Stochastic Processes and their Applications 20 (1985) 307-314
North-Holland

AN ASYMPTOTIC REPRESENTATION FOR PRODUCTS OF RANDOM MATRICES

C.C. HEYDE

Department of Statistics, University of Melbourne, Parkville, Vic. 3052, Australia

Received 1 October 1984
Revised 3 December 1984

Let $\{X_k\}$ be a stationary ergodic sequence of nonnegative matrices. It is shown in this paper that, under mild additional conditions, the logarithm of the i,jth element of $X_t \cdots X_1$ is well approximated by a sum of t random variables from a stationary ergodic sequence. This representation is very useful for the study of limit behaviour of products of random matrices. An iterated logarithm result and an estimation result of use in the theory of demographic population projections are derived as corollaries.

products of random matrices * sums of random variables

1. Introduction and principal result

Let $(\Omega, \mathscr{F}, P)$ denote a probability space on which are defined an ergodic measure-preserving transformation T and a stationary sequence

$$\{X_k(\omega) = X_0(T^k\omega),\ k = 1, 2, \ldots,\ \omega \in \Omega\}$$

of $K \times K$ matrices with nonnegative elements. Also, for a matrix M denote by M_{ij} the element in the ith row and jth column. It is our object in this paper to show that, under mild additional conditions, $\log(X_tX_{t-1} \cdots X_1)_{ij}$ is well approximated by a sum of stationary random variables. This provides a convenient route for the investigation of limit properties of products of stationary nonnegative matrices since much is known about the limit behaviour of sums of stationary sequences.

We shall suppose that the matrices $\{X_k\}$ satisfy the three assumptions:

A1. There exists an integer n_0 such that any product $X_{j+n_0} \cdots X_{j+1}X_j$ of n_0+1 of the matrices has all its elements positive with probability one.

A2. For some constant C, $1 < C < \infty$, and each matrix X_k,

$$1 \leq M(X_k)/m(X_k) \leq C$$

with probability one where $M(X)$ and $m(X)$ are, respectively, the maximum and minimum *positive* elements of X.

A3. $E|\log M(X_1)| < \infty$.

Assumptions A1 and A2 are familiar within the theory of nonnegative matrices and define an 'ergodic set' in the sense of Hajnal [2]. A discussion of these

0304-4149/85/$3.30

assumptions is given in Heyde and Cohen [5]. Assumption A3 allows an ergodic theorem for the associated stationary process shortly to be defined and is used in Furstenberg and Kesten [1].

Under Assumptions A1–A3 we shall establish the following representation which makes essential use of ideas in [1]. This refers to a new probability space with probability measure P_1 such that P_1 agrees with P for any events defined by the process $\{X_k\}$ alone.

Representation. *For $t > n_0$ and any $1 \leqslant i, j \leqslant K$,*

$$|\log(X_t X_{t-1} \cdots X_1)_{ij} - \log(X_{t-1} \cdots X_1)_{ij} - y_t^{(i)}| = o(\nu^t)$$
$$\text{a.s. as } t \to \infty \text{ for some } 0 < \nu < 1, \tag{1}$$

where $\{y_t^{(i)}\}$ is a stationary ergodic sequence of random variables with $E|y_1^{(i)}| < \infty$ and

$$\gamma = \lim_{t\to\infty} t^{-1} \log(X_t \cdots X_1)_{ij} = Ey_1^{(i)} \quad \text{a.s.}$$

not depending on i. Furthermore,

$$\limsup_{t\to\infty} \left| \log(X_t \cdots X_1)_{ij} - \sum_{l=1}^{t} y_l^{(i)} \right| < \infty \quad \text{a.s.} \tag{2}$$

2. Proof of Representation

Write, for $t > s$,

$${}^tY^s = X_t \cdots X_{s+1} X_s. \tag{3}$$

Under Assumptions A1, A2, it follows from the theory of nonnegative matrices that, for $t > n_0$,

$$\left| \frac{({}^tY^1)_{jk}}{({}^tY^1)_{ji}} - \alpha_{ki} \right| \leqslant D\delta^t \quad \text{a.s.} \tag{4}$$

where α_{ki} is positive and random, D and δ are positive constants and δ (which depends on C of Assumption A2) satisfies $\delta < 1$.

