STUDIA MATHEMATICA, T. XLVI. (1973)

On strong ergodicity of inhomogeneous products of finite stochastic matrices

by

E. SENETA (Canberra)

Abstract. For a sequence $\{P_k\}$ of $n \times n$ stochastic matrices, where $P_k = \{p_{ij}(k)\}$, such that: (1) $T_{r,k} = P_{r+1}P_{r+2}\ldots P_{r+k}$ is regular for each $r \geqslant 0$, $k \geqslant 1$; (2) $\min^+_{i,j} p_{ij}(k) \geqslant \delta > 0$ uniformly for all $k \geqslant 1$, where $\min^+$ refers to the minimum among all strictly positive elements; and (3) the sequence $\{P_k\}$ is asymptotically homogeneous, it is shown strong ergodicity obtains. This generalizes a result of Bernstein [1], who assumes in addition that all P_k are Markov matrices; and of Sarymsakov [8] and Sarymsakov and Mustafin [9] who show (1) and (2) are sufficient for weak ergodicity. Relationship to the results of Kozniewska [3] on necessary and sufficient conditions for strong ergodicity is also briefly discussed.

1. Introduction. In this note, all matrices are of fixed size: $n \times n$; and all vectors contain n elements. Let $\{P_k\}$, $k \geqslant 1$ be a sequence of stochastic matrices (i.e. matrices with non-negative entries and unit row sums), where $P_k = \{p_{ij}(k)\}$ and let $T_{r,k} = \{t_{i,j}^{(r,k)}\}$ be the stochastic matrix defined by

$$T_{r,k} = P_{r+1}P_{r+2}\ldots P_{r+k}$$

for $r \geqslant 0$, $k \geqslant 1$.

The sequence of matrices $\{P_k\}$ is said to be *weakly ergodic* (in the sense of Kolmogorov) if for all $i, j, s = 1, \ldots, n$ and $r \geqslant 0$

$$[t_{i,s}^{(r,k)} - t_{j,s}^{(r,k)}] \to 0$$

as $k \to \infty$. The sequence is said to be *strongly ergodic* if

$$\lim_{k\to\infty} T_{r,k} = \mathbf{1}\boldsymbol{D}_r', \quad r \geqslant 0$$

elementwise, where $\boldsymbol{D}_r$ is necessarily a probability vector (i.e. $\boldsymbol{D}_r \geqslant \mathbf{0}$, $\boldsymbol{D}_r'\mathbf{1} = 1$), but may depend on r.

The sequence $\{P_k\}$ is said to be *asymptotically homogeneous* (in the sense of Bernstein) if there exists a probability vector $\boldsymbol{D}$ such that

$$\lim_{k\to\infty} \boldsymbol{D}' P_k = \boldsymbol{D}'.$$

Koźniewska [3] has introduced the related concept of *asymptotic stationarity*: $\{P_k\}$ is said to be asymptotically stationary if there exists a probability vector $\boldsymbol{D}$ such that

$$\lim_{k\to\infty} \boldsymbol{D}' T_{r,k} = \boldsymbol{D}', \quad r \geqslant 0,$$

and has shown that, just as strong ergodicity implies weak ergodicity, asymptotic stationarity implies asymptotic homogeneity. She has, further, shown e.g. that weak ergodicity and asymptotic stationarity are together necessary and sufficient for strong ergodicity.

There are several necessary and sufficient conditions known for weak ergodicity also (see Koźniewska [3] for one such and her references to earlier work, especially Hajnal [2]), but from a practical viewpoint a useful sufficient condition is given by the following theorem of Sarymsakov [8] and Sarymsakov and Mustafin [9], where G_1 is the class of $(n\times n)$ *regular* stochastic matrices (i.e. stochastic matrices having a single eigenvalue(1) of modulus unity; or, in Markov chain terms, transition matrices corresponding to homogeneous Markov chains whose essential states form a single aperiodic class).

