Statistics and Probability Letters 78 (2008) 3082–3085

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Conditions for weak ergodicity of inhomogeneous Markov chains

Don Coppersmith[a], Chai Wah Wu[b,*]

[a] IDA Center for Communications Research, Princeton, NJ 08540, USA

[b] IBM Research Division, Thomas J. Watson Research Center, P. O. Box 704, Yorktown Heights, NY 10598, USA

ARTICLE INFO

Article history:
Received 10 January 2005
Received in revised form 6 September 2007
Accepted 1 May 2008
Available online 18 May 2008

ABSTRACT

A classical result of Wolfowitz states that an inhomogeneous Markov chain is weakly ergodic if the transition matrices are drawn from a finite set of indecomposable and aperiodic matrices and the products of transition matrices are also indecomposable and aperiodic. Since products of indecomposable and aperiodic matrices can be decomposable, any finite set of indecomposable and aperiodic transition matrices does not guarantee weak ergodicity. We present conditions for weak ergodicity which are simpler to verify and are related to properties of the graph of the transition matrices.

1. Introduction

Consider an inhomogeneous Markov chain whose transition matrix at time i is denoted A_i. Each A_i is an n by n stochastic matrix, i.e. a nonnegative matrix whose row sums are equal to 1. Let $A^{i:j} = A_iA_{i+1}\cdots A_j$ if $i \leq j$ and $A^{i:j} = A_iA_{i-1}\cdots A_j$ if $i \geq j$.

Definition 1 (*Hajnal (1958)*). The Markov chain is *weakly ergodic* if

$$|A^{i:j}_{k,l} - A^{i:j}_{k',l}| \to 0$$

as $j \to \infty$ for all i, k, k', l. The Markov chain is *strongly ergodic* if it is weakly ergodic and for any i, the matrix product $A^{i:j}$ converges as $j \to \infty$.

Remark 1. Weak ergodicity as defined in Definition 1 in general does not imply strong ergodicity. However, if the products of matrices are formed in the backward direction, i.e. replacing $A^{i:j}$ with $A^{j:i}$ in Definition 1, then weak ergodicity is equivalent to strong ergodicity (Chatterjee and Seneta, 1977).

Definition 2. A stochastic matrix is said to be *Markov* if there exists a column with only positive entries. A stochastic matrix A is said to be *scrambling* if for $i \neq j$, there exists k such that $A_{ik} > 0$ and $A_{jk} > 0$.

Clearly Markov matrices are scrambling but not necessarily vice versa.

Definition 3. A stochastic matrix A is said to be indecomposable and aperiodic (SIA) if $\lim_{n\to\infty} A^n = \mathbf{1}c^{\mathrm{T}}$ where $\mathbf{1}$ is the vector of all 1's.

This definition is equivalent to classical definitions of indecomposability and aperiodicity of stochastic matrices (Wolfowitz, 1963).

* Corresponding author. Tel.: +1 914 784 7231.
E-mail addresses: dcopper@idaccr.org (D. Coppersmith), cwwu@us.ibm.com, chaiwahwu@ieee.org (C.W. Wu).

0167-7152/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.spl.2008.05.012

Some early sufficient conditions show that the chain is weakly ergodic if the set of transition matrices is compact and either Markov or scrambling (Seneta, 1973; Hajnal, 1958). In Wolfowitz (1963) the following sufficient condition is given for a chain to be weakly ergodic:

Theorem 1. *Let P be a finite set of matrices. If all finite products of matrices in P are SIA, then the Markov chain with transition matrices in P is weakly ergodic.*

The key to this result is the following lemma:

Lemma 1. *Let P be a set of SIA matrices whose finite products are SIA. There is a constant t, which depends only on the order of the matrices, such that all products of matrices from P of length at least t is scrambling.*

In Anthonisse and Tijms (1977) it was shown that the condition in Theorem 1 implies left convergence of matrix products (see Remark 1):

Theorem 2. *If the hypothesis of Theorem 1 is satisfied and $A_i \in P$ for all i, then $A_nA_{n-1}\dots A_1$ converges exponentially to a rank-one matrix $\mathbf{1}c^{\mathrm{T}}$ as $n \to \infty$.*

The finiteness condition for the set P can be replaced with compactness. This is because the only place finiteness of P is used in Wolfowitz (1963); Anthonisse and Tijms (1977) is to ensure that the *ergodicity coefficient* $\gamma(A) = \min_{i,j}\sum_{k=1}^{n}\min(A_{ik}, A_{jk})$, which is positive for scrambling matrices and is continuous with respect to the matrix elements, has a nonzero lower bound for products of matrices of length t.

