ON THE LIMITING SET OF STOCHASTIC PRODUCTS $xA_1 \dots A_k$

D. J. HARTFIEL

ABSTRACT. In this paper we study the sequential limiting set L of the sequence $R_1, \dots, R_k = \{xA_1 \dots A_k | x \text{ is a stochastic vector, and } A_i \text{ a stochastic matrix with } P \leq A_i \leq Q, \text{ for nonnegative matrices } P \text{ and } Q\}, \dots$. We characterize L and provide some work on approximating L .

Let P and Q be nonnegative $n \times n$ matrices. For each positive integer k , let $C_k = \{A_1 \dots A_k | P \leq A_i \leq Q \text{ with } A_i \text{ an } n \times n \text{ stochastic matrix for } i = 1, \dots, k\}$, and

$$R_k = \{xA_1 \dots A_k | A_1 \dots A_k \in C_k \text{ with } x \text{ a stochastic vector in } R^n\}.$$

We will assume throughout the paper that P and Q are chosen so that $C_1 \neq \emptyset$. Hence, for all k , C_k and R_k are nonempty.

In this paper we study the sets $R_1, \dots, R_k, \dots$. We will show that each of these sets is a convex polytope. Further, we will find the sequential limiting set L for the sequence. Then, some work on approximating L is given. Finally, we discuss some application for this work.

RESULTS. We begin this study on the sets R_k by showing that C_1 is a convex polytope. For this, let e_i denote the $(0, 1)$ -vector in R^n with a one only in the i th position. Further, let $e = \sum_{i=1}^n e_i$.

LEMMA 1. *Let $a \leq b$ be nonnegative vectors with $b \leq e$. Then*

$$\Omega = \{\text{stochastic vectors } y \text{ so that } a \leq y \leq b\}$$

is a convex polytope with vertices those vectors $y \in \Omega$ so that there is at most one component y_i , called free, with $a_i < y_i < b_i$.

PROOF. It is clear that Ω is compact and convex so that it has vertices. Pick any $y \in \Omega$ with at least two free components, say y_i and y_j . Pick $\epsilon > 0$ so that $x = y + \epsilon e_i - \epsilon e_j$ and $z = y - \epsilon e_i + \epsilon e_j$ are in Ω . Then $\frac{1}{2}x + \frac{1}{2}z = y$ so that y is not a vertex.

Suppose now that $y \in \Omega$ has at most one free component. Set $y = \frac{1}{2}x + \frac{1}{2}z$ where x and z are in Ω . If $y_i = a_i$ for some i then $x_i = z_i = a_i$. Similarly if $y_i = b_i$ for some i then $x_i = z_i = b_i$. Hence, if y has no free coordinates, neither does x and z so $y = x = z$. If y has a free coordinate, say y_t , then since $y_t = 1 - \sum_{k \neq t} y_k$, $x_t = 1 - \sum_{k \neq t} x_k$, and $z_t = 1 - \sum_{k \neq t} z_k$ it follows that $y_t = x_t = z_t$. Hence, $y = x = z$. Thus, y is a vertex.

Received by the editors October 25, 1979 and, in revised form, February 20, 1980. Presented at the Auburn Matrix Theory Conference, March 19–22, 1980.

1980 *Mathematics Subject Classification*. Primary 15A51.

Finally, as we have already shown, if y is a vertex then y has at most one free component, say y_i . However, since y is stochastic, this component must be $1 - \sum_{k \neq i} y_k$ and so is completely determined by the nonfree components. Hence, Ω has finitely many vertices. $\square$

As a consequence, we have that C_1 is a convex polytope.

COROLLARY 1. *The set C_1 is a convex polytope, in the vector space of $n \times n$ matrices, with vertices those stochastic matrices in C_1 which have at most one free component in each row.*

An example is provided below.

EXAMPLE 1. Let $P = \begin{pmatrix} 1/3 & 1/3 \\ 1/3 & 1/3 \end{pmatrix}$ and $Q = \begin{pmatrix} 2/3 & 1/2 \\ 1/2 & 2/3 \end{pmatrix}$. Then the vertices of C_1 are $E_1 = \begin{pmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{pmatrix}$, $E_2 = \begin{pmatrix} 2/3 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$, $E_3 = \begin{pmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{pmatrix}$, and $E_4 = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$.

