SIAM J. MATRIX ANAL. APPL.
Vol. 15, No. 4, pp. 1157–1166, October 1994

117

ROW SUMS AND INVERSE ROW SUMS FOR NONNEGATIVE MATRICES *

SHMUEL FRIEDLAND†1, ROHAN HEMASINHA‡, HANS SCHNEIDER§1, JEFFREY STUART¶1, AND JAMES WEAVER‡1

Abstract. For a nonnegative, irreducible matrix A, the grading of the row sums vector and the grading of the Perron vector are used to predict the grading of the row sums vector of $(I-A)^{-1}$. This has applications to Markov chains.

Key words. row sums, inverse row sums, Markov chain, nonnegative matrix

AMS subject classifications. 15A48, 15A42

0. Motivation. Let T be a row stochastic matrix. It is well known that the matrix T is the transition matrix associated with an absorbing Markov chain if and only if T is permutation similar to a matrix of the form

$$T=\begin{bmatrix} I & 0 \\ B & A \end{bmatrix},$$

where A is a square matrix with $\rho(A)<1$ [BP, Thm. 8.3.21]. Furthermore, if F is the set of indices corresponding to the nonabsorbing, i.e., transient, states then the expected number of steps until absorption when starting in the nonabsorbing state i is given by

$$\sum_{j\in F}[(I-A)^{-1}]_{ij}$$

[BP, Thm. 8.4.27]. This leads to the natural question of what can be said about the row sums of the matrix $(I-A)^{-1}$ given some knowledge about the matrix A. In particular, what can we predict about the maximum and minimum row sums of $(I-A)^{-1}$ given the row sums of A and the Perron vector for A?

1. Notation. For an $n\times n$ matrix A, let $\rho=\rho(A)$ denote the spectral radius of A. The real matrix A will be called nonnegative, denoted $A\geq 0$, if each entry of A is nonnegative. If A is nonnegative and irreducible, let $X=X_A$ denote the Perron vector of euclidean norm one for A; that is, X is the unique strictly positive eigenvector of norm one corresponding to the eigenvalue $\rho(A)$. Unless otherwise specified, the matrix A will always be an $n\times n$ nonnegative matrix.

---

* Received by the editors July 27, 1992; accepted for publication December 8, 1993.

† Department of Mathematics, University of Illinois at Chicago, Chicago, Illinois 60680

‡ Department of Mathematics and Statistics, University of West Florida, Pensacola, Florida 32514

§ Department of Mathematics, University of Wisconsin, Madison, Wisconsin 53706. The work of this author was partially supported by National Science Foundation grants DMS 8901445, DMS 9123318, and ECS 8718971

¶ Department of Mathematics, University of Southern Mississippi, Hattiesburg, Mississippi 39406

1 The research for this paper was performed while these authors were visiting the Institute for Mathematics and Its Applications, University of Minnesota, Minneapolis, Minnesota 55455

Let $v = (v_1, v_2, \dots, v_n)^t \in \mathbb{R}^n$. There exists a permutation σ such that $v_{\sigma(1)} \geq v_{\sigma(2)} \geq \cdots \geq v_{\sigma(n)}$. The integer vector $(\sigma^{-1}(1), \sigma^{-1}(2), \dots, \sigma^{-1}(n))^t$ is called a grading of v. If the entries of v are pairwise distinct, then v has a unique grading and v is called a strictly graded vector. A set of vectors is said to share a common grading if the intersection of their sets of gradings is nonempty.

For $1 \leq i \leq n$, let e_i denote the ith standard basis vector for $\mathbb{R}^n$. Let $u = u_n$ denote the vector of ones. That is,

$$u_n = \sum_{i=1}^{n} e_i.$$

Let $D = D_n$ denote the cone in $\mathbb{R}^n$ generated by the vectors $e_1, e_1 + e_2, e_1 + e_2 + e_3, \dots, u_n$; that is, $D = \{v \in \mathbb{R}^n\colon\ v_1 \geq v_2 \geq \cdots \geq v_n \geq 0\}$. Let $\Pi(D)$ denote the class of D-preserving matrices: $\Pi(D) = \{A \in \mathcal{M}_n(\mathbb{R})\colon\ A(D) \subseteq D\}$.

Note that a nonnegative vector $v \in \mathbb{R}^n$ has its entries in decreasing order if and only if $v \in D$, and that v has its entries in strictly decreasing order if and only if $v \in \operatorname{int}(D)$, where $\operatorname{int}(D)$ denotes the interior of D. Also note that if $A \in \Pi(D)$, then $A^k \in \Pi(D)$ for all positive integers k. Finally note that the row sums of the matrix A are precisely the entries of the vector Au.

