ДОКЛАДЫ РОССИЙСКОЙ АКАДЕМИИ НАУК. МАТЕМАТИКА, ИНФОРМАТИКА, ПРОЦЕССЫ УПРАВЛЕНИЯ, 2020, том 490, с. 16–19

МАТЕМАТИКА

УДК 519.217

О СКОРОСТИ СХОДИМОСТИ ДЛЯ ОДНОРОДНЫХ ЦЕПЕЙ МАРКОВА

© 2020 г. А. Ю. Веретенников[1,2,3,*], М. А. Веретенникова[2,]**

Представлено академиком РАН А.Н. Ширяевым 21.06.2019 г.
Поступило 01.07.2019 г.
После доработки 01.07.2019 г.
Принято к публикации 31.10.2019 г.

Изучаются улучшенные оценки скорости сходимости для эргодических однородных цепей Маркова. Даны примеры сравнения с классическими оценками.

Ключевые слова: цепи Маркова, эргодичность, скорость сходимости

DOI: 10.31857/S2686954320010087

Классическая оценка скорости сходимости для однородной дискретной неприводимой ациклической цепи Маркова $(X_n,\ n \ge 0)$ с конечным фазовым пространством S к единственной инвариантной мере μ имеет вид

$$\|\mu_n - \mu\|_{TV} \le 2(1-\kappa)^n \quad \forall n, \tag{1}$$

и она убывает экспоненциально быстро при условии Маркова–Добрушина (см. [1])

$$\kappa = \min_{i,i'} \sum_{j\in S} p_{ij} \wedge p_{i'j} > 0,$$

равномерно относительно начального распределения, где p_{ij} – переходные вероятности, $\|\cdot\|_{TV}$ – расстояние между мерами по вариации. Аналогичная оценка справедлива и для общих цепей Маркова (см, например, [2]). Другой, “спектральный” способ оценки скорости сходимости основан на собственных числах переходной матрицы, см. [3, глава XIII, (96); 1, теорема 1.2]. Для общего фазового пространства этот метод может быть применен, если собственные функции переходного оператора образуют полный базис в пространстве $L_2(S)$. В работах [2, 4] было предложено использовать определенный способ марковской склейки (каплинга – см. [5, 6] и др.) для получения лучшей оценки скорости сходимости, чем (1). Целью данной работы является отказ от технического предположения в [2, 4] о наличии единой доминирующей меры (в [4] это мера Лебега) для всех переходных ядер $Q(x,\cdot) = P(X_1 \in \cdot | X_0)|_{X_0=x}$ и построение примеров. Заметим, что если спектральный метод применим, то он дает наилучшую возможную асимптотическую скорость сходимости. В то же время оценка (1) справедлива и в общих фазовых пространствах [2, 4], когда использование спектрального метода может быть затруднено. Новый подход всегда гарантирует как минимум такую же, но, как правило, лучшую асимптотику по сравнению с оценкой типа (1), и в ряде примеров его асимптотика совпадает с асимптотикой спектральной оценки. Есть все основания думать, что новый метод применим к неоднородным цепям, в отличие от спектрального.

ЭРГОДИЧЕСКАЯ ТЕОРЕМА, КАПЛИНГ, ОБЩИЙ СЛУЧАЙ

Рассматривается однородный марковский процесс (МП) в дискретном времени $(X_n,\ n \ge 0)$ на общем фазовом пространстве S с топологией и борелевской σ-алгеброй $\mathcal{B}(S)$. Индекс x в выражениях $\mathbb{E}_x$ и $\mathbb{P}_x$ означает неслучайное начальное значение процесса $X_0 = x$; оно может быть также случайным с распределением μ; в таком случае могут использоваться обозначения $\mathbb{E}_\mu$ и $\mathbb{P}_\mu$ (см. [7]). Если фазовое пространство S конечно, то $|S|$ обозначает число его элементов, и тогда $\mathcal{P} := (p_{ij})_{1\le i,j\le |S|}$ – матрица переходных вероятностей процесса. Об истории оценки Маркова–Добрушина (МД) см. [1] и др.

