MAXIMAL COUPLING PROCEDURE AND STABILITY OF DISCRETE MARKOV CHAINS. I

M. V. KARTASHOV AND V. V. GOLOMOZYĬ

Abstract. Two discrete Markov chains whose one-step transition probabilities are close to each other in the uniform total variation norm or in the V-norm are considered. The problem of stability of the transition probabilities for an arbitrary number of steps is investigated. The main assumption is either the uniform mixing or V-mixing condition. In particular, we prove that the uniform distance between the distributions of the chains after an arbitrary number of steps does not exceed $\varepsilon/(1-\rho)$, where ε is the uniform distance between the transition matrices and where ρ is the uniform mixing coefficient. A number of general examples are considered. The proofs are based on the maximal coupling procedure that maximizes the one-step coupling probabilities.

1. Introduction

A detailed investigation of the stability of distributions is presented in the monograph of the first coauthor [11] for general Markov chains under rather general assumptions on the mixing. The approach used in [11] is based on analytical operator methods including several new inequalities for the asymptotic behavior of renewal processes and solutions of renewal equations.

The main results of the theory of stability of stochastic models are given in the monograph by Zolotarev [28]. Some important achievements in the theory of stability are described in the book by Meyn and Tweedie [18].

The so-called *coupling method* was developed by several authors to study the asymptotic behavior of Markov processes. Based on the classical method of a common probability space, the coupling method proved its usefulness for many practical problems.

At present, the coupling method is a widely used tool for proving various limit theorems, for Markov processes for example. The idea of the coupling method was proposed by Doeblin [1] in the first half of the twentieth century.

The idea was essentially developed starting in the 1970s in papers by Nummelin [20, 21], Nummelin and Tuominen [22], Nummelin and Tweedie [23], Lindvall [16], Kovalenko and Kuznetsov [15] for essentially multidimensional queueing systems, and Ney [19] for renewal processes.

The main definitions and results of the coupling method are presented in the book by Lindvall [17] which became a classic textbook in this topic. The notion of an abstract coupling is introduced in [17], and different constructions are described. We mention a few such constructions: weak coupling, maximal coupling, Ornstein coupling, Mineka

coupling, etc. Another well-known book devoted to the coupling method was written by Thorisson [25].

The coupling method is used in the papers by Tuominen and Tweedie and Corcoran [26] and Tweedie [27].

A splitting technique is used by Meyn and Tweedie [18] along with the coupling method. The main idea of the splitting technique is close to that of the coupling method and is developed by Nummelin [20].

Many recent papers devoted to the stability and ergodicity of Markov chains use the coupling method as the main tool (see [10] and [2]–[6]). The so-called C-coupling can be described as follows. Choose a certain set C, and assume that the chains develop independently outside C. When entering C, the chains are merged with probability α and, with probability $1 - \alpha$, they are separated according to a certain distribution. This variant of the coupling method is studied by Thorisson [25].

We briefly describe the main results obtained in the papers [10] and [2]–[6]. Jarner and Roberts [10] extend the results of the fundamental paper by Tuominen and Tweedie [26] devoted to subgeometric ergodicity. The results of the paper [10] allow one to study subgeometric ergodicity (more precisely, polynomial ergodicity as in the paper [10]) by solving certain equations with the help of test functions.

The latter results by Jarner and Roberts [10] are generalized by Douc, Moulines, and Soulier [6]. The paper [6] is devoted to studies of the stability of homogeneous Markov chains in the uniform norm and in the f-norm. Under stability, we understand here the property that the transition probabilities over n steps are closed for Markov chains with the same transition probabilities but with different initial distributions. The main tool in [6] is the C-coupling method described above.

The coupling method is used not only for homogeneous Markov chains. Douc, Moulines, and Rosenthal [4] obtained estimates of the stability for nonhomogeneous Markov chain with the help of the coupling method.

It is worth mentioning that all the papers mentioned above use the coupling method for merging two copies of a certain chain with different initial distributions. A modification of the C-coupling method is used in [8] for the first time to analyze the stability of different Markov chains. Under stability, we understand the distance between the transition probabilities over n steps for two different Markov chains if their transition probabilities are close to each other in a certain sense. The results of the paper [8] are generalized in [9] for the case of nonuniform ergodic Markov chains.

Some results concerning the coupling times for renewal processes used in the theorems on stability are presented in [14].

In the current paper, the coupling method is used to study the stability of distributions of discrete Markov chains under small perturbations of the one-step transition probabilities. We prove the stability for transition probabilities for an arbitrary number of steps.

The defining property of the maximal coupling is that, at each one-step transition time, the chains evolve pasted together with a maximal probability α being equal to the weight of the maximal common component of the corresponding distribution of transitions and the distribution of the joint evolution. The distribution of the joint evolution is determined by the common component mentioned above; the chains evolve independently of each other with probability $1 - \alpha$.

The first part of the paper contains the statements of the results and some general examples of applications.

2. Main results, corollaries, and examples

Consider a discrete space $E = \{i, j, k, \dots\}$ equipped with the σ-algebra of all its subsets $\mathcal{E} = 2^E$. Let $P = (P_{ij}, i, j \in E)$ and $P' = (P'_{ij}, i, j \in E)$ be two stochastic matrices. Denote by P_i and E_i the conditional probabilities and mathematical expectations in the probability space, respectively, and let $X = (X_n, n \geq 0)$ be a Markov chain defined in this space. Denote by P the one-step transition probabilities and by $X_0 = i$ the initial condition. The similar notation P'_i and E'_i is used for the Markov chain $X' = (X'_n, n \geq 0)$ with the matrix of one-step transition probabilities P'.