LEMMA 1.1. *If A is a nonnegative, primitive matrix with $\rho(A) < 1$, such that $Au, (I - A)^{-1}u$, and X_A share a common (strict) grading, then there exists a permutation matrix P such that $PAP^tu, (I - PAP^t)^{-1}u$, and PX_A are all in* $(\operatorname{int}(D))D$. *Furthermore, $PX_A = X_{PAP^t}$.*

Proof. Let $v = Au$. Let the permutation matrix P correspond to the common permutation σ in the definition of grading. Then $PAP^tu = PAu \in D$. Since $\rho < 1, (I-A)^{-1}$ exists. Since $P(I-A)^{-1}P^t = (I-PAP^t)^{-1}$, and since Au and $(I-A)^{-1}u$ share the common grading $\sigma, (I-PAP^t)^{-1}u = (I-PAP^t)^{-1}P^tu = P(I-A)^{-1}u \in D$. Finally, X is an eigenvector for ρ for A if and only if PX is an eigenvector for ρ for PAP^t. Since multiplication by P is norm preserving and since PAP^t is nonnegative and primitive, $X_{PAP^t} = PX_A$. Note that σ is a common grading for Au and X_A, so $PX_A \in D$. □

One immediate consequence of this lemma is that we can always assume that a graded vector has its entries in decreasing order. Thus questions about graded vectors are transformed to questions about vector membership in the cone D.

Finally, recall the Neumann expansion for the inverse of the matrix $I - A$.

THEOREM 1.2 [O]. *Let A be an $n \times n$ real matrix with $\rho(A) < 1$. Then $(I - A)^{-1}$ exists, and*

$$(I - A)^{-1} = I_n + \sum_{k=1}^{\infty} A^k.$$

2. Empirical evidence. If A is nonnegative and primitive, then by the power method, $A^ku \approx c_kX_A$ for large k. Furthermore, if $\rho(A) < 1$, then $c_k \to 0$ as $k \to \infty$. Also, $\rho(A) < 1$ implies

$$(I - A)^{-1}u = u + \sum_{k=1}^{\infty} A^k u.$$

This suggests that the grading for $(I - A)^{-1}u$ should be linked to the grading for X_A, and that the early terms in the summation should be the most important. Since

$(I-A)^{-1}u$ and $(I-A)^{-1}u-u$ have the same grading, and since

$$(I-A)^{-1}u-u = Au + \sum_{k=2}^{\infty} A^k u,$$

the importance of the grading of Au is immediately apparent. When Au and X share a common grading, it remains to be seen how much of an effect the remaining low order summands have on the grading of $(I-A)^{-1}u$.

Motivated by numerical experiments conducted using MATLAB on an APOLLO workstation, we were led to several conjectures. The first was that if Au, X, and $(I-A)^{-1}u$ all share a common grading, then that grading is shared by A^k for all positive k. The second and more interesting conjecture was that if Au and X share a common grading, then $(I-A)^{-1}u$ also shares that grading. Unfortunately, neither conjecture holds.

If

$$A = \begin{bmatrix} 0.0783 & 0.2999 & 0.2421 & 0.0089 \\ 0.0305 & 0.0003 & 0.1814 & 0.2272 \\ 0.0013 & 0.1196 & 0.1426 & 0.1305 \\ 0.0008 & 0.0005 & 0.0009 & 0.0003 \end{bmatrix},$$

then

$$Au = \begin{bmatrix} 0.6292 \\ 0.4394 \\ 0.3940 \\ 0.0025 \end{bmatrix}, \quad X = \begin{bmatrix} 0.8954 \\ 0.3166 \\ 0.3131 \\ 0.0043 \end{bmatrix}, \quad \text{and} \quad (I-A)^{-1}u = \begin{bmatrix} 2.0101 \\ 1.5695 \\ 1.5411 \\ 1.0041 \end{bmatrix}.$$

Hence, Au, X, and $(I-A)^{-1}u$ are all in D. However,

$$A^2u = \begin{bmatrix} 0.2765 \\ 0.0914 \\ 0.1099 \\ 0.0011 \end{bmatrix},$$

which is not in D since its second and third entries are not in decreasing order.

Discovering a counterexample to the second conjecture proved to be very difficult. The following matrix was one of only four counterexamples generated during a run of 25,000 randomly generated, rank four, strictly positive 4×4 matrices with spectral radius less than one. In fact, for 67% of the matrices generated in that run, all three vectors—Au, X, and $(I-A)^{-1}u$—shared a common grading.

$$A = \begin{bmatrix} 0.2042 & 0.2837 & 0.0196 & 0.2356 \\ 0.1522 & 0.2149 & 0.0320 & 0.2848 \\ 0.1060 & 0.2072 & 0.1289 & 0.2415 \\ 0.2750 & 0.1958 & 0.0805 & 0.0187 \end{bmatrix},$$

then

$$Au = \begin{bmatrix} 0.7431 \\ 0.6839 \\ 0.6836 \\ 0.5700 \end{bmatrix}, \quad X = \begin{bmatrix} 0.5513 \\ 0.4991 \\ 0.4987 \\ 0.4452 \end{bmatrix}, \quad \text{and} \quad (I-A)^{-1}u = \begin{bmatrix} 3.2356 \\ 3.0330 \\ 3.0338 \\ 2.7799 \end{bmatrix}.$$

Hence Au and X are in D, however, $(I-A)^{-1}u$ is not in D.

A run of 100,000 randomly generated, rank three, strictly positive 4×4 matrices with spectral radius less than one yielded only one counterexample. Furthermore, for this run, 91% of the matrices had all vectors sharing a common grading, and an additional 3% had the Perron vector and the inverse row sums vector (but not the row sums vector) sharing a common grading.