Theorem 3. *Let P be a compact set of matrices with $A_i \in P$. If all finite products of matrices in P are SIA, then the Markov chain with transition matrices in P is weakly ergodic and $A_nA_{n-1}\dots A_1$ converges exponentially to a rank-one matrix $\mathbf{1}c^{\mathrm{T}}$ as $n \to \infty$.*

It should be noted that if a subchain forms a weakly ergodic Markov chain, then the original Markov chain is also weakly ergodic:

Theorem 4. *Let there be a sequence $m_0 < m_1 < \cdots$ and $B_i = A^{m_i:m_{i+1}-1}$, $C_i = A^{m_{i+1}-1:m_i}$. If the Markov chain with B_i as the transition matrices is weakly ergodic, then so is the Markov chain with A_i as the transition matrices. If $C_nC_{n-1}\cdots C_1$ converges to a rank-one matrix $\mathbf{1}c^{\mathrm{T}}$ as $n \to \infty$, then so does $A_nA_{n-1}\dots A_1$.*

Proof. Follows from Lemma 3 in Shen (2000) and from Anthonisse and Tijms (1977). □

Simple examples show that products of SIA matrices can be decomposable (Hajnal, 1958). Therefore to show weak ergodicity it is not enough that the transition matrices are SIA. The number of matrix products in Theorem 1 that needs to be checked for the property of SIA is less than $\frac{1}{2}(3^n - 2^{n+1} + 1)$ where n is the order of the matrices (Paz, 1971).

2. Matrices with positive diagonal elements

The *reversal* of a directed graph (digraph) is defined as the directed graph obtained by reversing the orientation of all the edges. If we restrict ourselves to matrices with nonzero diagonal elements, then SIA matrices correspond to stochastic matrices whose corresponding digraph contains a spanning directed tree after reversal (Wu, 2005b). Furthermore, products of SIA matrices remain SIA. This is due to the fact that the property of SIA depends only on the locations of the nonzero elements, and not their values, and multiplication with nonnegative matrices with nonzero diagonal elements can only add more nonzero elements and thus preserves the property of the graph containing a spanning directed tree. Thus the following result follows from Theorem 3.

Theorem 5. *If P is a compact set of SIA matrices with nonzero diagonal elements, then the Markov chain with transition matrices in P is weakly ergodic. Furthermore, if $A_i \in P$, then $A_nA_{n-1}, \dots A_1$ converges exponentially to a rank-one matrix $\mathbf{1}c^{\mathrm{T}}$ as $n \to \infty$.*

Remark 2. In Anthonisse and Tijms (1977) another property of matrices is defined such that if all transition matrices satisfy this property, then the corresponding Markov chain is weakly ergodic.

Definition 4. The set V consists of stochastic matrices A with positive diagonal elements such that

- The nonzero elements within each row are the same,
- $A_{ij} > 0 \Leftrightarrow A_{ji} > 0$.

In other words, the nonzero elements of matrices in V are of the from $\frac{1}{k}$ and form a symmetric pattern. The purpose of this paper is to consider Markov chains where the transition matrices are not necessarily SIA, but whose products are. A convergence result regarding such Markov chains is given by Jadbabaie et al. (2003):

Theorem 6. *If there exists a finite set $P \subset V$, a constant N and an infinite sequence $m_1 < m_2 < \cdots$ such that for all i,*

- *$A_i \in P$,*

- $m_{i+1} - m_i \le N$,
- $A^{m_{i+1}-1:m_i}$ *is irreducible*

then $A_n \cdots A_2 A_1$ *converges to a rank-one matrix* $\mathbf{1}c^{\mathrm{T}}$ *as* $n \to \infty$.

Definition 5. Let $S(\delta)$ be the set of n by n stochastic matrices with nonzero diagonal elements such that all nonzero elements are larger than or equal to δ.

Lemma 2. *Let A_i be stochastic matrices with nonzero diagonal elements. Then $\frac{1}{n}\sum_{i=1}^{n} A_i$ is SIA if and only if $A^{1:n}$ is SIA. In particular, $A^{1:n}$ is SIA if and only if $A^{n:1}$ is SIA.*