To obtain (4) we make use of the various results from Chapter 3 of Seneta [7]. Write $({}^tY^1)_{ij} = h_{ij}^{(t)}$. Then, from Theorem 3.3 and the Corollary to Theorem 3.2 of [7] we have that

$$1 - A\delta^t \leqslant \min_{i,j,k,l} \frac{h_{ik}^{(t)} h_{jl}^{(t)}}{h_{il}^{(t)} h_{jk}^{(t)}} \leqslant 1 \tag{5}$$

for positive constants A and δ with $\delta < 1$. But, for fixed r, s, α,

$$\min_{i,j,k,l} \frac{h_{ik}^{(t)} h_{jl}^{(t)}}{h_{il}^{(t)} h_{jk}^{(t)}} \leqslant \frac{h_{\alpha r}^{(t)}}{h_{\alpha s}^{(t)}} \min_{l} \frac{h_{ls}^{(t)}}{h_{lr}^{(t)}} \leqslant 1. \tag{6}$$

Furthermore, from the proof of Lemma 3.4 of [7] (in transposed form since it refers to forward products)

$$\min_l \frac{h_{ls}^{(t)}}{h_{lr}^{(t)}} \uparrow W_{sr},$$

say, as $t\to\infty$, where $W_{sr}>0$ a.s. and hence, using (5) and (6),

$$\frac{h_{\alpha r}^{(t)}}{h_{\alpha s}^{(t)}} \geqslant \frac{1}{W_{sr}}(1-A\delta^t) \quad \text{a.s.} \tag{7}$$

Now, since r and s are arbitrary in (7) we also have

$$\frac{h_{\alpha s}^{(t)}}{h_{\alpha r}^{(t)}} \geqslant \frac{1}{W_{rs}}(1-A\delta^t) \quad \text{a.s.} \tag{8}$$

while

$$\frac{h_{\alpha s}^{(t)}}{h_{\alpha r}^{(t)}} \to W_{rs} = \frac{1}{W_{sr}} \quad \text{a.s.}$$

as $t\to\infty$, using the result of Exercise 3.4 of [7]. The result (4) then follows from (7) and (8) since $W_{rs}\leqslant (KC)^{2n_0}$ suing Lemma 2′ of Heyde and Cohen [5].

Note that, using (4) it suffices to establish (1) for $j=1$. Now if A, B, C are $K\times K$ matrices, then

$$\begin{aligned}(ABC)_{i1}/(BC)_{i1} &= \sum_{k=1}^{K} A_{ik}[(BC)_{k1}/(BC)_{i1}] \\ &= \sum_{k=1}^{K} A_{ik}[B_{k\cdot}/B_{i\cdot}] \\ &\quad + \left[\sum_{k=1}^{K} A_{ik}B_{k\cdot} \sum_{l=1}^{K} (B_{kl}/B_{k\cdot} - B_{il}/B_{i\cdot})C_{l1}\right]\Big/(BC)_{i1}\end{aligned} \tag{9}$$

where for a matrix M, $M_{i\cdot}=\sum_{j=1}^{k} M_{ij}$. Thus, putting $A=X_t$, $B={}^{t-1}Y^2$, $C=X_1$ in (9) and using Lemmas 2′ and 3′ of [5] together with the identity

$$\log(\alpha+\beta)=\log\alpha+\log(1+\beta/\alpha)=\log\alpha+\mathrm{O}(\beta/\alpha)$$

as $\beta/\alpha\to 0$,

$$\log\frac{({}^tY^1)_{i1}}{({}^{t-1}Y^1)_{i1}} = \log\sum_{k=1}^{K}(X_t)_{ik}\frac{({}^{t-1}Y^2)_{k\cdot}}{({}^{t-1}Y^2)_{i\cdot}} + \mathrm{o}(\nu^t) \tag{10}$$

for $1-(KC)^{-4n_0}<\nu<1$.