Extensive numerical experiments with matrices of sizes up to 50×50 lead to the following observations. First, for low rank matrices, the grading of X is a good predictor for the grading of $(I-A)^{-1}u$. Second, even when the vectors do not share a common grading, they share a roughly blocked common grading in the sense that the grading vectors differ within blocks corresponding to closely sized entries of the vectors. In particular, the set of indices for the smallest (largest) row sums of A correspond roughly to the set of indices of the smallest (largest) entries of the Perron vector and to the set of indices of the smallest (largest) row sums of $(I-A)^{-1}$.

3. An analytic approach. In this section, we present several different types of results including an examination of certain classes of matrices for which the grading of the Perron vector and the row sums vector do determine the grading for the inverse row sums vector.

PROPOSITION 3.1. *Let A be an $n \times n$ nonnegative, irreducible matrix with $\rho = \rho(A) < 1$. Suppose that $X = X_A \in D$. For $1 \le i \le n$,*

$$(1-\rho)^{-1}\frac{x_i}{x_1} \le [(I-A)^{-1}u]_i \le (1-\rho)^{-1}\frac{x_i}{x_n}.$$

Proof. Since $A \ge 0$, and $\rho < 1, I-A$ is an invertible M-matrix. Thus $(I-A)^{-1} \ge 0$ [BP, Thm. 6.2.3]. Since $X \in D$ and X is strictly positive, $x_1 \ge \cdots \ge x_n > 0$. Note that $(I-A)^{-1}X = (1-\rho)^{-1}X$. Thus for $1 \le i \le n$,

$$\begin{aligned}(1-\rho)^{-1}x_i &= \sum_j [(I-A)^{-1}]_{ij}x_j \\ &\le \sum_j [(I-A)^{-1}]_{ij}x_1 = [(I-A)^{-1}u]_i x_1.\end{aligned}$$

Similarly, the other bound holds. □

THEOREM 3.2. *Let A be a nonnegative, irreducible matrix with $\rho(A) < 1$. If $A \in \Pi(D)$, then Au, X_A and $(I-A)^{-1}u$ are all in D.*

Proof. Since $u \in D$, and since $A^k \in \Pi(D)$ for all positive $k, A^k u \in D$ for all positive k. Since $\rho(A) < 1$, it follows from Theorem 1.2 that $(I-A)^{-1}u = u + \sum_{k=1}^{\infty} A^k u \in D$. Finally since A is nonnegative and irreducible, X_A exists, and by the Krein–Rutman Theorem [BP, Thm. 1.3.2], $A \in \Pi(D)$ implies $X_A \in D$. □

COROLLARY 3.3. *Let A be a nonnegative, irreducible $n \times n$ matrix with $\rho(A) < 1$. If $a_{ij} \ge a_{i+1,j}$ for $1 \le i \le n-1$ and $1 \le j \le n$, then Au, X_A, and $(I-A)^{-1}u$ are all in D.*

Proof. Pick $z \in D$. Note that $z \ge 0$ and $A \ge 0$. Thus for $1 \le i \le n-1$,

$$(Az)_i = \sum_j a_{ij}z_j \ge \sum_j a_{i+1,j}z_j = (Az)_{i+1} \ge 0.$$

Hence $Az \in D$. □

THEOREM 3.4. *Let A be a nonnegative, irreducible $n \times n$ matrix with $\rho = \rho(A) < 1$. Suppose that the minimum polynomial of A is $m_A(\lambda) = \lambda^k(\lambda - \rho)$. Then*

$(I-A)^{-1} = I + A + \cdots + A^{k-1} + (1-\rho)^{-1}A^k$. *Suppose that* $Au, \ldots, A^{k-1}u$ *are in* D. *If either of* X_A *and* $A^k u$ *is in* D, *then all three of* $X_A, A^k u$, *and* $(I-A)^{-1}u$ *are in* D. *Finally, if at least one of* $X_A, Au, \ldots, A^k u$ *is in* $\operatorname{int}(D)$, *then* $(I-A)^{-1}u$ *is in* $\operatorname{int}(D)$.

Proof. Let $X = X_A$. Then $A = \rho XY^t + N$ where Y^t is the strictly positive row eigenvector for ρ such that $Y^t X = 1$, and where N is the nilpotent matrix of index k satisfying $NX = 0$ and $Y^t N = 0^t$. For all nonnegative $r, A^{k+r} = \rho^{k+r} XY^t = \rho^r A^k$. Hence

$$\sum_{r=k}^{\infty} A^r = \sum_{r=0}^{\infty} \rho^r A^k = (1-\rho)^{-1} A^k.$$

Thus $(I-A)^{-1} = I + A + \cdots + (1-\rho)^{-1}A^k$. Since $A^k u = \rho^k (Y^t u) X, X$ is in D if and only if $A^k u$ is in D. Since $u \in D$, it follows that $(I-A)^{-1}u \in D$ when $Au, \ldots, A^k u$ are in D. Furthermore, if one of the summands is in $\operatorname{int}(D)$, then it is clear that $(I-A)^{-1}u$ is in $\operatorname{int}(D)$. □

COROLLARY 3.5. *Let* A *be a nonnegative, irreducible* $n \times n$ *matrix with* $\rho = \rho(A) < 1$. *If* $\operatorname{rank}(A) = 1$ *and if* $Au \in D$, *then* X_A *and* $(I-A)^{-1}u$ *are in* D.

