离散时间 Markov 链在全变差距离下的收敛速度

宋延红

中南财经政法大学统计与数学学院, 武汉 430073
E-mail: songyh@mail.bnu.edu.cn

收稿日期: 2013-11-11; 接受日期: 2015-01-22

国家自然科学基金 (批准号: 11426219)、国家留学基金 (批准号: 201407085008) 和中央高校基本科研业务费专项基金 (批准号: 31541311214) 资助项目

摘要 本文研究离散时间遍历 Markov 链在全变差距离下的收敛速度. 当 Markov 链的初分布满足一定的可积条件时, 本文分别利用 Poincaré、弱 Poincaré、对数 Sobolev 和弱对数 Sobolev 不等式给出收敛速度的估计. 进一步, 利用截断方法, 减弱对初分布的可积性要求.

关键词 Markov 链 收敛速度 泛函不等式

MSC (2010) 主题分类 60J05, 26D15

1 引言

设 $(\Phi_n)_{n \geq 0}$ 是一般状态空间 X 上的离散时间遍历 Markov 链, 在 X 上赋予可数生成的 σ 代数 $\mathcal{B}(X)$. 令 μ 为 Markov 链的平稳分布, $P^n(x, A)$ 为其 n 步转移核:

$$P^n(x, A) = \mathbb{P}_x\{\Phi_n \in A\}, \quad n \in \mathbb{Z}_+, \quad x \in X, \quad A \in \mathcal{B}(X),$$

其中 $\mathbb{P}_x$ 是给定 $\Phi_0 = x$ 的条件下 Markov 链的条件分布. 平稳分布 μ 以及转移核 P^n 作用在非负可测函数 f 上分别定义为 (如果相应的积分存在)

$$\mu(f) = \int_X f(x) \mu(dx), \quad P^n f(x) = \int_X f(y) P^n(x, dy), \quad n \in \mathbb{Z}_+, \quad x \in X.$$

假设 Markov 链 Φ_n 非周期且 ψ 不可约, 其中 ψ 为极大不可约测度 (参见文献 [1, 第 4 章]).

令 ν 为 Markov 链的任意初分布, 且 ν 关于平稳分布 μ 绝对连续. 令 $d\nu/d\mu = h$, 则

$$\nu P^n(\cdot) = \int_X P^n(x, \cdot) h(x) \mu(dx) = P^{n*} h d\mu$$

就是 Φ_n 的以 ν 为初分布的分布律, 其中 P^{n*} 为 P^n 的共轭算子. 测度 νP^n 与 μ 之间的全变差距离定义为

$$\|\nu P^n - \mu\|_{\text{TV}} = 2 \sup_{A \in \mathcal{B}(X)} |\nu P^n(A) - \mu(A)|,$$

则显然有

$$\|\nu P^n - \mu\|_{\text{TV}} = \|P^{n*} h - 1\|_1,$$

且分别由 Cauchy-Schwarz 不等式和 Pinsker 不等式^[2] 知,

$$\|\nu P^n - \mu\|_{\text{TV}} \leq \|P^{n*}h - 1\|_2, \quad (1.1)$$

$$\|\nu P^n - \mu\|_{\text{TV}} \leq \sqrt{2\text{Ent}_\mu(P^{n*}h)}, \quad (1.2)$$

其中 $\|\cdot\|_1$ 和 $\|\cdot\|_2$ 分别是 $L^1(\mu)$ 和 $L^2(\mu)$ 中的范数, $\text{Ent}_\mu(f)$ 为 f 相对于 μ 的相对熵, 定义为

$$\text{Ent}_\mu(f) = \mu(f \log f) - \mu(f) \log \mu(f), \quad f \geq 0.$$

对任意的 $f \in L^2(\mu)$, Dirichlet 型 $\mathcal{E}$ 定义为

$$\mathcal{E}(f, f) = \langle (I - P^*P)f, f \rangle = \|f\|_2^2 - \|Pf\|_2^2, \quad (1.3)$$

其中 $\langle \cdot, \cdot \rangle$ 为 $L^2(\mu)$ 中的内积.

本文的目的就是利用泛函不等式研究在全变差距离下 νP^n 收敛于 μ 的速度. 泛函不等式是研究 Markov 链的各种分析与概率性质的重要工具, 已有越来越多的概率学家致力于泛函不等式及其应用的研究, 关于这方面的系统介绍可参见专著 [3-5]. 利用连续时间泛函不等式研究收敛速度的通常思路是, 先对连续时间转移半群的方差或者熵求导, 然后再利用 Gronwall 不等式得到收敛速度的估计. 然而本文不能应用此方法. 问题在于: 一是全变差距离无法求导; 二是离散时间 Markov 链的转移核 P^n 也无法求导. 为解决问题一, 本文分别利用 (1.1) 和 (1.2) 将全变差距离下的收敛转化为 L^2 收敛以及熵意义下的收敛; 为解决问题二, 本文利用差分运算代替求导运算. 基于这种思想, 本文分别利用离散时间的 Poincaré、弱 Poincaré、对数 Sobolev 和弱对数 Sobolev 不等式, 得到了在全变差距离下 νP^n 收敛于 μ 的速度估计 (见定理 1、3、5 和 6).

