The Spectrum of a Nonlinear Operator Associated with a Matrix*†

M. V. MENON
University of Missouri, Columbia, Missouri

AND

HANS SCHNEIDER
University of Wisconsin, Madison, Wisconsin

1. INTRODUCTION

1.1. *Notation and conventions.* In this paper A (or B) will always denote an $m \times n$ matrix with non-negative elements and with no zero row or column, r will denote an $m \times 1$ column vector (!) with all elements positive and c a $1 \times n$ row vector.

Following the notation of [2], we write $M \geqslant 0$ if M is a matrix and all $m_{ij} \geqslant 0$; if $M \geqslant 0$ but $M \neq 0$, and we call M a *positive* matrix (thus a positive matrix may have zero entries). We write $M \gg 0$ if all $m_{ij} > 0$, and in this case we call M *strictly positive*.

1.2. *Introduction.* Let (A, r, c) be as in Section 1.1. We shall define a nonlinear homogeneous operator $T = T(A; r, c)$ on the positive cone $\mathscr{P} = \{x = (x_1, \ldots, x_n) : x_i \geqslant 0\}$ and we determine the spectrum of T and all positive-zero patterns of eigenvectors of T. Since, by definition, T is an operator of $\mathscr{P}$ into itself, all eigenvalues are necessarily non-negative and all eigenvectors lie in $\mathscr{P}$. Our operator T is an obvious modification, to take into account the vectors r and c, of the operator introduced by Menon in [4] for the case that A is square and strictly positive. This

* Dedicated to Alexander Ostrowski, on the occasion of his 75th birthday.

† Sponsored, in part, by the Mathematics Research Centre, United States Army, University of Wisconsin, Madison, Wisconsin, under contract No: DA-31-124-ARO-D-462 and by NSF grant NSF GP-7073.

Linear Algebra and Its Applications 2(1969), 321–334

operator was exploited by Brualdi, Parter, and Schneider in [2] in the case that A may have zero entries.

We feel that some explanation should be given for considering a special problem of this kind. Our first justification is that our results resemble the Perron-Frobenius theory. It is well known that, for an irreducible square matrix M with non-negative elements, there is a unique eigenvector x in $\mathscr{P}$ (except for scalar multiples), that all x_i are positive, and, of course, the associated eigenvalue ρ (the Perron-Frobenius root of M) is positive. Given the location of the zeros of a reducible square matrix M, it is then possible to determine the positive zero pattern of each eigenvector x of M with non-negative elements (i.e., whether $x_i > 0$ or $x_i = 0$ for given index i) and also the associated eigenvalues (cf. Schneider [8]). Our main theorem (3.6) is an analog. This similarity is not too surprising. One may define a more general operator $T(A, B; r, c)$ associated with a pair of matrices (cf. (2.2)) and then $T(A, I; e, e)$, where e is a vector with all $e_i = 1$, is the linear operator usually represented by the matrix A. Properties of the operator $T(A, B; r, c)$ are investigated by Menon in a recent paper [6].

Our second justification is perhaps more important. The operator $T = T(A; r, c)$ is so constructed that T has a positive eigenvalue λ with strictly positive eigenvector x: $Tx = \lambda x$ if and only if there exist diagonal matrices Y and X with positive diagonal elements for which YAX has row sums r_i, $i = 1, \ldots, m$, and column sums λc_i, $i = 1, \ldots, n$. Several authors have recently worked on the problem of the existence of such Y and X. We wish to mention Sinkhorn's paper [9] which solved the problem for A strictly positive, square and $r = c = e$. Necessary and sufficient conditions for the existence of Y and X for a square A, which may have zero elements, and $r = c = e$, were found by Sinkhorn and Knopp [11] and by Brualdi, Parter, and Schneider [2]. Recent work that should be mentioned is due to Menon [5], Sinkhorn [10], Mirsky and Perfect [7], and Brualdi [1]. Employing methods used in flow and network problems, Brualdi proved a result which was part of the inspiration of the present paper (cf. Section 4).

2. THE OPERATOR $T(A; r, c)$

2.1. DEFINITIONS. Let A, B, r, and c be as in Section 1.1. We shall call (A, r, c) a *matrix-rowsum-columnsum triple* or *mrc* for short. Similarly

(A, B, r, c) will be called a *matrix-matrix-rowsum quadruple* or *mmrc*. The following definition is also given in [5].

