Austrian Journal of Statistics
2025, Volume 54, 166–176.
http://www.ajs.or.at/
doi:10.17713/ajs.v54i1.2000

Stability of Geometrically Recurrent Time-inhomogeneous Markov Chains

Vitaliy Golomoziy
Taras Shevchenko National
University of Kyiv

Abstract

This paper is devoted to establishing upper bounds for a difference of n-step transition probabilities for two time-inhomogeneous Markov chains with values in a locally compact space when their one-step transition probabilities are close. This stability result is applied to the functional autoregression in $\mathbb{R}^n$.

Keywords: stability, functional autoregression, inhomogeneous Markov chain, drift condition.

1. Introduction

The functional autoregression model plays an important role in many applications. In particular, diffusion models are widely used in modern generative learning; see, for example Ho, Jain, and Abbel (2020) and Ho, Saharia, Chan, Fleet, Norouzi, and Salimans (2022). The typical approach used in such models involves Langevin diffusion in both discrete-time and continuous-time settings (see earlier works Parisi (1981), Grenander and Miller (1994) and modern work Song, Sohl-Dickstein, Kingma, Kumar, Ermon, and Poole (2020)). At the same time, inserting time-inhomogeneity into such models may improve their quality. For example, time-inhomogeneous simulating annealing algorithms (see Bras and Pages (2023)) demonstrate a potential for global optimisation or faster convergence time, while time-inhomogeneous functional autoregression processes in machine learning allow more expressive and complex models. The stability of such models plays a crucial role, for example, in actuarial mathematics, they are used in a non-trivial modelling, such as calculating premium for widow pension with inhomogeneous impact factor (see Golomoziy, Kartashov, and Kartashov (2016)).

That is the main motivation for us to study in this paper the stability of general discrete-time, inhomogeneous Markov chains with application to functional autoregression. The stability of homogeneous chains is well studied, and there are classical books that provide a modern treatment of the subject (Meyn and Tweedie (1993), Douc, Moulines, Priouret, and Soulier (2018)). At the same time, the inhomogeneous case is much more difficult and classical stability results often can not be applied, or require corrections (see Golomoziy (2014) and Golomoziy (2020)).

In the present work, we use the coupling method as a main tool for our research. The coupling method is well described in the classical books of Thorisson (2000) and Lindvall (1991). Our approach, however, is slightly different from that of Lindvall and other researchers (see Douc, Fort, Moulines, and Soulier (2004a), Douc, Moulines, and Rosenthal (2004b), Andrieu, Fort, and Vihola (2015) and Fort and Roberts (2005)), since we are estimating proximity of two different n-steps transition probabilites versus stability of a single Markov chain with regard to different initial distributions. To this end, there is a modified coupling framework (see Golomoziy and Kartashov (2013a), Golomoziy and Kartashov (2013b) and Golomoziy and Kartashov (2016) for Markov chains with discrete state space and Golomoziy (2020) and Golomoziy (2014) for general state spaces) which heavily relies on a renewal theory adapted to the inhomogeneous case (see Golomoziy (2015), Golomoziy and Kartashov (2012), Golomoziy (2022b) and Golomoziy (2017)).

Our stability result is obtained under the condition of geometric recurrence of corresponding time-inhomogeneous Markov chains. The notion of geometric recurrence in a homogeneous case is well described in Douc *et al.* (2018), while the works of Golomoziy (2022a) and Golomoziy (2022b) extend it to the time-inhomogeneous case and provide the necessary instruments to verify such recurrence using so-called drift condition (or Foster-Lyapunov criterion).

This paper is organized as follows. In Section 2, we define two time-inhomogeneous Markov chains with values in some locally compact phase space, introduce key notations and conditions and state the result about stability on n-steps transition probability. In Section 3, we introduce the coupling schema for such chains and provide proof of the main result. Section 4 is devoted to the functional autoregression on $\mathbb{R}^n$. In this section, we establish its geometric recurrence and demonstrate how to apply the stability result from Section 2. Finally, Appendix includes the statement of an auxiliary Lemma 1 from Golomoziy (2023), which plays an essential role in the proof of the main result.

2. Stability of general Markov chains

In this section, we study a pair of independent, discrete-time Markov chains $\{X_n^{(i)}, n \geq 0\}$, $i \in \{1, 2\}$ with values in a locally compact space E equipped with a σ-field $\mathcal{E}$. Since E is locally compact, there exists a sequence of compact sets $\{K_n, n \geq 0\}$, such that $K_n \subset K_{n+1},\ n \geq 0$ and $\bigcup_{n\geq 0} K_n = E$.

For any signed measure μ on $(E, \mathcal{E})$ we define a total variation norm

$$||\mu||_{TV} = |\mu|(E) = \mu^+(E) + \mu^-(E),$$

where $\mu = \mu^+ - \mu^-$ is a Hahn-Jordan decomposition.