THEOREM A. *If for each $r \geqslant 0$, $T_{r,k}\in G_1$ for all $k \geqslant 1$, and*

$$\min_{i,j}{}^{+} p_{ij}(k) \geqslant \delta > 0$$

uniformly for all $k\geqslant 1$ (where $\min^+$ refers to the minimum non-zero entry), then weak ergodicity obtains (uniformly for all r).

Actually, Sarymsakov and Mustafin have, instead of the condition $T_{r,k}\in G_1$, the stronger condition that each $P_k \in G_2$, where G_2 is a suitable subset of G_1; and their proof is rather concise. However it is not difficult to see from their work that the theorem as stated is valid; a complete proof will be given by the present author in [10].

The strength of Theorem A may be seen from the fact that it subsumes later, and quite practically useful, theorems of Wolfowitz [11] and Paz [6], [7].

The main purpose of the present note is to prove a strong ergodicity analogue of Theorem A, viz.

THEOREM B. *Under the conditions of Theorem A on the sequence $\{P_k\}$, and the additional assumption of asymptotic homogeneity, strong ergodicity obtains.*

We shall need also the concept of a *Markov matrix*. A Markov matrix is a stochastic matrix $P = \{p_{ij}\}$ satisfying $\lambda(P) \equiv \max_j(\min_i p_{ij}) > 0$: i.e. having a positive column. A sequence $\{P_k\}$ of stochastic matrices is said to be *uniformly Markov* if $\lambda(P_k) \geqslant \lambda_0 > 0$ for all k. The product of any stochastic matrix, in either order, with a Markov matrix, is Markov; and

(1) counting repeated eigenvalues as distinct.

a Markov matrix clearly is a member of G_1. Thus if M is the class of Markov matrices, $M \subset G_1$.

Theorems A and B for a sequence $\{P_k\}$ of uniformly Markov matrices were proved by Bernstein [1]; Theorem A in essence being due to Markov [4] himself, in this restricted situation. Sarymsakov and Mustafin [9] have in essence shown that these classical arguments may be extended to give Theorem A as it stands; and we shall do the same in the case of Theorem B.

Finally the reader should note that Koźniewska [3], in her Theorem 4, has given a necessary and sufficient condition for strong ergodicity in terms of asymptotic homogeneity. Her arguments could be adapted and expanded for our purposes, but it seems simpler to proceed directly.

2. Preliminary Results.

LEMMA 1. *Let $P = \{p_{ij}\}$ be a stochastic matrix and $\boldsymbol{\delta} = \{\delta_i\}$ a vector of real elements satisfying $\boldsymbol{\delta} \neq \boldsymbol{0}$, $\boldsymbol{\delta}'\boldsymbol{1} = 0$. Let $\Delta_0 = \Sigma|\delta_i|$* and *$\Delta_1 = \Sigma|\delta_i^*|$ where $\boldsymbol{\delta}^* = \{\delta_i^*\}$ is defined by $(\boldsymbol{\delta}^*)' = \boldsymbol{\delta}'P$. Then*

$$\Delta_1 \leqslant \big(1-\lambda(P)\big)\Delta_0$$

where $\lambda(P)\geqslant 0$ is as defined in § 1.

This result is in essence due to Markov [4], although rediscovered many times (see e.g. Mott [5]).

Lemma 2 is due to Sarymsakov and Mustafin [9], although the reader may prefer the simpler approach in Wolfowitz [11] (Lemmas 3 and 4, where the word "scrambling" may be replaced by "Markov" without altering the proofs).

LEMMA 2. *If for each $r\geqslant 0$, $T_{r,k}\in G_1$ for all $k\geqslant 1$, then $T_{r,k}\in M$ for $k\geqslant t+1$, where t is the number of distinct types(2) of matrix in G_1.*

LEMMA 3. *If the sequence $\{P_i\}$ is asymptotically homogeneous in respect to probability vector $\boldsymbol{D}$, then*

$$\lim_{r\to\infty} \boldsymbol{D}' T_{r,k} = \boldsymbol{D}', \quad \textit{for each } k\geqslant 1.$$

