Set hitting times in Markov chains

Ross J. Kang*

Utrecht University

Birmingham Combinatorics Seminar
11/2013

*With S. Griffiths (Oxford), R. Oliveira (IMPA) and V. Patel (Birmingham).
Support from NWO.

Introduction

$X = (X_t)_{t=1}^{\infty}$ irreducible discrete-time Markov chain on finite state space Ω, transition matrix P, stationary dist. π; law of X from $x \in \Omega$ is $\mathbb{P}_x(\cdot)$.

Introduction

$X = (X_t)_{t=1}^{\infty}$ irreducible discrete-time Markov chain on finite state space Ω, transition matrix P, stationary dist. π; law of X from $x \in \Omega$ is $\mathbb{P}_x(\cdot)$.

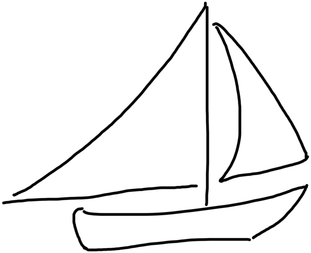
The *hitting time* τ_A of $A \subseteq \Omega$ is $\min\{t : X_t \in A\}$.

Extremal problem of max mean hitting time over 'large enough' A: for $0 < \alpha < 1$,

$$T(\alpha) = \max_{x \in \Omega, A \subseteq \Omega} \{\mathbb{E}_x(\tau_A) : \pi(A) \geq \alpha\}.$$

A fanciful example

Imagine meandering (compassless, mapless, drunken) sailor X.



What is worst-case time expected to reach some island A?
A large island? A continent?

Mixing time

Many other (more relevant) examples from statistical physics, network analysis, machine learning, card shuffling, ...

Fundamental property if X ergodic is time to get near stationarity, *mixing time*[†]

$$t_{\mathrm{mix}} = \min\left\{ t : \forall x \in \Omega,\ \forall A \subseteq \Omega,\ |P^t(x,A) - \pi(A)| \leq \frac{1}{4} \right\}.$$

Usually for applications, the faster the better.

(If X periodic, we use weaker notion, *Cesàro mixing time*.)

[†] The choice of constant $\frac{1}{4}$ is essentially irrelevant.

Hitting and mixing

For lazy, reversible[‡] X, mixing time is equivalent to the following hitting time parameter:

$$t_{\mathrm{prod}} = \max_{x\in\Omega, A\subseteq\Omega} \left\{\pi(A)\mathbb{E}_x(\tau_A) : A \neq \emptyset\right\}.$$

Theorem (Aldous, 1982)

$\exists C > 0$ *such that* $\frac{1}{C} t_{\mathrm{prod}} \leq t_{\mathrm{mix}} \leq C t_{\mathrm{prod}}$ *if* X *lazy, reversible.*

Later expanded (including Cesàro analogue without laziness, reversibility) by Aldous, Lovász & Winkler (1997), Lovász & Winkler (1998).

[‡] X is *reversible* if $\exists\pi$, $\pi(i)P^t(i,j) = \pi(j)P^t(j,i)$ for all $i,j \in \Omega$ and $t \geq 0$. By lazy, we mean in the sense that $P_{xx} \geq \frac{1}{2}$ for all $x \in \Omega$.

Hitting and mixing

Intuitively, X has not mixed until it has hit all large enough sets.

Does mixing time depend on hitting times of arbitrarily small sets?

Hitting and mixing

Intuitively, X has not mixed until it has hit all large enough sets.

Does mixing time depend on hitting times of arbitrarily small sets?

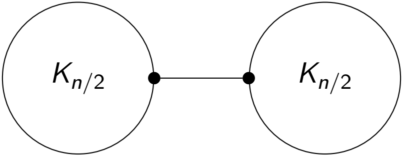
No, can restrict attention to large enough sets. . .

Theorem (Oliveira, 2012, and Peres & Sousi, 2011+/14?)

