A SUBGEOMETRIC ESTIMATE OF THE STABILITY FOR TIME-HOMOGENEOUS MARKOV CHAINS

UDC 519.21

V. V. GOLOMOZJI

Abstract. Estimates for the stability of time-homogeneous Markov chains are obtained with the help of the coupling method. The results are proved for both the uniform metric and $\| \cdot \|_v$.

1. Introduction

In this paper, we consider the stability of time-homogeneous Markov chains. Let $(E, \mathcal{E})$ be a measurable space and $X$ and $X'$ two independent chains defined in this space. The transition probabilities of the chains are $P(x, A)$ and $P'(x, A)$, respectively. In order to obtain an estimate of the stability, we assume that the chains are close to each other in the sense that

$$
P = (1 - \varepsilon)Q + \varepsilon R, \quad P' = (1 - \varepsilon)Q + \varepsilon R',
$$

where $Q$ is the "common" part of two transition probabilities. The smallness of the parameter $\varepsilon$ measures the closeness of two chains. We show that as $\varepsilon \to 0$, the difference of the transition probabilities after a sufficient number of steps tends to zero not faster than $\varepsilon$ does.

We also obtain an estimate for the nonuniform metric $\| \cdot \|_v$. The proof of this result uses the methods introduced in the proof of Theorem 2.1 in [2]. However, an essentially different result is studied in [2], and this explains why we repeat some reasoning from [2] for the case under consideration.

A thorough investigation of the stability of time-homogeneous Markov chains is carried out in [1]. However, the operator techniques are used in [1], while the key method of this paper is the coupling. This method allows us to generalize the results of [1] to the case of nonhomogeneous Markov chains.

For a measure $\mu$ on $(E, \mathcal{E})$ and a mapping $v: E \to \mathbb{R}$, we introduce the norm

$$
\| \mu \|_v = \sup_{|g| \leq v} \left| \int_E \mu(dx)g(x) \right|.
$$

In particular, if $v = 1$, then

$$
\| \mu \| = |\mu|(E).
$$

Recall that the mapping $Q: E \times \mathcal{E} \to \mathbb{R}$ is a transition kernel if $Q(x, \cdot)$ is a measure on $\mathcal{E}$ for all $x \in E$ and $Q(\cdot, A)$ is an $\mathcal{E}$-measurable function for all $A \in \mathcal{E}$. The product

2010 Mathematics Subject Classification. Primary 60J05.
Key words and phrases. Markov chain, coupling method, stability.

©2010 American Mathematical Society

35
of two kernels is given by

$$PQ(x,A) = \int_E Q(x,dy)P(y,A).$$

By $Q^k$ we denote the $k$-fold product of $Q$ by itself.

We define a measurable function $Qf: E \to \mathbb{R}$ by

$$Qf(x) = \int_E Q(x,dy)f(y),$$

where $f: E \to \mathbb{R}$ is an arbitrary measurable function.

Given a measure $\mu \in bE_+$, we define another measure $\mu Q$ by

$$\mu Q(A) = \int_E \mu(dx)Q(x,A).$$

2. **Main results**

The following conditions are used below to derive estimates:

(A1) Minorization condition (see [3, Chapter 5]): There are a set $C \in \mathcal{E}$, a probability measure $\nu$, and a constant $\alpha > 0$ such that

$$Q(x,\cdot) \geq \alpha \nu(\cdot)$$

for all $x \in C$.

(A2) For a stochastic kernel $Q$, there exists a finite invariant measure $\pi$ such that

$$K = \sum_{k \geq 0} k\|\nu Q^k - \pi\| < \infty,$$

$$K(x) = \sum_{k \geq 0} \|Q^k(x,\cdot) - \pi(\cdot)\| < \infty \text{ for all } x \in E.$$

We introduce the following kernels:

(2) $P_\alpha(x,dy) = \frac{P(x,dy) - \alpha \nu(dy)}{1 - \alpha}, \quad P'_\alpha(x,dy) = \frac{P'(x,dy) - \alpha \nu(dy)}{1 - \alpha},$

$$T(x,x';dy,dy') = (1 - \alpha)\mathbb{1}_{C \times C}(x,x')P_\alpha(x,dy)P'_\alpha(x',dy'),$$

(3) $+ \mathbb{1}_{(C \times C)'}(x,x')P(x,dy)P'(x,dy).$

Put

(4) $\lambda(dy,dy') = \int_E \int_E \pi(dx)R(x,dy) \times R'(x,dy').$

Consider the following condition:

(A3) Set

(5) $m = \sum_{k \geq 0} k\lambda T^k(E,E).$

Then $m$ is finite.

This condition implies that

(6) $\bar{L} = \sum_{k \geq 0} \lambda T^k(E,E) < \infty.$

Now we introduce the transition probability

(7) $\hat{P}(x,x';dy,dy') = T(x,x';dy,dy') + \alpha \mathbb{1}_{C \times C}(x,x')\nu(dy \cap dy').$

Our proof is based on the coupling method. This method can be described in short as follows. Two independent chains start. If they belong to $C \times C$, then we toss a coin
that results in a head with probability \( \alpha \) or a tail with probability \( 1 - \alpha \). If the coin lands a tail up, then the chains make the next step with probabilities \( P_\alpha \) and \( P_\alpha' \) and further develop independently. If the coin lands a head up, then the chains are pasting together and develop according to the transition probability \( Q \) and initial distribution \( \nu \).

At every consecutive step, the pasting chains disconnect with probability \( \epsilon \) and form two independent chains with the initial distribution \( R(x, dy) \) and \( R'(x, dy') \).

Then the above definitions can be stated in terms of the coupling; namely, \( T \) is the probability that the chains do not couple, while \( T_k(x, x') \) is the probability that the chains do not couple during \( k \) steps if they started disconnected from the point \( (x, x') \).

Further, \( \bar{L} \) is the mathematical expectation of the coupling moment for a disconnected chain with initial distribution \( \lambda \). Finally, \( \hat{P} \) describes the evolution of a disconnected chain.

