21-801 Chapter 10

[Hitting Times]

☆def: given a markov chain with state space Ω, the hitting time τ_A of a subset $A \subseteq \Omega$ is the first time one of the nodes in A is visited by the chain.

So if (X_t) is a random walk,

$$\tau_A = \min\{t \geq 0 : X_t \in A\}$$

→ recall $\tau^x = \min\{t \geq 1 : X_t = x\}$ from section 1.5.3 (called the first return time when $X_0 = x$) was used to build a stationary distribution.

10.2 Random Target Times:

lemma 10.1: (Random Target Lemma)

for an irreducible Markov Chain with state space Ω, transition matrix P, and stationary distribution π, the quantity $\sum_{x \in \Omega} E_a(\tau_x)\pi(x)$ does not depend on $a \in \Omega$.

proof: let $h_x(a) = E_a(\tau_x)$.

if $x \neq a$,

$$h_x(a) = \sum_{y \in \Omega} E_a(\tau_x \mid X_1 = y)P(a,y) = \sum_{y \in \Omega}(1 + h_x(y))P(a,y) = (Ph_x)(a) + 1$$

so that $(Ph_x)(a) = h_x(a) - 1$

if $x = a$, then

$$E_a(\tau_a^+) = \sum_{y \in \Omega} E_a(\tau_a^+ \mid X_1 = y)P(a,y) = \sum_{y \in \Omega}(1 + h_a(y))P(a,y) = 1 + (Ph_a)(a)$$

cont. →

proof (continued)

since $E_a(\tau_a^+) = \pi(a)^{-1}$,

$$(Ph_a)(a) = \frac{1}{\pi(a)} - 1$$

thus letting $h(a) := \sum_{x\in\Omega} h_x(a)\pi(x)$ we have that

$$(Ph)(a) = \sum_{x\in\Omega} (Ph_x)(a)\pi(x) = \sum_{x\neq a} (h_x(a)-1)\pi(x) + \pi(a)\left(\frac{1}{\pi(a)} - 1\right)$$

now setting $h_a(a) = 0$ and simplifying, we have $(Ph)(a) = h(a)$ and thus h is harmonic.

→ from lemma 1.16 we can show h is a constant function.

∎

note: lemma 10.1 implies that if we choose a state $y \in \Omega$ according to π, the expected time to hit the "random target" state y starting from a state a does not depend on a.

☆ def: the target time of an irreducible chain is defined to be

$$t_\odot := \sum_{x\in\Omega} E_a(\tau_x)\pi(x) = E_\pi(\tau_\pi)$$

→ since $t_\odot$ doesn't depend on starting state a,

$$t_\odot = \sum_{x,y\in\Omega} \pi(x)\pi(y)E_x(\tau_y) = E_\pi(\tau_\pi)$$

→ we define the worst-case hitting times between states of the chain to be

$$t_{hit} := \max_{x,y\in\Omega} E_x(\tau_y)$$

□ lemma 10.2: for an irreducible Markov Chain with state space Ω and stationary distribution π,

$$t_{hit} \leq 2 \max_{w} E_\pi(\tau_w)$$

<u>proof</u>: for all $a, y \in \Omega$ we have

$E_a(\tau_y) \leq E_a(\tau_\pi) + E_\pi(\tau_y)$ since the chain goes from a to y via a random state x chosen according to π. By the previous lemma we have that

$$E_a(\tau_\pi) = E_\pi(\tau_\pi) \leq \max_{w} E_\pi(\tau_w)$$

and putting these two equations together we get

$$E_a(\tau_y) \leq E_a(\tau_\pi) + E_\pi(\tau_y) = E_\pi(\tau_\pi) + E_\pi(\tau_\pi) = 2E_\pi(\tau_\pi)$$

$$\leq 2\max_{w} E_\pi(\tau_w)$$

so we get the desired result. ▨

<u>note</u>: for a transitive chain, for any b

$$t_\odot = E_\pi(\tau_\pi) = \sum_{x \in \Omega} E_a(\tau_x)\pi(x) = \sum_{x,y \in \Omega} \pi(y) E_y(\tau_x)\pi(x) = E_\pi(\tau_b)$$

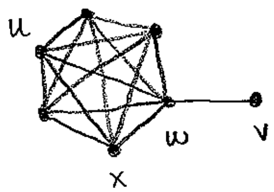
□ corollary: for an irreducible transitive markov chain,

$$t_{hit} \leq 2t_\odot$$

<u>note</u>: when the chain isn't transitive, t_{hit} might be a lot larger than $t_\odot$

[ex 10.4]

consider a random walk on a complete graph on n vertices with a leaf attached to one vertex.



let v be the leaf and let w be the neighbor of the leaf.