2.2. DEFINITION. Let (A, B, r, c) be an mmrc. For a positive column or zero vector $x = (x_1, \ldots, x_n)$, we define

$$S(A; r)x = y \tag{2.2.1}$$

where

$$y_i = r_i \Big/ \left(\sum_{j=1}^{n} a_{ij} x_j\right), \qquad i = 1, \ldots, m, \tag{2.2.2}$$

and then

$$T(A, B; r, c)x = z, \tag{2.2.3}$$

where

$$z_j = \frac{c_j}{\sum_{i=1}^{n} y_i b_{ij}}, \qquad j = 1, \ldots, n. \tag{2.2.4}$$

Except for one lemma, we shall be concerned with the case $A = B$. We write

$$T(A, A; r, c) = T(A; r, c). \tag{2.2.5}$$

When no confusion should arise, we write $T(A; r, c) = T$, and similarly $S(A; r) = S$. Here S is an operator of $\mathscr{P} = \{x : x \geqslant 0\}$ into the set

$$\mathscr{P}_\infty = \{y = (y_1, \ldots, y_n) : \ y_i \geqslant 0 \quad \text{or} \quad y_i = \infty\},$$

but T is a homogeneous operator of $\mathscr{P}$ into itself. We use the conventions $0^{-1} = \infty$, $\infty^{-1} = 0$, $\infty + \infty = \infty$, $0 \cdot \infty = 0$, and $a \cdot \infty = \infty$, for $a > 0$; see Section 3 of [2]. Further, T is a continuous operator of $\mathscr{P}$ into itself. The proof is essentially the same as that of (3.4) of [2].

2.3. LEMMA. *Let (A, r, c) be an* mrc. *Then $T = T(A; r, c)$ defined by Definition* 2.2 *has a positive eigenvalue*

$$\rho = \sup\{\lambda : \ \exists x > 0, \quad Tx \geqslant \lambda x\}. \tag{2.3.1}$$

Further, let E be the matrix with all $e_{ij} = 1$, and for $\varepsilon \geqslant 0$, set

$$A_\varepsilon = A + \varepsilon E \tag{2.3.2}$$

and

$$T_\varepsilon = T(A_\varepsilon, A; r, c). \tag{2.3.3}$$

If

$$\rho_\varepsilon = \sup\{\lambda\colon x > 0, \quad T_\varepsilon x \geqslant \lambda x\}, \tag{2.3.4}$$

then

$$\rho_\varepsilon \downarrow \rho \quad \text{as} \quad \varepsilon \to 0.$$

(By $\rho_\varepsilon \downarrow \rho$, we mean that ρ_ε is a decreasing function of ε in some interval $0 < \varepsilon < \partial$, and $\lim_{\varepsilon\to 0+} \rho_\varepsilon = \rho$.

Proof. For $\varepsilon > 0$, we have $A_\varepsilon \gg 0$, and so

$$T_\varepsilon x \gg T_\varepsilon x', \qquad \text{for} \qquad x > x' \geqslant 0. \tag{2.3.5}$$

We give a familiar argument (cf., e.g., Section 4 of [2]) to show that ρ_ε is an eigenvalue with a strictly positive eigenvector. Let

$$K = \{x > 0\colon \sum_i x_i = 1\}. \tag{2.3.6}$$

Since T_ε is homogeneous, (2.3.4) may be replaced by $\rho_\varepsilon = \sup\{\lambda\colon x \in K, T_\varepsilon x \geqslant \lambda x\}$, and it now follows from the compactness of K that there is a $u_\varepsilon \in K$ for which $T_\varepsilon u_\varepsilon \geqslant \rho_\varepsilon u_\varepsilon$. Suppose that $T_\varepsilon u_\varepsilon > \rho_\varepsilon u_\varepsilon$. In that case, by (2.3.5), $T_\varepsilon u_\varepsilon' \gg \rho u_\varepsilon'$, where $u_\varepsilon' = T_\varepsilon u_\varepsilon$, and so also $T_\varepsilon u_\varepsilon' \gg (\rho + \alpha) u_\varepsilon'$, for some $\alpha > 0$, contrary to (2.3.4). Hence

$$T_\varepsilon u_\varepsilon = \rho_\varepsilon u_\varepsilon. \tag{2.3.7}$$

Since $u_\varepsilon > 0$, it follows by (2.3.5) that $T_\varepsilon u_\varepsilon \gg 0$, whence also $u_\varepsilon \gg 0$ by (2.3.7).