We now intend to show that R_k for each positive integer k is convex polytope. This result requires the following lemma.

LEMMA 2. *Let $d_1 = (d_1^{(1)}, \dots, d_n^{(1)})$ and $d_2 = (d_1^{(2)}, \dots, d_n^{(2)})$ be nonnegative vectors. Set $D_1 = \text{diag } d_1$ and $D_2 = \text{diag } d_2$. For A and B in C_1 , α and β positive numbers, let*

$$M = (\alpha D_1 + \beta D_2)^+ (\alpha D_1 A + \beta D_2 B) + F$$

where $(\alpha D_1 + \beta D_2)^+$ is the Moore-Penrose inverse [1, p. 7] of $\alpha D_1 + \beta D_2$ and $F = (f_{ij})$ where $f_{ij} = 0$ if $d_i^{(1)} + d_i^{(2)} > 0$ and $f_{ij} = a_{ij}$ otherwise. Then $M \in C_1$ and

$$(\alpha d_1 + \beta d_2)M = \alpha d_1 A + \beta d_2 B.$$

PROOF. Note that if $d_i^{(1)} + d_i^{(2)} > 0$ then

$$m_{ij} = \frac{\alpha d_i^{(1)} a_{ij} + \beta d_i^{(2)} b_{ij}}{\alpha d_i^{(1)} + \beta d_i^{(2)}}.$$

Thus

$$p_{ij} \leq m_{ij} \leq q_{ij}.$$

If $d_i^{(1)} + d_i^{(2)} = 0$ then $m_{ij} = f_{ij} = a_{ij}$ so that

$$p_{ij} \leq m_{ij} \leq q_{ij}.$$

Hence, $P \leq M \leq Q$. Further, from the above calculations it is easily seen that M is stochastic. Thus, $M \in C_1$.

For the second part,

$$\begin{aligned} (\alpha d_1 + \beta d_2)M &= (\alpha d_1 + \beta d_2)[(\alpha D_1 + \beta D_2)^+ (\alpha D_1 A + \beta D_2 B) + F] \\ &= e(\alpha D_1 + \beta D_2)(\alpha D_1 + \beta D_2)^+ (\alpha D_1 A + \beta D_2 B) \\ &= e(\alpha D_1 A + \beta D_2 B) = \alpha d_1 A + \beta d_2 B. \quad \square \end{aligned}$$

As a consequence of this lemma we have the following.

THEOREM 1. *For each k , R_k is a convex polytope.*

PROOF. We first show that R_k is convex. For this, pick $x A_1 \dots A_k$ and $y B_1 \dots B_k$ in R_k . Consider

$$\alpha(x A_1 \dots A_k) + \beta(y B_1 \dots B_k) = z$$

where $\alpha > 0, \beta > 0$ and $\alpha + \beta = 1$. We need to show $z \in R_k$. Set

$M_0 = (\alpha \text{diag } x + \beta \text{diag } y)^+ (\alpha(\text{diag } x)A_1 + \beta(\text{diag } y)B_1) + F_0$
 where $f_{ij}^{(0)} = 0$ if $x_i + y_i > 0$ and $f_{ij}^{(0)} = a_{ij}^{(1)}$ otherwise. Further, for $l = 1, 2, \dots$ set

$$M_l = (\alpha \text{diag}(xA_1 \dots A_l) + \beta \text{diag}(yB_1 \dots B_l))^+ \\ \cdot (\alpha(\text{diag}(xA_1 \dots A_l))A_{l+1} + \beta(\text{diag}(yB_1 \dots B_l))B_{l+1}) + F_l$$

where $f_{ij}^{(l)} = 0$ if $(xA_1 \dots A_k)_i + (yB_1 \dots B_k)_i > 0$ and $f_{ij}^{(l)} = a_{ij}^{(l)}$ otherwise. Then by the lemma, $M_l \in C_1$ for all l . Further

$$\begin{aligned} (\alpha x + \beta y)M_0 \dots M_{k-1} &= (\alpha xA_1 + \beta yB_1)M_1 \dots M_{k-1} \\ &\vdots \\ &= \alpha xA_1 \dots A_k + \beta yB_1 \dots B_k = z. \end{aligned}$$

Hence $z \in R_k$ and so R_k is convex.