Let μ_1 and μ_2 be any non-negative measures on $(E, \mathcal{E})$, and denote by $\mathcal{D}(\mu_1, \mu_2)$ a set of all non-negative measures η on $(E, \mathcal{E})$ such that $\eta(A) \leq \mu_1(A)$ and $\eta(A) \leq \mu_2(A)$, for all $A \in \mathcal{E}$. Proposition D.2.8 from Douc *et al.* (2018) states that there exists a maximal element $\eta^* \in \mathcal{D}(\mu_1, \mu_2)$. This maximal element we call a minimum of two measures and denote by

$$\eta^* = \mu_1 \wedge \mu_2. \tag{1}$$

The minimum of two measures can also be defined as

$$\eta^* = \int_E \left(\frac{d\mu_1}{d(\mu_1 + \mu_2)}(x) \wedge \frac{d\mu_2}{d(\mu_1 + \mu_2)}(x) \right) (\mu_1 + \mu_2)(dx),$$

where $\frac{d\mu_i}{d(\mu_1+\mu_2)}$ is a Radon-Nikodym derivative and $\wedge$ is minimum of two real numbers.

Let us introduce Markov kernels

$$P_{in}(x, A) = \mathsf{P}\left\{ X_{n+1}^{(i)} \in A \middle| X_n^{(i)} = x \right\}, n \geq 1, x \in E, A \in \mathcal{E}.$$

Since we are interested in stability, the kernels P_{1n} and P_{2n} should be close in some sense which we are going to define next. To this end, we introduce substochastic kernels

$$Q_n(x,\cdot) = (P_{1n} \wedge P_{2n})(x,\cdot), \tag{2}$$

where $\wedge$ should be understood as a minimum of two measures defined in (1), and put

$$\varepsilon := 1 - \inf_{n,x} Q_n(x,E) \le 1. \tag{3}$$

Going forward, we assume that $\varepsilon < 1$. We denote the residual substochastic kernels by

$$R_{in}(x,A) = P_{in}(x,A) - Q_n(x,A),$$

so that

$$R_{in}(x,E) \le \varepsilon.$$

Here, $Q_n(x,\cdot)$ can be thought as a "common part" of probabilities $P_{1n}(x,\cdot)$ and $P_{2n}(x,\cdot)$, $x \in E$. Clearly, if $Q_n(x,E) = 1$ then $P_{1n}(x,\cdot) = P_{2n}(x,\cdot)$ and $\varepsilon = 0$.

For every pair of initial states $x, y \in E$, the family of Markov kernels $\{P_{in}, i \in \{1,2\}, n \ge 1\}$ defines a canonical probability space $(\Omega, \mathcal{F}, \mathbb{P}_{xy})$ (see Douc *et al.* (2018), Theorem 3.1.2). Denote its expectation as $\mathbb{E}_{xy}$. In case an event in probability $\mathbb{P}_{xy}\{\cdot\}$ is related to only one chain, $X^{(1)}$ or $X^{(2)}$, we will write $\mathbb{P}_x$ and omit the second initial value. For example, the notation $\mathbb{P}_x\left\{X_n^{(1)} \in A\right\}$ means the probability of $X_n^{(1)} \in A$, assuming $X_0^{(1)} = x$ and arbitrary initial value for $X_0^{(2)}$.

Our goal is to establish the upper bound of the form

$$\sup_n \left\|\mathbb{P}_x\left\{X_n^{(1)} \in \cdot\right\} - \mathbb{P}_x\left\{X_n^{(2)} \in \cdot\right\}\right\|_{TV} \le C(x)\varepsilon, \tag{4}$$

where $C : E \to \mathbb{R}$ is some finite function. Inequality (4) means that uniform in n proximity of one-step transition probabilities implies uniform in n proximity of n-steps transition probabilities.

When studying inhomogeneous chains, it is important to consider the probabilities of the form $\mathbb{P}\left\{X_{n+k}^{(1)} \in A, X_{n+j}^{(2)} \in B \middle| X_n^{(1)} = x, X_n^{(2)} = y\right\}$. We denote them by $\mathbb{P}_{xy}^n$ and follow the notation $\mathbb{P}_x^n$ for events that depend on a single chain $X^{(1)}$ or $X^{(2)}$.

Next, we introduce conditions that are essential in proving the stability result for chains $X^{(1)}$ and $X^{(2)}$.

Condition M (Minorization condition) Assume that there exist a set $C \in \mathcal{E}$, a sequence of real numbers $\{a_n, n \ge 1\}, a_n \in (0,1)$ and a sequence of probability measures ν_n on $(E, \mathcal{E})$ such that:

$$\inf_{x \in C} P_{in}(x,A) \ge a_n \nu_n(A),\ i \in \{1,2\},$$

$$\inf_n \nu_n(C) > 0,$$

$$0 < a_* := \inf_n a_n \le a_n \le a^* := \sup_n a_n < 1,$$

for all $A \in \mathcal{E}$ and $n \ge 1$.