We now turn to the proof that ρ is an eigenvalue of T. Since K is compact, there is a $u \in K$ and a sequence $\varepsilon(1), \varepsilon(2), \ldots$ with $\varepsilon(s) \downarrow 0$ and $u_{\varepsilon(s)} \to u$ as $s \to \infty$. We shall investigate the behavior of $T_{\varepsilon(s)}$. First, for each fixed $x \in K$, the operation $\varepsilon \to A_\varepsilon x$ is continuous for $0 \leqslant \varepsilon \leqslant 1$, and it then follows by an argument similar to that of Lemma 3.4 of [1] that $\varepsilon \to T_\varepsilon x$ is continuous in that domain. Since $T_\varepsilon x \downarrow Tx$ as $\varepsilon \downarrow 0$ ($\mathscr{P}$ being partially ordered as in [1]), it follows by a well-known theorem (Hobson [3], Vol. II, p. 116) that $T_\varepsilon x$ converges to Tx uniformly in ε over K. Clearly, therefore, $T_{\varepsilon(s)} x$ converges to Tx uniformly in s over K, and, since $u_{\varepsilon(s)} \to u$, as $s \to \infty$, it follows easily that $T_{\varepsilon(s)}u_{\varepsilon(s)} \to Tu$ as $s \to \infty$. But ρ_ε decreases as ε decreases (since T_ε is monotonic in ε), whence $\rho_\varepsilon \downarrow \sigma$, say, as $\varepsilon \downarrow 0$, where $\sigma \geqslant 0$. We now deduce from $T_{\varepsilon(s)}u_{\varepsilon(s)} = \rho_{\varepsilon(s)}u_{\varepsilon(s)}$ that $Tu = \sigma u$. We must still show that $\sigma = \rho$ and that $\rho > 0$. Since $\rho_\varepsilon \geqslant \rho$, if $\varepsilon > 0$, clearly $\sigma \geqslant \rho$. But, by (2.3.1), with $\varepsilon = 0$ and $Tu = \sigma u$, also $\rho \geqslant \sigma$. Hence $\rho = \sigma$. If $x \gg 0$, then $Tx \gg 0$, since A has no zero row or column. Thus, for λ sufficiently small and positive $Tx \geqslant \lambda x$, whence $\rho \geqslant \lambda > 0$.

2.4. *Remark and Example.* It may be worthwhile to elucidate the remark that $\varepsilon \to T_\varepsilon x$ is continuous on (0, 1). Using the notation of [1], we write $\mathscr{P}_\infty = \{x = (x_1, \ldots, x_k)\colon\ x_i \geqslant 0 \text{ or } x_i = \infty\}$, where $k = m$ or $k = n$, and let U be the mapping of $\mathscr{P}_\infty$ into itself given by $Ux = y$, where $y_i = x_i^{-1}$, $i = 1, \ldots, k$. Then we have a sequence of mappings

$$\varepsilon \to A_\varepsilon x \to RUA_\varepsilon x \to ARUA_\varepsilon x \to CUARUA_\varepsilon x = T_\varepsilon x, \qquad (2.4.1)$$

where $R = \operatorname{diag}(r_1, \ldots, r_m)$ and $C = \operatorname{diag}(c_1, \ldots, c_n)$. Now observe that addition is continuous on $(0, \infty)$ and so are $x \to cx$, $0 \leqslant c < \infty$, and $x \to x^{-1}$. However, $a \to ac$ is not continuous at 0 when $c = \infty$, nor is $(a, b) \to ab$ continuous at $(0, \infty)$. An inspection of the sequence (2.4.1) shows that lim ab with $a \to 0$, $b \to \infty$ does not occur there, and continuity follows. For contrast consider the operation $T_\varepsilon^* = T(A_\varepsilon, A_\varepsilon; r, c)$. We have a sequence

$$\varepsilon \to A_\varepsilon x \to RUA_\varepsilon x \to ARUA_\varepsilon x \to CUA_\varepsilon RUA_\varepsilon x = T_\varepsilon^* x, \qquad (2.4.2)$$

and the last operation may not be continuous in ε. As an example, let

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad r = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad c = [2, 1], \qquad \text{and let} \qquad x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Then

$$Tx = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \qquad T_\varepsilon x = \begin{bmatrix} 2(1+\varepsilon) \\ \dfrac{\varepsilon(\varepsilon+1)}{2\varepsilon+1} \end{bmatrix}$$

while

$$T_\varepsilon^* x = \begin{bmatrix} 1 \\ \dfrac{\varepsilon}{\varepsilon+1} \end{bmatrix}.$$

3. EIGENVALUES AND EIGENVECTORS OF $T(A; r, c)$

3.1. Notation and definitions. We shall put $M = \{1, \ldots, m\}$ and $N = \{1, \ldots, n\}$, and use I, J to denote nonempty subsets of M, N, respectively. We write I', J' for the complements of I, J in M, N, respectively (the M, N being understood from the context).

