A Note on Light Matrices

Robert E. Hartwig
Mathematics Department
North Carolina State University
Raleigh, North Carolina 27695-8205

Submitted by Hans Schneider

ABSTRACT

Some properties of light matrices are derived, and their relation to Perron matrices is investigated.

1. INTRODUCTION

A key procedure is the theory of non-negative matrices is the "lifting" of an inequality of the form

$$A\mathbf{u} \underset{\neq}{\geqslant} \mathbf{b}$$

to a strict inequality $A\mathbf{v} > \mathbf{c}$. This is usually done by premultiplication of the inequality by a *positive* matrix X which commutes with A, yielding $AX\mathbf{u} > X\mathbf{b}$. Such matrices we shall call "liftable" or simply "light". To be precise:

DEFINITION. $A \in \mathbf{C}_{n \times n}$ is called light if there exists a positive $n \times n$ matrix X such that $AX = XA$.

In this note we wish to investigate (possibly real or nonnegative) light matrices. This class of matrices generalizes the so called Perron matrices, which are defined by [7, 2]:

DEFINITION. $A \in \mathbf{C}_{n \times n}$ is a Perron matrix if there exists a polynomial $p(\lambda)$ such that $p(A) > 0$.

It should be clear that if A is a real Perron matrix, then the polynomial $p(\lambda)$ may be taken to be real. Likewise, if A is a nonnegative Perron matrix,

52 Vanderbilt Ave., New York, NY 10017 0024-3795/87/$3.50

then the coefficients in $p(\lambda)$ may without loss of generality be taken to be nonnegative, as well as that $p(\lambda)$ is nonconstant. It is evident that a (real/nonnegative) Perron matrix is (real/nonnegative) light.

An $n \times n$ matrix is called reducible (under permutation similarity) if

$$P^{-1}AP = \begin{bmatrix} A_1 & A_3 \\ 0 & A_4 \end{bmatrix}$$

for some permutation matrix P, with A_1 and A_4 square and nonempty. It is well known that A is irreducible (i.e. not reducible) exactly when the adjacency graph G_A is strongly connected. For a nonnegative matrix this implies that $(I+A)^{n-1}$ is a positive matrix which commutes with A. Hence, it is clear that if $A \geqslant 0$, then

$$\text{Perron} \Leftrightarrow \text{irreducible} \Rightarrow \text{light}. \tag{1.1}$$

There are some obvious differences between irreducible and light matrices. For example, if A is irreducible, and $p(\lambda)$ is a polynomial, then $p(A)$ need not be irreducible—e.g.

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad p(\lambda) = 1 + \lambda^2;$$

on the other hand, if A is light, so is $p(A)$.

The spectral radius of a complex matrix A will be denoted by $r(A)$, and any nonnegative eigenvector corresponding to $r(A)$ will be called a Perron vector. The principal idempotent associated with $\lambda_1 = r(A)$ will be denoted by $Z_1^\circ(A)$, and the characteristic and minimal polynomials of A will be given by $\Delta_A(\lambda)$ and $\Psi_A(\lambda)$ respectively. As usual, permutation similarity will be indicated by $\underset{p}{\approx}$, while the algebraic and geometric multiplication of λ_0 will be indicated by $n(\lambda_0)$ and $\nu(\lambda_0)$ respectively. For convenience we call a polynomial nonnegative if its coefficients are nonnegative, and we shall use the usual notation: $A \geqslant B$ if $a_{ij} \geqslant b_{ij}$ for all i, j; $A \underset{\neq}{\geqslant} B$ if $A \geqslant B$ and $A \neq B$; and $A > B$ if $a_{ij} > b_{ij}$ for all i and j.

2. PRELIMINARIES

First we shall need some preliminary results.

PROPOSITION 1. *Let* $A \geqslant 0$, $X > 0$. *Then* $\mathrm{Tr}(AX) = 0 \Leftrightarrow A = 0$.

Proof. $\Leftarrow$: Clear. $\Rightarrow$: If some $a_{ij} > 0$, then $\mathrm{Tr}(AX) \geqslant e_i^T AX e_i = \sum_k a_{ik} x_{ki} \geqslant a_{ij} x_{ji} > 0$. ■

COROLLARY 1. *Let*

$$A = \begin{bmatrix} A_1 & A_3 \\ 0 & A_4 \end{bmatrix} \underset{\neq}{\geqslant} 0 \quad \text{and} \quad X = \begin{bmatrix} X_1 & X_3 \\ X_2 & X_4 \end{bmatrix} > 0.$$

Then $AX = XA \Rightarrow A_3 = 0$ *and* $r(A_1) = r(A_4)$.

Proof.

$$AX = \begin{bmatrix} A_1X_1 + A_3X_2 & ? \\ ? & ? \end{bmatrix} = XA = \begin{bmatrix} X_1A_1 & ? \\ ? & ? \end{bmatrix}.$$

Hence, on equating (1,1) blocks we have $A_1X_1 - X_1A_1 = -A_3X_2$ and thus $\mathrm{Tr}(A_3X_2) = \mathrm{Tr}(A_1X_1 - X_1A_1) = 0$. By Proposition 1, $A_3 = 0$. Now equating (1,2) blocks yields $A_1X_3 = X_3A_4$ with $X_3 > 0$. An application of the following result ensures that $r(A_1) = r(A_4)$. ■

PROPOSITION 2. *If* $A \geqslant 0$, $B \geqslant 0$, *then each of the following conditions implies the next*:

(a) $AX = XB$ *has a positive solution* X.
(b) $r(A) = r(B)$.
(c) $AX = XB$ *has a nonnegative rank-one solution.*

Proof. (a) $\Rightarrow$ (b): Suppose $\alpha = r(A) > r(B) = \beta$. Then $(B/\alpha)^N \to 0$, and thus $(A/\alpha)^N X = X(B/\alpha)^N \to 0$. Since $X > 0$, this forces $(A/\alpha)^N \to 0$, which is impossible. Thus $\alpha \leqslant \beta$. Symmetry now yields $\alpha = \beta$.

