

Pacific Journal of Mathematics

**A STUDY OF CONVEX SETS OF STOCHASTIC MATRICES
INDUCED BY PROBABILITY VECTORS**

DARALD JOE HARTFIEL

A STUDY OF CONVEX SETS OF STOCHASTIC MATRICES INDUCED BY PROBABILITY VECTORS

D. J. HARTFIEL

This paper contains a study of convex sets of stochastic matrices induced by probability vectors. The vertices and dimension of each such convex set is found. Some topological properties of these sets are also given. Finally, the relationship between these sets and Markov chain theory is considered.

The primary motivation for this work is derived from studies generalizing the classical results concerning final transition probabilities in the theory of Markov chains. References dealing with such generalizations may be found in [3]. As an example of one such result, we provide the following.

THEOREM A. [Theorem 1,3] *Suppose $A_1, A_2, \dots, A_k, \dots$ is a sequence of stochastic matrices so that*

(1) $Y_k A_k = Y_k$ for probability vectors Y_1, Y_2, \dots ,

(2) $\lim_{k \rightarrow \infty} Y_k = Y_0 > 0$ and

(3) *given any $\varepsilon > 0$, there is an integer $N > 0$ so that $\rho(A_{k+1} A_{k+2} \cdots A_{k+N}) < \varepsilon$ for all k sufficiently large, where $\rho(A) = \max_{i_1, i_2, j} |a_{i_1 j} - a_{i_2 j}|$.*

Then $\lim_{k \rightarrow \infty} A_1 A_2 \cdots A_k = Y_0 Y_0$.

This result generalizes the classical Markov chain problem concerning $\lim_{k \rightarrow \infty} A^k$ where A is primitive and stochastic [2, p. 94], the generality being that one need not have the same transition matrix from step to step but may choose matrices in

$S[Y] = \{A \mid A \text{ is stochastic, } YA = Y \text{ with } Y \text{ a probability vector}\}$ which meet the criteria specified in the above theorem.

This paper then concerns a study of $S[Y]$. The objectives of the work are as follows.

(1) We hope to indicate how much freedom one has in selecting the sequence A_1, A_2, \dots if final transition probabilities are desired.

(2) In terms of $S[Y]$, we hope to illuminate the truth of Theorem A.

(3) Although no explicit problems are stated, we also hope to provide some feeling as to what future generalizations can be expected.

Finally we state that all matrices herein derive their entries from the real number field.

Results. It is easily verified that $S[Y]$ is a compact convex set with the property that if $A \in S[Y]$ and $B \in S[Y]$ then $AB \in S[Y]$. Of course if $Y = (1/n, 1/n, \dots, 1/n)$ then $S[Y]$ denotes the set of doubly stochastic matrices which, along with the stochastic matrices, have been studied extensively.

1. The vertices of $S[Y]$. In this section we give a procedure for finding the vertices of $S[Y]$. The inductive procedure of Jurkat and Ryser [4] for finding the vertices of $U(R, S) = \{m \times n \text{ matrices } A \geq 0 \text{ with } i\text{th row sum } r_i \text{ and } j\text{th column sum } c_j \text{ where } R = (r_1, r_2, \dots, r_m) \text{ and } S = (c_1, c_2, \dots, c_n)\}$ is utilized for finding vertices in $S[Y]$, for $Y > 0$. This is done by establishing an isomorphism between $S[Y]$ and $U(Y, Y)$.

LEMMA 1.1. *For $Y > 0$, $S[Y]$ is isomorphic to $U(Y, Y)$.*

Proof. Let $D = \text{diag. } (y_1, y_2, \dots, y_n)$. Then, if $A \in S[Y]$, $DA \in U(Y, Y)$. Further, if $A \in U(Y, Y)$, $D^{-1}A \in S[Y]$. Finally, it is easily seen that this one-to-one correspondence is in fact an isomorphism between $S[Y]$ and $U(Y, Y)$.

Thus to find the vertices of $S[Y]$ we can find the vertices of $U(Y, Y)$ by the Jurkat-Ryser procedure and then map these vertices by D^{-1} back to the vertices of $S[Y]$. For the sake of completeness, we shall include a summary of the Jurkat-Ryser procedure for finding the vertices in $U(R, S)$ for $R > 0$, $S > 0$, and

$$r_1 + \dots + r_m = c_1 + \dots + c_n.$$

To construct a vertex $A \in U(R, S)$, pick a position (i, j) in an $m \times n$ array. Compute $a_{ij} = \min\{r_i, c_j\}$. If $a_{ij} = r_i$, then complete the i th row with 0's and delete the i th row obtaining a smaller size matrix which must be a vertex of $U[(r_1, \dots, r_{i-1}, r_{i+1}, \dots, r_m), (c_1, \dots, c_{j-1}, c_j - r_i, c_{j+1}, \dots, c_n)]$. If $a_{ij} = c_j$, then complete the j th column with 0's and delete the j th column obtaining a smaller size matrix which must be a vertex of $U[(r_1, \dots, r_{i-1}, r_i - c_j, r_{i+1}, \dots, r_m), (c_1, \dots, c_{j-1}, c_{j+1}, \dots, c_n)]$.

This procedure is then reproduced on the smaller sized array until a vertex is found. Further, all vertices may be found by applying this inductive procedure.

