stichting
mathematisch
centrum

AFDELING BESLISKUNDE BW 58/75 DECEMBER

JAC.M. ANTHONISSE & H.C. TIJMS

ON THE STABILITY OF PRODUCTS OF STOCHASTIC MATRICES

Prepublication

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

AMS(MOS) subject classification scheme (1970): 60J10

On the Stability of Products of Stochastic Matrices *)

by

Jac. M. Anthonisse & H.C. Tijms

ABSTRACT.

This paper considers a finite set of stochastic matrices of finite order. Conditions will be given under which any product of matrices from this set converges to a constant stochastic matrix. Also, it will be shown that the convergence is exponentially fast.

KEY WORDS & PHRASES: *Stochastic matrices, products, exponential convergence.*

---

*) This paper is not for review; it is meant for publication elsewhere.

1. INTRODUCTION

This paper deals with a finite set $\mathcal{P}$ of $N \times N$ stochastic matrices, i.e., for each $P = (p_{ij}) \in \mathcal{P}$, $p_{ij} \geq 0$ and $\Sigma_{j=1}^{N} p_{ij} = 1$ for all $i,j = 1,\ldots,N$. Non-homogeneous Markov chains were studied amongst others in [3], [4] and [7], see also [5] and [6].

Consider the following condition introduced in WOLFOWITZ [7].

C1. For each integer $k \geq 1$ and any $P_i \in \mathcal{P}(1 \leq i \leq k)$ the stochastic matrix $P_k \ldots P_1$ is aperiodic and has a single ergodic class.

This conditon is equivalent to each of the following two conditions.

C2. There is an integer $\nu \geq 1$ such that for each $k \geq \nu$ and any $P_i \in \mathcal{P}(1 \leq i \leq k)$ the matrix $P_k \ldots P_1$ is scrambling, i.e. any two rows of $P_k \ldots P_1$ have a positive entry in a same column (cf.[3]).

C3. There is an integer $\mu \geq 1$ such that for each $k \geq \mu$ and any $P_i \in \mathcal{P}(1 \leq i \leq k)$ the matrix $P_k \ldots P_1$ has a column with only positive entries.

We remark that in C2(C3) it suffices to require the condition imposed on the matrix products only for those of length $\nu(\mu)$. The equivalencies C1 $\Leftrightarrow$ C2 $\Leftrightarrow$ C3 can be seen as follows. Using the fact that a stochastic matrix Q such that Q^n is scrambling for some $n \geq 1$ is aperiodic and has a single ergodic class, we have C3 $\Rightarrow$ C2 $\Rightarrow$ C1. WOLFOWITZ [7] proved that C1 $\Rightarrow$ C2. However, an examination of the proof of Lemma 3 in [7] shows that this lemma remains true when we replace its conclusion that P_1 is scrambling by the conclusion that P_1 has a column with only positive entries. Using this, the proof of Lemma 4 in [7] next shows that C1 $\Rightarrow$ C3.

The purpose of this paper is to show that under C1 for any sequence $\{P_i, i \geq 1\}$ of matrices from $\mathcal{P}$ the matrix product $P_n \ldots P_1$ converges to a constant stochastic matrix as $n \to \infty$. Also, it will be shown that the convergence is exponentially fast. Further, we shall give conditions imposed on the individual matrices from $\mathcal{P}$ such that C1 holds. This paper may have applications amongst others in Markov decision theory (cf. BROWN [1]).

2. CONVERGENCE OF THE MATRIX PRODUCTS

The following theorem generalizes the Theorem in WOLFOWITZ [7] and is related to Theorem 2 in PAZ & REICHAW [4]. Theorem 1 below shows not only that under C1 for any sequence $\{P_i\}$ of matrices from $\mathcal{P}$ the product matrix $P_n \dots P_1$ converges to a constant stochastic matrix as $n \to \infty$ but its proof which was suggested by the one given on pp. 173-174 in DOOB [2] shows also that the convergence is exponentially fast where the convergence rate is uniformly bounded in all sequences $\{P_i\}$.

THEOREM 1. *Suppose that* C1 *holds. Then there is an integer* $\nu \geq 1$, *a number* α *with* $0 \leq \alpha < 1$ *and for any sequence* $\{P_i, i \geq 1\}$ *of matrices from* $\mathcal{P}$ *there is a probability distribution* $\{\pi_j, 1 \leq j \leq N\}$ *such that, for all* $i,j = 1,\dots,N$,

$$(1) \qquad |(P_n \dots P_1)_{ij} - \pi_j| \leq \alpha^{[n/\nu]} \qquad \textit{for all } n \geq 1,$$

where $[x]$ *is the largest integer less than or equal to* x.