Proof. This is by induction. The proposition is true for $k=1$ by assumption since $T_{r,1} = P_{r+1}$. Assume it is true for some $k\geqslant 1$. Then

$$\boldsymbol{D}'T_{r,k+1} = \boldsymbol{D}'T_{r,k}P_{r+k+1}$$

and, writing $\boldsymbol{D}'T_{r,k} = \boldsymbol{D}'+\boldsymbol{E}'_{r,k}$ where $\boldsymbol{E}_{r,k}\to\boldsymbol{0}$ as $r\to\infty$ (by induction hypothesis), we have

$$\boldsymbol{D}'T_{r,k+1} = \boldsymbol{D}'P_{r+k+1}+\boldsymbol{E}'_{r,k}P_{r+k+1}$$

(2) i. e. with regard to location of positive elements, but not their actual values.

so, using asymptotic homogeneity, and the fact that P_{r+k+1}, being stochastic, is uniformly elementwise bounded as $k \to \infty$,

$$\lim_{r\to\infty} \boldsymbol{D}' T_{r,k+1} = \boldsymbol{D}'$$

which completes the proof.

LEMMA 4. *For a given sequence $\{P_k\}$ of stochastic matrices, suppose*

$$\boldsymbol{\delta}'(k+1) = \boldsymbol{\delta}'(k)P_{k+1} + \boldsymbol{r}'(k), \qquad k \geqslant 0$$

where $\boldsymbol{\delta}'(k)\mathbf{1} = 0 = \boldsymbol{r}'(k)\mathbf{1}$. Then

$$\Delta(k+1) \leqslant \Delta(0)\prod_{i=1}^{k+1}\big(1-\lambda(P_i)\big) + \prod_{i=0}^{k-1}\Gamma(i)\prod_{j=i+2}^{k+1}\big(1-\lambda(P_i)\big) + \Gamma(k)$$

for $k \geqslant 0$, where $\Delta(k) = \sum_i |\delta_i(k)|$, $\Gamma(k) = \sum_i |r_i(k)|$, and $\sum_{i=0}^{-1}$ is to be interpreted as 0.

Proof. By induction. For $k = 0$

$$\boldsymbol{\delta}'(1) - \boldsymbol{r}'(0) = \boldsymbol{\delta}'(0)P_1$$

so that, by Lemma 1, together with the elementary inequality $|a-b| \geqslant |a|-|b|$

$$\Delta(1) \leqslant \Delta(0)\big(1-\lambda(P_1)\big) + \Gamma(0)$$

so the assertion is true for $k = 0$. For arbitrary $k \geqslant 0$, similarly

$$\boldsymbol{\delta}'(k+2) - \boldsymbol{r}'(k+1) = \boldsymbol{\delta}'(k+1)P_{k+2}$$

so that

$$\Delta(k+2) \leqslant \Delta(k+1)\{1-\lambda(P_{k+2})\} + \Gamma(k+1).$$

Applying the induction hypothesis to $\Delta(k+1)$, the result follows.

COROLLARY. *If all P_k in $\{P_k\}$ are uniformly Markov (i.e. $\lambda(P_k) \geqslant \lambda_0 > 0$ all k), and also $\boldsymbol{r}(k) \to \mathbf{0}$ elementwise as $k \to \infty$, then*

$$\Delta(k) \to 0.$$

Proof. In the first place

$$\begin{aligned}\sum_{i=0}^{k-1}\Gamma(i)\prod_{j=i+2}^{k+1}\big(1-\lambda(P_i)\big) &\leqslant \sum_{i=0}^{k-1}\Gamma(i)(1-\lambda_0)^{k-i}\\ &= \sum_{i=0}^{j}\Gamma(i)(1-\lambda_0)^{k-i} + \sum_{i=j+1}^{k-1}\Gamma(i)(1-\lambda_0)^{k-i}\\ &\leqslant (1-\lambda_0)^{k-j}\sum_{i=0}^{j}\Gamma(i) + \sum_{i=j+1}^{k-1}\Gamma(i)(1-\lambda_0)^{k-i}.\end{aligned}$$

Now select $j \equiv j(\varepsilon)$ such that $\Gamma(i) < \varepsilon$ for $i > j$; and make use of the fact that $\Gamma(i) < C = \text{const.}$ for all i. It follows that, by judicious choice of ε, the right hand side will be arbitrarily small if k is sufficiently large.