If A is a matrix, then $A[I|J]$ is defined to be the submatrix of A lying in all rows i and all columns j, with $i \in I$ and $j \in J$. Thus $A[M|N] = A$. If $A[I|J] = 0$ we call $A[I|J]$ a *zero submatrix* of A, and we call $A[I|J]$ a *maximal zero submatrix* if, in addition, $A[I_1|J_1] \neq 0$ when $I_1 \times J_1 \supset I \times J$ (we use $\supset$ for proper containment).

3.2. Definition. Let (A, r, c) be an mrc and, for all nonempty subsets I, J of M, N, respectively, let

$$\omega(I, J) = \sum_J c_j / \sum_I r_i. \tag{3.2.1}$$

Then (A, r, c) is called *consistent* if and only if for all nonempty proper subsets I, J of M, N, respectively, for which $A[I'|J] = 0$ we have

$$\omega(I, J) < \omega(M, N) \qquad \text{whenever} \qquad A[I|J'] \neq 0 \tag{3.2.2}$$

and

$$\omega(I, J) = \omega(M, N) \qquad \text{whenever} \qquad A[I|J'] = 0. \tag{3.2.3}$$

3.3. Lemma. *Let (A, r, c) be an* mrc. *If $T = T(A; r, c)$ has a strictly positive eigenvector u, then (A, r, c) is consistent, and the corresponding eigenvalue λ equals $\omega(M, N)$.*

Proof. Let $u \gg 0$ and $Tu = \lambda u$. If $v = Su$, then

$$\sum_N v_i a_{ij} u_j = r_i, \tag{3.3.1}$$

whence

$$\sum_{M,N} v_i a_{ij} u_j = \sum_M r_i. \tag{3.3.2}$$

Similarly, if $w = Tu$, then

$$\sum_M v_i a_{ij} w_j = c_j \tag{3.3.3}$$

and so

$$\sum_{M,N} v_i a_{ij} w_j = \sum_N c_j. \tag{3.3.4}$$

But $w = \lambda u$, so

$$\lambda = \frac{\sum_{M,N} v_i a_{ij} w_j}{\sum_{M,N} v_i a_{ij} u_j} = \frac{\sum_N c_j}{\sum_M r_i} = \omega(M, N). \tag{3.3.5}$$

Suppose now that I, J are nonempty proper subsets of M, N, respectively, with $A[I'|J] = 0$. Again, by (3.3.1),

$$\sum_{I,N} v_i a_{ij} u_j = \sum_I r_i, \tag{3.3.6}$$

while, by (3.3.3),

$$\sum_{I,J} v_i a_{ij} w_j = \sum_{M,J} v_i a_{ij} w_j = \sum_J c_j,$$

since $A[I'|J] = 0$. Hence

$$\sum_{I,N} v_i a_{ij} w_j \geqslant \sum_J c_j \tag{3.3.7}$$

with equality if and only if $A[I|J'] = 0$. From (3.3.6) and (3.3.7) we obtain by $w = \lambda u = \omega(M, N)u$ that

$$\omega(I, J) < \frac{\sum_{I,N} v_i a_{ij} w_j}{\sum_{I,N} v_i a_{ij} u_j} = \omega(M, N),$$

provided that $A[I|J'] > 0$, and

$$\omega(I, J) = \frac{\sum_{I,N} v_i a_{ij} w_j}{\sum_{I,N} v_i a_{ij} u_j} = \omega(M, N),$$

provided that $A[I|J'] = 0$. Hence (A, r, c) is consistent and the lemma is proved.

3.4. LEMMA. *Let (A, r, c) be an* mrc *and suppose that for $T = T(A; r, c)$ and $u > 0$, we have $Tu = \lambda u$. Suppose J is the (necessarily)* *nonempty subset of N such that* $u_J \gg 0$ *while* $u_{J'} = 0$, *and let* I' *be the subset of* M *defined by*

$$I' = \{i \in M : A[i|J] = 0\}. \tag{3.4.1}$$

Then

$$\text{either (a)} \quad I \times J = M \times N \quad \text{or} \qquad \text{(b)} \quad \phi \subset J \subset N, \quad \phi \subset I \subset M \tag{3.4.2}$$

and $A[I'|J]$ *is a maximal zero submatrix of* A,

$$T^0 u_J = \lambda u_J, \qquad \text{where} \qquad T^0 = T(A[I|J]; r_I, c_J) \tag{3.4.3}$$

$$(A[I|J], r_I, c_J) \quad \text{is consistent,} \tag{3.4.4}$$

and

$$\lambda = \omega(I, J). \tag{3.4.5}$$

Proof. (a) Suppose first that $J = N$. Since A has no zero row, $I = M$ and this proves (3.4.2). In this case (3.4.3) is trivial and (3.4.4) and (3.4.5) reduce to Lemma 3.3.