In what follows, the sum without an index of summation or without upper bound for the summation is understood as the summation over all elements of the space E. If $Q = (Q_{ij})$ is a matrix, then $Q^{(n)}$ stands for its n-th power, while $Q_i. = (Q_{ij}, j \in E)$ denotes the i-th row of the matrix Q. The symbol $x^{\pm}$ stands for the positive and negative parts of a number x, and let $\delta_{ij} = 1_{i=j}$ be the Kronecker symbol.

2.1. Uniform stability.
Denote by $\|\mu\| = \sum_j |\mu_j|$ the full variation norm in the space of summable sequences

$$l_1(E) = \{\mu = (\mu_j, j \in E) \colon \|\mu\| < \infty\}$$

and let $\mu Q_j = \sum_i \mu_i Q_{ij}$ be the product of the measure μ and matrix Q.

The property of *uniform stability over one step* is that the transition matrices P and P' become closer in the uniform metrics, namely

$$(1) \qquad \exists \varepsilon \in (0, 1) \colon \quad r(P, P') \equiv \frac{1}{2} \sup_i \|P_i. - P'_i.\| \leq \varepsilon.$$

Note that the latter condition is equivalent to the inequality $\|\mu P - \mu P'\| \leq 2\varepsilon \|\mu\|$ for all measures μ.

In what follows, we assume, for a scheme of series, that $P' \to P$ as $\varepsilon \to 0$. Note nevertheless that the results below are stated in the form of certain inequalities and thus can be used in the cases where $\varepsilon > 0$ is fixed.

The condition of *uniform mixing* means that the cross-mixing coefficient is separated from the unity, namely

$$(2) \qquad \exists \rho \in (0, 1) \colon \quad \rho(P, P') \equiv \frac{1}{2} \sup_{i \neq k} \|P_i. - P'_k.\| \leq \rho.$$

Remark 1. Condition (2) holds for sufficiently small ε if and only if P satisfies the usual uniform mixing condition

$$\rho(P, P) \leq \rho_0 < 1,$$

since $|\rho(P', P) - \rho(P, P)| \leq r(P, P')$ and $\rho(P, P') \leq \rho_0 + \varepsilon < 1$ for sufficiently small $\varepsilon > 0$. In turn, the latter condition is equivalent to the operator contraction inequality

$$(3) \qquad \|\mu P\| \leq \rho_0 \|\mu\|, \qquad \forall \mu \in l_1^0(E) \equiv l_1(E) \cap \{\mu \colon \mu(E) = 0\}.$$

Theorem 1. *Let condition* (1) *of one-step stability hold. If condition* (2) *of uniform mixing is satisfied and $\varepsilon < 1 - \rho$, then*

$$(4) \qquad \sup_{B \subset E} \left| \mathsf{P}_i(X_n \in B) - \mathsf{P}'_i(X'_n \in B) \right| \leq \varepsilon \frac{1 - \rho^n}{1 - \rho} < \frac{\varepsilon}{1 - \rho}$$

for all $n \geq 1$ and uniformly with respect to $i \in E$.

Corollary 1. *If all the assumptions of Theorem 1 hold, then*

$$(5) \qquad \sup_{i,k} \sup_{B \subset E} \left| \mathsf{P}_i(X_n \in B) - \mathsf{P}'_k(X'_n \in B) \right| \leq \rho^n + \varepsilon \frac{1 - \rho^n}{1 - \rho}$$

for all $n \geq 1$.

Corollary 2. *If all the assumptions of Theorem* 1 *hold, then the Markov chains* X *and* X' *are ergodic with invariant and limit measures* π *and* π', *respectively. Moreover*

$$\text{(6)} \qquad \sup_{B \subset E} |\pi(B) - \pi'(B)| \leq \frac{\varepsilon}{1 - \rho}.$$

The following result provides an estimate for the stability in terms of a somewhat sharper index.

Theorem 2. *Let the matrices* h *and* T *be given below by equalities* (30) *and* (35), *respectively, in the definition of a coupling chain. Assume that the chain* X *has an invariant measure* π. *If the condition of minimal mixing*

$$\text{(7)} \qquad \forall i, k \in E: \quad \lim_{n \to \infty} T^{(n)} 1_{ik} = 0$$

holds, then

$$\text{(8)} \qquad \sup_{B \subset E} \left| \mathsf{P}'_{\pi}(X'_n \in B) - \pi(B) \right| \leq \sup_{B \subset E} \sum_i \pi_i \left| \mathsf{P}_i(X_n \in B) - \mathsf{P}'_i(X'_n \in B) \right| \leq \varepsilon_T$$

for all $n \geq 1$, *where*

$$\text{(9)} \qquad \varepsilon_T \equiv \pi h (I - T)^{-1} 1.$$

If all assumptions of Theorem 1 *hold, then relation* (7) *is satisfied. Moreover condition* (8) *implies that*

$$\text{(10)} \qquad \sup_{B \subset E} |\pi(B) - \pi'(B)| \leq \varepsilon_T \leq \frac{\varepsilon}{1 - \rho}.$$

Remark 2. Condition (2) means that the coupling occurs almost surely irrespective of the initial state. The number ε_T defined by relation (9) is equal to the average time until the coupling occurs provided that the one-step decoupling occurs. According to property (1), one can expect that ε_T is small. If ε_T is finite, then inequality (8) holds. On the other hand, (8) implies that ε_T is finite.