PROOF. We first introduce some notation. For any $N \times N$ stochastic matrix Q, let

$$\gamma(Q) = \min_{i_1,i_2} \sum_{j=1}^{N} \min(q_{i_1 j}, q_{i_2 j})$$

and, for $j = 1,\dots,N$, let

$$M_j(Q) = \max_i q_{ij} \quad \text{and} \quad m_j(Q) = \min_i q_{ij}$$

Observe that $\gamma(Q) > 0$ if and only if Q is scrambling. By Lemma 4 in WOLFOWITZ [7], we can choose an integer $\nu \geq 1$ such that the matrix $P_\nu \dots P_1$ is scrambling for any $P_i \in \mathcal{P}(1 \leq i \leq \nu)$. Then, by the finiteness of $\mathcal{P}$,

$$\gamma = \min\{\gamma(P_\nu \dots P_1) \mid P_i \in \mathcal{P}(1 \leq i \leq \nu)\} > 0.$$

Now choose any sequence $\{P_i, i \geq 1\}$ of matrices from $\mathcal{P}$. For any $n \geq m \geq 1$, put for abbreviation $P_{n,m} = P_n \dots P_m$. From $(P_{n+1,1})_{ij} = \Sigma_k (P_{n+1})_{ik} (P_{n,1})_{kj}$ it follows that for all $j = 1,\dots,N$,

$$(2) \qquad M_j(P_{n+1,1}) \leq M_j(P_{n,1}) \quad \text{and} \quad m_j(P_{n+1,1}) \geq m_j(P_{n,1}) \qquad \text{for all } n \geq 1.$$

Now, fix i,h and $n > \nu$. For any number a, let $a^+ = \max(a,0)$ and $a^- = -\min(a,0)$, so, $a = a^+ - a^-$ and $a^+, a^- \geq 0$. Using the fact that

$(a-b)^+ = a - \min(a,b)$ and that $\Sigma_1^N a_j^+ = \Sigma_1^N a_j^-$ when $\Sigma_1^N a_j = 0$, we get for any $j = 1,\ldots,N$,

$$(P_{n,1})_{ij} - (P_{n,1})_{hj} = \sum_{k=1}^{N} \{(P_{n,n-\nu+1})_{ik} - (P_{n,n-\nu+1})_{hk}\} (P_{n-\nu,1})_{kj} =$$

$$= \sum_{k=1}^{N} \{(P_{n,n-\nu+1})_{ik} - (P_{n,n-\nu+1})_{hk}\}^+ (P_{n-\nu,1})_{kj} +$$

$$- \sum_{k=1}^{N} \{(P_{n,n-\nu+1})_{ik} - (P_{n,n-\nu+1})_{hk}\}^- (P_{n-\nu,1})_{kj} \leq$$

$$\leq \sum_{k=1}^{N} \{(P_{n,n-\nu+1})_{ik} - (P_{n,n-\nu+1})_{hk}\}^+ \{M_j(P_{n-\nu,1}) - m_j(P_{n-\nu,1})\} =$$

$$= \{1 - \sum_{k=1}^{N} \min[(P_{n,n-\nu+1})_{ik}, (P_{n,n-\nu+1})_{hk}]\} \{M_j(P_{n-\nu,1}) - m_j(P_{n-\nu,1})\} \leq$$

$$\leq (1-\gamma) \{M_j(P_{n-\nu,1}) - m_j(P_{n-\nu,1})\}.$$

Since i and h were arbitrarily chosen, it follows that for all $j = 1,\ldots,N$

$$M_j(P_{n,1}) - m_j(P_{n,1}) \leq (1-\gamma) \{M_j(P_{n-\nu,1}) - m_j(P_{n-\nu,1})\} \quad \text{for all } n > \nu$$

A repeated application of this inequality and the fact that $M_j(Q) - m_j(Q) \leq 1$ for any stochastic matrix Q show that, for all $j = 1,\ldots,N$,

$$(3) \qquad M_j(P_{n,1}) - m_j(P_{n,1}) \leq (1-\gamma)^{[n/\nu]} \qquad \text{for all } n \geq 1.$$

Together (2) and (3) prove that for any $j = 1,\ldots,N$ there is a finite number $\pi_j \geq 0$ such that $M_j(P_{n,1})$ is monotone decreasing to π_j as $n \to \infty$ and $m_j(P_{n,1})$ is monotone increasing to π_j as $n \to \infty$. Next this result, inequality (3) and the definitions of M_j and m_j imply (1) with $\alpha = 1 - \gamma$. Clearly, $\Sigma\pi_j = 1$ since $P_n \ldots P_1$ is a stochastic matrix for all n. □

