Maximal Coupling Procedure and Stability of Continuous-time Markov Chains

Mokaedi V. Lekgari

School of Mathematics and Statistics, Wuhan University, Wuhan 430072, China
November 13, 2014

Keywords: Stability; Markov Chains; Coupling.

Abstract. In this study we first investigate the stability of subsampled discrete Markov chains through the use of the maximal coupling procedure. This is an extension of the available results on Markov chains and is realized through the analysis of the subsampled chain $\Phi_{T_n}$, where \( \{T_n, n \in \mathbb{Z}_+\} \) is an increasing sequence of random stopping times. Then the similar results are realized for the stability of countable-state Continuous-time Markov processes by employing the skeleton-chain method.

1. Introduction

Loosely speaking, coupling may be referred to as the practice of constructing two(or more) probability measures(or random processes) within a single measurable space for the purposes of investigating similarities and individual characteristics of each of the coupled measures. Coupling techniques which include maximal coupling(also known as Vasershtein coupling) and other forms of coupling like weak coupling(often referred to as distributional coupling) amongst others have had a very profound influence in the studies of stability and ergodicity of Markov processes. Of particular interest to our study is the maximal coupling technique which appeared in Griffeth [5] and has since been employed in its various forms by many researchers amongst them Connor [2], Roberts and Rosenthal [10] and Kartashov and Golomozyi [7]. We first discuss discrete Markov chains(DTMCs) and important information about coupling in general before going into the maximal coupling specifics and later countable-state Continuous-time Markov processes.

We let $\mathbb{Z}_+ = \{0,1,2,\ldots\}$, $\mathbb{N}_+ = \{1,2,\ldots\}$ and $\mathbb{R}_+ = [0,\infty)$, then we denote by $(\Phi_n)_{n \in \mathbb{Z}_+}$ a DTMC on a state space $(\mathcal{X},\mathcal{B}(\mathcal{X}))$ with $q$-matrix $Q = \{q_{ij}; i,j \in \text{mathcal}X\}$. $Q^{(k)}$ denotes the $k$-th power of this matrix. We also let $\mu Q = \Sigma_j \mu_j Q_{ij}$ be the product of the measure $\mu$ and the matrix $Q$. The transition function of the DTMC is denoted by $P^n_{ij} = P_i(\Phi_n = j) = E_i[1_{\Phi_n = j}]$, where here and hereafter $1_C$ is the indicator function of set $C$ and $P_i$ and $E_i$ respectively denote the probability and expectation of the chain under the condition that $\Phi_0 = i$. For a chain with 1-step transition probabilities $P'$ the similar notation of $P'$ and $E'$ respectively will be used for the chain $\Phi'_n$.

Let $T$ be the random time such that $\Phi_n = \Phi'_n$ for $n \geq T$. Further let $\hat{\Phi} = (\Phi_n, \Phi'_n, d_n) \in X \times X \times D$ be the coupling of $\Phi_n$ and $\Phi'_n$, then for the coupling time $T$ we get the coupling inequality

$$||P(\Phi_n \in (\cdot)) - P'(\Phi'_n \in (\cdot))|| \leq P(T > n),$$

(1)

where $D$ is a sequence of random variables on $\{0,1\}$ (independent of both $\Phi_n$ and $\Phi'_n$), and the variable $d_n$ is set equal to 0 starting from $T$.

Coupling is said to be successful if $P(T < \infty) = 1$ and $\{T > n\}$ is called the coupling event. We note that chain $\Phi_n$ is said to be weakly ergodic if $\lim_{n \to \infty} \sum_{k \in \mathcal{X}} |\hat{\delta}_i P^n(k) - \delta_i P^n(k)| = 0, \forall i,j \in \mathcal{X}$, where $\hat{\delta}_i$ is the Dirac point mass at $i$. Which implies that $\Phi_n$ is weakly ergodic if for any pair of initial states $i,j \in \mathcal{X}$ successful coupling exists. Maximal coupling is achieved when we get
equality instead of the inequality in (1). The interested reader can consult Lindvall [8] for more study on coupling.

