The analytic hierarchy process, max algebra and multi-objective optimisation

Buket Benek Gursoy [a,1], Oliver Mason [a,*,1], Sergeĭ Sergeev [b,2]

[a] *Hamilton Institute, National University of Ireland, Maynooth Maynooth, Co. Kildare, Ireland*
[b] *University of Birmingham, School of Mathematics, Edgbaston B15 2TT, UK*

ABSTRACT

The analytic hierarchy process (AHP) is widely used for decision making involving multiple criteria. Elsner and van den Driessche (2004, 2010) [10,11] introduced a max-algebraic approach to the single criterion AHP. We extend this to the multi-criteria AHP, by considering multi-objective generalisations of the single objective optimisation problem solved in these earlier papers. We relate the existence of globally optimal solutions to the commutativity properties of the associated matrices; we relate min–max optimal solutions to the generalised spectral radius; and we prove that Pareto optimal solutions are guaranteed to exist.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

The analytic hierarchy process (AHP) is a method for ranking alternatives in multi-criteria decision making problems. Developed by Saaty [26], it consists of a three layer hierarchical structure: the overall goal is at the top; the criteria are in the next level; and the alternatives are in the bottom level. The AHP has been used in many different areas including manufacturing systems, finance, politics, education,

business and industry; for more details on the method, see the monographs by Saaty-Vargas and Vaidya-Kumar [27,30].

The essence of the AHP can be described as follows. Given n alternatives we construct a *pairwise comparison matrix* (PC-matrix), $A > 0$ for each criterion, in which a_{ij} indicates the strength of alternative i relative to alternative j for that criterion. A PC-matrix with the property that $a_{ij}a_{ji} = 1$ for all $i \neq j$ and $a_{ii} = 1$ for all i is called a *symmetrically reciprocal* matrix (SR-matrix) [12]. (Note that this abbreviation might clash with the strongly regular matrices of Butkovič [5], but not in this paper.)

Once an SR-matrix A is constructed, the next step in the AHP is to derive a vector $(w_1, \ldots, w_n)$ of positive weights, which can be used to rank the alternatives, with w_i quantifying the weight of alternative i. As observed by Elsner and van den Driessche [10], the ideal situation is where $a_{ij} = w_i/w_j$, in which case the SR-matrix is *transitive*. In practice, this will rarely be the case and it is necessary to approximate A with a transitive matrix T, where $t_{ij} = w_i/w_j$ for some positive weight vector $w = (w_1, \ldots, w_n)$. The problem is then how to construct T given A. Several approaches have been proposed including Saaty's suggestion to take w to be the Perron vector of A, or the approach of Farkas et al. [12], which chooses w to minimise the Euclidean error $\sum_{i,j}(a_{ij} - w_i/w_j)^2$. Elsner and van den Driessche [10,11] suggested selecting w to be the max algebraic eigenvector of A. This is similar in spirit to Saaty's approach and also generates a transitive matrix that minimises the maximal relative error $\max_{i,j} |a_{ij} - w_i/w_j|/a_{ij}$. As noted in [11], minimising this functional is equivalent to minimising

$$e_A(x) = \max_{1 \leqslant i,j \leqslant n} a_{ij}x_j/x_i. \tag{1}$$

The different approaches to approximating an SR-matrix A with a transitive matrix T will in general produce different rankings of the alternatives. The question of how these rankings are affected by the choice of scheme is considered in the recent paper of Tran [29].

In the classical AHP involving multiple criteria, a set of SR-matrices is constructed: one for each criterion. One additional SR-matrix is constructed based on comparisons of the different criteria. Once weight vectors are obtained for each individual criterion, these are then combined using the entries of the weight vector for the criteria-comparison matrix. As an illustration, we take the following numerical example from Saaty [26] and show how the Perron vectors of the comparison matrices are used to construct a weight vector.

Example 1.1. The problem considered is deciding where to go for a one week vacation among the alternatives: 1. Short trips, 2. Quebec, 3. Denver, 4. California. Five criteria are considered: 1. Cost of the trip, 2. Sight-seeing opportunities, 3. Entertainment, 4. Means of travel and 5. Dining. The PC-matrix for the criteria and its Perron vector are given by

$$C = \begin{bmatrix} 1 & 1/5 & 1/5 & 1 & 1/3 \\ 5 & 1 & 1/5 & 1/5 & 1 \\ 5 & 5 & 1 & 1/5 & 1 \\ 1 & 5 & 5 & 1 & 5 \\ 3 & 1 & 1 & 1/5 & 1 \end{bmatrix} \quad \text{and} \quad c = \begin{bmatrix} 0.179 \\ 0.239 \\ 0.431 \\ 0.818 \\ 0.237 \end{bmatrix}.$$

The above matrix C describes the pairwise comparisons between the different *criteria*. For instance, as $c_{21} = 5$, criterion 2 is rated more important than criterion 1; $c_{32} = 5$ indicates that criterion 3 is rated more important than criterion 2 and so on. The vector c contains the weights of the criteria; in this method, criterion 4 is given most weight, followed by criterion 3 and so on.

The SR-matrices, $A_1, \ldots, A_5$, for each of the 5 criteria, their Perron vectors and corresponding ranking schemes are given below. For instance, for criterion 1, the first alternative is preferred to the second as the $(1, 2)$ entry of A_1 is 3. Similarly, for criterion 3, the 4th alternative is preferred to the 1st as the $(4, 1)$ entry of A_3 is 2.

For the cost of the trip:

$$
A_1 = \begin{bmatrix} 1 & 3 & 7 & 9 \\ 1/3 & 1 & 6 & 7 \\ 1/7 & 1/6 & 1 & 3 \\ 1/9 & 1/7 & 1/3 & 1 \end{bmatrix}, \quad v^{(1)} = \begin{bmatrix} 0.877 \\ 0.46 \\ 0.123 \\ 0.064 \end{bmatrix}, \quad 1 > 2 > 3 > 4
$$

For the sight-seeing opportunities:

$$
A_2 = \begin{bmatrix} 1 & 1/5 & 1/6 & 1/4 \\ 5 & 1 & 2 & 4 \\ 6 & 1/2 & 1 & 6 \\ 4 & 1/4 & 1/6 & 1 \end{bmatrix}, \quad v^{(2)} = \begin{bmatrix} 0.091 \\ 0.748 \\ 0.628 \\ 0.196 \end{bmatrix}, \quad 2 > 3 > 4 > 1
$$

For the entertainment:

$$
A_3 = \begin{bmatrix} 1 & 7 & 7 & 1/2 \\ 1/7 & 1 & 1 & 1/7 \\ 1/7 & 1 & 1 & 1/7 \\ 2 & 7 & 7 & 1 \end{bmatrix}, \quad v^{(3)} = \begin{bmatrix} 0.57 \\ 0.096 \\ 0.096 \\ 0.81 \end{bmatrix}, \quad 4 > 1 > 2 = 3
$$

For the means of travel:

$$
A_4 = \begin{bmatrix} 1 & 4 & 1/4 & 1/3 \\ 1/4 & 1 & 1/2 & 3 \\ 4 & 2 & 1 & 3 \\ 3 & 1/3 & 1/3 & 1 \end{bmatrix}, \quad v^{(4)} = \begin{bmatrix} 0.396 \\ 0.355 \\ 0.768 \\ 0.357 \end{bmatrix}, \quad 3 > 1 > 4 > 2
$$

For the dining:

$$
A_5 = \begin{bmatrix} 1 & 1 & 7 & 4 \\ 1 & 1 & 6 & 3 \\ 1/7 & 1/6 & 1 & 1/4 \\ 1/4 & 1/3 & 4 & 1 \end{bmatrix}, \quad v^{(5)} = \begin{bmatrix} 0.723 \\ 0.642 \\ 0.088 \\ 0.242 \end{bmatrix}, \quad 1 > 2 > 4 > 3
$$

To obtain the overall weight vector, we compute the weighted sum $\sum_{i=1}^{5} c_i v^{(i)}$. This gives

$$
w = \begin{bmatrix} 0.919 \\ 0.745 \\ 0.862 \\ 0.757 \end{bmatrix}
$$

with the associated ranking: $1 > 3 > 4 > 2$.

Our work here is inspired by the max-algebraic approach to the AHP introduced by Elsner and van den Driessche [10,11] and extends it in the following manner. In [10,11], the max eigenvector is used as a weight vector for a *single criterion* and it is shown to be optimal in the sense of minimising the maximal relative error as discussed above. This work naturally raises the question of how to treat multiple criteria within the max-algebraic framework. We address this question here by considering the multi-criteria AHP as a multi-objective optimisation problem, in which we have an objective function of the form (1) for each criterion (and associated *SR*-matrix). Rather than combining individual weight vectors as in Example 1.1, we consider three approaches within the framework of multi-objective optimisation, and use the optimal solution as a weight vector in each case. The advantage of this approach is that the weight vector can be interpreted in terms of the maximal relative error functions (1) associated with the *SR*-matrices given as data for the problem. The optimisation problems we consider are the following. First, we investigate the existence of a single *transitive* matrix with a minimum distance to all matrices in the set simultaneously. We remark that this amounts to finding a common subeigenvector of the given matrices. Clearly, this will not in general be possible. The second problem we consider is to obtain a transitive matrix that minimises the maximal distance to any of the given *SR*-matrices. The third problem concerns the existence of a transitive matrix that is Pareto optimal for the given set of matrices. To illustrate our results, we revisit Example 1.1 towards the end of the paper.

