The Perron Root of a Weighted Geometric Mean of Nonnegative Matrices

L. ELSNER

Fakultät für Mathematik der Universität Bielefeld, Postfach 86 40, 4800 Bielefeld, Fed. Rep. Germany

CHARLES R. JOHNSON*

Mathematics Department, College of William and Mary, Williamsburg, VA 23185

and

J. A. DIAS DA SILVA

Departamento de Matemática, Universidade de Lisboa, 134 4° Av. 24 de Julho, 1300 Lisboa, Portugal

For k nonnegative n-by-n matrices $A_1, \ldots, A_k$ we consider the matrix

$$C = A_1^{(\alpha_1)} \circ \cdots \circ A_k^{(\alpha_k)},$$

where $\alpha_i > 0$, $i = 1, \ldots, k$, "$\circ$" is the (entry-wise) Hadamard product and $A^{(\alpha)} = (a_{ij}^{\alpha})$ for $A = (a_{ij})$; i.e. C is the component-wise weighted geometric mean of $A_1, A_2, \ldots, A_k$ if $\sum \alpha_i = 1$. It is shown that for $\sum_{i=1}^{k} \alpha_i \geq 1$ the inequality

$$\rho(A_1^{(\alpha_1)} \circ \cdots \circ A_k^{(\alpha_k)}) \leq \rho(A_1)^{\alpha_1} \cdots \rho(A_k)^{\alpha_k}$$

holds. Here ρ denotes the spectral radius. The case of equality is characterized and it is shown that $\rho(C)$, considered as a function of $\alpha = (\alpha_1, \ldots, \alpha_k)$, is convex. This generalizes recent results of Schwenk, and of Karlin-Ost. Similarly, we consider for $A \geq 0$ the comparison matrix $M(A)$, where $M(A)_{ij} = a_{ij}$ for $i = j$, and $= -a_{ij}$ for $i \neq j$. If $\sigma(A)$ denotes the minimal real eigenvalue of $M(A)$ then it is shown that if $\sigma(A_i) > 0$, $i = 1, \ldots, k$ and $\sum_{i=1}^{k} \alpha_i \geq 1$ the dual inequality,

$$\sigma(A_1^{(\alpha_1)} \circ \cdots \circ A_k^{(\alpha_k)}) \geq \sigma(A_1)^{\alpha_1} \cdots \sigma(A_k)^{\alpha_k}$$

holds. Certain other inequalities, some already known, are related to these, and several characterizations are given for another quantity associated with a nonnegative matrix.

1. INTRODUCTION AND DEFINITIONS

The Hadamard (or component-wise) product of two n-by-n matrices $A = (a_{ij})$ and $B = (b_{ij})$ is the n-by-n matrix defined and denoted by $A \circ B = (a_{ij}b_{ij})$. For a component-wise nonnegative matrix $A = (a_{ij})$ (denoted $A \geq 0$), and $\alpha > 0$, we define

$A^{(\alpha)} = (a_{ij}^{\alpha})$. To have continuity in α, for $\alpha = 0$ also, we adopt the convention that $0^0 = 0$. Let $\rho(A)$ denote the spectral radius of A, which, for $A \geqslant 0$, is an eigenvalue (the Perron root) of A. Our first result is the following:

If $A_1, \ldots, A_k$ are nonnegative n-by-n matrices, and $\alpha_1, \ldots, \alpha_k$ are positive numbers such that $\sum_{i=1}^{k} \alpha_i \geqslant 1$, then

$$\rho(A_1^{(\alpha_1)} \circ \cdots \circ A_k^{(\alpha_k)}) \leqslant \rho(A_1)^{\alpha_1} \cdots \rho(A_k)^{\alpha_k}. \tag{1.1}$$

We initially prove the case $k = 2$ and $\alpha_1 + \alpha_2 = 1$ (Lemma 1)

$$\rho(A^{(\alpha)} \circ B^{(1-\alpha)}) \leqslant \rho(A)^{\alpha} \rho(B)^{1-\alpha} \tag{1.2}$$

and then note the case $k = 1$ (Lemma 2)

$$\rho(A^{(t)}) \leqslant \rho(A)^t, \qquad t \geqslant 1. \tag{1.3}$$

Then, the general result follows by induction (Theorem 1).

A special case is $\rho(A \circ B) \leqslant \rho(A)\rho(B)$, $A, B \geqslant 0$ which does not hold for the usual matrix product. Another case is

$$\rho(A^{(\frac{1}{2})} \circ A^{T(\frac{1}{2})}) \leqslant \rho(A) \tag{1.4}$$

which was proved in [12] and [9]. Slightly more general is

$$\rho(A^{(\frac{1}{2})} \circ B^{(\frac{1}{2})}) \leqslant \rho(A)^{\frac{1}{2}} \rho(B)^{\frac{1}{2}} \tag{1.5}$$

which was proved independently and from a point of view different from our work in [10]. The reference [10], which appeared after our proof of (1.1), also alludes, without proof, to the special case of (1.1) in which $\sum \alpha_i = 1$. We then turn to the question of equality in (1.1). In case all A_i are irreducible and $\sum \alpha_i = 1$, the necessary and sufficient condition is that all A_i are essentially diagonally similar, i.e. there exist $\gamma_i > 0$ and positive diagonal matrices D_i such that

$$\gamma_i A_i = D_i^{-1} A_1 D_i, \qquad i = 2, \ldots, k.$$

Using this we are able to give a characterization also in the general case.

