Stats 318: Lecture # 8

Agenda:

- Spectrum of reversible chains
- Spectral gap
- Speed of convergence of reversible chains
- Examples

Preliminaries

- Finite state space $\mathcal{X}$
- Invariant dist. $\pi \quad \pi(x) > 0 \quad \forall\, x \in \mathcal{X}$
- Vector/function $\{f(x)\}_{x \in \mathcal{X}}$
- Inner product

$$\langle f, g \rangle_\pi \; = \sum_x f(x) g(x) \pi(x)$$

- Norm: $\|f\| = \sqrt{\langle f, f \rangle_\pi}$
- $\mathbb{E}_\pi f = \sum f(x)\pi(x) = \langle \mathbb{1}, f \rangle_\pi$

Spectrum of reversible chains

P self adjoint if for all f, g

$$\langle Pf, g\rangle_\pi \;=\; \langle f, Pg\rangle_\pi$$

Proposition

P self adjoint $\iff$ P reversible wrt π

e.g. π uniform $\quad \cancel{\pi(x)}P(x,y) = P(y,x)\cancel{\pi(y)} \quad \Longrightarrow \quad P$ symmetric

Proof:

$$\langle Pf, g\rangle = \sum_{x,y} \pi(x)P(x,y)f(y)g(x)$$

$$\langle f, Pg\rangle = \sum_{x,y} \pi(y)P(y,x)f(y)g(x)$$

P reversible: $\pi(x)P(x,y) = \pi(y)P(y,x) \Longrightarrow \langle Pf, g\rangle = \langle f, Pg\rangle$

Conversely, P self adjoint:

Apply $\langle Pf, g\rangle \;=\; \langle f, Pg\rangle$ with $f = \delta_y,\ g = \delta_x \quad \Longrightarrow \quad P$ reversible

Eigenvalues of reversible chains

Proposition (P reversible wrt π)

- *Eigenvalues of P are real: $|\lambda_i| \leq 1$ & 1 always an eigenvalue*

$$1 = \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq -1$$

- *P irreducible $\Longrightarrow$ 1 is a single eigenvalue: $\lambda_2 < 1$*
- *P irreducible & aperiodic $\Longrightarrow$ -1 is not an eigenvalue: $\lambda_n > -1$*

Proof:

- P self adjoint $\Longrightarrow$ diagonalizable in orthobasis with real eigenvalues
- $\mathbb{1}(x) = 1$ eigenvector with eigenvalue 1
- f eigenvector with eigenvalue λ

$$|\lambda| \max_x |f(x)| = \max_x |Pf|(x) \leq \max_x \sum_y P(x,y)|f(y)| \leq \max_x |f(x)|$$

- P irreducible $\implies$ only eigenvector with $\lambda = 1$ [π is unique]
- Suppose -1 is an eigenvalue & f s.t. $Pf = -f \implies P^{2t+1} f = -f$

Convergence theorem

$$\mathbb{E}_x f(X_t) \underset{t\to\infty}{\longrightarrow} \mathbb{E}_\pi f$$

$$\mathbb{E}_x f(X_{2t+1}) = (P^{2t+1} f)(x) = -f(x) \qquad \Longrightarrow\Longleftarrow$$

Spectral decomposition of the transition matrix

- P self adjoint
- Spectral decomposition

$$Pf = \sum_{i=1}^{n} \lambda_i \; v_i \langle v_i, f \rangle_\pi \qquad \langle v_i, v_j \rangle_\pi = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \qquad v_i = \mathbb{1}$$

$$= \mathbb{1}\mathbb{E}_\pi f + \sum_{i \geq 2} \lambda_i \; v_i \langle v_i, f \rangle_\pi$$

- Shows convergence of Markov chain [3rd proof]

$$P^t f - (\mathbb{E}_\pi f)\mathbb{1} \; = \sum_{i \geq 2} \lambda_i^{2t+1} v_i \langle v_i, f \rangle_\pi$$

Speed of convergence of reversible chains

Theorem (P irreducible & aperiodic)

(i) $|\mathbb{E}_x f(X_t) - \mathbb{E}_\pi f|^2 \leq c(x)\rho^{2t-2}\mathsf{Var}_\pi f$