This condition can be understood as a local-mixing condition, where mixing occurs on a set C.

Let us introduce the following condition of geometric recurrence of the pair of chains $\left(X^{(1)}, X^{(2)}\right)$.

Condition GR (Geometric Recurrence) Assume independent chains $X^{(1)}$ and $X^{(2)}$ satisfy **Condition M** and C is a corresponding set. Then, there exist constant $\psi > 1$ such that

$$h(x,y) = \sup_n \mathbb{E}_{xy}^n\left[\psi^{\sigma_{C\times C}}\right] < \infty,$$

for all $x, y \in E$, where

$$\sigma_{C\times C} = \inf\left\{k \geq 1 \ : \ \left(X_k^{(1)}, X_k^{(2)}\right) \in C \times C\right\}.$$

When used in the context of $\mathbb{P}_{xy}^n$ by $\sigma_{C\times C}$ we mean

$$\sigma_{C\times C} = \inf\left\{k \geq n+1 \ : \ \left(X_{n+k}^{(1)}, X_{n+k}^{(2)}\right) \in C \times C\right\}.$$

To prove the main result, we will need some regularity conditions on Q_n.

Condition T (Tails condition). Denote by $A_m = K_{m+1} \setminus K_m$. Assume that there exist sequences $\{\hat{S}_n, n \geq 1\}$ and $\{\hat{r}_n, n \geq 1\}$, such that

$$\hat{m} = \sum_{m\geq 1} \hat{S}_m < \infty, \ \Delta = \sum_{m\geq 1} \hat{r}_m \hat{S}_m < \infty,$$

and

$$\begin{aligned} \left(\prod_{k=1}^n Q_{t+k}\right)(x, A_m) &\leq \left(\prod_{k=1}^n Q_{t+k}\right)(x, E)\hat{S}_m, \ x \in C, \\ \nu_t\left(\prod_{k=1}^n Q^{t+k}\right)(A_m) &\leq \nu_t\left(\prod_{k=1}^n Q_{t+k}\right)(E)\hat{S}_m, \ x \in C, \end{aligned} \tag{5}$$

and

$$\sup_{x,t\in A_m} \int_{E^2\setminus C\times C} \frac{R_{1t}(x,dy)R_{2t}(x,dz)}{1-Q_t(x,E)} h(y,z) \leq \hat{r}_m. \tag{6}$$

for all $t \geq 0$. Here, we understand the product of kernels as

$$(Q_n Q_{n+1})(x, A) = \int_E Q_n(x, dy) Q_{n+1}(y, A).$$

Remark 1. *This condition looks technical and difficult, but in fact, it is a reasonable condition that can be verified using the framework developed in the works Golomoziy (2022b), Golomoziy (2022a) and Golomoziy (2023). Note, that paper Golomoziy (2023) has a detailed discussion of such a condition and contains examples demonstrating how to check it in applications. We only mention that inequalities (5) are the inhomogeneous analogues of positivity (should Q be a homogeneous, positive Markov kernel equalities (5) would be a simple result of ergodicity). Regarding inequality (6), note that $1-Q_t(x,E)$ is approximately equal to ε, each $R_{it}(x,E) \leq \varepsilon$, and so the whole fraction $\frac{R_{1t}(x,E)R_{2t}(x,E)}{1-Q_t(x,E)}$ has a magnitude approximately equals to ε. Speaking about function $h(x,y)$, we can say that it is usually possible to estimate it using results from Golomoziy (2022b), and it very often has the form of $C_1(|x|+|y|)+C_2$, where C_1 and C_2 are some constants, and assuming $E = \mathbb{R}^n$. The value $\sup_{x,t\in A_m} \int_{E^2\setminus C\times C} \frac{R_{1t}(x,dy)R_{2t}(x,dz)}{1-Q_t(x,E)} h(y,z)$ in such typical applications will be $O(m)$, which makes condition $\sum_m \hat{S}_m \hat{r}_m \leq \infty$ to be a moment-type condition. See further discussion in Golomoziy (2023).*

Now, we are ready to state the main result related to the stability of general chains.