Example 1. Let $E = \{1, 2\}$, $P_{12} = \alpha_1$, $P_{21} = \alpha_2$, $\alpha_i \in (0, 1)$, and $P'_{12} = \alpha_1 - \varepsilon d_1$, $P'_{21} = \alpha_2 - \varepsilon d_2$. Condition (1) is equivalent to $|d_i| \leq 1$. With accuracy $O(\varepsilon^2)$, inequality (6) can be rewritten as

$$\frac{\varepsilon(d_2 \alpha_1 - d_1 \alpha_2)}{(\alpha_1 + \alpha_2)^2} \leq \frac{\varepsilon}{1 - |1 - \alpha_1 - \alpha_2|}.$$

Estimate (6) is precise for this example if considered with respect to the order of magnitude of ε (the first order of magnitude for the case under consideration). Moreover, if $\alpha_1 + \alpha_2 \leq 1$, then the constant on the right hand side of the latter inequality is equal to $1/(\alpha_1 + \alpha_2)$. This is the value of the left hand side if $d_1 = -1$ and $d_2 = 1$. Thus the constant on the right hand side of (6) is precise for this particular example.

On the other hand, if $\alpha_1 + \alpha_2 > 1$, then the relative error in the constant can be arbitrarily large as $\alpha_1 + \alpha_2 \uparrow 2$. This reflects the fact that inequality (6) for limit probabilities is a corollary of the stability of prelimit probabilities (4). The latter property (4) fails for the transition matrix $P_{ij} = 1 - \delta_{ij}$ if $\alpha_1 + \alpha_2 = 2$ in view of the periodicity. Indeed, consider the perturbation $P'_{12} = 1 - \varepsilon < 1$. Then

$$\sup_{n > 0} \left(P_{11}^{(2n)} - P_{11}'^{(2n)} \right) = \sup_{n > 0} \left(1 - \frac{1 + (1 - \varepsilon)^{2n+1}}{2 - \varepsilon} \right) = \frac{1 - \varepsilon}{2 - \varepsilon}$$

and the right hand side does not approach zero as $\varepsilon \to 0$ in contrast to the one step transition probabilities.

Therefore the errors in the constant are explained by the method of the proof where the mixing coefficient has been used in a nonoptimal way to estimate the stability of the

limit probabilities. More accurate analytic methods to estimate the stability for this case are discussed in [11].

Remark 3. If some powers $P^{(m)}$ and $P'^{(m)}$ replace the matrices P and P' in the uniform mixing condition (2), then the inequalities of Theorem 1 and Corollaries 1 and 2 remain valid if the expression $r\big(P^{(m)}, P'^{(m)}\big)$ replaces ε and if the additive correction $(m-1)\varepsilon$ is used on the right hand sides. Note also that $r\big(P^{(m)}, P'^{(m)}\big) \leq m\varepsilon$.

Example 2. Let the chains X and X' satisfy the Kolmogorov condition: there exist $o \in E$, $m \geq 1$, and $d > 0$ such that $P_{io}^{(m)} \geq d$ and $P_{io}'^{(m)} \geq d$ for all $i \in E$. Then

$$\sup_{n \geq 1} \sup_i \sup_{B \subset E} \big|\mathsf{P}_i(X_n \in B) - \mathsf{P}_i'(X_n' \in B)\big| \leq \frac{\varepsilon m}{d} + \varepsilon(m-1).$$

2.2. V-stability. Let $V = (V_j, j \in E)$ be a positive test function (we do not assume that V is bounded). In what follows, we assume that

$$(11) \qquad V_i \geq 1, \qquad \sum_j P_{ij} V_j < \infty, \qquad \sum_j P_{ij}' V_j < \infty, \qquad i \in E.$$

The one-step V-stability condition means that the operator norm

$$(12) \qquad \exists \varepsilon_V > 0: \ \|P - P'\|_V \equiv \sup_i V_i^{-1} \sum_j \big|P_{ij} - P_{ij}'\big| V_j \leq \varepsilon_V$$

is small (see [11, p. 1]).

Instead of the uniform mixing condition, we use the *strong mixing condition* expressed in terms of the function V as

$$(13) \qquad \exists \rho_V < 1: \ \sum_j \big|P_{ij} - P_{kj}'\big| V_j \leq \rho_V (V_i + V_k), \qquad \forall i \neq k \in E.$$

Remark 4. In general, condition (12) does not follow from (1), and (13) does not follow from (2). Nevertheless, these two sets of conditions are equivalent if $V \equiv 1$. Moreover, Remark 1 also holds in the case under consideration; namely if $\varepsilon_V > 0$ is small, then one can impose condition (13) on the matrix P only.

Theorem 3. *Let condition* (12) *of the one-step V-stability and strong mixing condition* (13) *hold. Then*

$$(14) \qquad \sup_{|f| \leq V} \big|\mathsf{E}_i f(X_n) - \mathsf{E}_i' f(X_n')\big| \leq \varepsilon_V K_i^{(n)} \frac{1 - \rho_V^n}{1 - \rho_V}$$

for all $i \in E$ and $n \geq 1$, where

$$(15) \qquad K_i^{(n)} = \sup_{t < n} \mathsf{E}_i V(X_t).$$

Remark 5. Assumption (11) implies that the left hand side of (4) does not exceed a half of the left (right, as well) hand side of (14).