THEOREM 3.6. *Let* A *be a nonnegative, irreducible* $n \times n$ *matrix with* $\rho = \rho(A) < 1$. *Suppose that the minimum polynomial of* A *is either* $m_A(\lambda) = \lambda(\lambda - \rho)(\lambda - \lambda_1)$ *or else* $m_A(\lambda) = (\lambda - \rho)(\lambda - \lambda_1)$, *where* $\lambda_1 \neq 0$. *If* X_A *and* Au *are in* D, *then* $(I-A)^{-1}u$ *is in* D. *Furthermore, if* Au *is in* $\operatorname{int}(D)$, *then* $(I-A)^{-1}u$ *is in* $\operatorname{int}(D)$.

Proof. Let $X = X_A$. Then $A = \rho XY^t + \lambda_1 E$, where Y^t is the strictly positive row eigenvector for ρ satisfying $Y^t X = 1$, and where $E^2 = E, EX = 0$, and $Y^t E = 0^t$. For each positive $k, A^k = \rho^k XY^t + \lambda_1^k E$. Since $\rho < 1, (I-A)^{-1} = I + \rho(1-\rho)^{-1}XY^t + \lambda_1(1-\lambda_1)^{-1}E$. Then $(I-A)^{-1}u = u + \rho(1-\rho)^{-1}(Y^t u)X + \lambda_1(1-\lambda_1)^{-1}Eu$. Since $\lambda_1 < \rho < 1, 0 < (1-\lambda_1)^{-1} < (1-\rho)^{-1}$. Clearly, $Au \geq 0$. Now $Au \in D$ if and only if for $1 \leq i \leq n-1$,

$$\begin{aligned}
&(Au)_i \geq (Au)_{i+1} \\
&\quad \Leftrightarrow \rho\,(Y^t u)\,X_i + \lambda_1\,(Eu)_i \geq \rho\,(Y^t u)\,X_{i+1} + \lambda_1\,(Eu)_{i+1} \\
&\quad \Leftrightarrow \rho\,(Y^t u)\,[X_i - X_{i+1}] \geq \lambda_1\,\left[(Eu)_{i+1} - (Eu)_i\right] \\
&\quad \Leftrightarrow \rho\,(1-\rho)^{-1}\,(Y^t u)\,[X_i - X_{i+1}] \geq \lambda_1\,(1-\rho)^{-1}\,\left[(Eu)_{i+1} - (Eu)_i\right].
\end{aligned}$$

Since $X \in D$, the left-hand side of the last inequality is nonnegative; hence the inequality remains valid when $(1-\rho)^{-1}$ is replaced with $(1-\lambda_1)^{-1}$ on the right-hand side. Thus it holds that

$$\begin{aligned}
&(Au)_i \geq (Au)_{i+1} \\
&\quad \Rightarrow \rho\,(1-\rho)^{-1}\,(Y^t u)\,[X_i - X_{i+1}] \geq \lambda_1\,(1-\lambda_1)^{-1}\,\left[(Eu)_{i+1} - (Eu)_i\right] \\
&\quad \Leftrightarrow \rho\,(1-\rho)^{-1}\,(Y^t u)\,X_i + \lambda_1\,(1-\lambda_1)^{-1}\,(Eu)_i \\
&\qquad \geq \rho\,(1-\rho)^{-1}\,(Y^t u)\,X_{i+1} + \lambda_1\,(1-\lambda_1)^{-1}\,(Eu)_{i+1}.
\end{aligned}$$

That is, $Au \in D$ and $X \in D$ together imply $\rho(1-\rho)^{-1}(Y^t u)X + \lambda_1(1-\lambda_1)^{-1}Eu \in D$. That is, $(I-A)^{-1}u - u \in D$. Since $u \in D, (I-A)^{-1}u \in D$. Furthermore, if $Au \in \operatorname{int}(D)$, then for all i, each inequality in the argument above can be replaced with a strict inequality, hence $(I-A)^{-1}u \in \operatorname{int}(D)$. □

Remark. Let A be an $n \times n$, nonnegative, irreducible matrix with $\rho(A) < 1$. For $n \leq 3$, there are only two cases for A that are not covered in the results above: when

A is 3×3, nonsingular, and either A is not diagonalizable or A has three distinct eigenvalues. All cases for $n = 3$ are contained in Theorem 4.6.

Let A be a nonnegative, irreducible matrix with $\rho(A) < 1$. In view of the preceding results, several natural open questions arise. Suppose that the minimum polynomial for A has degree k. If each of $X_A, Au, A^2u, \ldots, A^{k-1}u$ are in D, does that imply that $(I - A)^{-1}u$ is in D? Does it imply that $A^r u$ is in D for all positive r? If not, what additional restrictions might be sufficient on A or on the minimum polynomial?

