Mixing and hitting times for Markov chains

Overview 1) Equivalence (up to constants) between mixing times and hitting times of large sets

2) Hitting times: comparison for different sizes of sets

3) Refined mixing and hitting equivalence.

Let X be an irreducible Markov chain in a finite state space S.

Let P be the transition matrix of X.

$P^t(i,j) = \mathbb{P}_i(X_t = j) \quad \forall\, i,j \in S.$

π : invariant distr. , $\pi = \pi P$.

if X is also aperiodic, then $P^t(x,y) \longrightarrow \pi(y)$ as $t \to \infty$, $\forall\, x,y$

Let μ and ν be 2 prob. distr. on S.

$$\|\mu - \nu\|_{TV} = \max_{A \subset S} |\mu(A) - \nu(A)|$$

$$d(t) = \max_x \|P^t(x,\cdot) - \pi\|_{TV}.$$

$\forall\, \varepsilon \in (0,1) \qquad t_{mix}(\varepsilon) = \min\{t \geq 0 : d(t) \leq \varepsilon\}$

$t_{mix} = t_{mix}\left(\frac{1}{4}\right)$.

X is called reversible $\forall\, x,y \quad \pi(x)P(x,y) = \pi(y)P(y,x)$

$t_H(\alpha) = \max_{x, A : \pi(A) \geq \alpha} \mathbb{E}_x[\tau_A]$, where $\tau_A = \min\{t \geq 0 : X_t \in A\}$.

lazy version of X $\quad P_L = \frac{P+I}{2}$.

Theorem 1 (Oliveira, Peres-S., 2012)

$\forall\, \alpha < \frac{1}{2}$ $\exists$ positive constants c_α and c'_α s.t. for all reversible lazy Markov chains $\quad c_\alpha t_H(\alpha) \leq t_{mix} \leq c'_\alpha t_H(\alpha)$.

$[\ t_{mix} \asymp_\alpha t_H(\alpha)\ ]$

Proof of lower bound $t_{mix} \geq C_\alpha t_H(\alpha)$ $\alpha = \frac{1}{8}$

Let $t = t_{mix}\left(\frac{1}{16}\right) \leq 3\, t_{mix}$

$\forall x, A$ $P^t(x,A) \geq \pi(A) - \frac{1}{16}$.

Take A with $\pi(A) \geq \frac{1}{8}$, then $P^t(x,A) \geq \frac{1}{16}$ $\forall x$.

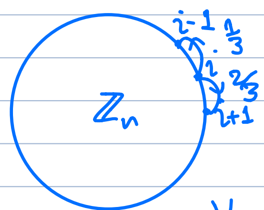
So $\tau_A \leq t \cdot \mathrm{Geo}\left(\frac{1}{16}\right) \Rightarrow \max_x \mathbb{E}_x[\tau_A] \leq 16\, t$. $\square$

Aldous '82 For all reversible lazy MC's

$$t_{mix} \asymp \max_{x, A} \pi(A)\, \mathbb{E}_x[\tau_A]$$

Remark Reversibility is essential!

Exercise 1 Consider a biased RW on $\mathbb{Z}_n$ (+ laziness)

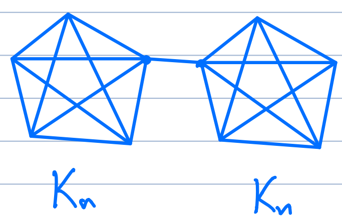


$P_L = \frac{P+I}{2}$

Show $t_{mix} \asymp n^2$ and $\forall \alpha$ $t_H(\alpha) \asymp n$.

Remark If $\alpha > \frac{1}{2}$, then the theorem is false.

Exercise 2



Consider a SRW on the 2 cliques joined by an edge.

Take its lazy version.

Show that $t_{mix} \asymp n^2$ and $t_H(\alpha) \asymp n$ if $\alpha > \frac{1}{2}$.

Proof of Theorem 1 (Upper bound)

Definition Mixing at a geometric time

Let Z_t be a geometric r.v. of parameter $\frac{1}{t}$ taking values in $\{1, \ldots\}$ and indep. of X.

Define $d_G(t) = \max_x \| P_x(X_{Z_t} = \cdot) - \pi \|_{TV}$

and $t_G = \min\{t \geq 0 : d_G(t) \leq \frac{1}{4}\}$: geometric mixing.

Remark If instead of geometric, we take U_t to be uniform on $\{1, \dots, t\}$, then this gives rise to the Cesaro mixing time.

Exercise 3 Show that $d_G(t)$ is decreasing in t.

Theorem 2 For all lazy reversible chains, $t_G \asymp t_{mix}$. $\leftarrow$ Ideas — Aldous; Lovász and Winkler

Theorem 3 For all chains, $t_G \asymp_\alpha t_H(\alpha)$ $\forall \alpha < \frac{1}{2}$.

Pf of Thm 1 Immediate from Thm's 2 and 3. $\square$

Pf of Thm 3 $t_G \underset{\text{up to constants}}{\gtrsim_\alpha} t_H(\alpha)$: easy.