注意到这四个定理对初分布的可积性要求比较高, 在实践中不便于应用, 因此, 我们采用截断方法来减弱对初分布的可积性要求. 截断方法的思想是, 插入一个任意的常数 $K > 0$, 从而对有界函数 $h \wedge K$ 应用泛函不等式. 具体过程是, 设函数 $\varphi : [0, \infty) \rightarrow [0, \infty)$ 连续且单调递增趋于无穷大. 若 $\int h\varphi(h)d\mu < \infty$, 则对于任意的常数 $K > 0$,

$$\begin{aligned} \|\nu P^n - \mu\|_{\text{TV}} &= \int |P^{n*}h - 1|d\mu \\ &\leq \int |P^{n*}(h \wedge K) - P^{n*}h|d\mu + \int \left| P^{n*}(h \wedge K) - \int (h \wedge K)d\mu \right|d\mu \\ &\quad + \left| \int (h \wedge K)d\mu - 1 \right| \\ &\leq \int \left| P^{n*}(h \wedge K) - \int (h \wedge K)d\mu \right|d\mu + 2 \int (h - K)1_{\{h \geq K\}}d\mu \\ &\leq \int \left| P^{n*}(h \wedge K) - \int (h \wedge K)d\mu \right|d\mu + \frac{2}{\varphi(K)} \int h\varphi(h)d\mu. \end{aligned} \quad (1.4)$$

由于 $h \wedge K$ 是有界函数, 因此可以直接利用定理 1、3、5 和 6 来估计 $\int |P^{n*}(h \wedge K) - \int (h \wedge K)d\mu|d\mu$. 最后再将所得的估计关于 K 取最优. 对连续时间 Markov 过程的收敛速度, 文献 [6, 7] 采用类似的截断方法进行了研究.

本文的结构如下: 第 2 节将应用 Poincaré 和弱 Poincaré 不等式来研究离散时间转移核 P^n 的 L^2 收敛, 进而得到在全变差距离下 νP^n 收敛于 μ 的速度; 第 3 节将利用对数 Sobolev 和弱对数 Sobolev 不等式研究转移核 P^n 在熵意义下的收敛, 从而也给出在全变差距离下 νP^n 收敛于 μ 的速度.

注意, 在全文中, 令 $M = \int h\varphi(h)d\mu$, 其中 $h = d\nu/d\mu$ 是概率密度函数, $\varphi: [0, \infty) \rightarrow [0, \infty)$ 为连续且单调递增趋于无穷大的函数.

2 Poincaré 和弱 Poincaré 不等式

Poincaré 不等式是最经典的泛函不等式之一. 连续时间的 Poincaré 不等式等价于转移半群的 L^2 指数收敛, 且与生成元的谱性质有密切联系, 详细内容可参见文献 [3]. 本文考虑下述离散时间的 Poincaré 不等式: 存在某个常数 $\lambda_P > 0$, 使得对任意的 $f \in L^2(\mu)$,

$$\text{Var}_\mu(f) \leq \lambda_P \mathcal{E}(f, f), \quad (2.1)$$

其中方差 $\text{Var}_\mu(f) = \mu(f^2) - \mu(f)^2$, Dirichlet 型 $\mathcal{E}$ 如 (1.3) 所示. 根据文献 [8, 引理 3.11 和 3.12] 知, (2.1) 等价于转移核 P^n 的 L^2 几何收敛. 由此以及不等式 (1.1) 可得下述定理.

定理 1 对于某个常数 $\lambda_P > 0$, 以下两条论述等价:

- (1) Poincaré 不等式 (2.1) 成立.
- (2) 对任意的 $f \in L^2(\mu)$,

$$\|P^n f - \mu(f)\|_2^2 \leq \|f - \mu(f)\|_2^2 \left(1 - \frac{1}{\lambda_P}\right)^n, \quad n \in \mathbb{Z}_+. \quad (2.2)$$

进一步, 若 Poincaré 不等式 (2.1) 成立且 $h \in L^2(\mu)$, 则

$$\|\nu P^n - \mu\|_{\text{TV}} \leq \|h - 1\|_2 \left(1 - \frac{1}{\lambda_P}\right)^{n/2}, \quad n \in \mathbb{Z}_+. \quad (2.3)$$

但定理 1 仅能得到 $h \in L^2(\mu)$ 时的收敛速度. 若 h 的可积性较弱, 则有下列的推论.

推论 2 设 $M < \infty$ 且 $x \mapsto \varphi(x)/x$ 单减. 若 Poincaré 不等式 (2.1) 成立, 则有

$$\|\nu P^n - \mu\|_{\text{TV}} \leq \beta^{-1} \left(4M \left(1 - \frac{1}{\lambda_P}\right)^{-n}\right) \left(1 - \frac{1}{\lambda_P}\right)^n, \quad n \in \mathbb{Z}_+,$$

其中 β^{-1} 为 β 的逆, $\beta(x) = x\varphi(x)$.