4. A second analytic approach.

THEOREM 4.1. *Let B be an $n \times n$ complex matrix with $\rho(B) = 1$. Then there exists a unique positive integer k with $k < n$, and there exist $n \times n$ complex matrices $B_1, \ldots, B_k$ with $B_k \neq 0$ such that* $\operatorname{adj}(I - xB) = I + xB_1 + x^2B_2 + \cdots + x^kB_k$. *Furthermore,*

(i) $k = n - 1$ *if and only if* $\operatorname{rank}(B) \geq n - 1$;

(ii) *if $k < n - 1$, then $k = m + t - 1$, where m is the number of nonzero eigenvalues of B (counting multiplicities), and where t is the size of the largest Jordan block corresponding to the eigenvalue 0 for B.*

Proof. Since each entry of $\operatorname{adj}(I - xB)$ is either zero or (± 1) times an $(n - 1) \times (n - 1)$ minor of $(I - xB)$, it follows that $k \leq n - 1$. Thus $\operatorname{adj}(I - xB) = B_0 + xB_1 + x^2B_2 + \cdots + x^{n-1}B_{n-1}$. Setting $x = 0, B_0 = \operatorname{adj}(I - 0B) = I$. Note that the coefficient matrix for x^{n-1} is generated only by terms from $-xB$. That is, $B_{n-1} = \operatorname{adj}(-B)$. Note that $\operatorname{adj}(-B) = 0$ if and only if every $(n-1) \times (n-1)$ minor of B is zero. That is, if and only if $\operatorname{rank}(B) < n - 1$. Thus (i) is proven.

If S is an invertible matrix, then $S = [\det(S)]\operatorname{adj}(S^{-1})$. Thus $S \operatorname{adj}(I - xB)S^{-1} = \operatorname{adj}(S^{-1})\operatorname{adj}(I - xB)\operatorname{adj}(S) = \operatorname{adj}(S(I - xB)S^{-1}) = \operatorname{adj}(I - xSBS^{-1})$. Thus

$$\operatorname{adj}(I - xSBS^{-1}) = I + xSB_1S^{-1} + x^2SB_2S^{-1} + \cdots + x^{n-1}SB_{n-1}S^{-1}.$$

Choose S so that SBS^{-1} is the Jordan canonical form of B. That is, $SBS^{-1} = J_1 \oplus \cdots \oplus J_r \oplus J_{r+1} \oplus \cdots \oplus J_s$, where the J_α for $1 \leq \alpha \leq r$ are the Jordan blocks corresponding to nonzero eigenvalues, and the J_α for $r < \alpha \leq s$ are the Jordan blocks corresponding to the eigenvalue zero. Then

$$\operatorname{adj}(I - xSBS^{-1}) = \bigoplus_{\alpha=1}^{s} \left[\operatorname{adj}(I - xJ_\alpha) \prod_{\beta \neq \alpha} \det(I - xJ_\beta) \right].$$

Consider the adjoint for a single Jordan block: $J = \lambda I_h + N_h$, where N_h is the $h \times h$ matrix whose only nonzero entries are ones down the superdiagonal. $\operatorname{Adj}(I - xJ)$ has diagonal entries $(1 - x\lambda)^{h-1}$, and the nonzero off-diagonal terms are of the form $(-x)^j(1 - x\lambda)^{h-j-1}$ for $1 \leq j \leq h - 1$. When $\lambda \neq 0$, the maximum degree of x in $\operatorname{adj}(I - xJ)$ is $h - 1$. When $\lambda = 0, (-x)^{h-1}$ is the only type of nonzero term. Thus $\operatorname{adj}(I - xJ)$ is always of degree $h - 1$ in x. Note that $\det(I - xJ) = (1 - x\lambda)^h$. When $\lambda \neq 0, \det(I - xJ)$ is of degree h in x. When $\lambda = 0, \det(I - xJ)$ is of degree zero in x. Consequently, the maximum degree of x in $\Pi_{\alpha=1}^s \det(I - xJ_\alpha)$ is precisely the sum of the sizes of $J_1, J_2, \ldots, J_r$. That is, $\Pi_{\alpha=1}^s \det(I - xJ_\alpha)$ is of degree m in x. Hence for $1 \leq \alpha \leq r, \operatorname{adj}(I - xJ_\alpha) \Pi_{\beta \neq \alpha} \det(I - xJ_\beta)$ is of degree $(h_\alpha - 1) + (m - h_\alpha)$, where h_α is the size of J_α. For $r < \alpha \leq s, \operatorname{adj}(I - xJ_\alpha) \Pi_{\beta \neq \alpha} \det(I - xJ_\beta)$ is of degree $(h_\alpha - 1) + (m - 0)$, where h_α is the size of J_α. Thus the maximum degree of x in $\operatorname{adj}(I - xSBS^{-1})$ and, hence, in $\operatorname{adj}(I - xB)$ is $m + t - 1$, where t is the size of the largest Jordan block for the eigenvalue zero. □

COROLLARY 4.2. *Let A be a nonnegative, irreducible matrix. The following are equivalent:*

(i) $\operatorname{adj}(I - xA) = I + xA_1 + x^2A_2$,

(ii) *at least one of the following holds;*

(a) $n \leq 3$,

(b) $n > 3$ *and* $\operatorname{rank}(A) \leq 2$,

(c) $n > 3$ *and* $\operatorname{rank}(A^2) = 1$.

