On Convexity Properties of the Spectral Radius of Nonnegative Matrices

L. Elsner
Fakultät für Mathematik
Universität Bielefeld
Postfach 86 40
4800 Bielefeld 1, Federal Republic of Germany

Submitted by Hans Schneider

ABSTRACT

Elementary matrix-theoretic proofs are given for the following well-known results: $r(D)=\max\{\operatorname{Re}\lambda:\lambda$ an eigenvalue of $A+D\}$ and $s(D)=\ln\rho(e^{D}A)$ are convex. Here D is diagonal, A a nonnegative $n\times n$ matrix, and ρ the spectral radius.

---

1. INTRODUCTION

In this note we give new proofs of two recent results which can be formulated as follows: Let $\mathscr{D}_n$ denote the set of real $n\times n$ diagonal matrices, and $I\in\mathscr{D}_n$ the unit matrix. A function $\varphi:\mathscr{D}_n\to\mathbf{R}$ is *convex* if

$$\varphi(\alpha D_1+(1-\alpha)D_2)\leqslant\alpha\varphi(D_1)+(1-\alpha)\varphi(D_2)\tag{1}$$

holds for $0\leqslant\alpha\leqslant 1$, $D_i\in\mathscr{D}_n$, $i=1,2$. φ is *s-convex* if it is convex and for $0<\alpha<1$ equality in (1) holds iff D_1-D_2 is a multiple of I.

Let $A=(a_{ij})\geqslant 0$ be a fixed nonnegative $n\times n$ matrix. Denote by $\rho(B)$ the spectral radius of a matrix B.

THEOREM 1. *Define* $r:\mathscr{D}_n\to\mathbf{R}$ *by*

$$r(D)=\max\{\operatorname{Re}\lambda:\lambda \textit{ an eigenvalue of } A+D\}.$$

Then r is convex. r is s-convex if A is irreducible.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/84/$3.00

THEOREM 2. *Define* $s : \mathscr{D}_n \to \mathbf{R}$ *by*

$$s(D) = \ln \rho(e^D A).$$

Then s is convex. If A is fully indecomposable, then s is s-convex.

Both results have been proved in [6] by S. Friedland, who used the Donsker-Varadhan variational principle. The convexity of r was first proved by Cohen [3] using tools from the theory of random evolutions. A purely matrix-theoretic proof was given by Deutsch and Neumann [4]. We feel that our proofs are more elementary, simpler, and shorter. We give essentially two versions of the proof. Firstly, we relate the convexity of r and s to the convexity of certain sets of M-matrices, a result which was established by Carlson and Varga in 1973 [2] and by the author in 1970 [5]. Then we give, by adapting the ideas in [2] and [5], direct proofs of Theorems 1 and 2, including the strictness results.

2. RESULTS

We define for $A \geqslant 0$ the sets

$$\mathscr{M} = \{ D \in \mathscr{D}_n : D - A \text{ is an } M\text{-matrix} \}$$

and

$$\mathscr{N} = \{ D \in \mathscr{D}_n : e^{-D} - A \text{ is an } M\text{-matrix} \}.$$

Recall that $B = \kappa I - C$, $C \geqslant 0$, is an M-matrix if $\kappa \geqslant \rho(C)$. A set $S \subset \mathscr{D}_n$ is *strictly convex* if $D_1, D_2 \in S$, $D_1 \neq D_2$, $0 < \alpha < 1$ implies $\alpha D_1 + (1-\alpha) D_2 \in \mathring{S}$, where $\mathring{S}$ denotes the interior of S relative to $\mathscr{D}_n$. We have the following

THEOREM 3. $\mathscr{M}$ *and* $\mathscr{N}$ *are convex. For A irreducible* $\mathscr{M}$ *is strictly convex. For A fully indecomposable* $\mathscr{N}$ *is strictly convex.*

The results on $\mathscr{M}$ are proved in [5, Satz 3] (observe however that the definitions of M-matrix are different). The convexity of $\mathscr{N}$ is equivalent to

$$D_1, D_2 \in \mathscr{M} \quad \Rightarrow \quad D_1^{\alpha} D_2^{1-\alpha} \in \mathscr{M} \quad \text{for } 0 \leqslant \alpha \leqslant 1, \tag{2}$$

which can be found in [2, proof of Theorem 3]. The last result is new and follows from the subsequent proof of Theorems 1 and 2.