Proof. A_i's having nonzero diagonal elements implies that if the (k, l)th element of A_j is nonzero for some j, then so is the corresponding element in $A_1 A_2 \cdots A_n$. This means that if the reversal of the graph of $\frac{1}{n}\sum_{i=1}^{n} A_i$ contains a spanning directed tree, then so does the reversal of the graph of $A_1 A_2 \cdots A_n$. On the other hand, suppose the reversal of the graph of $A_1 A_2 \cdots A_n$ contains a spanning directed tree. A graph containing a spanning directed tree is equivalent to the property of quasi-strongly connected, i.e. for vertices j and k, there exists a vertex i with directed paths from i to j and from i to k (Swamy and Thulasiraman, 1981). For each edge (i, j) in the graph of $A_1 A_2 \cdots A_n$ which is not in the graph of $\frac{1}{n}\sum_{i=1}^{n} A_i$, there is a directed path from i to j in the graph of $\frac{1}{n}\sum_{i=1}^{n} A_i$. Therefore the reversal of the graph of $\frac{1}{n}\sum_{i=1}^{n} A_i$ is also quasi-strongly connected. □

Theorem 6 is generalized in the next result (Wu, 2004, 2005a), noting that by Lemma 2 the matrix product $A^{m_i:m_{i+1}-1}$ is SIA if and only if $A^{m_{i+1}-1:m_i}$ is SIA.

Theorem 7. *If there exists constants N, $\delta > 0$ and an infinite sequence $m_1 < m_2 < \cdots$ such that for all i*

- $A_i \in S(\delta)$,
- $m_{i+1} - m_i \le N$,
- $A^{m_i:m_{i+1}-1}$ *is SIA,*

then the Markov chain is weakly ergodic and $A_n A_{n-1} \cdots A_1$ converges to a rank-one matrix $\mathbf{1}c^{\mathrm{T}}$ as $n \to \infty$.

3. Block diagonal matrices with irreducible blocks

In Theorem 7 there is a uniform bound N on the matrix product lengths. In Moreau (2003) an example is given which shows that Theorem 7 is false if this condition of the uniform bound N (i.e. $m_{i+1} - m_i \le N$) is omitted. The reason is that if $m_{i+1} - m_i$ is arbitrarily large, then even though $A^{m_i:m_{i+1}-1}$ is SIA, products of such SIA matrices might have arbitrarily small coefficients of ergodicity γ.

The main result of this paper is a sufficient condition for weak ergodicity which does not require this uniform bound. It turns out that this bound N is not needed if the graphs of A_i are undirected.

Definition 6. Let M_{bd} denote the set of matrices which can be written (possibly after simultaneous row and column permutation) in block diagonal form, where each diagonal block is irreducible.

Equivalently, matrices in M_{bd} are those matrices whose digraph consists of disjoint strongly connected components. M_{bd} can also be characterized as the matrices whose digraph does not contain a subgraph which is weakly connected but not strongly connected. An important subclass of matrices in M_{bd} are the stochastic matrices A such that $A_{ij} > 0 \Leftrightarrow A_{ji} > 0$, i.e. the graph of A (ignoring the weights on the edges) is undirected.

Lemma 3. *If $\delta \le 1$ and $A_i \in S(\delta) \cap M_{bd}$ for all i, then $A^{1:k} \in S(\delta^{n-1})$ for any k, where n denotes the order of the matrices A_i.*

Proof. Fix an index l. Let $x(k)$ be the lth row of $A^{1:k}$. We show by induction on k that the sum of any j elements of $x(k)$ is either 0 or greater than or equal to δ^{n-j} for each $j = 1, \ldots, n$. This is clearly true for $k = 1$. Note that $x(k) = x(k-1)A_k$. Pick n_j elements of $x(k)$, say with index set $J \subset \{1, \ldots, n\}$. They correspond to the matrix product of $x(k-1)$ with n_j columns of A_k. Let K be the indices of the rows of A_k which have nonzero intersection with some of these columns. Clearly $J \subset K$ since A_k has positive diagonal elements. If $K = J$, then $A_k(i, j) = 0$ for $i \notin J, j \in J$. Since $A_k \in M_{bd}$ this implies that $A_k(i, j) = 0$ for $i \in J, j \notin J$. This implies that the elements of $A_k(i, j)$ for $i, j \in J$ correspond to a n_j by n_j stochastic submatrix, and thus the sums of the n_j elements in $x(k)$ is the same as the sum of these corresponding elements in $x(k-1)$.