证明 对任意的 $K > 0$, 由 (1.4)、Cauchy-Schwarz 不等式和 (2.2) 知,

$$\begin{aligned} \|\nu P^n - \mu\|_{\text{TV}} &\leq \int \left| P^{n*}(h \wedge K) - \int (h \wedge K) d\mu \right| d\mu + \frac{2M}{\varphi(K)} \\ &\leq \left[\int \left(P^{n*}(h \wedge K) - \int (h \wedge K) d\mu \right)^2 d\mu \right]^{1/2} + \frac{2M}{\varphi(K)} \\ &\leq \sqrt{\text{Var}_\mu(h \wedge K)} \left(1 - \frac{1}{\lambda_P}\right)^{n/2} + \frac{2M}{\varphi(K)} \\ &\leq \sqrt{\int (h \wedge K)^2 d\mu} \left(1 - \frac{1}{\lambda_P}\right)^{n/2} + \frac{2M}{\varphi(K)}. \end{aligned} \quad (2.4)$$

若 $h \geq K$, 则由 φ 的单增性知,

$$(h \wedge K)^2 \leq hK \leq hK \frac{\varphi(h)}{\varphi(K)}.$$

若 $h < K$, 则由于 $x \mapsto \varphi(x)/x$ 单减, 故而

$$(h \wedge K)^2 \leq h^2 \leq h^2 \frac{K}{h} \frac{\varphi(h)}{\varphi(K)} = hK \frac{\varphi(h)}{\varphi(K)}.$$

因此, 由这两个不等式和 (2.4) 可得

$$\|\nu P^n - \mu\|_{\text{TV}} \leq \sqrt{\frac{KM}{\varphi(K)}} \left(1 - \frac{1}{\lambda_P}\right)^{n/2} + \frac{2M}{\varphi(K)}. \quad (2.5)$$

下面选取恰当的 $K > 0$, 以使 (2.5) 的右端达到最优. 注意到 $a + b \geq 2\sqrt{ab}$ ($a, b \geq 0$), 且等号成立当且仅当 $a = b$. 于是令

$$\sqrt{\frac{KM}{\varphi(K)}} \left(1 - \frac{1}{\lambda_P}\right)^{n/2} = \frac{2M}{\varphi(K)},$$

解得

$$K = \beta^{-1} \left(4M \left(1 - \frac{1}{\lambda_P}\right)^{-n}\right).$$

此 K 就能使 (2.5) 右端达到最优, 将此 K 代入 (2.5) 便证得所需结论. $\square$

例 1 若 Poincaré 不等式 (2.1) 成立且 $h \in L^q(\mu)$, $1 < q \leq 2$, 则由推论 2 (取 $\varphi(x) = x^{q-1}$) 可得

$$\|\nu P^n - \mu\|_{\text{TV}} \leq \left(4 \int h^q d\mu\right)^{1/q} \left(1 - \frac{1}{\lambda_P}\right)^{(q-1)n/q}, \quad n \in \mathbb{Z}_+.$$

当 $q = 2$ 时, 其收敛速度恰好与 (2.3) 中的收敛速度相同, 都为 $(1 - 1/\lambda_P)^{n/2}$.

由例 1 可知, 对任何满足 $h \in L^q(\mu)$, $1 < q \leq 2$ 的初分布, 我们都可以得到在全变差距离下 νP^n 收敛到平稳分布的速度. 考虑如下对称的 Metropolis 算法 (具体内容可参见文献 [9, 10]): 设 μ^{Leb} 为 $\mathbb{R}_+$ 上的 Lebesgue 测度. 给定概率密度函数 $\pi(x) = e^{-x}$, $x \in \mathbb{R}_+$, 并由此定义概率测度

$$\pi(dx) = \pi(x) \mu^{\text{Leb}}(dx).$$

设 $Q(x, \cdot)$ 为任意转移核, 且关于 Lebesgue 测度 μ^{Leb} 绝对连续, 密度函数为 $q(x, y) = q(|x - y|)$, 其支撑包含在 $[x - a, x + a]$ (常数 $a > 0$) 内. 于是, Metropolis 算法按照如下的转移核 P 进行:

$$\begin{aligned} P(x, dy) &= q(x, y) \alpha(x, y) \mu^{\text{Leb}}(dy), \quad y \neq x, \\ P(x, \{x\}) &= \int q(x, y) [1 - \alpha(x, y)] \mu^{\text{Leb}}(dy) + Q(x, \{x\}), \end{aligned}$$

其中

$$\alpha(x, y) = \begin{cases} \min \left\{ \frac{\pi(y)}{\pi(x)}, 1 \right\}, & \pi(x)q(x, y) > 0, \\ 1, & \pi(x)q(x, y) = 0. \end{cases}$$