Note that conditions (b) and (c) imply that the size of the largest possible Jordan block for the eigenvalue zero is two.

Proof (i) $\to$ (ii). If $A_2 \neq 0$ then either $2 = n-1$, hence $n = 3$, or else $2 < n - 1$ and $2 = m + t - 1$. That is, $n > 3$ and $m + t = 3$. Since A is irreducible, $m \geq 1$. Since $t = 0$ implies $m = n, t \geq 1$. Thus either $m = 2$ and $t = 1$, implying $\operatorname{rank}(A) = 2$ or $m = 1$ and $t = 2$, implying $\operatorname{rank}(A^2) = 1$.

If $A_2 = 0$, but $A_1 \neq 0$, then either $1 = n - 1$, hence $n = 2$, or $1 < n - 1$ and $1 = m + t - 1$. In the latter case, $n > 2$ and $m + t = 2$, implying $m = t = 1$. That is, $\operatorname{rank}(A) = 1$.

If $A_2 = A_1 = 0$, then $A = 0$, which contradicts the irreducibility of A. □

Proof (ii) $\to$ (i) If $n \leq 3$, (i) is immediate. If $n > 3$ and $\operatorname{rank}(A) \leq 2$, then since $m + (t-1) \leq \operatorname{rank}(A)$ always, $k = m + t - 1 \leq 2$. Now apply Theorem 4.1. Finally, if $n > 3$ and $\operatorname{rank}(A^2) = 1$, then $\rho(A)$ is the unique nonzero eigenvalue and $m = 1$. Clearly, $\operatorname{rank}(A^2) \leq 1$ implies $t \leq 2$. Again, $k = m + t - 1 \leq 2$. Apply Theorem 4.1. □

LEMMA 4.3. *Let B be a nonnegative, irreducible matrix with $\rho(B) = 1$. If $Bu \in$* $\operatorname{int}(D)$, *then there exists a maximal $\omega = \omega(B)$ such that $0 < \omega \leq 1$, and such that $(I - xB)^{-1}u \in \operatorname{int}(D)$ for $0 < x < \omega$.*

Proof. Since $\rho(B) = 1, 0 < \|B\|_2 \leq 1$. Then $\|B\|_2^k \leq 1$ for all $k \geq 0$. Assume that $0 < x < 1$. Then for $1 \leq i \leq n$,

$$\left| \left[x^2B^2 \left[\sum_{k=0}^{\infty} x^kB^k \right] u \right]_i \right| \leq \left\| \left[x^2B^2 \sum_{k=0}^{\infty} x^kB^k \right] u \right\|_2 \leq x^2 \|B\|_2^2 \left[\sum_{k=0}^{\infty} x^k \|B\|_2^k \right] \|u\|_2$$

$$\leq x^2 \|B\|_2^2 \left[\sum_{k=0}^{\infty} x^k \right] \|u\|_2 = x^2 \|B\|_2^2 (1-x)^{-1} \|u\|_2 .$$

Since $0 \leq x < 1$, Lemma 1.2 yields

$$(I - xB)^{-1} u = u + xBu + x^2B^2 \left[\sum_{k=0}^{\infty} x^kB^k \right] u.$$

For $1 \leq j < n$,

$$[(I - xB)^{-1} u]_j \geq 1 + x [Bu]_j$$

and

$$[(I - xB)^{-1} u]_{j+1} = 1 + x [Bu]_{j+1} + \left[\left[x^2B^2 \sum_{k=0}^{\infty} x^kB^k \right] u \right]_{j+1}$$

$$\leq 1 + x [Bu]_{j+1} + x^2 \|B\|_2^2 (1-x)^{-1} \|u\|_2 .$$

It follows that

$$[(I - xB)^{-1} u]_j - [(I - xB)^{-1} u]_{j+1}$$

$$\geq x \left[[Bu]_j - [Bu]_{j+1} - x \|B\|_2^2 (1-x)^{-1} \|u\|_2 \right].$$

Since $Bu \in \text{int}(D)$, the difference $[Bu]_j - [Bu]_{j+1}$ is strictly positive for $1 \le j < n$. Thus for sufficiently small, positive x, the terms $[(I-xB)^{-1}u]_j$ are strictly decreasing. Since $I - xB$ is a nonsingular M-matrix for $0 < x < 1, (I - xB)^{-1}u \ge 0$ [BP, Thm. 6.2.3]. Thus $(I - xB)^{-1}u \in \text{int}(D)$ for sufficiently small positive x. Thus ω exists. □

LEMMA 4.4. *Let B be a nonnegative, irreducible matrix with $\rho(B) = 1$. If $X_B \in \text{int}(D)$, then there exists a minimal $\tau = \tau(B)$ such that $0 \le \tau < 1$, and such that $(I - xB)^{-1}u \in \text{int}(D)$ for $\tau < x < 1$.*

Proof. Let $f(x)$ be the matrix valued function $f(x) = \text{adj}(I - xB)$. Note that $f(x)$ is continuous for all real x. For $0 < x < 1, I - xB$ is an irreducible, nonsingular M-matrix, hence $\det(I-xB) > 0$ and $(I-xB)^{-1}$ is strictly positive [BP, Thm. 6.2.3]. Since $(I - xB)^{-1} = \det(I - xB)\text{adj}(I - xB)$, it follows that $f(x)$ is strictly positive for $0 < x < 1$. Note also, for each x in $0 < x < 1, f(x)u \in \text{int}(D)$ if and only if $(I - xB)^{-1}u \in \text{int}(D)$.

