arXiv:2303.04239v1 [math.PR] 7 Mar 2023

又一个定量哈里斯定理

CHRISTOPHER DUPRE

Abstract. 本文提出一个定量哈里斯定理，并对其中的常数进行有效控制。我们的方法有一个优点，即将小集与李雅普诺夫-福斯特漂移条件解耦。只要满足所谓的“定量小性”条件，该方法便允许使用任意小集以及李雅普诺夫-福斯特条件中的任意集合。该定理依赖于一种新颖的定量肯德尔型定理证明，其灵感来自一般状态空间上马尔可夫链的相关技术。我们将此技术应用于混合过程的马尔可夫链近似。

1. Introduction

本文将提出一种针对一般状态空间上马尔可夫链的定量混合定理的新颖证明。对于一般状态空间上马尔可夫链演化的这类极限结论，统称为哈里斯定理，以纪念 T.E. Harris 在二十世纪五十年代的开创性工作 [1]。我们沿袭 Meyn 和 Tweedie 在 [2] 中的工作，将 Doeblin 构造“小”集的思想推广到一般状态空间。随后，我们结合一种称为 Numellin 分裂 [3] 的技术，将问题转化为一个更新理论命题。我们采用的方法基于关于加权范数的混合。这项技术最早见于 [4]。

在混合情形下，这种方法依赖于 Kendall [5]开创的一项关键定理。文献中的大多数证明似乎都依赖复分析技术 [2, 6]。这些证明本身都十分优美，但我们给出了肯德尔定理中一项重要推论的新证明，这正是我们估计所需的。该证明依赖于为一般状态空间马尔可夫链发展出的技术以及一种耦合方法。据我们所知，这似乎是文献中首次采用这种耦合证明。

这项技术已成功用于推导关于混合速率的定性 [2, 7] 和定量[6,8, 9] 估计。此外，这些思想还被多次出色地应用于马尔可夫链蒙特卡罗算法的严格或“诚实”误差界 [10, 11, 12]。

这项工作的动机源于我们考虑将 [13] 扩展到紧状态空间上的一族随机微分同胚，并研究其一致混合速率。在我们的情形中，小集与次水平集的对应关系并不自然，因此我们创建了这些工具。我们希望它们本身也能引起关注。

2. Setting and Main Theorem

本文将研究一般状态空间上的离散时间马尔可夫链。这里指的是，我们有一个可测空间 $(X, \mathcal{B})$，以及一个核 $P(x, \cdot)$，其定义满足

(1) 对每个 $x \in X$， $P(x, \cdot)$ 都是一个概率测度；(2) 对每个 $A \in \mathcal{B}$， $P(x, A)$ 都是一个非负可测函数。

此核编码了我们的马尔可夫过程的一步转移概率。随后通过 Chapman-Kolmogorov方程归纳定义 n 步转移概率。

$$P^n(x, \cdot) := \int_X P^{n-1}(x, dy)P(y, \cdot).$$

我们可以定义 核对可测函数 f 和测度 λ 的作用如下。

$$P^n(\lambda, \cdot) := \lambda P^n(\cdot) := \int_X \lambda(dy)P^N(y, \cdot) \qquad P^n(x, f) := P^n f(x) := \int_X P^N(x, dy)f(y).$$

给定任意初始测度 λ，我们将马尔可夫链 $\{\Phi^n\}_{n\in\mathbb{N}}$ 定义为满足以下分布律条件的随机过程：对于每个递增的指标序列 $i_1 < i_2 < ...i_n$，

$$\mathbb{P}\left(\Phi^{i_1} \in A_1, ...\Phi^{i_n} \in A_n\right) = \int_{A_1}\int_{A_2}...\int_{A_{n-1}} P^{i_1}(\lambda, dy)P^{i_2-i_1}(y, dz)...P^{i_n-i_{n-1}}(w, A_n).$$

我们将使用记号 $\mathbb{P}_x$ 表示初始测度为 δ_x的马尔可夫链所诱导的分布律。我们用 $\mathbb{E}_x$ 表示相对于此分布律积分。对于任意 $C \in \mathcal{B}$，令 τ_C 表示首次到达时间。

我们的目标是为马尔可夫链的混合速率获得定量估计。若$f : X \to \mathbb{R}$ 是我们空间上的可测函数，我们将用 $f(\Phi^n)$表示马尔可夫链经过第 n 步后与函数 f. 的复合。本文的一个重要构造是所谓的 V-加权范数。对于每个可测的 $V\colon X \to [1, \infty)$，我们定义一个作用于 X 上的可测函数的范数如下。

$$\|\phi\|_V := \sup_{x\in X} \frac{|\phi(x)|}{V(x)}.$$

特别地，每个有界可测函数也是 $\|\cdot\|_V$有界的。

定理 2.1（定量Harris定理，[2]）。设 P 是定义在可测空间（X，$\mathcal{B}$）上的马尔可夫转移核，并假设满足以下条件：

（a）（小集）存在一个可测集 U、一个 $\delta > 0$以及 X上的概率测度 μ，使得 $\mu(U) = 1$ 且

$$P(x, B) \geq \delta\mu(B) \text{ for all } x \in U, B \in \mathcal{B}.$$

（b）（漂移条件）存在一个可测函数 $V\colon X \to [1, \infty)$、一个常数 $0 < \lambda < 1$、一个常数 $b < \infty$以及一个集合 $C \in \mathcal{B}$，使得