**Theorem 2.1.** Let conditions (A1), (A2), and (A3) hold. Put \( \varepsilon_0 = (m + 5K)^{-1} \). Then

\[
(8) \quad \sup_{n > 0, A \in \mathcal{E}} \left| P^n(x, A) - P'^n(x, A) \right| \leq \varepsilon M(x, \varepsilon)
\]

for all \( x \in E \) and \( \varepsilon < \varepsilon_0 \). Moreover,

\[
M(x, \varepsilon) = K(x) + \bar{L} + \left( K(x) + 1 + \frac{2K(1 - \varepsilon m) + 3}{1 - \varepsilon \varepsilon_0} \right) (\varepsilon K + \bar{L})
\]

and \( M(x, \varepsilon) \leq (\bar{L} + 1)(K(x) + 2K + 5) \) starting with a certain small positive number \( \varepsilon_1 \). This means that \( M(x, \varepsilon) \) is bounded from above and the upper bound does not depend on \( \varepsilon \) for sufficiently small \( \varepsilon \).

The following reasoning is needed to prove a similar result in the nonuniform metric.

The transition probability \( \hat{P} \) (see equality (7)) determines the probability \( \hat{P}_{x,x'} \) and expectation \( \hat{E}_{x,x'} \). Put

\[
(9) \quad \sigma_{C \times C} = \inf\{n \geq 0; (X_n, X'_n) \in C \times C\}
\]

Let the functions \( \alpha(u) \) and \( \beta(u) \) and a real number \( \rho \in (0, 1) \) be such that

\[
(10) \quad \alpha(u)\beta(v) \leq \rho u + (1 - \rho)v.
\]

The pair of functions \((c_0 \beta, c_1 u^{1 - \beta})\), for example, satisfies the latter condition.

Consider the sequences \( r(n) \) of real numbers such that \( r(n) \geq n \). Let

\[
R(n) = \sum_{k=0}^{n} r(k).
\]

Assume that

\[
(\text{B1}) \quad \bar{R} = \sum_{n \geq 0} \alpha(R(n))^{-1} < \infty.
\]

In particular, if \( r(n) = n \) and \( \beta > 1/2 \), then \( R(n) \sim n^2 \) and \( \alpha(R(n))^{-1} \sim n^{-2\beta} \), whence we derive that the series in condition (B1) converges.

Consider the function \( U(x, y) = \hat{E}_{x,y}[R(\sigma_{C \times C})] \) and introduce two further conditions:

**(U1)** For all \( x, y \in E \) \( U(x, y) < \infty \),

\[
(11) \quad b_U = \sup_{x, y \in C} \hat{P}U(x, y) < \infty, \quad \lambda U < \infty.
\]
(V1) There exists a function \( v: E \times E \to [0, \infty) \) such that

\[
b_v = \sup_{x,y \in C} \hat{P}V(x,y) < \infty, \quad \lambda V < \infty
\]

for all \( x, y \in E \) with \( V(x, y) < \infty \), where

\[
V(x, y) = \hat{E}_{x,y} \left[ \sum_{k=0}^{\sigma} v(X_k, X_k') \right].
\]

We also assume that there is a function \( g: X \to \mathbb{R} \) such that

\[
g(x) + g(y) \leq \beta(V(x, y)).
\]

For \( \alpha \) defined in condition (A1), let

\[
M_U = \sup_{n \geq 1} \left[ \left( b_U r(n) \frac{1-\alpha}{\alpha} - R(n+1) \right) + \right],
\]

\[
M_V = b_V \frac{1-\alpha}{\alpha}.
\]

**Theorem 2.2.** Let conditions (A1), (A2), and (A3) hold. We also assume that (B1), (U1), and (V1) hold. Let a function \( g \) satisfy inequality (13) and let \( \varepsilon_0 = (m + 5K)^{-1} \). Then

\[
\sup_{n > 0, A \in \mathcal{L}} \| P^n(x, A) - P^n(x, A) \|_g \leq \varepsilon M_g(x, \varepsilon)
\]

for all \( x \in E \) and all \( \varepsilon < \varepsilon_0 \), where \( \bar{L}_g = (\rho(p_{\nu} + M_u) + (1-p)\rho(p_{\nu} + M_\nu)) \bar{R}, \bar{R} \) is defined in condition (B2), and

\[
M_g(x, \varepsilon) = K(x) + \bar{L}_g + \left( K(x) + 1 + \frac{2K(1-\varepsilon m) + 3}{1-\varepsilon / \varepsilon_0} \right)(\varepsilon K + \bar{L}_g).
\]

Moreover, \( M_g(x, \varepsilon) \leq (\bar{L}_g + 1)(K(x) + 2K + 5) \) for \( \varepsilon < \varepsilon_1 \), where \( \varepsilon_1 \) is a sufficiently small number.

### 3. Main Lemmas

Now we formally define the coupling procedure.

Consider the space \( Z = (E \times E \times \{0, 1, 2\}) \) and a Markov chain \( (Z_n, n \geq 0) \) defined on \( Z \). We define \( Z_n = (X_1, X_2, d_n) \) as follows: \( d_0 = 0 \), \( X_1 = X_0 \), and \( X_2 = X_0' \); if \( Z_n \) is already defined, then \( Z_{n+1} \) is defined as follows.

If \( d_n = 0 \) and \( (X_n, X_n') \in C \times C \), then we toss a coin landing head up with probability \( \alpha \) and tail up otherwise. If the coin lands head up, then we put \( d_{n+1} = 1 \) and \( X_{1_{n+1}} = X_{2_{n+1}} = X \), where \( X \sim \nu(\cdot) \). Otherwise, if the coin lands tail up, then we set \( d_{n+1} = 0 \) and \( (X_{1_{n+1}}, X_{2_{n+1}}) \sim T(X_1, X_2, (\cdot, \cdot)). \)

If \( d_n = 0 \) and \( (X_n, X_n') \notin C \times C \), then we set \( d_{n+1} = 0 \) and

\[
(X_{1_{n+1}}, X_{2_{n+1}}) \sim P \times P'(\cdot, \cdot).
\]

If \( d_n \in \{1, 2\} \), then we toss another coin landing head up with probability \( \varepsilon \). If the coin lands head up, then we set \( Z_{n+1} = (X, X', 0) \), where \( X \sim R(X_1, \cdot) \) and \( X' \sim R'(X_2, \cdot) \). Otherwise, if the coin lands tail up, we set \( Z_{n+1} = (X, X, 2) \), where \( X \sim Q(X_1) \).