Since B is nonnegative and irreducible, $\rho(B) = 1$ is a simple eigenvalue for B. Thus $\text{rank}(I - B) = n - 1$. Thus $\text{adj}(I - B) \ne 0$. Since $I - B$ has nullity one, its column null space has basis $\{X_B\}$ and its row null space has basis $\{[X_{B^t}]^t\}$. Since $\det(I - B) = 0$, for $x = 1$,

$$(I - xB)\text{adj}(I - xB) = [\text{adj}(I - B)]\,(I - B) = 0.$$

Thus $f(1) = \text{adj}(I - B) = cX_B[X_{B^t}]^t$ for some nonzero scalar c. Since $f(x)$ is continuous at $x = 1$, and since $f(x)$ is strictly positive for $0 < x < 1, f(1)$ is nonnegative. That is, $c > 0$. Then $f(1)u = [c[X_{B^t}]^t u]X_B$ is a positive multiple of X_B, hence $f(1)u \in \text{int}(D)$. Again using continuity at $x = 1$, it follows that τ exists such that $\tau < 1$ and $f(x)u \in \text{int}(D)$ for $\tau < x < 1$. Observe that $\tau \ge 0$ since $f(0)u = Iu \notin \text{int}(D)$. □

The following theorem is an immediate consequence of Lemmas 4.3 and 4.4.

THEOREM 4.5. *Let B be a nonnegative, irreducible matrix with $\rho(B) = 1$. Suppose that Bu and X_B are in* $\text{int}(D)$. *Let $\omega(B)$ be defined as in Lemma* 4.3, *and let $\tau(B)$ be defined as in Lemma* 4.4. *If $\tau(B) < \omega(B)$, then: $\tau(B) = 0, \omega(B) = 1, (I - xB)^{-1} \in \text{int}(D)$ for $0 < x < 1$, and $\text{adj}(I - xB)u \in \text{int}(D)$ for $0 \le x \le 1$.*

We currently have no useful general characterization of which matrices B satisfy the condition $\tau(B) < \omega(B)$. The numerical evidence presented in the second section, however, suggests that a substantial portion of the matrices B, such that Bu and X_B share a common grading, do satisfy this condition.

THEOREM 4.6. *Let A be a nonnegative, irreducible matrix with $\rho(A) < 1$. Suppose that A satisfies either* (i) *or* (ii) *in Corollary* 4.2. *If both Au and X_A are in* $\text{int}(D)$, *then $(I - A)^{-1}u$ is in* $\text{int}(D)$.

Proof. Let $\rho = \rho(A)$. Let $B = \rho^{-1}A$. Then B is a nonnegative, irreducible matrix with $\rho(B) = 1$. Clearly, $X_B = X_A$. Use (i): $\text{adj}(I - xB) = I + xB_1 + x^2B_2 = I + x(B_1 + xB_2)$. For $x > 0, \text{adj}(I - xB)u$ and $(B_1 + xB_2)u$ have the same gradings. In particular, $\text{adj}(I - xB)u \in \text{int}(D)$ if and only if $(B_1 + xB_2)u \in \text{int}(D)$. Note that for each $i, [(B_1 + xB_2)u]_i$ is a linear function in x. Pick i with $1 \le i < n$. Let α and β be real numbers such that $\alpha < \beta$. If $[(B_1 + \alpha B_2)u]_i \ge [(B_1 + \alpha B_2)u]_{i+1}$ and if $[(B_1 + \beta B_2)u]_i \ge [(B_1 + \beta B_2)u]_{i+1}$ hold, then by linearity in $x, [(B_1 + xB_2)u]_i \ge [(B_1 + xB_2)u]_{i+1}$ holds for $\alpha \le x \le \beta$. Furthermore, if the inequality is strict at either endpoint, then it is strict for $\alpha < x < \beta$.

By hypothesis, $Au \in \text{int}(D)$, hence $Bu \in \text{int}(D)$. Applying Lemma 4.3, there exists ω with $0 < \omega < 1$ such that $(I - xB)^{-1}u \in \text{int}(D)$ for $0 < x < \omega$. For

$0 < x < 1, I - xB$ is a nonsingular M-matrix, and as argued in the proof of Lemma 4.4, $(I - xB)^{-1}u$ is a positive scalar multiple of $\operatorname{adj}(I - xB)u$. Thus, $\operatorname{adj}(I - xB)u \in \operatorname{int}(D)$ for $0 < x < \omega$. Also by hypothesis, $X_B \in \operatorname{int}(D)$, hence from Lemma 4.4 and its proof, there exists a positive τ with $\tau < 1$ such that $\operatorname{adj}(I - xB)u \in \operatorname{int}(D)$ for $\tau < x \leq 1$.