(ii) $\|P^t(x,\cdot) - \pi\|_{\mathsf{TV}}^2 \leq \frac{1}{4}c(x)\rho^{2t-2}$

$c(x) = \frac{P^2(x,x)}{\pi(x)}$

$\rho = \mathsf{max}(\lambda_2, |\lambda_n|) < 1$

Proof

(i) Replacing f with $f - (\mathbb{E}_\pi f)\mathbb{1}$, can always assume $\mathbb{E}_\pi f = 0$

Note $\|f\|_\pi^2 \;=\; \mathsf{Var}_\pi f = \sum_x \pi(x) f^2(x)$

$$\|P^t f\|_\pi^2 \;=\; \sum_{i \geq 2} \lambda_i^{2t} \langle f, v_i \rangle_\pi^2 \leq \rho^{2t} \sum_{i \geq 2} \langle f, v_i \rangle_\pi^2 \;=\; \rho^{2t} \|f\|_\pi^2$$

Set $g = P^{t-1} f$: $\quad Pg(x) = \sum_y P(x,y) g(y) \;= \sum_y P(y,x) \frac{\pi(y)}{\pi(x)} g(y) = \langle q, g \rangle_\pi$

$$q(y) = \frac{P(y,x)}{\pi(x)}$$

$$\|q\|_\pi^2 = \sum_y \left[\frac{P(y,x)}{\pi(x)} \right]^2 \pi(y) = \sum_y \frac{\pi(x) P(x,y) P(y,x)}{\pi^2(x)} = \frac{1}{\pi(x)} P(x,x) = c(x)$$

$\|g\|_\pi^2 \leq \rho^{2t-2}\|f\|_\pi^2$

Cauchy-Schwarz $\implies$ $|Pg(x)|^2 \leq \|q\|_\pi^2 \ \|g\|_\pi^2 \leq c(x)\rho^{2t-2}\|f\|_\pi^2$

(ii) Apply (i) with $f = \mathbb{1}_A$ & $\mathsf{Var}_\pi \mathbb{1}_A = \pi(A)(1-\pi(A))$

Relaxation time

$$\rho = 1 - \gamma^* \qquad \& \qquad t_{\mathsf{relax}} = \tfrac{1}{\gamma^*}$$

$$d(t) = \max_x \|P^t(x,\cdot) - \pi\|_{\mathsf{TV}} \leq \frac{1}{2} \max \sqrt{\frac{P^2(x,x)}{\pi(x)}}\ \rho^{t-1}$$

$$\Longrightarrow \quad d(t) \leq \tfrac{1}{\sqrt{4\pi_{\mathsf{min}}}} e^{-\gamma^*(t-1)}$$

$$\therefore \qquad \gamma^*[t_{\mathsf{mix}}(\epsilon) - 1] \leq \log \frac{1}{\epsilon\sqrt{4\pi_{\mathsf{min}}}} \quad \Leftrightarrow \quad t_{\mathsf{mix}}(\epsilon) \leq 1 + \{\log[\quad]\}\ t_{\mathsf{relax}}$$

On the other hand, for a reversible & aperiodic chain [Theorem 12.4]

$$t_{\mathsf{mix}}(\epsilon) \geq (t_{\mathsf{relax}} - 1) \log \frac{1}{2\epsilon}$$

$$\Longrightarrow \qquad \lim_{t \to \infty} \frac{1}{t} \log d(t) = \log \rho$$

Spectral Gap

Definition

Spectral gap of P is smallest positive eigenvalue of $I - P$

$$\gamma(P) = 1 - \lambda_2$$

Lemma

Same assumptions $+$ $P(x,x) \geq \frac{1}{2}$ *for all* x

$\Longrightarrow$ *Same conclusion with* $1 - \gamma$ *instead of* ρ

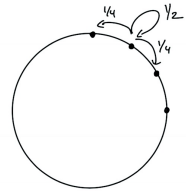
Proof: Diag elements $\geq \frac{1}{2} \Longrightarrow$ eigenvalues of $P \geq 0$ [Gershgorin's open circle thm]

$$\Longrightarrow \lambda_n = |\lambda_n| \leq \lambda_2$$

$$\Longrightarrow \rho = \lambda_2 = 1 - \gamma$$

Example: RW on n-cycle

- $$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & 0 & \cdots & 0 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & \cdots & 0 \\ & & & \cdots & & \end{bmatrix}$$