由文献 [3, 定理 9.15] 类似的证明和文献 [10, 例 4] 知, 概率测度 π 为转移核 P 的平稳分布, 且 P 满足 Poincaré 不等式 (2.1). 设 ν 为任意初分布, 且 $d\nu/d\pi = h$. 若 $h \in L^2(\pi)$, 则由定理 1 可得

$$\|\nu P^n - \pi\|_{\text{TV}} \leq \|h - 1\|_2 \left(1 - \frac{1}{\lambda_P}\right)^{n/2}, \quad n \in \mathbb{Z}_+.$$

若 $h \in L^q(\pi)$, $1 < q \leq 2$, 则由例 1 知,

$$\|\nu P^n - \pi\|_{\text{TV}} \leq \left(4 \int h^q d\mu\right)^{1/q} \left(1 - \frac{1}{\lambda_P}\right)^{(q-1)n/q}, \quad n \in \mathbb{Z}_+.$$

接下来, 我们考虑弱 Poincaré 不等式. 连续时间的弱 Poincaré 不等式最早是在 2001 年由 Röckner 和 Wang^[11] 引入的, 他们证明了此类不等式可以刻画转移半群的次指数收敛. 在这之后又涌现出了许多研究成果, 如文献 [4, 5, 12] 等, 但这些成果都是围绕连续时间弱 Poincaré 不等式展开的. 本文要研究的是离散时间弱 Poincaré 不等式, 即存在单减函数 $\alpha_{\text{WP}} : (0, \infty) \rightarrow [0, \infty)$, 使得对任意的 $r > 0$ 以及有界可测函数 f ,

$$\text{Var}_\mu(f) \leq \alpha_{\text{WP}}(r) \mathcal{E}(f, f) + r \text{Osc}^2(f), \quad (2.6)$$

其中 $\text{Osc}(f) = \sup f - \inf f$. 若函数 α_{WP} 有界, 则弱 Poincaré 不等式就退化为 Poincaré 不等式. 因此, 不失一般性, 我们假设 $\lim_{r \downarrow 0} \alpha_{\text{WP}}(r) = \infty$.

对于这类离散时间的弱 Poincaré 不等式, 我们将研究其与转移核 P^n 的收敛速度之间的关系, 并进而估计在全变变距离下 νP^n 收敛到 μ 的速度.

定理 3 (1) 设弱 Poincaré 不等式 (2.6) 成立, 则对任意的有界可测函数 f ,

$$\|P^n f - \mu(f)\|_2^2 \leq \text{Osc}^2(f) \xi(n), \quad n \in \mathbb{Z}_+, \quad (2.7)$$

其中

$$\xi(n) = \inf \left\{ \frac{1}{4} \left(\frac{\alpha_{\text{WP}}(r) - 1}{\alpha_{\text{WP}}(r)} \right)^n + r : \alpha_{\text{WP}}(r) > 1 \right\}. \quad (2.8)$$

特别地, 当 $n \rightarrow \infty$ 时, $\xi(n) \rightarrow 0$.

(2) 反之, 假设 $P = P^*$. 若存在单减函数 $\xi : \mathbb{Z}_+ \rightarrow (0, \infty)$, 满足当 $n \rightarrow \infty$ 时, $\xi(n) \rightarrow 0$, 且对任意的有界可测函数 f ,

$$\|P^n f - \mu(f)\|_2^2 \leq \text{Osc}^2(f) \xi(n), \quad n \in \mathbb{Z}_+, \quad (2.9)$$

则弱 Poincaré 不等式 (2.6) 对于 $\alpha_{\text{WP}}(r) = \inf\{n \in \mathbb{Z}_+ : \xi(n) \leq r\}$ 成立.

(3) 若弱 Poincaré 不等式 (2.6) 成立且 h 有界, 则有

$$\|\nu P^n - \mu\|_{\text{TV}} \leq \text{Osc}(h) \xi(n)^{1/2}, \quad n \in \mathbb{Z}_+, \quad (2.10)$$

其中 $\xi(n)$ 由 (2.8) 给出.

证明 不失一般性, 设 $\mu(f) = 0$. 如若不然, 则用 $f - \mu(f)$ 代替 f 即可.

(1) 任意固定 $r > 0$, 使得 $\alpha_{\text{WP}}(r) > 1$. 令 $t(n) = \|P^n f\|_2^2$, $n \in \mathbb{Z}_+$, 则由 (2.6) 和不等式 $\text{Osc}^2(P^{n-1}f) \leq \text{Osc}^2(f)$ 知,