After that we derive similar results for M-matrices. For $A \geqslant 0$ we denote by $\sigma(A)$ the minimal real eigenvalue of $M(A)$, where the comparison matrix $M(A)$ is defined by

$$M(A)_{ij} = \begin{cases} a_{ii}, & i = j \\ -a_{ij}, & i \neq j \end{cases}, \qquad i, j = 1, \ldots, n.$$

We prove a result dual to (1.1)

$$\sigma(A_1^{(\alpha_1)} \circ \cdots \circ A_k^{(\alpha_k)}) \geqslant \sigma(A_1)^{\alpha_1} \cdots \sigma(A_k)^{\alpha_k} \tag{1.6}$$

under the condition that $\sigma(A_i) > 0$, $i = 1, \ldots, k$, and $\sum_{i=1} \alpha_i \geqslant 1$.

Also the question of equality is treated. The cases of equality in (1.6) are quite similar to those mentioned above for (1.1).

We also study the question of convexity of the left-hand side of (1.1) as a function of $x_1, x_2, \ldots, x_k$, and derive several related inequalities. We close by showing that several functions of the entries of a nonnegative matrix, some of which have been studied elsewhere, are equal.

This note is a by-product of [9]. There (1.4) was proved in two different ways and the case of equality also discussed in the reducible case. We also observed that the more general inequality (1.1) holds. As the emphasis in [9] was upon various symmetric matrices associated with a nonnegative matrix, we felt that this observation deserved a separate note. Also, although there is an overlap with [10], in which (1.5) is proven, together with a discussion of equality in the irreducible case, we give a much more complete and rather different treatment of (1.1) as well as a number of related results.

We conclude this introduction by mentioning some well known or simple results, which we refer to later on. Use will be made of the Hölder inequality [1] in the form

$$\sum_{i=1}^{n} \xi_i^{\alpha} \eta_i^{1-\alpha} \leqslant \left(\sum_{i=1}^{n} \xi_i\right)^{\alpha} \left(\sum_{i=1}^{n} \eta_i\right)^{1-\alpha},$$

$$\eta_i \geqslant 0, \quad \xi_i \geqslant 0, \quad i = 1, \ldots, n, \quad 0 \leqslant \alpha \leqslant 1 \quad (1.7)$$

where for $0 < \alpha < 1$ equality holds iff $\xi = (\xi_1, \ldots, \xi_n)$ and $\eta = (\eta_1, \ldots, \eta_n)$ are linearly dependent.

In the set of all nonnegative n-by-n matrices we make use of the equivalence relation of diagonal similarity:

$$A \overset{D}{\sim} B \Leftrightarrow \exists \Delta = \mathrm{diag}(\delta_1, \ldots, \delta_n) \text{ with } \delta_i > 0, i = 1, \ldots, n \text{ such that } A = \Delta^{-1} B \Delta.$$

The following facts are easily observed for $A, B \geqslant 0$:

(i) $A \overset{D}{\sim} B, \alpha \geqslant 0 \Rightarrow A^{(\alpha)} \overset{D}{\sim} B^{(\alpha)}$

(ii) $A \overset{D}{\sim} B, C \overset{D}{\sim} E \Rightarrow A \circ C \overset{D}{\sim} B \circ E$ ⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀⠀(1.8)

(iii) $A \overset{D}{\sim} B \Rightarrow \rho(A) = \rho(B)$

(iv) $A \geqslant 0$ irreducible $\Rightarrow \exists B \geqslant 0, B \overset{D}{\sim} A$ and B has constant row sums.

2. THE FIRST INEQUALITY

In this section we prove the following inequality.

THEOREM 1 ⠀ *Let* $A_1, \ldots, A_k$ *be nonnegative n-by-n matrices, and* $\alpha_1, \ldots, \alpha_k$ *positive numbers such that*

$$\sum_{i=1}^{k} \alpha_i \geqslant 1.$$

Then

$$\rho(A_1^{(\alpha_1)} \circ \cdots \circ A_k^{(\alpha_k)}) \leqslant \rho(A_1)^{\alpha_1} \cdots \rho(A_k)^{\alpha_k}. \quad (2.1)$$

The proof of (2.1) is organized as follows: We first give two proofs of (2.1) for the case

$$\sum_{i=1}^{k} \alpha_i = 1.$$

Then we extend this result to

$$\sum_{i=1}^{k} \alpha_i \geq 1$$

with the help of the subsequent Lemma 2.

Let $R_+^k = \{x \in R^k, x_i \geq 0, i = 1, \ldots, k\}$. For $\alpha = (\alpha_1, \ldots, \alpha_k) \in R_+^k$ define

$$A(\alpha) = A_1^{(\alpha_1)} \cdots A_k^{(\alpha_k)}$$

$$\varphi(\alpha) = \rho(A(\alpha))$$

and as long as $\varphi(\alpha) > 0$

$$\psi(\alpha) = \log \varphi(\alpha).$$

The entries $a_{ij}(\alpha)$ of $A(\alpha)$ are either identically zero or log-convex on R_+^k. The result of [11] tells us that either $\varphi \equiv 0$ in R_+^k or $\varphi(\alpha) > 0$ throughout and ψ (and hence φ) is convex in R_+^k. (Although only stated in [11] for functions of one variable, it is obvious that the conclusions are also valid for functions of several variables.) If $\varphi \equiv 0$ then (2.1) holds. If $\varphi > 0$ and $\alpha = (\alpha_1, \ldots, \alpha_k) \in R_+^k$ with $\sum \alpha_i = 1$, then

$$\alpha = \sum_{i=1}^{k} \alpha_i e_i$$

where e_i are the usual unit vectors. Hence

$$\psi(\alpha) \leq \sum_{i=1}^{k} \alpha_i \psi(e_i) \leq \sum_{i=1}^{k} \alpha_i \log \rho(A_i) \tag{2.2}$$

which is just (2.1) in logarithmic form. Here we have used in the second inequality that

$$A(e_i) = A_1^{(0)} \circ \cdots A_i^{(1)} \cdots \circ A_k^{(0)} \leq A_i \quad \text{elementwise}.$$

This approach does not, however, yield the possibility of discussing the case of equality in (2.1).