Kartashov and Golomozy [7] studied the stability of distributions of discrete chains under small perturbations of 1-step transition probabilities, then proved for an arbitrary number of steps, the stability for the transition probabilities. This was achieved through the use of the maximal coupling technique. In this study, we prove that the stability holds not only for discrete Markov chains as in [7], but also for the subsampled chain \( \Phi_{T_n} \), where \( \{T_n, n \in \mathbb{Z}_+\} \) is a non-decreasing sequence of stopping times, with \( T_0 = 0 \). Then prove that the same results hold for Continuous-time Markov chains.

The paper is organized as follows. In Section 2, we present some minor results similar to previous contributions on stability with focus on the stability of Discrete-time Markov chains under maximal coupling and the results of [7] and related references will play a pivotal role here. The main results of our study are availed in Section 3. In Section 4 we have the conclusions. The Appendix is the 5th Section which contains the proofs.

2. Maximal Coupling and \( \Phi_{T_n} \) Stability

Theorem 1. [Theorem 4 in [5]] Any Markov chain \( \Phi_n \) has a maximal coupling, that is, it achieves equality in formula (1). Thus the maximal coupling \( \Phi_{\Pi_n} \) is successful if and only if the chain \( \Phi_n \) is weakly ergodic.

As noted in [7], at each 1-step transition time, for maximal coupling the chains evolve coupled with maximal probability of which is equal to the weight of the maximal common component of the corresponding transition distributions and the joint evolution distributions. Further the chains evolve independently of each other with probability \( 1 - \alpha \). According to Theorem 1 we proceed with the rest of the study with the understanding that existence of maximal coupling is a given. Assumption 2 below also reinforces the fact that irrespective of the initial state coupling occurs almost surely. We desire to prove in this Section that the stability holds for the subsampled chain \( \Phi_{T_n} \), where \( \{T_n, n \in \mathbb{Z}_+\} \) is an increasing sequence of random stopping times.

Assumption 1. There exists \( \varepsilon \in (0,1) \) such that

\[
\varepsilon(P,P') \equiv \frac{1}{2} \sup_i |P_i - P'_i| \leq \varepsilon.
\]

Assumption 2. There exists \( \rho \in (0,1) \) such that

\[
\rho(P,P') \equiv \frac{1}{2} \sup_{i \neq k} |P_i - P'_k| \leq \rho.
\]

Assumptions 1 and 2 are respectively referred to as 1-step uniform stability condition and uniform mixing condition.

Theorem 2. Suppose that both Assumptions 1 and 2 hold and \( \varepsilon < 1 - \rho \), then for some(and then for all) random stopping time \( T_n, n \in \mathbb{N}_+ \) we have

\[
\sup_{c \in \mathcal{C}} |P_{\Phi_n}(\Phi_{T_n} \in \mathcal{C}) - P'_{\Phi_n}(\Phi'_{T_n} \in \mathcal{C})| < \frac{\varepsilon}{1 - \rho}.
\]
Proof
See Proof 1 in the Appendix.

Theorem 2 is a modified version of Theorem 1 of [7]. Consequently we get the following Corollary.

**Corollary 1.** Suppose that both Assumptions 1 and 2 hold and \( \varepsilon < 1 - \rho \), then for some (and then for all) random stopping time \( T_n, n \in \mathbb{N}_+ \) we have

\[
\sup_{i, k} \sup_{C \in \mathcal{C}} \left| P_i(\Phi_{T_n} \in C) - P_i' k(\Phi_{T_n} \in C) \right| < \rho T_n + \frac{\varepsilon}{1 - \rho}.
\]  

(3)

**Assumption 3.** For the positive test function \( V, j \in \mathcal{X} \), there exists \( \varepsilon_V > 0 \) such that

\[
\left| P - P' \right|_V \equiv \sup_i V_i^{-1} |P_{ij} - P'_{ij}| V_j \leq \varepsilon_V
\]

is small. Where \( V_i \geq 1, \sum_j P_{ij} V_j < \infty \) and \( \sum_j P'_{ij} V_j < \infty, i \in \mathcal{X} \).

**Assumption 4.** There exists \( \rho_V < 1 \) such that

\[
\sum_j |P_{ij} - P'_{kj}| V_j \leq \rho_V (V_i + V_k), \quad \forall \ i \neq k \in \mathcal{X}.
\]

We refer to Assumptions 3 and 4 as the 1-step V-stability condition and the strong mixing condition respectively. Next we state Theorem 3 which is a modified version of Theorem 3 in [7].