$$t(n-1) \leq \alpha_{\text{WP}}(r)(t(n-1) - t(n)) + r \text{Osc}^2(f),$$

从而可得

$$\begin{aligned} t(n) &\leq \frac{\alpha_{\text{WP}}(r) - 1}{\alpha_{\text{WP}}(r)} t(n-1) + \frac{r}{\alpha_{\text{WP}}(r)} \text{Osc}^2(f) \\ &\leq \frac{\alpha_{\text{WP}}(r) - 1}{\alpha_{\text{WP}}(r)} \left(\frac{\alpha_{\text{WP}}(r) - 1}{\alpha_{\text{WP}}(r)} t(n-2) + \frac{r}{\alpha_{\text{WP}}(r)} \text{Osc}^2(f) \right) + \frac{r}{\alpha_{\text{WP}}(r)} \text{Osc}^2(f) \\ &= \left(\frac{\alpha_{\text{WP}}(r) - 1}{\alpha_{\text{WP}}(r)} \right)^2 t(n-2) + \sum_{k=0}^1 \left(\frac{\alpha_{\text{WP}}(r) - 1}{\alpha_{\text{WP}}(r)} \right)^k \frac{r}{\alpha_{\text{WP}}(r)} \text{Osc}^2(f) \end{aligned}$$

$$\begin{aligned}
 &\leq \dots \\
 &\leq \left(\frac{\alpha_{\text{WP}}(r) - 1}{\alpha_{\text{WP}}(r)} \right)^n t(0) + \sum_{k=0}^{n-1} \left(\frac{\alpha_{\text{WP}}(r) - 1}{\alpha_{\text{WP}}(r)} \right)^k \frac{r}{\alpha_{\text{WP}}(r)} \text{Osc}^2(f) \\
 &\leq \left(\frac{\alpha_{\text{WP}}(r) - 1}{\alpha_{\text{WP}}(r)} \right)^n t(0) + r \text{Osc}^2(f).
 \end{aligned}$$

注意到 $t(0) = \text{Var}_\mu(f) \leq \text{Osc}^2(f)/4$, 故而

$$\|P^n f\|_2^2 \leq \text{Osc}^2(f) \left[\frac{1}{4} \left(\frac{\alpha_{\text{WP}}(r) - 1}{\alpha_{\text{WP}}(r)} \right)^n + r \right].$$

对上式右端关于 r 取最优即得所需结论, 且显然当 $n \rightarrow \infty$ 时, $\xi(n) \rightarrow 0$.

(2) 假设 (2.9) 成立. 由于 $P = P^*$, 且 P^n 在 $L^2(\mu)$ 中具有压缩性, 故而

$$\begin{aligned}
 \|f\|_2^2 &= \sum_{i=0}^{n-1} (\|P^i f\|_2^2 - \|P^{i+1} f\|_2^2) + \|P^n f\|_2^2 \\
 &= \sum_{i=0}^{n-1} \|(I - P^2)^{1/2} P^i f\|_2^2 + \|P^n f\|_2^2 \\
 &\leq n \langle (I - P^2) f, f \rangle + \xi(n) \text{Osc}^2(f),
 \end{aligned}$$

其中 $(I - P^2)^{1/2}$ 是由谱理论给出的对称平方根. 令 $\xi(n) \leq r$, 则由上式得

$$\|f\|_2^2 \leq n \mathcal{E}(f, f) + r \text{Osc}^2(f).$$

再对上式右端关于 n 取最优即可得证.

(3) 由 Cauchy-Schwarz 不等式以及 (1) 便可证得所需结论. □

推论 4 设 $M < \infty$. 若弱 Poincaré 不等式 (2.6) 成立, 则有

$$\|\nu P^n - \mu\|_{\text{TV}} \leq 2\gamma^{-1} \frac{2M}{\sqrt{\xi(n)}} \xi(n)^{1/2}, \quad n \in \mathbb{Z}_+,$$

其中 $\xi(n)$ 如 (2.8) 所示, γ^{-1} 为 γ 的逆, $\gamma(x) = x\varphi(x)$.

证明 对任意的 $K > 0$, 由 (1.4)、Cauchy-Schwarz 不等式和 (2.7) 知,

$$\begin{aligned}
 \|\nu P^n - \mu\|_{\text{TV}} &\leq \left[\int \left(P^{n*}(h \wedge K) - \int (h \wedge K) d\mu \right)^2 d\mu \right]^{1/2} + \frac{2M}{\varphi(K)} \\
 &\leq [\text{Osc}^2(h \wedge K) \xi(n)]^{1/2} + \frac{2M}{\varphi(K)} \\
 &\leq \sqrt{\xi(n)} K + \frac{2M}{\varphi(K)}.
 \end{aligned} \tag{2.11}$$

下面选取恰当的 $K > 0$, 以使 (2.11) 最后一个不等式的右端达到最优. 令

$$\sqrt{\xi(n)} K = \frac{2M}{\varphi(K)},$$

计算得

$$K = \gamma^{-1} \frac{2M}{\sqrt{\xi(n)}}.$$

将其代入 (2.11) 便证得所需结论. $\square$

例 2 若弱 Poincaré 不等式 (2.6) 成立且 $h \in L^q(\mu)$, $q > 1$, 则在推论 4 中取 $\varphi(x) = x^{q-1}$ 可得

$$\|\nu P^n - \mu\|_{\text{TV}} \leq 2^{(q+1)/q} \left(\int h^q d\mu \right)^{1/q} \xi(n)^{(q-1)/(2q)}, \quad n \in \mathbb{Z}_+.$$

当 $q \rightarrow \infty$ 时, 上式的收敛速度恰好为 $\xi(n)^{1/2}$, 与 (2.10) 中的收敛速度一致.

