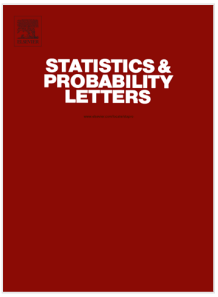
A generalization of a result concerning the asymptotic behavior of finite Markov chains

Alina Nicolaie



A generalization of a result concerning the asymptotic behavior of finite Markov chains

Alina Nicolaie

PII: S0167-7152(08)00322-2
DOI: 10.1016/j.spl.2008.07.003
Reference: STAPRO 5135

To appear in: *Statistics and Probability Letters*

Received date: 28 February 2008
Revised date: 4 July 2008
Accepted date: 9 July 2008

Please cite this article as: Nicolaie, A., A generalization of a result concerning the asymptotic behavior of finite Markov chains. *Statistics and Probability Letters* (2008), doi:10.1016/j.spl.2008.07.003

A GENERALIZATION OF A RESULT CONCERNING THE ASYMPTOTIC BEHAVIOR OF FINITE MARKOV CHAINS

ALINA NICOLAIE

Abstract. In this paper we give, in a more general context than previous studies, sufficient conditions for weak, strong and C-strong ergodicity of a finite state nonhomogeneous Markov chain in terms of similar properties of a certain Markov chain of smaller size.
AMS Subject Classification: 60J10.
Key words: weak ergodicity; strong ergodicity; C-strong ergodicity; finite Markov chain.

1. PRELIMINARIES

Consider a finite homogeneous Markov chain with state space $S = \{1, ..., r\}$ and transition matrix P. We shall refer to it as the finite Markov chain P. Assume that the stochastic matrix P has $p \geq 1$ irreducible aperiodic closed classes or equivalently, ergodic classes, and, perhaps transient states, so that it has the form

$$P = \begin{pmatrix} S_{(1)} & 0 & ... & 0 & 0 \\ 0 & S_{(2)} & ... & 0 & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & ... & S_{(p)} & 0 \\ L_{(1)} & L_{(2)} & ... & L_{(p)} & T \end{pmatrix}, \tag{1.1}$$

where $S_{(i)}$ are $r_i \times r_i$ transition matrices, $1 \leq i \leq p$, associated with the p irreducible aperiodic closed classes, T concerns the transitions of the chain as long as it stays in the $r - \sum_{t=1}^{p} r_t$ transient states and the $L_{(i)}$ concern transitions from the transient states into the ergodic sets S_i, $1 \leq i \leq p$.

Definition 1.1. A probability distribution $\mu = (\mu_1, ..., \mu_r)$ is said to be invariant with respect to an $r \times r$ stochastic matrix P if $\mu P = \mu$.

We shall need the following result

Department of Mathematical Analysis and Probability, Transilvania University of Braşov, Romania.
Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, The Netherlands.
E-mail: M.A.Nicolaie@lumc.nl.

THEOREM 1.2. *Consider a finite homogeneous Markov chain with state space S and transition matrix P of the form* (1.1). *Then*

$$\lim_{n\to\infty} P^n = \begin{pmatrix} \Gamma_1 & 0 & ... & 0 & 0 \\ 0 & \Gamma_2 & ... & 0 & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & ... & \Gamma_p & 0 \\ \Omega_1 & \Omega_2 & ... & \Omega_p & 0 \end{pmatrix}, \tag{1.2}$$

where Γ_i is a strictly positive $r_i \times r_i$ matrix, $1 \leq i \leq p$; each row of the matrix Γ_i is the invariant probability vector $\mu^{(i)} = (\mu_1^{(i)}, ..., \mu_{r_i}^{(i)})$ with respect to $S_{(i)}$, $1 \leq i \leq p$, and

$$\Omega_i = \begin{pmatrix} \mu_1^{(i)} a_{r_1+r_2+...+r_p+1,i} & ... & \mu_{r_i}^{(i)} a_{r_1+r_2+...+r_p+1,i} \\ ... & ... & ... \\ \mu_1^{(i)} a_{r,i} & ... & \mu_{r_i}^{(i)} a_{r,i} \end{pmatrix}$$

is an $(r - \sum_{t=1}^{p} r_t) \times r_i$ matrix, where a_{ji}= probability that the chain will enter and thus, will be absorbed in S_i given that the initial state is j, $\sum_{t=0}^{p} r_t \leq j \leq r$, $1 \leq i \leq p$ [with convention $r_0 = 1$].

Proof. For the form of Γ_i, $1 \leq i \leq p$, see, e.g., Iosifescu (1980, p. 123) and for Ω_i, $1 \leq i \leq p$, see, e.g., Karlin et Taylor (1975, p. 91). □

Remark 1.3. Clearly, $a_{ji} \geq 0$, $\sum_{t=0}^{p} r_t \leq j \leq r$, $1 \leq i \leq p$ and

$$\sum_{i=1}^{p} a_{ji} = 1, \sum_{t=0}^{p} r_t \leq j \leq r. \tag{1.3}$$

A vector $x \in \mathbf{C}^n$ will be understood as a row vector and $x^{'}$ denotes the transpose of x. Set $\mathbf{e} = (1, 1, ..., 1)$ and let $(e_i)_{1\leq i\leq r}$ be the canonical basis of the linear space $\mathbf{R}^r$.

THEOREM 1.4 (Nicolaie (2008a)). *Let $A = -I_r + P$ with P of the form* (1.1). *Then there exists a nonsingular complex $r \times r$ matrix Q such that*

$$A = QJQ^{-1}, \tag{1.4}$$

where J is a Jordan $r \times r$ matrix, and Q reads as

$$Q = \begin{array}{c} \\ r_1 \\ \\ r_2 \\ \\ \vdots \\ \\ r_p \\ \\ r - \sum_{t=0}^{p} r_t \end{array} \begin{pmatrix} 1 & 0 & ... & 0 & ... \\ ... & ... & ... & ... & ... \\ 1 & 0 & ... & 0 & ... \\ 0 & 1 & ... & 0 & ... \\ ... & ... & ... & ... & ... \\ 0 & 1 & ... & 0 & ... \\ 0 & 0 & ... & 0 & ... \\ ... & ... & ... & ... & ... \\ 0 & 0 & ... & 0 & ... \\ 0 & 0 & ... & 1 & ... \\ ... & ... & ... & ... & ... \\ 0 & 0 & ... & 1 & ... \\ a_{r_1+r_2+...+r_p+1,1} & a_{r_1+r_2+...+r_p+1,2} & ... & a_{r_1+r_2+...+r_p+1,p} & ... \\ ... & ... & ... & ... & ... \\ a_{r,1} & a_{r,2} & ... & a_{rp} & ... \end{pmatrix},$$