For example, by applying the procedure to find the vertices of $S[(1/2, 1/3, 1/6)]$, we obtain

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ for } i = 3, j = 3.$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} \frac{2}{3} & 0 & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 \end{pmatrix} \text{ for } i = 3, j = 2.$$

$$\begin{pmatrix} \frac{2}{3} & 0 & \frac{1}{3} \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 1 & 0 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \end{pmatrix} \text{ for } i = 3, j = 1.$$

This then specifies the entire list of vertices of $S[(1/2, 1/3, 1/6)]$.

To extend our work to finding the vertices of $S[Y]$ where $Y \geq 0$ we proceed as follows. As $YA = Y$ if and only if $YPP^tAP = YP$ i.e., $P^tS[Y]P = S[YP]$, for any permutation matrix P we may assume without loss of generality that $Y = (y_1, y_2, \dots, y_r, 0, \dots, 0)$ where $y_k > 0$ for $k = 1, 2, \dots, r$. Now $YA = Y$ implies that $A = \begin{pmatrix} A_1 & 0 \\ A_{21} & A_2 \end{pmatrix}$ where A_1 is of order r . Hence $(y_1, \dots, y_r)A_1 = (y_1, \dots, y_r)$ and so $A_1 \in S[(y_1, \dots, y_r)]$. Further, if $A_1 \in S[(y_1, \dots, y_r)]$ then $A = \begin{pmatrix} A_1 & 0 \\ A_{21} & A_2 \end{pmatrix} \in S[Y]$ if and only if $(A_{21} \ A_2) \geq 0$ has row sums equal to one. Thus the vertices of $S[Y]$ may be found as follows.

$A \in S[Y]$ is a vertex of $S[Y]$ if and only if A_1 is a vertex of $S[(y_1, \dots, y_r)]$ and $(A_{21} \ A_2)$ has precisely one 1 in each row. For example the vertices of $S[(1, 0, 0)]$ are as follows.

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

2. Moving in $S[Y]$. This section considers the kind of changes that can be made among the entries of a matrix $A \in S[Y]$ to obtain another matrix $B \in S[Y]$. In light of Theorem A, our curiosity is over the various choices for each A_i in constructing the sequence A_1, A_2, \dots . The first result related to this question requires the following definition. An $n \times n$ matrix N is called a loop matrix if N has a collection of nonzero entries say $n_{i_1j_1}, n_{i_1j_2}, n_{i_2j_2}, \dots, n_{i_{s-1}j_s}, n_{i_sj_s} = n_{i_1j_1}$, with

$$n_{ij} = \begin{cases} \frac{-\varepsilon}{y_i} & \text{if } (i, j) = (i_k, j_{k+1}) \\ \frac{\varepsilon}{y_i} & \text{if } (i, j) = (i_k, j_k) \\ 0 & \text{otherwise} \end{cases}$$

for $\varepsilon \neq 0$ and $Y = (y_1, y_2, \dots, y_n) > 0$ some probability vector. For example,

$$N = \begin{matrix} & \begin{matrix} r & s & t \end{matrix} \\ \begin{matrix} i \\ j \\ k \end{matrix} & \begin{pmatrix} \frac{\varepsilon}{y_i} & 0 & \frac{-\varepsilon}{y_i} \\ \frac{-\varepsilon}{y_j} & \frac{\varepsilon}{y_j} & 0 \\ 0 & \frac{-\varepsilon}{y_k} & \frac{\varepsilon}{y_k} \end{pmatrix} \end{matrix}.$$

THEOREM 2.1. *Suppose A and B are in $S[Y]$ with $Y > 0$. Then there is a sequence of loop matrices N_1, N_2, \dots, N_t so that $A = A_1$, $A_1 + N_1 = A_2$, \dots , $A_t + N_t = A_{t+1} = B$ where $A_k \in S[Y]$.*

Proof. The proof is a matter of translating [Theorem 3.1,1], by using Lemma 1.1, into the current result.

For the extension of this theorem to the case where $Y \geq 0$ we introduce the following definition. A matrix S is called a shift matrix if S has row sums zero and precisely two nonzero entries in some one of its row, say $a_{ij_1} = \varepsilon$ and $a_{ij_2} = -\varepsilon$. For example,

$$S = i \begin{pmatrix} & j_1 & & j_2 & & \\ 0 & 0 & 0 & 0 & 0 & \\ 0 & \varepsilon & 0 & -\varepsilon & 0 & \\ 0 & 0 & 0 & 0 & 0 & \end{pmatrix}.$$

THEOREM 2.2. *Suppose A and B are in $S[Y]$ with $Y \geq 0$. Then there is a sequence of loop matrices N_1, N_2, \dots, N_t and shift matrices S_1, S_2, \dots, S_l so that $A = A_1$, $A_1 + N_1 = A_2$, \dots , $A_t + N_t = A_{t+1}$, $A_{t+1} + S_1 = A_{t+2}$, \dots , $A_{t+l} + S_l = A_{t+l+1} = B$.*

Proof. Without loss of generality we assume $Y = (y_1, \dots, y_r, 0, \dots, 0)$ has precisely r nonzero components. Then

$$A = \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix} \quad B = \begin{pmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{pmatrix}$$

with A_{11} and B_{11} of order r . Now as $A_{11}, B_{11} \in S[(y_1, \dots, y_r)]$ we see from Theorem 2.1 that there are loop matrices N_1, N_2, \dots, N_t so that $A = A_1$, $A_1 + N_1 = A_2$, \dots , $A_t + N_t = A_{t+1}$ where each $A_k \in S[Y]$ with B and A_{t+1} agreeing in each entry in the first r rows. Now we may add a sequence of suitable shift matrices to A_{t+1} yielding A_{t+2} whose $(r+1, 1)$ entry is precisely that of B . Similarly without altering this entry we may again add a sequence of shift matrices to A_{t+2}

whose $(r+1, 2)$ entry is precisely that of B . Continuing in this manner it follows that there is a sequence of shift matrices say S_1, S_2, \dots, S_l so that $A_{t+1} + S_1 = A_{t+2}, \dots, A_{t+l} + S_l = A_{t+l+1} = B$ with each $A_{t+k} \in S[Y]$. This then is the result of the theorem.