Let us indicate that Theorem 3 and Theorems 1, 2 are equivalent. We need only apply the following

PROPOSITION. *For $\kappa \in \mathbf{R}$ and $D \in \mathscr{D}_n$ the following hold:*

(a) $\kappa \geqslant r(D)$ *iff* $\kappa I - D \in \mathscr{M}$,
(b) $\kappa > r(D)$ *iff* $\kappa I - D \in \mathring{\mathscr{M}}$,
(c) $\kappa \geqslant s(D)$ *iff* $D - \kappa I \in \mathscr{N}$,
(d) $\kappa > s(D)$ *iff* $D - \kappa I \in \mathring{\mathscr{N}}$.

These are easily established. For example, $r(D_i)I - D_i \in \mathscr{M}$. From the convexity of $\mathscr{M}$ we infer $B_\alpha = [\alpha r(D_1)+(1-\alpha)r(D_2)]I - [\alpha D_1 + (1-\alpha)D_2] \in \mathscr{M}$ and (a) gives $r(\alpha D_1 + (1-\alpha)D_2) \leqslant \alpha r(D_1)+(1-\alpha)r(D_2)$, i.e., the convexity of $\mathscr{M}$ implies the convexity of r.

If however equality holds in (1) for $\varphi = r$, then $B_\alpha \notin \mathring{\mathscr{M}}$. Then the strict convexity of $\mathscr{M}$ implies $r(D_1)I - D_1 = r(D_2)I - D_2$, i.e., r is s-convex. Similarly we can prove

$$r\,(s)\text{ convex} \quad \Leftrightarrow \quad \mathscr{M}\,(\mathscr{N})\text{ convex}$$
$$r\,(s)\ s\text{-convex} \quad \Leftrightarrow \quad \mathscr{M}\,(\mathscr{N})\text{ strictly convex.}$$

REMARK. By applying the inequality

$$\xi^\alpha \eta^{1-\alpha} \leqslant \alpha\xi + (1-\alpha)\eta \qquad \xi, \eta \geqslant 0, \quad 0 \leqslant \alpha \leqslant 1, \tag{3}$$

which is related to the Hölder inequality

$$\sum_{i=1}^{n} \xi_i^\alpha \eta_i^{1-\alpha} \leqslant \left(\sum \xi_i\right)^\alpha \left(\sum \eta_i\right)^{1-\alpha} \qquad \xi_i, \eta_i \geqslant 0, \quad 0 \leqslant \alpha \leqslant 1, \tag{4}$$

to (2), we see that the convexity of $\mathscr{N}$ implies the convexity of $\mathscr{M}$.

REMARK. For later use we state that for $0 < \alpha < 1$ equality holds in (3) iff $\xi = \eta$ and equality holds in (4) iff the vectors $\xi = (\xi_1, \ldots, \xi_n)$ and $\eta = (\eta_1, \ldots, \eta_n)$ are linearly dependent.

3. PROOFS

We turn now to the direct proofs of Theorems 1 and 2.

Proof of Theorem 1. To prove (1) for $\varphi = r$ it suffices to assume that A is irreducible and $A + D_i \geqslant 0$. Then by the Perron-Frobenius theorem (e.g. [7, p. 30]) there exist positive vectors $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$ such that $(A + D_1)x = r(D_1)x$ and $(A + D_2)y = r(D_2)y$. Denoting the diagonal elements of D_j by $d_{i,j}$, $j = 1,2$ we have

$$d_{i,1} + \sum_{k=1}^{n} a_{ik}\frac{x_k}{x_i} = r(D_1), \quad d_{i,2} + \sum_{k=1}^{n} a_{ik}\frac{y_k}{y_i} = r(D_2) \qquad i = 1,\dots,n. \tag{5}$$

Defining $z_i = x_i^{\alpha} y_i^{1-\alpha}$ and using (3), we infer

$$\alpha d_{i,1} + (1-\alpha) d_{i,2} + \sum_{k=1}^{n} a_{ik}\frac{z_k}{z_i} \leqslant \alpha r(D_1) + (1-\alpha) r(D_2) \tag{6}$$

which by the Collatz quotient theorem (e.g. [7, Theorem 2.2]) implies (1) for $\varphi = r$. Hence φ is convex.