Hence a second proof follows. It is based on

LEMMA 1   *If $A \geq 0$, $B \geq 0$, $0 \leq \alpha \leq 1$, then*

$$\rho(A^{(\alpha)} \circ B^{(1-\alpha)}) \leq \rho(A)^{\alpha} \rho(B)^{1-\alpha}. \tag{2.3}$$

Observe that this is just the special case $k = 2$ and $\sum \alpha_i = 1$ of (2.1).

Proof of Lemma 1   The inequality (2.3) is obviously true for $\rho(A) = 0$ or $\rho(B) = 0$ and also for $\alpha = 0$ or $\alpha = 1$. Using a continuity argument we may assume that A and B are irreducible and by (1.8) that both have row sums 1. We have to show that under these assumptions

$$C = A^{(\alpha)} \circ B^{(1-\alpha)} \quad \text{satisfies} \quad \rho(C) \leq 1.$$

If $z \geq 0$ is a nonnegative eigenvector of C, $Cz = \rho(C)z$ and k is such that

$$z_k = \max_i z_i,$$

then by (1.7)

$$\rho(C)z_k = (Cz)_k = \sum_{i=1}^{n} c_{ki}z_i = \sum_{i=1}^{n} a_{ki}^\alpha b_{ki}^{1-\alpha} z_i \leqslant \left(\sum_{i=1}^{n} a_{ki}^\alpha b_{ki}^{1-\alpha} \right) z_k$$

$$\leqslant \left(\sum_{i=1}^{n} a_{ki} \right)^\alpha \left(\sum_{i=1}^{n} b_{ki} \right)^{1-\alpha} z_k. \tag{2.4}$$

Hence $\rho(C) \leqslant 1$ and the lemma is proved.

A simple induction leads now to a proof of (2.1) for $\sum \alpha_i = 1$ for general k. Assume that (2.1) holds with k replaced by $k-1$ and define $B \geqslant 0$ by

$$B^{(1-\alpha_k)} = A_1^{(\alpha_1)} \circ \cdots \circ A_{k-1}^{(\alpha_{k-1})};$$

hence with $\beta_v = \alpha_v(1-\alpha_k)^{-1}$, $v = 1, \ldots, k-1$, $\sum_{v=1}^{k-1} \beta_v = 1$ and

$$B = A_1^{(\beta_1)} \circ \cdots \circ A_{k-1}^{(\beta_{k-1})}. \tag{2.5}$$

Now

$$\rho(A(\alpha)) = \rho(B^{(1-\alpha_k)} \circ A_k^{(\alpha_k)}) \leqslant \rho(B)^{1-\alpha_k} \rho(A_k)^{\alpha_k} \quad \text{by (2.3)}$$

$$\leqslant [\rho(A_1)^{\beta_1} \cdots \rho(A_{k-1})^{\beta_{k-1}}]^{1-\alpha_k} \cdot \rho(A_k)^{\alpha_k} \quad \text{by assumption}$$

$$= \rho(A_1)^{\alpha_1} \cdots \rho(A_k)^{\alpha_k}. \tag{2.6}$$

This finishes the second proof of (2.1) in the case $\sum \alpha_i = 1$. The general case then follows from

LEMMA 2 *If $A \geqslant 0$ and $t \geqslant 1$ then*

$$\rho(A^{(t)}) \leqslant \rho(A)^t. \tag{2.7}$$

Proof As usual we need to prove (2.7) only for irreducible A and by (1.8) we may assume that A has row sums 1. But then for $t \geqslant 1$, $A^{(t)} \leqslant A$ (elementwise) and (2.7) holds because of the monotonicity of the Perron root.

We observe that for $t \leqslant 1$ the inequality that reverses (2.7) is valid, i.e. $\rho(A^{(t)}) \geqslant \rho(A)^t$. As for $t < 1$ we have generically the strict inequality, so we see that (2.1) is not true for $\sum \alpha_i < 1$.

We return to the proof of (2.1) and consider the case $t = \sum_{i=1}^{k} \alpha_i \geqslant 1$. Define $\gamma_i = \alpha_i t^{-1}$, $\sum_{i=1}^{k} \gamma_i = 1$. Then by Lemma 2

$$\rho(A(\alpha)) = \rho((A_1^{(\gamma_1)} \circ \cdots \circ A_k^{(\gamma_k)})^{(t)}) \leqslant \rho(A_1^{(\gamma_1)} \circ \cdots \circ A_k^{(\gamma_k)})^t \tag{2.8}$$

and applying the already proved part of Theorem 1 the right-hand side is

$$\leqslant (\rho(A_1)^{\gamma_1} \cdots \rho(A_k)^{\gamma_k})^t = \rho(A_1)^{\alpha_1} \cdots \rho(A_k)^{\alpha_k}. \tag{2.9}$$

