Review of lecture 1

Lecture 2, Wednesday February 21

Three perspectives on Markov chains

I. A sequence of random variables $(X_0, X_1, \ldots)$ with associated stochastic matrix $P \in \mathbb{M}_{|\mathcal{X}|^2}$ such that

$$\begin{aligned} &\mathbb{P}(X_{t+1} = y \mid X_t = x, X_{t-1} = x_{t-1}, \ldots, X_0 = x_0) \\ =&\mathbb{P}(X_{t+1} = y \mid X_t = x) \\ =&P(x, y) \end{aligned}$$

Three perspectives on Markov chains

I. A sequence of random variables $(X_0, X_1, \ldots)$ with associated stochastic matrix $P \in \mathbb{M}_{|\mathcal{X}|^2}$ such that

$$\begin{aligned} &\mathbb{P}(X_{t+1} = y \mid X_t = x, X_{t-1} = x_{t-1}, \ldots, X_0 = x_0) \\ =&\mathbb{P}(X_{t+1} = y \mid X_t = x) \\ =&P(x, y) \end{aligned}$$

II. A measurable map $\mathcal{X} \to \mathcal{P}(\mathcal{X})$

Three perspectives on Markov chains

I. A sequence of random variables $(X_0, X_1, \ldots)$ with associated stochastic matrix $P \in \mathbb{M}_{|\mathcal{X}|^2}$ such that

$$\begin{aligned} &\mathbb{P}(X_{t+1} = y \mid X_t = x, X_{t-1} = x_{t-1}, \ldots, X_0 = x_0) \\ =&\mathbb{P}(X_{t+1} = y \mid X_t = x) \\ =&P(x, y) \end{aligned}$$

II. A measurable map $\mathcal{X} \to \mathcal{P}(\mathcal{X})$

III. A map

$$\begin{aligned} \mathcal{P}(\mathcal{X}) \times \mathbb{M}_{|\mathcal{X}|^2} &\to \mathcal{P}(\mathcal{X})^\infty \\ (\mu_0, P) &\mapsto (\mu_0, \mu_0 P, \mu_0 P^2, \ldots) = (\mu_0, \mu_1, \mu_2, \ldots), \end{aligned}$$

with μ_k the distribution of X_k in $\mathcal{X}$

Recall

- §1.2: Given P, we can always find a *random mapping representation*, namely a function f and random variable Z satisfying

$$P(x,y) = \mathbb{P}(f(x,Z) = y)$$

E.g.: walk on $\mathbb{Z}_n$, $P(j,k) = \mathbb{P}(j + Z = k \bmod n)$, $Z \sim \text{Unif}\{-1,1\}$.

Recall

- §1.2: Given P, we can always find a *random mapping representation*, namely a function f and random variable Z satisfying

$$P(x, y) = \mathbb{P}(f(x, Z) = y)$$

  E.g.: walk on $\mathbb{Z}_n$, $P(j, k) = \mathbb{P}(j + Z = k \bmod n)$, $Z \sim \text{Unif}\{-1, 1\}$.

- §1.3: We will mostly care about *irreducible* ("connected") and *aperiodic* Markov chains. If P is periodic, can replace it with the lazy chain $\frac{1}{2}(I + P)$.

Recall

- §1.5: a row vector $\pi \in \mathcal{P}(\mathcal{X})$ is a *stationary distribution* if $\pi P = \pi$.
  - If $\lim_{t\to\infty} \mu_t = \nu$ exists, ν is necessarily stationary
  - If $\mu_0 = \pi$, then $\mu_k = \pi$ for all k

Recall

- §1.5: a row vector $\pi \in \mathcal{P}(\mathcal{X})$ is a *stationary distribution* if $\pi P = \pi$.
    - If $\lim_{t\to\infty} \mu_t = \nu$ exists, ν is necessarily stationary
    - If $\mu_0 = \pi$, then $\mu_k = \pi$ for all k
- For the simple random walk on (the vertices) of a graph, for $x \in V$,

$$\pi(x) = \frac{\deg(x)}{2|E|}$$

Recall

- §1.5: a row vector $\pi \in \mathcal{P}(\mathcal{X})$ is a *stationary distribution* if $\pi P = \pi$.
  - If $\lim_{t\to\infty} \mu_t = \nu$ exists, ν is necessarily stationary
  - If $\mu_0 = \pi$, then $\mu_k = \pi$ for all k
- For the simple random walk on (the vertices) of a graph, for $x \in V$,