Example 3. There exist finite Markov chains for which the uniform mixing condition (1) fails but the strong mixing condition (13) holds if a test function V is chosen appropriately. Let, for example, a chain have the phase space $E = \{0, 1, 2\}$ and transition matrix

$$P = \begin{pmatrix} 0 & 1 & 0 \\ \alpha & \beta & \gamma \\ 1 & 0 & 0 \end{pmatrix}.$$

Then $r(P, P) = 1$. On the other hand, if $V = (1, 1, v)$, where a constant v belongs to the interval $(1, 1 + 2\beta/\gamma)$, then condition (13) holds with $P' = P$.

Corollary 3. *Assume that a chain X with a transition matrix P is such that there exist a set of states $O \subset E$ and constant $\rho_O < 1$ for which*

$$(16) \qquad \sum_{j \in O} P_{ij} V_j + \sum_{j \notin O} P_{ij} V_j \le \rho_O V_i, \qquad \forall i \notin O, \ V_i \ge 1, \ i \in E,$$

$$(17) \qquad \sum_j |P_{ij} - P_{kj}| V_j \le \rho_O(V_i + V_k), \qquad \forall i \in E, \ k \in O, \ i \ne k,$$

for some test function

$$V = (V_j, j \in O \cup \overline{O}).$$

If condition (12) *of the one-step V-stability holds and if $\varepsilon_V < 1 - \rho_O$, then, for all $i \in E$ and $n \ge 1$,*

$$(18) \qquad \sup_{|f| \le V} \left| \mathsf{E}_i \, f(X_n) - \mathsf{E}'_i \, f(X'_n) \right| \le \frac{\varepsilon_V K_i^{(n)}}{1 - \rho_O - \varepsilon_V},$$

where

$$(19) \qquad K_i^{(n)} = \sup_{t < n} \mathsf{E}_i \, V(X_t) \le \max \left(V_i, \sup_{i \in O} \sum_j P_{ij} \frac{V_j}{1 - \rho_O} \right).$$

Remark 6. Inequalities (16) become equalities if $E = \mathbb{Z}_+$ and if

$$V_i = \begin{cases} \mathsf{E}_i \, v^{\theta_0}, & i \notin O, \\ 1, & i \in O, \end{cases}$$

for some $v > 1$, where θ_0 is the first passage time, $\theta_0 = \inf(n \ge 1 : X_n \in O)$ (we assume that V_i is finite).

Put $\rho_O = v^{-1}$. In general, conditions (16) and (17) are written in a form that is convenient for checking them by applying the following two-step procedure. First, we fix the values $(V_i, i \in O)$ on the left hand side of (16) and solve this system of inequalities with respect to $(V_i, i \notin O)$. This can be done, for example, by using the first passage times for the set O. Second, we check inequalities (17) as explained in the following example.

Example 4. We find some sufficient conditions for the assumptions of Corollary 3 in the case where $E = \mathbb{Z}_+$ and $O = \{0\}$.

Let X be a random walk with majorized jumps,

$$(20) \qquad \exists \beta > 1 : \quad \sup_{i \ne 0} \sum_j P_{ij} \left(\beta^{j-i} + (i-j)^+ \right) < \infty.$$

We do not assume at the moment that X is spatially homogeneous. One can expect that the first passage time

$$\theta_0 = \inf(n \ge 1 : X_n = 0)$$

possesses the geometrical moment, that is, $V_i(u) = \mathsf{E}_i \, u^{\theta_0} < \infty$ for some $u > 1$. The function $V_i(u)$ is such that $V_i(u)u^{-1} = P_{i0} + \sum_{j \ne 0} P_{ij} V_j(u)$, $i \ne 0$. If we further assume that the chain is spatially homogeneous and continuous from below, then $V_i(u) = v^i$, $i \ne 0$, where $v = \mathsf{E}_1 \, u^{\tau_{10}}$ and where the first times $\tau_{i,i-1}$ when the chain reaches the states $i-1$ from the state i are independent and identically distributed. Note that

$$\theta_0 = \sum_{i=1}^{X_0} \tau_{i,i-1}.$$

Thus the function $V = (V_i = v^i, i \neq 0)$ is such that condition (16) of Corollary 3 holds with the constant $\rho_O = u^{-1} < 1$. Then we derive from (16) that $\sum_j P_{ij} V_j \leq \rho_O V_i$ for all $i \neq 0$ if $V_0 = 1$.

All the above assumptions are covered by the following one:

$$\text{(21)} \qquad \exists \beta_0 > 1: \quad \sup_{i \neq 0} \beta_0^{-i} \sum_j P_{ij} \beta_0^j < 1.$$

Condition (21) follows from (20) if

$$\text{(22)} \qquad \sup_{i \neq 0} \Delta_i < 0, \quad \text{where } \Delta_i \equiv \sum_j (j - i) P_{ij}$$

(see [11, 12]).

Since the expression in (21) is convex with respect to β_0, inequalities (21) hold for the function $V = (v^i, i \geq 0)$ for all $v \in (1, \beta_0]$.

Finally, we assume that at the moment when the chain X leaves the state $i \neq 0$, it has a positive drift to 0 with a positive speed or that the distribution of the exit from the state couples with the distribution of the exit time from 0, namely

$$\text{(23)} \qquad \sup_{i \neq 0} \min \left(\Delta_i + \Delta_0, \sum_j |P_{ij} - P_{0j}| - 2 \right) < 0.$$

Conditions (20), (21), (22), and (23) imply that all the assumptions of Corollary 3 hold and that

$$\text{(24)} \qquad \rho_i(v) \equiv (1 + v^i)^{-1} \sum_j |P_{ij} - P_{0j}| v^j \leq \rho_v, \qquad i \neq 0,$$

for all sufficiently small $v - 1 > 0$ and some $\rho_v < 1$.