(b) $\Rightarrow$ (c): From the Perron-Frobenius theorem, it follows that there exist Perron vectors

$$\mathbf{u} \underset{\neq}{\geqslant} \mathbf{0}, \qquad \mathbf{v} \underset{\neq}{\geqslant} \mathbf{0}$$

such that $A\mathbf{u} = r\mathbf{u}$, $B^T\mathbf{v} = r\mathbf{v}$. Now $X = \mathbf{u}\mathbf{v}^T \geqslant 0$ will solve the desired equation.

REMARKS.

(i) The above conditions include the case $A = 0$ or $B = 0$.
(ii) In general (c) $\not\Rightarrow$ (b), as seen from the example

$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix},$$

in which $\alpha = r(A) = 2 < 3 = r(B) = \beta$. Yet

$$\mathrm{rank}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = 1.$$

This means that (c) $\nRightarrow$ (a) in general.

(iii) In general (b) $\nRightarrow$ (a). For example, the equation

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} X = X \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

has no positive solutions, yet $\alpha = \beta = 2$ and all nonnegative solutions have rank equal to one, and are of the form $X = \mathbf{e}\mathbf{e}_1^T$, with $\mathbf{e} = [1,1,\dots,1]^T$.

Proposition 3. *Let $A \geqslant 0$, $B \geqslant 0$, and $r(A) = r(B) = r$. If $p(\lambda)$ is a nonnegative polynomial then*

$$r(p(A)) = r(p(B)) = p(r).$$

Proof. If the distinct eigenvalues of A are $\{\lambda_1,\dots,\lambda_s\}$, then the distinct eigenvalues of $p(A)$ are among the $\{p(\lambda_k)\}$, $k = 1,2,\dots,s$. Now

$$|p(\lambda_k)| = \left| \sum_{i=0} p_i \lambda_k^i \right| \leqslant \sum p_i |\lambda_k|^i \leqslant \sum p_i r^i = p(r).$$

Thus $r(p(A)) = p(r) = r(p(B))$. Since A and B are nonnegative, it follows that $p(r)$ is an eigenvalue of $p(A)$ as well as $p(B)$. ■

It should be noted that this result is not true if some $p_i < 0$. For example,

$$r\begin{bmatrix} 0 & 1 \\ 4 & 0 \end{bmatrix} = r\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = 2,$$

while with $p(\lambda) = 1 - \lambda$, the spectral radii drop to 3 and 1 respectively.

Proposition 4. *If $A \underset{\neq}{\geqslant} 0$ is irreducible and $\mathbf{x} \geqslant 0$, then $A\mathbf{x} = \mathbf{0} \Rightarrow \mathbf{x} = \mathbf{0}$.*

Proof. $A\mathbf{x} = \mathbf{0} \Rightarrow (A + \cdots + A^n)\mathbf{x} = \mathbf{0}$. But $A + \cdots + A^n > 0$, and thus $\mathbf{x} = \mathbf{0}$. ■

Proposition 5. *If*

$$A = \begin{bmatrix} A_1 & & & & C_1 \\ & A_2 & & & C_2 \\ & & \ddots & & \vdots \\ & & & A_{s-1} & C_{s-1} \\ 0 & & & & A_s \end{bmatrix} \underset{p}{\approx} \begin{bmatrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_t \end{bmatrix} = B$$

and B_i, $i = 1,\dots,t$, are irreducible, then $C_j = 0$, $j = 1,2,\dots,s-1$.

Proof. Reduce each A_{ii} to its Frobenius normal form [3, p. 90],

$$\begin{bmatrix} A_1^{(i)} & & ? \\ & \ddots & \\ 0 & & A_{k_i}^{(i)} \end{bmatrix},$$

with $A_j^{(i)}$ irreducible. Now if some $C_i \neq 0$, then we get a node in the condensed graph of A which is *not* maximal. This is impossible, since all nodes are maximal on account of the form of B. ∎

PROPOSITION 6. *Suppose*

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ & A_{22} & & A_{2s} \\ & & \ddots & \vdots \\ 0 & & & A_{ss} \end{bmatrix}$$

is nonnegative, with $A_{ii} \neq 0$, irreducible, and square. Let $p(\lambda) = p_1\lambda + \cdots + p_N\lambda^N \not\equiv 0$ be a nonnegative polynomial. Also let A^k and $p(A)$ have corresponding blocks $A_j^{(k)}$ and $(p(A))_{ij}$. Then

(i) $A_{ij} \neq 0 \Rightarrow A_{ij}^{(k)} \neq 0.$
(ii) $A_{ij} \neq 0 \Rightarrow (p(A))_{ij} \neq 0.$

Proof. (i): We use induction on k. It is clearly true for $k = 1$, so let us assume $A^{(k)}ij \neq 0$. Then $A_{ij}^{(k+1)} = A_{i1}A_{ij}^{(k)} + \cdots \geqslant A_{ii}A_{ij}^{(k)}$. Using Proposition 4, we may conclude that $A_{ii}A_{ij}^{(k)} \neq 0$, as desired.