Для общей оценки (4), приведенной ниже ради удобства читателя, предполагается положительность постоянной Маркова–Добрушина:

$$\kappa := \inf_{x,x'} \int \left(\frac{P_{x'}(1,dy)}{P_x(1,dy)} \wedge 1 \right) P_x(1,dy) > 0. \tag{2}$$

[1] *Университет Лидса, Лидс, Великобритания*
[2] *Национальный исследовательский университет “Высшая школа экономики”, Москва, Россия*
[3] *Институт проблем передачи информации им. А.А. Харкевича Российской академии наук, Москва, Россия*
**E-mail: a.veretennikov@leeds.ac.uk*
***E-mail: mveretennikova@hse.ru*

В основном результате работы (теорема 1) это условие будет заменено на более слабое. В (2) выражение $\frac{P_{x'}(1,dy)}{P_x(1,dy)}$ понимается в смысле плотности абсолютно непрерывной компоненты одной меры по другой. Для краткости будем использовать сокращенную нотацию $P_x(dz)$ вместо $P_x(1,dz)$. Отметим, что интеграл от любой борелевской функции по ядру $P_x(dz)$ является борелевской функцией от x; это стандартное требование для марковского процесса, см. [7]. В силу линейности аналогичная измеримость по паре (x,x') будет иметь место и для интеграла от борелевской функции с мерой $\Lambda_{x,x'}$, определенной ниже формулой

$$\Lambda_{x,x'}(dz) := P_x(1,dz) + P_{x'}(1,dz).$$

Отметим, что $\Lambda_{x,x'}(dz) = \Lambda_{x',x}(dz)$. Справедливо следующее эквивалентное представление постоянной из формулы (2):

$$\kappa = \inf_{x,x'}\int\left(\frac{P_{x'}(1,dy)}{\Lambda_{x,x'}(dy)} \wedge \frac{P_x(1,dy)}{\Lambda_{x,x'}(dy)}\right)\Lambda_{x,x'}(dy) \quad (>0). \tag{3}$$

Положим

$$\kappa(x,x') := \int\left(\frac{P_{x'}(1,dy)}{P_x(1,dy)} \wedge 1\right)P_x(1,dy).$$

Имеет место равенство $\kappa(x,x') = \kappa(x',x)$ для любых $x,x' \in S$. Условие (2) введено самим А.А. Марковым для конечных однородных цепей [8]; позднее для неоднородных процессов Маркова его аналог был использован Добрушиным [9]. По предложению Сенеты оно названо условием Маркова–Добрушина. Заметим, что во всех ситуациях $\kappa \le 1$. Случай $\kappa = 1$ соответствует последовательности независимых, одинаково распределенных случайных величин $(X_n,\ n \ge 0)$. В противоположной экстремальной ситуации, когда переходные ядра сингулярны для различных x и x', имеем $\kappa = 0$. Условие МД (2), так же, как и (3), исключительно полезно тем, что обеспечивает эффективную количественную оценку сверху на скорость сходимости цепи Маркова к его (единственной) инвариантной мере в метрике полной вариации. Имеет место следующий классический результат: в предположении условия (3) процесс (X_n) является марковским-эргодическим, т.е. для него существует предельная и она же инвариантная вероятностная мера μ, причем справедлива равномерная оценка при всех n

$$\sup_x \sup_{A\in\mathcal{B}(S)} \left|P_x(n,A) - \mu(A)\right| \le (1-\kappa)^n. \tag{4}$$

Далее, напомним, как применить схему каплинга на основе “леммы о двух случайных величинах” [2, Лемма 44; 4, Лемма 1] к цепи Маркова в общем фазовом пространстве $(S,\mathcal{B}(S))$. С методом каплинга можно познакомиться по работам [5, 6, 10, 12] и др. Данная работа следует изложению в [4], которое, в свою очередь, основано на подходе [6]. Заметим, что фазовым пространством в [4] было $\mathbb{R}^1$; однако и в $\mathbb{R}^d$ все формулы остаются такими же.