2. Notation and mathematical background

The set of all nonnegative real numbers is denoted by $\mathbb{R}_+$; the set of all n-tuples of nonnegative real numbers is denoted by $\mathbb{R}_+^n$ and the set of all $n \times n$ matrices with nonnegative real entries is denoted by $\mathbb{R}_+^{n \times n}$. We denote the set of all n-tuples of positive real numbers by $int(\mathbb{R}_+^n)$. For $A \in \mathbb{R}_+^{n \times n}$ and $1 \le i, j \le n$, a_{ij} refers to the (i, j)th entry of A. The matrix $A = [a_{ij}]$ is nonnegative (positive) if $a_{ij} \ge 0$ ($a_{ij} > 0$) for $1 \le i, j \le n$. This is denoted by $A \in \mathbb{R}_+^{n \times n}$ ($A > 0$).

The weighted directed graph of A is denoted by $D(A)$. It is an ordered pair $(N(A), E(A))$ where $N(A)$ is a finite set of nodes $\{1, 2, \ldots, n\}$ and $E(A)$ is a set of directed edges, with an edge (i, j) from i to j if and only if $a_{ij} > 0$. A path is a sequence of distinct nodes $i_1, i_2, \ldots, i_k$ of length $k - 1$ with the weight $a_{i_1 i_2} a_{i_2 i_3} \cdots a_{i_{k-1} i_k}$, where (i_p, i_{p+1}) is an edge in $D(A)$ for $p = 1, \ldots, k - 1$. It is standard that A is an irreducible matrix if and only if there is a directed path between any two nodes in $D(A)$. A cycle Γ of length k is a closed path of the form $i_1, i_2, \ldots, i_k, i_1$. The kth root of its weight is called its cycle geometric mean. For a matrix $A \in \mathbb{R}_+^{n \times n}$, the maximal cycle geometric mean over all possible cycles in $D(A)$ is denoted by $\mu(A)$. A cycle with the maximum cycle geometric mean is called a critical cycle. Nodes that lie on some critical cycle are known as critical nodes and denoted by $N^C(A)$. The set of edges belonging to critical cycles are said to be critical edges and denoted by $E^C(A)$. The critical matrix of A [8,9], A^C, is formed from the submatrix of A consisting of the rows and columns corresponding to critical nodes as follows. Set $a_{ij}^C = a_{ij}$ if (i, j) lies on a critical cycle and $a_{ij}^C = 0$ otherwise. We use the notation $D^C(A)$ for the critical graph where $D^C(A) = D(A^C) = (N^C(A), E^C(A))$.

The max algebra consists of the set of nonnegative numbers together with the two basic operations $a \oplus b = \max(a, b)$ and $a \otimes b = ab$. This is isomorphic to the max-plus algebra [1,5] via the natural isomorphism $x \to \log(x)$. These operations extend to nonnegative matrices and vectors in the obvious manner [1,2,7]. For A in $\mathbb{R}_+^{n \times n}$, the eigenequation in the max algebra is given by $A \otimes x = \lambda x$, $x \ge 0$, $\lambda \ge 0$. $\mu(A)$ is the largest max eigenvalue of A [5]. If A is irreducible, then it is the unique max eigenvalue of A and there is a positive max eigenvector $x \in int(\mathbb{R}_+^n)$ corresponding to it [1,2]. The eigenvector v is unique up to a scalar multiple if and only if A^C is irreducible.

Observe that an *SR*-matrix A is irreducible and $\mu(A) \ge 1$ [10]. Although our primary interest is in *SR*-matrices, it is noteworthy that many of our results also hold true for non-zero reducible matrices.

For $A \in \mathbb{R}_+^{n \times n}$ with $\mu(A) \le 1$, $I \oplus A \oplus A_\otimes^2 \oplus \cdots$ converges to a finite matrix called the Kleene star of A given by $A^* = I \oplus A \oplus A_\otimes^2 \oplus \cdots \oplus A_\otimes^{n-1}$ where $\mu(A^*) = 1$ [1,7,28]. Here, a_{ij}^* is the maximum weight of a path from i to j of any length [16] (if $i \ne j$), and $A_\otimes^k$ denotes the kth max-algebraic power of A. Note that for each $A \in \mathbb{R}_+^{n \times n}$, if A^* is finite then the max-algebraic sum of all of the columns of A^* is positive.

For $A \in \mathbb{R}_+^{n \times n}$, the set of subeigenvectors of A associated with $\mu(A)$ is called a subeigencone of A and denoted by $V^*(A) = \{y \in \mathbb{R}_+^n \mid A \otimes y \leq \mu(A)y\}$ [28]. It was shown in Proposition 2.5 of [28] that for $A \in \mathbb{R}_+^{n \times n}$ with $\mu(A) = 1$, $V^*(A) = V(A^*) = \text{span}_\oplus(A^*)$ where $V(A^*)$ denotes the eigencone of A^* consisting of its max eigenvectors. Note that the above-mentioned max-algebraic sum of all columns of A^* is in $V^*(A)$ [28], so $V^*(A)$ contains positive vectors. Note that if $\mu(A) > 0$, we can normalise A by $\mu(A)$ and $V^* \left(\frac{A}{\mu(A)} \right) = V^*(A)$.

To the authors' knowledge, max-algebraic subeigenvectors appeared in the works of Gaubert [13, 15]. However, they can be traced back to earlier works on nonnegative matrix scaling, see references in Butkovič–Schneider [6].

For $A \in \mathbb{R}_+^{n \times n}$ we will consider the following set, which was introduced in [11]

$$\mathcal{C}_{A,r} = \{x \in \text{int}(\mathbb{R}_+^n) \mid A \otimes x \leq rx\}. \tag{2}$$

For the special case of $r = \mu(A)$, $\mathcal{C}_{A,\mu(A)}$ is denoted by $\mathcal{C}_A$ [11]. Obviously $\mathcal{C}_A$ is the positive part of $V^*(A)$ (which is non-empty as we argued above), and it coincides with $V^*(A)$ when A is irreducible. To be consistent with the notation of [11], we recall the definition of the normalised set

$$\mathcal{D}_{A,r} = \{x \in \mathcal{C}_{A,r} \mid x_1 = 1\}. \tag{3}$$

As above, $\mathcal{D}_A$ is used to denote the special case where $r = \mu(A)$.

The relations between the sets $\mathcal{C}_{A,r}$, the error function (1) and $\mu(A)$ were clarified by Elsner and van den Driessche [11] and are recalled in the following propositions, which we easily extend (based on [5,6,28]) to the general reducible case.

Proposition 2.1 (cf. [11] Lemma 2)**.** *Let $A \in \mathbb{R}_+^{n \times n}$ be nonzero. Then:*

(i) $\mathcal{C}_{A,r} \neq \emptyset \iff r > 0, r \geq \mu(A)$;
(ii) $x \in \mathcal{C}_{A,r} \iff x \in \text{int}(\mathbb{R}_+^n), e_A(x) \leq r$.

Proof. To prove(i), exploit [5] Theorem 1.6.29 stating that

$$\mu(A) = \min\{\lambda \mid A \otimes x \leqslant \lambda x, \ x \in \text{int}(\mathbb{R}_+^n), \}, \tag{4}$$

when $\mu(A) > 0$ (based on Butkovič-Schneider [6] Theorem 2.6). In the trivial case $r = \mu(A) = 0$, we have $V^*(A) = \{x \mid A \otimes x = 0\}$, which consists of all vectors x such that $x_i \neq 0$ if and only if the ith column of A is zero. In this case $\mathcal{C}_{A,r} = \mathcal{D}_{A,r} = \emptyset$, unless $A = 0$ (which we exclude).

The result of (ii) follows from the definitions of $e_A(x)$ and $\mathcal{C}_{A,r}$ given in (1) and (2). $\square$

See also Gaubert [13] Ch. IV Lemma 1.3.8, Krivulin [20] Lemma 1, and Nussbaum [24] Theorem 3.1 (in a more general nonlinear context) for closely related statements.