(ii): $(p(A))_{ij} = \sum_{k=1}^{N} p_k(A^k)_{ij} \neq 0$, since some $p_k > 0$ and $A_{ij}^{(k)} \neq 0$ for all $k \geqslant 1$.

PROPOSITION 7. *Let*

$$A = \left[\begin{array}{cccc|ccc} A_1 & & S_1 & & & C_1 & \\ & A_2 & S_2 & & & C_2 & \\ & & \ddots \ \vdots & & & \vdots & \\ 0 & & A_q & & & C_q & \\ \hline & & & & D_{q+1} & & 0 \\ & & 0 & & & \ddots & \\ & & & & & & D_s \end{array}\right] \tag{2.1}$$

be in Frobenius normal form with A_i, D_j *irreducible and* $[S_i, C_i] \neq 0$. *Suppose further that* $A_i \neq 0$, $D_j \neq 0$, *and* $p(\lambda) = p_1\lambda + \cdots + p_N\lambda^N \not\equiv 0$ *is a nonnegative polynomial.*

If $p(A) \underset{p}{\approx} \mathrm{dg}(B_1, \dots, B_t) = B$, *with* B_i *irreducible, then* $A_1, \dots, A_q$ *are absent and*

$$A \underset{p}{\approx} \mathrm{dg}(D_1, \dots, D_s).$$

Proof. Suppose $A_1, \dots, A_q$ are present. Then

$$p(A)=\left[\begin{array}{cccc|ccc}
p(A_1) & & T_1 & & & E_1 & \\
& p(A_2) & T_2 & & & E_2 & \\
& & \ddots\ \ \vdots & & & \vdots & \\
0 & & & p(A_q) & & E_q & \\ \hline
& & & & p(D_{q+1}) & & 0 \\
& 0 & & & & \ddots & \\
& & & & 0 & & p(D_s)
\end{array}\right] \underset{p}{\approx} B.$$

It now follows from Proposition 5 that $[T_i, E_i] = 0$, and hence by Proposition 6 that $[S_i, C_i] = 0$, which is a contradiction. ■

PROPOSITION 8. *Let* $p(\lambda)$ *be nonnegative and nonconstant. Let* $r_1 > 0$ *and* $r_2 > 0$. *If* $p(r_1) = p(r_2)$ *then* $r_1 = r_2$.

Proof. $0 = p(r_1) - p(r_2) = \Sigma_{k=1}^{t} p_k(r_1^k - r_2^k) = (r_1 - r_2)\Sigma_{k=1}^{t} p_k(r_1^{k-1} + r_1^{k-2}r_2 + \cdots + r_2^{k-1})$. Since all coefficients are nonnegative and some $r_k > 0$, the summation is positive. This forces $r_1 = r_2$. ■

3. NONNEGATIVE LIGHT MATRICES

We are now ready to start our investigation of nonnegative light matrices.

THEOREM 1. *Let* $A \geqslant 0$ *be* $n \times n$. *The following are equivalent:*

(i) A *is light.*

(ii) $A \underset{p}{\approx} \mathrm{dg}(A_1, A_2, \dots, A_s)$, *with* A_i *irreducible, and* $r(A_i) = r(A) = r$.

(iii) *A and A^T have positive eigenvectors.*

(iv) *For A there is a real eigenvalue λ_0 with positive left and right eigenvectors.*

(v) *$A = 0$ or there is a nonnegative polynomial $p(\lambda)$, with $p(0) = 0$, such that $p(A) \underset{p}{\approx} \mathrm{dg}(B_1, \ldots, B_t)$ with $B_i > 0$ and $r(B_i) = r(B_1)$, $i = 1, 2, \ldots, t$.*

In which case (a) $\lambda_0 = r(A)$, (b) $Z_1^\circ(A) \underset{p}{\approx} \mathrm{dg}(E_1, \ldots, E_s)$, *with* $E_i^2 = E_i > 0$ *and* $\operatorname{rank} E_i = 1$, (c)$\Psi_A = (\lambda - r)^1 \chi(\lambda)$, $\Delta_A(\lambda) = (\lambda - r)^s \phi(\lambda)$, $\chi(r) \neq 0 \neq \phi(r)$.

Proof. If $A \equiv 0$, all parts are trivial, so let $A \neq 0$.

(i) $\Rightarrow$ (ii): Without loss of generality let

$$A = \begin{bmatrix} A_1 & & ? \\ & \ddots & \\ 0 & & A_s \end{bmatrix} \quad \text{and} \quad X = \begin{bmatrix} X_{11} & \cdots & X_{1s} \\ \vdots & & \vdots \\ X_{s1} & \cdots & X_{ss} \end{bmatrix},$$

where the A_i are irreducible. We now use induction on s. For $s = 2$ the result follows from Corollary 1. Now assume that the result holds for $s - 1$ diagonal blocks. Partition A as

$$A = \begin{bmatrix} \tilde{A} & \tilde{T} \\ 0 & A_s \end{bmatrix} \underset{\neq}{\geqslant} 0,$$