Итак, рассмотрим две копии (X_n^1) и (X_n^2) одного и того же марковского процесса с начальными распределениями, соответственно, μ_0^1 и μ_0^2 (это не исключает случай неслучайных начальных значений). Введем новый векторнозначный марковский процесс $(\eta_n^1,\eta_n^2,\xi_n,\zeta_n)$. Если $\kappa_0 = 0$, то положим

$$\eta_0^1 := X_0^1, \quad \eta_0^2 := X_0^2, \quad \xi_0 := 0, \quad \zeta_0 := 1.$$

Если $0 < \kappa(0) < 1$, то применим лемму [4, Лемма 1] или [2, Лемма 44] с $\Lambda = \Lambda_{x,x'}$ при “условии” $(X_0^1, X_0^2) = (x,x')$ к случайным величинам X_0^1 и X_0^2, построив таким образом случайные величины η_0^1, η_0^2, ξ_0 и ζ_0 (они соответствуют величинам η^1, η^2, ξ и ζ в лемме); формально, так можно сделать и при $\kappa_0 = 1$; отметим, что упомянутая лемма справедлива и в общих фазовых пространствах. Далее по индукции покажем, как строить вектор $(\eta_{n+1}^1,\eta_{n+1}^2, \xi_{n+1}, \zeta_{n+1})$, если уже построены случайные величины $(\eta_n^1,\eta_n^2,\xi_n,\zeta_n)$. Для этого определим плотности переходных вероятностей φ относительно меры $\Lambda_{x^1,x^2} \times \Lambda_{x^1,x^2} \times \Lambda_{x^1,x^2} \times (\delta_0 + \delta_1)$ для нашего векторнозначного процесса следующим образом:

$$\varphi(x,y) := \varphi_1(x,y^1)\varphi_2(x,y^2)\varphi_3(x,y^3)\varphi_4(x,y^4), \tag{5}$$

где $x = (x^1,x^2,x^3,x^4)$, $y = (y^1,y^2,y^3,y^4)$, и если $0 < \kappa(x^1,x^2) < 1$, то

$$\begin{gathered}\varphi_1(x,u) := \frac{p(x^1,u) - p(x^1,u)\wedge p(x^2,u)}{1-\kappa(x^1,x^2)},\\ \varphi_2(x,u) := \frac{p(x^2,u) - p(x^1,u)\wedge p(x^2,u)}{1-\kappa(x^1,x^2)},\end{gathered} \tag{6}$$

$$\begin{aligned}\varphi_3(x,u) := 1(x^4 = 1)\frac{p(x^1,u)\wedge p(x^2,u)}{\kappa(x^1,x^2)} &+ \\ + 1(x^4 = 0)p(x^3,u)&,\end{aligned} \tag{7}$$

$$\begin{aligned}\varphi_4(x,u) := 1(x^4 = 1)(\delta_1(u)(1-\kappa(x^1,x^2)) &+ \\ + \delta_0(u)\kappa(x^1,x^2)) + 1(x^4 = 0)\delta_0(u)&,\end{aligned} \tag{8}$$

где $\delta_i(u)$ – символ Кронекера, $\delta_i(u) = 1\ (u = i)$, т.е. дельта-мера, сосредоточенная в точке i. Равенство $x^4 = 0$ означает, что каплинг уже имел место на предыдущем шаге, а $u = 0$ символизирует успешную склейку в данный момент перехода.

В вырожденных случаях, если $\kappa(x^1, x^2) = 0$ (каплинг в момент перехода невозможен), то вместо (7) положим, например,

$$\varphi_3(x,u) := 1(x^4 = 1)p(x^3,u) + \\ + 1(x^4 = 0)p(x^3,u) = p(x^3,u),$$

а если $\kappa(x^1, x^2) = 1$, то вместо (6) можно определить

$$\varphi_1(x,u) = \varphi_2(x,u) := p(x^1,u).$$

Формула (8), определяющая $\varphi_4(x,u)$, применима во всех случаях.

