Assaf Goldberger; Michael Neumann

An improvement of an inequality of Fiedler leading to a new conjecture on nonnegative matrices

Czechoslovak Mathematical Journal, Vol. 54 (2004), No. 3, 773–780

Persistent URL: http://dml.cz/dmlcz/127928

Czechoslovak Mathematical Journal, 54 (129) (2004), 773–780

AN IMPROVEMENT OF AN INEQUALITY OF FIEDLER LEADING TO A NEW CONJECTURE ON NONNEGATIVE MATRICES

Assaf Goldberger and Michael Neumann, Storrs

(Received January 10, 2002)

Abstract. Suppose that A is an $n \times n$ nonnegative matrix whose eigenvalues are $\lambda = \varrho(A), \lambda_2, \ldots, \lambda_n$. Fiedler and others have shown that $\det(\lambda I - A) \leqslant \lambda^n - \varrho^n$, for all $\lambda > \varrho$, with equality for any such λ if and only if A is the simple cycle matrix. Let a_i be the signed sum of the determinants of the principal submatrices of A of order $i \times i$, $i = 1, \ldots, n-1$. We use similar techniques to Fiedler to show that Fiedler's inequality can be strengthened to: $\det(\lambda I - A) + \sum\limits_{i=1}^{n-1} \varrho^{n-2i} |a_i| (\lambda - \varrho)^i \leqslant \lambda^n - \varrho^n$, for all $\lambda \geqslant \varrho$. We use this inequality to derive the inequality that: $\prod\limits_2^n (\varrho - \lambda_i) \leqslant \varrho^{n-2} \sum\limits_{i=2}^{n} (\varrho - \lambda_i)$. In the spirit of a celebrated conjecture due to Boyle-Handelman, this inequality inspires us to **conjecture** the following inequality on the nonzero eigenvalues of A: *If* $\lambda_1 = \varrho(A), \lambda_2, \ldots, \lambda_k$ *are (all) the nonzero eigenvalues of* A, *then* $\prod\limits_2^k (\varrho - \lambda_i) \leqslant \varrho^{k-2} \sum\limits_{i=2}^{k} (\varrho - \lambda)$. We prove this conjecture for the case when the spectrum of A is real.

Keywords: nonnegative matrices, M-matrices, determinants

MSC 2000: 15A48, 15A15

1. Background and main results

Suppose that A is an $n \times n$ nonnegative matrix whose spectral radius is $\varrho = \varrho(A)$. In [5] Fiedler has shown that

$$\det(\lambda I - A) \leqslant \lambda^n - \varrho^n, \quad \forall \lambda > \varrho \tag{1.1}$$

with equality for any such λ if and only if A is the simple cycle matrix. It should

Research supported in part by NSF Grant No. DMS9973247.

be pointed out that in addition Fiedler's manuscript, Keilson and Styan [6] and Ashley [2] have also published similar results.

In Boyle and Handelman, [4, Question, p. 311], the authors conjecture a stronger result than (1.1), namely, if $\lambda_1, \ldots, \lambda_k$, with $\lambda_1 = \varrho$, are all **the nonzero eigenvalues** of A, then

$$\prod_{i=1}^{k}(\lambda - \lambda_i) \leqslant \lambda^k - \varrho^k, \quad \forall \lambda \geqslant \varrho. \tag{1.2}$$

In [7] it is shown that (i) the conjecture is true for $k \leqslant 5$, (ii) the conjecture is true when all the λ_i's are real, and (iii) in general there is a sequence of numbers $c_k > \varrho$, with $c_k \to \varrho$ as $k \to \infty$, such that (1.2) holds for all $\lambda \geqslant c_k$. In [1], Ambikkumar and Drury prove that (1.2) is true when $\operatorname{rank}(A) \geqslant n/2$.