$$PV \leq \lambda V + b_{\mathbb{I}_C}.$$

（c）（定量小性）存在一个 $N_0 \in \mathbb{N}$ 和一个 $c > 0$，使得

$$\inf_{x\in C} \mathbb{P}_x \{\tau_U \leq N_0\} \geq c.$$

令 $M_U = \sup_{x\in U} V(x)$，并令 $M_C = \sup_{x\in C} V(x)$。则存在一个不变测度 π 和一对有效常数 $D < \infty$、$0 < \gamma < 1$，它们仅（且明确地）取决于 δ、λ、b、N_0、c、M_U、M_C，使得

$$\left| P^n\phi(x) - \int \phi d\pi \right| \le DV(x)\|\phi\|_V \gamma^n.$$

请注意，该定理的若干条件并非必要。U 可以是 n 步链的一个小集，其中 n 可以是任意正整数。测度 μ 不必使 U 具有正质量，只需满足非周期性要求即可。一般而言，证明可以扩展到读者能够想到的任意程度。为清晰起见，我们采用上述条件。

研究一般状态马尔可夫链遍历性的一个关键技术，是将问题归约为更新理论中的一个问题 [2]。因此，我们针对一个与肯德尔定理相关的关键结果，给出了一种新的证明。虽然肯德尔定理通常具有比我们的结果更丰富的结构，但这一结果已足以满足我们的目的。

定理 2.2. 设 $p(n)$ 是一个满足 $p(1) \ge \beta > 0$ 的增量分布，并设 $u(n)$ 是相应的、延迟为 0 的更新序列。若存在一个 $r > 1$ 使得

$$\sum_{n=1}^{\infty} p(n) r^n \le B$$

那么，前向复发时间链 $e(n)$ 存在一个不变分布、一个常数 $\pi(1) = e * u(n)$，以及一个仅（且明确地）取决于 r、β、B 的有效界 ρ，使得对于任意 $r_2 \le \rho$，我们都能找到一个仅（且明确地）取决于 r_2、B、β、ρ 的有效界 L，使得

$$\sum_{n=1}^{\infty} \left| u(n) - \pi(1) \right| r_2^n \le L.$$

这里同样是出于清晰性而非必要性的考虑而作了限制。事实上，对于互素的 n、m，条件 $p(n) > 0$ 且 $p(m) > 0$ 就已经足够。这一点从我们的证明中可以清楚看出，并且不会实质性地影响寻找必要的定量界。我们的证明似乎具有新颖性，因为据我们所知，这是文献中首个通过耦合证明来证明肯德尔定理这一蕴含关系的证明。

Appendix A. Main Tools

本节将介绍我们在两个定理的表述中都会用到的主要工具。第一个工具给出了 r^{τ_C} 的矩母函数的显式界，其中 C 是我们的李雅普诺夫–福斯特漂移条件中的集合。该结果实际上是邓肯公式的一个推论。

定理 2.3（[2] 中的定理 15.2.5）。如果 $V : X \to [1, \infty)$ 满足李雅普诺夫-福斯特漂移条件

对于 $\lambda < 1$、$b < \infty$、$C \in \mathcal{B}$，则对任意 $PV \le \lambda V + b\mathbb{I}_C$ $r \in \left(1, \frac{1}{\lambda}\right)$

$$\frac{\mathbb{E}_x[r^{\tau_C}] - 1}{r - 1} \le \mathbb{E}_x\left[\sum_{k=0}^{\tau_C - 1} V(\Phi^k) r^k\right] \le \frac{1 + rb}{1 - \lambda r} V(x).$$

这使我们能够利用李雅普诺夫-福斯特漂移条件中函数的界，得到到达漂移条件中集合的首次到达时间的显式界。不过，我们需要将这类界从一个集合转移到另一个集合。为此，我们将使用以下结果，它是对 [2] 中定理 15.2.1 的仔细研读所得。

推论 2.4。假设存在一个 $r_1 > 1$ 使得

$$\sup_{x\in C} \mathbb{E}_x \left[\sum_{k=0}^{\tau_C - 1} V(\Phi^k) r_1^k \right] \leq M_0 < \infty.$$

此外，假设对于 $B \in \mathcal{B}$，存在一个 $N_0 \geq 1$ 和一个 $c > 0$，使得

$$\inf_{x\in C} \mathbb{P}_x \{\tau_B \leq N_0\} \geq c.$$

那么，存在一个仅（且明确地）取决于 r_1、M_0、c、N_0 的有效 ρ，使得对于所有 $r_2 < \rho$，我们都可以找到一个仅（且明确地）取决于 r_1、M_0、c、N_0、r_2 的有效常数 D，使得

$$\frac{\mathbb{E}_x \left[r_2^{\tau_B} \right] - 1}{r_2 - 1} \leq \mathbb{E}_x \left[\sum_{k=0}^{\tau_B - 1} V(\Phi^k) r_2^k \right] \leq DV(x)$$