Proposition 2.2 (cf. [11], Theorem 1 part 6)**.** *Let $A \in \mathbb{R}_+^{n \times n}$. Then $x \in \mathcal{D}_A$ is unique if and only if A^C is irreducible and $N^C(A) = N(A)$.*

Proof. Proposition 2.1(i) implies that we have $\mathcal{D}_A = \mathcal{C}_A = \emptyset$ when $\mu(A) = 0$, so we can assume $\mu(A) > 0$ and, further, $\mu(A) = 1$. According to [28] Theorem 2.8,

$$V^*(A) = V(A^*) = \left\{ \bigoplus_{i \in M(A)} \lambda_i g^i \oplus \bigoplus_{i \notin N^C(A)} \lambda_j g^j \mid \lambda_i, \lambda_j \in \mathbb{R}_+ \right\}. \tag{5}$$

Here g^i is the ith column of A^*. The subset $M(A) \subset \{1, \ldots, n\}$ is such that for each (maximal) irreducible submatrix of A^C there is a unique index of that submatrix in $M(A)$. By the same theorem of [28],

based on the classical results in [1,7], columns of A^* with indices in the same irreducible submatrix of A^C are proportional to each other, and there is no proportionality between different g^i appearing in (5) (moreover, these g^i are strongly linearly independent [5]).

"If": As A^C is irreducible and all nodes are critical, all columns of A^* are proportional to each other, and (5) shows that $V^*(A)$ is just one ray. $A \geqslant A^C$ is irreducible as well, so $V^*(A) = \mathcal{C}_A$ and $\mathcal{D}_A$ is the unique vector on that ray with $x_1 = 1$.

"Only if": Let A^C be reducible, or let $N^C(A) \neq N(A)$. The vector $z = \bigoplus_{i=1}^n g^i$ is positive. Consider $g^i \oplus \epsilon z$ for all i, taking small enough ϵ so that all entries of ϵz are less than any nonzero entry of A^*. In this case the positive entries of g^i do not change, and if g^i and g^j are not proportional then neither are $g^i \oplus \epsilon z$ and $g^j \oplus \epsilon z$. After suitably normalising these vectors, we obtain two different vectors in $\mathcal{D}_A$. $\square$

Next, let $\Psi \subset \mathbb{R}_+^{n \times n}$ be a finite set of nonnegative matrices given by

$$\Psi = \{A_1, A_2, \ldots, A_m\}, \quad \exists i \colon A_i \neq 0 \tag{6}$$

Given Ψ, let $\Psi_\otimes^p$ denote the set of all products of matrices from Ψ of length $p \geqslant 1$. Formally, $\Psi_\otimes^p = \{A_{j_1} \otimes \cdots \otimes A_{j_p} : 1 \leqslant j_k \leqslant m \text{ for } 1 \leqslant k \leqslant p\}$. Using this, the max version of the generalised spectral radius [21,25] is defined by

$$\hat{\mu}(\Psi) = \limsup_{p \to \infty} (\max_{\psi \in \Psi_\otimes^p} \mu(\psi))^{\frac{1}{p}}. \tag{7}$$

Before stating the next theorem, let S be the matrix given by

$$S = \bigoplus_{A \in \Psi} A = A_1 \oplus A_2 \oplus \cdots \oplus A_m. \tag{8}$$

Note that $S > 0$ if at least one $A_i > 0$. Moreover, $\mu(S) > 0$ if at least one $\mu(A_i) > 0$ and S is irreducible if at least one A_i is irreducible: these represent the main cases in which we are interested.

Theorem 2.1. *Let Ψ be given by (6) and S be given by (8). Then, $\hat{\mu}(\Psi) = \mu(S)$. (Gaubert [14], Benek Gursoy and Mason [3])*

Inspired by the approach to the single criterion AHP adopted in [11], we associate a set $\mathcal{C}_{\Psi,r}$ with the set Ψ of nonnegative matrices and show how the geometric properties of $\mathcal{C}_{A,r}$ discussed in [11] extend to this new setting.

Define

$$\mathcal{C}_{\Psi,r} = \{x \in \mathrm{int}(\mathbb{R}_+^n) \mid e_{A_i}(x) \leq r \text{ for all } A_i \in \Psi\} = \bigcap_{i=1}^m \mathcal{C}_{A_i,r}. \tag{9}$$

We also consider the set of normalised vectors:

$$\mathcal{D}_{\Psi,r} = \{x \in \mathcal{C}_{\Psi,r} \mid x_1 = 1\}. \tag{10}$$

We will use the notations $\mathcal{C}_\Psi$ for $\mathcal{C}_{\Psi,\hat{\mu}(\Psi)}$ and $\mathcal{D}_\Psi$ for $\mathcal{D}_{\Psi,\hat{\mu}(\Psi)}$. The following result shows that the set $\mathcal{C}_{\Psi,r}$ and the set $\mathcal{C}_{S,r}$ are equal. This will allow us to readily extend properties of $\mathcal{C}_{A,r}$ established in [11] to sets of matrices.

Theorem 2.2. *Consider the set $\Psi \subset \mathbb{R}_+^{n \times n}$ given by (6), and let S be given by (8). Then:*

$$\mathcal{C}_{\Psi,r} = \mathcal{C}_{S,r}.$$

Proof. (i): Let x in $\mathcal{C}_{\Psi,r}$ be given. Then, $e_{A_i}(x) \leqslant r$ which implies $A_i \otimes x \leqslant rx$ for each $A_i \in \Psi$ from (ii) in Proposition 2.1. Taking the maximum of both sides from 1 to m, we see that $S \otimes x \leqslant rx$. It follows that $x \in \mathcal{C}_{S,r}$. Thus, $\mathcal{C}_{\Psi,r} \subset \mathcal{C}_{S,r}$.

Now choose some $x \in \mathcal{C}_{S,r}$. Then $e_S(x) \leqslant r$ from (ii) in Proposition 2.1. Since $e_{A_i}(x) \leqslant e_S(x) \leqslant r$ for all $1 \leqslant i \leqslant m$, we obtain $x \in \mathcal{C}_{\Psi,r}$. Thus, $\mathcal{C}_{S,r} \subset \mathcal{C}_{\Psi,r}$. Hence $\mathcal{C}_{\Psi,r} = \mathcal{C}_{S,r}$. $\square$

The following corollary extends Proposition 2.1 to a set of nonnegative matrices Ψ. Since $\hat{\mu}(\Psi) = \mu(S)$ by Theorem 2.1 (ii), the corollary is an immediate consequence of Theorem 2.2.

Corollary 2.1. *Consider the set $\Psi \subset \mathbb{R}_+^{n \times n}$ given by (6), and let S be given by (8). Then:*

(i) $\mathcal{C}_{\Psi,r} \neq \emptyset \iff r > 0, r \geq \hat{\mu}(\Psi)$;
(ii) $x \in \mathcal{C}_{\Psi,r} \iff x \in \text{int}(\mathbb{R}_+^n), e_S(x) \leq r.$

Theorem 2.2 establishes that $\mathcal{C}_{\Psi,r} = \mathcal{C}_{S,r}$. It is immediate that we also have $\mathcal{D}_{\Psi,r} = \mathcal{D}_{S,r}$, $\mathcal{C}_\Psi = \mathcal{C}_S$ and $\mathcal{D}_\Psi = \mathcal{D}_S$. Therefore, studying these sets for a collection of nonnegative matrices reduces to studying the sets associated with the single matrix S. This fact means that the properties of $\mathcal{C}_{A,r}$ discussed in Theorem 1 of [11] can be directly extended to $\mathcal{C}_{\Psi,r}$. We state some of these in the following theorem, noting that property (i) does not require irreducibility.

Theorem 2.3 (cf. [11] Theorem 1). *Let $\Psi \subset \mathbb{R}_+^{n \times n}$ be given in (6), and let S be given by (8). Then:*

(i) *$\mathcal{C}_{\Psi,r}$ and $\mathcal{D}_{\Psi,r}$ are convex and max-convex;*
(ii) *$\mathcal{D}_\Psi$ consists of only one vector if and only if S^C is irreducible and $N^C(S) = N(S)$;*
(iii) *If all matrices in Ψ are irreducible then $\mathcal{D}_{\Psi,r}$ is compact.*

Proof.

(i) (see [28] Proposition 3.1) $\mathcal{C}_{\Psi,r}$ is convex and max-convex, since

$$\mathcal{C}_{\Psi,r} = \bigcap_{i,j} \{x \in \text{int}(\mathbb{R}_+^n) \mid s_{ij}x_j \leqslant rx_i\}, \tag{11}$$

where the sets whose intersection is taken are convex and max-convex. $\mathcal{D}_{\Psi,r}$ inherits these properties as a coordinate section of $\mathcal{C}_{\Psi,r}$.
(ii) follows from Proposition 2.2 applied to S;
(iii) follows from [11] Theorem 1 applied to S. $\square$

We remark that sets that are both convex and max-convex have appeared under various names like Kleene cones [28], polytropes [18], or zones [23].

