The Perron Root of a Weighted Geometric Mean of Nonnegative Matrices

L. ELSNER

Fakultät für Mathematik der Universität Bielefeld, Postfach 86 35 00, 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 4ª Av. 24 de Julho, 1300 Lisboa, Portugal

(Received April 13, 1987; in final form August 18, 1987)

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

$$C = A_1^{x_1} \circ \cdots \circ A_k^{x_k},$$

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

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

the inequality

$$\rho(A_1^{x_1} \circ \cdots \circ A_k^{x_k}) \leq \rho(A_1)^{x_1} \cdots \rho(A_k)^{x_k}$$

holds. Here $\rho$ denotes the spectral radius. The case of equality is characterized and it is shown that $\rho(C)$, considered as a function of $x = (x_1, \ldots, x_k)$, 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 $\sum_{i=1}^{k} x_i \geq 1$ the dual inequality,

$$\sigma(A_1^{x_1} \circ \cdots \circ A_k^{x_k}) \geq \sigma(A_1)^{x_1} \cdots \sigma(A_k)^{x_k}$$

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

---

* The work of this author was supported in part by National Science Foundation grant DMS-8713762 and by Office of Naval Research contracts N00014-86-K-0012 and N00014-87-K-0661.
$A^{(a)} = (a_{ij}^x)$. To have continuity in $x$, for $x = 0$ also, we adopt the convention that $0^0 = 0$. Let $\rho(A)$ denote the spectral radius of $A$, which, for $A \geq 0$, 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 $\sum_{i=1}^k \alpha_i \geq 1$, then

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

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

$$\rho(A^{(\alpha)} \circ B^{(1 - \alpha)}) \leq \rho(A)^{\alpha} \rho(B)^{1 - \alpha} \quad (1.2)$$

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

$$\rho(A^{(\alpha)}) \leq \rho(A)^t, \quad t \geq 1. \quad (1.3)$$

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

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

$$\rho(A^{(\alpha)} \circ A^{(\alpha)}) \leq \rho(A) \quad (1.4)$$

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

$$\rho(A^{(\alpha)} \circ B^{(\alpha)}) \leq \rho(A)^{\alpha} \rho(B)^{\alpha} \quad (1.5)$$

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

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

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

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

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

We prove a result dual to (1.1)

$$\sigma(A_1^{(\alpha_1)} \circ \cdots \circ A_k^{(\alpha_k)}) \geq \sigma(A_1)^{\alpha_1} \cdots \sigma(A_k)^{\alpha_k} \quad (1.6)$$

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

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
\[
\sum_{i=1}^{n} \xi_i \eta_i^{1-\alpha} \leq \left( \sum_{i=1}^{n} \xi_i \right)^{\alpha} \left( \sum_{i=1}^{n} \eta_i \right)^{1-\alpha} \quad \eta_i \geq 0, \quad i = 1, \ldots, n, \quad 0 \leq \alpha \leq 1 \quad (1.7)
\]
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:
\[
A \sim B \iff \exists \Delta = \text{diag}(\delta_1, \ldots, \delta_n) \text{ with } \delta_i > 0, i = 1, \ldots, n \text{ such that } A = \Delta^{-1} B \Delta.
\]
The following facts are easily observed for \(A, B \geq 0\):

(i) \(A \sim B, \alpha \geq 0 \Rightarrow A^{(\alpha)} \sim B^{(\alpha)}\)

(ii) \(A \sim B, C \sim E \Rightarrow A \circ C \sim B \circ E\)

(iii) \(A \sim B \Rightarrow \rho(A) = \rho(B)\)

(iv) \(A \geq 0 \text{ irreducible} \Rightarrow \exists B \geq 0, B \sim A \text{ and } B \text{ has constant row sums}.

\[
(1.8)
\]

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
\[
\sum_{i=1}^{k} \alpha_i \geq 1.
\]

Then
\[
\rho(A_1^{(\alpha_1)} \cdots A_k^{(\alpha_k)}) \leq \rho(A_1)^{\alpha_1} \cdots \rho(A_k)^{\alpha_k}. \quad (2.1)
\]

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
\[ \sum_{i=1}^{k} x_i \geq 1 \]
with the help of the subsequent Lemma 2.