Denote by \( \hat{P} \) and \( \hat{E} \) the probability measure and the corresponding expectation generated by the chain \( (Z_n, n \geq 0) \).
Below we write the transition probability for the chain $Z_n$:

\[
P(x, x', 0; A \times A' \times \{0\}) = (1 - \alpha)T(x, x'; A \times A'),
\]

\[
P(x, x', 0; A \times A' \times \{1\}) = \alpha \nu(A \cap A'),
\]

\[
P(x, x', 0; E \times E \times \{2\}) = 0,
\]

\[
P(x, x', 1; A \times A' \times \{0\}) = \varepsilon I_{\{x = x'\}}R(x, A)R'(x', A'),
\]

\[
P(x, x', 1; A \times A' \times \{1\}) = 0,
\]

\[
P(x, x', 1; A \times A' \times \{2\}) = (1 - \varepsilon)I_{\{x = x'\}}Q(x, A \cap A'),
\]

\[
P(x, x', 2; A \times A' \times D) = P(x, x', 1; A \times A' \times D).
\]

We introduce the following notation:

(17) \[ h_n(x) = \sum_{k=1}^{n-1} (1 - \varepsilon)\varepsilon \int Q^{k-1}(x, dy, dz)R(y, dz)R'(y, dz')T^{n-k}(z, z'), \]

(18) \[ h_n = \int \nu(dx)h_n(x), \]

(19) \[ g_n = \sum_{k=1}^{n-1} (1 - \varepsilon)\varepsilon \int \nu(dx)Q^{k-1}(x, y)R(y, dz)R'(y, dz')T^{n-k}(z, z'), \]

(20) \[ a_n(x) = \sum_{k=1}^{n-1} (1 - \varepsilon)\varepsilon \alpha \int Q^{k-1}(x, dy, dz)R(y, dz)R'(y, dz')T^{n-k}(z, z', C \times C), \]

(21) \[ g_n = \sum_{k=1}^{n-1} (1 - \varepsilon)\varepsilon \alpha \int \nu Q^{k-1}(x, dy, dz)R(y, dz)R'(y, dz')T^{n-k}(z, z', C \times C), \]

(22) \[ u_n = \sum_{t \geq 0} (g^t)_n, \]

\[ s_n = \int \pi(dy, dz)R'(y, dz')T^n(z, z', E, E) = \lambda T^n(E, E), \]

(23) \[ S_n = \sum_{k \geq 0} s_k. \]

Therefore, $\bar{L} = \sum_{k \geq 0} s_k$ and $m = \sum_{k \geq 0} S_k = \sum_{k \geq 0} k s_k$.

The above variables have a clear probabilistic meaning. Namely, $h_n(h_n(x))$ is the probability that the chains with initial distribution $\nu(\delta_x)$ start in couple, then disconnect, and after that do not couple again up to the moment $n$.

Further, $g_n$ is the probability that the chains with initial distribution $\nu$ start in couple, then disconnect, and after that couple again at the moment $n$. The sequence $u_n$ represents the renewals generated by $g_n$.

Finally, $s_n$ is the probability that the chains with distribution $\lambda$ start disconnected and do not couple during the time $n$.

We denote by $\tau$ the first pasting moment and by $\tau_1$ the moment of the next coupling.

Remark 3.1. The term $T^k(x, x')$ appeared in the above definitions should be understood as $T^k(x, x', E, E)$. Sometimes we write $h_n(x, dy, dy')$ and $\bar{L}(dy, dy')$. In this case, the operators $T^k(x, x')$ appeared in the definition of $h_n(x)$ and $\bar{L}$ should be understood as $T^k(x, x', dy, dy')$. 

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
Lemma 3.1.

\[ \sup_{n \geq 0} \| \delta_x P^n - \delta_x P^m \| \leq h_n(x) + \left( \sup_{n \geq 0} a_n(x) + \sup_{n \geq 1} u_n \right) \sum_{k \geq 0} h_n, \]

\[ \sup_{n \geq 0} \| \delta_x P^n - \delta_{x'} P^n \|_v \leq \int h_n(x, dy, dy')(v(x) + v(y)) + \left( \sup_{n \geq 0} a_n(x) + \sup_{n \geq 1} u_n \right) \int h_n(dv, dy')(v(x) + v(y)). \]

**Proof.** We have

\[ P(x, A) = \bar{P}(x, x', d; A \times E \times \{0, 1, 2\}), \]

\[ P'(x', A') = \bar{P}(x, x', d; E \times A' \times \{0, 1, 2\}) \]

for all \( d \in \{0, 1, 2\} \).

It is sufficient to prove the lemma for \( d \in \{0, 1\} \), since the transition probabilities for \( d = 2 \) coincide with those for \( d = 1 \):

\[ \bar{P}(x, x', 0; A \times E \times \{0, 1, 2\}) = \bar{P}(x, x', 0; A \times E \times \{0\}) + \bar{P}(x, x', 0; A \times E \times \{1\}) \]

\[ = (1 - \varepsilon)T(x, x'; A \times E) + \varepsilon \nu_e(A, E) \]

\[ = (1 - \varepsilon) \frac{P(x, A) - \varepsilon \nu_e(A) P'(x, E)}{1 - \varepsilon} + \varepsilon \nu_e(A) \]

\[ = (1 - \varepsilon) \frac{P(x, A) - \varepsilon \nu_e(A)}{1 - \varepsilon} + \varepsilon \nu_e(A) \]

\[ = P(x, A), \]

\[ \bar{P}(x, x, 1; A \times E \times \{0, 1, 2\}) = \bar{P}(x, x, 1; A \times E \times \{0\}) + \bar{P}(x, x, 1; A \times E \times \{2\}) \]

\[ = \alpha R(x, A) R'(x, E) + (1 - \alpha)Q(x, A) \]

\[ = \alpha R(x, A) + (1 - \alpha)Q(x, A) \]

\[ = P(x, A). \]

The equalities for \( P' \) are obtained similarly.