and let

$$X = \begin{bmatrix} \tilde{X} & ? \\ ? & X_{ss} \end{bmatrix} > 0.$$

Then by Corollary 1, $AX = XA \Rightarrow \tilde{T} = 0$ and $r(\tilde{A}) = r(A_s)$. Hence $\tilde{A}\tilde{X} = \tilde{X}\tilde{A}$ and $A_s X_{ss} = X_{ss} A_s$. By the hypothesis,

$$\tilde{A} \underset{p}{\approx} \mathrm{dg}(A_1, \ldots, A_{s-1})$$

with A_i irreducible and $r(A_i) = r(A)$, $i = 1, 2, \ldots, s-1$. This means that $r(A_i) = r(\tilde{A}_i)$ for all $i = 1, \ldots, s$, and so $A \underset{p}{\approx} \mathrm{dg}(A_1, \ldots, A_s)$ with all A_i irreducible and $r(A_i) = r(A)$. In the special case where A is light and nilpotent, we have $r(A)=0$. The above proof then shows that each $A_i=0$, and nilpotent, we have $r(A)=0$.

The equivalence of (ii), (iii), and (iv) follows at once from Theorem 6, p. 92 of [3], or Theorem 3.14, p. 41 of [9].

(iv) $\Rightarrow$ (i): If $A\mathbf{u}=\lambda_0\mathbf{u}$ and $\mathbf{v}^TA=\lambda_0\mathbf{v}^T$, with $\mathbf{u}>\mathbf{0}$, $\mathbf{v}>\mathbf{0}$, then $X=\mathbf{u}\mathbf{v}^T>\mathbf{0}$ suffices.

(ii) $\Rightarrow$ (v): If $r(A_i)=0$, then $A=0$ and we are done. So assume $r(A_i)=r>0$. Now because each A_i is irreducible, the polynomial $p(\lambda)=\lambda(1+\lambda)^{n-1}$ will do. Also, because $r(A_i)=r(A)=r$, we have by Proposition 3 that $r(p(A_i))=p(r)$, where $r=r(A)$. Hence each of the matrices $B_i=p(A_i)$ has the same spectral radius.

(v) $\Rightarrow$ (ii): Suppose $A\neq 0$, and let $p(A)\underset{p}{\approx}\mathrm{dg}(B_1,\dots,B_t)=B$, where $B_i>0$, $r(B_i)=s_1$, and $p(0)=0$. Also let $r(A)=r$. Then clearly $p(\lambda)\not\equiv 0$, and by Proposition 3, $s_1=r[p(A)]=p(r)$. Now let A be permuted to its Frobenius normal form $\tilde{A}$, as given in (2.1). Thus $p(\tilde{A})\underset{p}{\approx}B$. Since $p(0)\neq 0$ and $B_i>0$, we may conclude that $p(A_i)$ and $p(D_j)$ have positive diagonal elements and hence $A_i\neq 0$, $D_j\neq 0$. This means that Proposition 7 may be applied to $\tilde{A}$, to yield

$$A\underset{p}{\approx}\tilde{A}=\mathrm{dg}(D_1,\dots,D_s).$$

Hence $p(A)\underset{p}{\approx}\mathrm{dg}(p(D_1),\dots,p(D_s))$, in which each $p(D_j)$ is light [recall (1.1)], yet need not be irreducible. Hence, using (i) $\Rightarrow$ (ii), we have for each $j=1,\dots,s$, $p(D_j)\underset{p}{\approx}\mathrm{dg}(G_1^{(j)},\dots,G_{l_j}^{(j)})$, in which the $G_k^{(j)}$ are all irreducible and have the same spectral radius for $k=1,2,\dots,l_j$. Also, because each $G_k^{(j)}$ is similar to some B_i, we may conclude that for all j and k, $r(G_k^{(j)})=s_1=r(B_1)$ and hence $r(p(D_j))=s_1\ \forall_j$. Now let $r_j=r(D_j)$. Then $p(r_j)=r(p(D_j))=r(p(D_i))=p(r_i)$, for all i and j. By Proposition 8 we have $r_i=r_j$, and (ii) follows.

Suppose now that the above hold. Then:

(a): From (iii) and (iv) it follows that $\lambda_0=r$.

(b): If A is as in (ii), then $Z_1^\circ(A)\underset{p}{\approx}\mathrm{dg}(E_1,\dots,E_s)$, where $E_i=Z_1^\circ(A_1)=\mathbf{u}_i\mathbf{v}_i^T/\mathbf{v}_i^T\mathbf{u}_i$ is the rank-one principal idempotent of A_i corresponding to $\lambda=r$, with $\mathbf{u}_i>\mathbf{0}$, $\mathbf{v}_i>\mathbf{0}$.

(c): This is clear, and completes the proof. ∎

Several remarks are now in place.

REMARKS.

(1) Theorem 1, part (v) generalizes the concept of a nonnegative Perron matrix, for which there exists a nonnegative polynomial $p(\lambda)$ with $p(0)=0$, so that $p(A)>0$.

(2) In Theorem 1, part (v), we cannot drop the condition that $r(B_i) = r(B_1)$. For example, if

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

then

$$p(A) = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix},$$

yet A is not light, since it has no positive Perron vector. Similarly, we cannot drop the condition that $p(0) = 0$, as seen from the example where

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad p(\lambda) = 1 - \lambda^2.$$

In this case $p(A) = I$ satisfies the conditions of (v), yet A is not light, since it is nilpotent.