By Theorem 4.7 on p. 90 in PAZ [5] the integer ν in condition C2 can always be taken less than or equal to $\nu^* = (1/2)\ (3^N - 2^{N+1} + 1)$. Hence, by

$C1 \Leftrightarrow C2$, it is decidable whether C1 holds by checking all matrix products of at most length ν^*. This may be practically impossible when N is large. We shall now discuss conditions imposed on the individual matrices from $\mathcal{P}$ such that C1 holds. Before doing this, we first remark that it was pointed out on p. 235 in HAJNAL [3] that C1 does not generally hold when each $P \in \mathcal{P}$ is aperiodic and has a single ergodic class. Clearly, C1 holds when each $P \in \mathcal{P}$ is scrambling since in that case any product of P's is scrambling. The next theorem gives sufficient condition for a strong version of C3 under the assumption that the set $\mathcal{P}$ has the following "closedness" property.

C. If $P_1, P_2 \in \mathcal{P}$ then, for any $1 \le i \le N$, the matrix obtained from P_1 by replacing the ith row of P_1 by the ith row of P_2 belongs to $\mathcal{P}$.

THEOREM 2. *Suppose that the set* $\mathcal{P}$ *has property* C. *Further, assume that each* $P \in \mathcal{P}$ *has a single ergodic class and that there is an integer* s *with* $1 \le s \le N$ *such that, for each* $P \in \mathcal{P}$, $p_{ss} > 0$ *and* s *is an ergodic state of* P. *Then there is an integer* μ *with* $1 \le \mu \le N - 1$ *such that for all* $k \ge \mu$ *and any* $P_i \in \mathcal{P}(1 \le i \le k)$ *the* sth *column of the matrix* $P_k \ldots P_1$ *has only positive entries.*

PROOF. Let $S(0) = \{s\}$. Define the sets $R(k-1)$ and $S(k)$ for $k \ge 1$ by

$$R(k-1) = \bigcup_{j=0}^{k-1} S(j) \text{ and } S(k) = \{i \mid i \notin R(k-1), \sum_{j \in R(k-1)} p_{ij} > 0 \quad \text{for all } P \in \mathcal{P}\}.$$

From this definition it follows that there is a first integer μ with $1 \le \mu \le N - 1$ such that $R(\mu) = \{1,\ldots,N\}$ when we can prove that $S(k) \ne \phi$ when $R(k-1) \ne \{1,\ldots,N\}$. To do this , assume to the contrary that there is an integer $k \ge 1$ such that $S(k) = \phi$ and $R(k-1) \ne \{1,\ldots,N\}$. Then, for each $i \notin R(k-1)$, we can find a matrix $P^{(i)} \in \mathcal{P}$ such that $p_{ij}^{(i)} = 0$ for all $j \in R(k-1)$. Now, by property C, there is a matrix $P^* \in \mathcal{P}$ whose ith row is equal to the ith row of $P^{(i)}$ for all $i \notin R(k-1)$. Then, $p_{ij}^* = 0$ for all $i \notin R(k-1)$ and $j \in R(k-1)$. However, this is a contradiction since $s \in R(k-1)$ and it is assumed that P^* has a single ergodic class and that s is ergodic under P^*. This proves the existence of the above integer μ. Now, choose $k \ge \mu$, $P_i \in \mathcal{P}(1 \le i \le k)$ and $j \ne s$. By the construction of the sets S(h), we have $(P_k \ldots P_{k-m+1})_{js} > 0$ for some m with $1 \le m \le \mu$. Now since $p_{ss} > 0$ for all P, we get $(P_k \ldots P_1)_{is} > 0$ for all i which proves the desired result.

REFERENCES

[1] BROWN, B.W., On the Iterative Method of Dynamic Programming on a Finite Space Discrete Time Markov Process, *Ann. Math. Statist.*, Vol. 36 (1965), 1279-1285.

[2] DOOB, J.L., *Stochastic Processes*, Wiley, New York, 1953.

[3] HAJNAL, J., Weak Ergodicity in Nonhomogeneous Markov-chains, *Proc. Cambridge Philos. Soc.*, Vol. 54 (1958), 233-246.

[4] PAZ, A. & REICHAW, M., Ergodic Theorems for Sequences of Infinite Stochastic Matrices, *Proc. Cambridge Philos. Soc.*, Vol. 63 (1967), 777-784.

[5] PAZ, A., *Introduction to Probabilistic Automata*, Academic Press, New York, 1971.

[6] SENETA, E., *Non-negative Matrices*, George Allen & Unwin Ltd., London, 1973.

[7] WOLFOWITZ, J., Products of Indecomposable, Aperiodic, Stochastic Matrices, *Proc. Amer. Math. Soc.*, Vol. 14 (1963), 733-737.