(b) Now suppose that $\phi \subset J \subset N$. Observe that $\phi \subseteq I' \subset M$, since A has no zero column. Thus $I \neq \phi$. Put $S^0 = (S(A[I|J]; r_I)$, (cf. Definition 2.2). By direct computation,

$$(Su)_I = S^0 u_I \tag{3.4.6}$$

since $u_{J'} = 0$, and since each row of $A[I|J]$ is nonzero $(Su)_I$ is a finite (strictly positive) vector.

If $I' \neq \phi$, then

$$(Su)_{I'} = \infty, \tag{3.4.7}$$

a vector each of whose elements is ∞.

Next, since $A[I'|J] = 0$ if $I' \neq \phi$ and, by (3.4.6),

$$(Tu)_J = T^0 u_J, \tag{3.4.8}$$

and, by (3.4.7), we obtain for $j \in J'$ that

$$(Tu)_j = 0 \qquad \text{if and only if} \quad I' \neq \phi \quad \text{and} \quad A[I'|j] > 0. \tag{3.4.9}$$

But $(Tu)_j = \lambda u_j = 0$ if $j \in J'$, whence $\phi \subset I \subset M$ and $A[I'|j] > 0$ for $j \in J'$. Since we have supposed $J \subset N$, this proves (3.4.2). Returning to (3.4.8), we obtain (3.4.3), where u_J is strictly positive.

We now immediately deduce (3.4.4) and (3.4.5) from Lemma 3.3.

3.5. THEOREM. *Let (A, r, c) be an* mrc. *Then $T = T(A; r, c)$ has a strictly positive eigenvector if and only if (A, r, c) is consistent. In this case, the corresponding eigenvalue is $\omega(M, N)$, and $\omega(M, N)$ is the largest eigenvalue of T.*

Proof. If $u \gg 0$ and $Tu - \lambda u$, then by Lemma 3.3 (A, r, c) is consistent and $\lambda = \omega(M, N)$. By Lemma 3.4, every other eigenvalue of T is of form $\omega(I, J)$, where $\phi \subset I \subset M$ and $\phi \subset J \subset N$. By consistency $\omega(I, J) \leqslant \omega(M, N)$, whence $\omega(M, N)$ is the greatest eigenvalue. To prove the converse, suppose that (A, r, c) is consistent. We introduce two auxiliary operators: the operator $T_\varepsilon = T(A_\varepsilon, A; r, c)$ of (2.3.2), and (as in Section 2.4)

$$T_\varepsilon^* = T(A_\varepsilon; r, c) \tag{3.5.1}$$

where A_ε is again given by (2.3.1).

Observe that

$$\rho_\varepsilon^* = \sup\{\lambda\colon\ x > 0, \quad T_\varepsilon^* x \geqslant \lambda x\}$$

is an eigenvalue of T_ε^* by Lemma 2.3 and, by Lemma 3.4, $\rho_\varepsilon^* = \omega(M, N)$ since A_ε has no zero submatrix. But by direct computation $T_\varepsilon^* \leqslant T_\varepsilon$ (i.e., $T_\varepsilon^* \leqslant T_\varepsilon x$ for all $x \geqslant 0$), whence

$$\omega(M, N) = \rho_\varepsilon^* \leqslant \rho_\varepsilon, \tag{3.5.2}$$

where ρ_ε is defined by (2.3.3). By Lemma 2.3, $\rho_\varepsilon \downarrow \rho$, as $\varepsilon \to 0$, and ρ is an eigenvalue of T. Hence $\rho \geqslant \omega(M, N)$. But, by Lemma 3.4, $\rho = \omega(I, J)$ where either $A[I'|J]$ is a zero submatrix and $I \subset M$, $J \subset N$, or $I = M$, $J = N$. By consistency, $\omega(I, J) \leqslant \omega(M, N)$, whence $\rho = \omega(M, N)$. Next suppose that A is indecomposable. (The matrix A is *indecomposable* if and only if $A[I'|J] = 0$ implies that $A[I|J'] \neq 0$). Then under the stated conditions $\omega(I, J) = \omega(M, N)$ only if $I \times J = M \times N$. Hence, by Lemma 3.4, the corresponding eigenvector u is strictly positive. If A is decomposable, then $A = A_1 \oplus A_2 \oplus \cdots \oplus A_\sigma$, where $A_\alpha = A[I_\alpha|J_\alpha]$, $\alpha = 1, \ldots, \sigma$, each A_α is indecomposable, and the I_α,