$$\pi(x) = \frac{\deg(x)}{2|E|}$$

- Note: if $\pi P = \pi$,

$$\pi \frac{I + P}{2} = \pi,$$

  so "lazyfication" does not change π

Recall

- *Lemma 1.13.* For irreducible P (and $|\mathcal{X}| < \infty$),

$$\mathbb{E}_x(\tau_y^+) < \infty$$

for any $x, y \in \mathcal{X}$

Recall

- *Lemma 1.13.* For irreducible P (and $|\mathcal{X}| < \infty$),

$$\mathbb{E}_x(\tau_y^+) < \infty$$

  for any $x, y \in \mathcal{X}$
- Excursus: not true for MC's on infinite state spaces. Consider simple random walk on $\mathbb{Z}$ with $X_0 = 0$.

Recall

- *Lemma 1.13.* For irreducible P (and $|\mathcal{X}| < \infty$),

$$\mathbb{E}_x(\tau_y^+) < \infty$$

for any $x, y \in \mathcal{X}$

- Excursus: not true for MC's on infinite state spaces. Consider simple random walk on $\mathbb{Z}$ with $X_0 = 0$. Claim: $\mathbb{E}_0(\tau_1) = +\infty$.

Recall

- *Lemma 1.13.* For irreducible P (and $|\mathcal{X}| < \infty$),

$$\mathbb{E}_x(\tau_y^+) < \infty$$

  for any $x, y \in \mathcal{X}$
- Excursus: not true for MC's on infinite state spaces. Consider simple random walk on $\mathbb{Z}$ with $X_0 = 0$. Claim: $\mathbb{E}_0(\tau_1) = +\infty$.
  - Write $X_k = \sum_{j=1}^{k} Y_j$, where Y_j are iid Unif$\{-1, 1\}$ R.V.'s

Recall

- *Lemma 1.13.* For irreducible P (and $|\mathcal{X}| < \infty$),

$$\mathbb{E}_x(\tau_y^+) < \infty$$

  for any $x, y \in \mathcal{X}$
- Excursus: not true for MC's on infinite state spaces. Consider simple random walk on $\mathbb{Z}$ with $X_0 = 0$. Claim: $\mathbb{E}_0(\tau_1) = +\infty$.
  - Write $X_k = \sum_{j=1}^k Y_j$, where Y_j are iid Unif$\{-1, 1\}$ R.V.'s
  - If $\mathbb{E}_0(\tau_1) < \infty$, then by Wald's equation (see Durrett 5th edition §2.6),

$$1 = \mathbb{E}_0(X_{\tau_1}) = \mathbb{E}_0\left(\sum_{j=1}^{\tau_1} Y_j\right) = \mathbb{E}_0(\tau_1)\mathbb{E}(Y_1) = 0$$

Recall

- *Lemma 1.13.* For irreducible P (and $|\mathcal{X}| < \infty$),

$$\mathbb{E}_x(\tau_y^+) < \infty$$

  for any $x, y \in \mathcal{X}$
- Excursus: not true for MC's on infinite state spaces. Consider simple random walk on $\mathbb{Z}$ with $X_0 = 0$. Claim: $\mathbb{E}_0(\tau_1) = +\infty$.
  - Write $X_k = \sum_{j=1}^k Y_j$, where Y_j are iid Unif$\{-1, 1\}$ R.V.'s
  - If $\mathbb{E}_0(\tau_1) < \infty$, then by Wald's equation (see Durrett 5th edition §2.6),

$$1 = \mathbb{E}_0(X_{\tau_1}) = \mathbb{E}_0\left(\sum_{j=1}^{\tau_1} Y_j\right) = \mathbb{E}_0(\tau_1)\mathbb{E}(Y_1) = 0$$

  - In fact, $\mathbb{P}_0(\tau_1 > t) = O(t^{-1/2})$, whereas in the finite case, $\mathbb{P}_x(\tau_y^+ > t) = O(\delta^t)$ for some $\delta < 1$.