由例 2 可以发现, 相对定理 3(3) 而言, 推论 4 减弱了对 h 的可积性要求.

3 对数 Sobolev 和弱对数 Sobolev 不等式

本节将分别利用对数 Sobolev 和弱对数 Sobolev 不等式, 研究在全变差距离下 νP^n 收敛于 μ 的速度.

首先, 我们考虑对数 Sobolev 不等式. 对这类不等式的研究可以追溯到 Gross^[13,14], 他证明了连续时间的对数 Sobolev 不等式与转移半群在熵意义下的指数收敛有密切联系, 且等价于转移半群的超压缩性. 关于对数 Sobolev 不等式的其他研究成果, 读者可以参见文献 [3–5, 7, 15–17] 等. 本文研究的是离散时间对数 Sobolev 不等式, 即存在某个常数 $\lambda_{\text{LS}} > 0$, 使得对任意的 $f \in L^2(\mu)$,

$$\text{Ent}_\mu(f^2) \leq \lambda_{\text{LS}} \mathcal{E}(f, f), \quad (3.1)$$

其中 Dirichlet 型 $\mathcal{E}$ 如 (1.3) 所示, 即 $\mathcal{E}(f, f) = \langle (I - P^*P)f, f \rangle$. 由于 $I - P^*P$ 是有界算子, 故 (3.1) 成立当且仅当状态空间 X 有限. 对于这种离散时间的对数 Sobolev 不等式, Miclo 在 1997 年进行了系统地研究, 证明了 (3.1) 可以刻画转移核 P^n 在熵意义下的几何收敛性 (参见文献 [15, 推论 7]). 由此以及 Pinsker 不等式 (1.2) 可得如下定理.

定理 5 (1) 若对数 Sobolev 不等式 (3.1) 成立, 则对任意非负可测函数 f , 当 $\text{Ent}_\mu(f) < \infty$ 时,

$$\text{Ent}_\mu(P^n f) \leq \text{Ent}_\mu(f) \left(1 - \frac{1}{\lambda_{\text{LS}}} \right)^n, \quad n \in \mathbb{Z}_+;$$

(2) 若对数 Sobolev 不等式 (3.1) 成立且 $\text{Ent}_\mu(h) < \infty$, 则有

$$\|\nu P^n - \mu\|_{\text{TV}} \leq \sqrt{2\text{Ent}_\mu(h)} \left(1 - \frac{1}{\lambda_{\text{LS}}} \right)^{n/2}, \quad n \in \mathbb{Z}_+.$$

最后, 我们考虑弱对数 Sobolev 不等式. 对于这类不等式已有的研究成果, 读者可参见文献 [6, 7, 18] 等. 特别地, Cattiaux^[6] 研究了连续时间的弱对数 Sobolev 不等式与转移半群在熵意义下的次指数收敛之间的关系. 本文将研究的是, 离散时间弱对数 Sobolev 不等式与转移核在熵意义下的次几何收敛性之间的关系, 并由此给出在全变差距离下 νP^n 收敛到 μ 的速度. 这里的离散时间弱对数 Sobolev 不等式指的是, 存在单减函数 $\alpha_{\text{WLS}} : (0, \infty) \rightarrow [0, \infty)$, 使得对任意的 $r > 0$ 以及有界可测函数 f ,

$$\text{Ent}_\mu(f^2) \leq \alpha_{\text{WLS}}(r) \mathcal{E}(f, f) + r \text{Osc}^2(f). \quad (3.2)$$

若函数 α_{WLS} 有界, 则弱对数 Sobolev 不等式就退化为对数 Sobolev 不等式. 因此, 不失一般性, 假设 $\lim_{r \downarrow 0} \alpha_{\text{WLS}}(r) = \infty$.

定理 6 (1) 设弱对数 Sobolev 不等式 (3.2) 成立, 则对任意的非负有界可测函数 f ,

$$\text{Ent}_\mu(P^n f) \leq \text{Osc}^2(\sqrt{f}) \xi(n), \quad n \in \mathbb{Z}_+, \quad (3.3)$$

其中

$$\xi(n) = \inf \left\{ \frac{1}{e} \left(\frac{\alpha_{\text{WLS}}(r) - 1}{\alpha_{\text{WLS}}(r)} \right)^n + r : \alpha_{\text{WLS}}(r) > 1 \right\}. \quad (3.4)$$

特别地, 当 $n \rightarrow \infty$ 时, $\xi(n) \rightarrow 0$.

(2) 若弱对数 Sobolev 不等式 (3.2) 成立且 h 有界, 则有

$$\|\nu P^n - \mu\|_{\text{TV}} \leq \sqrt{2} \text{Osc}(\sqrt{h}) \xi(n)^{1/2}, \quad n \in \mathbb{Z}_+, \quad (3.5)$$

其中 $\xi(n)$ 由 (3.4) 给出.