**Theorem 3.** Suppose that both Assumptions 3 and 4 hold, then for some (and then for all) random stopping time \( T_n, n \in \mathbb{N}_+ \) we have

\[
\sup_{f \in \mathcal{C}} \left| E_i f(\Phi_{T_n}) - E_i f(\Phi'_{T_n}) \right| \leq \varepsilon_V K_i^{(T_n)} \frac{1 - \rho_T}{1 - \rho_V}
\]

\( \forall \) initial values \( i = \Phi_0 \in \mathcal{X} \) where

\[
K_i^{(T_n)} = \sup_{t < T_n} E_i V(\Phi_t).
\]

3. Main Results

1. CTMCs

We suppose that for the chain \( \Phi_n \) there exists a state \( i \in \mathcal{X} \) such that \( P_i(\xi_i < \infty) = 1 \), where \( \xi_i = \inf \{ t \geq 0: \Phi_t = i \} \), so that the sojourn time in state \( i \) has an exponential distribution with some parameter \( \lambda \in (0, \infty) \) owing to homogeneity and Markovian properties of the process. Thus state \( i \) can be viewed as having the same behavior as a state in a countable-state Continuous-time Markov process. We let \( (\Phi_{nh})_{n \in \mathbb{N}_+} \), for some \( h \in \mathbb{R}_+ \), be the \( h \)-skeleton chain of the Continuous-time Markov Chain (CTMC) \( (\Phi_n)_{n \in \mathbb{N}_+} \). By setting \( t = nh \) we show that the results of DTMCs in Section 2 can be extended to CTMCs in this study (see Proposition 1).

Analogous to DTMCs we denote by \( (\Phi_t)_{t \geq 0} \) a CTMC on a countable state space \( (\mathcal{X}, \mathcal{B}(\mathcal{X})) \), with transition function denoted by \( P^t(i,j) = P_i(\Phi_t = j) = E_i [1_{\Phi_t = j}] \). The rest of the notation follows in a
similar manner. For the CTMCs we define non-decreasing sequence of stopping times \( \{ \tau_n : n \in \mathbb{Z}_+ \} \) with \( \tau_0 = 0 \) in a way similar to the Assumption 2.1 of Spieksma [11] as follows. Recursively we define \( \tau_n = 0 \) and \( \tau_{n+1} = \inf \{ t > \tau_n : \Phi_t \neq \Phi_{\tau_n} \} \), if \( \Phi_{\tau_n} \) is not an absorbing state (that is, \( q_{\Phi_{\tau_n}} \neq 0 \)). We put \( \tau_n, k > n \) if \( \Phi_{\tau_n} \) is an absorbing state then we have \( \Phi_{\tau_n} = \Phi_{\tau_n} \). In this sense the sequence \( \{ \tau_n \}_{n \in \mathbb{Z}_+} \) is a non-decreasing sequence of stopping times representing successive jump times.

We expect the distance \( ||\mu_P|| \) to be a decreasing function of \( t \in \mathbb{R}_+ \). We also note that \( p_t^{nh}(i) \geq e^{-th} \), for any \( h > 0 \) thus \( i \) is an atom and the singleton set \( \{ i \} \) is a small set for the chain \( \Phi_{nh} \). Thus some small set \( C \) exists for the chain \( \Phi_{nh} \). Further we conclude that all the assumptions that we had for the discrete case also hold for the chain \( \Phi_{nh} \).

**Proposition 1.** Let all the conditions of Theorem 2 hold then for any \( h \in \mathbb{R}_+ \) and for some (and then for all) stopping times \( \{ \tau_n \}_{n \in \mathbb{N}_+} \) we have,

\[
\sup_{C \in \mathcal{X}} || P_{\Phi_n}(\Phi_{\tau_n} \in C) - P'_{\Phi_n}(\Phi'_{\tau_n} \in C) || \leq \frac{\varepsilon}{1 - \rho}.
\]  

**Proof**

See Proof 2 in the Appendix.

From Proposition 1 it is obvious that the rest of the DTMCs results like Theorem 3 follow in a similar manner, hence the omission of such results. Furthermore we give the following results which are closely related to the CTMCs in this study.