Remark By linearization of (2.3), i.e. by replacing A by $I + \varepsilon A$, B by $I + \varepsilon B$, and considering only the terms of first order in ε on both sides of (2.3), we get

$$\rho(C(\alpha, A, B)) \leqslant \alpha \rho(A) + (1-\alpha)\rho(B) \tag{2.10}$$

in which $0 \leqslant \alpha \leqslant 1$, $A, B \geqslant 0$ and $C(\alpha, A, B) = (C_{ij})$ is given by

$$C_{ij} = \begin{cases} \alpha a_{ii} + (1-\alpha)b_{ii}, & i = j \\ a_{ij}^\alpha b_{ij}^{1-\alpha} & i \neq j. \end{cases}$$

A corresponding result for more than two factors follows from (2.1). Observe that (2.10) is wrong if we replace the nondiagonal geometric means by arithmetic means, as in general $\rho(\alpha A + (1 - \alpha)B)$ does not compare with $\alpha\rho(A) + (1 - \alpha)\rho(B)$. In the special case $a_{ij} = b_{ij}$ $(i \neq j)$ the inequality (2.10) is well known and exhibits the convexity of the spectral radius considered as a function of the diagonal. See [3] [5], [7].

3. THE CASE OF EQUALITY

We study now the case of equality in the preceding results. It is obvious that we may assume $\rho(A_i) \neq 0, i = 1, \ldots, k$ as otherwise equality holds trivially. We start with

LEMMA 3   If $0 < \alpha < 1$, and A and B are irreducible nonnegative n-by-n matrices, and

$$\rho(A^{(\alpha)} \circ B^{(1-\alpha)}) = \rho(A)^\alpha \rho(B)^{1-\alpha}, \tag{3.1}$$

then there exists $\gamma > 0$ such that $\gamma A \overset{D}{\sim} B$.

Remark   This result generalizes results in [9] and [10] (and corrects a slight error in the latter).

Proof   As in the proof of Lemma 1 we assume, without loss of generality, that A and B both have row sums 1. Then in (2.4) all inequalities are equalities. By the equality condition in Hölder's inequality (1.7) we get $a_{ki} = \gamma_k b_{ki}, i = 1, \ldots, n$ and by the equality of the row sums $\gamma_k = 1$. In addition $a_{ki} \neq 0$ implies $z_i = z_k$ (= max!): hence, by the irreducibility of A (since each i is connected to k by a path in the graph of A), $z_i = z_k$ for $i = 1, \ldots, n$ and by the reasoning above $A = B$.

Before stating the general equality condition we introduce the following notation. For $\emptyset \neq \mu \subset \{1, 2, \ldots, n\}$, $A[\mu]$ denotes the principal submatrix of A consisting of the rows and columns of A with indices in μ.

THEOREM 2   Let A_i be a nonnegative n-by-n matrix with $\rho(A_i) > 0, \alpha_i > 0, i = 1, \ldots, k$ and

$$\sum_{i=1}^{k} \alpha_i = 1. \tag{3.2}$$

Then the following are equivalent:

(I) equality holds in (2.1), i.e.

$$\rho(A_1^{(\alpha_1)} \circ \cdots \circ A_k^{(\alpha_k)}) = \rho(A_1)^{\alpha_1} \cdots \rho(A_k)^{\alpha_k} \tag{3.3}$$

and

(II) $\exists \mu \subset \{1, \ldots, n\}, \mu \neq \emptyset$ such that the following holds for $i = 1, \ldots, k$.

  (i) $A_i[\mu]$ is irreducible,

  (ii) $\rho(A_i[\mu]) = \rho(A_i)$, \hfill (3.4)

  (iii) $\exists \gamma_i > 0$ such that $\gamma_i A_i[\mu] \overset{D}{\sim} A_1[\mu]$.

Proof Assume that (II) holds. Then

$$\rho(A(\alpha)) \geqslant \rho(A(\alpha)[\mu]) = \gamma_2^{\alpha_2} \cdots \gamma_k^{\alpha_k} \rho(A_1[\mu])$$

by (3.4)(iii) and a short calculation

$$= \rho(A_1[\mu])^{\alpha_1} \cdots \rho(A_k[\mu])^{\alpha_k} = \rho(A_1)^{\alpha_1} \cdots \rho(A_k)^{\alpha_k}$$

by (ii). Together with (2.1) we have (I).

To show that (I) implies (II), first note that $\rho(A(\alpha)) > 0$ by (3.3). Let $E = A(\alpha)$. We infer from well-known results on reducible matrices (see, e.g., [2, p. 39]) that there exists $\mu \neq \varnothing, \mu \subset \{1, \ldots, n\}$ such that $\rho(E) = \rho(E[\mu])$ and $E[\mu]$ is irreducible. From

$$E[\mu] = A_1[\mu]^{(\alpha_1)} \circ \cdots \circ A_k[\mu]^{(\alpha_k)}$$

it follows that (3.4)(i) holds. Also from (3.3) and

$$\rho(E) = \rho(E[\mu]) \leqslant \rho(A_1[\mu])^{\alpha_1} \cdots \rho(A_k[\mu])^{\alpha_k} \leqslant \rho(A_1)^{\alpha_1} \cdots \rho(A_k)^{\alpha_k}$$

we infer (3.4)(ii), because equality must occur in both inequalities. To prove (3.4)(iii) we can now restrict ourselves to the case $\mu = \{1, \ldots, n\}$. Going through the proof of (2.1), equality holds in (2.6). Using Lemma 3 we get that $\bar{\gamma}_k A_k \overset{D}{\sim} C$. But by the induction assumption we have that $\gamma_\nu A_\nu \overset{D}{\sim} A_1$ ($\nu = 1, \ldots, k-1$) and hence $C \overset{D}{\sim} A_1$. This shows $\gamma_k A_k \overset{D}{\sim} A_1$.