Let \( R_+^k = \{ x \in R^k, x_i \geq 0, i = 1, \ldots, k \} \). For \( x = (x_1, \ldots, x_k) \in R_+^k \), define
\[ A(x) = A_1^{(x_1)} \cdots A_k^{(x_k)} \]
\[ \varphi(x) = \rho(A(x)) \]
and as long as \( \varphi(x) > 0 \)
\[ \psi(x) = \log \varphi(x). \]
The entries \( a_{ij}(x) \) of \( A(x) \) 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 \( \varphi(x) > 0 \) throughout and \( \psi \) (and hence \( \varphi \)) 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 \( x = (x_1, \ldots, x_k) \in R_+^k \) with \( \sum x_i = 1 \), then
\[ x = \sum_{i=1}^{k} x_i e_i \]
where \( e_i \) are the usual unit vectors. Hence
\[ \psi(x) \leq \sum_{i=1}^{k} x_i \psi(e_i) \leq \sum_{i=1}^{k} x_i \log \rho(A_i) \]
(2.2)
which is just (2.1) in logarithmic form. Here we have used in the second inequality that
\[ A(e_i) = A_1^{(0)} \cdots A_i^{(1)} \cdots A_k^{(0)} \leq A_i \ 	ext{elementwise}. \]
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
\[ \rho(A^{(\alpha)} \circ B^{(1-\alpha)}) \leq \rho(A)^{\alpha} \rho(B)^{1-\alpha}. \]
(2.3)
Observe that this is just the special case \( k = 2 \) and \( \sum x_i = 1 \) 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
\[ C = A^{(\alpha)} \circ B^{(1-\alpha)} \] satisfies \( \rho(C) \leq 1 \).

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)

\[
\rho(C)z_k = (Cz)_k = \sum_{i=1}^{n} c_{ki}z_i = \sum_{i=1}^{n} a_{ki}^2b_{ki}^{1-\alpha}z_i \leq \left( \sum_{i=1}^{n} a_{ki}^2b_{ki}^{1-\alpha} \right)z_k
\]

\[
\leq \left( \sum_{i=1}^{n} a_{ki} \right)^2 \left( \sum_{i=1}^{n} b_{ki} \right)^{1-\alpha}z_k.
\]

(2.4)

Hence \( \rho(C) \leq 1 \) 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 \( B = A^{(1-\alpha_k)} \).

\[ B^{(1-\alpha_k)} = A^{(\alpha_1)} \circ \cdots \circ A^{(\alpha_{k-1})}; \]

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

\[ B = A^{(\beta_1)} \circ \cdots \circ A^{(\beta_{k-1})}. \]

(2.5)

Now

\[
\rho(A(x)) = \rho(B^{(1-\alpha_k)} \circ A^{(\alpha_k)}) \leq \rho(B)^{1-\alpha_k} \rho(A)^{\alpha_k} \quad \text{by (2.3)}
\]

\[
\leq \left[ \rho(A_1)^{\gamma_1} \cdots \rho(A_{k-1})^{\gamma_{k-1}} \right]^{1-\alpha_k} \cdot \rho(A_1)^{\alpha_k} \quad \text{by assumption}
\]

\[
= \rho(A_1)^{\gamma_1} \cdots \rho(A_k)^{\gamma_k}.
\]

(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 \( A \geq 0 \) and \( t \geq 1 \) then

\[
\rho(A^{(t)}) \leq \rho(A)^t.
\]

(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 \( t \geq 1 \), \( A^{(t)} \leq A \) (elementwise) and (2.7) holds because of the monotonicity of the Perron root.

We observe that for \( t \leq 1 \) the inequality that reverses (2.7) is valid, i.e. \( \rho(A^{(t)}) \geq \rho(A)^t \). 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 \( t = \sum_{i=1}^{k} \alpha_i \geq 1 \). Define \( \gamma_i = \alpha_i t^{-1}, \sum_{i=1}^{k} \gamma_i = 1 \). Then by Lemma 2

\[
\rho(A^{(t)}) = \rho((A^{(\gamma_1)} \circ \cdots \circ A^{(\gamma_k)})^{(t)}) \leq \rho(A^{(\gamma_1)} \circ \cdots \circ A^{(\gamma_k)})^t
\]

(2.8)

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

\[
\leq (\rho(A_1)^{\gamma_1} \cdots \rho(A_k)^{\gamma_k})^t = \rho(A_1)^{\gamma_1} \cdots \rho(A_k)^{\gamma_k}.
\]