3. Globally optimal solutions

The application of the max algebra to the AHP is motivated in [10,11] by the following considerations. First, it is observed that, for an *SR*-matrix A, vectors in the set $\mathcal{C}_A$ minimise the function (1) and hence the relative error. Based on this observation, these vectors are used to construct transitive matrices to obtain an overall ranking of the alternatives in the decision process. In light of the properties of $\mathcal{C}_A$, this is justified by the fact that the transitive matrices constructed in this way are *closest* to the original *SR*-matrix A in the sense of the relative error.

Thus, the approach to construct a ranking vector for a single *SR*-matrix taken in [10,11] amounts to solving the following optimisation problem.

$$\min_{x \in \text{int}(\mathbb{R}_+^n)} \{e_A(x)\}. \tag{12}$$

In this and the following section, we are concerned with extending the above approach to the general AHP with n alternatives and m criteria.

Formally, we are given m SR-matrices; one for each criterion. Let Ψ in (6) denote the set of these matrices. For each $A_i \in \Psi$, there is an error function $e_{A_i} : \text{int}(\mathbb{R}_+^n) \to \mathbb{R}_+$ defined as in (1). In contrast to the approach taken in the classical AHP, we view the construction of a ranking vector for the m criteria as a multi-objective optimisation problem for the error functions e_{A_i}, $1 \leqslant i \leqslant m$.

To begin with, we seek a vector that simultaneously minimizes all of the functions e_{A_i}. Such a vector is said to be a *globally optimal solution* for the multi-objective optimisation problem.

Note that for each $A_i \in \Psi$, the set of vectors that minimise $e_{A_i} : \text{int}(\mathbb{R}_+^n) \to \mathbb{R}_+$ is precisely $\mathcal{C}_{A_i}$ [11]: formally,

$$\mathcal{C}_{A_i} = \{x \in \text{int}(\mathbb{R}_+^n) \mid e_{A_i}(x) = \min_{w \in \text{int}(\mathbb{R}_+^n)} e_{A_i}(w)\}, \quad i = 1, 2, \ldots, m. \tag{13}$$

Hence, the problem of finding a vector $x \in \text{int}(\mathbb{R}_+^n)$ that simultaneously minimises all the error functions e_{A_i} amounts to determining when

$$\bigcap_{i=1}^m \mathcal{C}_{A_i} \neq \emptyset.$$

Equivalently, x simultaneously minimises all the error functions if and only if it is a common subeigenvector of A_i for all $i \in \{1, 2, \ldots, m\}$. The remainder of this section is divided into two parts: we first consider the existence of common subeigenvectors for arbitrary nonnegative matrices in the next subsection; we then specialise to sets of SR-matrices and globally optimal solutions.

3.1. Common max-algebraic subeigenvectors of nonnegative matrices

First of all, we consider the general problem of finding a common subeigenvector for a set of nonnegative matrices (not necessarily SR-matrices). Our results are clearly related to the work in [19] concerning the intersection of eigencones of commuting matrices over the max and nonnegative algebra.

In the next result, we adopt the notation $\hat{A}_i = \frac{A_i}{\mu(A_i)}$ for $1 \leqslant i \leqslant m$ and, in an abuse of notation,
$\hat{S} = \bigoplus_{i=1}^m \hat{A}_i$.

Theorem 3.1. *Consider the set $\Psi \subset \mathbb{R}_+^{n \times n}$ in (6). The following assertions are equivalent.*

- (i) $\mu(\hat{S}) = 1$;
- (ii) *There exists some $x \in \text{int}(\mathbb{R}_+^n)$ with $A_i \otimes x \leqslant \mu(A_i)x$ for all $A_i \in \Psi$;*
- (iii) $\mu(A_{j_1} \otimes \cdots \otimes A_{j_p}) \leqslant \mu(A_{j_1}) \cdots \mu(A_{j_p})$ *where $1 \leqslant j_k \leqslant m$ for $1 \leqslant k \leqslant p$. (We say that μ is submultiplicative on Ψ).*

Proof. (i)$\Rightarrow$(ii): First, assume that $\mu(\hat{S}) = 1$. Then, there exists $x \in \text{int}(\mathbb{R}_+^n)$ such that $\hat{S} \otimes x \leqslant x$. Thus, $\hat{A}_i \otimes x \leqslant \hat{S} \otimes x \leqslant x$ for all $A_i \in \Psi$. Hence, $A_i \otimes x \leqslant \mu(A_i)x$ for all i.

(ii)$\Rightarrow$(iii): Suppose that there exists some $x \in \text{int}(\mathbb{R}_+^n)$ with $A_i \otimes x \leqslant \mu(A_i)x$ for all $A_i \in \Psi$. Pick some $\psi \in \Psi_\otimes^p$ such that $\psi = A_{j_1} \otimes A_{j_2} \otimes \cdots \otimes A_{j_p}$ where $1 \leqslant j_k \leqslant m$ for $1 \leqslant k \leqslant p$. Then,

$$
\begin{aligned}
\psi \otimes x &= A_{j_1} \otimes A_{j_2} \otimes \cdots \otimes A_{j_p} \otimes x \\
&\leqslant \mu(A_{j_p})A_{j_1} \otimes A_{j_2} \otimes \cdots \otimes A_{j_{p-1}} \otimes x \\
&\quad \vdots \\
&\leqslant \mu(A_{j_p})\mu(A_{j_{p-1}}) \cdots \mu(A_{j_2})A_{j_1} \otimes x \\
&\leqslant \mu(A_{j_p})\mu(A_{j_{p-1}}) \cdots \mu(A_{j_2})\mu(A_{j_1})x.
\end{aligned}
$$

Writing $r = \mu(A_{j_1})\mu(A_{j_2})\cdots\mu(A_{j_p})$, we see that $x \in \mathcal{C}_{\psi,r}$ from the definition (2). Hence, $\mathcal{C}_{\psi,r} \neq \emptyset$. Point (i) in Proposition 2.1 implies $r \geqslant \mu(\psi)$. Thus, $\mu(A_{j_1})\mu(A_{j_2})\cdots\mu(A_{j_p}) \geqslant \mu(A_{j_1}\otimes A_{j_2}\otimes\cdots\otimes A_{j_p})$. Note that we essentially used (4).

(iii)$\Rightarrow$(i): Consider the set of normalised matrices

$$\hat{\Psi} = \{\hat{A}_1, \hat{A}_2, \ldots, \hat{A}_m\}$$

where $\mu(\hat{A}_i) = 1$ for all $i \in \{1, 2, \ldots, m\}$. Pick some $\psi \in \hat{\Psi}^p_\otimes$. As μ is submultiplicative on Ψ, it is also submultiplicative on $\hat{\Psi}$. Thus, we have $\mu(\psi) \leqslant 1$. As this is true for any $\psi \in \hat{\Psi}^p_\otimes$, it follows that

$$\max_{\psi \in \hat{\Psi}^p_\otimes} \mu(\psi) \leqslant 1.$$

Taking the pth root and $\lim\sup_{p\to\infty}$ of both sides, we see that

$$\hat{\mu}(\hat{\Psi}) \leqslant 1.$$

Theorem 2.1(ii) then implies that $\mu(\hat{S}) \leqslant 1$. Furthermore, since $\mu(\hat{S}) \geqslant \mu(\hat{A}_i) = 1$ for all $i \in \{1, 2, \ldots, m\}$ we obtain $\mu(\hat{S}) = 1$. $\square$

Note that the equivalence (i)$\Leftrightarrow$(ii) can be regarded as a special case of Hershkowitz-Schneider [17] Theorem 2.5, see also [6] Theorem 3.5 for an extension. In these works, the problem of simultaneous nonnegative matrix scaling is considered; this amounts to finding a diagonal matrix X such that $XA_kX^{-1} \leqslant B_k$ for $k = 1, \ldots, m$. For our case, take $B_k = \mathbf{1}_{n\times n}$ (the all-ones matrix) and impose $\mu(A_k) = 1$. However, condition (iii) does not appear in [17] or [6].

In the terminology of Butkovič et al. [5,28], there exists a *simultaneous visualisation* of all of the matrices in Ψ, meaning that $X^{-1}A_iX \leqslant \mu(A_i) \cdot \mathbf{1}_{n\times n}$ for all i, and in particular $(X^{-1}A_iX)_{kl} = \mu(A_i)$ for all i and $(k, l) \in E^C(A_i)$. The following result for general nonnegative matrices will be useful in the next subsection to clarify the relationship between commutativity and the existence of globally optimal solutions for 3×3 *SR*-matrices.