Appendix B. Proof of Theorem 2.2

我们首先介绍更新理论的语言。令 $p(n)\colon \mathbb{N} \to \mathbb{R}_{\geq 0}$ 和 $d(n)\colon \mathbb{N} \cup \{0\} \to \mathbb{R}_{\geq 0}$ 为满足以下条件的非负数列

$$\sum_{n=1}^{\infty} p(n) = \sum_{n=0}^{\infty} d(n) = 1.$$

我们将 $p(n)$ 称为增量分布，将 $d(n)$ 称为延迟分布。

令 $\{Y_i\}_{i=1}^{\infty}$ 为相互独立且同分布的离散随机变量，其分布律为 $\mathbb{P}\{Y_i = n\} = p(n)$；令 Y_0 为取值于 $\mathbb{N} \cup \{0\}$、分布律为 $d(n)$ 的随机变量，且它与所有其他 Y_i 相互独立。然后定义 $S_n^d = \sum_{i=0}^{n} Y_i$，并且

$$V_n^d = \begin{cases} 1 & S_m^d = n \text{ for some m} \\ 0 & \text{else} \end{cases}.$$

相应的更新过程为

$$v_d(n) := \mathbb{P}\left\{ V_n^d = 1 \right\} = \sum_{m=0}^{\infty} d * p^{*m}(n).$$

特别重要的是满足 $d(0) = 1$ 的更新序列。我们用符号 $u(n) := v_0(n)$ 表示这样的序列。于是，我们可以用零延迟序列将任意延迟的更新序列写成 $v_d(n) = d * u(n)$。关于更新序列的更多详情，请读者参阅[14]。

注意， $\sum_{n=1}^{\infty} p(n) r^n < \infty$ 意味着 $\sum_{n=1}^{\infty} p(n) n < \infty$，因为特别地，我们有$\frac{r}{r-1} r^n \geq n$ 对所有 $r > 1$ 都成立。令 $e(n)$ 表示分布律为

$$e(n) := \frac{\sum_{j=n+1}^{\infty} p(j)}{\sum_{n=1}^{\infty} p(n) n} = \frac{1 - \sum_{j=1}^{n} p(j)}{\sum_{n=1}^{\infty} p(n) n}, n \in \mathbb{N} \cup \{0\}.$$

容易看出，这就是前向复发时间链的不变概率测度（更多详情见 [2, 14] ）。特别地，我们有

$$\pi(1) := \mathbb{P}\{V_n^e = 1\} = e * u(n) = \frac{1}{\sum_{n=1}^{\infty} n p(n)}$$

根据更新序列的定义，我们可以将其化简为

$$\left|u(n) - \pi(1)\right| = \left|\mathbb{P}\left\{V^0(n) = 1\right\} - \mathbb{P}\left\{V^e(n) = 1\right\}\right|.$$

为了控制这一差异，我们引入一种耦合。

Defi义 2.5。 $V^0(n)$ $V^e(n)$的耦合时间定义为停止时 e

$$T_{0,e} = \min\left\{n \in \mathbb{N} \cup \{0\} \middle| V^0(n) = V^e(n) = 1\right\}.$$

耦合时间表示两个序列首次同时更新的时刻。注意，更新之后，下一个更新时间的分布由增量序列唯一确定。因此我们将看到，耦合时间使我们能够控制任意给定时刻之后两个序列之间的差异。

为使这一点形式化，我们定义随机变量

$$V^{0,e}(n) = \begin{cases} V^0(n) & n < T_{0,e} \\ V^e(n) & n \geq T_{0,e}. \end{cases}$$

注意，对于任意 n （由于首次同时更新之后分布相同），我们有

$$\mathbb{P}\left(V^{0,e}(n) = 1\right) = \mathbb{P}\left(V^0(n) = 1\right).$$

于是我们有

$$\begin{aligned} \left|u(n) - \pi(1)\right| &= \left|\mathbb{P}\left\{V^0(n) = 1\right\} - \mathbb{P}\left\{V^e(n) = 1\right\}\right| = \left|\mathbb{P}\left\{V^{0,e}(n) = 1\right\} - \mathbb{P}\left\{V^e(n) = 1\right\}\right| \\ &= \Big|\mathbb{P}\left\{V^0(n) = 1, T_{0,e} > n\right\} + \mathbb{P}\left\{V^e(n) = 1, T_{0,e} \leq n\right\} \\ &\quad - \mathbb{P}\left\{V^e(n) = 1, T_{0,e} > n\right\} - \mathbb{P}\left\{V^e(n) = 1, T_{0,e} \leq n\right\}\Big| \\ &\leq \mathbb{P}\{T_{0,e} > n\}. \end{aligned}$$