However, if $\{Z_i\}$ is the sequence of random matrices introduced in Lemma 1 of [1] which is stationary under P_1, we can take $A=X_t$, $B={}^{t-1}Y^2$, $C=Z_1$ in (9) and obtain

$$\log\frac{({}^tY^2Z_1)_{i1}}{({}^{t-1}Y^2Z_1)_{i1}} = \log\sum_{k=1}^{K}(X_t)_{ik}\frac{({}^{t-1}Y^2)_{k\cdot}}{({}^{t-1}Y^2)_{i\cdot}} + \mathrm{o}(\nu^t) \tag{11}$$

as $t\to\infty$. Thus, from (10) and (11),

$$\log\frac{({}^tY^1)_{i1}}{({}^{t-1}Y^1)_{i1}}=\log\frac{({}^tY^2Z_1)_{i1}}{({}^{t-1}Y^2Z_1)_{i1}}+o(\nu^t) \tag{12}$$

as $t\to\infty$. However, by construction,

$$\|X_{n+1}Z_n\|Z_{n+1}=X_{n+1}Z_n,\quad n=1,2,\ldots,$$

where for a matrix M, $\|M\|=\max_i\sum_{j=1}^K|M_{ij}|$. Then, we find that

$$({}^tY^2Z_1)_{i1}/({}^{t-1}Y^2Z_1)_{i1}=(X_tZ_{t-1})_{i1}/(Z_{t-1})_{i1},$$

$(Z_1)_{ij}>0$ a.s. for all i,j being ensured by A1, and (1) follows from (12) upon taking

$$y_t^{(i)}=\log[(X_tZ_{t-1})_{i1}/(Z_{t-1})_{i1}].$$

Furthermore, $E|y_t^{(i)}|<\infty$ in view of A3 since

$$\frac{({}^tY^2Z_1)_{i1}}{({}^{t-1}Y^2Z_1)_{i1}}=\frac{\sum_{k=1}^K\sum_{l=1}^K(X_t)_{ik}({}^{t-1}Y^2)_{kl}(Z_1)_{l1}}{\sum_{l=1}^K({}^{t-1}Y^2)_{il}(Z_1)_{l1}}\leqslant KM(X_t)M({}^{t-1}Y^2)/m({}^{t-1}Y^2)$$

$$\leqslant K(KC)^{2n_0}M(X_t)$$

for $t\geqslant n_0+1$ using Lemma 2′ of [5].

The result (2) is easily obtained from (1) since, writing $x_t=\log(X_t\cdots X_1)_{ij}$, we have for $t>n_0$,

$$\left|x_t-\sum_{k=1}^t y_k\right|=\left|x_{n_0}+\sum_{k=n_0+1}^t(x_k-x_{k-1}-y_k)-\sum_{k=1}^{n_0}y_k\right|$$

$$\leqslant|x_{n_0}|+\left|\sum_{k=1}^{n_0}y_k\right|+\sum_{k=n_0+1}^t|x_k-x_{k-1}-y_k|.$$

Finally, to identify Ey_1 as γ we first note that (2) gives

$$\lim_{t\to\infty}t^{-1}\left|x_t-\sum_{k=1}^t y_k\right|=0\quad\text{a.s.}$$

and, since $E|y_1|<\infty$, the classical ergodic theorem gives

$$t^{-1}\sum_{k=1}^t y_k\xrightarrow{\text{a.s.}}Ey_1.$$

Now write $W_k={}^{kn_0}Y^{(k-1)n_0+1}$, $k\geqslant1$, and note that each W_k has all its elements positive in view of A1. The ergodic theorem for products of positive matrices of Furstenberg and Kesten ([1, p. 462]) gives

$$\lim_{t\to\infty}t^{-1}x_{tn_0}=n_0\gamma\quad\text{a.s.,} \tag{13}$$

say, a constant not depending on i, j. But, for $0 \leq k < n_0$ and $t \geq 1$,

$$(KC)^{2n_0}({}^{n_0t}Y^1)_{ij} \prod_{l=n_0t+1}^{n_0t+k} M(X_l) \geq ({}^{n_0t+k}Y^1)_{ij} = \sum_{l=1}^{K} ({}^{n_0t+k}Y^{n_0t+1})_{il}({}^{n_0t}Y^1)_{lj}$$