- P circulant matrix $\implies$ diag in fourier basis
  - Eigenvalues are DFT of first row

$$\lambda_k = \sum_{x=0}^{n-1} P(0,x) e^{-i2\pi kx/n} = \frac{1}{4} e^{i2\pi k/n} + \frac{1}{2} + \frac{1}{4} e^{-i2\pi k/n}$$

$$= \frac{1}{2}\left(1 + \cos\frac{2\pi k}{n}\right) \geq 0$$

- $\lambda_1 = 1$, $\lambda_2 = \frac{1}{2} + \frac{1}{2}\cos\frac{2\pi}{n}$, $\gamma = 1 - \lambda_2 = \frac{1-\cos 2\pi/n}{2} = \sin^2\frac{\pi}{n} = \frac{\pi^2}{n^2}$
- $t_{\mathsf{mix}} \geq \frac{n^2}{\pi^2}\log 2$
- Previous lecture: $t_{\mathsf{mix}} \leq n^2$

Example: RW on $\{1,1\}^n$

$$P(x,y) = \begin{cases} \frac{1}{2} & x = y \\ \frac{1}{2n} & x \sim y \end{cases} \qquad \text{Eigenfunctions: } f_J(x) = \prod_{j \in J} x_j \qquad J \subset \{1,\dots,n\}$$

$$(Pf_J)(x) = \frac{1}{2}\left[f_J(x) + \sum_{y \sim x} \frac{1}{n} f_J(y)\right] = \frac{1}{2}\left[2 - \frac{2|J|}{n}\right] f_J(x) = \left[1 - \frac{|J|}{n}\right] f_J(x)$$

$$y \sim x \;\&\; y_i = -x_i \quad : \quad f_J(y) = \begin{cases} f_J(x) & i \notin J \\ -f_J(x) & i \in J \end{cases}$$

$$\Longrightarrow \sum_{y \sim x} f_J(y) = (n - |J|) f_J(x) - |J| f_J(x) = (n - 2|J|) f_J(x)$$

$\therefore$ Eigenvalues $1 - k/n \qquad k = 0, 1, 2, \dots, n$

$\gamma = 1/n \quad \Leftrightarrow \quad t_{\mathsf{relax}} = n$

$\therefore \quad t_{\mathsf{mix}}(\epsilon) \geq (n-1)\log(1/2\epsilon)$

$$
\begin{aligned}
t_{\mathsf{mix}}(\epsilon) &\leq n \log \frac{1}{4\epsilon\sqrt{\pi_{\mathsf{min}}}} \qquad [\text{Because } P^2(x,x) = \frac{1}{4}] \\
&= n \log \frac{1}{4\epsilon 2^{n/2}} \\
&= \frac{n^2}{2}\log 2 + n \log \frac{1}{4\epsilon}
\end{aligned}
$$

Bound is not tight!

Improved Inequalities

For this chain

$$\|P^t(x,\cdot) - \pi\|_{\mathsf{TV}}^2 \le \frac{1}{4}\sum_{j\ge 2}\lambda_j^{2t} \qquad (*)$$

$$\begin{aligned}
\therefore \quad 4\|P^t(x,\cdot) - \pi\|_{\mathsf{TV}}^2 \le \sum_{k=1}^{n}\left(1-\frac{k}{n}\right)^{2t}\binom{n}{k} &\le \sum_{k=1}^{n} e^{-2tk/n}\binom{n}{k} \\
&= (1+e^{-2t/n})^n - 1 \\
t = \frac{1}{2}n\log n + cn \Longrightarrow 4\|P^t(x,\cdot) - \pi\|_{\mathsf{TV}}^2 &\le \left[1+\frac{1}{n}e^{-2c}\right]^n - 1 \\
&\le e^{e^{-2c}} - 1 \\
&\le 2e^{-2c} \quad \text{for } c > 1
\end{aligned}$$

Sharp dropoff to zero!

(∗)

$$P^t f - (\mathbb{E}_\pi f)\mathbb{1} = \sum_{J:\lambda_J \neq 1} \lambda_J^t \langle f, f_J \rangle f_J$$

- $f = \mathbb{1}_A$
- $|f_J(x)| = 1$

$$\begin{aligned} |P^t(x, A) - \pi(A)|^2 &\leq \sum_{J:\lambda_J \neq 1} \lambda_J^{2t} |f_J(x)|^2 \sum_J \langle f, f_J \rangle_\pi^2 \\ &\leq \sum \lambda_J^{2t} \ \mathsf{Var}_\pi f = \frac{1}{4} \lambda_J^{2t} \end{aligned}$$