证明 (1) 任意固定 $r > 0$, 使得 $\alpha_{\text{WLS}}(r) > 1$. 设弱对数 Sobolev 不等式 (3.2) 成立, 则对于任意的非负有界可测函数 f ,

$$\langle (I - P^*P)\sqrt{f}, \sqrt{f} \rangle \geq \frac{1}{\alpha_{\text{WLS}}(r)} (\text{Ent}_\mu(f) - r \text{Osc}^2(\sqrt{f})).$$

故而由文献 [15, 性质 6] 以及上式知,

$$\text{Ent}_\mu(Pf) \leq \text{Ent}_\mu(f) - \langle (I - P^*P)\sqrt{f}, \sqrt{f} \rangle \leq \frac{\alpha_{\text{WLS}}(r) - 1}{\alpha_{\text{WLS}}(r)} \text{Ent}_\mu(f) + \frac{r}{\alpha_{\text{WLS}}(r)} \text{Osc}^2(\sqrt{f}). \quad (3.6)$$

用 Pf 代替上式的 f , 并利用不等式

$$\text{Osc}^2(\sqrt{P^n f}) \leq \text{Osc}^2(\sqrt{f}),$$

可得

$$\begin{aligned} \text{Ent}_\mu(P^2 f) &\leq \frac{\alpha_{\text{WLS}}(r) - 1}{\alpha_{\text{WLS}}(r)} \text{Ent}_\mu(Pf) + \frac{r}{\alpha_{\text{WLS}}(r)} \text{Osc}^2(\sqrt{f}) \\ &\leq \frac{\alpha_{\text{WLS}}(r) - 1}{\alpha_{\text{WLS}}(r)} \left(\frac{\alpha_{\text{WLS}}(r) - 1}{\alpha_{\text{WLS}}(r)} \text{Ent}_\mu(f) + \frac{r}{\alpha_{\text{WLS}}(r)} \text{Osc}^2(\sqrt{f}) \right) + \frac{r}{\alpha_{\text{WLS}}(r)} \text{Osc}^2(\sqrt{f}) \\ &= \left(\frac{\alpha_{\text{WLS}}(r) - 1}{\alpha_{\text{WLS}}(r)} \right)^2 \text{Ent}_\mu(f) + \sum_{k=0}^1 \left(\frac{\alpha_{\text{WLS}}(r) - 1}{\alpha_{\text{WLS}}(r)} \right)^k \frac{r}{\alpha_{\text{WLS}}(r)} \text{Osc}^2(\sqrt{f}). \end{aligned}$$

依次类推有

$$\begin{aligned} \text{Ent}_\mu(P^n f) &\leq \left(\frac{\alpha_{\text{WLS}}(r) - 1}{\alpha_{\text{WLS}}(r)} \right)^n \text{Ent}_\mu(f) + \sum_{k=0}^{n-1} \left(\frac{\alpha_{\text{WLS}}(r) - 1}{\alpha_{\text{WLS}}(r)} \right)^k \frac{r}{\alpha_{\text{WLS}}(r)} \text{Osc}^2(\sqrt{f}) \\ &\leq \left(\frac{\alpha_{\text{WLS}}(r) - 1}{\alpha_{\text{WLS}}(r)} \right)^n \text{Ent}_\mu(f) + r \text{Osc}^2(\sqrt{f}), \quad n \in \mathbb{Z}_+. \end{aligned}$$

由此并注意到 $\text{Ent}_\mu(f) \leq \text{Osc}^2(\sqrt{f})/e$, 可知

$$\text{Ent}_\mu(P^n f) \leq \text{Osc}^2(\sqrt{f}) \left[\frac{1}{e} \left(\frac{\alpha_{\text{WLS}}(r) - 1}{\alpha_{\text{WLS}}(r)} \right)^n + r \right].$$

将上式右端关于 r 取最优即可证得所需结论, 且显然当 $n \rightarrow \infty$ 时, $\xi(n) \rightarrow 0$.

(2) 由 Pinsker 不等式 (1.2) 和 (1) 可证. □

推论 7 设 $M < \infty$. 若弱对数 Sobolev 不等式 (3.2) 成立, 则有

$$\|\nu P^n - \mu\|_{\text{TV}} \leq 2\sqrt{2} \left(\eta^{-1} \left(\frac{2M}{\sqrt{2\xi(n)}} \right) \right)^{1/2} \xi(n)^{1/2}, \quad n \in \mathbb{Z}_+,$$

其中 $\xi(n)$ 如 (3.4) 所示, η^{-1} 为 η 的逆, $\eta(x) = \sqrt{x}\varphi(x)$.