$$\geq ({}^{n_0t}Y^1)_{ij} \prod_{l=n_0t+1}^{n_0t+k} m(X_l) \geq C^{-k}({}^{n_0t}Y^1)_{ij} \prod_{l=n_0t+1}^{n_0t+k} M(X_l) \tag{14}$$

using Lemma 2′ of [5]. From (14) we deduce that

$$|\log({}^{n_0t}Y^1)_{ij} - \log({}^{n_0t+k}Y^1)_{ij}| \leq \log(KC)^{2n_0} + \sum_{l=n_0t+1}^{n_0t+k} M(X_l)$$

and $\lim_{t\to\infty} t^{-1}x_t = \gamma$ a.s. follows using (13). This completes the proof of the representation.

3. Corollaries

Various strong law and central limit type results for products of random matrices have been available for more than twenty years (e.g. Furstenberg and Kesten [1]). Iterated logarithm results, however, have not been available and use of the representation immediately leads to results of this type as indicated in the following corollary.

Corollary 1 (Law of the iterated logarithm). *Suppose that the stationary sequence* $\{y_t^{(i)}\}$ *satisfies* (1). *If the ergodic measure-preserving transformation T is such that for some* $\sigma > 0$,

$$\left\{\xi_t = (2\sigma^2 t \log\log t)^{-1/2}\left(\sum_{l=1}^{t} y_l^{(i)} - t\gamma\right)\right\}$$

has its a.s. limit points confined to $[-1, 1]$ *with* $\limsup_{t\to\infty} \xi_t = +1$ *a.s.*, $\liminf_{t\to\infty} \xi_t = -1$ *a.s., then the same result holds for*

$$\{\eta_t = (2\sigma^2 t \log\log t)^{-1/2}(\log(X_t \cdots X_1)_{ij} - t\gamma)\}$$

for any $1 \leq i, j \leq K$.

A wide variety of asymptotic independence conditions on σ-fields generated by T are sufficient to ensure that the above law of the iterated logarithm holds for $\{\xi_t\}$. For example, let $\{\mathcal{M}_a^b, -\infty \leq a \leq b \leq \infty\}$ be the family of sub σ-fields of $\mathcal{F}$ satisfying the conditions

(i) if $a \leq c \leq d \leq b$, then $\mathcal{M}_c^d \subseteq \mathcal{M}_a^b$,

(ii) for all $a \leq b$, $T^{-1}\mathcal{M}_a^b = \mathcal{M}_{a+1}^{b+1}$,

and write

$$\phi(n) = \sup_{k\geq 0}\{|P(B|A) - P(B)|;\ A \in \mathcal{M}_0^k,\ B \in \mathcal{M}_{k+n}^{\infty},\ P(A) > 0\}.$$

Then, the conditions $\sum_{n=1}^{\infty}[\phi(n)]^{1/2}<\infty$ and $E|\log M(X_1)|^2<\infty$ suffice (see e.g. Hall and Heyde [3, Corollary 5.4]). It is easily checked from (10) and (11) that the condition $E|\log M(X_1)|^2<\infty$ ensures $E(y_t^{(i)})^2<\infty$, each $1\le i\le K$.

The representation is of particular use in facilitating statistical inference for population projections based on products of random matrices. For example, in the paper Heyde and Cohen [5] population projections based on the model

$$Y_{t+1}=X_{t+1}Y_t \tag{15}$$

are extensively studied, where Y_t is the column vector of the number of individuals in each of K age classes at time t and X_t is a random matrix of vital rates. The matrices X_t satisfy the conditions of the present paper. Inference for population projections is based on a central limit result of the form

$$(t\sigma^2)^{-1/2}\{\log(X_t\cdots X_1)_{ij}-t\gamma\}\overset{\mathrm{d}}{\to}N(0,1) \tag{16}$$

for some $0<\sigma<\infty$ (see Theorem 1 of [5]) and efficient estimation of σ^2 is crucial. A crude procedure is developed in [5] but the representation result of the present paper offers the prospect of the use of standard time series methods, for example those based on smoothed versions of the periodogram of the approximating stationary process.

