Lecture 10

Bounds to Mixing via Multicommodity Flows.

Input: Ergodic P on state space Ω, stationary distn' π

Flow Problem:

- (directed) edge $e = (z, z')$ **capacity** $C(e) = \underbrace{\pi(z) \cdot P(z,z')}$

  ↳ also called the ergodic flow from z to z'

- **Demands:** for $(x,y) \in \Omega^2$:

$$D(x,y) = \pi(x) \cdot \pi(y)$$

different commodity for every pair (x,y)

- **Flow:** $f : \mathcal{P} \to \mathbb{R}^+ \cup \{0\}$ where $\mathcal{P} = \bigcup_{xy} \mathcal{P}_{xy}$

  ↳ simple paths from x to y

s.t. $\sum_{p \in \mathcal{P}_{xy}} f(p) = D(x,y)$

- **Cost of flow:** $\rho(f) = \max_e \frac{f(e)}{C(e)}$ ← $\sum_{p \ni e} f(p)$

- **length of flow:** $\ell(f) = \max_{p : f(p) > 0} |p|$. (length of longest flow carrying path)

Theorem: For any lazy, ergodic MC $P = \frac{1}{2} I + \frac{1}{2} \hat{P}$ ($\hat{P}$ is ergodic) and any flow f we have:

$$\tau_{mix} \leq O\left(\rho(f) \cdot \ell(f) \cdot \ln \pi_{min}^{-1}\right),$$

where $\pi_{min} = \min_{x \in \Omega} \pi(x)$.

Remark 1: Usually above bound dominated by $\rho(f)$, since $\ell(f)$, $\ln \pi_{min}^{-1}$ are typically polynomial in the problem size.

Remark 2: We'll show a stronger result, replacing π_{min} w/ $\pi(x)$ where x is starting state.

Remark 3: Flows provide upper bound of mixing. Their dual problem of sparsest cut provides a lower bound. Any SC provides a lower bound.

The Poincaré Constant

Def: Let $\varphi : \Omega \to \mathbb{R}$ be any function. Then

$$\mathbb{E}_\pi \varphi = \sum_x \pi(x)\varphi(x)$$

$$\begin{aligned}\mathrm{Var}_\pi \varphi &= \sum_x \pi(x)\left(\varphi(x) - \mathbb{E}_\pi\varphi\right)^2 = \sum_x \pi(x)\varphi(x)^2 - \left(\mathbb{E}_\pi\varphi\right)^2 = \\ &= \sum_x \pi(x)\varphi(x)^2 \sum_y \pi(y) - \sum_x \pi(x)\varphi(x)\cdot\sum_y \pi(y)\varphi(y) = \\ &= \sum_x \pi(x)\pi(y)\left(\varphi(x)^2 - \varphi(x)\varphi(y)\right) = \\ &= \frac{1}{2}\sum_{xy} \pi(x)\pi(y)\left(\varphi(x) - \varphi(y)\right)^2.\end{aligned}$$

Def: Local Variance:

$$\mathcal{E}_\pi(\varphi,\varphi) = \frac{1}{2}\sum_{xy} \pi(x)P(x,y)\left(\varphi(x) - \varphi(y)\right)^2$$

↳ also called the Dirichlet form (because $\varphi(x)-\varphi(y)$ is the discrete derivative of φ)

Def: The Poincaré constant is

$$\alpha := \inf_{\varphi \text{ non-constant}} \frac{\mathcal{E}_\pi(\varphi,\varphi)}{\mathrm{Var}_\pi\varphi}$$

Theorem 1: For any lazy ergodic P and any initial state $x \in \Omega$:

$$\tau_x(\varepsilon) \le \frac{1}{\alpha}\left(2\ln\varepsilon^{-1} + \ln\pi(x)^{-1}\right).$$

Theorem 2: For any ergodic P and any flow f for P:

[Sinclair 92; Diaconis, Stroock '91]

$$\alpha \ge \frac{1}{\rho(f)\cdot\ell(f)}.$$

Thm 1 + Thm 2 $\Rightarrow$ main theorem (choose $\varepsilon = \frac{1}{2e}$)

Proof of Thm 1: (following Jerrum '03, Mihail '89)

- $P = \frac{1}{2}(I + \hat{P})$ (P & $\hat{P}$ have same stationary π).