Л е м м а 1 [4]. *Пусть $X_0^1 = x$, $X_0^2 = x'$, и пусть $\tilde{X}_n^1$ и $\tilde{X}_n^2$, при $n \in \mathbb{Z}_+$ определены формулой*

$$\tilde{X}_n^1 := \eta_n^1 1(\zeta_n = 1) + \xi_n 1(\zeta_n = 0),$$
$$\tilde{X}_n^2 := \eta_n^2 1(\zeta_n = 1) + \xi_n 1(\zeta_n = 0).$$

Тогда $(\tilde{X}_n^1)_{n\ge 0}$ и $(\tilde{X}_n^2)_{n\ge 0}$ — однородные марковские процессы, и

$$(\tilde{X}_n^1)_{n\ge 0} \stackrel{d}{=} (X_n^1)_{n\ge 0}, \quad (\tilde{X}_n^2)_{n\ge 0} \stackrel{d}{=} (X_n^2)_{n\ge 0}.$$

Более того,

$$\mathbb{P}_{\mu_0^1,\mu_0^2}(\tilde{X}_n^1 \ne \tilde{X}_n^2) \le \mathbb{E}_{\nu_0^1,\nu_0^2} \prod_{i=0}^{n-1} (1 - \kappa(\eta_i^1, \eta_i^2)), \qquad (9)$$

где ν_0^j — распределение случайной величины η_0^j при $j = 1, 2$.

В [4] в лемме 1 было дополнительно предположено, что существует единая мера, доминирующая все переходные ядра. Это условие не является необходимым и доказательство проходит с мерами $\Lambda_{x,x'}$ на каждом шаге вместо единой Λ без дальнейших изменений.

Обозначим $\mathrm{diag}(S^2) := \{x \in S \times S\colon x^1 = x^2\}$ и $\hat{S}^2 := S \times S \backslash \mathrm{diag}(S^2)$. Рассмотрим процесс $X_n := (\tilde{X}_n^1, \tilde{X}_n^2)$ на S^2 и его условную версию $\hat{X}_n := (\tilde{X}_n^1, \tilde{X}_n^2) \mid \tilde{X}_n^1 \ne \tilde{X}_n^2$ на $\hat{S}^2$. Отметим, что процесс $\hat{X}_n$ тоже марковский. Рассмотрим оператор V, определенный на борелевских ограниченных функциях h на пространстве $\hat{S}^2$ по формуле

$$Vh(x) := (1 - \kappa(x^1, x^2))\mathbb{E}_{x^1,x^2} h(\hat{X}_1), \qquad (10)$$

при $x = (x^1, x^2) \in \hat{S}^2$. Из оценки (9) следует, что

$$\frac{1}{2}\|\mu_n^1 - \mu_n^2\|_{TV} \le \mathbb{P}_{\mu_0^1,\mu_0^2}(\tilde{X}_n^1 \ne \tilde{X}_n^2) \le \\ \le \int V^n 1(x)\nu_0^1(dx^1)\nu_0^2(dx^2). \qquad (11)$$

По формуле Гельфанда для спектрального радиуса оператора V (см., например, [11])

$$\limsup_{n\to\infty} (V^n 1(x))^{1/n} \le \limsup_{n\to\infty} \|V^n\|^{1/n} = r(V) \le \|V\|. \qquad (12)$$

Здесь $\|V\| = \sup\limits_{\|h\|\le 1} \sup\limits_{x\in\hat{S}^2} |Vh(x)|$, наиболее удобной нормой для h является $\|h\| = \sum\limits_{x\in\hat{S}^2} |h(x)|$ (хотя все нормы в конечномерном пространстве эквивалентны, и, стало быть, значение $r(V)$ однозначно определено). Соотношения (11) и (12) приводят к следующему результату.