This proves the stability (18).

In particular, the conditions mentioned above hold for a spatially homogeneous birth and death chain with a "soft" reflection, where $P_{00} = q$ and $P_{01} = p = 1 - q$ for $q > 2/3$. This observation follows by $\Delta_i + \Delta_0 = 2p - q < 0$, $i > 2$.

Otherwise, in the case of a "strong" reflection, where $P_{00} = 0$, $P_{01} = 1$, $q \in (0, 1)$, condition (23) fails, since $\Delta_i + \Delta_0 = p - q + 1 > 0$. Obvious reasoning shows that condition (23) is equivalent to condition (13) of Theorem 3 for each of the cases discussed above.

2.3. **The stability of finite-dimensional distributions.** The inequalities given above are related to the stability of one-dimensional marginal distributions $P_i(X_n \in B)$. It is not possible to extend this class essentially and to preserve the stability in the general case. Indeed, let P and P' be different transition matrices of irreducible chains with the uniform mixing. Consider the random events

$$A_{ij} = \left\{ \lim_{n \to \infty} n^{-1} \sum_{t < n} 1_{\{X_t = j\}} = P_{ij} \right\}.$$

The events A'_{ij} are defined similarly to A_{ij} but with X'_t in place of X_t. Then $\mathsf{P}_i(A_{ij}) = 1$ and $\mathsf{P}'_i(A'_{ij}) = 0$ if $P_{ij} \neq P'_{ij}$.

This shows that one cannot obtain the stability of probabilities for all events $A \in \sigma[X_t, t \geq 1]$ under small perturbations (1) without imposing additional restrictions.

Introduce the σ-algebras $\mathfrak{F}_n = \sigma[X_t, t \leq n]$ and $\mathfrak{F}'_n = \sigma[X'_t, t \leq n]$ in the spaces where the chains X and X' live.

Let an $(\mathfrak{F}_t)$-stopping time θ and $\mathfrak{F}_\theta$-measurable random variable φ be defined by nonrandom sets $B_n \subset E^n$ and functions $\varphi_n \colon E^n \to \mathbb{R}$ such that

$$(25) \qquad \{\theta = n\} = \{(X_1, \ldots, X_n) \in B_n\}, \qquad \varphi 1_{\{\theta = n\}} = \varphi_n(X_1, \ldots, X_n).$$

Definition. We say that a pair (θ', φ'), where θ' is an $(\mathfrak{F}'_t)$-stopping time and φ' is an $\mathfrak{F}'_{\theta'}$-measurable random variable, is related to a pair (θ, φ) if the random events $\{\theta' = n\}$ and random variables $\varphi' 1_{\{\theta' = n\}}$ are defined by the same sets B_n and functions φ_n as those involved in (25). Respectively, we say that a pair (θ', A'), where θ' is an $(\mathfrak{F}'_t)$-stopping time and A' is a random event such that $A' \in \mathfrak{F}'_{\theta'}$, is related to a pair (θ, A) if the pair $(\theta', 1_{A'})$ is related to the pair $(\theta, 1_A)$.

Theorem 4. *Assume that condition* (1) *of the one-step stability holds. We further assume that $\theta \geq 1$ is an integrable $(\mathfrak{F}_t)$-stopping time, the constants $\alpha, \beta \geq 1$ are such that $1/\alpha + 1/\beta = 1$, and that φ is a nonnegative $\mathfrak{F}_\theta$-measurable random variable such that $\varphi \in L_\alpha(\mathsf{P}_i)$. If a pair (θ', φ'), where θ' is an $(\mathfrak{F}'_t)$-stopping time and φ' is an $\mathfrak{F}'_{\theta'}$-measurable random variable, is related to a pair (θ, φ), then*

$$(26) \qquad \left| \mathsf{E}_i\, \varphi - \mathsf{E}'_i\, \varphi' \right| \leq (\varepsilon\, \mathsf{E}_i\, \theta)^{1/\beta}\, K_i^{(\alpha)}(\varphi),$$

where

$$K_i^{(\alpha)}(\varphi) = \left(\max(\mathsf{E}_i\, \varphi^\alpha, \mathsf{E}'_i\, \varphi'^\alpha) \right)^{1/\alpha}.$$

Corollary 4. *Assume that condition* (1) *of the one-step stability holds. We further assume that $\theta \geq 1$ is an integrable $(\mathfrak{F}_t)$-stopping time and A is a random event such that $A \in \mathfrak{F}_\theta$. If a pair (θ', A'), where θ' is an $(\mathfrak{F}'_t)$-stopping time and A' is a random event such that $A' \in \mathfrak{F}'_{\theta'}$, is related to a pair (θ, A), then*

$$(27) \qquad \left| \mathsf{P}_i(A) - \mathsf{P}'_i(A') \right| \leq \varepsilon\, \mathsf{E}_i\, \theta.$$

Example 5. Assume that the stability condition (1) holds, $o \in E$ is an absorbing state for the chains X and X', and θ is a stopping time such that $\theta = \inf(t \geq 1 \colon X_t = o)$. Then inequality (27) holds for all $A \in \mathfrak{F}_\infty$ and all corresponding $A' \in \mathfrak{F}'_\infty$.