The truth of the Corollary is now trivial.

LEMMA 5. *For a sequence $\{P_k\}$ of uniformly Markov matrices, asymptotic stationarity and asymptotic homogeneity are equivalent.*

Proof. As already noted, Koźniewska has shown that asymptotic stationarity always implies asymptotic homogeneity. We need to prove now that in the present situation, the converse is true.

By the assumption of asymptotic homogeneity, there is a probability vector $\boldsymbol{D}$ such that

$$\boldsymbol{D}'P_k + \boldsymbol{e}'(k) = \boldsymbol{D}'$$

where $\boldsymbol{e}(k) \to \mathbf{0}$ as $k \to \infty$, and $\boldsymbol{e}'(k)\mathbf{1} = 0$. Let us write for $k \geqslant 1$

$$\boldsymbol{D}'P_k = \boldsymbol{D}'T_{0,k} + \boldsymbol{\delta}'(k).$$

Then

$$\boldsymbol{D}'P_{k+1} = \boldsymbol{D}'T_{0,k+1} + \boldsymbol{\delta}'(k+1).$$

On the other hand

$$\begin{aligned}\boldsymbol{D}'P_{k+1} &= \big(\boldsymbol{D}'P_k + \boldsymbol{e}'(k)\big)P_{k+1}\\ &= \big(\boldsymbol{D}'T_{0,k} + \boldsymbol{\delta}'(k) + \boldsymbol{e}'(k)\big)P_{k+1}\end{aligned}$$

so that, finally, for all $k \geqslant 0$

$$\boldsymbol{\delta}'(k+1) = \boldsymbol{\delta}'(k)P_{k+1} + \boldsymbol{r}'(k)$$

where $\boldsymbol{r}'(k) = \boldsymbol{e}'(k)P_{k+1}$, so that, by stochasticity $\boldsymbol{r}'(k)\mathbf{1} = 0 = \boldsymbol{\delta}'(k)\mathbf{1}$.

Moreover $\boldsymbol{r}(k) \to \mathbf{0}$ as $k \to \infty$, since $\boldsymbol{r}'(k) = \boldsymbol{e}'(k)P_{k+1}$, so the Corollary to Lemma 4 gives

$$\lim_{k\to\infty}\boldsymbol{D}'P_k = \lim_{k\to\infty}\boldsymbol{D}'T_{0,k} = \boldsymbol{D}'.$$

The same reasoning follows for $\boldsymbol{D}'T_{r,k}$, for any fixed $r \geqslant 0$, since the assumptions on $\{P_k\}$ are invariant under shift.

Hence the result follows.

3. Proof of Theorem B. Consider the probability vector $\boldsymbol{D}$ with respect to which $\{P_i\}$ is asymptotically homogeneous, so that

$$\lim_{k\to\infty}\boldsymbol{D}'P_k = \boldsymbol{D}'$$

and consider for fixed r, and t having the meaning of Lemma 2, the sequence of stochastic matrices

$$\{T_{r+k(t+1),t+1}\}, \qquad k \geqslant 0.$$

This sequence contains only Markov matrices, by Lemma 2; which are in fact uniformly Markov, for since by assumption $\min_{i,j}^{+} p_{ij}(k) \geqslant \delta > 0$, it follows that $\min_{i,j}^{+} t_{i,j}^{(r+i(t+1),t+1)} \geqslant \delta^{t+1}$, so that

$$\lambda(T_{r+k(t+1),t+1}) \geqslant \delta^{t+1}.$$

Moreover by Lemma 3,

$$\lim_{k\to\infty} \boldsymbol{D}' T_{r+k(t+1),t+1} = \boldsymbol{D}'$$

so that the sequence $\{T_{r+k(t+1),t+1}\}$, $k \geqslant 0$, is asymptotically homogeneous with respect to $\boldsymbol{D}$, and so by Lemma 5 is asymptotically stationary; in particular

$$\lim_{k\to\infty} \boldsymbol{D}' T_{r,t+1} T_{r+(t+1),t+1} \ldots T_{r+k(t+1),t+1} = \boldsymbol{D}'.$$