An easy consequence of Theorem 2 is the following. If in addition the A_j's are irreducible, then (3.3) is equivalent to

$$\exists \gamma_i > 0 \quad \text{such that} \quad \gamma_i A_i \overset{D}{\sim} A_1, \qquad i = 2, \ldots, k. \tag{3.5}$$

This follows from (3.4), as (3.4)(ii) and A_i irreducible implies $\mu = \{1, \ldots, n\}$.

LEMMA 4   *If $A \geqslant 0$, $\rho(A) > 0$ and $t > 1$, then the following are equivalent:*

(I) $\rho(A^{(t)}) = \rho(A)^t$,

(II) $\exists \varnothing \neq \mu \subset \{1, \ldots, n\}$ *such that*

    (i) $A[\mu]$ *is irreducible,*

    (ii) $\rho(A[\mu]) = \rho(A)$, $\tag{3.6}$

    (iii) $\exists \gamma > 0$ *and a $|\mu|$-by-$|\mu|$-permutation matrix P with $\gamma P \overset{D}{\sim} A[\mu]$.*

Proof By using $P^{(t)} = P$ we see at once that (II) implies (I). Here (i) is not used.

To show (I) $\Rightarrow$ (II) choose any μ such that (II)(i) and (II)(ii) hold. It follows that (I) holds also for $A[\mu]$. Going through the proof of Lemma 2 we infer (II)(iii).

Theorem 2 and Lemma 4 together give a result on equality in (2.1) also in the case $\sum \alpha_i > 1$, which we cite here without proof.

THEOREM   *Under the assumptions of Theorem 2 let*

$$t = \sum_{i=1}^{k} \alpha_i > 1.$$

Then the following are equivalent:

(I) *Equality holds in* (2.1),

(II) $\exists \varnothing \neq \mu \subset \{1, \ldots, n\}$ *such that the following holds for* $i = 1, 2, \ldots, k$:

(i) $A_i[\mu]$ *irreducible*,

(ii) $\rho(A_i[\mu]) = \rho(A_i)$,

(iii) $\exists \gamma_i > 0$ *and a* $|\mu|$-*by*-$|\mu|$-*permutation matrix* P *with* $\gamma_i A_i[\mu] \overset{D}{\sim} P$.

4. GENERALIZATIONS TO *M*-MATRICES

It is possible to generalize the preceding results to results about M-matrices.

THEOREM 4   *Let* A_i *be nonnegative matrices with* $\sigma(A_i) > 0$, $i = 1, \ldots, k$, $\alpha_1, \ldots, \alpha_k$ *positive numbers such that*

$$t = \sum_{i=1}^{k} \alpha_i \geqslant 1.$$

Then

$$\sigma(A_1^{(\alpha_1)} \circ \cdots \circ A_k^{(\alpha_k)}) \geqslant \sigma(A_1)^{\alpha_1} \cdots \sigma(A_k)^{\alpha_k}. \tag{4.1}$$

Proof   (4.1) is a consequence of (1.1) and

$$[D(\alpha) - M(\alpha)]^{-1} \leqslant [(D_1 - M_1)^{-1}]^{(\alpha_1)} \circ \cdots \circ [(D_k - M_k)^{-1}]^{(\alpha_k)} \tag{4.2}$$

where $\alpha = (\alpha_1, \ldots, \alpha_k)$, $A_i = D_i + M_i$, $D_i = $ diagonal of A_i, $i = 1, \ldots, k$ and

$$D(\alpha) = D_1^{\alpha_1} \cdots D_k^{\alpha_k}, \qquad M(\alpha) = M_1^{(\alpha_1)} \circ \cdots \circ M_k^{(\alpha_k)}.$$

Hence we have to prove (4.2) only.

From $D(\alpha)^{-1} M(\alpha) = (D_1^{-1} M_1)^{(\alpha_1)} \circ \cdots \circ (D_k^{-1} M_k)^{(\alpha_k)}$, from $\rho(D_i^{-1} M_i) < 1$, $i = 1, \ldots, k$ and (1.1) we infer $\rho(D(\alpha)^{-1} M(\alpha)) < 1$, hence $M(A(\alpha))$ is an M-matrix. We consider first the case $t = 1$. Define $S = \{\psi: K \to R_+, \psi > 0 \text{ and log-convex}\} \cup \{0\}$ where $K = \{\alpha \in \mathbb{R}^k, \alpha \geqslant 0, \sum \alpha_i \geqslant 1\}$. It was proved in [11] that S is closed under addition, multiplication and limits. The relation

$$[D(\alpha) - M(\alpha)]^{-1} = D(\alpha)^{-1} \sum_{j=0}^{\infty} [M(\alpha) D(\alpha)^{-1}]^j \tag{4.3}$$

shows that the entries of $[D(\alpha) - M(\alpha)]^{-1}$ are in S, and this implies (4.2) for $t = 1$. Here we have used that $[D(e_i) - M(e_i)]^{-1} \leqslant (D_i - M_i)^{-1}$. For $t > 1$ we use the fact that for any nonnegative matrix $A = D + M$, $D = $ diagonal of A, $\sigma(A) > 0$, one has

$$[M(A^{(t)})]^{-1} \leqslant [M(A)^{-1}]^{(t)}. \tag{4.4}$$

This follows by considering the Neumann-Series of both sides of (4.4). As for any set of indices I and nonnegative x_i, $i \in I$

$$\sum_{i \in I} x_i^t \leqslant \left(\sum_{i \in I} x_i \right)^t$$

one gets by applying this inequality to each entry

$$(D^{-1})^t \sum_{j=0}^{\infty} [(MD^{-1})^{(t)}]^j \leqslant \left[D^{-1} \sum_{j=0}^{\infty} (MD^{-1})^j \right]^{(t)}$$

which is just (4.4).