Theorem 1. *Let $X^{(i)}$, $i \in \{1,2\}$, be two Markov chains defined above that satisfy **Condition M**, **Condition GR** and **Condition T**. Assume that $\varepsilon < 1$, where ε is deinfed in (3). Then there exist constants $M_1, M_2 \in \mathbb{R}$, such that for every $x \in C$*

$$\left|\left|\mathbb{P}_x^t\left\{X_n^{(1)} \in \cdot\right\} - \mathbb{P}_x^t\left\{X_n^{(2)} \in \cdot\right\}\right|\right| \leq \varepsilon \hat{m} M_1 + \Delta M_2, \tag{7}$$

*where $\hat{m}$ and Δ are defined in **Condition T**.*

For every $x \notin C$ the following inequality holds true

$$\left\|\mathbb{P}_x^t\left\{X_n^{(1)} \in \cdot\right\} - \mathbb{P}_x^t\left\{X_n^{(2)} \in \cdot\right\}\right\| \le \varepsilon(2\hat{m}M_1 + \hat{\mu}(x)) + 2\Delta M_2, \tag{8}$$

where

$$\hat{\mu}(x) = \sup_t \sum_{k\ge 1}\left(\prod_{j=0}^{k-1} Q_{t+j}\mathbb{1}_{E\setminus C}\right)(x, E\setminus C) \le 1.$$

Remark 2. *Note, that the bound (7) is note of the same form as (4), due to the term ΔM_2. However, in many applications, $\Delta = O(\varepsilon)$, so that we actually have a bound of the form (4).*

3. Coupling construction and proof of stability bounds

The key tool in proving Theorem 1 will be the modified coupling method. We define it as follows.

Assuming **Condition M** holds true, we define "noncoupling" operators

$$T_{it}(x, A) = \frac{P_{it}(x, A) - a_t\nu_t(A)}{1 - a_t},$$

$$T_{xy}^{(t)}(A, B) = T_{1t}(x, A)T_{2t}(y, B).$$

We define the Markov chain $\bar{Z}_n = \left(Z_n^{(1)}, Z_n^{(2)}, d_n\right)$ with values in $(E, E, \{0,1,2\})$ by setting its transition probabilities

$$\bar{P}_n(x, y, 1; A\times B\times\{2\}) = \mathbb{1}_{x=y}Q_n(x, A\cap B),$$

$$\bar{P}_n(x, y, 1; A\times B\times\{0\}) = \mathbb{1}_{x=y}\frac{R_{1n}(x, A)R_{2n}(y, B)}{1 - Q_n(x, E)},$$

we assume the latter probability is equal to zero if $Q_n(x, E) = 1$,

$$\begin{aligned}\bar{P}_n(x, y, 0; A\times B\times\{0\}) &= (1 - a_n)\mathbb{1}_{C\times C}(x, y)T_{1n}(x, A)T_{2n}(y, B)\\ &\quad + (1 - \mathbb{1}_{C\times C}(x, y))P_{1n}(x, A)P_{2n}(y, A)\end{aligned}$$

$$\bar{P}_n(x, y, 0; A\times B\times\{1\}) = \mathbb{1}_{C\times C}(x, y)a_n\nu_n(A\cap B),$$
$$\bar{P}_n(x, y, 2; \cdot) = \bar{P}_n(x, y, 1; \cdot).$$

All other probabilities are equal to zero.

It is straightforward that marginal distributions of the process $\bar{Z}_n$ equal to those of $X_n^{(i)}$. Indeed, for all $x \in E$

$$\bar{P}_n(x, x, 1, A\times E\times\{0,1,2\}) = Q_n(x, A) + \frac{R_{1n}(x, A)R_{2n}(x, E)}{1 - Q_n(x, E)} = Q_n(x, A) + R_{1n}(x, A) = P_{1n}(x, A),$$

for all $(x, y) \in C\times C$

$$\bar{P}_n(x, y, 0, A\times E\times\{0,1,2\}) = (1 - a_n)T_{1n}(x, A)T_{2n}(x, E) + a_n\nu_n(A)$$

$$= P_{1n}(x, A) - a_n\nu_n(A) + a_n\nu_n(A) = P_{1n}(x, A),$$

and for all $(x, y) \notin C\times C$

$$\bar{P}_n(x, y, 0, A\times E\times\{0,1,2\}) = P_{1n}(x, A)P_{2n}(y, E) = P_{1n}(x, A).$$

Similar equalities for P_{2n} and for $\bar{P}_n(x,y,2,A\times E\times\{0,1,2\})$ can be obtained in the same exact fashion.

We will use canonical probability $\bar{\mathbb{P}}^t_{x,y,d}$ and expectation $\bar{\mathbb{E}}^t_{x,y,d}$, $x,y\in E$, $d\in\{0,1,2\}$ in the same sense as before.

Let us denote by

$$\bar{\sigma}_{C\times C}=\bar{\sigma}_{C\times C}(1)=\inf\left\{n\geq 1:\left(Z^{(1)}_n,Z^{(2)}_n\right)\in C\times C\right\},$$

$$\bar{\sigma}_{C\times C}(m)=\inf\left\{n\geq\bar{\sigma}_{C\times C}(m-1):\left(Z^{(1)}_n,Z^{(2)}_n\right)\in C\times C\right\},m\geq 2,$$

$$\bar{\sigma}=\inf\{t>0:Z^{(1)}_t=Z^{(2)}_t\in C,d_1=\ldots=d_t=2\}.$$

the first and m-th return times to $C\times C$ by the pair $\left(Z^{(1)}_n,Z^{(2)}_n\right)$.