Now let $E_1, \dots, E_r$ denote the vertices of C_1 . We will show that R_k is the convex polytope with vertices in $\{e_i E_{i_1} \dots E_{i_k} \text{ over all choices } i \text{ and } i_1, \dots, i_k\}$. Pick an $xA_1 \dots A_k \in R_k$. Write $x = \sum_{i=1}^n x_i e_i$ and $A_i = \sum_{k=1}^r \alpha_{ki} E_k$ where $\alpha_{ki} \geq 0$ and $\sum_k \alpha_{ki} = 1$. Then

$$\begin{aligned} xA_1 \dots A_k &= \left(\sum_{i=1}^n x_i e_i \right) \left(\sum_{k_1=1}^r \alpha_{k_1 1} E_{k_1} \right) \dots \left(\sum_{k_r=1}^r \alpha_{k_r r} E_{k_r} \right) \\ &= \sum_{i, k_1, \dots, k_r} x_i \alpha_{k_1 1} \dots \alpha_{k_r r} e_i E_{k_1} \dots E_{k_r}. \end{aligned}$$

Since $\sum_{i, k_1, \dots, k_r} x_i \alpha_{k_1 1} \dots \alpha_{k_r r} = 1$, the result follows. $\square$

Considering now the sequence $R_1, \dots, R_k, \dots$. We will say that a set L is the sequential limiting set of $R_1, \dots, R_k, \dots$ if $p \in L$ if and only if for each open set θ containing p there is a positive integer K so that for all $k \geq K$, $R_k \cap \theta \neq \emptyset$ [5]. The following result describes the sequential limiting set of $R_1, \dots, R_k, \dots$.

THEOREM 2. *Let L be the sequential limiting set of $R_1, \dots, R_k, \dots$. Then $L = \cap R_k$.*

PROOF. Pick any $p \in \cap R_k$. Then it follows from the definition of L that $p \in L$. Now pick any $p \in L$. Suppose there is some R_K so that $p \notin R_K$. Since R_K is closed, there is an open set θ containing p so that $\theta \cap R_K = \emptyset$. But now, since $R_{k+1} \subseteq R_k$ for all k , $\theta \cap R_k = \emptyset$ for all $k \geq K$. But now, from the definition of L , it follows that $p \notin L$. This is a contradiction. Hence $p \in R_k$ for all k and consequently $p \in \cap R_k$. $\square$

Since R_k is a convex polytope for each k , it follows that L is a compact convex set. An example in which this set is computed is below.

EXAMPLE 2. Let $P = (\frac{1}{3} \frac{1}{3})$ and $Q = (\frac{2}{3} \frac{2}{3})$. The vertices of C_1 are $E_1 = (\frac{1}{3} \frac{2}{3})$, $E_2 = (\frac{1}{3} \frac{2}{3})$, $E_3 = (\frac{2}{3} \frac{1}{3})$, and $E_4 = (\frac{2}{3} \frac{1}{3})$. Since E_1 and E_4 are rank one matrices, and each row of $E_{i_1} \dots E_{i_k}$ is a convex sum of $(1/3, 2/3)$ and $(2/3, 1/3)$ it follows that $R_1 = R_2 = \dots = L$. Thus

$$L = \text{conv}\{(1/3, 2/3), (2/3, 1/3)\}.$$

The above example is special since in general $R_1 = R_2 = \dots$ does not hold. However, it is easily argued that for any open set θ containing L , there is a positive integer K so that for $k \geq K$, $L \subseteq R_k \subseteq \theta$. Hence L can be approximated arbitrarily close by R_k , for sufficiently large k . If each of $E_1, \dots, E_r$ is fully indecomposable [4, p. 123], a measure of closeness of an R_k to L may be obtained as follows.