J_α form partitions of M, N, respectively. By consistency $\omega(I_\alpha, J_\alpha) = \omega(M, N)$ and, if $T_\alpha = T(A[I_\alpha, J_\alpha], r_{I_\alpha}, c_{J_\alpha})$, then we have already proved that there is a u_α which must be strictly positive, such that $T_\alpha u_\alpha = \omega(I_\alpha, J_\alpha)u_\alpha = \omega(M, N)u_\alpha$. If $u = u_1 \oplus \cdots \oplus u_\sigma$, then $u \gg 0$, and, since A is a direct sum,

$$Tu = Tu_1 \oplus \cdots \oplus T_\sigma u_\sigma = \omega(M, N)(v_1 \oplus \cdots \oplus u_\sigma) = \omega(M, N)u.$$

The theorem is proved.

3.6. Main Theorem. *Let (A, r, c) be an* mrc. *The spectrum of $T = T(A; r, c)$ consists of all λ for which there exist nonempty subsets I, J of M, N, respectively, such that*

either (a) $I \times J = M \times N$, *or*

(b) $\phi \subset I \subset M$, $\phi \subset J \subset N$ *and* $A[I'|J]$ *is a maximal zero submatrix of* A, (3.4.2)

$$(A[I|J], r_I, c_J) \quad \textit{is consistent} \tag{3.4.4}$$

and

$$\lambda = \omega(I, J). \tag{3.4.5}$$

If all these conditions are satisfied, then there is an associated eigenvector with $u_J \gg 0$ and (for $J \subset N$) $u_{J'} = 0$.

Proof. By Lemma 3.4, we know that each λ in the spectrum of T satisfies (3.4.2), (3.4.4), and (3.4.5). Conversely, let (3.4.2), (3.4.4), and (3.4.5) hold. If (3.4.2(a)) holds, then our theorem reduces to Theorem 3.5. So suppose (3.4.2(b)) is satisfied. Since $A[I'|J]$ is a maximal zero submatrix of A, and A has no zero columns, it follows that $A[I|J]$ has no zero row or column; hence, by Theorem 3.5, $T^0 = T(A[I|J]; r_I, c_J)$ has a strictly positive eigenvector u_J with associated eigenvalue $\lambda = \omega(I, J)$. Let $u_{J'} = 0$, and set $u = u_J \oplus u_{J'}$. If $S^0 = S(A[I|J]; r_I)$, then

$$(Su)_I = S^0 u_I \tag{3.6.1}$$

and

$$(Su)_{I'} = \infty \tag{3.6.2}$$

whence

$$(Tu)_J = T^0 u_J = \omega(I, J) u_J,$$

where T^0 is defined by (3.4.3). Further, $A[I'|J']$ has no zero column since $A[I'|J]$ is a maximal zero submatrix of A, whence

$$(Tu)_{J'} = 0.$$

Hence $Tu = \omega(I, J)u$, and the theorem is proved.

3.7. Definition and Remark. *The operator T of $\mathscr{P}$ into itself is called strongly monotonic on the open cone $\mathscr{P}^0 = \{x: x_i > 0\}$ if $0 \ll x < x'$ implies $Tx < Tx'$ and, for some integer m, $T^m x \ll T^m x'$.*

Observe that, for $x > 0$, $T^m x \gg 0$ for strong monotonic T. Clearly T is strongly monotonic on $\mathscr{P}^0$ if $x \ll x'$ implies $Tx \ll Tx'$ and, for $\phi \subset J \subset N$, also $x_J \ll x_J'$, $x_{J'} = x_{J'}'$ imply $(Tx)_J \ll (Tx')_J, (Tx)_{J'} < (Tx')_{J'}$.