3. The size of $S[Y]$. This section discusses the size of the set $S[Y]$. In particular, we find the size of the largest simplex contained in $S[Y]$. The work may be considered an extension of a result of Marcus and Minc which provides that $\dim S[1/n, 1/n, \dots, 1/n] = (n-1)^2$.

Let $\mathcal{N}(m, n) = \{m \times n \text{ matrices } A \text{ so that } A \text{ has its } i\text{th row sum and } j\text{th column sum being zero}\}$. $\mathcal{N}(m, n)$, of course, is a vector space.

LEMMA 3.1. $\dim \mathcal{N} = (m-1)(n-1)$.

Proof. Let e_i be the $(0, 1)$ -row vector with m coordinates having precisely one 1, in its i th position. Let e be a row vector with n coordinates having a one in each position. Let $A^{(i)}$ be the i th column of A and

$$X = \begin{pmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(n)} \end{pmatrix}. \quad \text{Set } M = \begin{pmatrix} e & 0 & \dots & 0 \\ 0 & e & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & e \\ e_1 & e_1 & \dots & e_1 \\ e_2 & e_2 & \dots & e_2 \\ \dots & \dots & \dots & \dots \\ e_m & e_m & \dots & e_m \end{pmatrix}$$

an $(n+m) \times mn$ matrix. Then $A \in \mathcal{N}(m, n)$ if and only if $MX = 0$. Thus, as $\text{rank } M = m+n-1$, $\dim \mathcal{N}(m, n) = mn - [m+n-1] = (m-1)(n-1)$.

LEMMA 3.2. If $R = (r_1, r_2, \dots, r_m) > 0$ and $S = (c_1, c_2, \dots, c_n) > 0$ with $\tau = r_1 + r_2 + \dots + r_m = c_1 + c_2 + \dots + c_n$, then $\dim U(R, S) = (m-1)(n-1)$.

Proof. First note that $A = \tau^{-1}(r_i c_j) > 0$ and $A \in U(R, S)$. Let $E_1, E_2, \dots, E_{(m-1)(n-1)}$ be a basis of $\mathcal{N}(m, n)$. Then $A, A + \varepsilon E_1, A + \varepsilon E_2, \dots, A + \varepsilon E_{(m-1)(n-1)}$ provide the vertices of an $(m-1)(n-1)$ dimensional simplex of $U(R, S)$ for ε sufficiently small. To see that this is a largest simplex in $U(R, S)$, suppose B_0, B_1, \dots, B_r are the

vertices of any simplex of $U(R, S)$. Then $\{B_1 - B_0, B_2 - B_0, \dots, B_r - B_0\}$ is linearly independent. But each $B_k - B_0 \in \mathcal{N}(m, n)$ and so $r \leq (m-1)(n-1)$.

These two lemmas provide the initial result.

THEOREM 3.1. *For $Y > 0$, $\dim S[Y] = (n-1)^2$.*

Proof. An application of Lemma 1.1 and Lemma 3.2.

For the $\dim S[Y]$ when $Y \geq 0$, we proceed as follows.

Let $\mathcal{N}(m, \cdot) = \{m \times n \text{ matrices } A \text{ with } i\text{th row sum zero}\}$. Let $U(R, \cdot) = \{m \times n \text{ matrices } A \geq 0 \text{ with } i\text{th row sum } r_i\}$.

LEMMA 3.3. $\dim \mathcal{N}(m, \cdot) = m(n-1)$.

Proof. As in Lemma 3.1.

LEMMA 3.4. *If $R = (r_1, r_2, \dots, r_m) > 0$ then $\dim U(R, \cdot) = m(n-1)$.*

Proof. As in Lemma 3.2.

The major result of this section may now be stated as follows.

THEOREM 3.2. $\dim S[Y] = (r-1)^2 + (n-r)(n-1)$ for Y having precisely r nonzero values.

Proof. Without loss of generality, we assume that $Y = (y_1, y_2, \dots, y_r, 0, \dots, 0)$. Recall $A = \begin{pmatrix} A_1 & 0 \\ A_{21} & A_2 \end{pmatrix} \in S[Y]$ if and only if $A_1 \in S[(y_1, y_2, \dots, y_r)]$ and $(A_{21} \ A_2) \geq 0$ has row sums one. Let $B_0, B_1, \dots, B_{(r-1)^2}$ be the vertices of simplex in $S[(y_1, y_2, \dots, y_r)]$. Applying Lemma 3.4, let $C_0, C_1, \dots, C_{m(n-1)}$ be $(n-r) \times n$ matrices which are the vertices of a simplex in $U(R, \cdot)$ where $R = (1, 1, \dots, 1)$. Let

$$D_0 = \begin{pmatrix} B_0 & 0 \\ C_0 \end{pmatrix}, \quad D_1 = \begin{pmatrix} B_0 & 0 \\ C_1 \end{pmatrix}, \quad D_{(n-r)(n-1)} = \begin{pmatrix} B_0 & 0 \\ C_{(n-r)(n-1)} \end{pmatrix}, \quad \dots$$

$$D_{(n-r)(n-1)+1} = \begin{pmatrix} B_1 & 0 \\ C_0 \end{pmatrix}, \quad \dots, \quad D_{(n-r)(n-1)+(r-1)^2} = \begin{pmatrix} B_{(r-1)^2} & 0 \\ C_0 \end{pmatrix}.$$

It is easily verified that these matrices form the vertices of an $(r-1)^2 + (n-r)(n-1)$ dimensional simplex of $S[Y]$. Finally, by an argument similar to that in Theorem 3.1, $S[Y]$ can have no larger simplex.