Proposition 3.1. *Let $A, B \in \mathbb{R}^{n\times n}_+$ have $\mu(A) = \mu(B) = 1$. If $\mu(A \oplus B) = 1$, then*

(i) $a_{ij} = b_{ij}$ *for all edges $(i, j) \in E^C(A) \cap E^C(B)$;*
(ii) $a_{ij}b_{ji} = 1$ *for $(i, j) \in E^C(A)$ and $(j, i) \in E^C(B)$.*

Proof. As, $\mu(A) = \mu(B) = 1$, it follows that $\hat{A} = A$ and $\hat{B} = B$. Thus, $\hat{S} = A \oplus B$. From the assumption, we obtain

$$\mu(\hat{S}) = 1.$$

It now follows from Theorem 3.1 that there exists some $x > 0$ with $A \otimes x \leqslant x$, $B \otimes x \leqslant x$. Let $X = \mathrm{diag}(x)$ and consider the diagonally scaled matrices

$$X^{-1}AX, \quad X^{-1}BX.$$

From the choice of X it is immediate that

$$X^{-1}AX \leqslant \mathbf{1}_{n\times n}, \quad X^{-1}BX \leqslant \mathbf{1}_{n\times n}. \tag{14}$$

Furthermore, from well-known facts about diagonal scaling (see Proposition 2.10 of [28]), it follows that

$$1 = \mu(A) = \mu(X^{-1}AX), \quad 1 = \mu(B) = \mu(X^{-1}BX) \tag{15}$$

and that

$$E^C(A) = E^C(X^{-1}AX), \quad E^C(B) = E^C(X^{-1}BX). \tag{16}$$

We prove (i): Let $(i, j) \in E^C(A) \cap E^C(B)$ be given. It follows from (16) that (i, j) is also a critical edge in the digraphs of $X^{-1}AX$ and $X^{-1}BX$. (14) and (15) now imply that

$$\frac{a_{ij}x_j}{x_i} = \frac{b_{ij}x_j}{x_i} = 1.$$

Hence

$$a_{ij} = b_{ij} = \frac{x_i}{x_j}$$

and this completes the proof.

We prove (ii): Let $(i, j) \in E^C(A)$ and $(j, i) \in E^C(B)$. It follows from (16) that (i, j) is also a critical edge in the digraph of $X^{-1}AX$ and (j, i) is a critical edge in the digraph of $X^{-1}BX$. Then

$$\frac{a_{ij}x_j}{x_i} = \frac{b_{ji}x_i}{x_j} = 1,$$

and hence

$$a_{ij}b_{ji} = \frac{a_{ij}x_j}{x_i} \cdot \frac{b_{ji}x_i}{x_j} = 1. \quad \square$$

We next recall the following result, which was established in [19] and shows that commutativity is a sufficient condition for the existence of a common eigenvector for irreducible matrices.

Proposition 3.2 [19]. *Consider the set $\Psi \subset \mathbb{R}_+^{n \times n}$ in (6). Assume that each $A_i \in \Psi$ is irreducible and moreover that*

$$A_i \otimes A_j = A_j \otimes A_i \quad \text{for } 1 \leqslant i, j \leqslant m. \tag{17}$$

Then there exists some $x \in \text{int}(\mathbb{R}_+^n)$ with $A_i \otimes x = \mu(A_i)x$ for $1 \leqslant i \leqslant m$.

The next corollary is an immediate consequence of Proposition 3.2 and the fact, which we recalled in Section 1, that for an irreducible matrix A, the set C_A is the subeigencone $V^*(A)$, which coincides with the eigencone $V(\hat{A}^*)$.

Corollary 3.1. *Consider the set $\Psi \subset \mathbb{R}_+^{n \times n}$ in (6). Assume that each $A_i \in \Psi$ is irreducible and moreover that*

$$\hat{A}_i^* \otimes \hat{A}_j^* = \hat{A}_j^* \otimes \hat{A}_i^* \quad \text{for } 1 \leqslant i, j \leqslant m. \tag{18}$$

Then there exists some $x \in \text{int}(\mathbb{R}_+^n)$ with $A_i \otimes x \leqslant \mu(A_i)x$ for $1 \leqslant i \leqslant m$.

Note that (17) implies (18).

3.2. SR-matrices and globally optimal solutions

In the remainder of this section, we will only focus on *SR*-matrices. We first present the following corollary of Theorem 3.1, which develops the concept of simultaneous visualization for *SR*-matrices. Before stating the corollary, define the *anticritical graph* of an *SR*-matrix to consist of the edges $E^{\overline{C}}(A)$ given by:

$$(i, j) \in E^{\overline{C}}(A) \Leftrightarrow (j, i) \in E^{C}(A). \tag{19}$$

Corollary 3.2. *Consider the set $\Psi \subset \mathbb{R}_+^{n \times n}$ in (6). Assume that each $A_i \in \Psi$ is an SR-matrix. If any of the equivalent statements of Theorem 3.1 holds, then there exists some $x \in int(\mathbb{R}_+^n)$ such that for $X = \mathrm{diag}(x)$ we have*

$$\mu^{-1}(A_i) \cdot \mathbf{1}_{n \times n} \leqslant X^{-1} A_i X \leqslant \mu(A_i) \cdot \mathbf{1}_{n \times n} \tag{20}$$

In particular,

$$\begin{aligned} (k, l) \in E^{C}(A_i) &\Rightarrow (X^{-1} A_i X)_{kl} = \mu(A_i), \\ (k, l) \in E^{\overline{C}}(A_i) &\Rightarrow (X^{-1} A_i X)_{kl} = \mu^{-1}(A_i), \end{aligned} \tag{21}$$

Proof. The right-hand side inequality of (20) is the same as Theorem 3.1 (ii). For the remaining left-hand side inequality of (20) we observe that $x_i^{-1} a_{ij} x_j \leqslant \mu(A)$ is equivalent to $x_j^{-1} a_{ij}^{-1} x_i \geqslant \mu^{-1}(A)$. Then we apply $a_{ij}^{-1} = a_{ji}$. $\square$

We next show that two distinct *SR*-matrices A, B in $\mathbb{R}_+^{2 \times 2}$ cannot have a common subeigenvector. Let

$$A = \begin{bmatrix} 1 & a \\ 1/a & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & b \\ 1/b & 1 \end{bmatrix}$$

and assume that $A \neq B$. Clearly, $\mu(A) = \mu(B) = 1$ and $\hat{S} = A \oplus B$. If $a > b$, then $1/a < 1/b$ and $\mu(\hat{S}) = a/b > 1$. If $b > a$, then $1/b < 1/a$ and $\mu(\hat{S}) = b/a > 1$. In both cases, $\mu(\hat{S}) \neq 1$. Hence by Theorem 3.1, A and B do not have a common subeigenvector.

Proposition 3.2 shows that commuting irreducible matrices possess a common max eigenvector. We now show that for 3×3 *SR*-matrices, commutativity is both necessary and sufficient for the existence of a common subeigenvector.

Remark 3.1. For an *SR*-matrix $A \in \mathbb{R}_+^{3 \times 3}$, it is immediate that all cycle products of length one and two in $D(A)$ are equal to 1. Further, there are two possible cycle products of length 3 in $D(A)$: $a_{12} a_{23} a_{31}$ and $a_{13} a_{32} a_{21}$. As A is an *SR*-matrix, it follows that

$$a_{12} a_{23} a_{31} = \frac{1}{a_{13} a_{32} a_{21}}$$

and hence at least one of the above products must be greater than or equal to 1. Since $\mu(A) \geqslant 1$, one of the cycles of length three is critical, and the other cycle is anticritical. Thus, $N^C(A) = N(A)$ and A^C is irreducible. Hence, it follows from Proposition 2.2 that A has a unique subeigenvector up to a scalar multiple in $\mathcal{C}_A$ which is its max eigenvector. Observe that each edge (i, j) with $i \neq j$ belongs either to the critical or to the anticritical graph.

Our next result characterises when two *SR*-matrices in $\mathbb{R}_+^{3 \times 3}$ have a common subeigenvector.

Theorem 3.2. *Let a set* $\{A, B\} \subset \mathbb{R}_+^{3 \times 3}$ *of SR-matrices be given. Write* $\hat{S} = \hat{A} \oplus \hat{B}$. *The following are equivalent.*

(i) $\mu(\hat{S}) = 1$;
(ii) *A and B commute*;
(iii) *There exists a vector* $x \in int(\mathbb{R}_+^n)$ *with* $A \otimes x \leqslant \mu(A)x$, $B \otimes x \leqslant \mu(B)x$.

Proof. The equivalence of (i) and (iii) follows immediately from Theorem 3.1 so we will show that (i) and (ii) are also equivalent.