where the first column contains 1 *in the first* r_1 *rows, the next* $p-1$ *columns contain* 1 *in the* $r_{i-1}+1, \ldots, r_i$ *rows,* $2 \le i \le p$*, and the last* $r-p$ *columns comprise complex numbers. For* a_{ji}*,* $\sum_{t=0}^{p} r_t \le j \le r$*,* $1 \le i \le p$*, we have the meaning given in Theorem* 1.2*. The inverse* Q^{-1} *has the form*

$$Q^{-1} = \begin{pmatrix} \mu_1^{(1)} & ... & \mu_{r_1}^{(1)} & 0 & ... & 0 & 0 & ... & 0 & 0 & ... & 0 \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ 0 & ... & 0 & 0 & ... & 0 & \mu_1^{(p)} & ... & \mu_{r_p}^{(p)} & 0 & ... & 0 \\ q_{p+1,1} & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & q_{p+1,r} \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ q_{r,1} & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & q_{rr} \end{pmatrix},$$

where $\mu^{(i)} = (\mu_1^{(i)}, \ldots, \mu_{r_i}^{(i)})$ *is the invariant probability vector with respect to* $S_{(i)}$*,* $1 \le i \le p$*, and the last* $r-p$ *rows comprise complex numbers.*

Proof. See Nicolaie (2008a) and also Gidas (1985). □

Remark 1.5 (Nicolaie (2008a)). (i) We shall need some spectral properties of A, where $A = -I_p + P$ with P of the form (1.1). We have

$$\lambda_1 = 0 \tag{1.5}$$

is an eigenvalue of A whose algebraic multiplicity is equal to its geometric multiplicity and equal to p. All other distinct eigenvalues $\lambda_2, ..., \lambda_{l+s}$ of A satisfy

$$|\lambda_i + 1| < 1,\ 2 \le i \le l+s. \tag{1.6}$$

(ii) From (1.5) (see Horn et Johnson (1985, pp. 129-131)) we have

$$J = \begin{pmatrix} J_1 & 0 & ... & ... & ... & ... & 0 \\ 0 & J_2 & ... & ... & ... & ... & 0 \\ ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & ... & J_l & 0 & ... & ... \\ 0 & 0 & ... & 0 & J_{l+1} & ... & ... \\ ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & ... & 0 & 0 & ... & J_{l+s} \end{pmatrix},$$

where $J_1 = \mathbf{0}_{p\times p}$, J_k is a diagonal $m_k \times m_k$ matrix with entries the eigenvalues λ_k whose algebraic and geometric multiplicities are identical, $2 \leq k \leq l$, and

$$J_{l+i} = \begin{pmatrix} \lambda_{l+i} & \varepsilon_1^{(i)} & 0 & ... & ... & 0 \\ 0 & \lambda_{l+i} & \varepsilon_2^{(i)} & ... & ... & 0 \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & ... & ... & \lambda_{l+i} & \varepsilon_{m_{l+i}-1}^{(i)} \\ 0 & 0 & ... & ... & 0 & \lambda_{l+i} \end{pmatrix}$$

are $m_{l+i} \times m_{l+i}$ matrices corresponding to eigenvalues whose geometric multiplicities are smaller than their algebraic multiplicities and $\varepsilon_t^{(i)} \in \{0,1\}$, $1 \leq t \leq m_{l+i}-1$, $1 \leq i \leq s$. Clearly, $p + m_2 + ... + m_{l+s} = r$.

If $A = (A_{ij})$ is an $m \times n$ matrix, then for $M \subseteq \{1,...,m\}$, $N \subseteq \{1,...,n\}$, $M, N \neq \emptyset$, define $A_{M\times N} = (A_{ij})_{i\in M, j\in N}$. Define, also, the matrix norm (see, e.g., Horn et Johnson (1985, p. 295)) $|||A|||_\infty = \max_{1\leq i\leq m} \sum_{j=1}^n |A_{ij}|$. Note that such a norm tends to zero if and only if all elements of A tend to zero.

Let $(P_n)_{n\geq 1}$ be a sequence of stochastic $r \times r$ matrices. Set $S = \{1, 2, ..., r\}$. For all integers $m \geq 0$, $n > m$, define $P_{m,n} = P_{m+1}P_{m+2}...P_n = ((P_{m,n})_{ij})_{i,j\in S}$. In the context of finite Markov chains, the matrix $P_{m,n}$ has a very simple probabilistic interpretation, namely being the $n - m$ step transition matrix starting at m.

Definition 1.6 (see, e.g., Isaacson et Madsen (1976, p. 144)). Let P be a stochastic $r \times r$ matrix. The ergodic coefficient of P, denoted by $\delta(P)$, is defined by $\delta(P) = 1 - \min_{1\leq i,k\leq r} \sum_{j=1}^r \min(P_{ij}, P_{kj})$.

THEOREM 1.7 (see, e.g., Isaacson et Madsen (1976, p. 147)). *Let R be a real $m \times n$ matrix with $R\mathbf{e}' = \mathbf{0}$ and P be a stochastic $n \times p$ matrix. Then $|||RP|||_\infty \leq |||R|||_\infty \delta(P)$.*

Proof. See, e.g., Isaacson et Madsen (1976, p. 147). □

Definition 1.8 (see, e.g., Iosifescu (1980, p. 217)). A sequence of stochastic $r \times r$ matrices $(P_n)_{n\geq 1}$ is said to be weakly ergodic if $\forall m \geq 0$, $\forall i, j, k \in S$, $\lim_{n\to\infty}[(P_{m,n})_{ik} - (P_{m,n})_{jk}] = 0$.

A stochastic matrix whose rows are identical is said to be stable.