因此，将其相乘乘以适当的 r 次幂（稍后确定）并求和，我们得到 t

$$\sum_{n=1}^{\infty} \left|u(n) - \pi(1)\right| r^n \leq \sum_{n=1}^{\infty} r^n \mathbb{P}\{T_{0,e} > n\} \leq \frac{r\mathbb{E}[r^{T_{0,e}}]}{r-1} = \frac{r}{r-1} \sum_{n=1}^{\infty} e(n) \mathbb{E}[r^{T_{0,n}}]. \tag{2.1}$$

因此，我们只需证明零延迟分布与常数延迟分布的首次耦合时间具有几何有界尾部。

受一般状态空间上的马尔可夫链理论（例如定理 2.3 和推论 2.4）中的方法启发，我们注意到，返回时间的矩母函数的有界性与李雅普诺夫函数的存在性有关。因此，我们希望将问题表述为某个状态空间上的马尔可夫链问题，然后找到一个适当的李雅普诺夫函数。

我们首先考虑所谓的前向复发时间链。这是定义在状态空间 $\mathbb{N}$ 上、带有离散 σ-代数的马尔可夫链，其转移核为

$$P(n, n-1) = 1 \text{ if } n > 1, P(1, m) = p(m).$$

分别考察两个更新序列的演化，我们可以将系统的状态想象为 $\mathbb{N} \times \mathbb{N}$ 中的一个点。初始延迟为 0、n 时，对应的初始点为 $(1, n+1)$。于是，系统在 $\mathbb{N} \times \mathbb{N}$ 上诱导出一个转移核如下定义的马尔可夫链

$$\begin{aligned} P\left((a,b),(a-1,b-1)\right) &= 1 \text{ if } a, b \neq 1 \\ P\left((1,b),(n,b-1)\right) &= p(n) \text{ if } a = 1, b \neq 1 \\ P\left((a,1),(a-1,n)\right) &= p(n) \text{ if } a \neq 1, b = 1 \\ P\left((1,1),(n,m)\right) &= p(n)p(m) \text{ if } a, b = 1. \end{aligned}$$

同时更新等价于该马尔可夫链达到 $(1,1)$。因此我们有$\mathbb{E}[r^{T_{0,n}}] = \mathbb{E}_{1,\ n+1}\left[r^{\tau_{1,1}}\right]$.

现在考虑函数 $V(a,b) = \frac{r^{a-1}+r^{b-1}}{2}$。通过直接计算，我们得到以下关系。

$$\begin{aligned} P\left((a,b),V\right) &= \frac{1}{r}V(a,b) && \text{若 } a, b \neq 1 \\ P\left((1,b),V\right) &= \frac{1}{2r}\left(\sum_{n=1}^{\infty} p(n)r^n + r^{b-1}\right) && a = 1, b \neq 1 \\ P\left((a,1),V\right) &= \frac{1}{2r}\left(r^{a-1} + \sum_{n=1}^{\infty} p(n)r^n\right) && a = 1, b \neq 1 \\ P\left((1,1),V\right) &= \frac{1}{r}\sum_{n=1}^{\infty} p(n)r^n && a = 1, b \neq 1. \end{aligned}$$

我们有 $\sum_{n=1}^{\infty} p(n)r^n < \infty$，因此至少可以确定这些都定义良好。我们希望在李雅普诺夫函数中构造的集合 C 是有限的，原因将在后文讨论：我们需要控制从集合 C 到 $(1, 1)$ 的访问。因此，点 $(1, 1)$ 无需担心，但射线 $(1, n)$、$(n, 1)$ 似乎都存在问题。不过，我们有

$$\frac{P\left((a,1),V\right)}{V(a,1)} = \frac{r^{a-1} + \sum_{n=1}^{\infty} p(n)r^n}{r^a + r}.$$

对所有 $r>1$，利用 $\sum_{n=1}^{\infty}p(n)r^n\geq r$，我们可以将这个分数小于$\frac{1}{r}<\eta<1$ 的条件改写如下。

$$\begin{aligned}\frac{P\left((a,1),V\right)}{V(a,1)}\leq\eta &\iff \frac{r^{a-1}+\sum_{n=1}^{\infty}p(n)r^n}{r^a+r}\leq\eta\\ &\iff \sum_{n=1}^{\infty}p(n)r^n-\eta r\leq r^{a-1}\left(\eta r-1\right)\\ &\iff a\geq\frac{1}{\ln(r)}\ln\left(\frac{\sum_{n=1}^{\infty}p(n)r^n-\eta r}{\eta r-1}\right)+1.\end{aligned}$$

在 $a=1,b\neq 0$的情况下，同样的计算也适用。因此，我们得到如下李雅普诺夫-福斯特漂移条件。

$$\begin{aligned}M&:=\left\lceil\frac{1}{\ln(r)}\ln\left(\frac{\sum_{n=1}^{\infty}p(n)r^n-\eta r}{\eta r-1}\right)+1\right\rceil\\ C&:=\{(1,b)\in\mathbb{N}^2\text{ such that }b\leq M\}\cup\{(a,1)\in\mathbb{N}^2\text{ such that }a\leq M\}\\ PV&\leq\eta V+\frac{\max\left\{2\sum_{n=1}^{\infty}p(n)r^n,\sum_{n=1}^{\infty}p(n)r^n+r^{M-1}\right\}}{2r}\mathbb{I}^C.\end{aligned}$$

