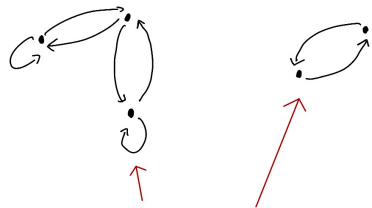
Stats 318: Lecture 3

Agenda:

- Irreducibility
- Uniqueness of Invariant Distributions
- Hitting Times and Times of First Return
- First Ergodic Theorem
- Implications of Ergodic Theorem

Invariant probabilities may not be unique

Suppose transition graph is disconnected



If we start here or here, we will stay there forever

At least one invariant measure for $\left\langle \begin{array}{l} \text{left } \pi^L \\ \text{right } \pi^R \end{array} \right\rangle$ any convex combination is invariant

Irreducibility

Definition (Irreducible)

Chain with transition matrix P is irreducible if for all $x, y \in \mathcal{X}$, $\exists\ t$

$$P(X_t = y \mid X_0 = x) > 0 \qquad [P^t(x, y) > 0]$$

- Means transition graph is connected
- For all $x, y \in \mathcal{X}$, there is a positive prob. of ever reaching y when starting from x
- Equivalently, from each state, there is a sequence of arrows leading to any other state

Theorem (Perron Frobenius)

(i) *If P is irreducible, there is a unique invariant dist.*

(ii) $\pi(x) > 0 \qquad \forall\ x \in \mathcal{X}$

Proof

Positivity: Pick x s.t. $\pi(x) > 0$

Take $y \in \mathcal{X}$: $\pi(y) = (\pi P^t)(y) \geq \pi(x) P^t(x, y) > 0$

Unicity: Consider

$$\mathcal{E}(u) = \sum_{x,y} (u(y) - u(x))^2 \pi(x) P(x, y)$$

Because $\pi > 0$ and P irreducible, any u s.t. $\mathcal{E}(u) = 0$ is constant

$$\mathcal{E}(u) = 0 \qquad \Longrightarrow \qquad (u(y) - u(x))^2 P(x, y) \qquad \forall\ x, y \in \mathcal{X}$$

Fix x, y. There is a path $x_0 = x, x_1, \ldots, x_t = y$ and $P(x_s, x_{s+1}) > 0\ \forall\ s$

$$\therefore \qquad u(x_s) = u(x_{s-1})\ \forall\ s \Rightarrow u(y) = u(x)$$

We show that $Pu = u \quad \Longrightarrow \quad \mathcal{E}(u) = 0$

$$\begin{aligned}\mathcal{E}(u) &= \sum_{x,y}\{(u^2(y)\pi(x)P(x,y) - 2u(y)u(x)\pi(x)P(x,y) + u^2(x)\pi(x)P(x,y)\} \\ &= \sum_y u^2(y)\pi(y) - 2\sum_x u^2(x)\pi(x) + \sum_x u^2(x)\pi(x) = 0\end{aligned}$$

Consider another invariant measure μ and set

$$\begin{aligned}Q(x,y) &= P(y,x)\pi(y)/\pi(x) \\ u(x) &= \mu(x)/\pi(x)\end{aligned}$$

Q irreducible and

$$\sum_y Q(x,y)u(y) = \sum_y P(y,x)\frac{\mu(y)}{\pi(x)} = \frac{\mu(x)}{\pi(x)} = u(x)$$

Hence $Qu = u \quad \Longrightarrow \quad \mu = \pi$

Hitting Times & Return Times

Return Time

$$\tau_x = \min(t \geq 1 : X_t = x) \qquad [X_0 = x]$$

with the convention that $\tau_x = +\infty$ if chain never visits x again

Hitting Time

$$\tau_{x,y} = \min(t \geq 1 : X_t = y) \qquad [X_0 = x]$$

Proposition

For any $x, y \in \mathcal{X}$ of irreducible chain, $\mathbb{E}_x(\tau_y) < \infty$

- $\exists\ r > 0$ s.t. for any states $(z, w)\ \exists\ s \leq r \qquad P^s(z, w) > \epsilon > 0$
- For any state x_t, prob. of hitting y between t & $t + r$ is at least ϵ

$$\mathbb{P}_x(\tau_y > kr) \leq (1 - \epsilon)\mathbb{P}_x(\tau_y > (k-1)r)$$
$$\therefore \qquad \mathbb{P}_x(\tau_y > kr) \leq (1 - \epsilon)^k$$

- Y nonnegative RV:

$$\mathbb{E}Y = \sum_{t \geq 0} \mathbb{P}(Y \geq t)$$

$$\therefore \qquad \mathbb{E}_x(\tau_y) = \sum_{t \geq 0} \mathbb{P}_x(\tau_y > t) \leq \sum_{k \geq 0} r\mathbb{P}_x(\tau_y > kr) \leq r \sum_{k \geq 0} (1 - \epsilon)^k < \infty$$

Waiting Times

Markov Chain with initial state x

$$\tau_x^{(0)} = 0 \qquad \tau_x^{(1)} = \inf\{t > 0 : X_t = x\} \qquad \tau_x^{(k+1)} = \inf\{t > \tau_x^{(k)} : X_t = x\}$$