To prove the second fact, we assume A irreducible and equality in (1) for some $\alpha \in (0,1)$. We want to show that $D_1 = D_2 + \gamma I$. We apply Theorem 2.2 in [7] again and see that equality holds in (6). Considering the case of equality in (3), we infer that $a_{ik} \neq 0$ implies $x_k/x_i = y_k/y_i$. Equation (5) yields $d_{i,1} - r(D_1) = d_{i,2} - r(D_2)$ for $i = 1,\dots,n$ or $D_1 = D_2 + \gamma I$. ■

Proof of Theorem 2. Consider $\tilde{s}(D) = \rho(e^{D}A)$. It suffices to show that for A irreducible and $0 \leqslant \alpha \leqslant 1$

$$\tilde{s}(D_1)^{\alpha}\tilde{s}(D_2)^{1-\alpha} \geqslant \tilde{s}(\alpha D_1 + (1-\alpha) D_2). \tag{7}$$

There exist $x > 0$, $y > 0$ such that $\tilde{s}(D_1)x = e^{D_1}Ax$, $\tilde{s}(D_2)y = e^{D_2}Ay$. Hence by (4)

$$\begin{aligned}\tilde{s}(D_1)^{\alpha}\tilde{s}(D_2)^{1-\alpha} &= \left(e^{d_{i,1}}\sum_k a_{ik}\frac{x_k}{x_i}\right)^{\alpha}\left(e^{d_{i,2}}\sum_k a_{ik}\frac{y_k}{y_i}\right)^{1-\alpha} \\ &\geqslant e^{\alpha d_{i,1} + (1-\alpha)d_{i,2}}\sum_k a_{ik}^{\alpha}\left(\frac{x_k}{x_i}\right)^{\alpha} a_{ik}^{1-\alpha}\left(\frac{y_k}{y_i}\right)^{1-\alpha} \\ &= e^{\alpha d_{i,1} + (1-\alpha)d_{i,2}}\sum_k a_{ik}\frac{z_k}{z_i}, \qquad i = 1,\dots,n, \quad z_i = x_i^{\alpha}y_i^{1-\alpha}.\end{aligned} \tag{8}$$

Again the quotient theorem implies (7).

To establish the second part of Theorem 2, we assume A to be fully indecomposable. It suffices to show that for $0 < \alpha < 1$ equality in (7) implies $D_1 = D_2 + \gamma I$.

Recall that A is fully indecomposable if PA is irreducible for all permutations P. There exists a permutation π such that $B = (b_{ij})$, $b_{ij} = a_{\pi(i),j}$, is irreducible and has a positive diagonal [1]. Equality in (7) implies, as before, equality in (8) for $i = 1,\ldots,n$. By the equality condition for the Hölder inequality we get

$$e^{d_{i,1}} a_{ik} \frac{x_k}{x_i} = c_i e^{d_{i,2}} a_{ik} \frac{y_k}{y_i}, \qquad i,k = 1,\ldots,n. \tag{9}$$

If we sum (9) over k, we get $\tilde{s}(D_1) = c_i \tilde{s}(D_2)$; hence $c_i = c$ independent of i. Setting $W = \operatorname{diag}(w_i)$, $w_i = ce^{d_{i,2} - d_{i,1}}$, we see that A and WA are diagonally similar. Let $z_k = y_k / x_k$. From (9) $a_{ik} = w_i a_{ik} z_k / z_i$. If $b_{ij} \neq 0$, then $w_{\pi(i)} z_j / z_{\pi(i)} = 1$. But also $b_{\pi(i),i} \neq 0$; hence $w_{\pi(i)} z_i / z_{\pi(i)} = 1$, and $z_j = z_i$.

As B is irreducible, any two indices can be connected in B, and therefore $z_i = z_j$ for all i, j. Hence $w_i = 1$, which implies $D_1 = D_2 + \gamma I$, $\gamma = s(D_1) - s(D_2)$. ∎

REFERENCES

1 R. A. Brualdi, S. V. Parter, and H. Schneider, The diagonal equivalence of a non-negative matrix to a stochastic matrix, *J. Math. Anal. Appl.* 16:31–50 (1966).

2 D. H. Carlson and R. S. Varga, Minimal G-functions, *Linear Algebra Appl.* 6:97–117 (1973).

3 J. E. Cohen, Random evolutions and the spectral radius of a non-negative matrix, *Math. Proc. Cambridge Philos. Soc.* 86:345–350 (1979).

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.*, to appear.

5 L. Elsner, Über Eigenwerteinschließungen mit Hilfe von Gerschgorin-Kreisen, *Z. Angew. Math. Mech.* 50:381–384 (1970).

6 S. Friedland, Convex spectral functions, *Linear and Multilinear Algebra* 9:299–316 (1981).

7 R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N.J., 1962.

Received 21 March 1983