We will also need a special notation for the sets

$$\begin{aligned}D_n&=\{d_1=d_2=\ldots=d_n=0\},\\D_{nk}&=\{d_1=d_2=\ldots=d_n=0,\bar{\sigma}_{C\times C}(k)=n\},\\B_{nk}&=\{d_k\in\{1,2\},d_{k+1}=0,\ldots d_n=0\},\end{aligned}\tag{9}$$

and for the values

$$\rho_{nk}=\sup_{x,y\in C,t}\bar{\mathbb{P}}^t_{x,y,0}\left(D_{nk}\right).\tag{10}$$

Proof of Theorem 1. The proof of this theorem follows the line of reasoning of Theorem 2 from Golomoziy (2023) with only minor deviations.

First, from Theorem 2 in Golomoziy (2023) we have an inquality

$$\left|\mathbb{P}^t_x\left\{X^{(1)}_n\in A\right\}-\mathbb{P}^t_x\left\{X^{(2)}_n\in A\right\}\right|\leq\bar{\mathbb{P}}^t_{x,x,1}\left\{d_n=0\right\}\leq\sum_{k=1}^{n-1}\sum_{m\geq 0}\Lambda^t_k(x,A_m)\sup_{y\in A_m}\bar{\mathbb{P}}^{t+k}_{y,y,1}(D_{n-k}),\tag{11}$$

where

$$\Lambda^t_k(x,A)=\bar{\mathbb{P}}^t_{x,x,1}\left\{d_k\in\{1,2\},Z^{(1)}_k=Z^{(2)}_k\in A\right\}.$$

and $A_m=K_{m+1}\setminus K_m$.

Second, from the proof of same Theorem 2 in Golomoziy (2023) we know that

$$\Lambda^t_k(x,A_m)\leq\mathbb{1}_C(x)\hat{S}_m+\mathbb{1}_{E\setminus C}(x)\left(2\hat{S}_m+q_{k,m}(x)\right)$$

where

$$q_{k,m}(x)=\bar{\mathbb{P}}^t_{x,x,1}\{d_2=\ldots=d_{k-1}=2,\bar{\sigma}\geq k,Z^{(1)}_{k-1}=Z^{(2)}_{k-1}\in A_m\},$$

and

$$\sum_{k,m\geq 0}q_{k,m}(x)=\hat{\mu}(x).$$

Denote by

$$H^{(m)}_t=\sup_{x,t\in A_m}\int_{E^2\setminus C\times C}\frac{R_{1t}(x,dy)R_{2t}(x,dz)}{1-Q_t(x,E)}h(y,z).$$

Note that by **Condition T** $H^{(m)}_t\leq\hat{r}_m$.

Let us first consider the case $x \in C$. Since $\Lambda_k^t(x, A_m) \le \hat{S}_m$ in this case, we can apply Lemma 1 to (11) and obtain

$$\left|\mathbb{P}_x^t\left\{X_n^{(1)} \in A\right\} - \mathbb{P}_x^t\left\{X_n^{(2)} \in A\right\}\right| \le \sum_{m\ge 0} \hat{S}_m \sum_{k\ge 1} \sup_{y\in A_m, t} \bar{\mathbb{P}}_{y,y,1}^t(D_k)$$
$$\le \varepsilon \hat{m} S \frac{\psi(1+\rho)}{\psi - 1} + \left(\sum_{m\ge 0} \hat{S}_m H_t^{(m)}\right)\left(\frac{\psi S(1+\rho\psi)}{(\psi-1)^2} + \frac{\psi(1+S)}{\psi-1}\right)$$
$$\le \varepsilon \hat{m} S \frac{\psi(1+\rho)}{\psi - 1} + \Delta\left(\frac{\psi S(1+\rho\psi)}{(\psi-1)^2} + \frac{\psi(1+S)}{\psi-1}\right),$$

where $\hat{m}$ is defined in **Condition T**.

In case $x \notin C$ we can follow the same reasoning and use the equality $\sum_{k,m\ge 0} q_{k,m}(x) = \hat{\mu}(x)$ along with obvious $\bar{\mathbb{P}}_{y,y,1}^{t+k}(D_{n-k}) \le \varepsilon$ to obtain

$$\left|\mathbb{P}_x^t\left\{X_n^{(1)} \in A\right\} - \mathbb{P}_x^t\left\{X_n^{(2)} \in A\right\}\right| \le \sum_{m\ge 0} \hat{S}_m \sum_{k\ge 1} \sup_{y\in A_m, t} \bar{\mathbb{P}}_{y,y,1}^t(D_k)$$
$$\le 2\varepsilon \hat{m} S \frac{\psi(1+\rho)}{\psi - 1} + 2\Delta\left(\frac{\psi S(1+\rho\psi)}{(\psi-1)^2} + \frac{\psi(1+S)}{\psi-1}\right) + \varepsilon\hat{\mu}(x).$$