$\forall \alpha \in (0, \frac{1}{2})$, $\exists C > 0$ *such that* $\frac{1}{C} T(\alpha) \leq t_{\mathrm{mix}} \leq C T(\alpha)$ *if* X *lazy, reversible.*

Hitting and mixing

...but not too large. Consider SSRW on



$t_{\mathrm{mix}} = \Omega(n^2)$ while $T(\frac{1}{2} + \varepsilon) = O(n)$

$\implies$ hitting/mixing connection fails if all sets too large, i.e. statement in previous theorem false when $\alpha > \frac{1}{2}$.

Hitting and mixing

What about $\alpha = \frac{1}{2}$?

Question (Peres, 2007–)

$\exists C > 0$ *such that* $\frac{1}{C} T(\frac{1}{2}) \leq t_{\mathrm{mix}} \leq C T(\frac{1}{2})$ *if* X *lazy, reversible?*

Hitting large sets

These connections and this question led us to study

$$T(\alpha) = \max_{x\in\Omega, A\subseteq\Omega} \left\{\mathbb{E}_x(\tau_A) : \pi(A) \geq \alpha\right\}.$$

Note $T(\alpha) \geq T(\beta)$ for $0 < \alpha < \beta < 1$.

Question ('Extremal ratio problem')

Let $0 < \alpha < \beta < 1$. *Over all* X, *how large can* $T(\alpha)/T(\beta)$ *be?*

Hitting large sets

These connections and this question led us to study

$$T(\alpha) = \max_{x\in\Omega, A\subseteq\Omega} \{\mathbb{E}_x(\tau_A) : \pi(A) \geq \alpha\}.$$

Note $T(\alpha) \geq T(\beta)$ for $0 < \alpha < \beta < 1$.

$\Longrightarrow$ 'It can take longer to reach smaller islands.'

Question ('Extremal ratio problem')

Let $0 < \alpha < \beta < 1$. *Over all* X, *how large can* $T(\alpha)/T(\beta)$ *be?*

$\Longrightarrow$ 'How much longer?'

Hitting large sets

Question ('Extremal ratio problem')

Let $0 < \alpha < \beta < 1$*. Over all* X*, how large can* $T(\alpha)/T(\beta)$ *be?*

Hitting large sets

Question ('Extremal ratio problem')

Let $0 < \alpha < \beta < 1$*. Over all* X*, how large can* $T(\alpha)/T(\beta)$ *be?*

Cesàro version of Oliveira/Peres & Sousi result implies

Corollary

Let $0 < \alpha < \beta < \frac{1}{2}$*.* $\exists C_\beta > 0$ *s.t.* $T(\alpha) \leq C_\beta \dfrac{T(\beta)}{\alpha}$ *for any* X*.*

Main theorem

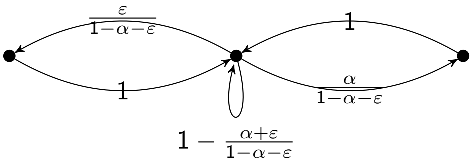
Theorem (Griffiths, K., Oliveira & Patel, 2012+/2014?)

Let $0 < \alpha < \beta \leq \frac{1}{2}$. *For any* X,

$$T(\alpha) \leq T(\beta) + \left(\frac{1}{\alpha} - 1\right) T(1-\beta) \leq \frac{T(\beta)}{\alpha}. \qquad (\star)$$

Sharpness of the theorem

Given $0 < \alpha < \beta \leq \frac{1}{2}$, for some small $\varepsilon > 0$, consider

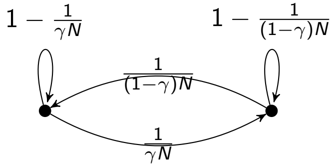


Check $\pi = (\varepsilon, 1 - \alpha - \varepsilon, \alpha)$, $T(\beta) = 1$ and $T(\alpha) = \frac{1}{\alpha}$.

$\implies T(\alpha) = \dfrac{T(\beta)}{\alpha}$, meeting $(\star)$ with equality.

Sharpness of the theorem

Given $0 < \alpha < \beta < 1$, $\beta > \frac{1}{2}$, let $\max\{\alpha, \frac{1}{2}\} < \gamma < \beta$ and N large. Consider



Check $\pi = (\gamma, 1-\gamma)$, $T(\beta) = 0$ and $T(\alpha) \geq (1-\gamma)N$.

$\implies$ no constant bound in extremal ratio problem when $\beta > \frac{1}{2}$.