(ii) $\Rightarrow$ (i) follows immediately from Proposition 3.2. We prove (iii)$\Rightarrow$(ii). First note that it follows from Remark 3.1 that for distinct i, j, k, the edges (i, j), (j, k) are either both critical or both anti-critical for A. The same is true of B. Calculating $X^{-1}AX$ and $X^{-1}BX$ where $X = \text{diag}(x)$, it follows from Theorem 3.1, Corollary 3.2 and the identities $\mu(A)\mu(B) = \mu(B)\mu(A)$, $\mu(A)\mu^{-1}(B) = \mu^{-1}(B)\mu(A)$ that

$$a_{ij}b_{jk} = b_{ij}a_{jk} \tag{22}$$

for any distinct i, j, k. It now follows from (22) that for $i \neq j$

$$\begin{aligned}
(A \otimes B)_{ij} &= a_{ii}b_{ij} \oplus a_{ij}b_{jj} \oplus a_{ik}b_{kj} \\
&= b_{ij} \oplus a_{ij} \oplus b_{ik}a_{kj} \\
&= (B \otimes A)_{ij}
\end{aligned}$$

where $k \neq i$, $k \neq j$. Rewriting (22) as $a_{ji}b_{ik} = b_{ji}a_{ik}$, it follows readily that $b_{ik}a_{ki} = a_{ij}b_{ji}$ and $a_{ik}b_{ki} = b_{ij}a_{ji}$. It now follows that for $1 \leqslant i \leqslant 3$,

$$\begin{aligned}
(A \otimes B)_{ii} &= a_{ii}b_{ii} \oplus a_{ij}b_{ji} \oplus a_{ik}b_{ki} \\
&= b_{ii}a_{ii} \oplus b_{ij}a_{ji} \oplus b_{ik}a_{ki} \\
&= (B \otimes A)_{ii}
\end{aligned}$$

Thus, $A \otimes B = B \otimes A$ as claimed. $\square$

It is now straightforward to extend the above result to an arbitrary finite set of *SR*-matrices in $\mathbb{R}_+^{3 \times 3}$.

Theorem 3.3. *Let a set* $\{A_1, \ldots, A_m\} \subset \mathbb{R}_+^{3 \times 3}$ *of SR-matrices be given. Write* $\hat{S} = \hat{A}_1 \oplus \cdots \oplus \hat{A}_m$. *The following are equivalent.*

(i) $\mu(\hat{S}) = 1$;
(ii) $A_i \otimes A_j = A_j \otimes A_i$ *for all* i, j;
(iii) *There exists a vector* $x \in int(\mathbb{R}_+^n)$ *with* $A_i \otimes x \leqslant \mu(A_i)x$ *for all* i.

Proof. As above, the equivalence of (i) and (iii) follows immediately from Theorem 3.1 and (ii) $\Rightarrow$ (i) follows immediately from Proposition 3.2. To show that (i) $\Rightarrow$ (ii), suppose $\mu(\hat{S}) = 1$. Then it follows that for all i, j,

$$\hat{A}_i \oplus \hat{A}_j \leqslant \hat{S}$$

and hence that $\mu(\hat{A}_i \oplus \hat{A}_j) \leqslant 1$. As $\mu(\hat{A}_i \oplus \hat{A}_j) \geqslant 1$, it is immediate that

$$\mu(\hat{A}_i \oplus \hat{A}_j) = 1$$

for all i, j in $\{1, \ldots, m\}$. It follows immediately from Theorem 3.2 that

$$A_i \otimes A_j = A_j \otimes A_i$$

for $1 \leqslant i, j \leqslant m$ as claimed. $\quad \square$

We note with the following example that commutativity is not a necessary condition for 4×4 *SR*-matrices to possess a common subeigenvector.

Example 3.1. Consider the *SR*-matrices given by

$$A = \begin{bmatrix} 1 & 8 & 1/4 & 7 \\ 1/8 & 1 & 6 & 1/4 \\ 4 & 1/6 & 1 & 4 \\ 1/7 & 4 & 1/4 & 1 \end{bmatrix} B = \begin{bmatrix} 1 & 4 & 5 & 9 \\ 1/4 & 1 & 1/8 & 9 \\ 1/5 & 8 & 1 & 1/8 \\ 1/9 & 1/9 & 8 & 1 \end{bmatrix}$$

where $\mu(\hat{S}) = 1$. Here, $x = \begin{bmatrix} 1 & 0.721 & 0.693 & 0.667 \end{bmatrix}^T$ is a common subeigenvector. However, it can be readily verified that $A \otimes B \neq B \otimes A$.

4. Min–max optimal points and the generalised spectral radius

In general, it will not be possible to find a single vector x that is globally optimal for the set Ψ of *SR*-matrices given by (6). With this in mind, in this short section we consider a different notion of optimal solution for the multiple objective functions $e_{A_i} : \text{int}(\mathbb{R}^n_+) \rightarrow \mathbb{R}_+$, $1 \leqslant i \leqslant m$. In fact, we consider the following optimisation problem.

$$\min_{x \in \text{int}(\mathbb{R}^n_+)} \left(\max_{1 \leqslant i \leqslant m} e_{A_i}(x) \right). \tag{23}$$

In words, we are seeking a weight vector that minimises the maximal relative error where the maximum is taken over the m criteria (*SR*-matrices).

Corollary 2.1 has the following interpretation in terms of the optimisation problem given in (23).

Proposition 4.1. *Consider the set Ψ given by (6). Then:*

(i) $\hat{\mu}(\Psi) = \min_{x \in \text{int}(\mathbb{R}^n_+)} \left(\max_{1 \leqslant i \leqslant m} e_{A_i}(x) \right);$

(ii) *x solves (23) if and only if $x \in \mathcal{C}_\Psi$.*

Proof. Corollary 2.1 shows that there exists some $x \in \text{int}(\mathbb{R}^n_+)$ with

$$\max_{1 \leqslant i \leqslant m} e_{A_i}(x) \leqslant r$$

if and only if $r \geqslant \hat{\mu}(\Psi)$. (i) follows from this observation. The result of (ii) is then immediate from the definition of $\mathcal{C}_\Psi$. $\quad \square$

5. Pareto optimality and the AHP

Thus far, we have considered two different approaches to the multi-objective optimisation problem associated with the AHP. In this section we turn our attention to what is arguably the most common

framework adopted in multi-objective optimisation: Pareto Optimality [4,22]. As above, we are concerned with the existence of optimal points for the set of objective functions e_{A_i}, for $1 \leqslant i \leqslant m$ associated with the set Ψ (6) of *SR*-matrices. We first recall the notion of *weak Pareto optimality*.

Definition 5.1 [22]. $w \in \text{int}(\mathbb{R}^n_+)$ is said to be a *weak Pareto optimal point* for the functions e_{A_i} : $\text{int}(\mathbb{R}^n_+) \to \mathbb{R}_+ (1 \leqslant i \leqslant m)$ if there does not exist $x \in \text{int}(\mathbb{R}^n_+)$ such that

$$e_{A_i}(x) < e_{A_i}(w)$$

for all $i = 1, 2, \ldots, m$.

The next lemma shows that every point in the set $\mathcal{C}_\Psi$ is a weak Pareto optimal point for $e_{A_1}, \ldots, e_{A_m}$.

Lemma 5.1. *Let $\Psi \subset \mathbb{R}^{n \times n}_+$ be given by (6). Any $w \in \mathcal{C}_\Psi$ is a weak Pareto optimal point for $e_{A_1}, \ldots, e_{A_m}$.*

Proof. Let $w \in \mathcal{C}_\Psi$ be given. Then $e_{A_i}(w) \leqslant \hat{\mu}(\Psi)$ for $1 \leqslant i \leqslant m$. If there exists some $x \in \text{int}(\mathbb{R}^n_+)$ such that $e_{A_i}(x) < e_{A_i}(w)$ for $1 \leqslant i \leqslant m$, then for this x

$$e_{A_i}(x) < \hat{\mu}(\Psi)$$

for $1 \leqslant i \leqslant m$. This contradicts Proposition 4.1. $\square$

We next recall the usual definition of a Pareto optimal point.

Definition 5.2 [22]. $w \in \text{int}(\mathbb{R}^n_+)$ is said to be a *Pareto optimal point* for the functions e_{A_i} : $\text{int}(\mathbb{R}^n_+) \to \mathbb{R}_+ (1 \leqslant i \leqslant m)$ if $e_{A_i}(x) \leqslant e_{A_i}(w)$ for $1 \leqslant i \leqslant m$ implies $e_{A_i}(x) = e_{A_i}(w)$ for all $1 \leqslant i \leqslant m$.

We later show that the multi-objective optimisation problem associated with the AHP always admits a Pareto optimal point. We first present some simple facts concerning such points.

Theorem 5.1. *Let $\Psi \subset \mathbb{R}^{n \times n}_+$ be given by (6). Then:*

(i) *If $w \in \mathcal{C}_\Psi$ is unique up to a scalar multiple, then it is a Pareto optimal point for $e_{A_1}, \ldots, e_{A_m}$;*
(ii) *If $w \in \mathcal{C}_{A_i}$ is unique up to a scalar multiple for some $i \in \{1, 2, \ldots, m\}$, then it is a Pareto optimal point for $e_{A_1}, \ldots, e_{A_m}$.*

Proof. Observe that both conditions imply $\hat{\mu}(\Psi) > 0$.