THEOREM 1.9. *Let $(P_n)_{n\geq 1}$ be a sequence of stochastic $r\times r$ matrices. Then the following statements are equivalent*:

(i) *the sequence $(P_n)_{n\geq 1}$ is weakly ergodic*;

(ii) *there exist stable stochastic $r\times r$ matrices $\Pi_{m,n}$, $m\geq 0$, $n\geq 1$ such that* $\lim_{n\to\infty}(P_{m,n}-\Pi_{m,n})=\mathbf{0}$, $\forall m\geq 0$;

(iii) $\lim_{n\to\infty}\delta(P_{m,n})=0$, $\forall m\geq 0$.

Proof. For (i)⇔(ii) see, e.g., Iosifescu (1980, p. 218) and for (i)⇔(iii) see, e.g., Isaacson et Madsen (1976, p. 149). □

Definition 1.10 (see, e.g., Iosifescu (1980, p. 223)). A sequence of stochastic $r\times r$ matrices $(P_n)_{n\geq 1}$ is said to be strongly ergodic if $\forall m\geq 0$, $\forall i,j\in S$, the limit $\lim_{n\to\infty}(P_{m,n})_{ij}=:(\pi_m)_j$ exists and does not depend on i.

Remark 1.11 (see, e.g., Iosifescu (1980, p. 223)). It is easy to prove that if $\lim_{n\to\infty}(P_{m,n})_{ij}$ exists as stated in Definition 1.10, then the $(\pi_m)_j$ are also independent of $m\geq 0$. Therefore, a sequence of stochastic matrices $(P_n)_{n\geq 1}$ is strongly ergodic if and only if there exists a stable stochastic matrix Π such that $\lim_{n\to\infty}(P_{m,n}-\Pi)=0$, $\forall m\geq 0$.

Definition 1.12 (Isaacson et Madsen (1976)). A sequence of stochastic matrices $(P_n)_{n\geq 1}$ is said to be C-strongly ergodic if there exists a stable stochastic matrix Π such that $\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}P_{0,k}=\Pi$.

THEOREM 1.13 (Isaacson et Madsen (1976)). *Let $(P_n)_{n\geq 1}$ be a sequence of stochastic matrices such that, for some fixed d, $\lim_{n\to\infty}P_{0,nd+t}=W_t$, $1\leq t\leq d$. Then the limit $\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}P_{0,k}$ exists and equals $W=\frac{1}{d}\big[W_1+...+W_d\big]$.*

Proof. See Isaacson et Madsen (1976, p. 184). □

Remark 1.14. Note that in Theorem 1.13 the limiting matrices W_t of the Cesaro averages are not necessarily stable matrices. But if they are stable, this theorem provides sufficient conditions for the C-strong ergodicity of the sequence $(P_n)_{n\geq 1}$.

The next two results will only be used in Section 2 of this paper.

PROPOSITION 1.15 (see, e.g., Isaacson et Madsen (1976, p. 29)). *Let $(a_{nk})_{n,k\geq 1}$ be a doubly indexed sequence of real numbers such that $\lim_{n\to\infty}a_{nk}=a_k$ exists, for all $k\geq 1$. If there exists a sequence of nonnegative numbers $(b_k)_{k\geq 1}$ such that $|a_{nk}|\leq b_k$,* for all *$n\geq 1$, $k\geq 1$ and $\sum_{k=1}^{\infty}b_k<\infty$, then $\lim_{n\to\infty}\sum_{k=1}^{\infty}a_{nk}=\sum_{k=1}^{\infty}a_k$.*

Proof. See, e.g., Isaacson et Madsen (1976, p. 29). □

PROPOSITION 1.16 (see, e.g., Isaacson et Madsen (1976, pp. 34-36)). *Let* $(a_n)_{n\geq 0}$ *be a sequence of real numbers convergent to* 0 *and* $\sum_{n=0}^{\infty} b_n$ *an absolute convergent series. Then* $\lim_{n\to\infty}\sum_{i=0}^{n} a_i b_{n-i} = 0$.

Proof. See, e.g., Isaacson et Madsen (1976, pp. 34-36). □

2. SOME RESULTS ON RECURRENCE RELATIONS

In this section we give some results related to sequences defined by recurrence relations. First, we shall need the following

PROPOSITION 2.1. *Let* $(a_n)_{n\geq 0}$ *and* $(b_n)_{n\geq 0}$ *be two sequences of nonnegative real numbers with* $\lim_{n\to\infty} b_n = 0$. *If*

$$a_{s+d} \leq \alpha a_s + b_s,\ s \geq 0, \tag{2.1}$$

where d *stands for a natural number,* $d \geq 1$ *and* $\alpha \in [0,1)$, *then*

$$\lim_{n\to\infty} a_n = 0. \tag{2.2}$$

Proof. Applying the recurrence inequality (2.1) successively we obtain

$$a_{td+l} \leq \alpha^t a_l + \sum_{k=0}^{t-1} \alpha^{t-k-1} b_{kd+l},\ t \geq 0,\ 0 \leq l \leq d-1. \tag{2.3}$$

Since $\alpha \in [0,1)$, we have $\lim_{t\to\infty} \alpha^t = 0$. Moreover, the series $\sum_{t=0}^{\infty} \alpha^t < \infty$ (the geometrical series). Also, since $\lim_{n\to\infty} b_n = 0$, all the subsequences of $(b_n)_{n\geq 0}$ are convergent with limit zero; in particular, $\lim_{t\to\infty} b_{td+l} = 0$, $0 \leq l \leq d-1$. Applying Proposition 1.16 it follows that $\lim_{t\to\infty}\sum_{k=0}^{t-1} \alpha^{t-k-1} b_{kd+l} = 0$. Next, letting $t \to \infty$ in (2.3), it follows $a_{td+l} \to 0$, $0 \leq l \leq d-1$. The conclusion follows. □

The next result is a generalization of Proposition 2.2 from Nicolaie (2008b).