An application of the theorem

Corollary

$\exists C > 0$ such that $\frac{1}{C} T(\frac{1}{2}) \le t_{\mathrm{mix}} \le C T(\frac{1}{2})$ if X lazy, reversible.

An application of the theorem

Corollary

$\exists C > 0$ *such that* $\frac{1}{C} T(\frac{1}{2}) \leq t_{\mathrm{mix}} \leq C T(\frac{1}{2})$ *if* X *lazy, reversible.*

Proof.

Note $\frac{T(\frac{1}{2})}{2} \leq \max_{x \in \Omega, A \subseteq \Omega} \left\{ \pi(A) \mathbb{E}_x(\tau_A) : \pi(A) \geq \frac{1}{2} \right\} \leq t_{\mathrm{prod}}$.

Also, $\pi(A)\mathbb{E}_x(\tau_A) \leq \pi(A) T(\pi(A)) \leq T(\frac{1}{2})$ for all $A \subseteq \Omega$ (by theorem if $\pi(A) \leq \frac{1}{2}$ and monotonicity of T otherwise). □

An ergodic property

Given $A, C \subseteq \Omega$, define

$$d^+(A,C) = \max_{x\in A} \mathbb{E}_x(\tau_C) \quad \text{and} \quad d^-(C,A) = \min_{x\in C} \mathbb{E}_x(\tau_A).$$

Lemma

For any chain X and $A, C \subseteq \Omega$,

$$\pi(A) \le \frac{d^+(A,C)}{d^+(A,C) + d^-(C,A)}.$$

Proof of theorem

Fix $x \in \Omega$, $A \subseteq \Omega$ with $\pi(A) \geq \alpha$. Suffices to prove

$$\mathbb{E}_x(\tau_A) \leq T(\beta) + \left(\frac{1}{\alpha} - 1\right) T(1-\beta).$$

Define set

$$C = \left\{ y \in \Omega : \mathbb{E}_y(\tau_A) > \left(\frac{1}{\alpha} - 1\right) T(1-\beta) \right\}$$

By definition, 'hard' to get from C to A. Also, $\pi(C) < 1 - \beta$:

Proof of theorem

Fix $x \in \Omega$, $A \subseteq \Omega$ with $\pi(A) \geq \alpha$. Suffices to prove

$$\mathbb{E}_x(\tau_A) \leq T(\beta) + \left(\frac{1}{\alpha} - 1\right) T(1-\beta).$$

Define set

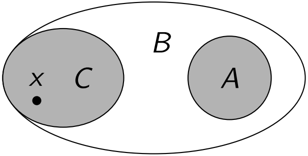
$$C = \left\{ y \in \Omega : \mathbb{E}_y(\tau_A) > \left(\frac{1}{\alpha} - 1\right) T(1-\beta) \right\}$$

By definition, 'hard' to get from C to A. Also, $\pi(C) < 1 - \beta$:

Suppose, for ※, that $\pi(C) \geq 1 - \beta$. Then $d^+(A, C) \leq T(1-\beta)$ while $d^-(C, A) > \left(\frac{1}{\alpha} - 1\right) T(1-\beta)$. Lemma implies $\pi(A) < \alpha$, ※.

Proof of theorem

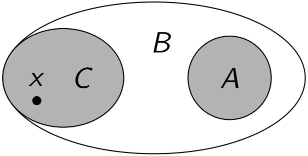
Let $B = \Omega \setminus C$.



- 'easy' to get from x to B: $\mathbb{E}_x(\tau_B) \leq T(\beta)$ as $\pi(B) > \beta$;
- 'easy' to get from B to A: $d^+(B, A) \leq \left(\frac{1}{\alpha} - 1\right) T(1 - \beta)$;
- from x to A, we must hit B.

Proof of theorem

Let $B = \Omega \setminus C$.



- 'easy' to get from x to B: $\mathbb{E}_x(\tau_B) \leq T(\beta)$ as $\pi(B) > \beta$;
- 'easy' to get from B to A: $d^+(B, A) \leq \left(\frac{1}{\alpha} - 1\right) T(1-\beta)$;
- from x to A, we must hit B.