☆ def: we call the other vertices ordinary.

let the initial state of the walk be v. Now the first return time to v satisfies both

$$E_v \tau_v^+ = E_v \tau_w + E_w \tau_v = 1 + E_w \tau_v \quad \text{and}$$

$$E_v \tau_v^+ = \frac{1}{\pi(v)} = \frac{2\left(\binom{n}{2}+1\right)}{1} = n^2 - n + 2$$

therefore $E_w \tau_v = n^2 - n + 1 \leq t_{hit}$.

Now by the random target lemma we can start to estimate $t_\odot$. So starting at v, $E_v \tau_v = 0$ while $E_v \tau_w = 1$ and $E_v \tau_u = 1 + E_w \tau_u$ for ordinary vertex u.

→ how long does it take to get from w to u on average?

let x be another ordinary vertex. by conditioning on the first step of the walk and exploiting symmetry,

$$E_w \tau_u = 1 + \frac{1}{n}\left(E_v \tau_u + (n-2) E_x \tau_u\right)$$

$$= 1 + \frac{1}{n}\left(1 + E_w \tau_u + (n-2) E_x \tau_u\right) \quad \text{and}$$

$$E_x \tau_u = 1 + \frac{1}{n-1}\left(E_w \tau_u + (n-3) E_x \tau_u\right).$$

so this gives us a system of equations of two "variables" $E_w \tau_u$ and $E_x \tau_u$

(continued)

solving this yeilds

$$E_w T_u = \frac{n^2-n+4}{n} = \Theta(n) \quad \text{and} \quad E_x T_u = \frac{n^2-n+2}{n} = \Theta(n)$$

so now considering only $E_w T_u$

$$t_\odot = E_v T_\pi = \pi(v)(0) + \pi(w)(1) + (n-1)\pi(u)\Theta(n)$$

$$= \frac{1(0) + n(1) + (n-1)^2\Theta(n)}{2\left(\binom{n}{2}+1\right)} = \Theta(n) \ll t_{hit}$$

10.3 Commute Time:

☆ def: the <u>commute time</u> between nodes a and b in a network is the time it takes to move from a to b then back to a:

$$T_{a,b} = \min\{t \geq T_b : X_t = a\}$$

assuming $X_0 = a$.

→ commute time can be computed/estimated using resistance (commute time identity)

<u>note</u>: in graphs for which $E_a(T_b) = E_b(T_a)$, the expected hitting time is half the commute time.

⇒ estimating the commute time can estimate the hitting time.

(Commute Time Identity)

□ lemma 10.5: If τ is a stopping time for a finite and irreducible Markov chain satisfying $P_a\{X_\tau = a\} = 1$ and Green's function $G_\tau(a,x)$, then $\forall x$

$$\frac{G_\tau(a,x)}{E_a(\tau)} = \pi(x) .$$

(exercise 10.1 asks us to prove this).

→ recall $G_\tau(a,x) = E_a(\#\text{ of visits to } x \text{ before } \tau)$

□ proposition 10.6 (commute time identity)
let $(G, \{c(e)\})$ be a network and let (X_t) be the random walk on this network. For any nodes a and b in V, let $\tau_{a,b}$ be the commute time between a and b. Then

$$E_a(\tau_{a,b}) = E_a(\tau_b) + E_b(\tau_a) = c_G R(a \leftrightarrow b)$$

→ recall $c(x) = \sum_{y \sim x} c(x,y)$ and $c_G = \sum_{x \in V} c(x) = 2 \sum_{e \in E} c(e)$

<u>proof</u>:
by lemma 10.5, $\dfrac{G_{\tau_{a,b}}(a,a)}{E_a(\tau_{a,b})} = \pi(a) = \dfrac{c(a)}{c_G}$ so by definition, after visiting b, the chain does not visit a until time $\tau_{a,b}$ so $G_{\tau_{a,b}}(a,a) = G_{\tau_b}(a,a)$. Now by lemma 9.6 we have $G_{\tau_b}(a,a) = c(a) R(a \leftrightarrow z)$.

cont →

proof (continued)

So by putting this together we have that

$$\frac{G_{\tau_b}(a,a)}{E_a(\tau_{a,b})} = \frac{c(a)}{c_G} \quad \text{and} \quad G_{\tau_b}(a,a) = c(a)R(a \leftrightarrow b)$$

and therefore $E_a(\tau_{a,b}) = c_G R(a \leftrightarrow b)$ which can be found by solving for $E_a(\tau_{a,b})$.