If K is strictly larger than J, then the sum of the n_j elements in $x(k)$ is larger than or equal to the sum of the elements in $x(k-1)$ with indices in K multiplied by nonzero elements in A_k each of which is at least δ. Since the sum of the elements in $x(k-1)$ with indices in K is at least δ^{n-j-1} or zero, the result follows. □

Remark 3. The bound δ^{n-j} is the best possible in the proof of Lemma 3. Let A_i be the identity matrix except that $A_i(i, i) = A_i(i+1, i+1) = 1 - \delta$, $A_i(i, i+1) = A(i+1, i) = \delta$. Then $A_1 A_2 \ldots A_{n-1}$ has an element of size δ^{n-1} and the jth row of

$A_1A_2 \ldots A_{n-j}$ has j elements whose sum is δ^{n-j}. If we remove the condition of positive diagonal elements, then Lemma 3 is false. For instance, take

$$A_1 = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{pmatrix}.$$

Then $(A_1A_2)^m$ has elements of size $\frac{1}{4^m}$.

Theorem 8. *If there exists $\delta > 0$ and an infinite sequence $m_1 < m_2 < \cdots$ such that $A_i \in S(\delta) \cap M_{bd}$ and $A^{m_i:m_{i+1}-1}$ is SIA for all i, then the Markov chain is weakly ergodic and $A_nA_{n-1}\ldots A_1$ converges to a rank-one matrix $\mathbf{1}c^{\mathrm{T}}$ as $n \to \infty$.*

Proof. Without loss of generality, we can assume that $\delta \leq 1$. By Lemma 3, $P_i = A^{m_i:m_{i+1}}$ are elements of $S(\delta^{n-1})$. By Lemma 1, $B_i = P_{ti+1}P_{ti+2}\cdots P_{t(i+1)} \in S(\delta^{t(n-1)})$ is scrambling. Since $\gamma(B_i) > 0$, this means that $\gamma(B_i) \geq \delta^{t(n-1)} > 0$. Since the lower bound $\delta^{t(n-1)}$ does not depend on the length of the products in B_i, the arguments in Wolfowitz (1963); Anthonisse and Tijms (1977) along with Theorem 4 and Lemma 2 can be used to prove the result. □

Note that there is no uniform bound on the length of the matrix product in Theorem 8, i.e., $m_{i+1} - m_i$ can be arbitrarily large. In particular, this means that the condition $m_{i+1} - m_i \leq N$ is not necessary in Theorem 6. Theorem 8 can be described in terms of the graphs of A_i as follows: The Markov chain is weakly ergodic if for all i,

- $A_i \in S(\delta)$;
- The digraph of A_i consists of disjoint strongly connected components;
- The union of the reversal of the digraphs of $A_{m_i}, A_{m_i+1}, \ldots, A_{m_{i+1}-1}$ contains a spanning directed tree.

4. Conclusions

We study Markov chains with transition matrices with nonzero diagonal elements and give conditions for weak ergodicity that depends on the properties of the graph of the transition matrices.

References

Anthonisse, J.M., Tijms, H., 1977. Exponential convergence of products of stochastic matrices. Journal of Mathematical Analysis and Applications 59, 360–364.
Chatterjee, S., Seneta, E., 1977. Towards consensus: Some convergence theorems on repeated averaging. Journal of Applied Probability 14, 89–97.
Hajnal, J., 1958. Weak ergodicity in non-homogeneous Markov chains. Proc. Cambridge Philos. Soc. 54, 233–246.
Jadbabaie, A., Lin, J., Morse, A.S., 2003. Coordination of groups of mobile autonomous agents using nearest neighbor rules. IEEE Transactions on Automatic Control 48 (6), 988–1001.
Moreau, L., 2003. Leaderless coordination via bidirectional and unidirectional time-dependent communication.
Paz, A., 1971. Introduction to Probabilistic Automata. Academic Press.
Seneta, E., 1973. On the historical development of the theory of finite inhomogeneous Markov chains. Proceedings of the Cambridge Philosophical Society 74, 507–513.
Shen, J., 2000. A geometric approach to ergodic non-homogeneous Markov chains. In: Wavelet Analysis and Multiresolution Methods. In: Lecture Notes in Pure and Applied Math., vol. 212. Marcel Dekker, pp. 341–366.
Swamy, M.N.S., Thulasiraman, K., 1981. Graphs, Networks, and Algorithms. John Wiley & Sons.
Wolfowitz, J., 1963. Products of indecomposable, aperiodic, stochastic matrices. Proc. Amer. Math. Soc. 14, 733–737.
Wu, C.W., 2004. Agreement and consensus problems in groups of autonomous agents with linear dynamics, Research Report RC23403, IBM.
Wu, C.W., 2005a. Agreement and consensus problems in groups of autonomous agents with linear dynamics, in: Proceedings of 2005 IEEE International Symposium on Circuits and Systems, pp. 292–295.
Wu, C.W., 2005b. On bounds of extremal eigenvalues of irreducible and m-reducible matrices. Linear Algebra and its Applications 402, 29–45.