Now for fixed $m, 1 \leqslant m < (t+1)$

$$\lim_{k\to\infty} \boldsymbol{D}' P_{r+(k+1)(t+1)+1} P_{r+(k+1)(t+1)+2} \ldots P_{r+(k+1)(t+1)+m} = \lim_{k\to\infty} \boldsymbol{D}' T_{r+(k+1)(t+1),m} = \boldsymbol{D}'$$

by Lemma 3. Hence, easily for $0 \leqslant m < (t+1)$

$$\lim_{k\to\infty} \boldsymbol{D}' T_{r,t+1} T_{r+(t+1),t+1} \ldots T_{r+k(t+1),t+1} T_{(m,k)} = \boldsymbol{D}'$$

where

$$T_{(m,k)} = \begin{cases} T_{r+(k+1)(t+1),m} & \text{for } m > 0, \\ I & \text{for } m = 0. \end{cases}$$

Thus

$$\boldsymbol{D}' T_{r,(k+1)(t+1)+m} \to \boldsymbol{D}'$$

as $k \to \infty$ for each m, $0 \leqslant m < t+1$, and each fixed $r \geqslant 0$.
Thus

$$\lim_{k\to\infty} \boldsymbol{D}' T_{r,k} = \boldsymbol{D}'$$

so that the sequence $\{P_i\}$ is asymptotically stationary in terms of the matrix $\boldsymbol{D}$.

We thus have asymptotic stationarity; and weak ergodicity from Theorem A. Strong ergodicity is thus a simple consequence e.g. by the theorem of Koźniewska mentioned earlier; in fact $T_{r,k} \to \boldsymbol{1}\boldsymbol{D}'$ for all $r \geqslant 0$ as $k \to \infty$.

References

[1] С. Н. Бернстейн, *Теория вероятностей*, Moscow–Leningrad, 1946.

[2] J. Hajnal, *Weak ergodicity in non-homogeneous Markov chains*, Proc. Cambridge Philos. Soc. 54 (1958), pp. 233–246.

[3] I. Koźniewska, *Ergodicité et stationnarité des chaînes de Markoff variables à un nombre fini d'états possibles*, Colloq. Math. 9 (1962), pp. 333–346.

[4] A. A. Markov, *Investigation of an important case of dependent trials*, [in Russian]. Izv. Akad. Nauk SPB (6), 1 (1907), pp. 61–80.

[5] J. L. Mott, *Conditions for the ergodicity of non-homogeneous finite Markov chains*, Proc. Roy. Soc. Edinburgh. Section A., 64 (1957), pp. 369–380.

[6] A. Paz, *Graph-theoretic and algebraic characterizations of some Markov processes*, Israel J. Math. 1 (1963), pp. 169–180.

[7] — *Definite and quasi-definite sets of stochastic matrices*, Proc. Amer. Math. Soc. 16 (1965), 634–641.

[8] T. A. Sarymsakov, *On the ergodic principle for inhomogeneous Markov chains*, [in Russian]. Dokl. Akad. Nauk SSSR 90 (1953), pp. 25–28.

[9] — and H. A. Mustafin, *On an ergodic theorem for inhomogeneous Markov chains*, [in Russian]. Sredneaziatskii Gos. Universitet im. V. I. Lenina: Trudy. Novaiia seria, vyp. 74, Fiz. — mat., nauki, kniga 15, pp. 1–38. Izd. AN Uzbekskoi SSR, Tashkent, 1957.

[10] E. Seneta, *Non-negative Matrices*. London, 1973 (to appear).

[11] J. Wolfowitz, *Products of indecomposable, aperiodic, stochastic matrices*, Proc. Amer. Math. Soc. 14 (1963), pp. 733–737.

Received April 4, 1972 (502)