4. Some intersection properties of $S[Y]$. This section considers the question of how well $S[Y]$ can be used to determine particular stochastic matrices. In [3], criteria (3) of Theorem A is used to show

that the product $A_1 A_2 \dots A_k$ gets closer to the set of rank one stochastic matrices as k increases. Criteria (2) is added to show that the product also gets closer to $S[Y_0]$. These two bits of information thus provide the desired result, as there is only one rank one matrix in $S[Y_0]$. Our results are intended to expand on this area.

THEOREM 4.1. *$A \in S[Y_1] \cap S[Y_2] \cap \cdots \cap S[Y_h]$ for some linearly independent set $\{Y_1, Y_2, \dots, Y_h\}$ if and only if A is reducible into at least h isolated submatrices.*

Proof. Suppose $\{Y_1, Y_2, \dots, Y_h\}$ is linearly independent and $A \in S[Y_1] \cap S[Y_2] \cap \dots \cap S[Y_h]$. Without loss of generality we may assume that

$$A = \begin{pmatrix} A_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A_g & 0 & \dots & 0 \\ A_{g+1\ 1} & A_{g+1\ 2} & \dots & \dots & A_{g+1} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ A_{l_1} & A_{l_2} & \dots & \dots & \dots & \dots & A_l \end{pmatrix}$$

is in normal form [2, p. 75] with g isolated irreducible submatrices. Now $\dim\{X \mid XA = X\} = g$. As $\{Y_1, Y_2, \dots, Y_h\} \subset \{X \mid XA = X\}$ it follows that $h \leq g$.

The converse argument is elementary.

As an application note that $\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ and $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$ are in precisely one

$S[Y]$ while $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$ is in infinitely many $S[Y]$. However, it does

follow that if $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} \in S[Y_1] \cap S[Y_2] \cap \dots \cap S[Y_h]$ then $\dim \text{span} \{Y_1, Y_2, \dots, Y_h\} \leq 2$. Also as a consequence, if $R = \{\text{rank one stochastic matrices}\}$ we have the following.

COROLLARY 4.1.

$$R \cap S[Y] = \left\{ \bar{Y} = \begin{pmatrix} y_1 & y_2 & \cdots & y_n \\ y_1 & y_2 & \cdots & y_n \\ \vdots & \vdots & & \vdots \\ y_1 & y_2 & \cdots & y_n \end{pmatrix} \right\}$$

i.e., the only rank one matrix in $S[Y]$ is \bar{Y} .

From this corollary we see that rank one stochastic matrices are completely determined by the particular $S[Y]$ they are in. To generalize this result to idempotents of higher rank we have the following. One should recall that idempotents play an important role in the study of final transition probabilities.

THEOREM 4.2. *Suppose $\{Y_1, Y_2, \dots, Y_h\}$ is a linearly independent set of probability vectors and $Y_1 + Y_2 + \dots + Y_h > 0$. Then there is at most one idempotent of rank h in $S[Y_1] \cap S[Y_2] \cap \dots \cap S[Y_h]$.*

Proof. Suppose A and B are idempotents of rank h in $S[Y_1] \cap S[Y_2] \cap \dots \cap S[Y_h]$. Without loss of generality by Theorem 4.1 we may assume that

$$A = \begin{pmatrix} A_1 & 0 & 0 \\ 0 & A_2 & 0 \\ & \ddots & \vdots \\ 0 & 0 & \dots A_h \end{pmatrix}$$

where each A_k ($k = 1, 2, \dots, h$) is rank one. Thus there is a linearly independent set $\{Z_1, \dots, Z_h\}$ of probability vectors so that

$$Z_k \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots A_k \dots 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} = Z_k \quad \text{for } k = 1, 2, \dots, h$$

and $\text{span}\{Z_1, \dots, Z_h\} = \text{span}\{Y_1, \dots, Y_h\}$. Therefore, $Z_k B = Z_k$ and hence by partitioning B as is A , say

$$B = \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1h} \\ B_{21} & B_{22} & \dots & B_{2h} \\ \dots & \dots & \dots & \dots \\ B_{h1} & B_{h2} & \dots & B_{hh} \end{pmatrix},$$

we have that $B_{ij} = 0$ for $i \neq j$. Hence each B_k ($k = 1, 2, \dots, h$) is rank one and again as $Z_k B = Z_k$ it follows that $A = B$.

It should be noted here that Theorem 4.2 does not imply that there is an idempotent of rank h in $S[Y_1] \cap S[Y_2] \cap \dots \cap S[Y_h]$. In fact $S[Y_1] \cap S[Y_2] \cap \dots \cap S[Y_h]$ may contain only I . Further we should mention that the condition $Y_1 + Y_2 + \dots + Y_h > 0$ may not in general be relaxed, for if the normal form of A is

$$\left(\begin{array}{ccc|c} A_1 & \cdots & 0 & 0 \\ \vdots & & & \\ 0 & \cdots & A_g & 0 \\ \hline & B & & 0 \end{array} \right)$$

then A is idempotent if and only if each A_k ($k = 1, 2, \dots, g$) is idempotent and $\text{rank } A = \text{rank } A_1 + \cdots + \text{rank } A_g$. Thus if $g > 1$ and B appears then there are infinitely many choices for B and hence infinitely many idempotents. Concerning the count of the idempotents in $S[Y]$ we do have the following.