PROPOSITION 2.2. *Let* $(X_n)_{n\geq 1}$ *and* $(R_n)_{n\geq 1}$ *be two sequences of real vectors, each vector* X_n *and* R_n *having* p *components,* $R_n\mathbf{e}' = \mathbf{0}$, $n \geq 1$ *and* $\sum_{n=1}^{\infty} |||R_n|||_\infty < \infty$. *Let* $(C_n)_{n\geq 1}$ *be a sequence of stochastic* $p \times p$ *matrices and set* $\widetilde{C}_n^{(l)} = C_{(n-1)d+l+1}$, *for some fixed* d, $n \geq 1$, $0 \leq l \leq d-1$. *Suppose that*

$$X_{s+d} = X_s C_s + R_s,\ s \geq 1, \tag{2.4}$$

Then the following statements hold:

(i) *If* $X_n\mathbf{e}' = 0$, $n \geq 1$, *and* $(\widetilde{C}_n^{(l)})_{n\geq 1}$ *is weakly ergodic,* $0 \leq l \leq d-1$, *then* $\lim_{n\to\infty} X_n = \mathbf{0}$.

(ii) *If* $X_n\mathbf{e}'=1$, $n\geq 1$ *and* $(\widetilde{C}_n^{(l)})_{n\geq 1}$ *is strongly ergodic with* $\lim_{n\to\infty}\widetilde{C}_{m,n}^{(l)}=\Pi=\mathbf{e}'\cdot\pi$, $m\geq 0$, $0\leq l\leq d-1$, *then* $\lim_{n\to\infty}X_n=\pi$.

(iii) *If* $X_n\mathbf{e}'=1$, $n\geq 1$ *and* $(\widetilde{C}_n^{(l)})_{n\geq 1}$ *is strongly ergodic with* $\lim_{n\to\infty}\widetilde{C}_{m,n}^{(l)}=\Pi_{(l+1)}=\mathbf{e}'\cdot\pi^{(l+1)}$, $m\geq 0$, $0\leq l\leq d-1$, *then* $\lim_{n\to\infty}X_{nd+l+1}=\pi^{(l+1)}$, $0\leq l\leq d-1$.

Proof. Applying the recurrence relation (2.4) successively we obtain

$$X_{(t+1)d+l+1}=X_{l+1}\widetilde{C}_{0,t+1}^{(l)}+\big[R_{td+l+1}+\sum_{k=0}^{t-1}R_{kd+l+1}\widetilde{C}_{k+1,t+1}^{(l)}\big],t\geq 1,0\leq l\leq d-1. \tag{2.5}$$

First, assuming that $(\widetilde{C}_n^{(l)})_{n\geq 1}$ is weakly ergodic, $0\leq l\leq d-1$, we shall prove that

$$\lim_{t\to\infty}\sum_{k=0}^{t-1}R_{kd+l+1}\widetilde{C}_{k+1,t+1}^{(l)}=\mathbf{0},0\leq l\leq d-1. \tag{2.6}$$

We have

$$|||\sum_{k=0}^{t-1}R_{kd+l+1}\widetilde{C}_{k+1,t+1}^{(l)}|||_\infty\leq\sum_{k=0}^{t-1}|||R_{kd+l+1}\widetilde{C}_{k+1,t+1}^{(l)}|||_\infty\leq$$

(using Theorem 1.7)

$$\leq\sum_{k=0}^{t-1}|||R_{kd+l+1}|||_\infty\delta(\widetilde{C}_{k+1,t+1}^{(l)}).$$

Next, choose $a_{tk}^{(l)}=|||R_{kd+l+1}|||_\infty\delta(\widetilde{C}_{k+1,t+1}^{(l)})$, $t,k\geq 1$ (take $a_{tk}^{(l)}=0$ if $k>t$), $0\leq l\leq d-1$. Then, by Theorem 1.9 (iii), it follows that $\lim_{t\to\infty}a_{tk}^{(l)}=0$, $k\geq 1$, $0\leq l\leq d-1$. Moreover, $|a_{tk}^{(l)}|\leq b_k^{(l)}:=|||R_{kd+l+1}|||_\infty$, $t\geq 1$, since $\delta(\widetilde{C}_{k,t}^{(l)})\leq 1$, $0\leq k<t$, $0\leq l\leq d-1$. The conditions of Proposition 1.15 are fulfilled, so $\lim_{t\to\infty}\sum_{k=1}^{t}a_{tk}^{(l)}=\lim_{t\to\infty}\sum_{k=1}^{\infty}a_{tk}^{(l)}=0$, $0\leq l\leq d-1$, which means (2.6).

(i) Let $l\in\{0,1,...,d-1\}$. By Theorem 1.9 (ii), it follows that there exists a sequence of stable stochastic matrices $\Pi_{m,t}^{(l)}$ such that $\lim_{t\to\infty}\big(\widetilde{C}_{m,t}^{(l)}-\Pi_{m,t}^{(l)}\big)=\mathbf{0}$.

Letting $t\to\infty$ in (2.5), using $X_t\mathbf{e}'=0$, $t\geq 1$, $\sum_{n=1}^{\infty}|||R_n|||_\infty<\infty$ and (2.6), we get $\lim_{t\to\infty}X_{(t+1)d+l+1}=X_{l+1}\lim_{t\to\infty}\big(\widetilde{C}_{0,t+1}^{(l)}-\Pi_{0,t+1}^{(l)}\big)=\mathbf{0}$. (because $X_{l+1}\Pi_{m,t}^{(l)}=\mathbf{0}$, $0\leq m<t$). Since l was chosen arbitrarily, it follows $\lim_{t\to\infty}X_{(t+1)d+l+1}=\mathbf{0},0\leq l\leq d-1$. Therefore, the sequence $(X_n)_{n\geq 1}$ is convergent and $\lim_{n\to\infty}X_n=\mathbf{0}$.

(ii) Letting $t\to\infty$ in (2.5), using $X_t\mathbf{e}'=1$, $t\geq 1$, $\sum_{n=1}^{\infty}|||R_n|||_\infty<\infty$ and (2.6), we get $\lim_{t\to\infty}X_{(t+1)d+l+1}=X_{l+1}\lim_{t\to\infty}\widetilde{C}_{0,t+1}^{(l)}=X_{l+1}\Pi=\pi$, $0\leq l\leq d-1$. Therefore, the sequence $(X_n)_{n\geq 1}$ is convergent and $\lim_{n\to\infty}X_n=\pi$.

