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 $X$.

\[ P^t(i,j) = P_i(X^t=j) \quad \forall i,j \in S. \]

$\pi$: invariant distr., $\pi = \pi P$.

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

Let $\mu$ and $\nu$ be 2 prob. distr. on $S$.

\[ d_{tv}(\mu, \nu) = \max_A |\mu(A) - \nu(A)|. \]

\[ d(t) = \max_x \| P^t(x, \cdot) - \pi \|_{tv}. \]

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

$\text{twix} = \text{twix}(\frac{\varepsilon}{2})$.

\[ X \text{ is called reversible } \iff \forall x, y \quad \pi(x) P(x, y) = \pi(y) P(y, x). \]

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

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

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

$\forall \alpha < \frac{1}{2}$, $\exists$ positive constants $C_{\alpha}$ and $C'_{\alpha}$ s.t. for all reversible lazy Markov chains $C_{\alpha} \tau_H(x) \leq \text{twix} \leq C'_{\alpha} \tau_H(x)$.

\[ \left[ \text{twix} \leq C_{\alpha} \tau_H(x) \right] \]
**Proof of lower bound**

Let $t = \text{twix}(\frac{1}{16}) < 3 \text{twix}$

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

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

So $\tau_A \leq t \cdot \text{Geo}(\frac{1}{16}) \Rightarrow \max_x E_x[\tau_A] \leq 16 \cdot t$. \( \square \)

**Remark** Reversibility is essential!

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

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

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

and $t_G = \min \{ t > 0 : d_G(t) \leq \frac{1}{4} \}$: geometric mixing.
Remark If instead of geometric, we take $U_t$ to be uniform on $[1,...,t^2]$ then this gives rise to the Cesaro mixing time.  

Exercise 3 Show that $d_{C_{\alpha}}(t)$ is decreasing in $t$. 

Theorem 2 For all reversible chains, $t_\alpha \geq t_{mix}$. \textit{Ideas Aldous Lovász and Winkler}  

Theorem 3 For all chains, $t_\alpha \geq t_{mix} \ orall \alpha < \frac{1}{8}$. 

Pf of Thm 1 Immediate from Thm's 2 and 3. \textbf{D}  

Pf of Thm 3 $t_{Ca} \geq t_{mix}(\alpha)$: easy, up to constants 

We prove $t_{Ca} \leq t_{mix}(\alpha)$, $\alpha = \frac{1}{8}$  

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

$$\max_{x} \mathbb{E}_x[\tau_{\text{mix}}] > \Theta t$$  

for some positive constant $\Theta$. 

$t < t_{Ca} \Rightarrow \exists \varepsilon, A$ s.t. $P_\varepsilon(X_{\tau_{\text{mix}}} \in A) < \pi(A) - \frac{1}{4}$  

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

$B = \{ y : P_y(X_{\tau_{\text{mix}}} \in A) > \pi(A) - \frac{1}{8} \}$  

Claim $\pi(B) > \frac{1}{8}$ 

$$\pi = \pi \mathcal{P} \Rightarrow \pi(A) = \sum_{y \in B} \pi(y) P_y(X_{\tau_{\text{mix}}} \in A) + \sum_{y \notin B} \pi(y) P_y(X_{\tau_{\text{mix}}} \in A)$$  

$$\leq 1 \pi(B) + \pi(A) - \frac{1}{8} \to \pi(B) > \frac{1}{8}.$$ \textbf{D}  

We will prove that assuming $\mathbb{E}_x[\tau_{\text{mix}}] < \Theta t$ for a suitable constant $\Theta$ leads to a contradiction.  

By Markov's ineq. $P_\varepsilon(\tau_{\text{mix}} > 2\Theta Mt) < \frac{1}{2M}$ \textbf{M} $\in \mathbb{N}$  

$$P_{\varepsilon}(X_{\tau_{\text{mix}}} \in A) \geq P_{\varepsilon}(X_{\tau_{\text{mix}}} \in A|Z_t \geq \tau_{\text{mix}}, \tau_{\text{mix}} < 2\Theta Mt) P_{\varepsilon}(Z_t \geq \tau_{\text{mix}}, \tau_{\text{mix}} < 2\Theta Mt)$$  

$\geq \min_{y \in B} P_y(X_{\tau_{\text{mix}}} \in A)$ memoryless property of $Z_t$ and strong Markov at $\tau_{\text{mix}}$
\[
\begin{align*}
\mathbb{P}(\exists t \geq \frac{\pi(A)}{8} \cdot \mathbb{P}(\mathcal{Z}_t \geq 2\theta M, \mathcal{Z}_B < 2\theta M)) &= \mathbb{P}(\mathcal{Z}_t \geq 2\theta M) \cdot \mathbb{P}(\mathcal{Z}_B < 2\theta M) \\
* &\geq \left(\pi(A) - \frac{1}{8}\right) \left(1 - \frac{1}{2M}\right) \\
\theta M > 1 &\geq \left(\pi(A) - \frac{1}{8}\right) \left(1 - 2\theta M\right) \left(1 - \frac{1}{2M}\right) \\
\text{Choosing } &\theta = \frac{1}{4M^2} \Rightarrow \mathbb{P}(\mathcal{Z}_t \in A) \geq \left(\pi(A) - \frac{1}{8}\right) \left(1 - \frac{1}{2M}\right)^2
\end{align*}
\]

Taking \(M\) large enough shows \(\mathbb{P}(\mathcal{Z}_t \in A) > \pi(A) - \frac{1}{4}\)

which is a contradiction. \(\square\)

**Idea of geometric mixing** : due to Oded Schramm

\(t_{\text{stop}} = \max \min \{ \mathbb{E}_x [\Lambda_x] : \Lambda_x \text{ is a randomised stopping time s.t. } \mathbb{P}_x(\mathcal{Z}_{\Lambda_x} = \cdot) = \pi(\cdot) \}\)

**Filling rule** Baxter and Chacon 76 Aldous, Lovász-Winkler.

Thus \(2\) reversible \(t_{\text{stop}} \leq \text{twix}\)

\(t_{\text{stop}} \leq 8\text{twix} \; \text{easy.}\)

The hard direction is to show \(t_{\text{stop}} \geq \text{twix}.\)

**Exercise 4** Prove that for reversible chains \(t_{\text{stop}} \leq 8\text{twix}.\)

**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 \(\pi\).

(i) Prove that for all \(x, y\)

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

Deduce that

\[
P^{2t_{\text{mix}}}(x, y) \geq \frac{1}{4} \pi(y)
\]

and that there exists a transition matrix \(\hat{P}\) such that

\[
P^{2t_{\text{mix}}}(x, y) = \frac{1}{4} \pi(y) + \frac{3}{4} \hat{P}(x, y)
\]

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

\[
t_{\text{stop}} \leq t_{\text{mix}}.
\]