If A is an $n \times n$ nonnegative and irreducible matrix, then it is well known that A is diagonally similar to a nonnegative and irreducible matrix with constant row sums. It will be simpler initially to develop some of our main results in this paper by restricting ourselves to the set of all $n \times n$ nonnegative matrices whose row sums are a constant 1 which is commonly known as *the set of all the row stochastic matrices.*

Let $\mathscr{P}$ be the set of all $n \times n$ row stochastic matrices. Then an approach along the lines which Fiedler used in [5] to prove (1.1) yields the following result:

Theorem 1.1. *Suppose that* $\Pi\colon \mathbb{R}^{n,n} \to \mathbb{R}$ *is a row multiaffine functional*[1]. *Denote the maximum* $\pi = \max_{A \in \mathscr{P}} \Pi(A)$. *Then there is a nonnegative matrix* $P \in \mathscr{P}$ *with one nonzero entry in every row (necessarily* 1*) such that* $\Pi(P) = \pi$.

Proof. Let $P' = (p'_{i,j}) \in \mathscr{P}$ be a matrix with the smallest possible number of nonzero entries such that $\Pi(P') = \pi$. We will show that, necessarily, P' has only one nonzero entry in every row. Suppose that P' has at least two nonzero entries, say in the first row, in the positions $(1, r)$ and $(1, s)$. Let $Y = (y_{i,j}) \in \mathbb{R}^{n,n}$ be defined by $y_{1,r} = 1$; $y_{1,s} = -1$ and $y_{i,j} = 0$ otherwise. For small values of $|\varepsilon|$, the matrices $P' + \varepsilon Y$ are still in $\mathscr{P}$ so that we have

$$\Pi(P' + \varepsilon Y) \leqslant \pi$$

with equality at $\varepsilon = 0$. Since by our assumption on Π, the above expression is affine in ε, it is therefore independent of ε. By choosing $\varepsilon = p'_{1,s}$ we arrive at new stochastic matrix $P'' = P' + \varepsilon Y$ with lesser amount of nonzero entries and still $\Pi(P'') = \pi$ which is a contradiction. $\square$

[1] A *row multiaffine functional* from $\mathbb{R}^{n,n} \to \mathbb{R}$ is a functional which is affine in every row of the matrices in its domain.

As an example for multiaffine functionals we give the generalized matrix or Schur functions

$$d_\chi^G(A) = \sum_{\sigma \in G \subseteq S_n} \chi(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)},$$

where G is a subgroup of the symmetric group S_n and χ is a real character of G. We comment that all Schur functions are invariant under diagonal similarity. Two well known particular cases of a Schur function are: (i) $G = S_n$ and $\chi = 1$ which yield the permanental function and (ii) $G = S_n$ and $\chi =$ sign which yield the deteminantal function. It is the determinantal function which is of interest to us in this paper.

For a matrix $A \in \mathbb{R}^{n,n}$ let us write:

$$\varphi_A(\lambda) := \det(\lambda I - A) = \lambda^n + \sum_{i=1}^{n} a_i(A)\lambda^{n-i}. \tag{1.3}$$

It is well known that $a_i(A)$ are the signed sum of the determinants of the principal submatrices of A of size $i \times i$, $i = 1, \ldots, n$, and viewed as the functions $a_i \colon \mathbb{R}^{n,n} \to \mathbb{R}$, the a_i's are multiaffine functions in the rows and columns of A.

To obtain some applications of Theorem 1.1 we first need the following technical results:

Lemma 1.2. *Let r_i and s_i be nonnegative real numbers, $1 \leqslant i \leqslant r$, with $s_i \geqslant r_i + 1$, for all $i = 1, \ldots, r$. Then*

$$\prod_{i=1}^{r}(r_i + 1) + \prod_{i=1}^{r}(s_i - 1) \leqslant \prod_{i=1}^{r} r_i + \prod_{i=1}^{r} s_i.$$

P r o o f. First suppose that $s_i = r_i + 1$. Then we have an equality. It will thus suffice to show that the expression $\prod\limits_i s_i - \prod\limits_i (s_i - 1)$ is increasing in each s_i. This is obviously seen if we substitute $\sigma_i = s_i - 1 \geqslant 0$ and expand the resulting expression. □

An immediate corollary to Lemma 1.2 is the following:

Corollary 1.3. *Let $n = \sum\limits_{i=1}^{r} n_i$, where $n_i \geqslant 1$, $i = 1, \ldots, r$, are integers. Then for all $\lambda \geqslant 1$,*

$$\prod_{i=1}^{r}[(\lambda - 1)^{n_i} - 1] + \prod_{i=1}^{r}[\lambda^{n_i} - 1] \leqslant \lambda^n + (\lambda - 1)^n.$$