□

4. Applicaton to functional autoregression

In this section we demonstrate an example of estimating function $h(x, y)$ in a model which plays an important role in applications, in particular in optimisation problems that arise in modern diffusion models of generative learning. Let $E = \mathbb{R}^d$ and $\mathcal{E}$ be a Borel σ-field. As before we consider a pair of independent time-inhomogeneous Markov chains $X_n^{(1)}, X_n^{(2)}$ with values in $\mathbb{R}^d$ of the form

$$X_n^{(i)} = f_n^{(i)}\left(X_{n-1}^{(i)}\right) + Z_n^{(i)}, \tag{12}$$

where $f_n^{(i)}$ - are locally bounded, measurable functions, and and $Z_n^{(i)}$ are independent random variables.

Theorem 2. *Let $X^{(1)}$ and $X^{(2)}$ be independent Markov chains defined in (12). Assume the following conditions hold*
1. $\mu_1 := \sup_{i,n} \mathbb{E}|Z_n^{(i)}| < \infty$.
2.

$$\sup_{i,n} \limsup_{|x|\to\infty} \frac{|f_n^{(i)}(x)|}{|x|} < 1.$$

Then $X^{(1)}$ and $X^{(2)}$ are geometrically recurrent and the following estimate holds for all $x \in \mathbb{R}^d$ and some $r_1 > 0$

$$\sup_n \mathbb{E}_{ix}^n\left[\lambda^{-\sigma_C^{(i)}}\right] \le 1 + |x| + b\mathbb{1}_{|x|\le r_1},$$

where

$$\lambda = \frac{1}{2}\left(1 + \sup_{i,n} \limsup_{|x|\to\infty} \frac{|f_n^{(i)}(x)|}{|x|}\right) < 1,$$

$$C = \{|x| \le r_1\},$$

and b some constant.

Proof. This result is well-known for time-homogeneous chains (see Douc *et al.* (2018), Example 11.4.3, p.257). In order to establish geometric recurrence we will use Theorem 1 from Golomoziy (2022a). Let us select a Foster-Lyapunov function

$$V(x) = 1 + |x|.$$

We have then

$$P_{in}V(x) = 1 + \mathbb{E}_{in}\left[\left|f_n^{(i)}(x) + Z_n^{(i)}\right|\right] \leq 1 + |f_n^{(i)}(x)| + \mu_1 = V(x) + |f_n^{(i)}(x)| - |x| + \mu_1$$

$$= V(x) + |x|\left(\frac{|f_n^{(i)}(x)|}{|x|} - 1\right) + \mu_1 = \lambda V(x) + (1-\lambda)(1+|x|) + |x|\left(\frac{|f_n^{(i)}(x)|}{|x|} - 1\right) + \mu_1$$

$$= \lambda V(x) + (1-\lambda) + |x|\left(\frac{|f_n^{(i)}(x)|}{|x|} - \lambda\right) + \mu_1$$

Clearly

$$\lim_{|x|\to\infty} (1-\lambda) + |x|\left(\frac{|f_n^{(i)}(x)|}{|x|} - \lambda\right) + \mu_1 = -\infty,$$

so we can select

$$r_1 := \inf_x \left\{(1-\lambda) + |x|\left(\frac{|f_n^{(i)}(x)|}{|x|} - \lambda\right) + \mu_1 < 0\right\} + 1.$$

The result now follows from Theorem 1 from Golomoziy (2022a) with set $C = \{|x| \leq r_1\}$. □

Theorem 3. *Let $X^{(1)}$ and $X^{(2)}$ be the chains defined above. Assume conditions of Theorem 2 hold. Assume, additionally, that* ***Condition M*** *holds true for both chains $X^{(1)}$, $X^{(2)}$ with the same measures ν_n, constants a_n, $n \geq 1$ and set $C = \{|x| \leq r_1\}$. Then there exist $\psi > 1$ and constants $C_0, C_1 \in \mathbb{R}$, such that for every $x, y \in \mathbb{R}^d$ and $n \geq 0$*

$$\mathbb{E}_{xy}^n\left[\psi_1^{\sigma_{C\times C}}\right] \leq C_0(|x| + |y|) + C_1,$$

where

$$\sigma_{C\times C} = \inf\left\{t \geq 1 : \left(X_t^{(1)}, X_t^{(2)}\right) \in C \times C\right\},$$
$$C = \{|x| \leq r_1\},$$

and r_1 is defined in Theorem 2.