Now let \( \phi \in \mathcal{E}_+, d \in \{0, 1, 2\} \), and \( x' \in E \). Then

\[ P^n \phi(x) = \int_{E \times E \times \{0, 1, 2\}} \bar{P}(x, x', d; dx_1 \times dx_1' \times i) \phi(x_1). \]

It is sufficient to check the above equality for \( \phi(x) = I_A(x) \), where \( A \in \mathcal{E} \):

\[ \int_{E \times E \times \{0, 1, 2\}} \bar{P}(x, x', d; dx_1 \times dx_1' \times i) I_A(x_1) \]

\[ = \bar{P}(x, x', 1; A \times E \times \{0, 1, 2\}) = P(x, A) = P \phi(x). \]

Then

\[ P^n \phi(x) = \int_{E \times E \times \{0, 1, 2\}} \bar{P}^n(x, x', d; dx_n \times dx_n' \times i) \phi(x_n) \]

for an arbitrary \( d \in \{0, 1, 2\} \) and all \( x' \in E \) and \( \phi \in \mathcal{E}_+ \).
Lemma 3.1 for $n = 1$ follows from [24]. Assume that the statement of Lemma 3.1 holds for some $n$ and check it for $n + 1$:

\[
\begin{aligned}
\int_{E \times E \times [0,1,2]} \bar{P}^n(x', d; dx_n \times dx_n' \times i_n) \int \bar{P}(x_n, x_n', i_n; dx_{n+1} \times dx_{n+1}' \times i_{n+1}) \phi(x_{n+1}) \\
&= \int_{E \times E \times [0,1,2]} \bar{P}^n(x, x', d; dx_n \times dx_n' \times i_n) P \phi(x_n) \\
&= P^{n+1} \phi(x),
\end{aligned}
\]

\[
\begin{aligned}
|P^n(x, A) - P^m(x, A)| &= |\bar{P}^n(x, x, 1; A \times E \times \{1, 2\}) - \bar{P}^m(x, x, 1; A \times E \times \{1, 2\})| \\
&= \left| \int \bar{P}(x, x, 1; dx_1 \times dx_2 \times \{0, 1, 2\}) \cdots \int \varepsilon(A \cap E) \\
&\quad - \int \bar{P}(x, x, 1; dx_1 \times dx_2 \times \{0, 1, 2\}) \cdots \int \varepsilon(A \cap E) \right| \\
&= 0.
\end{aligned}
\]

Thus

\[
\begin{aligned}
|P^n(x, A) - P^m(x, A)| &= |\bar{P}^n(x, x, 1; A \times E \times \{1\}) - \bar{P}^m(x, x, 1; A \times E \times \{1\})| \\
&\leq \max (\bar{P}^n(x, x, 1; A \times E \times \{1\}), \bar{P}^m(x, x, 1; A \times E \times \{1\})) \\
&\leq \bar{P}^n(x, x, 1; A \times E \times \{1\}),
\end{aligned}
\]

\[
\begin{aligned}
\int |P^n(x, dy) - P^m(x, dy)| v(y) \\
&= \left| \int \bar{P}^n(x, x, 1; dy \times E \times \{0\}) v(y) - \int \bar{P}^m(x, x, 1; dy \times E \times \{0\}) v(y) \right| \\
&\leq \max \left( \int \bar{P}^n(x, x, 1; dy \times E \times \{0\}) v(y), \int \bar{P}^m(x, x, 1; dy \times E \times \{0\}) v(y) \right) \\
&\leq \int_{E \times E} \bar{P}^n(x, x, 1; dy \times dy' \times \{0\}) (v(y) + v(y')).
\end{aligned}
\]

Now

\[
\begin{aligned}
P^n(x, x, 1; E \times E \times \{0\}) &= \bar{P}_{x,1}^{x,1}(\tau > n) + \bar{P}_{x,1}^{x,1}(\tau \leq n, \tau_1 > n) + \bar{P}_{x,1}^{x,1}(\tau_1 \leq n), \\
\bar{P}_{x,1}^{x,1}(\tau > n) &= h_n(x), \\
\bar{P}_{x,1}^{x,1}(\tau \leq n, \tau_1 > n) &= (a(x) \ast h)_n, \\
\bar{P}_{x,1}^{x,1}(\tau_1 \leq n) &= (a(x) \ast u \ast h)_n \leq \sup_{n \geq 1} u_n \sum_{k \geq 0} h_k \sum_{k \geq 0} a_k(x) \leq \sup_{n \geq 1} u_n \sum_{k \geq 0} h_k.
\end{aligned}
\]
Similarly, 
\[
\int \hat{P}_n(x, x, 1; dy \times dy' \times \{0\}) (v(y) + v(y')) \\
= \int \left[ \hat{P}_{x, 1}(\tau > n, (X_n, X'_n) \in (dy, dy')) \\
+ \hat{P}_{x, 1}(\tau \leq n, \tau_1 > n, (X_n, X'_n) \in (dy, dy')) \\
+ \hat{P}_{x, 1}(\tau_1 \leq n, (X_n, X'_n) \in (dy, dy')) \right] (v(y) + v(y')) \\
\int \hat{P}_{x, 1}(\tau > n, (X_n, X'_n) \in (dy, dy')) (v(y) + v(y')) = \int h_n(x, dy, dy')(v(y) + v(y')), \\
\int \hat{P}_{x, 1}(\tau \leq n, \tau_1 > n, (X_n, X'_n) \in (dy, dy')) = \left( a(x) * \int h(dy, dy')(v(y) + v(y')) \right)_n, \\
\int \hat{P}_{x, 1}(\tau_1 \leq n, (X_n, X'_n) \in (dy, dy')) = \left( a(x) * u * \int h(dy, dy')[v(y) + v(y')] \right)_n \\
\leq \sup_{n \geq 1} u_n \sum_{k \geq 0} h_k(dy, dy') [v(y) + v(y')] \sum_{k \geq 0} a_k(x) \\
\leq \sup_{n \geq 1} u_n \sum_{k \geq 0} h_k(dy, dy') [v(y) + v(y')]. \\
\]

The lemma is proved. \(\Box\)

**Lemma 3.2.**

(25) \[ \sup_{n \geq 1} u_n \leq \varepsilon \times \frac{2K + \frac{3}{1-\varepsilon m}}{1 - \varepsilon K \left( 2 + \frac{3\varepsilon}{1-\varepsilon m} \right)}. \]

**Proof.** We have

\[ a_n = \sum_{k=1}^{n-k} (1 - \varepsilon)^{k-1} \int \pi(dy) R(y, dz; R'(y, dz')T^{n-k}(z, z') = \varepsilon \sum_{k=1}^{n} (1 - \varepsilon)^{k-1} s_{n-k}, \]

\[ e_n = \varepsilon \sum_{k=1}^{n} (1 - \varepsilon)^{k-1} (\nu Q^{k-1} - \pi) \int R \times R'(dy, dy')T^{n-k}(y, y', E, E). \]