特别地，利用定理（2.3），我们得到以下界。

$$\sup_{x\in C}\mathbb{E}_x\left[\sum_{k=0}^{\tau_C-1}V(\Phi^k)r^k\right]\leq\frac{2+\max\left\{2\sum_{n=1}^{\infty}p(n)r^n,\sum_{n=1}^{\infty}p(n)r^n+r^{M-1}\right\}}{4(1-\eta r)}(r^{M-1}+1).$$

现在，我们希望将这个关于 τ_C 的界转化为关于$\tau_{1,\ 1}$的界。从马尔可夫链的构造可以看出，$\inf_{x\in C}\mathbb{P}_x\{\tau_{1,1}\leq M\}\geq p(1)^M$.这个界将起到定量小性条件的作用，正如哈里斯定理的陈述中所述；根据推论 2.4，该条件使我们能够将关于 τ_C 的界转化为关于$\tau_{1,\ 1}$ 的界。因此，我们现在有了一个有效常数 D 和一个有效的新r，并且有如下界

$$\mathbb{E}[r^{T_{0,n}}]=\mathbb{E}_{1,n+1}[r^{\tau_{1,1}}]\leq DV(1,n+1)=\frac{D(1+r^n)}{2}\leq\frac{D}{2}\left(\frac{1}{r-1}+1\right)r^n=\frac{Dr}{2(r-1)}r^n.$$

我们利用了以下事实：对于所有 $n\geq 1$，对所有 $r>1$，都有$\frac{r^n}{r-1}>1$。将这些结果相加，便得到如下结论。

$$\begin{aligned}\mathbb{E}[r^{T_{0,e}}]=\sum_{n=0}^{\infty}e(n)\mathbb{E}[r^{T_{0,n}}]&\leq\frac{Dr}{2(r-1)}\sum_{n=0}^{\infty}e(n)r^n=\frac{Dr}{2(r-1)}\frac{1}{\sum_{n=1}^{\infty}np(n)}\left[\sum_{n=0}^{\infty}\sum_{j=n+1}^{\infty}p(j)r^n\right]\\ &\leq\frac{Dr}{2(r-1)}\left[\sum_{j=1}^{\infty}\sum_{n=0}^{j-1}p(j)r^n\right]\leq\frac{Dr}{2(r-1)^2}\left[\sum_{j=1}^{\infty}p(j)r^j\right].\end{aligned}$$

将此结果代回 (2.1) 即可得到所需结论。

Appendix C. Proof of Theorem 2.1

首先注意到，通过缩放，只需考虑满足 $\|\phi\|_V \leq 1.$ 的 ϕ。我们采用耦合方法。我们希望控制两个不同的量

(1) τ_U上的几何尾部；(2) 构造一个所谓的“原子”，将我们的问题转化为更新理论问题。

定理 2.3 和推论 2.4 中的主要工具为实现第一个结果提供了一种有效方法。结合定理中的 M_U、M_C、c、N_0，我们可以得到从任意初始点到达 U 的首次到达时间具有几何尾部。

现在回到我们的第二个目标。我们回顾原子的定义。

定义 2.6. 设 $P(x, dy)$ 是可测空间 $(X, \mathcal{B})$ 上的一个转移核。如果存在一个概率测度 ν，使得 $\alpha \in \mathcal{B}$ 满足

$$P(x, B) = \nu(B) \text{ for all } x \in \alpha.$$

我们将保留符号 α 表示一个原子。当 α 是一个原子时，我们将记 $P(\alpha, B)$ 表示对于任意 $x \in \alpha$，$P(x, B)$，或等价地表示 $\nu(B)$。如果某个停时涉及一个原子，我们指的是带有原子的马尔可夫链。下一段中，我们将使用 Nummelin 分裂技术 [3] 构造这样一条链。随后，我们会将几何尾部转移到新构造的原子上。

接下来，我们在 U 处通过 $\delta/2$ 分裂来分裂该链。具体来说，我们考虑配备乘积 σ-代数的扩展空间 $X \times \{0, 1\}$。采用如下记号：$A_0 = A \times \{0\}$、$A_1 = A \times \{1\}$ $A \in \mathcal{B}$、$x_i = x \times \{i\}$，其中 $x \in X$。给定 $(X, \mathcal{B})$ 上的任意测度 λ，我们将在这个新空间上构造一个测度，如下所示：

$$\lambda^*(A_0) = \left(1 - \frac{\delta}{2}\right) \lambda(A \cap U) + \lambda(A \cap U^c)$$
$$\lambda^*(A_1) = \frac{\delta}{2} \lambda(A \cap U).$$