With $\{Y_t,\ t\ge 0\}$ defined according to (15), let $Z_t=e_i'Y_t$ where e_i is the vector with 1 in the ith component and zero elsewhere. We shall take $\{Z_t,\ t\ge 0\}$ as our data and note that Theorem 2 of [5] gives sufficient conditions (the same as for (16)) to ensure that

$$(t\sigma^2)^{-1/2}(\log Z_t-t\gamma)\overset{\mathrm{d}}{\to}N(0,1) \tag{17}$$

as $t\to\infty$ for some $\sigma>0$.

Write

$$\hat\gamma_t=t^{-1}\log Z_t,\qquad w_t=\log(Z_t/Z_{t-1})$$

and for $0\le n\le t-1$,

$$\hat c_t(n)=(t-n)^{-1}\sum_{l=1}^{t-n}(w_l-\hat\gamma_t)(w_{l+n}-\hat\gamma_t),$$

$$c_t(n)=(t-n)^{-1}\sum_{l=1}^{t-n}(y_l^{(i)}-\gamma)(y_{l+n}^{(i)}-\gamma),$$

$$\hat c_t(-n)=\hat c_t(n),\qquad c_t(-n)=c_t(n),$$

where $\{y_t^{(i)}\}$ satisfies (1). The following corollary is indicative of the way in which the representation can be used.

Corollary 2. *Suppose that* (17) *holds. If* $s\to\infty$, $t\to\infty$ *in such a way that* $st^{-1/2}\to 0$, *then*

$$\sum_{n=-s}^{s}\left(1-\frac{|n|}{s}\right)(\hat c_t(n)-c_t(n))\,\mathrm{e}^{-\mathrm{i}n\lambda}\overset{\mathrm{P}}{\to}0.$$

It should be noted that the smoothed correlogram

$$\hat{f}^B(\lambda) = \frac{1}{2\pi} \sum_{n=-s}^{s} \left(1 - \frac{|n|}{s}\right) c_t(n)\, e^{-in\lambda}$$

is Bartlett's estimator for the spectral density in the case where this exists, of the stationary process $\{y_t^{(i)} - \gamma\}$ (e.g. Hannan [4, p. 278]) and that it converges in probability to the true spectral density under relatively mild additional conditions (e.g. [4, Theorem 9, p. 280]). Further, under these conditions we also have

$$\lim_{t\to\infty} t^{-1} E\left(\sum_{l=1}^{t} (y_l^{(i)} - \gamma)\right)^2 = \sigma^2$$

with $0 < \sigma < \infty$ and

$$2\pi \hat{f}^B(0) \xrightarrow{\text{P}} \sigma^2$$

as $s, t \to \infty$.

Proof of Corollary 2. We use (4) and writing $Y_0' = (Y_1^{(0)}, \ldots, Y_K^{(0)})$, we have

$$\frac{e_i' Y_t}{({}^tY^1)_{i1}} = \sum_{j=1}^{K} \frac{({}^tY^1)_{ij}}{({}^tY^1)_{i1}} Y_j^{(0)} = \sum_{j=1}^{K} \alpha_{j1} Y_j^{(0)} + \dot{R}_t,$$

say, where $R_t = o(\delta^t)$ as $t \to \infty$ and

$$L = \sum_{j=1}^{K} \alpha_{j1} \beta_{j1} Y_j^{(0)} > 0 \quad \text{a.s.}$$

Thus,

$$|\log Z_t - \log({}^tY^1)_{i1} - \log L| = o(\delta^t) \quad \text{a.s.}$$

and, using (1),

$$\begin{aligned} |w_t - y_t^{(i)}| &= |(\log Z_t - \log({}^tY^1)_{i1} - \log L) - (\log Z_{t-1} - \log({}^{t-1}Y^1)_{i1} - \log L) \\ &\quad + (\log({}^tY^1)_{i1} - \log({}^{t-1}Y^1)_{i1} - y_t^{(i)})| \\ &= o(\delta_1^t) \quad \text{a.s.} \end{aligned} \tag{18}$$

where $\delta_1 = \max(\nu, \delta)$.

