Duality, Evolving Sets and Mixing

Saurabh Amin   Ambuj Tewari

November 30, 2006 / CS294-2 Paper Presentation

Outline

Conductance

- For $A, B \subset \Omega$, ergodic flow from A to B: $Q(A, B) := \sum_{x \in A} \sum_{y \in B} Q(x, y)$ with $Q(x, y) = \pi(x) P(x, y)$.
- Conductance

$$\Phi(S) := \frac{Q(S, \bar{S})}{\pi(S)} \text{ and } \Phi_* := \min_{S \subset \Omega} \{\Phi(S) : 0 < \pi(S) \leq 1/2\}$$

Theorem (From Lecture)

$$\tau_{\text{mix}} \geq \frac{1}{4\Phi^*} \qquad \text{for any Markov chain}$$

$$\tau_x(\epsilon) \leq \frac{2}{\Phi_*^2}\left(\log \frac{1}{\pi(x)} + \log \frac{1}{\epsilon}\right) \qquad \text{for lazy, reversible chains}$$

Main Result [Morris and Peres '03]

Define *conductance profile* as (where $\pi^* = \min_x \pi(x)$)

$$\Phi(r) := \begin{cases} \min_{S\in\Omega}\{\Phi(S) : 0 < \pi(S) \le r\} & r \in [\pi_*, 1/2] \\ \Phi_* & r > 1/2 \end{cases}$$

and $\epsilon-$uniform mixing time as

$$\tau^\infty(\epsilon) := \min\{n \ : \ \forall x, y \in \Omega, \left|\frac{P^n(x,y)}{\pi(y)} - 1\right| \le \epsilon\}$$

Theorem (MP03)

For lazy (not necessarily reversible) MC,

$$\tau^\infty(\epsilon) \le \int_{4\pi_*}^{4/\epsilon} \frac{4du}{u\Phi^2(u)}$$

Since $\Phi(u) \ge \Phi^*$, this is no worse than previous bound.

Profile bound can be better

- For the $n \times n$ 2-D grid, conductance profile satisfies

$$\Phi(u) \geq \frac{a}{n\sqrt{u}} .$$

- Profile bound is bounded by

$$C \log \frac{1}{\epsilon} + 4 \int_{1/n^2}^{1/2} \frac{du}{u \left(\frac{a}{n\sqrt{u}}\right)^2} = C \log \frac{1}{\epsilon} + O(n^2) .$$

- Previous bound would have given

$$C' n^2 \left(2 \log n + \log(1/\epsilon)\right) .$$

Various Mixing Times

- More generally, for $1 \leq p \leq \infty$, define

$$\tau_x^p(\epsilon) := \min\{n \ : \ \left\| \frac{P^n(x,\cdot)}{\pi(\cdot)} - \mathbf{1} \right\|_{L_p(\pi)} \leq \epsilon\}$$

- Note that $\tau_x^1(\epsilon) \leq \tau_x^2(\epsilon) \leq \tau_x^\infty(\epsilon)$.
- Define $\tau^p(\epsilon) = \max_{x \in \Omega} \tau_x^p(\epsilon)$.

The Time Reversed Chain

Given P time-reversed MC has transition matrix

$$P_R(x,y) := \frac{\pi(y)P(y,x)}{\pi(x)} .$$

It satisfies:

- $\pi(x)P_R^n(x,y) = \pi(y)P^n(y,x)$
- $\left|\frac{P^{(n+m)}(x,y)-\pi(y)}{\pi(y)}\right| \leq \left\|\frac{P^n(x,\cdot)}{\pi(\cdot)} - \mathbf{1}\right\|_{L_2(\pi)} \left\|\frac{P_R^m(y,\cdot)}{\pi(\cdot)} - \mathbf{1}\right\|_{L_2(\pi)}$

Corollary

$\tau^\infty(\epsilon) \leq \tau^2(\sqrt{\epsilon}) + \tau_R^2(\sqrt{\epsilon})$

Reversed chain has same profile

- Note that "conservation of mass" tells us

$$Q(S, \bar{S}) = Q(\bar{S}, S) \; .$$

- Since P_R also has stationary dist. π, we have

$$\Phi_R(S) := \frac{Q_R(S, \bar{S})}{\pi(S)} = \frac{Q(\bar{S}, S)}{\pi(S)} = \frac{Q(S, \bar{S})}{\pi(S)} =: \Phi(S)$$