THEOREM 4.3. *If $Y = (y_1, y_2, \dots, y_n)$ with $n \leq 2$, then $S[Y]$ has only finitely many idempotents.*

THEOREM 4.4. *If $Y = (y_1, y_2, \dots, y_n)$ with $n \geq 2$ and $Y \succ 0$ then $S[Y]$ has infinitely many idempotents.*

Proof. Without loss of generality, assume $Y = (y_1, y_2, \dots, y_{n-1}, 0)$. Now $A = \begin{pmatrix} I & 0 \\ B & 0 \end{pmatrix}$ where $B \geq 0$ a $n-1$ dimensional row vector with $\sum_{i=1}^{n-1} b_i = 1$ yields infinitely many idempotents in $S[Y]$ corresponding to infinitely many choices of B .

THEOREM 4.5. *If $Y > 0$ then $S[Y]$ has only finitely many idempotents.*

Proof. A is an idempotent of $S[Y]$ if and only if there is a permutation matrix P so that

$$B = P^t A P = \begin{pmatrix} A_1 & 0 & 0 \\ 0 & A_2 & 0 \\ & \ddots & \vdots \\ 0 & 0 \cdots A_g \end{pmatrix}$$

with each A_k of order n_k , rank one for $k = 1, 2, \dots, g$ and $B \in S[YP]$. Further, if $YP = \hat{Y}$ and $\sigma_k = \hat{y}_{n_1+\cdots+n_{k-1}+1} + \cdots + \hat{y}_{n_1+\cdots+n_{k-1}+n_k}$ then $A_k = \binom{(k)}{a_{ij}}$ where

$$a_{ij} = \frac{\hat{y}_{n_1+\cdots+n_{k-1}+1+\cdots+j}}{\sigma_k}.$$

This then implies that there are only finitely many idempotents in $S[Y]$.

It is easily established that if $Y \succ 0$ then $A \in S[Y]$ implies that

A is reducible. Further, if $Y > 0$ then $A \in S[Y]$ implies that A is irreducible or completely reducible [2, p. 78]. The remainder of this section then contains results concerning the 0 pattern of matrices in more than one $S[Y]$.

THEOREM 4.6. *Suppose Y has precisely r nonzero entries. If P is any permutation matrix so that $YP = (0, \dots, 0, y_{t+1}, \dots, y_{t+r}, 0, \dots, 0)$ then $A \in S[Y]$ implies that $P^tAP = B$ is such that $b_{ij} = 0$ for $t+1 \leq i \leq t+r$ and $j < t+1$ or $j > t+r$, i.e., B has an isolated submatrix in rows $t+1$ to $t+r$.*

Proof. By direct calculation.

COROLLARY 4.2. *If $\{Y_1, \dots, Y_r\}$ is a linearly independent set of probability vectors so that $Y_i Y_j^t = 0$ for $i \neq j$ with $Y_1 + Y_2 + \dots + Y_r > 0$ then $A \in S[Y_1] \cap S[Y_2] \cap \dots \cap S[Y_r]$ implies that there is a permutation matrix P so that $B = P^tAP$ is completely reducible into r isolated submatrices.*

THEOREM 4.7. *If $A \in S[X] \cap S[Y]$ then either*

(1) $XY^t = 0$ or

(2) *there is a probability vector Z so that $z_i > 0$ if and only if $x_i > 0$ and $y_i > 0$ with $A \in S[Z]$, i.e., the isolated submatrices corresponding to X and Y intersect in an isolated submatrix corresponding to Z .*

Proof. As a consequence of [2, p. 96] there is a positive integer m and a permutation matrix P so that $A_0 = \lim_{k \rightarrow \infty} P^t A^{mk} P = \begin{pmatrix} J & 0 \\ F & 0 \end{pmatrix}$ where

$$J = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A_s \end{pmatrix}$$

is of order t , each $A_i > 0$ of order t_i and rank one. Now of course $A_0 \in S(XP) \cap S(YP)$. Note that $XY^t = 0$ if and only if $(XP)(YP)^t = 0$. Thus suppose $XY^t \neq 0$. Let $XP = \hat{X}$, $YP = \hat{Y}$. As $A_0 \in S[\hat{X}] \cap S[\hat{Y}]$ it follows that $\hat{x}_i = \hat{y}_i = 0$ for $i > t$. Thus $J \in S[(\hat{x}_1, \dots, \hat{x}_t)] \cap S[(\hat{y}_1, \dots, \hat{y}_t)]$. One now sees (2) by direct calculation.

5. Topological properties of $S[Y]$. In this section we consider how close and how far apart matrices in $S[X]$ can be from matrices in $S[Y]$ in terms of the components of X and Y . Our first result

extends Lemma 2 in [3] concerning the closeness of matrices in $S[X]$ to those in $S[Y]$.