Adding and subtracting \(\pi\) in the definition (19) of \(g_n\), we get \(g_n = a_n + e_n\). Then \[ u_n = \delta_0(n) + (g \ast u)_n = \delta_0(n) + (a \ast u)_n + (e \ast u)_n. \]

Put \[ y_n = \delta_0(n) + (e \ast u)_n. \]

Hence \[ u_n = y_n + (a \ast u)_n. \]

Therefore, for an arbitrary \(n \geq 0\), \[ u_n = (y \ast b)_n = b_n + (e \ast b \ast u)_n, \]

where \[ b_n = \delta_0(n) + (a \ast b)_n = \sum_{t \geq 0} (a^t)_n. \]

The moment generating function \(b(z)\) of the sequence \(b_n\) is given by \[ b(z) = \frac{1}{1 - a(z)}, \]
where \( a(z) \) is the moment generating function of the sequence \( a_n \):

\[
a(z) = 0 + z \varepsilon s_0 + z^2 \varepsilon \sum_{k=1}^{2} (1 - \varepsilon)^k s_{2-k} + \cdots = \varepsilon z \sum_{n \geq 0} z^n \sum_{k=0}^{n} (1 - \varepsilon)^k s_{n-k} \\
= \varepsilon z \frac{s(z)}{1 - z(1 - \varepsilon)}.
\]

Then the function \( b(z) \) is such that

\[
b(z) = \frac{1}{1 - \varepsilon z s(z)} = \frac{1 - z(1 - \varepsilon)}{1 - z(1 - \varepsilon) - \varepsilon z s(z)} = \frac{1 - z(1 - \varepsilon)}{1 - z + \varepsilon(1 - z) - \varepsilon z s(z)}
\]

\[
= \frac{1 - z + \varepsilon z}{1 - z} \frac{1}{1 - \varepsilon z s(z)} = \left( 1 + \frac{\varepsilon z}{1 - z} \right) \left( \frac{1}{1 - \varepsilon z s(z)} \right),
\]

where \( S_n \) is defined by \( (23) \).

Set

\[
d(z) = \sum_{n \geq 0} z^n d_n = \frac{1}{1 + \varepsilon z S(z)} = \sum_{n \geq 0} z^n (-\varepsilon)^n S(z)^n = \sum_{n \geq 0} z^n (-\varepsilon)^n \sum_{k \geq 0} z^k S_k^{*n}
\]

\[
= \sum_{n \geq 0} z^n \sum_{k = 0}^{n} (-\varepsilon)^k S_{n-k}^{*k}.
\]

Thus

\[
d_0 = 1, \quad d_n = \sum_{k=0}^{n} (-\varepsilon)^k S_{n-k}^{*k}
\]

and

\[
b(z) = d(z) + \frac{\varepsilon z}{1 - z} d(z) = \sum_{k \geq 0} z^k d_k + \varepsilon \sum_{k \geq 0} z^k d_k = d_0 + \sum_{k \geq 1} z^k d_k + \varepsilon \sum_{k \geq 1} z^k d_k = d_0 + \sum_{k \geq 1} z^k d_k + \varepsilon \sum_{k \geq 1} z^k d_k
\]

\[
= d_0 + \sum_{k \geq 1} z^k \left( d_k + \varepsilon \sum_{i=0}^{k-1} d_i \right),
\]

whence we obtain a representation for \( b_n \):

\[
b_0 = d_0 = 1,
\]

\[
b_n = d_n + \varepsilon \sum_{k=0}^{n-1} d_k.
\]

Consider the following auxiliary sequence \( c_n \):

\[
c_0 = 1, \quad c_n = b_n - b_{n-1}.
\]
Note that
\[
b_n = \sum_{k=0}^{n} c_k \leq \sum_{k=0}^{n} |c_k|, \quad \sup_{n \geq 0} b_n \leq \sum_{k \geq 0} |c_k|,
\]
\[
\sup_{n \geq 1} b_n \leq \sum_{k \geq 2} |b_k - b_{k-1}| + |b_1| = \sum_{k \geq 0} |b_k - b_{k-1}| - 2 = \sum_{k \geq 0} |c_k| - 2.
\]
Here we have used the property that \( b_1 = 0 \).

Our aim is to show that
\[
\sum_{n \geq 0} |b_n - b_{n-1}| = \sum_{n \geq 0} |c_n| < \infty.
\]
For all \( n \geq 1 \), we have
\[
b_n - b_{n-1} = d_n - d_{n-1} + \varepsilon d_{n-1}.
\]
Therefore,
\[
\sum_{n \geq 0} |b_n - b_{n-1}| \leq \sum_{n \geq 0} |d_n - d_{n-1}| + \varepsilon \sum_{n \geq 0} |d_n|,
\]
where \( b_{-1} = d_{-1} = 0 \).

Now we evaluate \( \sum_{n \geq 0} |d_n - d_{n-1}| \). Note that
\[
S_{n}^{*k} - S_{n-1}^{*k} = S_{n}^{*k-1} - (S_{n}^{*k-1} * s)_{n-1}
\]
for \( k \geq 1 \) and \( n \geq 1 \), where \( S_n \) is defined by (23). Indeed,
\[
S_{n}^{*k} - S_{n-1}^{*k} = \sum_{i=0}^{n} S_{i}^{*k-1} S_{n-i} - \sum_{i=0}^{n-1} S_{i}^{*k-1} S_{n-1-i} = S_{n}^{*k-1} S_{0} + \sum_{i=0}^{n-1} S_{i}^{*k-1} (S_{n-i} - S_{n-1-i}) = S_{n}^{*k-1} + \sum_{i=0}^{n-1} S_{i}^{*k-1} s_{n-i-1} - (S_{n}^{*k-1} * s)_{n-1}.
\]
Note also that
\[
S_{0}^{*0} = 1, \quad S_{i}^{*0} = 0
\]
for all \( i > 0 \). Putting
\[
D_{1} = \sum_{n \geq 1} \sum_{k=0}^{n-1} \varepsilon^k |S_{n-k}^{*k} - S_{n-k-1}^{*k}|
\]
and taking into account that
\[
S_{0}^{*n} = S_{0} = 1^n = 1, \quad d_0 = 1,
\]
we get
\[
\sum_{n \geq 0} |d_n - d_{n-1}| = d_0 + \sum_{n \geq 1} |d_n - d_{n-1}| = d_0 + \sum_{n \geq 1} \left[ \sum_{k=0}^{n-1} \varepsilon^k |S_{n-k}^{*k} - S_{n-k-1}^{*k}| + \varepsilon^n S_{0}^{*0} \right] = 1 + \frac{\varepsilon}{1 - \varepsilon} + D_1.
\]
Now we evaluate $D_1$:

$$D_1 = \sum_{n \geq 1} \sum_{k=0}^{n-1} \epsilon^k |S_{n-k}^k - S_{n-k-1}^k|$$

$$= \sum_{n \geq 2} \sum_{k=0}^{n-1} \epsilon^k |S_{n-k}^k - S_{n-k-1}^k| + \sum_{k=0}^0 \epsilon^k |S_{1-k}^k - S_{1-k-1}^k|$$

$$= \sum_{n \geq 2} \sum_{k=0}^{n-1} \epsilon^k |S_{n-k}^k - S_{n-k-1}^k| + \epsilon^0 |S_{0}^0 - S_{0}^0| = 1 + D_2,$$

where

$$D_2 = \sum_{n \geq 2} \sum_{k=0}^{n-1} \epsilon^k |S_{n-k}^k - S_{n-k-1}^k|.$$

Using equality (21) we get

$$D_2 \leq \sum_{n \geq 2} \sum_{k=0}^{n-1} \epsilon^k S_{n-k}^{k-1} + \sum_{n \geq 2} \sum_{k=0}^{n-1} \epsilon^k (S_{k-1}^k * s)_{n-k-1} = D_{21} + D_{22},$$

where

$$D_{21} = \sum_{n \geq 2} \sum_{k=0}^{n-1} \epsilon^k S_{n-k}^{k-1},$$

$$D_{22} = \sum_{n \geq 2} \sum_{k=0}^{n-1} \epsilon^k (S_{k-1}^k * s)_{n-k-1}.$$

First we deal with $D_{21}$:

$$D_{21} = \epsilon \sum_{n \geq 0} \epsilon^n \sum_{k=1}^{\infty} S_k^{*n} = \epsilon \sum_{n \geq 0} \epsilon^n \sum_{k \geq 0} S_k^{*n} - \epsilon \sum_{n \geq 0} \epsilon^n S_0^{*n}$$

$$= \epsilon \sum_{n \geq 0} \epsilon^n \left( \sum_{k \geq 0} S_k \right)^n - \frac{\epsilon}{1 - \epsilon} = \epsilon \sum_{n \geq 0} (\epsilon m)^n - \frac{\epsilon}{1 - \epsilon} = \frac{\epsilon}{1 - \epsilon} - \frac{\epsilon}{1 - \epsilon}.$$
Hence we write
\[
\sum_{n \geq 0} |c_n| = \sum_{n \geq 0} |b_n - b_{n-1}| \leq \sum_{n \geq 0} |d_n - d_{n-1}| + \varepsilon \sum_{n \geq 0} |d_n|
\leq 2 + \frac{2\varepsilon}{1 - \varepsilon m} \leq 2 + \frac{3\varepsilon}{1 - \varepsilon m},
\]
(28)
\[
\sup_{n \geq 1} b_n \leq \sum_{n \geq 2} |b_n - b_{n-1}| \leq \frac{3\varepsilon}{1 - \varepsilon m}.
\]

Now we turn to the sequence \(\{u_n\}\):
\[
u_n = (y \ast b)_n = (\delta_0 \ast b)_n + (e \ast b \ast u)_n = b_n + (e \ast b \ast u)_n,
\]
\[
(e \ast b \ast u)_n = \sum_{k=0}^{n} u_k(e \ast b)_{n-k} \leq (e \ast b)_n + \sum_{k=1}^{n} (e \ast b)_{n-k} \sup_{n \geq 1} u_n
\]
\[
= (e \ast b)_n + (e \ast b \ast 1)_{n-1} \sup_{n \geq 1} u_n.
\]

Note that
\[
\nu = 1 \ast c.
\]

Since \(e\) is the difference of two probability distributions, we have
\[
\sum_{k \geq 0} e_k = 0,
\]
\[
E_n = (e \ast 1)_n = \sum_{k=0}^{n} e_k = - \sum_{k > n} e_k.
\]

Consider the series \(\sum_{n \geq 0} |E_n|\):
\[
\sum_{n \geq 0} |E_n| = \sum_{n \geq 0} \sum_{k > n} |e_k| \leq \varepsilon \sum_{n \geq 1} \sum_{k=1}^{n} (1 - \varepsilon)^{k-1} \int (\nu Q^{k-1} - \pi) \, (dy) s_{n-k}(y)
\leq \varepsilon \int \sum_{n \geq 1} n(1 - \varepsilon)^{n-1} (\nu Q^{n-1} - \pi) \, (dy) \times \sum_{n \geq 1} s_n(y)
\leq \varepsilon \sum_{n \geq 1} n(1 - \varepsilon)^{n-1} ||\nu Q^{n-1} - \pi|| \leq \varepsilon K.
\]

We have
\[
(e \ast b \ast 1)_n = (e \ast 1 \ast c \ast 1)_n \leq \sum_{n \geq 0} |E_n| \sum_{n \geq 0} |c_n| \leq \varepsilon K \times \left( 2 + \frac{3\varepsilon}{1 - \varepsilon m} \right).
\]

Inequality (28) implies that
(29)
\[
(e \ast b)_n \leq \sum_{n \geq 0} |E_n| \sup_{n \geq 0} |c_n| \leq \varepsilon K.
\]

Thus
\[
\sup_{n \geq 1} u_n \leq \sup_{n \geq 1} b_n + \varepsilon K + \varepsilon K \times \left( 2 + \frac{3\varepsilon}{1 - \varepsilon m} \right) \sup_{n \geq 1} u_n,
\]
\[
\sup_{n \geq 1} u_n \leq \frac{\varepsilon K + \frac{3\varepsilon}{1 - \varepsilon m}}{1 - \varepsilon K \times \left( 2 + \frac{3\varepsilon}{1 - \varepsilon m} \right)}.
\]

The lemma is proved. \(\square\)
4. Additional lemmas

Lemma 4.1.

\[ \sum_{k \geq 0} h_k \leq \varepsilon K + \bar{L}. \]