(iii) Letting $t\to\infty$ in (2.5), using $X_t\mathbf{e}'=1$, $t\geq 1$, $\sum_{n=1}^{\infty}|||R_n|||_\infty<\infty$ and (2.6), we get $\lim_{t\to\infty}X_{(t+1)d+l+1}=X_{l+1}\lim_{t\to\infty}\widetilde{C}_{0,t+1}^{(l)}=X_{l+1}\Pi_{(l+1)}=\pi^{(l+1)}$, $0\leq l\leq d-1$. □

3. WEAK, STRONG AND C-STRONG ERGODICITY RESULTS

In this section a previous study of the author from Nicolaie (2008a) and Nicolaie (2008b) is continued. We give sufficient conditions for weak, strong and C-strong ergodicity of a finite state nonhomogeneous Markov chain in terms of similar behavior of a certain nonhomogeneous Markov chain of smaller size. Our main result is given in Theorem 3.4.

In the sequel, we shall consider a nonhomogeneous Markov chain with state space $S = \{1, 2, ..., r\}$ and transition matrices $(P_n)_{n\geq 1}$. We shall refer to it as the finite Markov chain $(P_n)_{n\geq 1}$. Assume that $\exists d \geq 1$ such that $\lim_{n\to\infty} P_{n,n+d} = P$.

Suppose that P has $p \geq 1$ ergodic classes S_i, $1 \leq i \leq p$, and, perhaps, transient states, i.e., P is of the form (1.1). Let $\mu^{(i)}$ be the invariant probability vector with respect to $S_{(i)}$, $1 \leq i \leq p$, and a_{ji}, $\sum_{t=0}^{p} r_t \leq j \leq r$, $1 \leq i \leq p$ as in Theorem 1.2.

The following Remark will illustrate several aspects that characterize this type of nonhomogeneous Markov chains.

Remark 3.1. (i) Obviously, the condition $\lim_{n\to\infty} P_{n,n+d} = P$, for some $d > 1$, where P is of the form (1.1), does not imply the existence of $\lim_{n\to\infty} P_n$. An example is the chain $(P_n)_{n\geq 1}$, where

$$P_{2n-1} = \begin{pmatrix} \frac{6}{8} - \frac{1}{8n} & \frac{2}{8} + \frac{1}{8n} & 0 \\ \frac{1}{8} + \frac{2}{8n} & \frac{7}{8} - \frac{2}{8n} & 0 \\ 0 & \frac{1}{n} & 1 - \frac{1}{n} \end{pmatrix}, \; P_{2n} = \begin{pmatrix} \frac{4}{10} - \frac{1}{10n} & \frac{6}{10} + \frac{1}{10n} & 0 \\ \frac{3}{10} + \frac{3}{10n^2} & \frac{7}{10} - \frac{3}{10n^2} & 0 \\ 0 & \frac{1}{n} & 1 - \frac{1}{n} \end{pmatrix},$$

$\forall n \geq 1$. We see that $\lim_{n\to\infty} P_n$ does not exist, but there exists $\lim_{n\to\infty} P_{n,n+2} = P$, where

$$P = \begin{pmatrix} \frac{3}{8} & \frac{5}{8} & 0 \\ \frac{25}{80} & \frac{55}{80} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

(ii) A remarkable example for the case when $\lim_{n\to\infty} P_n$ exists is $\lim_{n\to\infty} P_n =: R$, where R has $p \geq 1$ irreducible closed periodic classes S_i, $1 \leq i \leq p$, and, possibly, transient states, S_i having period $d_i \geq 1$, $1 \leq i \leq p$. There exists d (e.g., $d = \text{lcm}\{d_i | 1 \leq i \leq p\} \geq 1$) such that $\lim_{n\to\infty} P_{n,n+d} = R^d$, where R^d has only ergodic classes and, perhaps, transient states, as the next example shows

Let $(P_n)_{n\geq 1}$ given by

$$P_n = \begin{pmatrix} 0 & 1-\frac{1}{n} & 0 & \frac{1}{n} \\ 1-\frac{1}{n} & 0 & 0 & \frac{1}{n} \\ 0 & 0 & 1-\frac{1}{n} & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{pmatrix}, \ n \geq 1.$$

In this section we prove an ergodic theorem for nonhomogeneous Markov chains of this types.

Let $m \geq 0$. By the Chapman-Kolmogorov equation we have $P_{m,n+d} = P_{m,n}P_{n,n+d}$, $n > m$. Subtracting $P_{m,n}$ from both sides, we obtain

$$P_{m,n+d} - P_{m,n} = P_{m,n}\big[-I_r + P_{n,n+d}\big], n > m. \tag{3.1}$$

Setting

$$z_{m,n}^{(i)} = ((P_{m,n})_{i,1}, ..., (P_{m,n})_{ir}), n > m, i \in S, \tag{3.2}$$

then equations (3.1) read as $z_{m,n+d}^{(i)} - z_{m,n}^{(i)} = z_{m,n}^{(i)}\big[-I_r + P_{n,n+d}\big]$, $n > m$, $i \in S$. We remark that $z_{m,n}^{(i)}$ defined in (3.2) are solutions of equations of the type

$$x_{m,n+d} - x_{m,n} = x_{m,n}\big[-I_r + P_{n,n+d}\big], \ n > m, \tag{3.3}$$

under the conditions

$$(x_{m,n})_i \in [0,1], i \in S, \ \sum_{i=1}^{r}(x_{m,n})_i = 1, \ n > m, \tag{3.4}$$

or

$$(x_{m,n})_i \in [-1,1], i \in S, \ \sum_{i=1}^{r}(x_{m,n})_i = 0, \ n > m, \tag{3.5}$$

where $x_{m,n} = [(x_{m,n})_1, ..., (x_{m,n})_r]$, $n > m$.

We are interested in the asymptotic behavior of the proposed solutions of (3.3), under conditions (3.4) or (3.5). A first result is given in the next

LEMMA 3.2.

$$\lim_{n\to\infty}(x_{m,n})_i = 0, \ p+1 \leq i \leq r, \tag{3.6}$$

uniformly with respect to $m \leq 0$.

