Exponential Convergence of Products of Stochastic Matrices

JAC. M. ANTHONISSE AND HENK TIJMS

Mathematisch Centrum, 2e Boerhaavestraat 49, Amsterdam, The Netherlands
Submitted by H. J. Kushner

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

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 $\sum_{i=1}^{N} p_{ij} = 1$ for all $i, j = 1, \ldots, N$. Non-homogeneous Markov chains were studied in among others, [3, 4, 9]; see also [5, 7].

Consider the following conditions introduced in [9].

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

This condition 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 $P_k \cdots P_i$ is scrambling; i.e., any two rows of $P_k \cdots P_i$ 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 $P_k \cdots P_i$ 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 equivalences C1 $\iff$ C2 $\iff$ 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 $\iff$ C2 $\iff$ C1. Wolfowitz [9] proved that C1 $\iff$ C2. However, an examination of the proof of Lemma 3 in [9] 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 [9] next shows that C1 $\iff$ C3.

* This paper is registered as Mathematical Centre Report BW 58/75.
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 \( P_n \cdots P_1 \) converges to a constant stochastic matrix as \( n \to \infty \). Also, it is shown that the convergence is exponentially fast. Further, we give conditions imposed on the individual matrices from \( \mathcal{P} \) such that C1 holds. This paper, among others, may have applications in Markov decision theory (see [1, 8]).

2. Convergence of the Matrix Products

The following theorem generalizes the theorem in [9] and is related to Theorem 2 in [4]. Theorem 1 below shows not only that under C1 for any sequence \( \{P_i\} \) of matrices from \( \mathcal{P} \) the product matrix \( P_n \cdots P_1 \) converges to a constant stochastic matrix as \( n \to \infty \), but its proof which was suggested by the one given in [2, pp. 173–174] 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 \( \alpha \) 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 \( i, j = 1, \ldots, N \),

\[
\left| (P_n \cdots P_1)_{ij} - \pi_j \right| \leq \alpha^{\lceil \nu/n \rceil} \quad \text{for all } n \geq 1,
\]

where \( \lceil x \rceil \) is the largest integer less than or equal to \( x \).

**Proof.** We first introduce some notation. For any \( N \times N \) stochastic matrix \( Q \), define its ergodic coefficient by

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

and, for \( j = 1, \ldots, N \), let

\[
M_n(Q) = \max_i q_{ij} \quad \text{and} \quad m_n(Q) = \min_i q_{ij}.
\]

Observe that \( \gamma(Q) > 0 \) if and only if \( Q \) is scrambling. By [9, Lemma 4] we can choose an integer \( \nu \geq 1 \) such that the matrix \( P_\nu \cdots P_1 \) is scrambling for any \( P_i \in \mathcal{P} (1 \leq i \leq \nu) \). Then, by the finiteness of \( \mathcal{P} \),

\[
\gamma = \min(\gamma(P_\nu \cdots P_1)) P_i \in \mathcal{P} (1 \leq i \leq \nu) > 0.
\]

Now choose any sequence \( \{P_i, i \geq 1\} \) of matrices from \( \mathcal{P} \). For any \( n \geq m \geq 1 \), put for abbreviation \( P_{n,m} = P_n \cdots P_m \). From \( (P_{n+1,1})_{ij} = \Sigma_k (P_{n+1})_{ik} (P_1)_{kj} \) it follows that for all \( j = 1, \ldots, N \),

\[
M_j(P_{n+1,1}) \leq M_j(P_{n,1}) \quad \text{and} \quad m_j(P_{n+1,1}) \geq m_j(P_{n,1}) \quad \text{for all } n \geq 1.
\]
Now, fix $i$, $h$ and $n > v$. For any number $a$, let $a^+ = \max(a, 0)$ and $a^- = -\min(a, 0)$, so that $a = a^+ - a^-$ and $a^+, a^- \geq 0$. Using the fact that $(a - b)^+ = a - \min(a, b)$ and that $\sum_{i=1}^N a_i^+ = \sum_{i=1}^N a_i^-$ when $\sum_{i=1}^N a_i = 0$, we get for any $j = 1, \ldots, N$,

$$(P_{n,1})_{ij} - (P_{n,1})_{kj}$$

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

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

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

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

$$= \left\{1 - \sum_{k=1}^N \min\{(P_{n,n-v+1})_{ik}, (P_{n,n-v+1})_{kh}\}\right\} \{M_j(P_{n-v,1}) - m_j(P_{n-v,1})\}$$

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

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

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

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 $j = 1, \ldots, N$,

$$M_j(P_{n,1}) - m_j(P_{n,1}) \leq (1 - \gamma)^{\lfloor n/v \rfloor} \quad \text{for all } n \geq 1. \quad (3)$$

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 $\pi_j$ as $n \to \infty$ and $m_j(P_{n,1})$ is monotone increasing to $\pi_j$ 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, $\sum \pi_j = 1$ since $P_{n} \cdots P_i$ is a stochastic matrix for all $n$. \[\square\]

We remark that C1 holds when relation (1) applies for any sequence $\{P_i\}$, so that C1 is both sufficient and necessary for the assertion of Theorem 1.

By [5, Theorem 4.7, p. 90] the integer $v$ in condition C2 can always be taken less than or equal to $v^* = (1/2)(3^N - 2^{N+1} + 1)$. Hence, by C1 $\Rightarrow$ C2, one may decide whether C1 holds by checking all matrix products of at most length $v^*$. This may be practically impossible when $N$ is large. We 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 in [3, p. 235] that C1 does not generally hold when each $P \in \mathcal{P}$ is aperiodic and has a single ergodic class (see also [6]). 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 conditions for a strong version of C3 under the assumption that the set $\mathcal{P}$ has the following “product” property.

A. The set $\mathcal{P}$ is the Cartesian product of finite sets of probability distributions.

**Theorem 2.** Suppose that the set $\mathcal{P}$ has property A. Further, assume that each $P \in \mathcal{P}$ has a single ergodic class and that there is an integer $s$ with $1 \leq s \leq N$ such that, for each $P \in \mathcal{P}$, $p_{ss} > 0$ and $s$ is an ergodic state of $P$. Then there is an integer $\mu$ with $1 \leq \mu \leq N - 1$ such that for all $k \geq \mu$ and any $P_i \in \mathcal{P} \{1 \leq i \leq k\}$ the $i$th column of the matrix $P_k \cdots P_1$ has only positive entries.

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

$$R(k - 1) = \bigcup_{j=0}^{k-1} S(j)$$

and

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

From this definition it follows that there is a first integer $\mu$ with $1 \leq \mu \leq N - 1$ such that $R(\mu) = \{1, ..., N\}$ when we can prove that $S(k) \neq \emptyset$ when $R(k-1) \neq \{1, ..., N\}$. To do this, assume to the contrary that there is an integer $k \geq 1$ such that $S(k) = \emptyset$ and $R(k - 1) \neq \{1, ..., N\}$. 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 A, there is a matrix $P^{*} \in \mathcal{P}$ whose $i$th row is equal to the $i$th 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 $i \in R(k - 1)$ 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 $\mu$. Now, choose $k \geq \mu$, $P_i \in \mathcal{P} \{1 \leq i \leq k\}$ and $j \neq s$. By the construction of the sets $S(h)$, we have $(P_k \cdots P_{k-m+1})_{is} > 0$ for some $m$ with $1 \leq m \leq \mu$. Now since $p_{is} > 0$ for all $P$, we get $(P_k \cdots P_1)_{is} > 0$ for all $i$, which proves the desired result.

**References**