Proof. We have

\[
\sum_{n \geq 0} h_n = \varepsilon \sum_{n \geq 0} \sum_{k=1}^{n} (1 - \varepsilon)^{k-1} \nu Q^{k-1} R \times R' T^{n-k}
\]

\[
= \varepsilon \sum_{n \geq 0} \sum_{k=1}^{n} (1 - \varepsilon)^{k-1} (\nu Q^{k-1} - \pi) R \times R' T^{n-k}
\]

\[
+ \varepsilon \sum_{n \geq 0} \sum_{k=1}^{n} (1 - \varepsilon)^{k-1} \pi R \times R' T^{n-k}
\]

\[
\leq \varepsilon K + \varepsilon \sum_{n \geq 0} (1 - \varepsilon)^{n} \sum_{k \geq 0} \lambda T^k \leq \varepsilon K + \bar{L}.
\]

The lemma is proved.  \(\square\)

Lemma 4.2.

\[ \sup_{n \geq 0} h_n(x) \leq \varepsilon \lfloor K(x) + \bar{L} \rfloor, \]

\[ \sup_{n \geq 0} \int h_n(x, dy, dy)(v(x) + v(y)) \leq \varepsilon \left[ K(x) + \int \bar{L}(dy, dy')(v(y) + v(y')) \right]. \]

Proof. We have

\[
h_n(x) = \varepsilon \sum_{k=1}^{n} (1 - \varepsilon)^{k-1} \delta_x Q^{k-1} R \times R' T^{n-k}
\]

\[
\leq \varepsilon \sum_{k=0}^{n} (1 - \varepsilon)^{k-1} \delta_x Q^{k-1} - \pi \right) R \times R' T^{n-k} + \varepsilon \sum_{k=1}^{n} (1 - \varepsilon)^{k-1} \pi R \times R' T^{n-k}
\]

\[
\leq \varepsilon \sum_{k \geq 0} |\delta_x Q^k - \pi| + \varepsilon \sum_{k \geq 0} \lambda T^k = \varepsilon \lfloor K(x) + \bar{L} \rfloor.
\]

Substituting \( T^k \) for \( T^k(dy, dy') \) and reasoning in the same way as before we prove the second equality.

The lemma is proved.  \(\square\)

Lemma 4.3.

\[ \sup_{n \geq 0} a_n(x) \leq \varepsilon \lfloor K(x) + 1 \rfloor. \]

Proof. We have

\[
a_n(x) = \varepsilon \sum_{k=1}^{n} (1 - \varepsilon)^{k-1} \delta_x Q^{k-1} R \times R' T^{n-k}
\]

\[
\leq \varepsilon \sum_{k=1}^{n} (1 - \varepsilon)^{k-1} (\delta_x Q^{k-1} - \pi) R \times R' T^{n-k} + \varepsilon \sum_{k=1}^{n} (1 - \varepsilon)^{k-1} \lambda T^{n-k}
\]

\[
\leq \varepsilon \left[ K(x) + \sum_{k \geq 0} \lambda s_k \right] = \varepsilon \lfloor K(x) + 1 \rfloor.
\]

The lemma is proved.  \(\square\)
5. The Proof of the Main Results

Proof of Theorem 2.1. Lemma 3.1 implies that
\[
\sup_{n \geq 0} \| P_n(x, \cdot) - P^n(x, \cdot) \| \leq \sup_{n \geq 0} h_n(x) + \left( \sup_{n \geq 0} a_n(x) + \sup_{n \geq 1} u_n \right) \sum_{k \geq 0} h_k.
\]

Lemmas 4.1–4.3 yield
\[
\sup_{n \geq 0} \| P_n(x, \cdot) - P^n(x, \cdot) \| \leq \varepsilon[K(x) + \bar{L}] + \left( \varepsilon[K(x) + 1] + \sup_{n \geq 1} u_n \right) (\varepsilon K + \bar{L}).
\]

Substituting the bound (26) for \( u_n \), we obtain
\[
\sup_{n \geq 0} \| P_n(x, \cdot) - P^n(x, \cdot) \|
\leq \varepsilon[K(x) + \bar{L}] + \varepsilon \left( K(x) + 1 \right) \left( 2K + \frac{3}{1 - \varepsilon m} \right) (\varepsilon K + \bar{L})
\]
\[
= \varepsilon[K(x) + \bar{L}] + \varepsilon \left( K(x) + 1 \right) \left( 2K - 2K \varepsilon m + 3 \right) \left( 1 - \varepsilon m \right) + \varepsilon \left( K + 1 \right) \left( 1 - \varepsilon m \right) + 3 \varepsilon K + \bar{L}
\]
\[
\leq \varepsilon[K(x) + \bar{L}] + \varepsilon \left( K(x) + 1 \right) \left( 2K + 1 \right) \left( 1 - \varepsilon m \right) + 3 \varepsilon K + \bar{L}.
\]

It is clear that \( M(x, \varepsilon) \) does not exceed
\[
K(x) + 5 + 2K + \bar{L}(5 + K(x) + 2K) = (\bar{L} + 1)(K(x) + 2K + 5)
\]
starting with some small number \( \varepsilon_1 \).

Proof of Theorem 2.2. It is sufficient to prove that \( \int L(dy, dy')((v(y) + v(y')) = \bar{L}_g \). The rest of the proof is the same as that of Theorem 2.1. We have
\[
\int LT^n(dx, dy)(g(x) + g(y)) \leq \bar{E}_\lambda \left[ \left( (g(X_n) + g(X'_n))(1 - \alpha)^{N_{n-1}} \right) \right]
\]
\[
\leq \bar{E}_\lambda \left[ \beta(V(X_n, X'_n))(1 - \alpha)^{N_{n-1}} \right].
\]
Here \( N_{n-1} \) is the number of visits to the set \( C \times C \) over the time \( n - 1 \). Further,
\[
\alpha(R(n) + M_U) \int LT^n(dx, dy)(g(x) + g(y))
\]
\[
\leq \bar{E}_\lambda \left[ \alpha(R(n) + M_U) \beta(V(X_n, X'_n))(1 - \alpha)^{N_{n-1}} \right]
\]
\[
\leq \rho \bar{E}_\lambda \left[ \beta(V(X_n, X'_n)(1 - \alpha)^{N_{n-1}} \right] + \beta \bar{E}_\lambda \left[ (1 - \alpha)^{N_{n-1}} \right] + \rho \bar{E}_\lambda \left[ (1 - \alpha)^{N_{n-1}} \right].
\]