For this, let $N = \{1, \dots, n\}$ and let M be an $n \times n$ nonnegative matrix. Define

$$\mu(M) = \min_{R, C \subseteq N} \left(\max_{\substack{i \in R \\ j \in C}} m_{ij} \right) \quad \text{with } |R| + |C| = n,$$

$|S|$ denoting the number of elements in set S and neither R nor C being empty [3]. Similar to the proof of [3, Lemma 3], it can be shown that if $M_1 > 0$ and $M_2 > 0$ are $n \times n$ matrices and $\alpha > 0, \beta > 0$, then $\mu(\alpha M_1 + \beta M_2) > \alpha \mu(M_1) + \beta \mu(M_2)$. Further, as is shown in [3, Theorem 2], that if $S_1, \dots, S_{n-1}$ is any sequence of $n \times n$ stochastic matrices with $\mu(S_k) > \mu$ for $k = 1, \dots, n-1$ then for $S = S_1 \dots S_{n-1}$ and any nonnegative vector x ,

$$\max_{i,j} \left| \sum_{k=1}^n s_{ik} x_k - \sum_{k=1}^n s_{jk} x_k \right| < (1 - 2\mu^{n-1})(\bar{x} - \underline{x})$$

where $\bar{x} = \max_i x_i$ and $\underline{x} = \min_i x_i$. We use these results to obtain the following.

THEOREM 3. Suppose $\mu(E_i) > \mu > 0$ for $i = 1, \dots, r$. Pick any $\epsilon > 0$. Let l be a positive integer so that $(1 - 2\mu^{n-1})^l < \epsilon$, and set $m = l(n-1)$. Then if $xA_1 \dots A_m \in R_m$ there is a $y \in L$ so that $|y - xA_1 \dots A_m|_\infty < \epsilon$.

PROOF. First, let $A \in C_1$. Write $A = \sum_{i=1}^r \alpha_i E_i$ where $\alpha_i > 0$ for each i and $\sum_{i=1}^r \alpha_i = 1$. Then, by the result mentioned above, $\mu(A) > \sum_{i=1}^r \alpha_i \mu(E_i) > \mu$. Hence, if $A \in C_1$, $\mu(A) > \mu$.

Now, let $A_1 \dots A_m = B \in C_m$. Set $B_1 = A_1 \dots A_{n-1}$, $B_2 = A_n \dots A_{2(n-1)}$, $\dots$, $B_l = A_{(l-1)(n-1)+1} \dots A_{ln-1}$. Let z be any stochastic column vector. Set $w_1 = B_1 z$, $w_2 = B_2 z$, $\dots$, $w_l = B_l z$. Then by the result in [3, Theorem 2] applied sequentially to $w_1 = B_1 w_2$, $w_2 = B_2 w_3$, $\dots$, $w_l = B_l z$, we have

$$\begin{aligned} |\bar{w}_1 - \underline{w}_1| &< (1 - 2\mu^{n-1})(\bar{w}_2 - \underline{w}_2) \\ &< (1 - 2\mu^{n-1})^2(\bar{w}_3 - \underline{w}_3) \\ &\vdots \\ &< (1 - 2\mu^{n-1})^l(\bar{z} - \underline{z}). \end{aligned}$$

Hence if $z = e_k^t$ then since $Be_k^t = B_1 \dots B_l e_k^t$ is the k th column of B , it follows that $\max_{i,j} |b_{ik} - b_{jk}| < (1 - 2\mu^{n-1})^l < \epsilon$ for all k .

Since $\mu(B) > 0$, B is primitive and thus $\lim_{k \rightarrow \infty} B^k = \bar{B}$ exists and is rank 1 [2, p. 88]. Suppose

$$\bar{B} = \begin{bmatrix} b \\ \vdots \\ b \end{bmatrix}$$

where b is a stochastic vector. Since b is a convex sum of the rows of B , and of $B^2, \dots$ it follows that $b \in R_m$, and $b \in R_{2m}, \dots$. Thus, $b \in L$. Now both $xA_1 \dots A_m = xB = z$ and b are convex sums of the rows of B . Hence, since $\max_{i,j} |b_{ik} - b_{jk}| < (1 - 2\mu^{n-1})' < \epsilon$ it follows that

$$|b_k - z_k| < \epsilon$$

for each k . Thus,

$$|b - xA_1 \dots A_m|_\infty < \epsilon.$$

This is the result of the theorem. $\square$

This theorem, and Theorem 1 can now be used to provide an approximation R_m arbitrarily close to L . Such approximations can be done computationally provided that m is not too large.