The argument in the preceding paragraph implies that when α is chosen as an arbitrarily small, positive number and when $\beta = 1$, the inequalities for successive entries of $(B_1 + xB_2)u$ are valid and strict for $\alpha < x < \beta$. Thus $\operatorname{adj}(I - xB)u \in \operatorname{int}(D)$ for $0 < x \leq 1$. Hence $(I - xB)^{-1}u \in \operatorname{int}(D)$ for $0 < x < 1$. Since $0 < \rho(A) < 1$, and since $\rho(A)B = A, (I - A)^{-1}u \in \operatorname{int}(D)$. □

The following example shows that the conclusion of Theorem 4.6 can be false if the condition that $Au \in \operatorname{int}(D)$ is dropped. Let $X = (3, 2, 1)^t$. Let B be the parameterized matrix

$$B = \begin{bmatrix} 1-r & r & r \\ r & r & 2-5r \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}.$$

For all $r, BX = X$ and $Bu = (1 + r, 2 - 3r, \frac{1}{2})^t$. For $0 < x < 1$ and $0 < r < \frac{1}{4}, A = xB$ is a nonnegative, irreducible matrix with $\rho(A) = x < 1$. Since A is 3×3, A satisfies condition (ii) of Corollary 4.2. Clearly $X_A \in \operatorname{int}(D)$. Note, however, that $Au \notin D$ when $0 < r < \frac{1}{4}$. From Theorem 1.2, $(I - xB)^{-1}u \approx Iu + xBu$ for small, positive x. Thus $(I - A)^{-1}u \notin D$ when $0 < r < \frac{1}{4}$.

5. Exploiting permutation invariance. In the context of this paper, circulant matrices have three important properties. If A is a circulant matrix, then u is an eigenvector for both A and $(I - A)^{-1}, \rho = \rho(A)$ is the unique row sum of A, and $(1 - \rho)^{-1}$ is the unique row sum of $(I - A)^{-1}$. Noting that the circulant matrices are precisely those matrices invariant under permutation similarity by the matrix for the full cycle permutation, we now examine how any permutation invariance can be exploited.

Let A be an $n \times n$, nonnegative, irreducible matrix with spectral radius $\rho(A)$. Suppose that P is a permutation matrix such that $PAP^t = A$. Clearly, $Pu = u$. Furthermore, $P(Au) = Au$ and $P[(I - A)^{-1}u] = (I - A)^{-1}u$. Also, $PAP^t = A$ implies PX_A is a positive eigenvector for $\rho(A)$ with norm one, hence $PX_A = X_A$. Thus the cycle structure of P is reflected in a pattern of constant blocks in the vectors X_A, Au, and $(I - A)^{-1}u$.

Assume that the permutation corresponding to P decomposes into k disjoint cycles. Let $\mathcal{V}$ denote the eigenspace for P for the eigenvalue $\lambda = 1$. Then $\mathcal{V}$ is a k-dimensional subspace of $\mathbb{R}^n$ with a natural basis consisting of certain $\{0, 1\}$ vectors. See [SW, §3]. Furthermore, $u_n \in \mathcal{V}$. Since A and P commute, $\mathcal{V}$ is an A-invariant space. Let M be the $k \times k$ matrix representing the restriction of A to $\mathcal{V}$ with respect to the natural basis. The following can be proven:

(i) M is a nonnegative, irreducible matrix with $\rho(M) = \rho(A)$.

(ii) Each entry of Mu_k is the value of all of the entries in the corresponding block of Au_n.

(iii) Each entry of $(I_k - M)^{-1}u_k$ is the value of all of the entries in the corresponding block of $(I_n - A)^{-1}u_n$.

(iv) There is a normalizing scalar $c > 0$ such that each entry of cX_M is the value of all of the entries in the corresponding block of X_A.

It follows immediately that the gradings for X_M, Mu, and $(I - M)^{-1}u$ lift naturally to gradings for X_A, Au, and $(I - A)^{-1}u$. Finally, when P has the form given by (3.1) of [SW], the matrix A naturally block partitions into blocks $A_{\langle i,j \rangle}$ for $1 \leq i, j \leq k$, and $M = [m_{ij}]$ is determined uniquely by $m_{ij} = (h_i)^{-1}[u_{h_i}]^t A_{\langle i,j\rangle} u_{h_j}$ for $1 \leq i, j \leq k$. See [SW, §§3 and 4.]

REFERENCES

[BP] A. BERMAN AND R. PLEMMONS, *Nonnegative Matrices in the Mathematical Sciences*, Academic Press, New York, 1979.

[O] R. OLDENBURGER, *Infinite powers of matrices and characteristic roots*, Duke Math. J., 6 (1940), pp. 357–361.

[SW] J. STUART AND J. WEAVER, *Matrices that commute with a permutation matrix*, Linear Algebra Appl., 150 (1991), pp. 255–265.