Proof. A key step is the choice of

$$A = -I_r + P. \tag{3.7}$$

We can benefit of the result given in Theorem 1.4. Let Q and Q^{-1} as in Theorem 1.4. A second key step is the choice of $V_n = P_{n,n+d} - P$, $n \geq 1$, and

$$\widetilde{V}_n = Q^{-1} V_n Q,\ n \geq 1. \tag{3.8}$$

Also, setting

$$y_{m,n} = x_{m,n} Q,\ n > m, \tag{3.9}$$

equations (3.3) amount to

$$y_{m,n+d} - y_{m,n} = y_{m,n} J + y_{m,n} \widetilde{V}_n,\ n > m. \tag{3.10}$$

We shall study the asymptotic behavior of $(y_{m,n})_i$, $p+1 \leq i \leq r$. First, note that

$$|||y_{m,n}|||_\infty \leq |||x_{m,n}|||_\infty |||Q|||_\infty \leq 2|||Q|||_\infty =: \beta, 0 \leq m < n.$$

Using the boundedness of $y_{m,n}$, $0 \leq m < n$, and $\lim_{n\to\infty} V_n = \mathbf{0}$, it follows that

$$(W_{m,n})_i := \sum_{j=1}^{r} (y_{m,n})_j (\widetilde{V}_n)_{ji} \to 0 \text{ as } n \to \infty, m \geq 0,\ p+1 \leq i \leq r.$$

Moreover, the convergence is uniform with respect to $m \geq 0$, for $p+1 \leq i \leq r$.

In the system (3.10) we shall be concerned with the equations corresponding to $i \in \{p+1, ..., p+\sum_{t=2}^{l} m_t + 1\}$. Equivalently, we can write $(y_{m,n+d})_i = (\lambda_2+1)(y_{m,n})_i + (W_{m,n})_i$, $n > m$, $p+1 \leq i \leq p+m_2$, and

$$(y_{m,n+d})_i = (\lambda_t+1)(y_{m,n})_i + (W_{m,n})_i,\ n > m,\ p + \sum_{s=2}^{t-1} m_s + 1 \leq i \leq p + \sum_{s=2}^{t} m_s,\ 3 \leq t \leq l,$$

respectively. The conditions of Proposition 2.1 are verified, so

$$\lim_{n\to\infty} (y_{m,n})_i = 0,\ p+1 \leq i \leq p + \sum_{t=2}^{l} m_t + 1.$$

Moreover, this convergence is uniform with respect to $m \geq 0$, for $p+1 \leq i \leq p+\sum_{t=2}^{l} m_t + 1$ since $\exists \beta \geq 0$ such that $|||y_{m,n}|||_\infty \leq \beta$, $0 \leq m < n$, and the convergence to zero of $(W_{m,n})_i$ is uniform with respect to $m \geq 0$, for $p+1 \leq i \leq r$.

Next, we pay attention to the equations corresponding to i from $p + \sum_{t=2}^{l} m_t + 2$ to $p+\sum_{t=2}^{l+1} m_t$. The case $\varepsilon_k^{(i)} = 0$ for some $1 \leq k \leq m_{l+1}-1$ is similar to the preceding one. We are interested now in cases for which $\varepsilon_k^{(i)} = 1$. Set $k_* = \min\{k \in \{1, \ldots, m_{l+1}-1\} | \varepsilon_k^{(i)} = 1\}$ (k_* is well defined since the set $\{k \in \{1, \ldots, m_{l+1}-1\} | \varepsilon_k^{(i)} = 1\}$ is nonempty, see Remark 1.5 (ii)). If $k_* = 1$, then $(y_{m,n+d})_{p+\sum_{t=2}^{l} m_t+2} = (\lambda_l+1)(y_{m,n})_{p+\sum_{t=2}^{l} m_t+2} + (y_{m,n})_{p+\sum_{t=2}^{l} m_t+1} + (W_{m,n})_{p+\sum_{t=2}^{l} m_t+2}$, $n > m$. Using $\lim_{n\to\infty} (y_{m,n})_{p+\sum_{t=2}^{l} m_t+1} = 0$ uniformly with respect to $m \geq 0$, the conditions of Proposition 2.1 are verified, so we get $\lim_{n\to\infty} (y_{m,n})_{p+\sum_{t=2}^{l} m_t+2} =$

0 uniformly with respect to $m \geq 0$. If $k_* > 1$, then $\lim_{n\to\infty}(y_{m,n})_{p+\sum_{t=2}^{l} m_t+k_*-1} = 0$ uniformly with respect to $m \geq 0$, since $\varepsilon^{(i)}_{k_*-1} = 0$. Also,

$$(y_{m,n+d})_{p+\sum_{t=2}^{l} m_t+k_*} = (\lambda_l + 1)(y_{m,n})_{p+\sum_{t=2}^{l} m_t+k_*} +$$
$$+(y_{m,n})_{p+\sum_{t=2}^{l} m_t+k_*-1} + (W_{m,n})_{p+\sum_{t=2}^{l} m_t+k_*},$$

$n > m$. The conditions of Proposition 2.1 are fulfilled, so $\lim_{n\to\infty}(y_{m,n})_{p+m_2+...+m_l+k_*} = 0$, uniformly with respect to $m \geq 0$. Similarly, $\lim_{n\to\infty}(y_{m,n})_i = 0$ uniformly with respect to $m \geq 0$, for $p + \sum_{t=2}^{l} m_t + 3 \leq i \leq p + \sum_{t=2}^{l+1} m_t$.