Т е о р е м а 1. *Во всех случаях имеет место цепочка неравенств*

$$\limsup_{n\to\infty} \frac{1}{n} \ln \|P_{x^1}(n,\cdot) - \mu(\cdot)\|_{TV} \le \\ \le \limsup_{n\to\infty} \frac{1}{n} \ln \int 2V^n 1(x^1, x^2)\mu(dx^2) \le \ln r(V).$$

З а м е ч а н и е 1. Если для процесса $\hat{X}_n$ выполнен аналог условия эргодичности (2), то строгое неравенство

$$r(V) < \|V\| = 1 - \kappa$$

справедливо тогда и только тогда, когда функция $1 - \kappa(\cdot)$ не является $\hat{\mu}$-п.н. постоянной, где $\hat{\mu}$ — единственная инвариантная мера процесса $\hat{X}_n$. Равенство $\|V\| = 1 - \kappa$ справедливо для выбранной выше нормы.

ПРИМЕРЫ: КОНЕЧНОЕ S

П р е д л о ж е н и е 1. *Рассмотрим переходную матрицу $\mathcal{P}$ размера 2×2 с элементами $p_{1,1} = a$, $p_{1,2} = 1 - a$, $p_{2,1} = (1 - b)$, $p_{2,2} = b$, предполагая, что $0 < a,b < 1$. Тогда $r = 1 - \kappa = |\lambda_2|$, где λ_2 — наименьшее собственное число матрицы $\mathcal{P}$.*

Доказательство несложное и будет приведено в полной версии работы.

П р и м е р 1 (три метода дают одинаковые оценки скорости сходимости). Рассмотрим цепь Маркова (ЦМ) с $S = \{1,2\}$ и с переходной матрицей

$$\mathcal{P} = \begin{pmatrix} 0.65 & 0.35 \\ 0.35 & 0.65 \end{pmatrix}.$$

Вычисления приводят к значениям $1 - \kappa = 0.3 = r(V)$ и $\lambda_2 = 0.3 = r(V)$, как и ожидалось, в силу равенства $\lambda_2 = 1 - \kappa$.

П р и м е р 2 (новый метод аналогичен МД; оценка спектрального метода лучше). Рассмотрим ЦМ на фазовом пространстве $S = \{1,2,3\}$ с переходной матрицей

$$\mathcal{P} = \begin{pmatrix} 0 & 0.3 & 0.7 \\ 0.3 & 0.7 & 0 \\ 0.7 & 0 & 0.3 \end{pmatrix}.$$

Простые вычисления показывают, что $r(V) = 0.7$, $1 - \kappa = 0.7$, и что собственные числа матрицы $\mathcal{P}$ равны $\lambda_1 = 1$, $\lambda_{2,3} = \pm\sqrt{37}/10 \cong \pm 0.6082763$. Стало быть, $|\lambda_{2,3}| < r(V)$.

Пример 3 (оценка нового метода аналогична классической МД; оценка спектрального метода лучше).

$$\mathcal{P} = \begin{pmatrix} 0 & 0.3 & 0.7 \\ 0.7 & 0 & 0.3 \\ 0.3 & 0.7 & 0 \end{pmatrix} \text{ -- двойная стохастическая;}$$

$$\pi = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) \text{ инвариантно;}$$

$\kappa(\cdot) \equiv 0.3$. Имеем, $1 - \kappa = 0.7$, и также $r(V) = 0.7$. Итак, в данном примере новая оценка аналогична классической. Анализ показывает, что

$$|\lambda_{2,3}| = \sqrt{0.25 + 0.12} = \sqrt{0.37} \approxeq 0.6082763 < 0.7.$$

Пример 4 (оценка нового метода лучше, чем МД, и аналогична оценке спектрального). Рассмотрим ЦМ с фазовым пространством $S = \{1, 2, 3\}$ и с переходной матрицей

$$\mathcal{P} = \begin{pmatrix} 0 & 0.3 & 0.7 \\ 1.0 & 0 & 0 \\ 0.8 & 0.1 & 0.1 \end{pmatrix}.$$

В этом случае код (будет приведен в полной версии изложения) дает значения

$$|\lambda_2| = 1 - \kappa = r(V) = \frac{9}{20} + \frac{\sqrt{65}}{20} \approx 0.85311289.$$

В этом примере классический метод Маркова–Добрушина бесполезен, поскольку $1 - \kappa = 1$. Замечание 1 здесь также применимо.