Given \( n \geq 0 \), we put
\[
U_n(x, x') = \bar{E}_\lambda \left[ \alpha_{C \times C} \sum_{k=0}^{\infty} r(n + k) \right].
\]
Then
\[
\bar{P}U_{n+1} \leq U_n - r(n) + b_U r(n) I_{C \times C}.
\]
Similarly,
\[
\bar{P}V \leq V - v + b_V I_{C \times C}.
\]
For \( n \geq 0 \), put
\[
W_n^0 = U_n(X_n, X_n') + \sum_{k=0}^{n-1} r(k) + M_U,
\]
\[
W_n^1 = V_n(X_n, X_n') + \sum_{k=0}^{n-1} v(X_k, X_k') + M_V.
\]
By construction,
\[
W_n^0 \geq R(n) + M_U,
\]
\[
W_n^1 \geq V(X_n, X_n'),
\]
whence
\[
\alpha(R(n) + M_U) \int \lambda T^n(dx, dy)(g(x) + g(y)) \leq \rho E_\lambda \left[ W_n^0(1 - \alpha)^{N_{n-1}} \right] + (1 - \rho) E_\lambda \left[ W_n^1(1 - \alpha)^{N_{n-1}} \right].
\]
Let
\[
T_n^0 = \prod_{i=0}^{n-1} \frac{W_i^0 + b_U r(i) I_{C \times C}(X_i, X_i')}{W_i^0},
\]
\[
T_n^1 = \prod_{i=0}^{n-1} \frac{W_i^1 + b_V U I_{C \times C}(X_i, X_i')}{W_i^1},
\]
\[
\sigma_0 = \sigma_{C \times C}, \quad \sigma_{j+1} = \inf \{ n \geq \sigma_j : (X_n, X_n') \in C \times C \}.
\]
If \( \alpha = 1 \), then \( T_n^0(1 - \alpha)^{N_{n-1}} = T_n^0 I_{\sigma_0 \geq n} = I_{\sigma_0 \geq n} \leq 1 \). If \( \alpha < 1 \), then by definition \( N_{n-1} = 0, T_n^1 = 1 \), and, for \( N_{n-1} > 0 \),
\[
T_n^0 = \prod_{i=0}^{N_{n-1}-1} \frac{W_{\sigma_i}^0 + b_U r(\sigma_i)}{W_{\sigma_i}^0},
\]
\[
T_n^1 = \prod_{i=0}^{N_{n-1}-1} \frac{W_{\sigma_i}^1 + b_V}{W_{\sigma_i}^1}.
\]
Since
\[
W_n^0 \geq R(n + 1) + M_U
\]
and
\[
\frac{1 + b_U r(n)}{R(n + 1) + M_U} \leq \frac{1}{1 - \alpha}, \quad N_{n-1} > 0,
\]
we obtain
\[
T_n^0(1 - \alpha)^{N_{n-1}} \leq \prod_{i=0}^{N_{n-1}-1} \left( 1 + \frac{b_U r(\sigma_i)}{R(\sigma_i + 1) + M_U} (1 - \alpha) \right) \leq 1.
\]
Taking into account that \( W_n^1 \geq M_V \) and \( 1 + b_V/M_V \leq (1 - \alpha)^{-1} \), we get
\[
T_n^1(1 - \alpha)^{N_{n-1}} \leq 1.
\]
Therefore,
\[
E_\lambda \left[ W_n^0(1 - \alpha)^{N_{n-1}} \right] \leq E_\lambda \left[ \frac{W_n^0}{T_n^0} \right],
\]
\[
E_\lambda \left[ W_n^1(1 - \alpha)^{N_{n-1}} \right] \leq E_\lambda \left[ \frac{W_n^1}{T_n^1} \right].
\]
Now we evaluate $\bar{E}_\lambda[W^0_n/T^0_n]$. By construction,

$$E^{(n)}_{\lambda} \left[ \frac{W^0_n}{T^0_n} \mid F_{n-1} \right] = \bar{E}_{\lambda} \left[ W^0_n \mid F_{n-1} \right] \frac{W^0_{n-1}}{W^0_{n-1} + b_U r(n-1) I_{C \times C}(X_{n-1}, X'_{n-1}) T^0_{n-1}},$$

where $F_n = \sigma([X_k, X'_k])_{k \leq n}$. The shift condition implies that

$$\bar{E}_{\lambda} \left[ W^0_n \mid F_{n-1} \right] \leq W^0_{n-1} + b_U r(n-1) I_{C \times C}(X_{n-1}, X'_{n-1}).$$

Thus $\bar{E}_{\lambda}[W^0_n/T^0_n]$ is a supermartingale, whence we deduce that

$$\bar{E}_{\lambda} \left[ W^0_n (1 - \alpha)^N \mid F_{n-1} \right] \leq \bar{E}_{\lambda}[W^0_n] = \lambda U + M_U.$$

Similarly,

$$\bar{E}_{\lambda} \left[ W^1_n (1 - \alpha)^N \mid F_{n-1} \right] \leq \bar{E}_{\lambda}[W^1_n] = \lambda V + M_V,$$

whence

$$\alpha(R(n) + M_V) \int \lambda T^n(dx, dy) (g(x) + g(y)) \leq \rho(\lambda U + M_U) + (1 - \rho)(\lambda V + M_V),$$

and this completes the proof. \qed

6. CONCLUDING REMARKS

The stability of time-homogeneous Markov chains is considered in this paper. The assumption imposed on transition probabilities is

$$P = (1 - \varepsilon)Q + \varepsilon R, \quad P' = (1 - \varepsilon)Q + \varepsilon R'.$$

Other assumptions are minorization condition (A1) and technical conditions (A2) and (A3). Under these conditions, we prove that the difference of transition probabilities after $n$ steps is of order $\varepsilon$. The results are obtained for the uniform metric and for $\| \cdot \|_v$ as well.

BIBLIOGRAPHY

1. N. V. Kartashov, *Strong Stable Markov Chains*, VSP/TViMS, Utrecht, the Netherlands/Kiev, Ukraine, 1996. MR1451375 (99e:60150)


Received 13/NOV/2009

Translated by S. KVASKO