- $[P\varphi](x) = \sum_y P(x,y)\,\varphi(y)$

  $[P^t\varphi](x) = \sum_y P^t(x,y)\,\varphi(y) \xrightarrow{t\to\infty} \sum_y \pi(y)\,\varphi(y) \equiv \mathcal{E}_\pi \varphi$

  hence $\mathrm{Var}_\pi P^t\varphi \xrightarrow{t\to\infty} 0$

- We are interested in the rate of convergence of $\mathrm{Var}_\pi P^t\varphi$

- **Main Lemma:** For any $\varphi : \Omega \to \mathbb{R}$

  proof postponed

  $$\mathrm{Var}_\pi P\varphi \leq \mathrm{Var}_\pi \varphi - \mathcal{E}_\pi(\varphi,\varphi) \leq \mathrm{Var}_\pi \varphi\,(1-\alpha).$$

- **Corollary:** $\mathrm{Var}_\pi P^t\varphi \leq (1-\alpha)^t \cdot \mathrm{Var}_\pi \varphi$.

- for any $A \subseteq \Omega$ define $\varphi_A(x) = \begin{cases} 1, & x \in A \\ 0, & \text{ow.} \end{cases}$

  - clearly $\mathrm{Var}_\pi \varphi_A \leq 1$, hence $\mathrm{Var}_\pi P^t\varphi_A \leq (1-\alpha)^t \leq e^{-\alpha t}$

    $$\leq \varepsilon^2 \cdot \pi(x) \qquad (*)$$

    setting $t = \frac{1}{\alpha}\left(\ln \pi(x)^{-1} + 2\ln \varepsilon^{-1}\right)$

  - on the other hand:

    $$\mathrm{Var}_\pi P^t\varphi_A \geq \pi(x)\Big(\underbrace{[P^t\varphi_A](x)}_{\text{III}①} - \underbrace{\mathbb{E}_\pi P^t\varphi_A}_{\text{III}}\Big)^2$$

    $$= \pi(x)\left(P_x^t(A) - \mathbb{E}_\pi \varphi_A\right)^2$$

    $$= \pi(x)\left(P_x^t(A) - \pi(A)\right)^2 \qquad (**)$$

$$\boxed{① \; [P^t\varphi_A](x) = \sum_y P^t(x,y)\,\varphi_A(y) = \sum_{y\in A} P^t(x,y) = P_x^t(A)}$$

$(*) + (**) \Rightarrow$ for $t = \frac{1}{\alpha}\left(\ln \pi(x)^{-1} + 2\ln \varepsilon^{-1}\right)$

$$|P_x^t(A) - \pi(A)| \leq \varepsilon$$

But A is arbitrary event $\Rightarrow \|P_x^t - \pi\|_{TV} \leq \varepsilon$, for t chosen as above.

Remains to show main lemma:

- $$[P\varphi](x) = \sum_y P(x,y)\varphi(y) =$$
$$= \frac{1}{2}\varphi(x) + \frac{1}{2}\sum_y \hat{P}(x,y)\varphi(y) =$$
$$= \frac{1}{2}\sum_y \hat{P}(x,y)\left(\varphi(x)+\varphi(y)\right).$$

- w.l.o.g. assume $E_\pi \varphi = 0$ (shifting by a constant won't change any of the variances computed below)

- $$\mathrm{Var}_\pi P\varphi = \sum_x \pi(x)\cdot\left([P\varphi](x)\right)^2$$
$$= \frac{1}{4}\sum_x \pi(x)\left(\sum_y \hat{P}(x,y)\left(\varphi(x)+\varphi(y)\right)\right)^2$$
$$\leq \frac{1}{4}\sum_{xy} \pi(x)\hat{P}(x,y)\left(\varphi(x)+\varphi(y)\right)^2 \qquad \text{(by Cauchy-Schwartz)}$$

- $$\mathrm{Var}_\pi \varphi = \frac{1}{2}\sum_x \pi(x)\varphi(x)^2 + \frac{1}{2}\sum_y \pi(y)\varphi(y)^2 \qquad \text{(using that } E_\pi\varphi = 0\text{)}$$
$$= \frac{1}{2}\sum_{xy} \pi(x)\varphi(x)^2\hat{P}(x,y) + \frac{1}{2}\sum_{xy}\pi(x)\hat{P}(x,y)\varphi(y)^2$$
$$= \frac{1}{2}\sum_{xy}\pi(x)\hat{P}(x,y)\left(\varphi(x)^2+\varphi(y)^2\right).$$