Now let $\gamma_i = \alpha_i t^{-1}$, $\sum_{i=1}^{k} \gamma_i = 1$. Then by (4.4)

$$[D(\alpha) - M(\alpha)]^{-1} \leqslant [(D(\gamma) - M(\gamma))^{-1}]^{(t)}$$

and applying the already proved part of (4.2), the right-hand side is

$$\leqslant \{ [(D_1 - M_1)^{-1}]^{(\gamma_1)} \circ \cdots \circ [(D_k - M_k)^{-1}]^{(\gamma_k)} \}^{(t)}$$

$$= [(D_1 - M_1)^{-1}]^{(\alpha_1)} \circ \cdots \circ [(D_k - M_k)^{-1}]^{(\alpha_k)}.$$

We remark finally that the discussion of equality in (4.1) turns out to be very simple and yields similar results as for the inequality (1.1). If $\sigma(A_i) > 0$, $i = 1, \ldots, k$ and $\sum \alpha_i = 1$ then equality in (4.1) is equivalent to the statements (II) of Theorem 2, with (ii) replaced by "$\sigma(A_i) = \sigma(A_i[\mu])$." Sufficiency follows by a simple calculation, while necessity follows from the step (4.2) → (4.1) and from Theorem 2. Similarly one can prove an analogue of Theorem 3: If $\sum \alpha_i > 1$ then equality in (4.1) is equivalent to the statement: "$\exists \mu \in \{1, \ldots, n\}$ such that $0 \neq \sigma(A_i) = A_i[\mu]$, $i = 1, \ldots, n$." We refrain from giving the details.

Remark Similarly an analog of (2.10) can be proved by linearizing (4.1) for $k = 2$, $\alpha_1 + \alpha_2 = 1$,

$$\sigma(C(\alpha, A, B)) \geqslant \alpha\sigma(A) + (1 - \alpha)\sigma(B). \tag{4.5}$$

5. CONVEXITY

We next give some observations regarding the convexity of the left-hand side of (2.1) as a function of $\alpha_1, \ldots, \alpha_k$. In the second section we have already shown that except for the case $\rho(A(\alpha)) = 0$ the mapping $\alpha \in R_+^k \to \log \rho(A(\alpha))$ is convex, and deduced from that result the inequality

$$\sum_{i=1}^{k} \alpha_i = 1 \quad \Rightarrow \quad \rho(A(\alpha)) \leqslant \prod_{i=1}^{k} \rho(A_i)^{\alpha_i}. \tag{5.1}$$

But also the reverse implication holds: We show that (5.1) implies the convexity of $\psi(\alpha) = \log \rho(A(\alpha))$. If $\alpha = \sum_{v=1}^{r} \gamma_v \mu_v$, $\mu_v \in R_+^k$, $\gamma_v \geqslant 0$, $\sum_{v=1}^{r} \gamma_v = 1$ and $B_v = A(\mu_v)$ then $A(\alpha) = B_1^{(\gamma_1)} \circ \cdots \circ B_r^{(\gamma_r)}$ and applying (5.1) (for r factors) we get $\rho(A(\alpha)) \leqslant \rho(B_1)^{\gamma_1} \cdots \rho(B_r)^{\gamma_r}$, i.e., the convexity of $\psi(\alpha)$. In particular, if $A, B \geqslant 0$ then

$$\tilde{\varphi}(\alpha) = \rho(A^{(\alpha)} \circ B^{(1-\alpha)}) \tag{5.2}$$

and

$$\tilde{\psi}(\alpha) = \log \tilde{\varphi}(\alpha) \tag{5.3}$$

are convex in $\alpha \in [0, 1]$.

The question of strict convexity is discussed in the next theorem. Here a convex

function f is called strictly convex if $s \neq t$, $0 < \alpha < 1$ implies

$$f(\alpha s + (1 - \alpha)t) < \alpha f(s) + (1 - \alpha)f(t).$$

It is tempting to conjecture that $\tilde{\psi}(\alpha)$ is either strictly convex or linear. This is, however, not even true for A, B irreducible. A counterexample is

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & \frac{1}{4} \end{pmatrix}, \qquad B = \begin{pmatrix} \frac{1}{4} & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}$$

for which

$$\tilde{\psi}(\alpha) = \begin{cases} (1 - 3\alpha)\log 2, & 0 < \alpha \leq \frac{1}{3} \\ 0, & \frac{1}{3} \leq \alpha \leq \frac{2}{3} \\ (3\alpha - 2)\log 2, & \frac{2}{3} \leq \alpha < 1 \end{cases}.$$

Hence we need obviously a further condition:

THEOREM 5 *Assume that for some* $\beta \in (0, 1)$, $A^{(\beta)} \circ B^{(1 - \beta)}$ *is irreducible. Then exactly one of the following properties holds*:

(i) $\tilde{\psi}$ *is linear on* $[0, 1]$,

(ii) $\tilde{\psi}$ *is strictly convex on* $[0, 1]$.