☆def: a network $\langle G, \{c(e)\} \rangle$ is transitive if for any pair of vertices $x, y \in V$, $\exists$ a permutation $\psi_{x,y} : V \to V$ with $\psi_{x,y}(x) = y$ and $c(\psi_{x,y}(u), \psi_{x,y}(v)) = c(u,v)$ $\forall u, v \in V$.

□ corollary: resistance R satisfies the triangle inequality; if a, b, c are vertices, then $R(a \leftrightarrow c) \le R(a \leftrightarrow b) + R(b \leftrightarrow c)$.

□ lemma 10.10: for any three states a, b, c of a reversible Markov chain,

$$E_a(\tau_b) + E_b(\tau_c) + E_c(\tau_a) = E_a(\tau_c) + E_c(\tau_b) + E_b(\tau_a)$$

or rewording this $E_a(\tau_{bca}) = E_a(\tau_{cba})$.

→ we're not going to prove this.

□ proposition 10.9:

for a random walk on a transitive connected graph G, for any vertices $a, b \in V$,

$$E_a(\tau_b) = E_b(\tau_a).$$

proof of prop 10.9 :

let ψ be the mapping previously defined, with $u=a$ and $v=b$. let $a_0=a$, $a_j=\psi^{(j)}(a_0)$ for $j\geq 1$, where $\psi^{(j)}$ denotes the jth iterate of ψ.

the sequence $a_0 a_1 \ldots$ will return to a_0 eventually; so say $a_m=a_0$ $m>0$. the function $\psi^{(j)}$ takes a,b to a_j, a_{j+1}. So for any j, $E_{a_j}(\tau_{a_{j+1}}) = E_a(\tau_b)$ and summing over j from 0 to $m-1$, we obtain $E_{a_0}(\tau_{a_1 a_2 \ldots a_{m-1} a_0}) = m E_a(\tau_b)$ and similarly $E_{a_0}(\tau_{a_{m-1} a_{m-2} \ldots a_1 a_0}) = m E_b(\tau_a)$.

So we can see the left hand side of these two equations are the same, and so our result follows.

■

10.5: Bounding Mixing Times via Hitting Times

GOAL: we want to prove

theorem 10.14: consider the finite reversible chain with transition matrix P and stationary distribution π on Ω

1) for all $m\geq 0$ and $x\in\Omega$

$$\|P^m(x,\cdot)-\pi\|_{TV}^2 \leq \frac{1}{4}\left[\frac{P^{2m}(x,x)}{\pi(x)} - 1\right]$$

11) if the chain satisfies $P(x,x)\geq \frac{1}{2}$ $\forall x$, then

$$t_{mix}\left(\tfrac{1}{4}\right) \leq 2\max_{x\in\Omega} E_\pi(\tau_x)+1$$

[remark]: note that by conditioning on X_0,

$$E_\pi(\tau_x) = \sum_{y\in\Omega} E_y(\tau_x)\pi(y) \leq \max_{y\in\Omega} E_y(\tau_x) \leq t_{hit}$$

□ proposition 10.18:

let P be the transition matrix for a finite reversible chain on state space Ω with stationary distribution π.

i) $\forall\ t \geq 0,\ x \in \Omega \quad P^{2t+2}(x,x) \leq P^{2t}(x,x)$

ii) If chain P_L is lazy, that is $P_L(x,x) \geq \frac{1}{2}\ \forall x$, then $\forall t \geq 0,\ x \in \Omega$ we have that $P_L^{t+1}(x,x) \leq P_L^{t}(x,x)$

proof:

i) since $P^{2t+2}(x,x) = \sum_{y,z \in \Omega} P^t(x,y)P^2(y,z)P^t(z,x)$ we have that

$$\pi(X)P^{2t+2}(x,x) = \sum_{y,z \in \Omega} P^t(y,x)\pi(y)P^2(y,z)P(z,x) = \sum_{y,z \in \Omega} \Psi(y,z)\Psi(z,y)$$

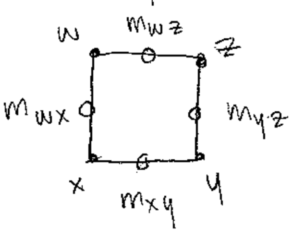
where $\Psi(y,z) = P^t(y,z)\sqrt{\pi(y)P^2(y,z)}$.

then by cauchy-schwartz, the right hand side is at most

$$\sum_{y,z \in \Omega} \Psi(y,z)^2 = \sum [P^t(y,X)]^2 \pi(y) = \pi(x)P^{2t}(x,x)$$

so this gives that $\pi(X)P^{2t+2}(x,x) \leq \pi(X)P^{2t}(x,x)$ and thus $P^{2t+2}(x,x) \leq P^{2t}(x,x)$.

ii) now given a lazy chain $P_L = (P+I)/2$, enlarge the state space by adding a new state $m_{xy} = m_{yx}$ for each pair $x,y \in \Omega$.