Example 6. If condition (1) holds with $A \in \mathfrak{F}_m$ and with a corresponding random event $A' \in \mathfrak{F}'_m$, then

$$\left| \mathsf{P}_i(A) - \mathsf{P}'_i(A') \right| \leq \varepsilon m.$$

3. The maximal coupling of chains

Let $D = \{0, 1\}$. We introduce the one-step transition probabilities and the trajectories of the coupling Markov chain $\overline{X}$ whose states are

$$\overline{X}_n = (X_n, X'_n, d_n) \in E \times E \times D$$

in such a way that its coordinates X_n and X'_n are Markov chains with transition matrices P and P', respectively, and have the maximal probability for the coupling

$$\{X_n = X'_n\}.$$

The conditional probabilities and conditional expectations for such a chain given $\overline{X}_0 = (i, k, d)$ are denoted by P_{ikd} and E_{ikd}, respectively.

In the definitions given below, there could appear a distribution $\{a_j/a\}$ in E for the case of $a = 0$. To avoid ambiguity, we agree that the distribution in this case coincides with a fixed distribution, say δ_{oj}, $j \in E$. This choice does not influence the result, since $a = \sum a_j$, which means that the corresponding random events do not occur almost surely.

Let the initial state be denoted by $\overline{X}_0 = (i, k, \delta_{ik})$.

3.1. **The one-step decoupling.** We introduce a substochastic matrix $Q = (Q_{ij})$ and vector of weights by

$$(28) \qquad Q_{ij} = \min(P_{ij}, P'_{ij}), \qquad q_i = \sum_j Q_{ij}.$$

Consider the substochastic matrices

$$R = (R_{ij}) = P - Q, \qquad R' = (R'_{ij}) = P' - Q,$$
$$(29) \qquad R_{ij} = (P_{ij} - P'_{ij})^+, \qquad R'_{ij} = (P'_{ij} - P_{ij})^+.$$

Recall the equality $x = \min(x, y) + (x - y)^+$.

We also introduce the substochastic $E \times E^2$ matrix with the entries

$$(30) \qquad h_{i,jl} = \frac{R_{ij} R'_{il}}{1 - q_i}, \qquad i, j, l \in E.$$

Then we choose the one-step transition probabilities for the chain $\overline{X}$ from the states with $d_0 = 1$:

$$(31) \qquad \mathsf{P}_{ii1}(\overline{X}_1 = (j, l, 1)) = Q_{ij}\delta_{jl}, \qquad j, l \in E,$$

$$(32) \qquad \mathsf{P}_{ii1}(\overline{X}_1 = (j, l, 0)) = h_{i,jl}, \qquad j, l \in E.$$

A simple algebra shows that

$$\mathsf{P}_{ii1}(X_1 = j, X'_1 \in E, d_1 \in D) = P_{ij} = \mathsf{P}_i(X_1 = j),$$
$$(33) \qquad \mathsf{P}_{ii1}(X_1 \in E, X'_1 = l, d_1 \in D) = P'_{il} = \mathsf{P}'_i(X'_1 = l),$$

that is, the marginal transition probabilities for the coordinates X_1 and X'_1 are defined by the matrices P and P', respectively. Moreover, definition (29) implies that $h_{i,jj} = 0$, which in turn results in the equality for events $\{d_1 = 1\} = \{X_1 = X'_1\}$.

3.2. **The one-step coupling.** For $i \neq k$, we define the substochastic matrices and vector of weights by

$$g_{ik,j} = \min(P_{ij}, P'_{kj}), \qquad q_{ik} = \sum_j g_{ik,j},$$

$$(34) \qquad S_{ik,j} = P_{ij} - g_{ik,j} = (P_{ij} - P'_{kj})^+, \qquad S'_{ik,l} = P'_{kl} - g_{ik,l} = (P'_{kl} - P_{il})^+.$$

Further, introduce the $E^2 \times E^2$ substochastic matrix

$$(35) \qquad T_{ik,jl} = \frac{S_{ik,j} S'_{ik,l}}{1 - q_{ik}}.$$

For $d_0 = 0$, $i \neq k$, we choose

$$(36) \qquad \mathsf{P}_{ik0}(\overline{X}_1 = (j, l, 1)) = g_{ik,j}\delta_{jl},$$

$$(37) \qquad \mathsf{P}_{ik0}(\overline{X}_1 = (j, l, 0)) = T_{ik,jl}.$$

As above, the summation proves that

$$\mathsf{P}_{ik0}(X_1 = j, X'_1 \in E, d_1 \in D) = P_{ij} = \mathsf{P}_i(X_1 = j),$$
$$(38) \qquad \mathsf{P}_{ik0}(X_1 \in E, X'_1 = l, d_1 \in D) = P'_{kl} = \mathsf{P}'_k(X'_1 = l).$$

Then definition (34) implies that $T_{ik,jj} = 0$, which in turn results in the equality for events $\{d_1 = 1\} = \{X_1 = X'_1\}$.

3.3. **The construction of trajectories.** By $\mathfrak{P}$ we denote a countable family of discrete distributions on E. The distributions of $\mathfrak{P}$ are described by (31), (32), (36), and (37).