Moreover, we can conclude that

$$\lim_{n\to\infty}(y_{m,n})_i = 0, \tag{3.11}$$

regardless of the initial data y_0 and uniformly with respect to $m \geq 0$, for $p+1 \leq i \leq r$. The conclusion follows. □

In the following, we shall prepare the context for the examination of the remaining components $(x_{m,n})_i$, $1 \leq i \leq p$, of the vector $x_{m,n}$. Set

$$(R_{m,n})_i = \sum_{j=p+1}^{r} (y_{m,n})_j (\widetilde{V}_n)_{ji}, n > m, 1 \leq i \leq p.$$

If the equality (3.9) is multiplied on the right side by $\widetilde{V}_n$, we get $y_{m,n}\widetilde{V}_n = x_{m,n}V_nQ$. Set $\widetilde{q}_i$ for the ith column of the matrix Q. From

$$\sum_{i=1}^{p} \big(y_{m,n}\widetilde{V}_n\big)_i = \sum_{j=1}^{r} \big(\sum_{i=1}^{p} (y_{m,n})_j (\widetilde{V}_n)_{ji}\big) =$$
$$= \sum_{j=1}^{p} (y_{m,n})_j \big(\sum_{i=1}^{p} (\widetilde{V}_n)_{ji}\big) + \sum_{j=p+1}^{r} \big(\sum_{i=1}^{p} (y_{m,n})_j (\widetilde{V}_n)_{ji}\big) =$$

(using $\sum_{i=1}^{p} (\widetilde{V}_n)_{ji} = 0$, $1 \leq j \leq p$ (see Nicolaie (2008b), Proposition 3.1))

$$= \sum_{i=1}^{p} (R_{m,n})_i, \ n > m,$$

and

$$\sum_{i=1}^{p} \big(x_{m,n}V_nQ\big)_i = \sum_{i=1}^{p} x_{m,n}V_n\widetilde{q}_i = x_{m,n}V_n \sum_{i=1}^{p} \widetilde{q}_i =$$
$$= x_{m,n}V_n\mathbf{e}' = x_{m,n}\mathbf{0}_{p\times 1} = 0, \ n > m,$$

it follows

$$R_{m,n}\mathbf{e}' = 0, \ n > m. \tag{3.12}$$

Setting $Y_{m,n} = ((y_{m,n})_1, ..., (y_{m,n})_p)$, $R_{m,n} = ((R_{m,n})_1, ..., (R_{m,n})_p)$, $n > m$, and

$$C_n = I_p + (\widetilde{V}_n)_{M\times M}, \ n \geq 1, \text{ where } M = \{1, ..., p\}, \tag{3.13}$$

we have

$$Y_{m,n+d} = Y_{m,n}C_n + R_{m,n}, n > m. \tag{3.14}$$

We need a preliminary result on C_n.

PROPOSITION 3.3. *$(C_n)_{n\geq 1}$ is a sequence of stochastic matrices.*

Proof. Similar to the proof of Proposition 3.1 from Nicolaie (2008b). □

We are now ready to give the main result of this paper. The next theorem is a generalization of Theorem 3.2 from Nicolaie (2008b).

THEOREM 3.4. *Let $(P_n)_{n\geq 1}$ be a Markov chain with state space S for which $\exists d \geq 1$ such that $\lim_{n\to\infty} P_{n,n+d} = P$, where P is of the form (1.1) with $p \geq 1$. Let $V_n = P_{n,n+d} - P$, $n \geq 1$, and $\widetilde{V}_n = Q^{-1}V_nQ$, $n \geq 1$, with Q and Q^{-1} as in Theorem 1.4. Then*

$$\lim_{n\to\infty}(P_{m,n})_{ij} = 0,\ i \in S,\ \sum_{t=0}^{p} r_t \leq j \leq r, \tag{3.15}$$

uniformly with respect to $m \geq 0$.

Moreover, let C_n, $n \geq 1$ as in (3.13) and set $\widetilde{C}_n^{(l)} = C_{(n-1)d+l+1}$, $n \geq 1$, $0 \leq l \leq d-1$. Suppose that

$$\sum_{n=1}^{\infty} |||(\widetilde{V}_n)_{(S\setminus M)\times M}|||_\infty < \infty. \tag{3.16}$$

Then the following statements hold:

(i) *If $(\widetilde{C}_n^{(l)})_{n\geq 1}$ is weakly ergodic, $0 \leq l \leq d-1$, then $(P_n)_{n\geq 1}$ is weakly ergodic, i.e., the chain $(P_n)_{n\geq 1}$ is weakly ergodic* ;

(ii) *If $(\widetilde{C}_n^{(l)})_{n\geq 1}$ is strongly ergodic, $0 \leq l \leq d-1$, such that $\lim_{n\to\infty}\widetilde{C}_{m,n}^{(l)} = \Pi$, $m \geq 0$, $0 \leq l \leq d-1$, then $(P_n)_{n\geq 1}$ is strongly ergodic, i.e., the chain $(P_n)_{n\geq 1}$ is strongly ergodic* .

(iii) *If $(\widetilde{C}_n^{(l)})_{n\geq 1}$ is strongly ergodic, $0 \leq l \leq d-1$, such that $\lim_{n\to\infty}\widetilde{C}_{m,n}^{(l)} = \Pi_{(l+1)} = \mathbf{e}'\pi^{(l+1)}$, $m \geq 0$, $0 \leq l \leq d-1$, then $(P_n)_{n\geq 1}$ is C-strongly ergodic, i.e., the chain $(P_n)_{n\geq 1}$ is C-strongly ergodic.*

Proof. For the first conclusion of the theorem, we will use Lemma 3.1 taking $x_{m,n} = z_{m,n}^{(i)}$ in equation (3.3), where $z_{m,n}^{(i)}$ is defined in (3.2), $n > m$, $i \in S$. From (3.9) we get $x_{m,n} = y_{m,n}Q^{-1}$, $n > m$. Next, letting $n \to \infty$ in the last equality, on account of (3.11), we get (3.15).

(i) We shall take $x_{m,n} = z_{m,n}^{(i)} - z_{m,n}^{(j)}$ in equation (3.3), $n > m$, $i, j \in S$, $i \neq j$, where $z_{m,n}^{(i)}$ is defined in (3.2). Starting from (3.9), using (1.3), we have (for $y_{m,n} := x_{m,n}Q$)

$$\sum_{k=1}^{p}(y_{m,n})_k = \sum_{k=1}^{p} x_{m,n}\widetilde{q}_k = x_{m,n}\sum_{k=1}^{p}\widetilde{q}_k = x_{m,n}\mathbf{e}' = 0,\ n > m.$$

Let $i, j \in S$. By (3.16) and the fact that $\exists \beta \geq 0$ such that $|||y_{m,n}|||_\infty \leq \beta$, $0 \leq m < n$, it follows that $\sum_{n=m}^{\infty} |||R_{m,n}|||_\infty < \infty$, $m \geq 0$; as concerns equation (3.14), by Proposition 2.2 (i), we get $\lim_{n\to\infty} Y_{m,n} = \mathbf{0}$, $m \geq 0$.