On the larger state space Ω_K, define a transition matrix k by

$$k(x, m_{xy}) = P(x,y) \quad x,y \in \Omega$$

$$k(m_{xy}, x) = k(m_{xy}, y) = 1/2 \quad x \neq y$$

$$k(m_{xy}, x) = 1 \quad \forall x$$

and other transitions having k-probability 0. Then k is reversible with stationary measure π_K given by $\pi_K(X) = \frac{\pi(X)}{2}$ for $x \in \Omega$ and $\pi_K(m_{xy}) = \pi(X)P(x,y)$.

cont →

proof of prop. 10.18 (continued)

So we can see $K^2(x,y) = P_L(x,y)$ for $x,y \in \Omega$, so $K^{2t}(x,y) = P_L^t(x,y)$ and so by monotonicity the claim follows. ∎

▫ proposition 10.19: (hitting time from stationary)
consider a finite irreducible aperiodic chain with transition matrix P with stationary distribution π on Ω. then for any $x \in \Omega$,

$$\pi(x) E_\pi(\tau_x) = \sum_{t=0}^{\infty} [P^t(x,x) - \pi(x)]$$

proof: (by stopping times)
define $\tau_x^{(m)} := \min\{t \geq m : X_t = x\}$ and write $\mu_m := P^m(x,\cdot)$.
Now by the convergence theorem (4.9) μ_m tends to π as $m \to \infty$. So by the previous lemma,

$$\sum_{k=0}^{m-1} P^k(x,x) = \pi(x) E_x(\tau_x^{(m)}) = \pi(x)[m + E_{\mu_m}(\tau_x)]$$

thus $\sum_{k=0}^{m-1} [P^k(x,x) - \pi(x)] = \pi(x) E_{\mu_m}(\tau_x)$.

now by taking $m \to \infty$, we get the desired result. ∎

proof of theorem!

1) by cauchy-schwartz we have

$$\left(\frac{1}{2}\sum_{y\in\Omega} \pi(y) \left|\frac{P^m(x,y)}{\pi(y)} - 1\right|\right)^2 \leq \frac{1}{4}\sum_{y\in\Omega} \pi(y)\left[\frac{P^m(x,y)}{\pi(y)} - 1\right]^2$$

therefore

$$\|P^m(x,\cdot) - \pi\|_{TV}^2 \leq \frac{1}{4}\sum_{y\in\Omega}\left[\frac{P^m(x,y)P^m(y,x)}{\pi(x)} - 2P^m(x,y) + 1\right]$$

$$= \frac{1}{4}\left[\frac{P^{2m}(x,x)}{\pi(x)} - 1\right]$$

cont. →

proof of theorem (continued)

ii) by prop. 10.19 and monotonicity in 10.18 ii), for any $m>0$ we have

$$\pi(x) E_\pi(\tau_x) \geq \sum_{k=1}^{2m} \left[P^k(x,x) - \pi(x)\right] \geq 2m\left[P^{2m}(x,x) - \pi(x)\right]$$

so dividing by $8m\pi(x)$ and using part i) we have

$$\frac{E_\pi(\tau_x)}{8m} \geq \| P^m(x,\cdot) - \pi \|_{TV}^2$$

and so for $m \geq 2E_\pi(\tau_x)$, $\quad \dfrac{E_\pi(\tau_x)}{8m} \leq \dfrac{1}{16}$

$\blacksquare$

[ex 10.20]

Lazy Random Walk on the Cycle:

so we've previously shown for a lazy random walk on $\mathbb{Z}_n$, $t_{mix} \leq n^2$, but now we can also use theorem 10.14.

So label the states of $\mathbb{Z}_n$ with $\{0,1,\dots,n-1\}$. Now by identifying the states 0 and n, we can see that $E_k(\tau_0)$ for the lazy simple random walk on the cycle must be the same as the expected time to ruin or success in a lazy gambler's ruin on the path $\{0,1,\dots,n\}$. Therefore for a lazy simple random walk on the cycle, by ex. 2.1 we have

$$t_{hit} = \max_{x,y} E_x(\tau_y) = \max_{0\leq k\leq n} 2k(n-k) = \left\lceil \frac{n^2}{2} \right\rceil$$

so now the theorem gives $\quad t_{mix} \leq n^2+1$