Consider a sequence of jointly independent random variables

$$\mathfrak{C} = \bigcup_{n \geq 1} \mathfrak{C}_n,$$

where the set

(39) $$\mathfrak{C}_n = \left\{ (\chi_i^n, \varkappa_{ik}^n, i, k \in E), (\lambda_n(p), \lambda_n'(p), p \in \mathfrak{P}) \right\}$$

contains jointly independent random variables such that

$$\chi_i^n, \varkappa_{ik}^n \in \{0, 1\}, \qquad \mathsf{P}(\chi_i^n = 1) = q_i, \qquad \mathsf{P}(\varkappa_{ik}^n = 1) = q_{ik},$$

(40) $$\mathsf{P}(\lambda_n(p) = j) = \mathsf{P}(\lambda_n'(p) = j) = p_j, \qquad j \in E.$$

We choose the initial state as follows: $\overline{X}_0 = (X_0, X_0', d_0) = (i, k, \delta_{ik})$, $i, k \in E$. Then we use induction for $n \geq 1$:

(0) On the set $\{X_{n-1} = i, X_{n-1}' = i, d_{n-1} = 1\}$, we choose

$$Y_n = \lambda_n(Q_i./q_i),$$
$$X_n = \chi_i^n Y_n + (1 - \chi_i^n)\lambda_n(R_i./(1 - q_i)),$$
$$X_n' = \chi_i^n Y_n + (1 - \chi_i^n)\lambda_n'(R_i'./(1 - q_i)),$$

(41) $$d_n = \chi_i^n = 1_{\{X_n = X_n'\}}.$$

(1) On the set $\{X_{n-1} = i, X_{n-1}' = k, d_{n-1} = 0\}$, $i \neq k$, we choose

$$Y_n = \lambda_n(g_{ik,.}/q_{ik}),$$
$$X_n = \varkappa_{ik}^n Y_n + (1 - \varkappa_{ik}^n)\lambda_n(S_{ik,.}/(1 - q_{ik})),$$
$$X_n' = \varkappa_{ik}^n Y_n + (1 - \varkappa_{ik}^n)\lambda_n'(S_{ik,.}'/(1 - q_{ik})),$$

(42) $$d_n = \varkappa_{ik}^n = 1_{\{X_n = X_n'\}}.$$

This definition implies that the sequence $\overline{X}$ satisfies a recurrence equation

$$\overline{X}_n = f(\overline{X}_{n-1}, \mathfrak{C}_n)$$

with a nonrandom function f. Since the random variables of the family $\mathfrak{C}$ are independent for different n, we derive that the sequence $\overline{X}$ possesses the Markov property. Further, the chain $\overline{X}$ is time homogeneous, since the vectors $\mathfrak{C}_n$ are identically distributed. Finally, we conclude that equalities (31), (32), (36), and (37) coincide by definition with the one-step transition probabilities in (41) and (42).

Taking into account relations (33) and (38) and using induction, we prove that

$$\mathsf{P}_{ikd}(X_n = j, X_n' \in E, d_n \in D) = P_{ij}^{(n)} = \mathsf{P}_i(X_n = j),$$
$$\mathsf{P}_{ikd}(X_n \in E, X_n' = l, d_n \in D) = P_{kl}'^{(n)} = \mathsf{P}_k'(X_n' = l),$$

(43) $$\mathsf{P}_{ikd}\big(\{X_n = X_n'\} \Delta \{d_n = 1\}\big) = 0$$

for all $n \geq 1$, $i, k, j, l \in E$ and $d = \delta_{ik}$.

As a result we obtain the following assertion.

Remark 7. The sequence $\widetilde{X} = (X_n, X_n') \in E^2$ is a Markov chain belonging to the family of chains $\widetilde{Y} = (Y_n, Y_n')$ such that

$$\mathsf{P}\big(Y_1 = j \mid \widetilde{Y}_0 = (i, k)\big) = P_{ij}, \qquad \mathsf{P}\big(Y_1' = l \mid \widetilde{Y}_0 = (i, k)\big) = P_{kl}'.$$

The maximal property of $\widetilde{X}$ is that

$$(44) \qquad\qquad \mathsf{P}(Y_1 = Y_1') \leq \mathsf{P}(X_1 = X_1')$$

for equal initial probabilities $\mathsf{P}(\widetilde{Y}_0 = (i, k)) = \mathsf{P}(\widetilde{X}_0 = (i, k))$.