证明 对任意的 $K > 0$, 由 (1.4)、Pinsker 不等式 (1.2) 和 (3.3) 知,

$$\begin{aligned} \|\nu P^n - \mu\|_{\text{TV}} &\leq \int \left| P^{n*}(h \wedge K) - \int (h \wedge K) d\mu \right| d\mu + \frac{2M}{\varphi(K)} \\ &\leq \sqrt{2 \left(\int (h \wedge K) d\mu \right) \text{Ent}_\mu(P^{n*}(h \wedge K))} + \frac{2M}{\varphi(K)} \\ &\leq \sqrt{2 \text{Osc}^2(\sqrt{h \wedge K}) \xi(n)} + \frac{2M}{\varphi(K)} \\ &\leq \sqrt{2K \xi(n)} + \frac{2M}{\varphi(K)}. \end{aligned} \quad (3.7)$$

选取 $K > 0$, 使上式最后一个不等式的右端两项相等, 从而解得

$$K = \eta^{-1} \left(\frac{2M}{\sqrt{2\xi(n)}} \right),$$

将其代入 (3.7) 便证得所需结论. □

例 3 若弱对数 Sobolev 不等式 (3.2) 成立且 $h \in L^q(\mu)$, $q > 1$, 则由推论 7 知,

$$\|\nu P^n - \mu\|_{\text{TV}} \leq 2^{(3q-1)/(2q-1)} \left(\int h^q d\mu \right)^{1/(2q-1)} \xi(n)^{(q-1)/(2q-1)}.$$

当 $q \rightarrow \infty$ 时, 收敛速度为 $\xi(n)^{1/2}$, 恰好与 (3.5) 中的收敛速度相同.

致谢 作者感谢毛永华教授对本文的有益指导, 并感谢审稿人对本文初稿提出的宝贵意见.

参考文献

- 1 Meyn S P, Tweedie R L. Markov Chains and Stochastic Stability. London: Springer-Verlag, 1993
- 2 Csiszár I. Information-type measures of difference of probability distributions and indirect observations. *Studia Sci Math Hungar*, 1967, 2: 299–318
- 3 Chen M F. From Markov Chains to Non-Equilibrium Particle Systems. Singapore: World Scientific, 1992
- 4 Chen M F. Eigenvalues, Inequalities and Ergodic Theory. New York: Springer-Verlag, 2004
- 5 Wang F Y. Functional Inequalities, Markov Semigroups and Spectral Theory. Beijing: Science Press, 2005
- 6 Cattiaux P, Gentil I, Guillin A. Weak logarithmic Sobolev inequalities and entropic convergence. *Probab Theory Related Fields*, 2007, 139: 563–603
- 7 Cattiaux P, Guillin A. Trends to equilibrium in total variation distance. *Ann Inst H Poincaré Probab Statist*, 2009, 45: 117–145
- 8 Sokal A D, Thomas L E. Exponential convergence to equilibrium for a class of random-walk models. *J Stat Phys*, 1989, 54: 797–828
- 9 Mengersen K L, Tweedie R L. Rates of convergence of the Hastings and Metropolis algorithms. *Ann Statist*, 1996, 24: 101–121
- 10 Roberts G O. General state space Markov chains and MCMC algorithms. *Probab Surv*, 2004, 1: 20–71
- 11 Röckner M, Wang F Y. Weak Poincaré inequalities and L^2 -convergence rates of Markov semigroups. *J Funct Anal*, 2001, 185: 564–603
- 12 Aida S. Weak Poincaré inequalities on domains defined by Brownian rough paths. *Ann Probab*, 2004, 32: 3116–3137
- 13 Gross L. Logarithmic Sobolev inequalities. *Amer J Math*, 1976, 97: 1061–1083
- 14 Gross L. Logarithmic Sobolev inequalities and contractivity properties of semigroups. In: *Lecture Notes in Math*, vol. 1563. Berlin: Springer-Verlag, 1993, 54–88
- 15 Miclo L. Remarques sur l'hypercontractivité et l'évolution de l'entropie pour des chaînes de Markov finies. *Sém Probab (Strasbourg)*, 1997, 31: 136–167

- 16 Aida S, Masuda T, Shigekawa I. Logarithmic Sobolev inequalities and exponential integrability. *J Funct Anal*, 1994, 126: 83–101
- 17 Diaconis P, Saloff-Coste L. Logarithmic Sobolev inequalities for finite Markov chains. *Ann Appl Probab*, 1996, 6: 695–750
- 18 Deng C S, Song Y H. Weak log-Sobolev and L^p weak Poincaré inequalities for general symmetric forms. *Front Math China*, 2012, 7: 1059–1071

Convergence rates for discrete-time Markov chains in total variation distance

SONG YanHong

Abstract In this paper, we study convergence rates for ergodic discrete-time Markov chains in total variation distance. When the initial distribution of a Markov chain satisfies some integrability conditions, convergence rates are estimated by using Poincaré, weak Poincaré, log-Sobolev and weak log-Sobolev inequalities. Moreover, integrability conditions of the initial distribution are weakened by a truncation method.

Keywords Markov chain, convergence rate, functional inequality

MSC(2010) 60J05, 26D15

doi: 10.1360/N012013-00168