Proof. From Golomoziy (2022b), Theorem 4.2 we can find such $\psi_0, \psi \in (1, 1/\lambda)$ (where λ is defined in Theorem 2) that

$$\mathbb{E}_{x,y}^n\left[\psi^{\sigma_{C\times C}}\right] \leq M\left(\mathbb{E}_x^n\left[\psi_0^{\sigma_C^{(1)}}\right] S_1(\psi_0) + \mathbb{E}_y^n\left[\psi_0^{\sigma_C^{(2)}}\right] S_2(\psi_0)\right),$$

where $M \in \mathbb{R}$ is some constant and

$$S_i(u) = \sup_{n, x\in C}\left\{\frac{1}{1-a_n}\left(\mathbb{E}_x^n\left[u^{\sigma_C^{(i)}}\right] - a_n \mathbb{E}_{\nu_n}^n\left[u^{\sigma_C^{(i)}}\right]\right)\right\}.$$

Note, that **Condition M** guarantees that $S_i(u)$ are positive and finite for all $\psi_0 < 1/\lambda$. From the estimate in Theorem 2 we get

$$\begin{aligned}\mathbb{E}_{x,y}^n\left[\psi^{\sigma_{C\times C}}\right] &\leq M\left((|x| + 1 + b)S_1(\psi_0) + (|y| + 1 + b)S_2(\psi_0)\right)\\ &\leq MB(|x| + |y|) + MB(1+b),\end{aligned}$$

where $B = \max\{S_1(\psi_0), S_2(\psi_0)\}$. □

Remark 3. *For the model* (12) *minorization condition takes the form*

$$\min_{|x|\le r_1,i,n} \Gamma_n^{(i)}\left(C - f_n^{(i)}(x)\right) > 0, \tag{13}$$

where $\Gamma_n^{(i)}$ *is a distribution of* $Z_n^{(i)}$, $\min$ *is a minimum of measures defined in* (1). *Condition* (13) *is a natural condition, which means all incremental distributions have some common probability mass around* 0 *(assuming some regularity of* $f_n^{(i)}$*).*

The final question to address before we can apply Theorem 1 to the model (12) is related to verifying **Condition T**. In order to verify **Condition T**, we need more information about functions $f_n^{(i)}$ and distributions of increments $Z_n^{(i)}$. For example, in Golomoziy (2023), it was shown that for a linear autoregression in $\mathbb{R}$ **Condition T** reduces to

$$\sup_{n,i} \mathbb{E}|Z_n^{(i)}|^p < \infty,$$

for any $p > 2$, and is valid, for example, for Guassian autoregression. The same exact reasoning can be applied to a linear autoregression in $\mathbb{R}^n$. For non-linear autoregressions, exact calculations depend on the form of $f_n^{(i)}$.

5. Appendix

Lemma 1. *In the notations of Section 2 the following inequality holds*

$$\sum_{n\ge 1} \sup_{x\in[m,m+1),t} \bar{\mathbb{P}}^t_{x,x,1}\{D_n\} \le$$

$$\le \varepsilon S \frac{\psi(1+\rho)}{\psi-1} + H_t^{(m)}(x)\left(\frac{\psi S(1+\rho\psi)}{(\psi-1)^2} + \frac{\psi(1+S)}{\psi-1}\right),$$

where

$$H_t^{(m)} = \sup_{x,t\in A_m} \int_{E^2\setminus C\times C} \frac{R_{1t}(x,dy)R_{2t}(x,dz)}{1-Q_t(x,E)} h(y,z),$$

and

$$S = (1-a_*) \sup_{x,y\in C, t\ge 0,} \int_{\mathbb{R}^2\setminus C\times C} T^{(t)}_{x,y}(du,dv)h(u,v).$$

This statement was proved in Lemma 9, Golomoziy (2023).

References

Andrieu C, Fort G, Vihola M (2015). "Quantitative Convergence Rates for Sugeometric Markov Chains." *Annals of Applied Probability*, **52**, 391–404.

Bras P, Pages G (2023). "Convergence of Langevin-simulated Annealing Algorithms with Multiplicative Noise II: Total Variation." *Monte Carlo Methods and Applications*, **29**(3), 203–219. doi:10.1515/mcma-2023-2009.

Douc R, Fort G, Moulines E, Soulier P (2004a). "Practical Drift Conditions for Subgeometric Rates of Convergence." *Annals of Applied Probability*, **14**(3), 1353–1377. doi:10.1214/105051604000000323.

Douc R, Moulines E, Priouret P, Soulier P (2018). *Markov Chains*. Springer, Switzerland. doi:10.1007/978-3-319-97704-1.

Douc R, Moulines E, Rosenthal S (2004b). "Quantitative Bounds on Convergence of Time-inhomogeneous Markov Chains." *Annals of Applied Probability*, **14**(4), 1643–1665. doi:10.1214/105051604000000620.

Fort G, Roberts GO (2005). "Subgeometric Ergodicity of Strong Markov Processes." *Annals of Applied Probability*, **15**, 1565–1589.

Golomoziy V (2014). "An Estimate of the Stability for Nonhomogeneous Markov Chains under Classical Minorization Condition." *Theory of Probability and Mathematical Statistics*, **88**, 34–59. doi:10.1090/S0094-9000-2014-00917-5.