2. **V-norm Stability**

The results that follow stated in the form of Propositions are focused on the \( V \)-norm. We did not assume in Assumptions 3 that the function \( V \geq 1 \) is bounded either on or on \( \mathcal{X} \) the small set \( C \). In the CTMCs results below we assume \( \sup_C V < \infty \). Further we are satisfied to confirm similarities of our results with the results of authors like [1], [7] etc.

**Proposition 2.** From the result of [Corollary 2 in [7]] we know that if the chains \( \Phi_x \) and \( \Phi'_x \) are ergodic with positive invariant limit measures \( \mu_x \) and \( \mu'_x \) then \( ||\mu_x - \mu'_x||_V \leq q < 1 \) where \( q = ace1 - \rho \). Then for some increasing function \( \varphi: \mathbb{R}_+ \rightarrow \mathbb{R}_+ \) such that \( t \mapsto \varphi(t)/t \) and for any function \( V: \mathcal{X} \rightarrow [1, \infty) \) we have

\[
||\mu_x - \mu'_x||_V \leq q \varphi^{-1}(\frac{\int (\varphi \circ V) (dx)}{q} + q_{\mu_x} + q_{\mu'_x})(dx).
\]  

**Proof**

See Proof 3 in the Appendix.

**Proposition 3.** [Borrowed from Proposition 2.2 in Connor and Fort[3]] Let the transition kernels \( P \) and \( P' \) be \( \Psi \)-irreducible and aperiodic. For a small set \( C \), measurable functions \( \chi, V: \mathcal{X} \rightarrow [1, \infty) \), and a konstant \( b < \infty \) such that \( \sup_C V < \infty \) and

\[
P V_i \leq V_i - \chi_i + b 1_{[\text{sec}]} \\
P \chi_i \leq \chi_i + b 1_{[\text{sec}]}
\]  

then there exist a constant \( \kappa < \infty \) such that for any pair of initial states \( i, k \) and for the increasing sequence of stopping times \( \{ \tau_n \}_{n \in \mathbb{N}_+} \) we have
If we have a continuously differentiable increasing concave function \( \varphi: [1, \infty) \to (0, \infty) \) with \( \chi \propto \varphi \circ V \) then we have a subgeometric drift inequality for the first inequality of (7).

**Proposition 4.** Let \( \Phi_t \) and \( \Phi'_t \) be CTMCs and let the measurable functions \( \chi, V: \mathcal{X} \to [1, \infty) \) be such that \( PV_i \leq V_i - \chi_i + b1_{[i \in \mathcal{E}]} \). Then for some positive constant \( \beta \) there exists a function \( \rho(t) = \beta / \chi(t) \) such that for any two initial states \( i, k \) and for the increasing sequence of stopping times \( T_n, n \in \mathbb{N}_+ \) we have

\[
||P_i(\Phi_{T_n} \in (\cdot)) - P'_k(\Phi'_{T_n} \in (\cdot))||_\chi \leq \kappa [V_i + V_k].
\]

**Proof**

See Proof 4 in the Appendix.

For the sake of consistency we note that inequality (9) still holds for the norm \( \chi \neq 1 \). The following Proposition 5 confirms the consistency of the Propositions 3 and 4. It generalizes Theorem 3.6 of [1].

We define the pair of inverse Young functions \( \Psi_1, \Psi_2: \mathcal{X} \to [1, \infty) \). All pairs \( \Psi_1, \Psi_2 \), satisfy \( \Psi_1(i)\Psi_2(j) \leq i + j \) and \( \Psi_1(i), \Psi_2(j) \to \infty \) as \( i, j \to \infty \). Let \( r: \mathbb{R}_+ \to [1, \infty) \) be a subgeometric rate function. For more on subgeometric rate functions the interested reader can consult [9]. We note that if \( \lim_{t \to \infty} \psi'(t) = 0 \) then \( r_\varphi \) is also a subgeometric rate function and so \( \Psi_1 \circ r_\varphi \) is provided \( \Psi_1(i) \leq \delta i \) for some constant \( \delta \).