Proof It suffices to show that $\tilde{\psi}$ not strictly convex implies that $A \overset{D}{\sim} \gamma B$ for some γ, because then by Theorem 2 $\tilde{\psi}$ is linear. Hence assume that there are $0 < s < t < 1$, $0 < \alpha < 1$ such that

$$\tilde{\psi}(\alpha s + (1 - \alpha)t) = \alpha\tilde{\psi}(s) + (1 - \alpha)\tilde{\psi}(t).$$

But as $A^{(\beta)} \circ B^{(1 - \beta)}$ is irreducible, so are $A^{(s)} \circ B^{(1 - s)}$ and $A^{(t)} \circ B^{(1 - t)}$. By Lemma 3 there exists $\bar{\gamma} > 0$ such that $\bar{\gamma}A^{(s)} \circ B^{(1 - s)} \overset{D}{\sim} A^{(t)} \circ B^{(1 - t)}$. This implies $A \overset{D}{\sim} \gamma B$ for suitable γ.

The case considered in Theorem 5 is not so special as it seems. The zero-pattern of $A^{(\beta)} \circ B^{(1 - \beta)}$ is constant throughout $(0, 1)$.

Hence there is a partition $\mu_1 \cup \cdots \cup \mu_r = \{1, \ldots, n\}$ such that $A^{(\beta)}[\mu_i] \circ B^{(1 - \beta)}[\mu_i]$ are irreducible or one-dimensional zero blocks for $0 < \beta < 1$ and $i = 1, \ldots, r$.

Define $\tilde{\psi}_i(\beta) = \log \rho((A^{(\beta)} \circ B^{(1 - \beta)})[\mu_i])$, then $\tilde{\psi}(\beta) = \text{Max}_i \, \tilde{\psi}_i(\beta)$ and each of the $\tilde{\psi}_i$ satisfies the assumptions of Theorem 5. Results analogous to Theorem 5 hold for $\tilde{\varphi}(\alpha)$.

THEOREM 6 *Under the conditions of Theorem 5 exactly one of the following properties holds*:

(i) $\tilde{\varphi}$ *is constant in* $[0, 1]$,

(ii) $\tilde{\varphi}$ *is strictly convex in* $[0, 1]$.

6. AN EIGENVALUE LIMIT

We conclude with some observations relating quantities suggested here with some quantities already occurring in the literature.

Let $A = (a_{ij}) \geqslant 0$ be an n-by-n nonnegative matrix. It is a consequence of Lemma 2 that

$$r \geqslant s \quad \Rightarrow \quad \rho(A^{(r)})^{1/r} \leqslant \rho(A^{(s)})^{1/s}.$$

Hence

$$\mu(A) = \lim_{r \to \infty} \rho(A^{(r)})^{1/r} \tag{6.1}$$

exists. This was established in [10] and a probabilistic interpretation of $\mu(A)$ was given in the case that A is a stochastic matrix. In [8] it was shown that $\mu(A)$ is equal to the functional $h(A)$, defined below in (6.5) and hence depends on cycle products from A. Earlier, it had been shown in [6] that $h(A)$ (and hence $\mu(A)$) is equal to $g(A)$, defined below in (6.4). Here we add another characterization namely

$$\mu(A) = \max_{\substack{B \geqslant 0 \\ \rho(B) > 0}} \frac{\rho(A \circ B)}{\rho(B)}. \tag{6.2}$$

In the subsequent Theorem 7 we prove all the above-mentioned results together. We feel that this procedure simplifies the proofs in the literature and stresses the interdependence of the results.

We introduce

$$f(A) = \max\{\rho(A \circ B): B \geqslant 0, \rho(B) \leqslant 1\} \tag{6.3}$$

$$g(A) = \inf\left\{\max_{i,k} a_{ik} \frac{d_i}{d_k}: d_j > 0, j = 1, \dots, n\right\}. \tag{6.4}$$

Let $\tilde{\Omega}_n$ be the set of all paths in the complete graph with vertices $\{1, \dots, n\}$, i.e.,

$$w \in \tilde{\Omega}_n \quad \leftrightarrow \quad w = \{i_1, i_2, \dots, i_k, i_{k+1}\}, \quad i_j \in \{1, \dots, n\}.$$

$|w| = k$ is the length of the path and $\prod_w (A) = a_{i_1 i_2} a_{i_2 i_3} \cdots a_{i_k i_{k+1}}$. Ω_n is the set of all cycles, i.e., paths where $i_1 = i_{k+1}$, and Ω_n^s the set of all simple cycles, i.e.,

$$w \in \Omega_n^s \quad \leftrightarrow \quad w = \{i_1, \dots, i_k, i_1\} \quad \text{and} \quad i_j \neq i_l \text{ if } j \neq l, l = 1, \dots, k.$$

Define

$$h(A) \equiv \max\left\{\prod_w (A)^{1/|w|}, w \in \Omega_n^s\right\}$$

$$= \max\left\{\prod_w (A)^{1/|w|}, w \in \Omega_n\right\}. \tag{6.5}$$

Here the second equality is easy to prove, as each cycle is the union of simple cycles. We are now ready to state

THEOREM 7 *If A is a nonnegative n-by-n matrix, then $\mu(A) = f(A) = g(A) = h(A)$.*

Proof We establish step by step the following chain of inequalities:

$$\mu \leqslant f \leqslant g \leqslant h \leqslant \mu.$$

(i) $\mu \leqslant f$: If $f(A) = 0$, then also $\mu(A) = 0$. Hence assume now $f(A) > 0$. Define

$\rho_r = \rho(A^{(r)})$. By definition of $f = f(A)$:

$$\rho_r \leqslant f\rho_{r-1},\qquad\qquad (6.6)$$

hence $\rho_r f^{-r} \leqslant \rho_{r-1} f^{-r+1} \leqslant \cdots \leqslant \rho_1 f^{-1}$, which implies

$$f^{-1}\mu(A) = \lim_{r\to\infty} (\rho_r f^{-r})^{1/r} \leqslant 1.$$

(ii) $f \leqslant g$: For any diagonal matrix D with positive diagonal entries $d_1, \ldots, d_n$ and any $B \geqslant 0$, we have

$$\rho(A \circ B) = \rho(DAD^{-1} \circ B) \leqslant \max_{i,k} a_{ik}\, \frac{d_i}{d_k} \cdot \rho(B)$$

and hence $f(A) \leqslant g(A)$.