- Therefore, it suffices to prove

$$\tau^2(\epsilon) \leq \int_{4\pi_*}^{4/\epsilon^2} \frac{2du}{u\Phi^2(u)} \; .$$

Notion of a Dual Chain

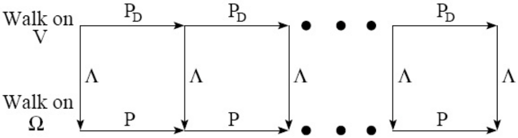
- If we have another MC with transition matrix P_D on some state space V (we consider $V = 2^{\Omega}$) such that

$$\underset{|V|\times|\Omega|}{\Lambda} \times \underset{|\Omega|\times|\Omega|}{P} = \underset{|V|\times|V|}{P_D} \times \underset{|V|\times|\Omega|}{\Lambda}$$

for some "link" matrix Λ.

- Then the evolution of the two chains will be related:

$$\forall n > 0, \ \Lambda P^n = P_D^n \Lambda \ .$$



Evolving Set

Definition

Evolving set process is a MC $\{S_n\}$ defined on 2^Ω as follows. If $S_n = A \subseteq \Omega$, pick $u \sim \mathrm{Unif}[0,1]$, and set $S_{n+1} = A_u$ where

$$A_u := \{y \in \Omega \ : \ Q(A,y) \geq u\pi(y)\} \ .$$

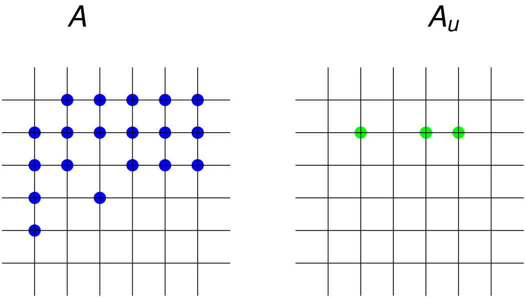
- Let $K(A,A') := \Pr(S_{n+1} = A' | S_n = A)$ be its transition matrix
- Simple but important observation:

$$\Pr(y \in S_{n+1} | S_n = A) = \Pr(Q(A,y) \geq u\pi(y)) = \frac{Q(A,y)}{\pi(y)} \ .$$

- $\emptyset$ and Ω are absorbing states of the evolving set process.
- Important martingale property:

$$\mathbb{E}\left[\pi(S_{n+1}) | S_n = A\right] = \sum_{y \in \Omega} \pi(y) \Pr(y \in S_{n+1} | S_n = A) = \sum_y Q(A,y) = \pi(A) \ .$$