ИСТОЧНИКИ ФИНАНСИРОВАНИЯ

Для первого автора исследования разделов 1 и 2 за исключением теоремы 1 и замечания 2 финансировались в рамках государственной поддержки ведущих университетов Российской Федерации “5-100”; теорема 1 и замечание 2 выполнены за счет гранта Российского научного фонда (проект 17–11–01098). Для второго автора исследования раздела 3 выполнены за счет гранта Российского научного фонда (проект 17–11–01098).

СПИСОК ЛИТЕРАТУРЫ

1. *Seneta E.* Non-negative Matrices and Markov Chains. N.Y.: Springer, 1981. https://doi.org/10.1007/0-387-32792-4
2. *Veretennikov A.Yu.* Ergodic Markov processes and Poisson equations (lecture notes). In: Modern problems of stochastic analysis and statistics – Selected contributions in honor of Valentin Konakov / Ed. by V. Panov. Springer, 2017. P. 457–511. https://doi.org/10.1007/978-3-319-65313-6
3. *Гантмахер Ф.Р.* Теория матриц. 5-е изд. М.: Физматлит, 2004.
4. *Butkovsky O.A., Veretennikov A.Yu.* On Asymptotics for Vaserstein Coupling of Markov Chains // Stochastic Processes and Their Applications. 2013. V. 123. № 9. P. 3518–3541.
5. *Lindvall T.* Lectures on the Coupling Method. Dover, 2002.
6. *Васерштейн Л.Н.* Марковские процессы на счетном произведении пространств, описывающие большие системы автоматов // Пробл. передачи информ. 1969. Т. 5. № 3. С. 64–72.
7. *Дынкин Е.Б.* Марковские процессы. М.: Физматгиз, 1963.
8. *Марков А.А.* Распространение закона больших чисел на величины, зависящие друг от друга // Изв. физ.-матем., Казан. унив. 1906. Т. 15. С. 135–156.
9. *Добрушин Р.Л.* Центральная предельная теорема для неоднородных цепей Маркова I, II // Теория вероятн. и ее применения. 1956. № 1. С. 72–89; № 4. С. 365–425.
10. *Калашников В.В.* Метод склеивания, его развитие и применения / В кн.: Э. Нуммелин. Общие неприводимые цепи Маркова и неотрицательные операторы. М.: Мир, 1989. С. 176–190.
11. *Красносельский М.А., Лифшиц Е.А., Соболев А.В.* Позитивные линейные системы. Метод положительных операторов. Сер.: Теория и методы системного анализа. М.: Наука, 1985.
12. *Nummelin E.* General Irreducible Markov Chains. CUP. Cambridge, 2008. https://doi.org/10.1017/S001309150001765X

ON CONVERGENCE RATES FOR HOMOGENEOUS MARKOV CHAINS

A. Yu. Veretennikov[1,2,3] and M. A. Veretennikova[2]

[1]*University of Leeds, Leeds, United Kingdom*

[2]*National Research University Higher School of Economics, Moscow, Russian Federation*

[3]*Institute for Information Transmission Problems, Moscow, Russian Federation*

Presented by Academician of the RAS A.N. Shiryaev June 21, 2019

Received July 01, 2019

New imporved rates of convergence for ergodic homogeneous Markov chains are studied. Examples of comparison with the classical rate bounds are provided.

Keywords: Markov chains, ergodicity, convergence rate