THEOREM 5.1. *If $A \in S[X]$ there is a $B \in S[Y]$ so that*

$$\max_{ij} |a_{ij} - b_{ij}| \leq \frac{2n(n+1)}{\delta} \max |x_i - y_i| \text{ where } \delta = \min_{y_i > 0} y_i.$$

Proof. We may assume without loss of generality that $Y = (y_1, \dots, y_r, 0, \dots, 0)$ with $y_1 > 0, \dots, y_r > 0$. Partition $A = \begin{pmatrix} A_1 & A_{12} \\ A_{21} & A_2 \end{pmatrix}$ with A_1 of order r . Compute $YA = Z$. Set $\hat{Y} = (y_1, \dots, y_r)$ and $\hat{A} = (A_1 \ A_{12})$. Then $\hat{Y}\hat{A} = Z$. Again without loss of generality, assume $y_1 \leq z_1, y_2 \leq z_2, \dots, y_s \leq z_s, y_{s+1} \geq z_{s+1}, \dots, y_r \geq z_r$. Let $\varepsilon_k \geq 0$ so that $(1 - \varepsilon_k)z_k = y_k$ for $k = 1, 2, \dots, s$. Thus as $\sum_{k=1}^r y_k = \sum_{k=1}^n z_k$ it follows that $\sum_{k=s+1}^r y_i = \sum_{k=1}^s \varepsilon_k z_k + \sum_{k=s+1}^n z_k$. Pick $0 \leq \delta_i^j \leq 1$ so that

$$\begin{aligned} z_{s+1} + \delta_1^{s+1} z_1 + \dots + \delta_s^{s+1} z_s + \delta_{r+1}^{s+1} z_{r+1} + \dots + \delta_n^{s+1} z_n &= y_{s+1} \\ \vdots \\ z_r + \delta_1^r z_1 + \dots + \delta_s^r z_s + \delta_{r+1}^r z_{r+1} + \dots + \delta_n^r z_n &= y_r, \end{aligned}$$

with

$$\begin{aligned} \delta_1^{s+1} z_1 + \dots + \delta_s^r z_s &= \varepsilon_1 z_1 \\ \delta_s^{s+1} z_s + \dots + \delta_s^r z_s &= \varepsilon_s z_s \\ \delta_{r+1}^{s+1} z_{r+1} + \dots + \delta_{r+1}^r z_{r+1} &= z_{r+1} \\ \delta_n^{s+1} z_n + \dots + \delta_n^r z_n &= z_n. \end{aligned}$$

Now let $\hat{A}^{(k)}$ denote the k th column of \hat{A} for $k = 1, 2, \dots, n$ and set

$$\begin{aligned} \hat{B} &= (\hat{A}^{(1)} - \varepsilon_1 \hat{A}^{(1)}, \dots, \hat{A}^{(s)} - \varepsilon_s \hat{A}^{(s)}, \hat{A}^{(s+1)} + \delta_1^{s+1} \hat{A}^{(1)} \\ &\quad + \dots + \delta_s^{s+1} \hat{A}^{(s)} + \delta_{r+1}^{s+1} \hat{A}^{(r+1)} + \dots + \delta_n^{s+1} \hat{A}^{(n)}, \dots, \hat{A}^{(r)} \\ &\quad + \delta_1^r \hat{A}^{(1)} + \dots + \delta_s^r \hat{A}^{(s)} + \delta_{r+1}^r \hat{A}^{(r+1)} + \dots + \delta_n^r \hat{A}^{(n)}, 0, \\ &\quad \dots, 0). \end{aligned}$$

Now \hat{B} has row sums equal to one and $\hat{Y}\hat{B} = \hat{Y}$. Therefore,

$$B = \begin{pmatrix} \hat{B} \\ A_{21} & A_2 \end{pmatrix} \in S[Y].$$

Now note that $\delta_1^k + \dots + \delta_s^k \leq n \max_{1 \leq i \leq s} \varepsilon_i$ for $k = s+1, \dots, r$ and that as $\sum_{k=1}^r y_k a_{kt} = z_t$ we have that $a_{kt} \leq z_t / \delta = (|z_t - y_t|) / \delta$ for $t > r$ and $k \leq r$. Hence

$$\begin{aligned}
|a_{ij} - b_{ij}| &\leq \max \left\{ \varepsilon_1, \dots, \varepsilon_s, n \left(\max_{1 \leq i \leq s} \varepsilon_i + \frac{1}{\delta} \max_{r < t} |z_t - y_t| \right) \right\} \\
&\leq n \left(\max_{1 \leq i \leq s} \varepsilon_i + \frac{1}{\delta} \max_{r < t} |z_t - y_t| \right) \\
&\leq n \left(\frac{1}{\delta} \max_{1 \leq i \leq s} |z_i - y_i| + \frac{1}{\delta} \max_{r < t} |z_t - y_t| \right) \\
&\leq \frac{2n}{\delta} \max |z_i - y_i|.
\end{aligned}$$

Finally set $x_i - y_i = \gamma_i$ for $i = 1, 2, \dots, n$. Then

$$\begin{aligned}
|y_i - z_i| &= \left| y_i - \sum_k a_{ki} y_k \right| = \left| y_i - \sum_k a_{ki} (x_k - \gamma_k) \right| \\
&= \left| y_i - x_i + \sum_k a_{ki} \gamma_k \right| \leq |y_i - x_i| + \left| \sum_k a_{ki} \gamma_k \right| \\
&\leq |x_i - y_i| + n \max |x_i - y_i| \leq (n+1) \max |x_i - y_i|
\end{aligned}$$

and so

$$|\hat{a}_{ij} - \hat{b}_{ij}| \leq \frac{2n(n+1)}{\delta} \max |x_i - y_i|.$$

Thus

$$|a_{ij} - b_{ij}| \leq \frac{2n(n+1)}{\delta} \max |x_i - y_i|.$$

Our results concerning how far matrices in $S[X]$ can be from matrices in $S[Y]$ rest on the following theorem.

