^*}\right) + x_3^*\left(\frac{1}{x_4^*} - \frac{1}{x_2^*}\right) + \frac{1}{x_3^*}\left(x_2^* - x_4^*\right). \end{aligned}$$

Let $g(\alpha) = \alpha^2 x_3^*\left(\frac{1}{x_1^*} - \frac{1}{x_4^*}\right) + \frac{1}{\alpha^2}\left(\frac{x_4^*}{x_3^*} - \frac{x_1^*}{x_3^*}\right) + x_3^*\left(\frac{1}{x_4^*} - \frac{1}{x_2^*}\right) + \frac{1}{x_3^*}\left(x_2^* - x_4^*\right)$, then we have $g(1) = x_3^*\left(\frac{1}{x_1^*} - \frac{1}{x_2^*}\right) + \frac{1}{x_3^*}\left(x_2^* - x_1^*\right) \leq 0$. Moreover, we have

$$\frac{dg(\alpha)}{d\alpha} = 2\alpha x_3^*\left(\frac{1}{x_1^*} - \frac{1}{x_4^*}\right) - \frac{2}{\alpha^3}\left(\frac{x_4^*}{x_3^*} - \frac{x_1^*}{x_3^*}\right) = 2\left(\frac{\alpha^4(x_3^*)^2 - x_1^* x_4^*}{\alpha^3 x_1^* x_3^* x_4^*}\right)(x_4^* - x_1^*). \tag{6.18}$$

It follows from (6.17) that $\alpha^2 x_3^* > \alpha = x_1^*$ and $\alpha^2 x_3 > \alpha = x_1^* > x_4^*$, and hence, $\frac{dg(\alpha)}{d\alpha} < 0$ for all $\alpha > 1$. Since $g(1) \leq 0$, we have $g(\alpha) < 0$ for all $\alpha > 1$, implying that $0 \leq (x_1^*)^2 - (x_2^*)^2 \leq g(\alpha) < 0$ for all $\alpha > 1$. This is contradiction. Therefore, we have $x_2^* > x_1^*$.

In the similar way, we can prove $x_4^* > x_3^*$. In fact, we have

$$\left(\frac{x_3^*}{x_4^*}\right)^2 = \frac{x_1^*/\alpha + x_2^*/\alpha + \alpha x_4^*}{\alpha/x_1^* + \alpha/x_2^* + 1/(\alpha x_4^*)} \cdot \frac{1/(x_1^*\alpha) + \alpha/x_2^* + \alpha/x_3^*}{\alpha x_1^* + x_2^*/\alpha + x_4^*/\alpha}.$$

Assume that $x_4^* \le x_3^*$, then

$$\begin{aligned}
0 \ \le\ & (x_3^*)^2 - (x_4^*)^2 = \frac{1}{\alpha^2} - \alpha^2 + \alpha^2\left(\frac{x_4^*}{x_2^*} + \frac{x_4^*}{x_3^*} - \frac{x_1^*}{x_2^*}\right) + \frac{1}{\alpha^2}\left(\frac{x_2^*}{x_1^*} - \frac{x_2^*}{x_4^*} - \frac{x_3^*}{x_4^*}\right) \\
& + \frac{x_4^*}{x_1^*} + \frac{x_1^*}{x_2^*} + \frac{x_1^*}{x_3^*} + \frac{x_2^*}{x_3^*} - \frac{x_2^*}{x_1^*} - \frac{x_3^*}{x_1^*} - \frac{x_3^*}{x_2^*} - \frac{x_1^*}{x_4^*} \\
=\ & \alpha^2\left(\frac{x_4^*}{x_2^*} + \frac{x_4^*}{x_3^*} - \frac{x_1^*}{x_2^*} - 1\right) + \frac{1}{\alpha^2}\left(1 + \frac{x_2^*}{x_1^*} - \frac{x_2^*}{x_4^*} - \frac{x_3^*}{x_4^*}\right) \\
& + \frac{x_4^*}{x_1^*} + \frac{x_1^*}{x_2^*} + \frac{x_1^*}{x_3^*} + \frac{x_2^*}{x_3^*} - \frac{x_2^*}{x_1^*} - \frac{x_3^*}{x_1^*} - \frac{x_3^*}{x_2^*} - \frac{x_1^*}{x_4^*} \\
\le\ & \alpha^2\left(\frac{x_4^*}{x_2^*} - \frac{x_1^*}{x_2^*}\right)\frac{1}{\alpha^2}\left(\frac{x_2^*}{x_1^*} - \frac{x_2^*}{x_4^*}\right) + \frac{x_4^*}{x_1^*} + \frac{x_1^*}{x_2^*} + \frac{x_1^*}{x_3^*} + \frac{x_2^*}{x_3^*} - \frac{x_2^*}{x_1^*} - \frac{x_3^*}{x_1^*} - \frac{x_3^*}{x_2^*} - \frac{x_1^*}{x_4^*}.
\end{aligned}$$