Now write $R_t = y_j^{(i)} - w_t$. We have, after some algebra,

$$\begin{aligned} \hat{c}_t(n) - c_t(n) &= (\gamma - \hat{\gamma}_t)^2 - (\gamma - \hat{\gamma}_t)(t-n)^{-1} \sum_{l=1}^{t-n} (R_l + R_{l+n}) \\ &\quad + (\gamma - \hat{\gamma}_t)(t-n)^{-1} \sum_{l=1}^{t-n} [(y_l^{(i)} - \gamma) + (y_{l+n}^{(i)} - \gamma)] \\ &\quad - (t-n)^{-1} \sum_{l=1}^{t-n} [R_{l+n}(y_l^{(i)} - \gamma) + R_l(y_{l+n}^{(i)} - \gamma) - R_l R_{l+n}]. \end{aligned}$$

Thus, with the aid of Schwarz' inequality,

$$\sup_{0\le n\le s} |\hat{c}_t(n) - c_t(n)| \le (\gamma - \hat{\gamma}_t)^2 + 2|\gamma - \hat{\gamma}_t|(t-s)^{-1} \sum_{l=1}^{t} |R_l|$$

$$+2|\gamma - \hat{\gamma}_t|(t-s)^{-1} \sum_{l=1}^{t} |y_l^{(i)} - \gamma|$$

$$+2(t-s)^{-1}\left(\sum_{l=1}^{t} (y_l^{(i)} - \gamma)^2\right)^{1/2}\left(\sum_{l=1}^{t} R_l^2\right)^{1/2}$$

$$+(t-s)^{-1} \sum_{l=1}^{t} R_l^2 \quad \text{a.s.}$$

and, using (19) and the ergodic theorem,

$$\left|\sum_{n=-s}^{s} \left(1 - \frac{|n|}{s}\right)(\hat{c}_t(n) - c_t(n))\,\mathrm{e}^{-in\lambda}\right| \le 2 \sum_{n=0}^{s} \left(1 - \frac{n}{s}\right)|\hat{c}_t(n) - c_t(n)|$$

$$\le s(\gamma - \hat{\gamma}_t)^2 + \mathrm{O}(s|\gamma - \hat{\gamma}_t|t^{-1}) + \mathrm{O}(|\gamma - \hat{\gamma}_t|s) + \mathrm{O}(st^{-1/2}) + \mathrm{O}(st^{-1}) \xrightarrow{\mathrm{p}} 0$$

since $st^{-1/2} \to 0$ and $s|\gamma - \hat{\gamma}_t| \xrightarrow{\mathrm{p}} 0$ by assumption. This completes the proof.

It may be expected that Corollary 2 will hold for values of s $(= \mathrm{o}(t))$ which are more rapidly increasing but it is not our purpose to provide a detailed exploration of a result which is incidental for this paper. An improvement might be sought using Baxter's maximal inequality for renewal sequence weighted sums of stationary random variables (e.g. Stout [8, Theorem 4.2.8, p. 247]). The sequence $\{\delta_1^l\}$ is a Kaluza sequence and hence a renewal sequence (e.g. Kingman [6, Theorem 1.8, p. 17]).

References

[1] H. Furstenberg and H. Kesten, Products of random matrices, Ann. Math. Statist. 31 (1960) 457–469.

[2] J. Hajnal, On products of non-negative matrices, Math. Proc. Cambridge Philos. Soc. 79 (1976) 521–530.

[3] P. Hall and C.C. Heyde, Martingale Limit Theory and its Application (Academic Press, New York, 1980).

[4] E.J. Hannan, Multiple Time Series (Wiley, New York, 1970).

[5] C.C. Heyde and J.E. Cohen, Confidence intervals for demographic projections based on products of random matrices, Theor. Popul. Biol., to appear (1985).

[6] J.F.C. Kingman, Regenerative Phenomena (Wiley, London, 1972).

[7] E. Seneta, Non-negative Matrices and Markov Chains (Springer-Verlag, New York, 1981).

[8] W.F. Stout, Almost Sure Convergence (Academic Press, New York, 1974).