(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 \( \varepsilon \) on both sides of (2.3), we get

\[
\rho(C(x, A, B)) \leq \varepsilon \rho(A) + (1 - \varepsilon) \rho(B)
\]

(2.10)

in which \( 0 \leq \varepsilon \leq 1, A, B \geq 0 \) and \( C(x, A, B) = (C_{ij}) \) is given by

\[
C_{ij} = \begin{cases} 
\alpha a_{ii} + (1-\alpha)b_{ii}, & i = j \\
\alpha^2 b_{ij}^{1-\alpha} & i \neq j.
\end{cases}
\]
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 \( \rho(xA + (1-x)B) \) does not compare with \( x\rho(A) + (1-x)\rho(B) \). 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 \( \rho(A_{ij}) \neq 0, i = 1, \ldots, k \) as otherwise equality holds trivially. We start with

**Lemma 3** If \( 0 < x < 1 \), and \( A \) and \( B \) are irreducible nonnegative \( n \)-by-\( n \) matrices, and

\[
\rho(A^{[\alpha]} \odot B^{[1-\alpha]}) = \rho(A)^\alpha \rho(B)^{1-\alpha},
\]

(3.1)

then there exists \( \gamma > 0 \) such that \( \gamma A \sim B \).

**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 \( \mu \).

**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

\[
\sum_{i=1}^{k} \alpha_i = 1.
\]

(3.2)

Then the following are equivalent:

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

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

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

(3.4)

(iii) \( \exists \gamma_i > 0 \) such that \( \gamma_i A_i[\mu] \sim A_i[\mu] \).
Proof Assume that (II) holds. Then

$$
\rho(A(x)) \geq \rho(A(x)[\mu]) = \gamma_2^{x_2} \cdots \gamma_k^{x_k} \rho(A_1[\mu])
$$

by (3.4)(iii) and a short calculation

$$
\rho(A_1[\mu])^{x_1} \cdots \rho(A_k[\mu])^{x_k} = \rho(A_1)^{x_1} \cdots \rho(A_k)^{x_k}
$$

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

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

$$
E[\mu] = A_1[\mu]^{(x_1)} \cdots A_k[\mu]^{(x_k)}
$$

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

$$
\rho(E) = \rho(E[\mu]) \leq \rho(A_1[\mu])^{x_1} \cdots \rho(A_k[\mu])^{x_k} \leq \rho(A_1)^{x_1} \cdots \rho(A_k)^{x_k}
$$

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 $$\tilde{\gamma}_k A_k \sim C$$. But by the induction assumption we have that $$\gamma_v A_v \sim A_1 (v = 1, \ldots, k - 1)$$ and hence $$C \sim A_1$$.

This shows $$\gamma_k A_k \sim A_1$$.

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

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

(3.5)

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

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

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

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

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

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

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

**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 $$\mu$$ 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) \( \exists \emptyset \neq \mu \subset \{1, \ldots, n\} \) 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 \( \gamma_i A_i[\mu] \sim P \).

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 \( \sigma(A_i) > 0 \), \( i = 1, \ldots, k \), \( x_1, \ldots, x_k \) positive numbers such that

\[
t = \sum_{i=1}^{k} x_i > 1.
\]

Then

\[
\sigma(A_1^{x_1} \cdots A_k^{x_k}) \geq \sigma(A_1 x_1 \cdots A_k x_k).
\]  

(4.1)

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

\[
[D(x) - M(x)]^{-1} \preceq [(D_1 - M_1)^{-1}]^{(x_1)} \cdots [(D_k - M_k)^{-1}]^{(x_k)}
\]

where \( x = (x_1, \ldots, x_k) \), \( A_i = D_i + M_i \), \( D_i \) diagonal of \( A_i, \ i = 1, \ldots, k \) and

\[
D(x) = D_1^{x_1} \cdots D_k^{x_k}, \quad M(x) = M_1^{(x_1)} \cdots M_k^{(x_k)}.
\]

Hence we have to prove (4.2) only.