We prove $t_G \lesssim_\alpha t_H(\alpha)$. $\alpha = \frac{1}{8}$

Let $t < t_G$. We want to find a set B with $\pi(B) \geq \frac{1}{8}$ s.t.

$$\max_x \mathbb{E}_x[\tau_B] \geq \vartheta t \text{ for some positive constant } \vartheta.$$

$$t < t_G \Rightarrow \exists z, A \text{ s.t. } \mathbb{P}_z(X_{Z_t} \in A) < \pi(A) - \frac{1}{4}$$

$$\Rightarrow \pi(A) > \frac{1}{4}.$$

$$B = \{y : \mathbb{P}_y(X_{Z_t} \in A) \geq \pi(A) - \tfrac{1}{8}\}.$$

<u>Claim</u> $\pi(B) \geq \frac{1}{8}$

$$\pi = \pi P \Rightarrow \pi(A) = \underbrace{\sum_{y \in B} \pi(y) \underbrace{\mathbb{P}_y(X_{Z_t} \in A)}_{\leq 1}}_{\leq \pi(B)} + \sum_{y \in B^c} \underbrace{\pi(y)}_{\leq 1} \underbrace{\mathbb{P}_y(X_{Z_t} \in A)}_{\leq \pi(A) - \frac{1}{8}}$$

$$\pi(A) \leq \pi(B) + \pi(A) - \frac{1}{8} \Rightarrow \pi(B) \geq \frac{1}{8}. \quad \square$$

We will prove that assuming $\mathbb{E}_z[\tau_B] \leq \vartheta t$ for a suitable constant ϑ leads to a contradiction.

By Markov's ineq. $\mathbb{P}_z(\tau_B \geq 2\vartheta M t) \leq \frac{1}{2M}$ $\star$

$$M \in \mathbb{N}$$

$$\mathbb{P}_z(X_{Z_t} \in A) \geq \underbrace{\mathbb{P}_z(X_{Z_t} \in A \mid Z_t \geq \tau_B, \tau_B < 2\vartheta M t)}_{\geq \min_{y \in B} \mathbb{P}_y(X_{Z_t} \in A)} \mathbb{P}_z(Z_t \geq \tau_B, \tau_B < 2\vartheta M t)$$

memoryless property of Z_t and strong Markov at τ_B

$$\geq \left(\pi(A) - \frac{1}{8}\right) \cdot \mathbb{P}_z\left(Z_t > 2\theta Mt, \tau_B < 2\theta Mt\right)$$

$$= \mathbb{P}_z(Z_t > 2\theta Mt) \cdot \mathbb{P}_z(\tau_B < 2\theta Mt)$$

$$\star \quad \geq \left(\pi(A) - \frac{1}{8}\right) \cdot \left(1 - \frac{1}{t}\right)^{2\theta Mt} \cdot \left(1 - \frac{1}{2M}\right)$$

$2\theta Mt > 1$

$$\geq \left(\pi(A) - \frac{1}{8}\right) \left(1 - 2\theta M\right) \left(1 - \frac{1}{2M}\right)$$

Choosing $\theta = \frac{1}{4M^2}$ $\rightsquigarrow$ $\mathbb{P}_z(X_{Z_t} \in A) \geq \left(\pi(A) - \frac{1}{8}\right)\left(1 - \frac{1}{2M}\right)^2$

Taking M large enough shows $\mathbb{P}_z(X_{Z_t} \in A) > \pi(A) - \frac{1}{4}$ which is a contradiction. $\square$

Idea of geometric mixing : due to Oded Schramm

$$t_{stop} = \max_x \min\left\{ \mathbb{E}_x[\Lambda_x] : \Lambda_x \text{ is a randomised stopping time s.t. } \mathbb{P}_x(X_{\Lambda_x} = \cdot) = \pi(\cdot) \right\}.$$

filling rule Baxter and Chacon '76

Aldous, Lovász - Winkler.

Thm 2 reversible $t_{stop} \asymp t_{mix}$

$t_{stop} \leq 8 t_{mix}$ easy.

The hard direction is to show $t_{stop} \gtrsim t_{mix}$.

Exercise 4 Prove that for reversible chains $t_{stop} \leq 8 t_{mix}$.

Hint : Use separation distance to define an appropriate stopping time.

4. Let X be a reversible Markov chain with transition matrix P and invariant distribution π.

(i) Prove that for all x, y

$$\frac{P^{2t}(x,y)}{\pi(y)} \geq \left(1 - \max_{z,w} \|P^t(z,\cdot) - P^t(w,\cdot)\|_{\mathrm{TV}}\right)^2.$$

Deduce that

$$P^{2t_{\mathrm{mix}}}(x,y) \geq \frac{1}{4}\pi(y)$$

and that there exists a transition matrix $\widetilde{P}$ such that

$$P^{2t_{\mathrm{mix}}}(x,y) = \frac{1}{4}\pi(y) + \frac{3}{4}\widetilde{P}(x,y)$$

(ii) Let $t_{\mathrm{stop}} = \max_x \min\{\mathbb{E}_x[\Lambda_x] : \Lambda_x \text{ is a stopping time s.t. } \mathbb{P}_x(X_{\Lambda_x} \in \cdot) = \pi(\cdot)\}$. By defining an appropriate stationary time, prove that

$$t_{\mathrm{stop}} \leq 8 t_{\mathrm{mix}}.$$