Let $h(\alpha) = \alpha^2\left(\frac{x_4^*}{x_2^*} - \frac{x_1^*}{x_2^*}\right)\frac{1}{\alpha^2}\left(\frac{x_2^*}{x_1^*} - \frac{x_2^*}{x_4^*}\right) + \frac{x_4^*}{x_1^*} + \frac{x_1^*}{x_2^*} + \frac{x_1^*}{x_3^*} + \frac{x_2^*}{x_3^*} - \frac{x_2^*}{x_1^*} - \frac{x_3^*}{x_1^*} - \frac{x_3^*}{x_2^*} - \frac{x_1^*}{x_4^*}$, then we have

$$\begin{aligned}
h(1) \ =\ & \frac{x_4^*}{x_2^*} - \frac{x_1^*}{x_2^*} + \frac{x_2^*}{x_1^*} - \frac{x_2^*}{x_4^*} + \frac{x_4^*}{x_1^*} + \frac{x_1^*}{x_2^*} + \frac{x_1^*}{x_3^*} + \frac{x_2^*}{x_3^*} - \frac{x_2^*}{x_1^*} - \frac{x_3^*}{x_1^*} - \frac{x_3^*}{x_2^*} - \frac{x_1^*}{x_4^*} \\
\le\ & \frac{x_4^*}{x_2^*} - \frac{x_3^*}{x_2^*} + \frac{x_2^*}{x_3^*} - \frac{x_2^*}{x_4^*} + \frac{x_4^*}{x_1^*} - \frac{x_3^*}{x_1^*} + \frac{x_1^*}{x_3^*} - \frac{x_1^*}{x_4^*} \\
=\ & \frac{1}{x_2^*}(x_4^* - x_3^*) + x_2^*\left(\frac{1}{x_3^*} - \frac{1}{x_4^*}\right)\frac{1}{x_1^*}(x_4^* - x_3^*) + x_1^*\left(\frac{1}{x_3^*} - \frac{1}{x_4^*}\right) \le 0.
\end{aligned}$$

Moreover, it follows from $\alpha x_1^* > \alpha x_3^* = \max\left\{\frac{x_1^*}{\alpha}, \alpha x_3^*, \alpha x_4^*\right\} > x_2^*$ that

$$\frac{dh(\alpha)}{d\alpha} = 2\alpha\left(\frac{x_4^* - x_1^*}{x_2^*}\right) - \frac{2x_2^*}{\alpha^3}\left(\frac{1}{x_1^*} - \frac{1}{x_4^*}\right) = 2(x_4^* - x_1^*)\frac{\alpha^4 x_1^* x_3^* - (x_2^*)^2}{\alpha^3 x_1^* x_2^* x_4^*} < 0$$

for all $\alpha > 1$. Since $h(1) \le 0$, we have $0 > h(\alpha) \ge (x_3^*)^2 - (x_4^*)^2 \ge 0$ for all $\alpha > 1$. This is contradiction. Therefore, we have $x_4^* > x_3^*$. □

Kazuyuki Sekitani
Department of Systems Engineering
Shizuoka University
Hamamatsu, Shizuoka, 432-8561
E-mail: sekitani@sys.eng.shizuoka.ac.jp