3.8. Lemma. *Let (A, r, c) be an* mrc. *If A is indecomposable, then $T = T(A; r, c)$ is strongly monotonic on $\mathscr{P}^0$.*

Proof. Suppose $0 \ll x \ll x'$. Then $Sx \ll Sx'$, whence $Tx \ll Tx'$. Now suppose that $\phi \subset J \subset N$ and $0 \ll x_J \ll x_J'$ but $0 \ll x_J = x_J'$. Let

$$I' = \{i \in M: A[i|J] = 0\}. \tag{3.8.1}$$

Possibly $I' = \phi$, but, since A contains no zero column, we have $I' \subset M$. We now have $(Sx)_I \gg (Sx')_I$ and, if $I' \neq \phi$, also $(Sx)_{I'} = (Sx')_{I'}$. Since $A[I|J]$ can contain no zero column, it now follows that $(Tx)_J \ll (Tx')_J$. If $I' = \phi$, then $A[I|J'] = A[M|J'] = 0$ since A has no zero column. If $I' \neq \phi$, then $A[I|J'] \neq 0$ since A is indecomposable. Hence $A[I|J'] \neq 0$, and we may deduce that $(Tx)_{J'} < (Tx')_{J'}$. It follows by Definition 3.7 that T is strongly monotonic on $\mathscr{P}^0$.

3.9. Theorem. *Let (A, r, c) be a consistent* mrc. *If A is indecomposable, then $T = T(A; r, c)$ has a unique eigenvector u (except for scalar multiples) associated with $\omega(M, N)$ and u is strictly positive.*

Proof. The existence of a strictly positive eigenvector belonging to $\omega(M, N)$ is assured by Theorem 3.5. Since A is indecomposable, $\omega(I, J) =$

$\omega(M, N)$ only if $I \times J = M \times N$, whence by Theorem 3.6 every eigenvector belonging to $\omega(M, N)$ is strictly positive. Further, by Lemma 3.8, T is strongly monotonic, so uniqueness follows by (4.4) of [1].

3.10. COROLLARY. *If* (A, r, c) *is a consistent* mrc, *then the eigenvectors of* T *belonging to* $\omega(M, N)$ *(together with* 0*) form a cone.*

Proof. Suppose $A = A \oplus \cdots \oplus A_\sigma$, where $A_\alpha = A[I_\alpha|J_\alpha]$ is indecomposable and the I_α, J_α form a partition of M, N, respectively. It is easy to see that $(A[I_\alpha|J_\alpha], r_{I_\alpha}, c_{J_\alpha})$ is consistent, for suppose that $I \times J \subset I_\alpha \times J_\alpha$ and $A[I_\alpha \setminus I, J] = 0$. Then $A[I_\alpha'|J] = 0$, but $A[I_\alpha|J \setminus J_\alpha] \neq 0$ and so $A[I_\alpha|J_\alpha'] \neq 0$. It follows that

$$\omega(I_\alpha|J_\alpha) < \omega(M, N) = \omega(I, J).$$

Thus by Theorem 3.9 $T_\alpha = T(A[I_\alpha|J_\alpha], r_{I_\alpha}, c_{J_\alpha})$ has a unique strictly positive eigenvector u_α. Since A is a direct sum, it follows by direct computation that any vector $\gamma_1 u_1 \oplus \cdots \oplus \gamma_\sigma u_\sigma$ with $\gamma_i \geqslant 0$ is an eigenvector of T belonging to $\omega(M, N)$ or 0.

Conversely, assume that $Tu = \omega(M, N)u$. Direct computation again shows that $T_\alpha u_{J_\alpha} = (Tu)_{J_\alpha} = \omega(M, N)u_{J_\alpha}$ whence, by Theorem 3.9, $u_{J_\alpha} = \gamma_\alpha u_\alpha$, for some $\gamma_\alpha \geqslant 0$. Hence $u = \gamma_1 u_1 \oplus \cdots \oplus \gamma_\sigma u_\sigma$, and the corollary is proved.

3.11. *Example.* Corollary 3.10 fails for eigenvalues $\lambda < \omega(M, N)$. For example, let

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad r = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \qquad c = [1, 1].$$

If $u = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, then $Tu = \frac{1}{2}u$, and if $v = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, then $Tv = \frac{1}{2}v$. But $u + v = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $T(u + v) = \frac{2}{3}v$.