Recall

Fix $z \in \mathcal{X}$ and define

$$\tilde{\pi}(x) := \mathbb{E}_z(\text{visits to } x \text{ before } \tau_z^+) = \mathbb{E}_z\left(\sum_{t=0}^{\infty} \mathbf{1}\{X_t = x, \tau_z^+ > t\}\right)$$

Proposition 1.14. (i) If $\mathbb{P}_z(\tau_z^+ < \infty) = 1$, then $\tilde{\pi} = \tilde{\pi} P$. (ii) If $\mathbb{E}_z(\tau_z^+) < \infty$, then $\tilde{\pi}/\mathbb{E}_z(\tau_z^+)$ is a stationary distribution.

Recall

Fix $z \in \mathcal{X}$ and define

$$\tilde{\pi}(x) := \mathbb{E}_z(\text{visits to } x \text{ before } \tau_z^+) = \mathbb{E}_z\left(\sum_{t=0}^{\infty} \mathbf{1}\{X_t = x, \tau_z^+ > t\}\right)$$

Proposition 1.14. (i) If $\mathbb{P}_z(\tau_z^+ < \infty) = 1$, then $\tilde{\pi} = \tilde{\pi} P$. (ii) If $\mathbb{E}_z(\tau_z^+) < \infty$, then $\tilde{\pi}/\mathbb{E}_z(\tau_z^+)$ is a stationary distribution.

Proof: Already saw (i) (modulo a few details).

Recall

Fix $z \in \mathcal{X}$ and define

$$\tilde{\pi}(x) := \mathbb{E}_z(\text{visits to } x \text{ before } \tau_z^+) = \mathbb{E}_z\left(\sum_{t=0}^{\infty} \mathbf{1}\{X_t = x, \tau_z^+ > t\}\right)$$

Proposition 1.14. (i) If $\mathbb{P}_z(\tau_z^+ < \infty) = 1$, then $\tilde{\pi} = \tilde{\pi} P$. (ii) If $\mathbb{E}_z(\tau_z^+) < \infty$, then $\tilde{\pi}/\mathbb{E}_z(\tau_z^+)$ is a stationary distribution.

Proof: Already saw (i) (modulo a few details). For (ii), normalize to get a probability measure:

$$\sum_{x \in \mathcal{X}} \tilde{\pi}(x) = \mathbb{E}_z\left(\sum_{t=0}^{\infty} \sum_{x \in \mathcal{X}} \mathbf{1}\{X_t = x\}\mathbf{1}\{\tau_z^+ > t\}\right)$$

$$= \mathbb{E}_z\left(\sum_{t=0}^{\infty} \mathbf{1}\{\tau_z^+ > t\}\right) = \mathbb{E}_z(\tau_z^+) < \infty. \quad \square$$

Conclusions for §1.5.3

- At state z, the stationary distribution is therefore

$$\pi(z) = \frac{\tilde{\pi}(z)}{\mathbb{E}_z(\tau_z^+)} = \frac{1}{\mathbb{E}_z(\tau_z^+)}$$

  And since $z \in \mathcal{X}$ was arbitrary, $\pi(x) = 1/\mathbb{E}_x(\tau_x^+)$ for all x

Conclusions for §1.5.3

- At state z, the stationary distribution is therefore

$$\pi(z) = \frac{\tilde{\pi}(z)}{\mathbb{E}_z(\tau_z^+)} = \frac{1}{\mathbb{E}_z(\tau_z^+)}$$

  And since $z \in \mathcal{X}$ was arbitrary, $\pi(x) = 1/\mathbb{E}_x(\tau_x^+)$ for all x

- By Lemma 1.13, irreducible chains therefore always have a stationary distribution

Conclusions for §1.5.3

- At state z, the stationary distribution is therefore

$$\pi(z) = \frac{\tilde{\pi}(z)}{\mathbb{E}_z(\tau_z^+)} = \frac{1}{\mathbb{E}_z(\tau_z^+)}$$

  And since $z \in \mathcal{X}$ was arbitrary, $\pi(x) = 1/\mathbb{E}_x(\tau_x^+)$ for all x
- By Lemma 1.13, irreducible chains therefore always have a stationary distribution
- For simple random walk on a graph, $\pi(x) = \deg(x)/(2|E|)$, and thus

$$\mathbb{E}_x(\tau_x^+) = \frac{2|E|}{\deg(x)}$$

  Surprising!