P r o o f. Use Lemma 1.2 with $r_i = (\lambda - 1)^{n_i}$ and $s_i = \lambda^{n_i}$. Clearly $s_i = ((\lambda - 1) + 1)^{n_i} \geqslant (\lambda - 1)^{n_i} + 1 = r_i + 1$. □

As an application of Theorem 1.1 we have now the following lemma:

Lemma 1.4. *Let $\varepsilon_i = \pm 1$, for $1 \leqslant i \leqslant n-1$. Then for every nonnegative matrix $A \in \mathbb{R}^{n,n}$ whose spectral radius is* 1,

$$\det(\lambda I - A) + \sum_{i=1}^{n-1} \varepsilon_i a_i(A)(\lambda - 1)^i \leqslant \lambda^n - 1, \quad \forall\, \lambda \geqslant 1. \tag{1.4}$$

P r o o f. We will prove the lemma by induction on n. For $n = 1$ this is obvious. It is enough to prove the lemma at n for row stochastic matrices. Having done this, the inequality extends by similarity to all irreducible nonnegative matrices with spectral radius 1 and then by continuity also to reducible matrices. Since the left hand side of (1.4) is multiaffine in the rows of A, then by the previous theorem it is enough to verify (1.4) for row stochastic matrices P with one nonzero entry at every row.

Suppose first that P is a permutation matrix in $\mathbb{R}^{n,n}$. Then P is similar to a diagonal block matrix whose blocks P_i are cyclic permutations matrices of size n_i, $1 \leqslant i \leqslant r$, $\sum\limits_{i=1}^{r} n_i = n$. In particular the characteristic polynomial is $f(\lambda) = \prod\limits_{i=1}^{r} (\lambda^{n_i} - 1)$, and $|a_i(P)| = |a_{n-i}(P)|$ is the coefficient of λ^i in the polynomial $g(\lambda) = \prod\limits_{i=1}^{r} (\lambda^{n_i} + 1)$. Now,

$$\begin{aligned} s := \sum_{i=1}^{n-1} \varepsilon_i a_i(P)(\lambda-1)^i &\leqslant \sum_{i=1}^{n-1} |a_i(P)|(\lambda - 1)^i \\ &= \prod_{i=1}^{r} [(\lambda-1)^{n_i} + 1] - 1 - (\lambda-1)^n. \end{aligned}$$

Thus it will suffice to show that

$$\prod_{i=1}^{r} [\lambda^{n_i} - 1] + \prod_{i=1}^{r} [(\lambda-1)^{n_i} - 1] - 1 - (\lambda-1)^n \leqslant \lambda^n - 1.$$

This follows from Corollary 1.3.

It remains to verify (1.4) for all row stochastic matrices P with only one nonzero entry in every row, whose rank is strictly less than n. Such a matrix can be factored as $P = AB$ with *nonnegative* $A \in \mathbb{R}^{n,r}$, $B \in \mathbb{R}^{r,n}$ and $r = \operatorname{rank}(P)$. Now, consider the matrix $Q = BA$. Then $s := \sum\limits_{i=1}^{r-1} \varepsilon_i a_i(Q)(\lambda-1)^i = \sum\limits_{i=1}^{n-1} \varepsilon_i a_i(P)(\lambda-1)^i$ and $\lambda^{n-r}\det(\lambda I - Q) = \det(\lambda I - P)$. By the induction hypothesis,

$$\det(\lambda I - Q) + s \leqslant \lambda^r - 1. \tag{1.5}$$

If $s < 0$, then (1.4) holds for P because of Fiedler's original inequality (1.1), so assume that $s \geqslant 0$. In this case multiply (1.5) by λ^{n-r} to have:

$$\det(\lambda I - P) + \lambda^{n-r} s \leqslant \lambda^n - \lambda^{n-r}.$$

As $\lambda \geqslant 1$, replacing $\lambda^{n-r}s$ with s diminishes the left hand side, and replacing $\lambda^n - \lambda^{n-r}$ with $\lambda^n - 1$ increases the right hand side. The lemma now follows for n. □

As a corollary we get an improvement to Fiedler's result given in (1.1):