THEOREM 5.2. *Given any probability vector Y , there is an $A \in S[Y]$ so that $a_{ii} = 0$ for some i .*

Proof. If $y \not\geq 0$, pick P a permutation matrix so that $YP = Y_0 = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_r, 0, \dots, 0)$ with $\hat{y}_1 > 0, \hat{y}_2 > 0, \dots, \hat{y}_r > 0$. Then

$$\bar{Y}_0 = \begin{pmatrix} \hat{y}_1 & \hat{y}_2 & \dots & \hat{y}_r & 0 & \dots & 0 \\ \hat{y}_1 & \hat{y}_2 & \dots & \hat{y}_r & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \hat{y}_1 & \hat{y}_2 & \dots & \hat{y}_r & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix} \in S[YP]$$

with 0 in the (n, n) position and hence $P^t \bar{Y}_0 P \in S[Y]$ has the desired property. If $Y \geq 0$ then let $y_{i_0} = \min y_k$. Now consider the loop matrix

$$N = \begin{matrix} & i_0 & j \\ i_0 & \begin{pmatrix} -y_{i_0} & y_{i_0} \\ \frac{y_{i_0}^2}{y_j} & -\frac{y_{i_0}^2}{y_j} \end{pmatrix} & j \end{matrix}. \quad \text{Let } \bar{Y} = \begin{pmatrix} y_1 & y_2 & \cdots & y_n \\ y_1 & y_2 & \cdots & y_n \\ \dots\dots\dots & \dots\dots\dots & \dots\dots\dots & \dots\dots\dots \\ y_1 & y_2 & \cdots & y_n \end{pmatrix} \in S[Y].$$

Then $\bar{Y} + N \in S[Y]$ and has the desired property.

As corollaries, we can see how far the matrices in $S[Y]$ are spread and also how far one can expect matrices in $S[X]$ to be from matrices in $S[Y]$.

COROLLARY 5.1. *Given any probability vector Y there is an $A \in S[Y]$ and a $B \in S[Y]$ so that $\max_{ij} |a_{ij} - b_{ij}| = 1$.*

Proof. Take A with the property of Theorem 5.2 and $B = I$.

COROLLARY 5.2. *Given any probability vectors X and Y there is an $A \in S[Y]$ and a $B \in S[X]$ so that $\max_{ij} |a_{ij} - b_{ij}| = 1$.*

Proof. As in Corollary 5.1.

6. Conclusion. Concluding this paper, we cite §§ 1, 2, and 3 as providing some answer to motivating point (1) in our introduction. For motivating point (2) we cite § 4 and § 5 as being significant. Concerning motivating point (3) we label the general areas of

(1) mean limiting transition probabilities [2, p. 96] and

(2) studies on sequences of transition matrices with some specified zero pattern,

as possibly fruitful areas in which to further generalize the classical work of Markov chains. Loosely speaking, as I see it, one can consider the defining properties of the final transition matrix A_0 , i.e., does it lie in $S[Y]$ or several $S[Y]$, is it idempotent, etc., and let these properties determine possible sequences of stochastic matrices A_1, A_2, \dots so that $\lim_{k \rightarrow \infty} A_1 A_2 \cdots A_k = A_0$, etc. Sections 4 and 5 may then be useful in determining various types of such sequences.

ACKNOWLEDGMENT. The author acknowledges that the referee's remarks were very useful in simplifying the material contained in §§ 1, 2, and 3.

REFERENCES

1. R. A. Brualdi, *Convex sets of nonnegative matrices*, Canad. J. Math., **20** (1968), 144-157.
2. F. R. Gantmacher, *The Theory of Matrices*, Volume 2, Chelsea Publishing Company, New, York, 1964.

3. D. J. Hartfiel, *A result concerning strongly erogodic nonhomogeneous Markov chains*, to appear in the J. Linear Algebra and Appl.
4. W. B. Jurkat and H. J. Ryser, *Term ranks and permanents of nonnegative matrices*, J. Algebra, **5** (1967), 342-357.
5. M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Allyn and Bacon, Inc., Boston, 1964.
6. S. Schwarz, *On the structure of the semigroup of stochastic matrices*, Magyar Tud. Akad. Mat. Kutots Int. Koze., **9** (1965), 297-311.
7. R. Sinkhorn, *On the factor space of the complex doubly stochastic matrices*, Abstract 62T-243, Notices Amer. Math. Soc., **9** (1962), 334-335.

Received July 17, 1973 and in revised form March 1, 1974. This research was supported by the College of Science Organized Research Funds of Texas A & M University.

TEXAS A & M University

PACIFIC JOURNAL OF MATHEMATICS

EDITORS

RICHARD ARENS (Managing Editor)
University of California
Los Angeles, California 90024

J. DUGUNDJI
Department of Mathematics
University of Southern California
Los Angeles, California 90007

R. A. BEAUMONT
University of Washington
Seattle, Washington 98105

D. GILBARG AND J. MILGRAM
Stanford University
Stanford, California 94305

ASSOCIATE EDITORS

E. F. BECKENBACH

B. H. NEUMANN

F. WOLF

K. YOSHIDA

SUPPORTING INSTITUTIONS

UNIVERSITY OF BRITISH COLUMBIA
CALIFORNIA INSTITUTE OF TECHNOLOGY
UNIVERSITY OF CALIFORNIA
MONTANA STATE UNIVERSITY
UNIVERSITY OF NEVADA
NEW MEXICO STATE UNIVERSITY
OREGON STATE UNIVERSITY
UNIVERSITY OF OREGON
OSAKA UNIVERSITY

UNIVERSITY OF SOUTHERN CALIFORNIA
STANFORD UNIVERSITY
UNIVERSITY OF TOKYO
UNIVERSITY OF UTAH
WASHINGTON STATE UNIVERSITY
UNIVERSITY OF WASHINGTON
* * *
AMERICAN MATHEMATICAL SOCIETY
NAVAL WEAPONS CENTER

The Supporting Institutions listed above contribute to the cost of publication of this Journal, but they are not owners or publishers and have no responsibility for its content or policies.