(iii) $g \leqslant h$: Here we use a technique introduced in [6] and also used in [8] to prove a slightly different result. We may assume here $\rho(A) > 0$ and A irreducible, the general case following by either a continuity argument or by considering the irreducible components of A in its reducible normal form.

We may also assume $h(A) = 1$. We show that there exists a diagonal scaling such that

$$a_{ij}\frac{d_i}{d_j} \leqslant 1, \qquad i, j = 1, \ldots, n. \qquad\qquad (6.7)$$

Define $d_1 = 1$ and

$$d_j = \max\left\{\prod_w (A),\ w = \{i_1, \ldots, i_k\} \in \tilde{\Omega}_n,\ i_1 = 1,\ i_k = j\right\}.$$

As all cyclic products are bounded by one this maximum exists and is positive due to the irreducibility of A. For $i = j$, (6.7) is obvious. Let $i \neq j$ and $w = \{1, \ldots, i\} \in \tilde{\Omega}_n$ such that $\prod_w (A) = d_i$. Let $w' = \{1, \ldots, i, j\}$. Then

$$d_j \geqslant \prod_{w'} (A) = a_{ij}d_i;$$

hence (6.7) holds.

(iv) $h \leqslant \mu$: Let $w \in \Omega_n^s$ be chosen such that $h(A) = \prod_w (A)^{1/|w|}$. Replacing all entries of A not appearing in $\prod_w (A)$ by zero, we obtain a matrix $A_0 \leqslant A$. Then $h(A) = \rho(A_0) = \rho(A_0^{(r)})^{1/r} \leqslant \rho(A^{(r)})^{1/r}$ for $r > 0$. Hence $h(A) \leqslant \mu(A)$.

Remark From $\mu(A) = h(A)$ it is easy to conclude that also

$$\mu(A) = \limsup\left\{\prod_w (A)^{1/|w|},\ w \in \tilde{\Omega}_n\right\}, \qquad\qquad (6.8)$$

the largest accumulation point of this countable set of real numbers. [Statement (6.8) has been established in [10] for A stochastic.] Thus, the right-hand side of (6.8) may also be added to the list in Theorem 7.

We list some simple properties of the functional μ:

$$\mu(A \circ B) \leqslant \mu(A) \cdot \mu(B) \qquad \text{by (6.2)} \tag{6.9}$$

$$\mu(A^{(t)}) = \mu(A)^t \qquad (\text{as } \mu = h) \tag{6.10}$$

$$\mu(A) \leqslant \rho(A) \leqslant n\mu(A). \tag{6.11}$$

Here the first inequality is obvious and equality holds iff $\rho(A^{(t)}) = \rho(A)$ for some (and hence all) $t > 1$ (see Lemma 4 for a description of the structure of A in this case). The second inequality is a consequence of (6.2) and can also be found in [8], where also the case of equality is characterized; equality occurs exactly when A is a (positive) rank 1 matrix, all of whose diagonal entries are equal. As simple examples show, neither $\mu(AB) \leqslant \mu(A)\mu(B)$ nor $\mu(A + B) \leqslant \mu(A) + \mu(B)$ are correct in general.

References

[1] E. F. Beckenbach and R. Bellman, *Inequalities*, Springer-Verlag, Berlin-Heidelberg-New York, 1965.
[2] A. Berman and R. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic Press, New York, 1979.
[3] J. E. Cohen, Random evolutions and the spectral radius of a non-negative matrix, *Math. Proc. Cambridge Phil. Soc.* **86** (1979), 345–370.
[4] E. Deutsch and M. Neumann, Derivatives of the Perron root at an essentially non-negative matrix and the group inverse of an M-matrix, *J. Math. Anal. Appl.* **102** (1984), 1–29.
[5] L. Elsner, On convexity properties of the spectral radius of nonnegative matrices, *Lin. Alg. Appl.* **61** (1984), 31–35.
[6] E. M. Engel and H. Schneider, Diagonal similarity and equivalence for matrices over groups with 0, *Czech. Math. J.* **25** (1975), 389–403.
[7] S. Friedland, Convex spectral functions, *Lin. Multilin. Alg.* **9** (1981), 299–316.
[8] S. Friedland, Limit eigenvalues of nonnegative matrices, *Lin. Alg. Appl.* **74** (1986), 173–178.
[9] C. R. Johnson and J. A. Dias da Silva, Symmetric matrices associated with a nonnegative matrix, *Circuits, Systems and Signal Processing*, to appear.
[10] S. Karlin and F. Ost, Some monotonicity properties of Schur powers of matrices and related inequalities, *Lin. Alg. Appl.* **68** (1985), 47–65.
[11] J. F. C. Kingman, A convexity property of positive matrices, *Quart. J. Math. Oxford, Ser. 2* **12** (1961), 283–284.
[12] A. Schwenk, Tight bounds on the spectral radius of asymmetric non-negative matrices, *Lin. Alg. Appl.* **75** (1986), 257–265.