This, (3.11) and $x_{m,n} = y_{m,n}Q^{-1}$ give us

$$\lim_{n\to\infty} \big[(P_{m,n})_{ik} - (P_{m,n})_{jk}\big] = 0,\ m \geq 0,\ k \in S.$$

Therefore, $(P_n)_{n\geq 1}$ is weakly ergodic.

(ii) We shall take $x_{m,n} = z_{m,n}^{(i)}$ in equation (3.3), where $z_{m,n}^{(i)}$ is defined in (3.2), $n > m$, $i \in S$. Starting from (3.9), using (1.3), we have (for $y_{m,n} := x_{m,n}Q$)

$$\sum_{k=1}^{p} (y_{m,n})_k = \sum_{k=1}^{p} x_{m,n}\widetilde{q}_k = x_{m,n}\sum_{k=1}^{p} \widetilde{q}_k = x_{m,n}\mathbf{e}' = 1,\ n > m.$$

Let $i \in S$. By (3.16) and the fact that $\exists \beta \geq 0$ such that $|||y_{m,n}|||_\infty \leq \beta$, $0 \leq m < n$, it follows that $\sum_{n=m}^{\infty} |||R_{m,n}|||_\infty < \infty$, $m \geq 0$; as concerns the equation (3.14), by Proposition 2.2 (ii), it follows that $\lim_{n\to\infty} Y_{m,n} = \pi$, $m \geq 0$.

This and $x_{m,n} = y_{m,n}Q^{-1}$ give us

$$\lim_{n\to\infty} \big((P_{m,n})_{i,r_0+\ldots+r_{t-1}}, ..., (P_{m,n})_{i,r_1+\ldots+r_t}\big) = (\mu_1^{(t)}\pi_t, ..., \mu_{r_t}^{(t)}\pi_t).$$

$m \geq 0$, $1 \leq t \leq p$. Therefore, using (3.15), $(P_n)_{n\geq 1}$ is strongly ergodic.

(iii) Making the same choice in equation (3.3) as in the proof of (ii) above and applying Proposition 2.3 (iii), it follows $\lim_{n\to\infty} Y_{m,nd+l+1} = \pi^{(l+1)}$, $m \geq 0$, $0 \leq l \leq d-1$. This and $x_{m,n} = y_{m,n}Q^{-1}$ give us

$$\lim_{n\to\infty} \big((P_{m,nd+l+1})_{i,r_0+\ldots+r_{t-1}}, ..., (P_{m,nd+l+1})_{i,r_1+\ldots+r_t}\big) = (\mu_1^{(t)}\pi_t^{(l+1)}, ..., \mu_{r_t}^{(t)}\pi_t^{(l+1)}).$$

$m \geq 0$, $1 \leq t \leq p$ and $0 \leq l \leq d-1$. Therefore, using (3.15), $\lim_{n\to\infty} P_{m,nd+l+1}$ exists and is a stable stochastic matrix for $m \geq 0$, $0 \leq l \leq d-1$. Set $W_{l+1} = \lim_{n\to\infty} P_{0,nd+l+1}$, $0 \leq l \leq d-1$. The conditions of Theorem 1.13 are fulfilled; it follows that $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} P_{0,k} = \frac{1}{d}\big[W_1 + \ldots + W_d\big]$, i.e., the chain $(P_n)_{n\geq 1}$ is C-strongly ergodic. □

Remark 3.5. (i) We can prove that the chains $(P_n)_{n\geq 1}$ from Remark 3.1 (i) and (ii), respectively, are strongly ergodic using Theorem 3.4 (ii) (for the second chain take $d = 2$).

(ii) Obviously, in the hypothesis (3.16), the condition that any subsequence $(\widetilde{C}_n^{(l)})_{n\geq 1}$ is strongly ergodic, $0 \leq l \leq d-1$, does not imply that $(P_n)_{n\geq 1}$ is strongly ergodic. An example is the chain $(P_n)_{n\geq 1}$, where

$$P_{2n-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ P_{2n} = \begin{pmatrix} 0 & 1 \\ 1-\frac{1}{2n} & \frac{1}{2n} \end{pmatrix},\ n \geq 1.$$

The chain $(\widetilde{C}_n^{(0)})_{n\geq 1}$ is strongly ergodic and

$$\lim_{n\to\infty} \widetilde{C}_{m,n}^{(0)} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} := \Pi_1,\ m \geq 0.$$

Also, the chain $(\widetilde{C}_n^{(1)})_{n\geq 1}$ is strongly ergodic and

$$\lim_{n\to\infty} \widetilde{C}_{m,n}^{(1)} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} := \Pi_2,\ m \geq 0.$$

The chain $(P_n)_{n\geq 1}$ is not strongly ergodic, but using Theorem 3.4 (iii) we can prove that the chain $(P_n)_{n\geq 1}$ is C-strongly ergodic. Moreover, $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} P_{0,k} = \frac{1}{2}\big[\Pi_1 + \Pi_2\big]$.

Acknowledgements

The author thanks the referees for their useful suggestions. The author gratefully acknowledges support from CNCSIS Grant PNII - IDEI 209 "Brownian motion: monotonicity and extremum properties" and from Netherlands Organization for Scientific Research Grant ZONMW-912-07-018 "Prognostic modeling and dynamic prediction for competing risks and multi-state models".

References

Gidas, B., 1985. Nonstationary Markov chains and convergence of the annealing algorithm. J. Statist. Phys. 39, 73-131.

Horn, R.A., Johnson, C.R., 1985. Matrix Analysis. Cambridge University Press, New York.

Iosifescu, M., 1980. Finite Markov Processes and Their Applications. Wiley, Chichester and Editura Tehnică, Bucureşti; reprinted by Dover, 2007.

Isaacson, D.L., Madsen, R.W., 1976. Markov Chains: Theory and Applications. Wiley, New York.

Karlin, S., Taylor, H.M., 1975. A First Course in Stochastic Processes. Academic Press, New York.

Nicolaie, A., 2008a. A sufficient condition for the convergence of a finite Markov chain. Math. Rep.(Bucur.) 10(60), 57-91.

Nicolaie, A., 2008b. On the asymptotic behavior of finite Markov chains. Carpathian J. Math. 24, no.2.