Mathematical papers intended for publication in the *Pacific Journal of Mathematics* should be in typed form or offset-reproduced, (not dittoed), double spaced with large margins. Underline Greek letters in red, German in green, and script in blue. The first paragraph or two must be capable of being used separately as a synopsis of the entire paper. Items of the bibliography should not be cited there unless absolutely necessary, in which case they must be identified by author and Journal, rather than by item number. Manuscripts, in duplicate if possible, may be sent to any one of the four editors. Please classify according to the scheme of Math. Rev. Index to Vol. **39**. All other communications to the editors should be addressed to the managing editor, or Elaine Barth, University of California, Los Angeles, California, 90024.

100 reprints are provided free for each article, only if page charges have been substantially paid. Additional copies may be obtained at cost in multiples of 50.

The *Pacific of Journal Mathematics* is issued monthly as of January 1966. Regular subscription rate: \$72.00 a year (6 Vols., 12 issues). Special rate: \$36.00 a year to individual members of supporting institutions.

Subscriptions, orders for back numbers, and changes of address should be sent to Pacific Journal of Mathematics, 103 Highland Boulevard, Berkeley, California, 94708.

PUBLISHED BY PACIFIC JOURNAL OF MATHEMATICS, A NON-PROFIT CORPORATION

Printed at Kokusai Bunkin Insatsusha (International Academic Printing Co., Ltd.), 270, 3-chome Totsuka-cho, Shinjuku-ku, Tokyo 160, Japan.

Copyright © 1973 by Pacific Journal of Mathematics
Manufactured and first issued in Japan

Harm Bart, <i>Spectral properties of locally holomorphic vector-valued functions</i>	321
J. Adrian (John) Bondy and Robert Louis Hemminger, <i>Reconstructing infinite graphs</i>	331
Bryan Edmund Cain and Richard J. Tondra, <i>Biholomorphic approximation of planar domains</i>	341
Richard Carey and Joel David Pincus, <i>Eigenvalues of seminormal operators, examples</i>	347
Tyrone Duncan, <i>Absolute continuity for abstract Wiener spaces</i>	359
Joe Wayne Fisher and Louis Halle Rowen, <i>An embedding of semiprime P.I.-rings</i>	369
Andrew S. Geue, <i>Precompact and collectively semi-precompact sets of semi-precompact continuous linear operators</i>	377
Charles Lemuel Hagopian, <i>Locally homeomorphic λ connected plane continua</i>	403
Darald Joe Hartfiel, <i>A study of convex sets of stochastic matrices induced by probability vectors</i>	405
Yasunori Ishibashi, <i>Some remarks on high order derivations</i>	419
Donald Gordon James, <i>Orthogonal groups of dyadic unimodular quadratic forms. II</i>	425
Geoffrey Thomas Jones, <i>Projective pseudo-complemented semilattices</i>	443
Darrell Conley Kent, Kelly Denis McKennon, G. Richardson and M. Schroder, <i>Continuous convergence in $C(X)$</i>	457
J. J. Koliha, <i>Some convergence theorems in Banach algebras</i>	467
Tsang Hai Kuo, <i>Projections in the spaces of bounded linear operations</i>	475
George Berry Leeman, Jr., <i>A local estimate for typically real functions</i>	481
Andrew Guy Markoe, <i>A characterization of normal analytic spaces by the homological codimension of the structure sheaf</i>	485
Kunio Murasugi, <i>On the divisibility of knot groups</i>	491
John Phillips, <i>Perturbations of type I von Neumann algebras</i>	505
Billy E. Rhoades, <i>Commutants of some quasi-Hausdorff matrices</i>	513
David W. Roeder, <i>Category theory applied to Pontryagin duality</i>	519
Maxwell Alexander Rosenlicht, <i>The nonminimality of the differential closure</i>	529
Peter Michael Rosenthal, <i>On an inversion theorem for the general Mehler-Fock transform pair</i>	539
Alan Saleski, <i>Stopping times for Bernoulli automorphisms</i>	547
John Herman Scheuneman, <i>Fundamental groups of compact complete locally affine complex surfaces. II</i>	553
Vashishtha Narayan Singh, <i>Reproducing kernels and operators with a cyclic vector. I</i>	567
Peggy Strait, <i>On the maximum and minimum of partial sums of random variables</i>	585
J. L. Brenner, <i>Maximal ideals in the near ring of polynomials modulo 2</i>	595
Ernst Gabor Straus, <i>Remark on the preceding paper: "Ideals in near rings of polynomials over a field"</i>	601
Masamichi Takesaki, <i>Faithful states on a C^*-algebra</i>	605
R. Michael Tanner, <i>Some content maximizing properties of the regular simplex</i>	611
Andrew Bao-hwa Wang, <i>An analogue of the Paley-Wiener theorem for certain function spaces on $SL(2, \mathbb{C})$</i>	617
James Juei-Chin Yeh, <i>Inversion of conditional expectations</i>	631