**Proposition 5.** Let \( \Phi_t \) and \( \Phi'_t \) be aperiodic and irreducible CTMCs and let the measurable functions \( \varphi: [1, \infty) \to (0, \infty) \), \( V: \mathcal{X} \to [1, \infty) \), be such that \( PV_i \leq V_i - \varphi \circ V_i + b1_{[i \in \mathcal{E}]} \) and \( \lim_{t \to \infty} \varphi'(t) = 0 \). Then there exists some positive constant \( \rho_\varphi \) such that for the pair of inverse Young functions \( \Psi_1, \Psi_2 \), any two initial states \( i, k \), and for the increasing sequence of stopping times \( T_n, n \in \mathbb{N}_+ \) we have

\[
\Psi_1(r_\varphi(T_n)) ||P_i(\Phi_{T_n} \in (\cdot)) - P'_k(\Phi'_{T_n} \in (\cdot))||_{\Psi_2(\varphi \circ V)} \leq \rho_\varphi [V_i + V_k].
\]

**4. Conclusions**

We note that the CTMCs results in this study hold also for \( \Phi_t \), where \( \Phi_t \) are continuous-time processes in the space of cadlag functions defined on \( \mathbb{R}_+ \) with values in the Polish space \( \mathcal{X} \), endowed with the Skorokhod topology. Also extending the results to general state Markov processes follows naturally because the coupling technique is the one tool for such an endeavor.

In the future we would like to consider ‘\( J_1 \) – step uniform stability’ and ‘\( J_1 \) – step \( V \) – stability condition’ as defined below and find out how they affect the results.

**Definition 1.** We define the ‘\( J_1 \) – step uniform stability’ as follows. Instead of the 1-step probability transitions as in Assumption 1 we now assume a 1 – jump probability transition for the uniform stability condition.

**Definition 2.** Similarly we define the ‘\( J_1 \) – step \( V \) – stability condition’ as follows. Instead of the 1-step probability transitions as in Assumption 3 we now assume a 1 – jump probability transition for the \( V \)-stability condition.
We would like to know if the $J_1$–step uniform stability and Assumption 2 hold and $\varepsilon < 1 - \rho$, then for all jumps $J_{n, n} \in N_+$

$$
\sup_{c \in \mathcal{X}} |P_{\Phi_{J_n}}(\Phi_{J_n} \in C) - P'_{\Phi_{J_n}}(\Phi_{J_n} \in C)| \leq \frac{\varepsilon}{1 - \rho},
$$

(11)

uniformly and if this is the same as Theorem 2. Similarly we would like to know if Theorem 3 follows naturally by assuming $J_1$–step $V$–stability condition and strong mixing condition. As already mentioned in [7], in general, the 1-step $V$-stability condition doesn’t follow from the -step uniform stability condition, though these conditions are equivalent for $V \equiv 1$. Consequently, we interested at similar conclusions for the $J_1$-step uniform stability and $J_1$-step $V$-stability.

We are also interested in investigating the stability of finite-dimensional distributions instead of only the 1-dimensional distributions as done in this study.

5. Appendix

**Proof 1.** *(For Theorem 2)* Suppose that $T_n = n, \forall n \in N_+$, then the proof is trivial. For the case when $T_n \neq n$, we have for some $n$ that $T_n = \eta_n$, where $\eta_n$ is some arbitrary number of steps for the chain $\Phi_n$. Also in this latter case the proof is trivial because $\eta_n \in N_+$ also.

**Proof 2.** *(For Proposition 1)* Let $t \in (0, \infty)$ be such that for some $n \in N_+$ then $(T_{n-1})h \leq t \leq T_nh$ for some $h>0$. Then we know for any measure $\mu$ that $||\mu P||$ is monotonic because the function $t \mapsto ||\mu P_{(t)}||$ a decreasing function. Therefore,

$$
sup_{t} |\mu P_{(t)} - \mu P'_{(t)}| \leq ||\mu P_{(T_{n-1}h)} - \mu P'_{(T_{n-1}h)}||
\leq 2\varepsilon||\mu||
$$

which confirms Assumption 1. It then remains only to prove that Assumption 2 also holds. Again because of the monotonicity of $||\mu P||$ we have for any measure $\mu$ and any pair of initial starting values $i, k$ that

$$
sup_{t} |\mu P_{i}(\Phi_t | \cdot) - \mu P'_{k}(\Phi'_t | \cdot)| \leq ||\mu P_{i}(\Phi_{T_{n-1}h} | \cdot) - \mu P'_{k}(\Phi'_{T_{n-1}h} | \cdot)||
\leq ||\mu||e/(1 - \rho) + (1 - \varepsilon(1 - \rho))\rho^n
< ||\mu||
$$

(12)

because $\varepsilon < 1 - \rho$. Formula (12) implies that

$$
sup_{i, k} |\mu P_{i}(\Phi_{T_{n-1}h} | \cdot) - \mu P'_{k}(\Phi'_{T_{n-1}h} | \cdot)| \leq \rho, \text{ for some } \rho \in (0, 1).
$$

(13)

Thus Assumption 2 is satisfied.