Acknowledgement

Bibliography

1. W. Doeblin, *Exposé de la théorie des chaînes simples constantes de Markov à un nombre fini d'états*, Mathematique de l'Union Interbalkanique **2** (1938), 77–105.
2. R. Douc, G. Fort, and A. Guillin, *Subgeometric rates of convergence of f-ergodic strong Markov processes*, Stoch. Process. Appl. **119** (2009), no. 3, 897–923. MR2499863 (2010j:60184)
3. R. Douc, E. Moulines, and J. S. Rosenthal, *Quantitative bounds for geometric convergence rates of Markov chains*, Ann. Appl. Probab. **14** (2004), 1643–1664. MR2099647 (2005i:60146)
4. R. Douc, E. Moulines, and J. S. Rosenthal, *Quantitative bounds on convergence of time-inhomogeneous Markov chains*, Ann. Appl. Probab. **14** (2004), no. 4, 1643–1665. MR2099647 (2005i:60146)
5. R. Douc, E. Moulines, and P. Soulier, *Practical drift conditions for subgeometric rates of convergence*, Ann. Appl. Probab. **14** (2004), no. 4, 1353–1377. MR2071426 (2005e:60156)
6. R. Douc, E. Moulines, and P. Soulier, *Computable convergence rates for subgeometrically ergodic Markov Chains*, Bernoulli **13** (2007), no. 3, 831–848. MR2348753 (2008j:60172)
7. W. Feller, *An Introduction to Probability Theory and its Applications*, vol. 1, John Wiley & Sons, New York, 1966. MR0210154 (35:1048)
8. V. V. Golomozyĭ, *Stability of non-homogeneous Markov chains*, Visnyk Kyiv Univ., Ser. Fiz. Mat. Nauk **4** (2009), 10–15. (Ukrainian)
9. V. V. Golomozyĭ, *A subgeometric estimate of the stability for time-homogeneous Markov chains*, Teor. Imovir. Matem. Statyst. **81** (2010), 31–46; English transl. in Theor. Probability and Math. Statist. **81** (2010), 35–50. MR2667308 (2011c:60232)
10. S. F. Jarner and G. O. Roberts, *Polynomial convergence rates of Markov chains*, Ann. Appl. Probab. **12** (2001), 224–247. MR1890063 (2003c:60117)
11. N. V. Kartashov, *Strong Stable Markov Chains*, VSP/"TViMS", Utrecht/Kiev, The Netherlands/Ukraine, 1996. MR1451375 (99e:60150)
12. N. V. Kartashov, *Exponential asymptotics of matrices of the Markov renewal*, Asymptotic Problems for Stochastic Processes, Preprint 77-24, Institute of Mathematics of Academy of Science off Ukraine, Kiev, 1977, 2–43. (Russian)
13. M. V. Kartashov, *Boundedness, limits, and stability of solutions of a perturbation of a non-homogeneous renewal equation on a semiaxis*, Teor. Imovir. Matem. Statyst. **81** (2009), 65–75; English transl. in Theor. Probability and Math. Statist. **81** (2010), 71–83. MR2667311 (2011f:60154)
14. M. V. Kartashov and V. V. Golomozyĭ, *The mean coupling time for independent discrete renewal processes*, Teor. Imovir. Matem. Statyst. **84** (2011), 78–85; English transl. in Theor. Probability and Math. Statist. **84** (2012), 79–86. MR2857418 (2012f:60306)
15. I. N. Kovalenko and N. Yu. Kuznetsov, *A construction of an embedded renewal process for essentially multidimensional processes of queueing theory and application for proving limit theorems*, Preprint 80-12, Institute of Cybernetics, Academy of Science of Ukraine, Kiev, 1980. (Russian) MR612478 (82i:60142)
16. T. Lindvall, *On coupling of discrete renewal sequences*, Z. Wahrsch. Verw. Gebiete **48** (1979), 57–70. MR533006 (80g:60091)
17. T. Lindvall, *Lectures on the Coupling Method*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons Inc., New York, 1992. A Wiley-Interscience Publication. MR1180522 (94c:60002)
18. S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*, Communications and Control Engineering Series, Springer-Verlag London Ltd., London, 1993. MR1287609 (95j:60103)
19. P. Ney, *A refinement of the coupling method in renewal theory*, Stoch. Process. Appl. **11** (1981), 11–26. MR608004 (82d:60169)

20. E. Nummelin, *A splitting technique for Harris recurrent chains*, Z. Wahrsch. Verw. Gebiete **43** (1978), 309–318. MR0501353 (58:18732)

21. E. Nummelin, *General Irreducible Markov Chains and Nonnegative Operators*, Cambridge University Press, Cambridge, 1984. MR776608 (87a:60074)

22. E. Numemelin and P. Tuominen, *Geometric ergodicity of Harris recurrent Markov chains with applications to renewal theory*, Stoch. Process. Appl. **12** (1982), 187–202. MR651903 (83f:60089)

23. E. Nummelin and R. L. Tweedie, *Geometric ergodicity and R-positivity for general Markov chains*, Ann. Probab. **6** (1978), 404–420. MR0474504 (57:14143)

24. S. T. Rachev, *The Monge–Kantorovich mass transference problem and its stochastic applications*, Teor. Veroyatnost. Primenen. **29** (1984), no. 4, 625–653; English transl. in Theory Probab. Appl. **29** (1984), no. 4, 647–676. MR773434 (86m:60026)

25. H. Thorisson, *Coupling, Stationarity, and Regeneration*, Springer, New York, 2000. MR1741181 (2001b:60003)

26. P. Tuominen and R. Tweedie, *Subgeometric rates of convergence of f-ergodic Markov chains*, Adv. Appl. Probab. **26** (1994), 775–798. MR1285459 (95m:60097)

27. R. L. Tweedie and J. N. Corcoran, *Perfect sampling of ergodic Harris chains*, Ann. Appl. Probab. **11** (2001), no. 2, 438–451. MR1843053 (2002g:60111)

28. V. M. Zolotarev, *Modern Theory of Summation of Independent Random Variables*, "Nauka", Moscow, 1986; English transl., VSP, Utrecht, the Netherlands, Tokyo, Japan, 1997.

DEPARTMENT OF PROBABILITY THEORY, STATISTICS, AND ACTUARIAL MATHEMATICS, FACULTY FOR MECHANICS AND MATHEMATICS, NATIONAL TARAS SHEVCHENKO UNIVERSITY, ACADEMICIAN GLUSHKOV AVENUE, 4E, KIEV 03127, UKRAINE

E-mail address: nkartashov@skif.com.ua

DEPARTMENT OF PROBABILITY THEORY, STATISTICS, AND ACTUARIAL MATHEMATICS, FACULTY FOR MECHANICS AND MATHEMATICS, NATIONAL TARAS SHEVCHENKO UNIVERSITY, ACADEMICIAN GLUSHKOV AVENUE, 4E, KIEV 03127, UKRAINE