(i) Assume that $w \in \mathcal{C}_\Psi$ is unique up to a scalar multiple. Pick some $x \in \text{int}(\mathbb{R}^n_+)$ such that $e_{A_i}(x) \leqslant e_{A_i}(w)$ for all i. Then, $e_{A_i}(x) \leqslant \hat{\mu}(\Psi)$ for all i which implies that $x \in \mathcal{C}_\Psi$. Thus, $x = \alpha w$ for some $\alpha \in \mathbb{R}_+$. Hence, $e_{A_i}(x) = e_{A_i}(w)$ for all i and w is a Pareto optimal point.
(ii) Assume that for some $A_i \in \Psi$, $w \in \mathcal{C}_{A_i}$ is unique up to a scalar multiple. Suppose $x \in \text{int}(\mathbb{R}^n_+)$ is such that $e_{A_j}(x) \leqslant e_{A_j}(w)$ for all $1 \leqslant j \leqslant m$. In particular, $x \in \mathcal{C}_{A_i}$, and this implies that $x = \alpha w$ for some $\alpha \in \mathbb{R}$. Further, it is immediate that for any other $A_j \in \Psi$ $(i \neq j)$, we have $e_{A_j}(x) = e_{A_j}(w)$. Thus, w is a Pareto optimal point. $\square$

By Proposition 2.2, condition (i) is equivalent to $N^C(S) = N(S)$ and S^C to be irreducible, and condition (ii) is equivalent to $N^C(A_i) = N(A_i)$ and A_i^C to be irreducible for some i.

Corollary 5.1. *Let the set $\Psi \subset \mathbb{R}^{n \times n}_+$ given by (6) consist of SR-matrices. For $n \in \{2, 3\}$, any $w \in \mathcal{C}_{A_i} (1 \leqslant i \leqslant m)$ is a Pareto optimal point for $e_{A_1}, \ldots, e_{A_m}$.*
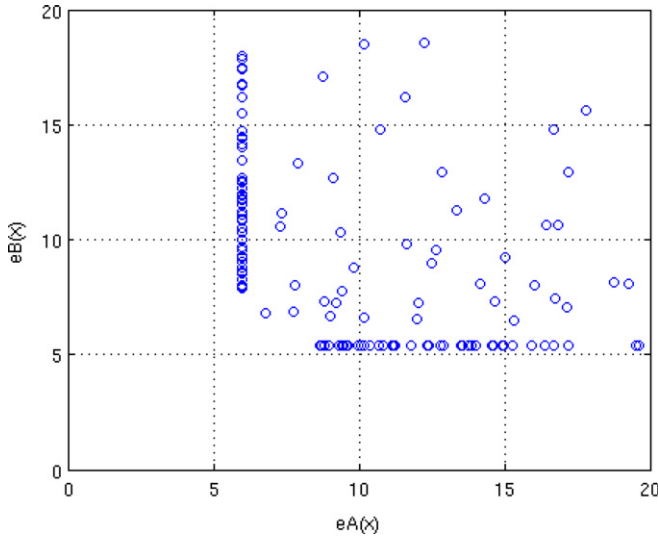
Fig. 1. Refer to Example 5.1.

Proof. Notice that from Remark 3.1 for the 3×3 case, there exists a unique subeigenvector up to a scalar multiple in each C_{A_i} for $1 \leqslant i \leqslant m$. This is also true for the 2×2 case because $N^C(A) = N(A)$ and A^C is irreducible. The result directly follows from (ii) in Theorem 5.1. $\square$

The following example demonstrates point (i) in Theorem 5.1.

Example 5.1. Consider the following matrices given by

$$
A = \begin{bmatrix} 1 & 9 & 1/4 & 2 \\ 1/9 & 1 & 6 & 3 \\ 4 & 1/6 & 1 & 1/4 \\ 1/2 & 1/3 & 4 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 1/2 & 4 & 1/8 \\ 2 & 1 & 3 & 2 \\ 1/4 & 1/3 & 1 & 5 \\ 8 & 1/2 & 1/5 & 1 \end{bmatrix}.
$$

The S matrix is obtained as follows

$$
S = \begin{bmatrix} 1 & 9 & 4 & 2 \\ 2 & 1 & 6 & 3 \\ 4 & 1/3 & 1 & 5 \\ 8 & 1/2 & 4 & 1 \end{bmatrix}
$$

where $N^C(S) = N(S)$ and S^C is irreducible. From Proposition 2.2, we have a unique vector (up to a scalar multiple) in C_Ψ: $w = [\,1\ 0.758\ 0.861\ 1.174\,]^T$. Fig. 1 represents the values of $e_A(x)$ and $e_B(x)$ at w and some points in C_A, C_B and $\text{int}(\mathbb{R}^n_+)$. Remark that Pareto optimality is observed at $w \in C_\Psi$ where $e_A(w) = e_B(w) = 6.817$.

Our objective in the remainder of this section is to show that the multi-objective optimisation problem associated with the AHP *always* admits a Pareto optimal solution. We first recall the following

general result giving a sufficient condition for the existence of a Pareto optimal point with respect to a set $E \subset \mathbb{R}^n$. Essentially, this is a direct application of the fact that a continuous function on a compact set always attains its minimum.

Theorem 5.2 [4]. *Let $E \subseteq \mathbb{R}^n$ be nonempty and compact. Let a set of continuous functions $\{f_1, \dots, f_m\}$ be given where*

$$f_i : E \to \mathbb{R}_+$$

for $1 \leqslant i \leqslant m$. There exists $w \in E$ such that $x \in E, f_i(x) \leqslant f_i(w)$ for $1 \leqslant i \leqslant m$ implies $f_i(x) = f_i(w)$ for $1 \leqslant i \leqslant m$.

This result follows from elementary real analysis and the observation that if w minimises the (continuous) weighted sum $\sum_{i=1}^{m} \alpha_i f_i(x)$ where $\alpha_i > 0$ for $1 \leqslant i \leqslant m$, then w must be Pareto optimal for the functions $f_1, \dots, f_m$.

A point w satisfying the conclusion of the Theorem 5.2 is said to be Pareto optimal for $\{f_1, \dots, f_m\}$ *with respect to E*. Thus, for any multi-objective optimisation problem with continuous objective functions defined on a compact set, there exists a point that is Pareto optimal with respect to the given set.

To apply the above result to the AHP, we first note that for a set Ψ of SR-matrices, $\mathcal{D}_\Psi$ is compact by Theorem 2.3 (recall that SR-matrices are positive and hence irreducible). We now show that any point in $\mathcal{D}_\Psi$ that is Pareto optimal with respect to $\mathcal{D}_\Psi$ is also Pareto optimal with respect to the set $\mathrm{int}(\mathbb{R}^n_+)$.

Although we don't specifically use the SR property in the following, we assume Ψ to consist of SR-matrices as we are primarily interested in the AHP application. Instead, we could just assume that A_i are irreducible implying that $\mathcal{D}_\Psi$ is compact and a Pareto optimal point exists.

Lemma 5.2. *Consider the set Ψ in (6) and assume that A_i is an SR-matrix for $1 \leqslant i \leqslant m$. Let w be a Pareto optimal point for $\{e_{A_1}, \dots, e_{A_m}\}$ with respect to $\mathcal{D}_\Psi$. Then, w is also Pareto optimal for $e_{A_1}, \dots, e_{A_m}$ with respect to $\mathrm{int}(\mathbb{R}^n_+)$.*

Proof. Assume that $w \in \mathcal{D}_\Psi$ is a Pareto optimal point with respect to $\mathcal{D}_\Psi$. Suppose $x \in \mathrm{int}(\mathbb{R}^n_+) \backslash \mathcal{C}_\Psi$. Then from the definition of $\mathcal{C}_\Psi$ (9), it follows that

$$e_{A_{i_0}}(x) > \hat{\mu}(\Psi) \quad \text{for some } i_0. \tag{24}$$

As $w \in \mathcal{D}_\Psi$, $e_{A_i}(w) \leqslant \hat{\mu}(\Psi)$ for $1 \leqslant i \leqslant m$. It follows immediately from (24) that for any $x \notin \mathcal{C}_\Psi$, it cannot happen that $e_{A_i}(x) \leqslant e_{A_i}(w)$ for $1 \leqslant i \leqslant m$.

Let x in $\mathcal{C}_\Psi$ be such that

$$e_{A_i}(x) \leqslant e_{A_i}(w) \quad \text{for } 1 \leqslant i \leqslant m.$$

As $e_{A_i}(\lambda x) = e_{A_i}(x)$ for all $\lambda > 0$, $1 \leqslant i \leqslant m$, and w is Pareto optimal with respect to $\mathcal{D}_\Psi$, it follows that $e_{A_i}(x) = e_{A_i}(w)$ for $1 \leqslant i \leqslant m$. $\square$

Our next step is to show that there exists a point $x \in \mathcal{D}_\Psi$ that is Pareto optimal with respect to $\mathcal{D}_\Psi$.