$W^{(k)} = \tau^{(k)} - \tau^{(k-1)}$: waiting time between consecutive visits

Proposition

Waiting times are iid

Note: by definition, $W^{(1)} \sim \tau^{(1)}$

Proof

$$\begin{aligned}\mathbb{P}_x(\tau^{(1)} = t_1, \tau^{(2)} = t_2) &= \mathbb{P}_x(\tau^{(1)} = t_1)\mathbb{P}_x(\tau^{(2)} = t_2 \mid \tau^{(1)} = t_1) \\ &= \mathbb{P}_x(W^{(1)} = t_1)\mathbb{P}_x(W^{(2)} = t_2 - t_1) \qquad \text{why?}\end{aligned}$$

Same argument with $W^{(1)}$, $W^{(2)}$, ..., $W^{(k)}$

Ergodic Theorem

Theorem

P irreducible with invariant distribution π.

$$\forall\, x \in \mathcal{X} \qquad \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} 1\{X_t = x\} = \pi(x) = \frac{1}{\mathbb{E}\tau(x)}$$

- Fraction of time chain is in state x is $\pi(x)$
- We have seen $\pi(x) > 0$ & $\mathbb{E}\tau(x) < \infty$
- Initial state of the chain does not matter
- Generalization of law of large numbers for dependent sequences

$$\text{(LLN)} \quad X_1, X_2, \ldots \text{ iid } \pi \quad \Longrightarrow \quad \lim_{T\to\infty} \frac{1}{T} \sum_{t} 1\{X_t = x\} = \pi(x) \text{ a.s.}$$

Recall SLLN

I_t iid RVs & $I_t \geq 0$

$$\lim_{T \to \infty} \frac{I_1 + I_2 + \ldots + I_T}{T} = \mathbb{E} I_1 \text{ a.s.} \quad \text{(always, even if } \mathbb{E} I_1 = \infty)$$

Proof of Ergodic Theorem

- $\tau_x^{(n)} = W^{(1)} + \ldots + W^{(n)}$ & $W^{(k)} \overset{i.i.d.}{\sim} \tau_x$
- By SLLN

$$\lim_{T\to\infty} \frac{\tau_x^{(T)}}{T} = \lim_{n\to\infty} \frac{W^{(1)} + \ldots + W^{(T)}}{T} = \mathbb{E}\ \tau_x$$

- For each $T \geq 1$, $\exists\ m$ s.t.

$$\tau_x^{(m)} \leq T < \tau_x^{(m+1)}$$

$$\implies \quad \frac{1}{T}\sum_{t=1}^{T} 1\{X_t = x\} = \frac{m}{T} \in\ ]\frac{m}{\tau_x^{(m+1)}}, \frac{m}{\tau_x^{(m)}}\ ]$$

As $T\ \uparrow\ \infty,\ m\ \uparrow\ \infty$ and $\therefore$ $\quad \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} 1\{X_t = x\} = \frac{1}{\mathbb{E}\tau_x}$

Concludes proof when $X_0 = x$

- If $X_0 \neq x$, $\{X_t\}$ visits x after time $\tau_{x_0}(x) = \tau$
- $\tau < \infty$ a.s.
- Apply previous argument to $\{X_{t+\tau}\}_{t \geq 0}$, which is a chain with initial state x_0 & trans. matrix P

Conclusion:

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} 1\{X_t = x\} = \mathbb{E} \frac{1}{\tau_x}$$

Remains to prove that $\pi(x) = \frac{1}{\mathbb{E}\tau_x}$

$$Y_T = \frac{1}{T}\sum_{t=1}^{T} 1\{X_t = x\} \xrightarrow[T\to+\infty]{} \frac{1}{\mathbb{E}\tau_x}$$

By DOM $\qquad \mathbb{E}[Y_T] = \frac{1}{T}\sum_{t=1}^{T} \mathbb{P}(X_t = x) \xrightarrow[T\to+\infty]{} \frac{1}{\mathbb{E}\tau_x}$

In other words, if $X_0 \sim \mu$

$$\lim_{T\to\infty} \frac{1}{T}(\mu P^t)(x) = \frac{1}{\mathbb{E}\tau_x}$$

$$\mu = \pi \qquad \Longrightarrow \qquad \pi P^t = \pi \qquad \Longrightarrow \qquad \pi(x) = \frac{1}{\mathbb{E}\tau_x}$$

Implication of Ergodic Theorem

Ergodic theorem asserts that one can run a chain & approximate theoretical means by sample averages

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} f(X_t) = \sum_{x} \pi(x) f(x) = \mathbb{E}_\pi f(X)$$

- Generalization of LLN for independent sampling
- Important applications in MCMC methods

Example: Ehrenfest Model

- State where all particles are on one side will be visited infinitely often
- A little paradoxical (at least, seems so) when n large because we expect about same # on each side

"Solution": Suppose n even & recall $\pi(x) = 2^{-n}\binom{n}{x}$

- Mean return to $n/2$ is

$$\frac{1}{\pi(n/2)} = \frac{2^n}{\binom{n}{n/2}} = \sqrt{\frac{\pi n}{2}}$$

- Mean return to n is

$$\frac{1}{\pi(n)} = 2^n \qquad \text{(very large)}$$

Can show that for large c, chain has almost no chance of ever leaving [$\frac{n-c\sqrt{n}}{2}, \frac{n+c\sqrt{n}}{2}$] in reasonable time.