(3) It is clear that Perron matrices are light. A larger class of matrices which contains the light matrices are those complex matrices for which there exists a polynomial $p(\lambda)$ with $p(0) = 0$ such that $p(A) \underset{p}{\approx} \mathrm{dg}(B_1, \ldots, B_s)$ with $B_i > 0$. It is not known whether these matrices can be characterized by means of a matrix equation.

(4) If A is light and $AX = XA$, $X > 0$, then X need not be a polynomial in A. That is, $\{\text{Perron}\} \underset{\neq}{\subset} \{\text{light}\}$. For example, if

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

then $\Delta_A = (\lambda + 1)^2(\lambda - 2)$, $\Psi_A = (\lambda + 1)(\lambda - 2)$, and $A^2 = A + 2I$. Now let

$$X = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 3 & 1 \\ 2 & 1 & 2 \end{bmatrix}.$$

Then $AX = XA$, $X > 0$, yet X cannot be a polynomial in A. Indeed, such a polynomial would have to be of the form

$$p_0 I + p_1 A = \begin{bmatrix} p_0 & p_1 & p_1 \\ & ? & \end{bmatrix},$$

which cannot equal X.

(5) The spectral component $Z_1^\circ(A)$ is always a polynomial in A, yet its polynomial coefficients need not be nonnegative. For example, if

$$A=\begin{bmatrix}\frac{3}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{3}{2}\end{bmatrix},$$

then

$$Z_1^\circ=\begin{bmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{1}{2}\end{bmatrix}=A-I.$$

Moreover, since $\Psi_A=(\lambda-1)(\lambda-2)$, no other polynomial representation of Z_1° can be nonnegative.

Corollary 2. *Let* $A \underset{\neq}{\geqslant} 0$ *be* $n\times n$. *The following are equivalent*:

(a) *If* A *is light and* $AX=XA$, $X>0$, *then* X *is a polynomial in* A.
(b) A *is irreducible and nonderogatory.*

Proof. (a) ⇒ (b): Since A is light, there exists $X>0$ such that $AX=XA$. By assumption we have $X=p(A)>0$ for some polynomial. Now by Theorem 1, we know that because A is light, $A \underset{p}{\approx} \mathrm{dg}(A_1,\dots,A_s)$, with A_i irreducible and $r(A_i)=r$. Hence $p(A)\underset{p}{\approx}\mathrm{dg}(p(A_1),\dots,p(A_s))$, which can never be positive if $s>1$. Hence $s=1$ and A is irreducible. Now suppose that A is derogatory. Then there exists a real matrix E such that $AE=EA$ and E is *not* a polynomial in A. If we set $X_\epsilon=(I+A)^{n-1}+\epsilon E$, then $AX_\epsilon=X_\epsilon A$, and for sufficiently small ϵ, $X_\epsilon>0$. Moreover, X_ϵ cannot be a polynomial in A. This would contradict (a), and thus A must be nonderogatory.

(b) ⇒ (a): From the standard theory we know that the only solutions to $AX=XA$ are polynomials in A. It is also clear that such matrices exist, e.g. $X=(I+A)^{n-1}$. ■

Let us now examine the related matrix equation $AX=XB$ with A and B nonnegative.

4. THE EQUATION $AX=XB$, WITH $A,B>0$

We now wish to investigate how the conditions of Theorem 1, for the existence of positive solutions to $AX=XA$, have to be generalized to guarantee the existence of positive solutions to $AX=XB$, with $A,B>0$. For convenience let us say that a matrix A has "cornered" Frobenius normal form if in (2.1), $r(A_i) < r = r(D_j)$, $i = 1,\dots,q$, $j = q+1,\dots,s$, that is, if the spectral radius of A is "cornered" in the (2.2) block of D. In other words, the singular vertices (classes) are precisely the maximal (also called final or essential) vertices (classes) under the access partial ordering.

PROPOSITION 9. *Let $A, B > 0$. The following are equivalent:*

(i) *$AX = XB$ has a positive rank-one solution X.*
(ii) *$AX = XB$ has a positive solution X.*
(iii) *A and B^T have positive eigenvectors and $r(A) = r(B)$.*
(iv) *A and B^T have cornered Frobenius normal forms and $r(A) = r(B)$.*

Proof. (i) $\Rightarrow$ (ii): Clear. (ii) $\Rightarrow$ (iii): From Proposition 2, we know that $r(A) = r(B) = r$. Using the Perron-Frobenius theorem, we may select $\mathbf{y}$ and $\mathbf{w}$ so that

$$A^T\mathbf{y} = r\mathbf{y} \text{ and } B\mathbf{w} = r\mathbf{w}, \qquad \text{with } \mathbf{y} \gneq \mathbf{0}, \quad \mathbf{w} \gneq \mathbf{0}.$$

Then $AX = XB$, $X > 0 \Rightarrow A\mathbf{u} = r\mathbf{u}$ and $B^T\mathbf{v} = r\mathbf{v}$, where $\mathbf{u} = X\mathbf{w} > 0$, $\mathbf{v}^T = \mathbf{y}^T X > \mathbf{0}^T$.

(iii) $\Rightarrow$ (i): If $A\mathbf{u} = r_1\mathbf{u}$, $B^T\mathbf{v} = r_2\mathbf{v}$, $\mathbf{u} > 0$, $\mathbf{v} > 0$, and $r(A) = r(B)$, then $r_1 = r(A) = r(B) = r_2$. Now use $X = \mathbf{u}\mathbf{v}^T > 0$.