Theorem 1.5. *For any nonnegative matrix A with spectral radius 1 and for every $\lambda \geqslant 1$,*

$$\det(\lambda I - A) + \sum_{i=1}^{n-1} |a_i(A)|(\lambda - 1)^i \leqslant \lambda^n - 1. \tag{1.6}$$

As a further corollary to Lemma 1.4 we obtain that:

Corollary 1.6. *Let A be an $n \times n$ nonnegative matrix with spectral radius 1 and whose eigenvalues are $\lambda_1 = 1, \lambda_2, \ldots, \lambda_n$. Then*

$$\varphi_A'(1) = \prod_{i=2}^{n}(1 - \lambda_i) \leqslant \operatorname{trace}(I - A), \tag{1.7}$$

where φ is given in (1.3).

Proof. By Lemma 1.4,

$$f(\lambda) := \det(\lambda I - A) + (\lambda - 1)\mathrm{tr(A)} - (\lambda^{\mathrm{n}} - 1) \leqslant 0$$

with equality at $\lambda = 1$. Thus at $\lambda = 1$ we have that

$$0 \geqslant f'(1) = \varphi_A'(1) - \operatorname{trace}(A) - n$$

from which (1.7) follows. □

We mention that inequality (1.7) is sharp as can be see by taking A to be the simple cycle matrix of order n. In that case both the left and right hand sides of (1.7) equal to n.

Corollary 1.6 leads us to a conjecture on the eigenvalues of a stochastic matrix which is in the spirit of the Boyle-Handelman conjecture stated in (1.2).

Conjecture 1.7. *Let A be an $n \times n$ nonnegative matrix with spectral radius* 1 *and whose nonzero eigenvalues are $\lambda_1 = 1, \lambda_2, \ldots, \lambda_k$. Then*

$$\prod_{i=2}^{k}(1-\lambda_i) \leqslant \sum_{i=2}^{k}(1-\lambda_i). \tag{1.8}$$

At this time we are unable to prove the conjecture, though we have tried it on many examples without failure. However, we can partially prove the conjecture for the case in which A has a **real** spectrum.

Theorem 1.8. *If A is an $n \times n$ nonnegative matrix whose eigenvalues are all real, with spectral radius* 1, *and whose nonzero eigenvalues are $\lambda_1 = 1, \lambda_2, \ldots, \lambda_k$, then* (1.8) *is true.*

Proof. We begin by noting that as A is a nonnegative matrix, $\operatorname{trace}(A) = \sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{k} \lambda_i \geqslant 0$. Actually we shall prove a stronger result than stated here, namely, *suppose $x_1, x_2, \ldots, x_k$ are real numbers such that $x_1 = 1 \geqslant |x_i|$, $i = 1, \ldots, k$ and such that $\sum_{i=1}^{k} x_i \geqslant 0$. Then*

$$\prod_{i=2}^{k}(1-x_i) \leqslant \sum_{i=2}^{k}(1-x_i). \tag{1.9}$$

To prove the claim note first that if all the x_i's are nonnegative, then (1.9) trivially holds. So assume that at least one of $x_2, \ldots, x_k$ is negative. As $1 \geqslant x_i$, for $i = 2, \ldots, k$, and $1 \geqslant -\sum_{i=2}^{k} x_i$, it follows from the arithmetic-geometric inequality that

$$\begin{aligned}\prod_{i=2}^{k}(1-x_i) &\leqslant \left[\frac{1}{k-1}\sum_{i=2}^{k}(1-x_i)\right]^{k-1} = \left[\frac{1}{k-1}\left((k-1)-\sum_{i=2}^{k}x_i\right)\right]^{k-1} \\ &\leqslant \left(1+\frac{1}{k-1}\right)^{k-1} \leqslant \mathrm{e} \leqslant 3.\end{aligned} \tag{1.10}$$

Suppose now that exactly one of the x_i's among $x_2, \ldots, x_k$, say x_2, that is negative. Then on letting $y := \prod_{i=3}^{k}(1-x_i)$, we see that $0 \leqslant y \leqslant 1$ and $y \leqslant \sum_{i=3}^{k}(1-x_i)$. But then

$$\prod_{i=2}^{k}(1-x_i) = (1-x_2)y \leqslant (1-x_2)+y \leqslant \sum_{i=2}^{k}(1-x_i).$$