3.12. *Remark.* The restriction that A has no zero row or column is somewhat technical. For, let (A, r, c) be an mrc, let A^ρ be the $(m + 1) \times n$ matrix obtained by adjoining a zero row to A as row $(m + 1)$, and let $r_M{}^\rho = r$, $r^\rho_{M+1} > 0$. Then $T(A^\rho; r^\rho, c)x = T(A; r, c)x$ for all $x \in \mathscr{P}$. Next, let A^γ be the $(n+1)\times m$ matrix obtained from A by adjoining a zero column to A as column $(n+1)$ and $c_N^\gamma = c$, $c_{n+1}^\gamma > 0$, and let $x_{n+1}^\gamma \geqslant 0$. If $z^\gamma = T(A^\gamma; r, c^\gamma)x^\gamma$ where $x \in \mathscr{P}$, then $z_N^\gamma = T(A; r, c)x$ and $z_{n+1}^\gamma = \infty$.

4. THE YAX PROBLEM

We shall briefly point out the applications of Theorem 3.6 to some problems mentioned in Section 1.

4.1. THEOREM. *Let (A, r, c) satisfy Definition 2.2. Then there exists a positive λ and diagonal matrices X and Y with positive diagonal elements such that YAX has row sum vector r and column sum vector c if and only if (A, r, c) is consistent. In this case*

$$\lambda = \sum_N c_j / \sum_M r_i. \tag{4.1.1}$$

Proof. If X and Y satisfy the conditions of Theorem 4.1, then $Tx = \lambda x$ for some $x = (x_1, \ldots, x_n) \gg 0$ and λ given by (4.1.1); and conversely with $X = \operatorname{diag}(x_1, \ldots, x_2)$ and $Y = ((Sx)_1, \ldots, (Sx)_m)$. The result follows by Theorem 3.5.

The positive $m \times n$ matrices A and B have the same *pattern* if $a_{ij} = 0$ if and only if $b_{ij} = 0$.

4.2. COROLLARY. (Brualdi [1], Theorem (2.1)). *Let A be an indecomposable positive $m \times n$ matrix and let r and c be strictly positive column and row vectors, respectively, such that $\sum_m r_i = \sum_N c_j$. There exists a positive matrix B with the same pattern as A, row sum vector r and column sum vector c if and only if $A[I'|J] = 0$ implies that $\sum_J c_j > \sum_I r_i$.*

Proof. Since A is indecomposable the condition on A is precisely that (A, r, c) be consistent. Also (B, r, c) is consistent if and only if (A, r, c) is consistent.

4.3. COROLLARY (Menon [5], Theorem 2). *Let A be a given positive $m \times n$ matrix; r, c be strictly positive column and row vectors. Suppose there exists at least one positive matrix B with the same pattern as A and having row and column sum vectors r and c, respectively. Then there exist diagonal matrices X and Y such that YAX has row sum vector r and column sum vector c.*

Proof. The assumptions imply that (B, r, c) is consistent and $\sum_M r_i = \sum_N c_j$. Hence (A, r, c) is also consistent and the result follows from Theorem 4.1.

Of course, the theorems of Brualdi and Menon (our Corollaries 4.2 and 4.3) together are essentially equivalent to Theorem 4.1.

REFERENCES

1 R. A. Brualdi, Convex sets of non-negative matrices. *Canad. J. Math.* **20**(1968), 144–157.

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

3 E. W. Hobson, The theory of functions of a real variable, 2nd ed., Cambridge Univ. Press, 1926.

4 M. V. Menon, Reduction of a matrix with positive elements to a doubly stochastic matrix. *Proc. Ann. Math. Soc.* **18**(1967), 244–247.

5 M. V. Menon, Matrix links, an extremisation problem and the reduction of a non-negative matrix to one with prescribed row and column sums. *Canad. J. Math.* **20**(1968), 225–232.

6 M. V. Menon, Some spectral properties of an operator associated with a pair of non-negative matrices. *Trans. Amer. Math. Soc.* **132**(1968), 369–376.

7 H. Perfect and L. Mirsky, The distribution of positive elements in double stochastic matrices. *J. London Math. Soc.* **40**(1965), 688–698.

8 H. Schneider, The elementary divisors associated with 0 of a singular M-matrix. *Proc. Edinburgh Math. Soc.* (2)**10**(1956), 108–122.

9 R. Sinkhorn, A relationship between arbitrary positive matrices and doubly stochastic matrices. *Ann. Math. Statist.* **35**(1964), 876–879.

10 R. Sinkhorn, Diagonal equivalence to matrices with prescribed row and column sums. *Ann. Math. Monthly* **74**(1967), 402–405.

11 R. Sinkhorn and P. Knopp, Concerning non-negative matrices and doubly stochastic matrices. *Pacific J. Math.* **21**(1967), 343–348.

Received July 8, 1968