**Proof 3.** *(For Proposition 2)* This result is not altogether new [see Section 4.4 in [6]]. By definition of $||\cdot||_p$ for some measurable function $V: X \rightarrow [1, \infty)$ and some constant $R>0$ we have

$$
||\mu - \mu'||_V \leq R||\mu - \mu'||_TV + \int_{V(x)>R} V(x)(\mu - \mu')(dx)
\leq RQ + \int_{V(x)>R} V(x)(\mu + \mu')(dx)
< \infty.
$$
Then for $t > R$ we have $t \leq \varphi(t) \frac{R}{r(R)}$ because $t \mapsto t/\varphi(t)$ is increasing. Thus for the formula

$$||\mu_s - \mu_s'||_{tv} \leq RQ + \int_{\mathcal{X}} (\varphi \circ V)(x)(\mu_s + \mu_s')(dx) \frac{R}{\varphi(R)}$$

optimizing over $R$ yields the desired result

$$||\mu_s - \mu_s'||_{tv} \leq q\varphi^{-1}\left(\int_{\mathcal{X}} (\varphi \circ V)(x)(\mu_s + \mu_s')(dx)\right).$$

\[\square\]

**Proof 4.** (For Proposition 4) Let $\zeta_C = \min \{ t \geq 0 : \Phi_{\tau_n} \in \mathcal{C} \}$. That is, we let $\mathcal{T}_n$ denote the $n^{th}$ hitting time on the small set $C$, with $\mathcal{F}_n$ as the $\sigma$-algebra of events generated by $\mathcal{T}_n$. Then by Dynkin’s inequality for all $\Phi_{\tau_n} \in \mathcal{C}$ we get

$$E\left[\sum_{n=0}^{\zeta_C-1} \chi(\Phi_{\tau_n})|\mathcal{F}_n\right] \leq V(\Phi_{\tau_C}) + bE\left[\sum_{n=0}^{\zeta_C-1} 1_{\Phi_{\tau_n} \in \mathcal{C}}\right]$$

$$\leq V(\Phi_{\tau_C}) + b$$

$$< \infty$$

(15)

because $\sup_C V < \infty$. We also note that $\zeta_B \leq \zeta_B$ for all accessible sets $B \in \mathcal{B}^+(\mathcal{X})$, because the chain can miss some of the visits to the set $B$. Thus,

$$E[\chi(\zeta_B), E[\chi(\zeta_B)] \leq \infty$$ or $$E[\chi(\zeta_B)] \leq E[\chi(\zeta_B)] \leq \infty \forall B \in \mathcal{B}^+(\mathcal{X})$$ (for the subgeometric case).

(16)

Which confirms that maximal coupling (that is, formula (1) holds) for the chains $\Phi_{\tau_n}$ and $\Phi_{\tau_n}$. Combining (16) with $||P_t(\Phi_{\tau_n} \in \cdot) - P'_k(\Phi_{\tau_n} \in \cdot)|| \leq \beta \{V_i + V_k\}$ from Proposition 3 for some constant $\beta > 0$ and $t \in \mathbb{R}_+$ we get

$$\int_0^\infty x'(t)||P_t(\Phi_{\tau_n} \in \cdot) - P'_k(\Phi_{\tau_n} \in \cdot)||dt \leq \beta \{V_i + V_k\}$$

(17)

which implies

$$||P_t(\Phi_{\tau_n} \in \cdot) - P'_k(\Phi_{\tau_n} \in \cdot)|| \leq \frac{\beta}{x(\tau_n)} \{V_i + V_k\},$$

for some $\tau_n, n \in \mathbb{N}_+$. (18)

6. Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this article.
References