Next, suppose that exactly two of the x_i's among $x_2, \ldots, x_k$, say x_2 and x_3, which are negative. Then on letting $z = \prod_{i=4}^{k}(1 - x_i)$, we see that $0 \leqslant z \leqslant 1$ and that $z \leqslant \sum_{i=4}^{k}(1 - x_i)$. In this case we can write that

$$\begin{aligned}(1 - x_2)(1 - x_3)z &\leqslant (1 - x_2)(1 - x_3) + z \\ &\leqslant (1 - x_2) + (1 - x_3) + z \\ &\leqslant \sum_{i=2}^{k}(1 - x_i),\end{aligned}$$

where we have used the fact that if a and b are numbers such that $ab \leqslant 1$, then $(1 + a)(1 + b) \leqslant (1 + a) + (1 + b)$. Finally, if three or more of the x_i, $i = 2, \ldots, k$, negative, then clearly $\sum_{i=2}^{k}(1 - x_i) \geqslant 3$ and so (1.9) holds because of (1.10). □

Reverting to general $n \times n$ nonnegative matrices we have proved in this paper the following results:

Theorem 1.9. *Let A be an $n \times n$ nonnegative matrix whose eigenvalues are $\lambda_1 = \varrho$, $\lambda_2, \ldots, \lambda_n$, where $\varrho = \varrho(A)$ is the spectral radius of A. Then*

$$\det(\lambda I - A) + \sum_{i=1}^{n-1} \varrho^{n-2i}\ |a_i(A)|(\lambda - \varrho)^i \leqslant \lambda^n - \varrho^n, \quad \forall \lambda \geqslant \varrho,$$

and

$$\prod_{i=2}^{n}(\varrho - \lambda_i) \leqslant \varrho^{n-2}\,\mathrm{trace}(\varrho I - A).$$

Theorem 1.9 and Conjecture 1.10 lead us to formulate the following conjecture:

Conjecture 1.10. *Let A be an $n \times n$ nonnegative matrix, whose spectral radius is ϱ, and whose nonzero eigenvalues are $\lambda_1 = \varrho$, $\lambda_2, \ldots, \lambda_k$. Then*

$$\prod_{i=1}^{k}(\lambda - \lambda_i) + \sum_{i=1}^{k-1} \varrho^{k-2i}|a_i(A)|(\lambda - \varrho)^{\leqslant}\lambda^k - \varrho^k, \quad \forall \lambda \geqslant \varrho, \tag{1.11}$$

and

$$\prod_{i=2}^{k}(\varrho - \lambda_i) \leqslant \varrho^{k-2}\sum_{i=2}^{k}(\varrho - \lambda_i). \tag{1.12}$$

Concerning the above conjecture, we comment that if (1.11) is true, then (1.12) follows as a corollary.

References

[1] *S. Ambikkumar and S. W. Drury*: Some remarks on a conjecture of Boyle and Handelman. Lin. Alg. Appl. *264* (1997), 63–99.

[2] *J. Ashley.*: On the Perron-Frobenius eigenvector for nonnegative integral matrices whose largest eigenvalue is integral. Lin. Alg. Appl. *94* (1987), 103–108.

[3] *A. Berman and R. J. Plemmons*: Nonnegative Matrices in the Mathematical Sciences. SIAM, Philadelphia, 1994.

[4] *M. Boyle and D. Handelman*: The spectra of nonnegative matrices via symbolic dynamics. Annals of Math. *133* (1991), 249–316.

[5] *M. Fiedler*: Untitled private communication. 1982.

[6] *J. Keilson and G. Styan*: Markov chains and M-matrices: Inequalities and equalities. J. Math. Anal. Appl. *41* (1973), 439–459.

[7] *I. Koltracht, M. Neumann and D. Xiao*: On a question of Boyle and Handelman concerning eigenvalues of nonnegative matrices. Lin. Multilin. Alg. *36* (1993), 125–140.

Authors' address: Department of Mathematics, University of Connecticut, Storrs, Connecticut 06269-3009, USA, `e-mail: m.neumann@uconn.edu`.