Golomoziy V (2015). "An Inequality for the Coupling Moment in the Case of Two Inhomogeneous Markov Chains." *Theory of Probability and Mathematical Statistics*, **90**, 43–56. doi:10.1090/tpms/948.

Golomoziy V (2017). "An Estimate of the Expectation of the Excess of a Renewal Sequence Generated by a Time-inhomogeneous Markov Chain if a Square-integrable Majorizing Sequence Exists." *Theory of Probability and Mathematical Statistics*, **94**, 53–62. doi:10.1090/tpms/1008.

Golomoziy V (2020). "Estimates of Stability of Transition Probabilities for Non-homogeneous Markov Chains in the Case of the Uniform Minorization." *Theory of Probability and Mathematical Statistics*, **101**, 85–101. doi:10.1090/tpms/1113.

Golomoziy V (2022a). "Computable Bounds of Exponential Moments of Simultaneous Hitting Time for Two Time-Inhomogeneous Atomic Markov Chains." In A Malyarenko, Y Ni, M Rančić, S Silvestrov (eds.), *Stochastic Processes, Statistical Methods, and Engineering Mathematics*, pp. 97–119. Springer International Publishing. doi:10.1007/978-3-031-17820-7_5.

Golomoziy V (2022b). "Exponential Moments of Simultaneous Hitting Time for Non-atomic Markov Chains." *Glasnik Matematicki*, **57**(1), 129–147. doi:10.3336/gm.57.1.09.

Golomoziy V (2023). "On Geometric Recurrence for Time-inhomogeneous Autoregression." *Modern Stochastics: Theory and Applications*, **10**(3), 313–341. doi:10.15559/23-VMSTA228.

Golomoziy V, Kartashov M (2012). "The Mean Coupling Time for Independent Discrete Renewal Processes." *Theory of Probability and Mathematical Statistics*, **84**, 79–86. doi:10.1090/S0094-9000-2012-00855-7.

Golomoziy V, Kartashov M (2013a). "Maximal Coupling Procedure and Stability of Discrete Markov Chains. I." *Theory of Probability and Mathematical Statistics*, **86**, 93–104. doi:10.1090/S0094-9000-2013-00891-6.

Golomoziy V, Kartashov M (2013b). "Maximal Coupling Procedure and Stability of Discrete Markov Chains. II." *Theory of Probability and Mathematical Statistics*, **87**, 65–78. doi:10.1090/S0094-9000-2014-00905-9.

Golomoziy V, Kartashov M (2016). "Maximal Coupling and V -stability of Discrete Nonhomogeneous Markov Chains." *Theory of Probability and Mathematical Statistics*, **93**, 19–31. doi:10.1090/tpms/992.

Golomoziy V, Kartashov M, Kartashov Y (2016). "Impact of the Stress Factor on the Price of Widow's Pensions. Proofs." *Theory of Probability and Mathematical Statistics*, **92**, 17–22. doi:10.1090/tpms/979.

Grenander U, Miller MI (1994). "Representations of Knowledge in Complex Systems." *Journal of the Royal Statistical Society: Series B (Methodological)*, **56**, 549–581. doi:10.1111/j.2517-6161.1994.tb02000.x.

Ho J, Jain A, Abbel P (2020). "Denoising Diffusion Probabilistic Models." pp. 6840–6852.

Ho J, Saharia C, Chan W, Fleet DJ, Norouzi M, Salimans T (2022). "Cascaded Diffusion Models for High Fidelity Image Generation." *The Journal of Machine Learning Research*, **23**, 22499–2281.

Lindvall T (1991). *Lectures on Coupling Method*. John Wiley and Sons.

Meyn S, Tweedie R (1993). *Markov Chains and Stochastic Stability*. Springer London, Switzerland. doi:10.1007/978-3-319-97704-1.

Parisi G (1981). "Correlation Functions and Computer Simulations." *Nuclear Physics B*, **180**(3), 378–384. doi:10.1016/0550-3213(81)90056-0.

Song Y, Sohl-Dickstein JN, Kingma DP, Kumar A, Ermon S, Poole B (2020). "Score-Based Generative Modeling through Stochastic Differential Equations."

Thorisson H (2000). *Coupling, Stationarity, and Regeneration*. Springer, New York.

Affiliation:

Vitaliy Golomoziy
Department of Probability Theory, Statistics and Actuarial Mathematics
Mechanics and Mathematics faculty,
Taras Shevchenko National University of Kyiv,
60 Volodymyrska Street, City of Kyiv, Ukraine, 01033,
E-mail: vitaliy.golomoziy@knu.ua

Austrian Journal of Statistics http://www.ajs.or.at/
published by the Austrian Society of Statistics http://www.osg.or.at/
Volume 54 *Submitted:* 2024-07-13
2025 *Accepted:* 2024-07-29