(iii) $\Leftrightarrow$ (iv): It was shown in [3, p. 92] that A has a positive eigenvector exactly when its Frobenius normal form is concerned. ■

REMARKS.

(a) It may be shown that if the Frobenius form (2.1) is cornered, then all positive Perron vectors are given by

$$\begin{bmatrix} (rI - A')^{-1}C\mathbf{u} \\ \mathbf{u} \end{bmatrix}, \qquad \text{where} \quad \mathbf{u} = \begin{bmatrix} \alpha_{q+1}\mathbf{u}_{q+1} \\ \vdots \\ \alpha_s \mathbf{u}_s \end{bmatrix},$$

with $\alpha_i > 0$ and $D_i\mathbf{u}_i = r\mathbf{u}_i$, $\mathbf{u}_i > 0$.

(b) If $B = A$, then the conditions (i), (ii), and (iii) of Theorem 1 are recovered. In particular, if both A and A^T have cornered normal form, then only the D_i blocks can be present.

(c) It is not known if the existence of a positive eigenvector can be replaced by a polynomial condition.

(d) The idea of a "cornered" matrix also plays an important role in the study of the zero nonzero pattern of a Perron vector of a nonnegative matrix. See for example [1], [4], [5], [6], and [8].

COROLLARY 3. *The systems $AX = XB$ and $BY = YA$ have positive solutions if and only if A and B are light and $r(A) = r(B)$.*

Proof. $\Rightarrow$: By Proposition 6, A and B^T as well as B and A^T have positive eigenvectors, and $r(A) = r(B)$. Hence by Theorem 1, A and B are light.

$\Leftarrow$: By Theorem 1, we may without loss of generality consider

$$\begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_s \end{bmatrix} X = X \begin{bmatrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_t \end{bmatrix}.$$

Since A_i and B_j are irreducible, they have positive Perron vectors. Hence by Proposition 6, the equations $A_i X_{ij} = X_{ij} B_j$ with $r(A_i) = r(B_j)$ always have positive solutions. Similarly for $BY = YA$. ■

PROPOSITION 10. *Let $A, B > 0$. If $AX = XB$ has a nonnegative solution and A is light then $r(A) \leqslant r(B)$.*

Proof. Let $\alpha = r(A)$. If $r(B) < \alpha$, then $(A/\alpha)^N X = X(B/\alpha)^N \to 0$. Now $X \gneqq 0$ has a nonzero column x_i. Hence $(A/\alpha)^N x_i \to 0$. We may lift this to yield $(A/\alpha)^N \mathbf{p} \to 0$, where $\mathbf{p} = Y\mathbf{x}_i > 0$ and $AY = YA$, $Y > 0$. This yields $r(A/\alpha) < 1$, which is impossible. Thus $r(A) \leqslant r(B)$. ■

REMARKS.

(i) This result is no longer true if we replace the assumption that A is light by the weaker assumption that A only has a positive eigenvector. For example let

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} \frac{3}{2} & 1 \\ 0 & 1 \end{bmatrix}, \qquad X = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Then $A\mathbf{e} = 2\mathbf{e}$, $[\frac{1}{2}, 1]B = \frac{3}{2}[\frac{1}{2}, 1]$, and $AX = XB$. Yet $r(A) = 2 > \frac{3}{2} = r(B)$.

(ii) Strict inequality is possible in Proposition 10, even if A is irreducible. This may be seen from the example where

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \text{and} \qquad X = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

Here $r(A) = 1 < 2 = r(B)$.

COROLLARY 4. *If A and B are nonnegative light, and $AX = XB$ has a nonnegative solution, then $r(A) = r(B)$.*

Combining Proposition 9 with Corollary 4, we obtain:

COROLLARY 5. *If $A, B > 0$ are light, then the following are equivalent:*

(i) *there is a positive rank-one solution to $AX = XB$.*
(ii) *There is a positive solution to $AX = XB$.*
(iii) *There is a nonnegative solution to $AX = XB$.*
(iv) $r(A) = r(B)$.

5. REAL LIGHT MATRICES

We may now characterize the *real* light matrices.

THEOREM 2. *Let $A \in \mathbb{R}_{n\times n}$. The following are equivalent:*

(a) *A is real light.*
(b) *A and A^T have positive eigenvectors.*
(c) *there is a real eigenvalue λ_1 such that $A\mathbf{u} = \lambda_1\mathbf{u}$ and $A^T\mathbf{v} = \lambda_1\mathbf{v}$, with $\mathbf{u} > \mathbf{0}$, $\mathbf{v} > \mathbf{0}$.*

Proof. (a) $\Rightarrow$ (b): Again suppose $A \neq 0$. Suppose $AX = XA$, $X > 0$, and set $\alpha = r(A)$, $r = r(X)$. Now every eigenspace of X contains an eigenvector of A. In particular, if $X\mathbf{u} = r\mathbf{u}$, with $\mathbf{u} > \mathbf{0}$, then $AX\mathbf{u} = X(A\mathbf{u}) = r(A\mathbf{u})$. Since $A\mathbf{u} \neq \mathbf{0}$, and the eigenspace $N(X - rI)$ has dimension *one*, we may conclude that $A\mathbf{u} = \lambda_1\mathbf{u}$ for some λ_1. Clearly λ_1 must be real. Likewise if $X^T\mathbf{v} = r\mathbf{v}$, with $\mathbf{v} > \mathbf{0}$, then $A^TX^T\mathbf{v} = X^T(A^T\mathbf{v}) = r(A^T\mathbf{v})$, and $A^T\mathbf{v} = \lambda_2\mathbf{v}$ with λ_2 real.