ˆ 类似地，给定任意函数 $V\colon X \to \mathbb{R}$，我们可以在分裂链上定义函数 $V\colon X \times \{0, 1\} \to \mathbb{R}$，定义为 $V(x_i) := V(x)$。现在，我们可以将测度扩展到这个新空间，因此接下来将转而扩展转移概率。这样，我们就能从原始链在这个新空间上构造一个马尔可夫链。

ˆ 我们定义分裂核 $P(x_i, A)$，用于 $x_i = (x, i) \in X \times \{0, 1\}$ 和 $A \in \mathcal{B} \times \{0, 1\}$，如下所示。

$$\hat{P}(x_i, A) = \begin{cases} P^*(x, A) & x_i \in X_0 \backslash U_0 \\ \frac{2P^*(x,A) - \delta\mu^*(A)}{2-\delta} & x_i \in U_0 \\ \mu^*(A) & x_i \in X_1. \end{cases}$$

这一构造中至关重要的一点是，我们原来的链恰好是所构造链的边缘分布，因此分裂链的 $\tau_{A_1 \cup A_0}$ 恰好等于原链的 τ_A。另一个至关重要的一点是，U_1 现在构成了这条新马尔可夫链的一个原子。有关更多

马尔可夫链语境下分裂方法的信息，请参阅 [2]第5章。

构造的另一个重要方面是以下估计。

$$\begin{aligned} x_i \in U_1 &\implies \hat{P}(x_i, U_1) = \mu^*(U_1) = \frac{\delta}{2} \\ x_i \in U_0 &\implies \hat{P}(x_i, U_1) = \frac{2P^*(x, U_1) - \delta\mu^*(U_1)}{2-\delta} \\ &= \frac{\delta\left(P(x,U) - \frac{\delta}{2}\mu(U)\right)}{2-\delta} \geq \frac{\delta}{2-\delta}\frac{\delta}{2}. \end{aligned}$$

无论哪种情况，我们都有 $\mathbb{P}_{x_i}\left(\tau_{U_1} = 1\right) \geq \frac{\delta}{2-\delta}\frac{\delta}{2}$（注意，这始终小于或等于 $\frac{\delta}{2}$，因为必然有 $0 < \delta \leq 1$）。

现在，我们从 $U_0 \cup U_1$ 到 U_1 获得了定量控制，因此可以再次应用推论 2.4，证明 U_1 是一个原子，且对所有 $x \in X_0 \cup X_1$ 都有 $\mathbb{E}_x\left[r^{\tau_{U_1}}\right] < \infty$，并且在 U_1 上有界，其中界由某个可显式计算的常数 r 给出。我们的目标是将对原子的首次占用时间的几何尾部控制，与对 V-加权距离的控制联系起来。证明这部分的关键工具是所谓的“再生”分解，亦称“首次进入—最后离开”分解。

定义 2.7. 给定一个转移核 $P(x, dy)$，它在可测空间 $(X, \mathcal{B})$ 上生成马尔可夫链 $\{\Phi^n\}_{n=1}^{\infty}$，以及一个 $B \in \mathcal{B}$，该集合的禁入概率是指在避开 B 的同时到达另一个集合的概率。更形式化地说，它定义为

$$ {}_B P^N(x, A) = \mathbb{P}_x\left\{\Phi^n \in A, \tau_B \geq n\right\}.$$

根据以下关系，也可以用归纳方式表示。

$$\begin{aligned} {}_B P(x, A) &= P(x, A) \\ {}_B P^{N+1}(x, A) &= \int_{B^c} {}_B P^{N+1}(x, dy) P(y, A). \end{aligned}$$

定义 2.8. 给定可测空间 $(X, \mathcal{B})$ 上的转移核 $P(x, dy)$，以及一个 $B \in \mathcal{B}$，我们将该核的再生分解定义如下

$$P^n(x, A) = {}_B P^n(x, C) + \sum_{j=1}^{n-1} \int_B \left[\sum_{k=1}^{j} \int_B {}_B P^k(x, dv) P^{j-k}(v, dw)\right] {}_B P^{n-j}(w, A).$$

直观而言，再生分解追踪进入给定集合的首次时刻和离开该集合的末次时刻。每条轨迹都可以分为两类：在 n 步内从未与该集合相交，或在不到n 步内与该集合相交。这些轨迹还可以根据离开前最后一次访问B 的时间长短进一步分类，并且离开后在前 n 次迭代中不再返回。请注意，末次离开时刻通常不是停时。因此，我们每次只考虑前 n 步，这一点至关重要。

如果集合 $B = \alpha$ 是一个原子，并且我们考虑的是在分裂链上定义的马尔可夫链，那么这一过程就可以从一系列积分简化为如下的求和。

$$\hat{P}^n(x, C) = {}_\alpha\hat{P}^n(x, C) + \sum_{j=1}^{n-1}\sum_{k=1}^{j} {}_\alpha\hat{P}^k(x, \alpha)\hat{P}^{j-k}(\alpha, \alpha)\,{}_\alpha\hat{P}^{n-j}(\alpha, C)$$

接下来，我们可以考虑它与一个固定测度（例如不变测度）之差。我们必须先验证不变测度存在。由于我们已经知道 $\mathbb{E}_\alpha[r_\alpha^\tau]$ 有界，因此 $\mathbb{E}_\alpha[\tau_\alpha]$ 也有限。于是，根据 [2], 中的定理 10.2.1，分裂链存在一个不变测度。另一个关键事实是 $\hat{}\pi(\alpha) = \mathbb{E}_\alpha[\tau_\alpha]^{-1}$.