Proposition 5.1. *Consider the set Ψ in (6) and assume that A_i is an SR-matrix for $1 \leqslant i \leqslant m$. There exists $x \in \mathcal{D}_\Psi$ that is Pareto optimal for $e_{A_1}, \dots, e_{A_m}$ with respect to $\mathcal{D}_\Psi$.*

Proof. First note that $\mathcal{D}_\Psi \neq \emptyset$ since $\hat{\mu}(\Psi) > 0$. Theorem 2.3 shows that $\mathcal{D}_\Psi$ is compact. Furthermore, for any irreducible matrix A, the function $e_A : \mathcal{D}_\Psi \to \mathbb{R}_+$ is a composition of continuous functions and hence continuous on a compact set. Theorem 5.2 implies that there exists w in $\mathcal{D}_\Psi$ that is Pareto optimal with respect to $\mathcal{D}_\Psi$. $\square$

Combining Proposition 5.1 with Lemma 5.2, we immediately obtain the following result.

Corollary 5.2. *Consider the set Ψ in (6) and assume that A_i is an SR-matrix for $1 \leqslant i \leqslant m$. There exists $x \in \mathcal{D}_\Psi$ that is Pareto optimal for $e_{A_1}, \ldots, e_{A_m}$ with respect to $int(\mathbb{R}^n_+)$.*

Corollary 5.2 means that there exists a vector x of positive weights that is simultaneously Pareto optimal and also optimal in the *min–max* sense of Section 4 for the error functions e_{A_i}, $1 \leqslant i \leqslant m$.

Finally, to illustrate the above results, we revisit Example 1.1.

Example 5.2. (Example 1.1 Revisited) Let $C, A_1, \ldots, A_5$ be as in Example 1.1. Taking α_i, $1 \leqslant i \leqslant 5$ to be the *i*th entry of the max eigenvector of C, normalised so that $\alpha_1 = 1$, we apply Theorem 5.2 to compute Pareto optimal solutions in the set $\mathcal{D}_\Psi$ by minimising the weighted sum

$$\sum_{i=1}^{m} \alpha_i e_{A_i}(x)$$

using the MATLAB function *fminsearch*.

Observe that $\mu(\hat{S}) = 4.985$, so there is no common subeigenvector in this case. Next, we calculate the max eigenvector of C:

$$\begin{bmatrix} 1 \\ 1.495 \\ 2.236 \\ 3.344 \\ 0.897 \end{bmatrix}$$

We find that there are multiple Pareto optimal points giving at least two possible distinct rankings: $1 > 3 > 4 > 2$ and $1 > 3 > 2 > 4$. Notice that first ranking scheme is the same as the one obtained from the classical method used in Example 1.1. The second ranking scheme is also reasonable, since if we analyse the local rankings associated with the set of *SR*-matrices in detail, we see that $2 > 4$ for A_1, A_2 and A_5. In particular, 2 is preferred to all other alternatives for A_2.

6. Conclusions

Building on the work of Elsner and van den Driessche [10,11], we have considered a max-algebraic approach to the multi-criteria AHP within the framework of multi-objective optimisation. Papers [10, 11] characterise the max eigenvectors and subeigenvectors of a *single SR*-matrix as solving an optimisation problem with a single objective. We have extended this work to the multi-criteria AHP by directly considering several natural extensions of this basic optimisation problem to the multiple objective case. Specifically, we have presented results concerning the existence of: globally optimal solutions; min–max optimal solutions; Pareto optimal solutions. The principal contribution of the paper is to draw attention to this max-algebraic perspective on the multi-criteria AHP, with the main results in this direction being: establishing the connection between the generalised spectral radius and min–max optimal solutions (Proposition 4.1); proving the existence of Pareto optimal solutions and showing that it is possible to simultaneously solve the Pareto and min–max optimisation problems (Proposition 5.1 and Corollary 5.2). We have also related the existence of globally optimal solutions to the existence of common subeigenvectors and highlighted connections between this question and commutativity (Theorem 3.2).

Acknowledgments

The authors would like to thank the anonymous reviewer for their helpful comments and suggestions, which have improved the paper.

References

[1] F. Baccelli, G. Cohen, G.J. Olsder, J.-P. Quadrat, Synchronization and Linearity: An Algebra for Discrete Event Systems, John Wiley & Sons, Chichester, New York, 1992.

[2] R.B. Bapat, A max version of the Perron–Frobenius theorem, Linear Algebra Appl. 275–276 (1998) 3–18.

[3] B. Benek Gursoy, O. Mason, P_{max}^1 and S_{max} properties and asymptotic stability in the max algebra, Linear Algebra Appl. 435 (2011) 1008–1018.

[4] T.F. Bewley, General Equilibrium, Overlapping Generations Models, and Optimal Growth Theory, Harvard University Press, Cambridge, MA, 2007.

[5] P. Butkovič, Max-linear Systems: Theory and Algorithms, Springer-Verlag, London, 2010.

[6] P. Butkovič, H. Schneider, Applications of max algebra to diagonal scaling of matrices, Electron. J. Linear Algebra 13 (2005) 262–273.

[7] R.A. Cuninghame-Green, Minimax algebra, Springer-Verlag, Berlin, 1979.

[8] L. Elsner, P. van den Driessche, On the power method in max algebra, Linear Algebra Appl. 302–303 (1999) 17–32.

[9] L. Elsner, P. van den Driessche, Modifying the power method in max algebra, Linear Algebra Appl. 332–334 (2001) 3–13.

[10] L. Elsner, P. van den Driessche, Max-algebra and pairwise comparison matrices, Linear Algebra Appl. 385 (2004) 47–62.

[11] L. Elsner, P. van den Driessche, Max-algebra and pairwise comparison matrices II, Linear Algebra Appl. 432 (2010) 927–935.

[12] A. Farkas, P. Lancaster, P. Rózsa, Consistency adjustment for pairwise comparison matrices, Numer. Linear Algebra Appl. 10 (2003) 689–700.

[13] S. Gaubert, Théorie des Systèmes Linéaires dans les Dioïdes, Ph.D. Thesis, L'École Nationale Supérieure des Mines de Paris, France, 1992.

[14] S. Gaubert, Performance evaluation of (max, +) automata, IEEE Trans. Auto. Contr. 40 (1995) 2014–2025.

[15] S. Gaubert, Resource optimization and (min, +) spectral theory, IEEE Trans. Automat. Contr. 40 (1995) 1931–1934.

[16] B. Heidergott, G.J. Olsder, J. van der Woude, Max Plus at Work: Modeling and Analysis of Synchronized Systems, A Course on Max-Plus Algebra and Its Applications, Princeton University Press, Princeton, NJ, 2006.

[17] D. Hershkowitz, H. Schneider, One-sided simultaneous inequalities and sandwich theorems for diagonal similarity and diagonal equivalence of nonnegative matrices, Electron. J. Linear Algebra 10 (2003) 81–101.

[18] M. Joswig, K. Kulas, Tropical and ordinary convexity combined, Adv. Geometry 10 (2010) 333–352.

[19] R.D. Katz, H. Schneider, S. Sergeev, On commuting matrices in max algebra and in classical nonnegative algebra, Linear Algebra Appl. 436 (2012) 276–292.

[20] N.K. Krivulin, Evaluation of bounds on the mean rate of growth of the state vector of a linear dynamical stochastic system in idempotent algebra, Vestnik St. Petersburg Univ.: Math. 38 (2005) 45–54.

[21] Y.-Y. Lur, A max version of the generalized spectral radius theorem, Linear Algebra Appl. 418 (2006) 336–346.

[22] K. Miettinen, Nonlinear Multiobjective Optimization, Kluwer Academic Publishers, Boston, MA, 1999.

[23] A. Miné, Weakly relational numerical abstract domains, Ph.D. Thesis, École Polytechnique, Palaiseau, France, 2004.

[24] R.D. Nussbaum, Convexity and log-convexity for the spectral radius, Linear Algebra Appl. 73 (1986) 59–122.

[25] A. Peperko, On the max version of the generalized spectral radius theorem, Linear Algebra Appl. 428 (2008) 2312–2318.

[26] T.L. Saaty, A scaling method for priorities in hierarchical structures, J. Math. Psychol. 15 (1977) 234–281.

[27] T.L. Saaty, L.G. Vargas, Models, Methods, Concepts and Applications of the Analytic Hierarchy Process, Kluwer Academic Publishers, Norwell, MA, 2001.

[28] S. Sergeev, H. Schneider, P. Butkovič, On visualization scaling, subeigenvectors and Kleene stars in max algebra, Linear Algebra Appl. 431 (2009) 2395–2406.

[29] N.M. Tran, Pairwise ranking: choice of method can produce arbitrarily different rank order, arXiv:1103.1110 (2011).

[30] O.S. Vaidya, S. Kumar, Analytic hierarchy process: an overview of applications, Eur. J. Oper. Res. 169 (2006) 1–29.