(b) $\Rightarrow$ (c): If $A\mathbf{u} = \lambda_1\mathbf{u}$, $A^T\mathbf{v} = \lambda_2\mathbf{v}$, with $\mathbf{u} > \mathbf{0}$, $\mathbf{v} > \mathbf{0}$, then $\lambda_2\mathbf{v}^T\mathbf{u} = \mathbf{v}^TA\mathbf{u} = \lambda_1\mathbf{v}^T\mathbf{u}$ and thus $\lambda_1 = \lambda_2$.

(c) $\Rightarrow$ (a): $X = \mathbf{u}\mathbf{v}^T$ will do. ■

In general $|\lambda_1|$ need not equal $r(X)$. Furthermore A may be reducible and the geometric and algebraic multiplicities of λ_1 may exceed unity. For example, let

$$A = \begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad X = \mathbf{e}\mathbf{e}^T > 0.$$

Then $AX = XA = 0$, $\mathbf{u} = \mathbf{v} = \mathbf{e}$, and $\lambda_1 = 0$. Moreover, $\Delta_X = \lambda^2(\lambda - 3)$, $r = 3$, $\Delta_A = \lambda^2(\lambda + 2)$. Since $\text{rank}(A) = 1$, the nullity of A equals 2 and thus both multiplicities of λ_1 equal 2. Unlike in the nonnegative light case, no canonical form seems possible for general real light matrices.

We now return to the Perron matrices. The following simplifies and extends Theorem 2.7 of [7].

PROPOSITION 11. *Let A be real $n \times n$ with $n \geqslant 2$. The following are equivalent:*

(a) *A is a real Perron matrix.*

(b) *A has a real eigenvalue λ_1 with algebraic multiplicity $n(\lambda_1) = 1$ and associated positive left and right eigenvectors $\mathbf{u} > \mathbf{0}$, $\mathbf{v} > \mathbf{0}$.*

(c) *There is an eigenvalue λ_1 of A for which $Z_1^\circ > 0$.*

In which case A is irreducible.

Proof. (a) $\Rightarrow$ (b): Let $X = p(A) > 0$, where $p(\lambda)$ is a polynomial (with real coefficient). Clearly $A \neq 0$ and $p(\lambda) \neq p_0$. From Theorem 2 we have $A\mathbf{u} = \lambda_1 \mathbf{u}\, A^T\mathbf{v} = \lambda_1 \mathbf{v}$, where $\mathbf{u}, \mathbf{v} > \mathbf{0}$ and $p(A)\mathbf{u} = r\mathbf{u}$, $p(A^T)\mathbf{v} = r\mathbf{v}$. Here $r = r(p(A))$ is a simple eigenvalue of $p(A)$, of algebraic multiplicity equal to unity. Now $A\mathbf{u} = \lambda_1 \mathbf{u} \Rightarrow p(A)\mathbf{u} = p(\lambda_1)\mathbf{u} = r\mathbf{u}$. Hence $r = p(\lambda_1)$. Since each eigenvalue of $p(A)$ is of the form $p(\lambda_1)$ for some eigenvalue λ_i of A, we may conclude that λ_1 must be a simple eigenvalue of A with algebraic multiplicity one.

(b) $\Rightarrow$ (c): Let $A\mathbf{u} = \lambda_1\mathbf{u}$, $A^T\mathbf{v} = \lambda_1\mathbf{v}$, $\mathbf{u} > \mathbf{0}$, $\mathbf{v} > \mathbf{0}$, and let $n(\lambda_1) = 1$. Then $Z_1^\circ(\lambda_1) = \mathbf{u}\mathbf{v}^T/\mathbf{v}^T\mathbf{u} > \mathbf{0}$.

(c) $\Rightarrow$ (a): $Z^\circ{}_1$ is always a polynomial in A, with real coefficients.

Lastly, if A were reducible, so would $p(A)$ be and hence it could not be positive. ∎

EXAMPLE. If

$$A = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \quad \text{and} \quad p(\lambda) = 1 + 2\lambda$$

then

$$p(A) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} > 0.$$

In this case $\mathbf{u} = \mathbf{v} = e$, and $\lambda_1 = 0 < 2 = r$. Clearly A is irreducible.

It should be noted that a Perron matrix need not be nonderogatory, as seen from the matrix

$$B=\begin{bmatrix}0&1&1\\1&0&1\\1&1&0\end{bmatrix}$$

with $\Psi_B(\lambda)=(\lambda+1)(\lambda-2)$ and $\Delta=(\lambda+1)^2(\lambda-2)$. Here $A^2>0$.

Let us now turn to the complex case.

PROPOSITION 12. *The following are equivalent:*

(a) *A is complex light.*
(b) *A and A^T have positive eigenvectors.*
(c) *A and A^* have positive eigenvectors.*

Proof. Let $A=B+iC$, with B,C real. (a) $\Rightarrow$ (b): If $AX=XA$, $X>0$, then $BX=XB$ and $CX=XC$. Hence by Theorem 2, if $X\mathbf{u}=r\mathbf{u}$, $\mathbf{u}>\mathbf{0}$, then $B\mathbf{u}=\mu_1\mathbf{u}$ and $C\mathbf{u}=\nu_1\mathbf{u}$. Setting $\lambda_1=\mu_1+i\nu_1$, we get $A\mathbf{u}=(B+iC)\mathbf{u}=\lambda_1\mathbf{u}$. Similarly, if $X^T\mathbf{v}=r\mathbf{v}$, $\mathbf{v}>\mathbf{0}$, then $B^T\mathbf{v}=\mu_1\mathbf{v}$, $C^T\mathbf{v}=\nu_1\mathbf{v}$, and $A^T\mathbf{v}=\lambda_1\mathbf{v}$.