直接应用 [2], 中的定理 10.4.9，可得

$$\pi(g) = \pi(\alpha)\mathbb{E}_\alpha\left[\sum_{k=0}^{\tau_\alpha - 1} g\left(\hat{\Phi}^k\right)\right] = \pi(\alpha)\sum_{k=1}^{\infty}\mathbb{E}_\alpha\left[g\left(\hat{\Phi}^k\right)\mathbb{I}^{\tau_\alpha \geq k}\right] = \pi(\alpha)\sum_{k=1}^{\infty} {}_\alpha\hat{P}^k(\alpha, g).$$

基于这两种分解，我们引入三个值序列。

$$\begin{aligned} a_x(n) &:= \mathbb{P}_x\{\tau_\alpha = n\} = {}_\alpha\hat{P}^n(x, \alpha) \\ u(n) &:= \mathbb{P}_\alpha\left\{\hat{\Phi}^n \in \alpha\right\} = \hat{P}^n(\alpha, \alpha) \\ t_g(n) &:= \mathbb{E}_\alpha\left[g\left(\hat{\Phi}^n\right)\mathbb{I}^{\tau_\alpha \geq n}\right] = {}_\alpha\hat{P}^n(\alpha, g) \end{aligned}$$

合并这些项，可得以下界。

$$\begin{aligned} \hat{P}^n(x, g) &= {}_\alpha\hat{P}^n(x, g) + a_x * u * t_g(n) \\ \hat{\pi}(g) = \hat{\pi}(\alpha)\sum_{k=1}^{\infty} {}_\alpha\hat{P}^k(\alpha, g) &= \hat{\pi}(\alpha) * t_g(n) + \hat{\pi}(\alpha)\sum_{k=n+1}^{\infty} {}_\alpha\hat{P}^k(\alpha, g) \end{aligned}$$

我们稍微滥用了记号，将 $\hat{}\pi(\alpha) * t_g(n)$ 称为 $t_g(n)$ 与值恒为 $\hat{}\pi(\alpha)$ 的常数序列的卷积。由此得到以下有用的界。

$$\left|\hat{P}^n(x, g) - \hat{\pi}(g)\right| \leq {}_\alpha P^n(x, g) + \left|a_x * u - \hat{\pi}(\alpha)\right| * t_g(n) + \hat{\pi}(\alpha)\sum_{k=n+1}^{\infty} {}_\alpha P^k(\alpha, g).$$

乘以 r^n 并求和，可得如下结果。

$$\begin{aligned} \sum_{n=1}^{\infty}\left|\hat{P}^n(x, g) - \hat{\pi}(g)\right| r^n \leq \sum_{n=1}^{\infty} {}_\alpha\hat{P}^n(x, g)r^n + \sum_{n=1}^{\infty}\left|a_x * u - \hat{\pi}(\alpha)\right| * t_g(n)r^n \\ + \hat{\pi}(\alpha)\sum_{n=1}^{\infty}\sum_{k=n+1}^{\infty} {}_\alpha\hat{P}^k(\alpha, g)r^n. \end{aligned}$$

我们现在的目标是对这些项进行界定。根据 [2], 第 360—361 页中的界，我们可对所有 $\|g\|_{\hat{V}} \leq 1$ 得到以下结果。

$$\sum_{n=1}^{\infty}\left|\hat{P}^n(x,g)-\hat{\pi}(g)\right| r^n \leq \mathbb{E}_x\left[\sum_{n=1}^{\tau_\alpha}\hat{V}(\Phi^n)r^n\right]+\frac{1}{r-1}\mathbb{E}_\alpha\left[\sum_{n=1}^{\tau_\alpha}\hat{V}(\Phi^n)r^n\right] + \mathbb{E}_\alpha[r^{\tau_\alpha}]\left[\frac{r}{r-1}+\sum_{n=1}^{\infty}\left|u(n)-\hat{\pi}(\alpha)\right| r^n\right]$$

现在我们的目标是对这些项进行界定。

首先注意到

$$\mathbb{E}_x\left[\sum_{n=1}^{\tau_\alpha}\hat{V}(\Phi^n)r^n\right] \leq \sup_{x\in\alpha}\hat{V}(x)\mathbb{E}_x[r^{\tau_\alpha}]+\mathbb{E}_x\left[\sum_{n=0}^{\tau_\alpha-1}\hat{V}(\Phi^n)r^n\right]$$

$$\mathbb{E}_\alpha\left[\sum_{n=1}^{\tau_\alpha}\hat{V}(\Phi^n)r^n\right] \leq \sup_{x\in\alpha}\hat{V}(x)\mathbb{E}_\alpha[r^{\tau_\alpha}]+\mathbb{E}_\alpha\left[\sum_{n=0}^{\tau_\alpha-1}\hat{V}(\Phi^n)r^n\right]$$

根据定义，

$$\sup_{x\in\alpha}\hat{V}(x)=\sup_{x\in U_1}\hat{V}(x)=\sup_{x\in U}V(x)=M_U.$$

其余各项均可由有效常数乘以 $V(x)$ 来界定 .