By Markovian property of X,

$$\mathbb{E}_x(\tau_A) \leq \mathbb{E}_x(\tau_B) + d^+(B, A) \leq T(\beta) + \left(\frac{1}{\alpha} - 1\right) T(1-\beta). \quad \square$$

An ergodic property

Lemma

$$\pi(A) \leq \frac{d^+(A,C)}{d^+(A,C) + d^-(C,A)},$$

where $d^+(A,C) = \max_{x\in A} \mathbb{E}_x(\tau_C)$, $d^-(C,A) = \min_{x\in C} \mathbb{E}_x(\tau_A)$.

Proof outline.

Martingale concentration + ergodic theorem

OR

Auxiliary chain simulates stationary hitting behaviour $A \leftrightarrow C$. $\square$

Shape problem

We already saw tightness in two senses, but we may ask more.

Question ('Shape problem')

Besides decreasing and $(\star)$*, what other constraints on* $T(\alpha)$*,* $\alpha \in (0, \frac{1}{2}]$*, for all chains* X *on at least two states?*

Shape problem

Let $\mathcal{F}$ be all decreasing functions $f : (0, \frac{1}{2}] \to \mathbb{R}$ given by $f(\alpha) = \dfrac{T(\alpha)}{T(\frac{1}{2})}$ for some X on at least two states.

Let $\overline{\mathcal{F}}$ be all decreasing functions $f : (0, \frac{1}{2}] \to \mathbb{R}$ which are obtained by the almost everywhere pointwise limit of functions from $\mathcal{F}$.

Question ('Shape problem')

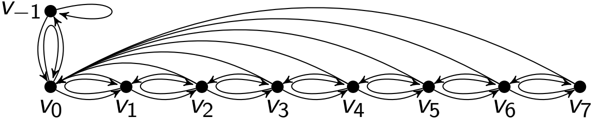
Does $(\star)$ characterise $\overline{\mathcal{F}}$?

Shape theorem

Theorem (GKOP)

Let $f : (0, \frac{1}{2}] \to \mathbb{R}$ *be a decreasing function.*
Then $f \in \overline{\mathcal{F}}$ *iff* $f(\frac{1}{2}) = 1$ *and* $f(\alpha) \leq \frac{1}{\alpha}$ *for all* $\alpha \in (0, \frac{1}{2})$.

L-shaped chains



Hitting time functions $\frac{T(\alpha)}{T(\frac{1}{2})}$ approximate any step function of form

$$f_n(x) = f\left(\frac{\lceil 2^n x\rceil}{2^n}\right),$$

where $f(\frac{1}{2}) = 1$ and $f(\alpha) \le \frac{1}{\alpha}$ for all $\alpha \in (0, \frac{1}{2})$. Then let $n \to \infty$.

An additional contraint

We restricted our domain to $(0, \frac{1}{2}]$, but what about domain $(0, 1)$?

Theorem (GKOP)

For any X, if $T(0.01) = 99.9T(0.02)$, then ____________________.

An additional contraint

We restricted our domain to $(0, \frac{1}{2}]$, but what about domain $(0, 1)$?

Theorem (GKOP)

For any X, if $T(0.01) = 99.9\,T(0.02)$, then ____________________.

Notes

- There are L-shaped X with $T(0.01) = 99.9\,T(0.02)$.
- $(\star)$ implies $T(0.01) \leq T(0.02) + 99\,T(0.98)$, and so

$$T(0.98) \geq \frac{98.9}{99} T(0.02).$$

An additional contraint

We restricted our domain to $(0, \frac{1}{2}]$, but what about domain $(0, 1)$?

Theorem (GKOP)

For any X, if $T(0.01) = 99.9\,T(0.02)$, then $T(0.99) \geq 0.1\,T(0.02)$.

Notes

- There are L-shaped X with $T(0.01) = 99.9\,T(0.02)$.
- $(\star)$ implies $T(0.01) \leq T(0.02) + 99\,T(0.98)$, and so

$$T(0.98) \geq \frac{98.9}{99}\,T(0.02).$$

Further investigation

1. Shape problem for domain $(0, 1)$?
2. Connect to (analogues of) other properties of Markov chains, e.g. cover time, blanket times, . . . ?