EXAMPLE 3. Let $P = (\frac{1}{3} \frac{1}{3})$ and $Q = (\frac{2}{3} \frac{1}{2})$. The vertices of C_1 are given in Example 1. Thus, $\mu = 1/3$ and $1 - 2\mu = 1/3$. Hence, for $\epsilon = 1/9$, we can take $l = 2$ and so R_2 will approximate L to within ϵ . Calculating R_2 by Theorem 1 yields that $R_2 = \text{conv}\{(11/18, 7/18), (7/18, 11/18)\}$. $\square$

For m large, R_m can be difficult to compute. For this reason we now show how to find a set S_m containing R_m but which can be calculated more easily. For this, note that

$$xP^k \leq xA_1 \dots A_k \leq xQ^k$$

for any stochastic vector x . Thus, setting $z = xA_1 \dots A_k$, it follows that

$$\min_i p_{ij}^{(k)} \leq z_j \leq \max_i q_{ij}^{(k)}$$

for all j . Now let

$$S_k = \left\{ \text{stochastic vectors } y \mid \min_i p_{ij}^{(k)} \leq y_j \leq \max_i q_{ij}^{(k)} \text{ for all } j \right\}.$$

Then S_k is a convex polytope and $R_k \subseteq S_k$. Two examples are provided below.

EXAMPLE 4. Let $P = (\frac{1}{3} \frac{1}{3})$ and $Q = (\frac{2}{3} \frac{1}{2})$. Then $P^2 = (\frac{2}{9} \frac{2}{9})$ and $Q^2 = (\frac{25}{36} \frac{2}{9})$ so

$$S_2 = \{ \text{stochastic vectors } y \mid 2/9 \leq y_1 \leq 25/36 \text{ and } 2/9 \leq y_2 \leq 25/36 \}.$$

Thus, $S_2 = \text{conv}\{(25/36, 11/36), (11/36, 25/36)\}$. This result may be compared to Example 3.

EXAMPLE 5. Let

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 1/3 & 0 & 1/3 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad Q = \begin{bmatrix} 0 & 1 & 0 \\ 2/3 & 0 & 2/3 \\ 1 & 0 & 0 \end{bmatrix}.$$

Then

$$P^5 = \begin{bmatrix} 2/7 & 1/9 & 1/9 \\ 4/27 & 6/27 & 1/27 \\ 1/9 & 1/3 & 1/9 \end{bmatrix} \quad \text{and} \quad Q^5 = \begin{bmatrix} 8/9 & 4/7 & 4/9 \\ 20/27 & 8/9 & 8/27 \\ 4/9 & 2/3 & 4/9 \end{bmatrix}$$

so that

$$S_5 = \{ \text{stochastic vectors } y \mid 1/9 \leq y_1 \leq 8/9, 1/9 \leq y_2 \leq 8/9, 1/27 \leq y_3 \leq 4/9 \}.$$

Then

$$S_5 = \text{conv}\{(1/9, 23/27, 1/27), (1/9, 4/9, 4/9), \\ (23/27, 1/9, 1/27), (4/9, 1/9, 4/9)\}$$

yielding some information about R_5 and consequently L . $\square$

From the examples it is seen that S_m is more easily calculated than R_m and gives at least some rough estimate of R_m and thus L .

In conclusion, some application for the work of this paper can be seen in a nonhomogeneous Markov chain setting in which the entries in each transition matrix of the chain can only be estimated within specified intervals. A specific situation of this type is given in [6].

REFERENCES

1. Adi Ben-Israel, *Generalized inverses: theory and applications*, Wiley, New York, 1974.
2. F. R. Gantmacher, *The theory of matrices*, Vol. 2, Chelsea, New York, 1964.
3. D. J. Hartfiel, *On infinite products of nonnegative matrices*, SIAM J. Appl. Math. **26** (1974), 297-301.
4. Marvin Marcus and Henryk Minc, *A survey of matrix theory and matrix inequalities*, Allyn and Bacon, Boston, Mass., 1964.
5. R. L. Moore, *Foundations of point set theory*, Amer. Math. Soc. Colloq. Publ., vol. 13, Amer. Math. Soc., Providence, R.I., 1962.
6. T. C. Wesselkamper, *A Markov model of cyclic structured programs* (preprint).

DEPARTMENT OF MATHEMATICS, TEXAS A & M UNIVERSITY, COLLEGE STATION, TEXAS 77843