(b) $\Rightarrow$ (a): If $A\mathbf{u}=\lambda_1\mathbf{u}$ and $A^T\mathbf{v}=\lambda_2\mathbf{v}$ with $\mathbf{u}>0$, $\mathbf{v}>\mathbf{0}$, then again $\lambda_1=\lambda_2$ and $X=\mathbf{u}\mathbf{v}^T$ will do in $AX=XA$.

(a) $\Rightarrow$ (c): As in the first section, we get $A\mathbf{u}=\lambda_1\mathbf{u}$, $\mathbf{u}>\mathbf{0}$ while $A\mathbf{v}=(B^T-iC^T)\mathbf{v}=\lambda_1\mathbf{v}$, $\mathbf{v}>\mathbf{0}$.

(c) $\Rightarrow$ (a): If $A\mathbf{u}=\lambda_1\mathbf{u}$ and $A^*\mathbf{v}=\lambda_1\mathbf{v}$, with $\mathbf{u}>\mathbf{0}$, $\mathbf{v}>\mathbf{0}$, then $\mathbf{v}^TA=\mathbf{v}^T\lambda_2$ and thus $\lambda_1\mathbf{v}^T\mathbf{u}=\mathbf{v}^TA\mathbf{u}=\mathbf{u}^TA^T\mathbf{v}=\mathbf{u}^TA^*\mathbf{v}=\lambda_2\mathbf{v}^T\mathbf{u}$, and hence $\lambda_2=\lambda_1$. Now $A\mathbf{u}\mathbf{v}^T=\lambda_1\mathbf{u}\mathbf{v}^T=\lambda_2\mathbf{u}\mathbf{v}^T=\mathbf{u}\mathbf{v}^TA$, as desired. ■

PROPOSITION 13. *Let $n\geqslant 2$. The following are equivalent:*

(a) *A is Perron.*
(b) *A has eigenvalue λ_1 with $n(\lambda_1)=1$ and positive left and right eigenvectors.*
(c) $Z_1^\circ(\lambda_1)>0$ *for some eigenvalue* λ_1.

In which case A is irreducible.

The proof is similar to that of Proposition 12 and is omitted.

6. DISCUSSION AND OPEN QUESTIONS

We have seen that the study of light matrices is closely related to the study of the matrix equation $AX=XB$, with $A,B>0$. The existence of positive solutions has been fairly well characterized in Proposition 9. On the other hand, it is not known how to characterize *all* positive solutions to this equation in a satisfactory manner, even if A and B are irreducible and nonnegative.

A related problem is the following: If $A, B \geqslant 0$, when does $AX = XB$ have *nonnegative nonzero* solutions X? If so, what do they look like? It suffices that $r(A) = r(B)$; however, this is not necessary. It is further not necessary for AX or XB to be a multiple of X, as seen from

$$\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}.$$

We may recapitulate this in the following table:

SOLUTIONS TO $AX = XB$

	Existence of $X > 0$	Characterize all $X > 0$	Existence of $X \underset{\neq}{\geqslant} 0$	Characterize all $X \underset{\neq}{\geqslant} 0$
$A, B \geqslant 0$	Known	Open	Open	Open
$A, B \geqslant 0$ irreducible	Known	Open	Known	Open

A related and more formidable problem is to find all positive (nonnegative) solutions to $AX - XB = C$ with $A, B \geqslant 0$.

Lastly, it would be of interest to know whether part (b) of Theorem 1 characterizes nonnegative light matrices.

The author wishes to thank Hans Schneider for several valuable suggestions, resulting in an improved presentation of Theorem 1.

REFERENCES

1 D. H. Carlson, A note on M-matrix equations, *SIAM J. Appl. Math.* 11:1027–1033 (1963).

2 J. J. Dionisio, Matrizes nao negatives, *Rev. Fac. Ciencias Lisboa Ser. A* 10:5–35.

3 F. R. Gantmacher, *The Theory of Matrices*, Vol. 2, Chelsea, New York, 1960.

4 D. J. Richman and H. Schneider, On the singular graph and the Weyr characteristic of an M-matrix, *Aequationes Math.* 17:208–234 (1978).

5 U. G. Rothblum, Algebraic eigenspaces of non-negative matrices, *Linear Algebra Appl.* 12:281–292 (1975).

6 H. Schneider, The elementary divisors associated with 0, of a singular M-matrix, *Proc. Edinburgh Math. Soc.* 10:108–122 (1956).
7 E. Seneta, *Nonnegative Matrices and Markov Chains*, 2nd ed., Springer, New York, 1981.
8 H. D. Victory, On non-negative solutions of Matrix Equations, *SIAM J. Algebraic Discrete Methods* 6:406–411 (1985).
9 A. Berman and R. J. Plemmons, *Nonnegative Matrices in Mathematical Sciences*, Academic, New York, 1979.

Received 9 July 1985; revised 25 March 1987