From \( D(x)^{-1}M(x) = (D_1^{-1}M_1)^{(x_1)} \cdots (D_k^{-1}M_k)^{(x_k)} \), from \( \rho(D_i^{-1}M_i) < 1 \), \( i = 1, \ldots, k \) and (1.1) we infer \( \rho(D(x)^{-1}M(x)) < 1 \), hence \( M(A(x)) \) is an \( M \)-matrix. We consider first the case \( t = 1 \). Define \( S = \{ \psi : K \to R^+, \ \psi > 0 \) and log-convex \} \cup \{0\} \) where \( K = (x_i \in R^k, x > 0, \Sigma x_i \geq 1 \}. \) It was proved in [11] that \( S \) is closed under addition, multiplication and limits. The relation

\[
[D(x) - M(x)]^{-1} = D(x)^{-1} \sum_{j=0}^{\infty} [M(x)D(x)^{-1}]^j
\]

(4.3)

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

\[
[M(A^{(0)})]^{-1} \leq [M(A)]^{-1}^{(0)}.
\]

(4.4)

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

\[
\sum_{i \in I} x_i^t \leq \left( \sum_{i \in I} x_i \right)^t
\]


one gets by applying this inequality to each entry

\[ (D^{-1} y) \sum_{j=0}^{\infty} [(MD^{-1} y)^{\ell}] j \leq \left[ D^{-1} \sum_{j=0}^{\infty} (MD^{-1} y)^{\ell} \right]^{\ell} \]

which is just (4.4).

Now let \( \gamma_i = \alpha_i t^{-1}, \sum_i \gamma_i = 1 \). Then by (4.4)

\[ [D(\gamma) - M(\gamma)]^{-1} \leq [(D(\gamma) - M(\gamma))^{-1}]^{\ell} \]

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

\[ \leq [(D_1 - M_1)^{-1}]^{(\gamma_1)} \cdots [(D_k - M_k)^{-1}]^{(\gamma_k)} \]

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

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) \( \rightarrow \) (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 \( k = 2 \), \( \alpha_1 + \alpha_2 = 1 \),

\[ \sigma(C(\alpha, A, B)) \geq \alpha \sigma(A) + (1 - \alpha) \sigma(B). \] (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 \( \alpha \in R^k \rightarrow \log \rho(A(\alpha)) \) is convex, and deduced from that result the inequality

\[ \sum_{i=1}^{k} \alpha_i = 1 \Rightarrow \rho(A(\alpha)) \leq \prod_{i=1}^{k} \rho(A_i)^{\alpha_i}. \] (5.1)

But also the reverse implication holds: We show that (5.1) implies the convexity of \( \psi(\alpha) = \log \rho(A(\alpha)) \). If \( \alpha = \sum_{\gamma=1}^{r} \gamma_\nu \mu_\nu, \mu_\nu \in R^+, \gamma_\nu \geq 0, \sum_{\gamma=1}^{r} \gamma_\nu = 1 \) and \( B_\nu = A(\mu_\nu) \) then \( A(\alpha) = B_1^{(\gamma_1)} \cdots B_r^{(\gamma_r)} \) and applying (5.1) (for \( r \) factors) we get \( \rho(A(\alpha)) \leq \rho(B_1)^{\gamma_1} \cdots \rho(B_r)^{\gamma_r}, \) i.e., the convexity of \( \psi(\alpha) \). In particular, if \( A, B_\nu \geq 0 \) then

\[ \tilde{\phi}(\alpha) = \rho(A(\alpha) \circ B(1-\alpha)) \] (5.2)

and

\[ \tilde{\psi}(\alpha) = \log \tilde{\phi}(\alpha) \] (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
\[
A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1/2 \end{pmatrix}, \quad B = \begin{pmatrix} 1/4 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}
\]

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) $\psi$ is strictly convex on $[0, 1]$.

**Proof** It suffices to show that $\tilde{\psi}$ not strictly convex implies that $A \sim_B^D \gamma B$ for some $\gamma$, because then by Theorem 2 $\psi$ 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 $\gamma > 0$ such that $\gamma A^{(s)} \circ B^{(1 - s)} \sim^D A^{(t)} \circ B^{(1 - t)}$. This implies $A \sim_B^D \gamma B$ for suitable $\gamma$.

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 $\tilde{\psi}(\beta) = \max_i \tilde{\psi}_i(\beta)$ and each of the $\tilde{\psi}_i$ satisfies the assumptions of Theorem 5. Results analogous to Theorem 5 hold for $\tilde{\psi}(\alpha)$.