- Hence: $$\mathrm{Var}_\pi \varphi - \mathrm{Var}_\pi P\varphi \geq \frac{1}{4}\sum_{xy}\pi(x)\hat{P}(x,y)\left(\varphi(x)-\varphi(y)\right)^2$$
$$= \frac{1}{2}\sum_{xy}\pi(x)P(x,y)\left(\varphi(x)-\varphi(y)\right)^2 = \mathcal{E}_\pi(\varphi,\varphi).$$

$\boxtimes$

(end of proof of Thm 1)

Proof of Thm 2:

$$
\begin{aligned}
2\,\mathrm{Var}_\pi \varphi &= \sum_{x,y} \pi(x)\pi(y)\big(\varphi(x)-\varphi(y)\big)^2 \\
&= \sum_{xy} \sum_{p\in P_{xy}} f(p)\big(\varphi(x)-\varphi(y)\big)^2 \\
&= \sum_{xy} \sum_{p\in P_{xy}} f(p)\Big(\sum_{(u,v)\in p}\big(\varphi(v)-\varphi(u)\big)\Big)^2 && \text{(massaging formula to make it up from local variances)} \\
&\le \sum_{xy} \sum_{p\in P_{xy}} f(p)\,|p|\cdot\sum_{(u,v)\in p}\big(\varphi(v)-\varphi(u)\big)^2 && \text{(Cauchy-Schwartz)} \\
&= \sum_{e=(u,v)} \big(\varphi(v)-\varphi(u)\big)^2 \sum_{p\ni(u,v)} f(p)\cdot|p| \\
&\le \ell(f)\sum_{e=(u,v)}\big(\varphi(v)-\varphi(u)\big)^2 \underbrace{\sum_{p\ni e} f(p)}_{=f(e)\le \rho(f)\cdot C(e)} \\
&\le \ell(f)\sum_{e=(u,v)}\big(\varphi(v)-\varphi(u)\big)^2 C(e)\cdot\rho(f) \\
&= \ell(f)\cdot\rho(f)\cdot\sum_{e=(u,v)}\big(\varphi(v)-\varphi(u)\big)^2\cdot\pi(u)P(u,v) \\
&= \ell(f)\cdot\rho(f)\cdot 2\cdot\mathcal{E}_\pi(\varphi,\varphi).
\end{aligned}
$$

$$
\underset{\varphi\text{ non-constant}}{\Longrightarrow} \qquad \frac{\mathcal{E}_\pi(\varphi,\varphi)}{\mathrm{Var}_\pi\varphi} \geq \frac{1}{\ell(f)\,\rho(f)} \qquad \boxtimes
$$

Some Further Intuition: Reversible Markov Chains and the Poincaré Constant

- **Known fact:** If P is reversible and ergodic then its eigenvalues are:

$$1 = \lambda_1 > \lambda_2 > \dots > \lambda_N > -1$$

and its spectral gap $1-\lambda_2$ satisfies:

$$1-\lambda_2 \equiv \inf_{\varphi \text{ non-constant}} \frac{\mathcal{E}_\pi(\varphi,\varphi)}{Var_\pi \varphi}.$$

[we: the first part in lecture 2; the second part follows from the variational characterization of eigenvalues of symmetric matrices, using the symmetrization operator we constructed in lec 2 for reversible chains]

- If P is also lazy, then $\lambda_N > 0$ and hence $\max_{i \geq 2} |1-\lambda_i| \equiv 1-\lambda_2$.

- Now suppose an arbitrary starting state $x \in \Omega$, and suppose

$$P_x^{(0)} = \sum_{i=1}^{N} \alpha_i e_i$$

, where e_i is the eigenvector corresponding to eigenvalue λ_i

then
$$P_x^t = P_x^0 \cdot P^t = \alpha_1 e_1 + \underbrace{\sum_{i \geq 2} \alpha_i \lambda_i^t e_i}_{\downarrow\, t\to\infty \;\; 0} \equiv \pi + \sum_{i \geq 2} \alpha_i \lambda_i^t e_i.$$

(by the fundamental Thm of Markov Chains

Hence:

$$\Delta_x(t) \leq \lambda_2^t \cdot \Big(\sum_{i \geq 2} \sum_{y} |\alpha_i| \cdot |e_i(y)| \Big)$$

$$\leq (1-(1-\lambda_2))^t \cdot (\text{-//-})$$

$$\leq e^{-(1-\lambda_2)\cdot t} \cdot (\text{-//-}).$$

Therefore, the rate of convergence is dictated by $1-\lambda_2 \equiv$ Poincaré constant