现在我们来处理最后一项。根据构造可知， $P(\alpha,\alpha) = p(1) \geq \frac{\delta}{2}$且$\mathbb{E}_\alpha[r^{\tau_\alpha}]<\infty.$。直接比较形式还可得， $\pi(1) = \hat{}\pi(\alpha)$。因此，由定理 2.2 可知，存在一个有效常数 $D<\infty$，使得 $\sum_{n=1}^{\infty}\left|u(n) - \hat{}\pi(\alpha)\right| r^n \leq D$。同样，我们可以用一个有效常数界定 $\mathbb{E}[r^{\tau_\alpha}]$。由于 $V(x) \geq 1$，我们可以将所有相关项乘以它而不失一般性，从而得到一个有效常数 D，使得对所有满足$\|g\|_{\hat{V}} \leq 1$ 的 g，都有如下界。

$$\sum_{n=1}^{\infty}\left|\hat{P}^n(x,g)-\hat{\pi}(g)\right| r^n \leq DV(x)$$

$$\sup_{g:\|g\|_{\hat{V}}\leq 1}\left|\hat{P}^n(x,g)-\hat{\pi}(g)\right| \leq DV(x)r^{-n}.$$

利用原始马尔可夫链是此链的边缘分布这一事实，我们得到的最终界与 [2] 第 327 页中的结论完全相同。 π 此处表示 $\hat{}\pi$ 的边缘分布。根据 [2] 中的定理 10.4.1，这是原始马尔可夫过程的一个不变测度。由此便得到该定理。

$$\sup_{g:\|g\|_V\leq 1}\left|P^n(x,g)-\pi(g)\right| \leq \sup_{g:\|g\|_{\hat{V}}\leq 1}\left|\hat{P}^n(x,g)-\hat{\pi}(g)\right| \leq DV(x)r^{-n}.$$

References

[1] Theodore E Harris. “某些马尔可夫过程的平稳测度存在性”。载于：Matematika 4.1 (1960)，第 131–143 页。

[2] Sean P Meyn 和 Richard L Tweedie。《马尔可夫链与随机稳定性》。Springer Science & Business Media，2012 年。[3] Esa Nummelin. “Harris常返马尔可夫链的分裂技术”。载于：Zeitschrift f ¨ur Wahrscheinlichkeitstheorie und verwandte Gebiete 43.4 (1978)，第 309–318 页。[4] Arie Hordijk 和 Flora Spieksma. “关于马尔可夫链的遍历性与常返性质及其在开放Jackson网络中的应用”。载于：应用概率进展 24.2 (1992)，第 343–376 页。[5] David G Kendall. “马尔可夫转移算子的酉扩张及转移概率矩阵相应的积分表示”。载于：概率与统计 (1959)，第 139–161 页。[6] Peter H Baxendale. “几何遍历马尔可夫链的更新理论与可计算收敛速率”。载于：应用概率年刊 15.1B (2005)，第 700–738 页。[7] Jeffrey S Rosenthal. “马尔可夫链蒙特卡罗的小化条件与收敛速率”。载于：美国统计学会会刊 90.430 (1995)，第 558–566 页。[8] Witold Bednorz. “Kendall定理及其在马尔可夫链几何遍历性中的应用”。英语。载于：数学应用 40.2 (2013)，第 129–165 页。 URL: http://eudml.org/doc

[9] Yu Hang Jiang 等人。《马尔可夫链收敛速率的耦合/小集/漂移方法》。载于《美国数学会通报》68.4（2020）。[10] Krzysztof Latuszy ´nski 和 Wojciech Niemiro。《几何漂移条件下马尔可夫链蒙特卡罗的严格置信界》。载于《复杂性杂志》27.1（2011），第 23–38 页。[11] Galin L Jones 和 James P Hobert。《通过马尔可夫链蒙特卡罗诚实地探索难以处理的概率分布》。载于《统计科学》（2001），第 312–334 页。[12] Luke Tierney。《用于探索后验分布的马尔可夫链》。载于《统计年刊》（1994），第 1701–1728 页。[13] Alex Blumenthal、Michele Coti Zelati 和 Rishabh S Gvalani。《随机动力系统的指数混合及 Pierrehumbert 的一个例子》。载于 arXiv 预印本 arXiv:2204.13651（2022）。[14] Torgny Lindvall。《耦合法讲义》。Courier Corporation，2002。

Department of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332
Email address: cdupre3@gatech.edu