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

(i) $\tilde{\psi}$ is constant in $[0, 1]$,
(ii) $\tilde{\psi}$ 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 $A = (a_{ij}) \geq 0$ be an $n$-by-$n$ nonnegative matrix. It is a consequence of Lemma 2 that

$$r \geq s \Rightarrow \rho(A^{(r)})^{1/r} \leq \rho(A^{(s)})^{1/s}.$$  

Hence

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

$$\mu(A) = \max_{B \geq 0, \rho(B) > 0} \frac{\rho(A \circ B)}{\rho(B)}.$$  \hspace{1cm} (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

$$f(A) = \max\{\rho(A \circ B); B \geq 0, \rho(B) \leq 1\}$$  \hspace{1cm} (6.3)

$$g(A) = \inf \left\{ \max_{i,k} \frac{d_i}{d_k}; d_j > 0, j = 1, \ldots, n \right\}.$$  \hspace{1cm} (6.4)

Let $\Omega_n$ be the set of all paths in the complete graph with vertices $\{1, \ldots, n\}$, i.e.,

$$w \in \Omega_n \leftrightarrow w = \{i_1, i_2, \ldots, i_k, i_{k+1}\}, \quad i_j \in \{1, \ldots, n\}.$$

$|w| = k$ is the length of the path and $\prod_w (A) = a_{i_1i_2}a_{i_2i_3} \ldots a_{i_{k}i_{k+1}}$. $\Omega_n$ is the set of all cycles, i.e., paths where $i_1 = i_{k+1}$, and $\Omega_n^* \subset \Omega_n$ the set of all simple cycles, i.e.,

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

Define

$$h(A) = \max\left\{\prod_w (A)^{1/|w|}, w \in \Omega_n^*\right\} = \max\left\{\prod_w (A)^{1/|w|}, w \in \Omega_n\right\}.$$  \hspace{1cm} (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:

$$\mu \leq f \leq g \leq h \leq \mu.$$

(i) $\mu \leq f$: If $f(A) = 0$, then also $\mu(A) = 0$. Hence assume now $f(A) > 0$. Define
\[ \rho_r = \rho(A^r). \] By definition of \( f = f(A) \):

\[ \rho_r \leq f \rho_{r-1}, \]  

(6.6)

hence \( \rho_r f^{-r} \leq \rho_{r-1} f^{-r+1} \leq \cdots \leq \rho_1 f^{-1} \), which implies

\[ f^{-1} \mu(A) = \lim_{r \to \infty} (\rho_r f^{-r})^{1/r} \leq 1. \]

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

\[ \rho(A \circ B) = \rho(DAD^{-1} \circ B) \leq \max_{i,k} \frac{d_i}{d_k} \cdot \rho(B) \]

and hence \( f(A) \leq g(A) \).

(iii) \( g \leq h \): 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

\[ a_{ij} \frac{d_i}{d_j} \leq 1, \quad i, j = 1, \ldots, n. \]  

(6.7)

Define \( d_1 = 1 \) and

\[ d_j = \max \left\{ \prod_w (A), \ w = \{i_1, \ldots, i_k\} \in \Omega_n, \ i_1 = 1, \ i_k = j \right\}. \]

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 \( w = \{1, \ldots, i\} \in \Omega_n \) such that \( \prod_w (A) = d_i \). Let \( w' = \{1, \ldots, i, j\} \). Then

\[ d_j \geq \prod_{w'} (A) = a_{ij}d_i; \]

hence (6.7) holds.

(iv) \( h \leq \mu \): Let \( w \in \Omega_n^* \) 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 \( A_0 \leq A \). Then \( h(A) = \rho(A_0) = \rho(A_0^{(r)})^{1/r} \leq \rho(A^{(r)})^{1/r} \) for \( r > 0 \). Hence \( h(A) \leq \mu(A) \).

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

\[ \mu(A) = \limsup \left\{ \prod_w (A)^{1/|w|}, \ w \in \Omega_n \right\}, \]  

(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 $\mu$:
\[\mu(A \circ B) \leq \mu(A) \cdot \mu(B)\quad\text{by (6.2)}\]  
\[\mu(A^t) = \mu(A)^t\quad\text{(as $\mu = h$)}\]  
\[\mu(A) \leq \rho(A) \leq \eta\mu(A).\]

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 $\mu(AB) \leq \mu(A)\mu(B)$ nor $\mu(A + B) \leq \mu(A) + \mu(B)$ are correct in general.

References