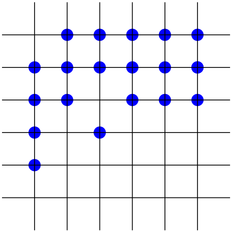
Example



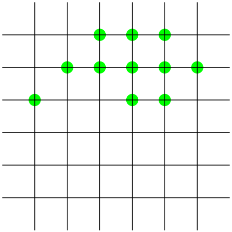
$u \in (\frac{7}{8}, 1]$

Example

A

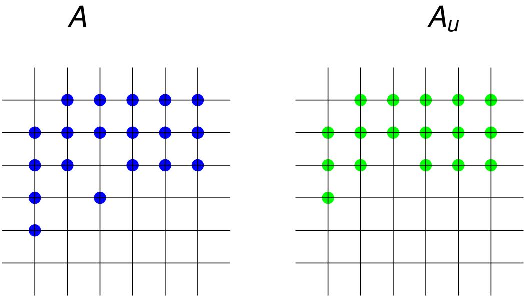


A_u



$u \in (\frac{6}{8}, \frac{7}{8}]$

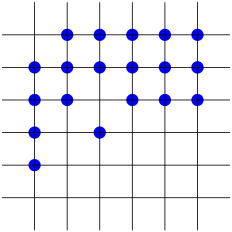
Example



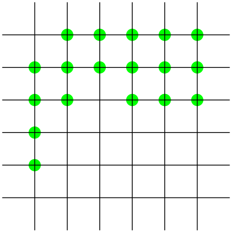
$u \in (\frac{5}{8}, \frac{6}{8}]$

Example

A

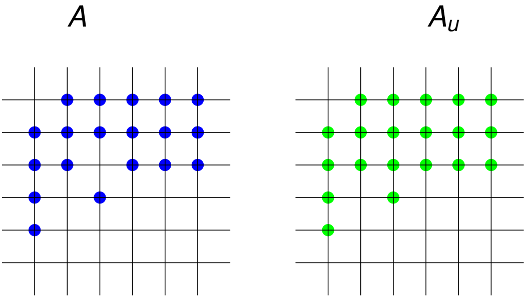


A_u



$u \in (\frac{4}{8}, \frac{5}{8}]$

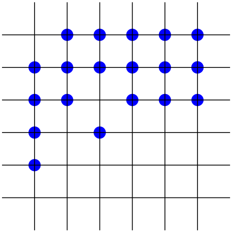
Example



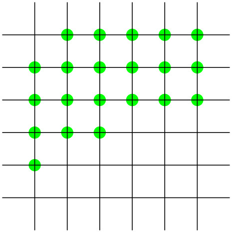
$u \in (\frac{3}{8}, \frac{4}{8}]$

Example

A

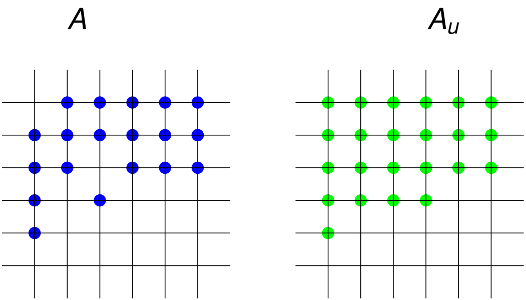


A_u



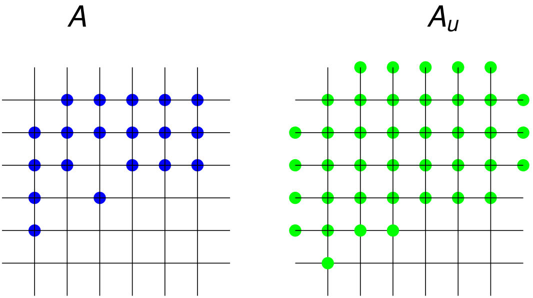
$u \in (\frac{2}{8}, \frac{3}{8}]$

Example



$u \in (\frac{1}{8}, \frac{2}{8}]$

Example



$u \in (0, \frac{1}{8}]$

Relating Original MC to evolving set process

Define another MC on sets with transition kernel $\hat{K}(S, S') = \frac{\pi(S')}{\pi(S)} K(S, S')$. Can show

- $\hat{K}$ is a valid transition matrix.
- $\hat{K}^n(S, S') = \frac{\pi(S')}{\pi(S)} K^n(S, S')$ and so, for any function f, if $S_0 = S$, then expectation $\hat{\mathbb{E}}^n$ under $\hat{K}^n$ is $\hat{\mathbb{E}}^n_S[f(S_n)] = \mathbb{E}^n_S\left[\frac{\pi(S_n)}{\pi(S)} f(S_n)\right]$

Lemma

If $S \subset \Omega$, $y \in \Omega$, and $\Lambda(S, y) = \frac{\pi(y)}{\pi(S)} 1_{(y \in S)}$, $[\Lambda P](S, y) = [\hat{K}\Lambda](S, y)$. That is, $\hat{K}$ is *dual* of P.

Relating n-step transition probability to evolving set process

Lemma

If $x \in \Omega$, $S_0 = \{x\}$, $P^n(x,y) = \hat{\mathbb{E}}^n_{\{x\}}[\pi_{S_n}(y)]$ with $\pi_S(y) := \frac{\mathbf{1}_S(y)\pi(y)}{\pi(S)} = \Lambda(S,y)$.

Proof.

$$P^n(x,y) \overset{(1)}{=} [\Lambda P^n](\{x\},y) \overset{(2)}{=} [\hat{K}^n\Lambda](\{x\},y) \overset{(3)}{=} \hat{\mathbb{E}}^n_{\{x\}}(\pi_{S_n}(y))$$

- $\Lambda(\{x\},y) = \mathbf{1}(x=y)\frac{\pi(y)}{\pi(x)} \Rightarrow (1)$
- Duality $\Rightarrow (2)$
- $[\hat{K}^n\Lambda](\{x\},y) = \sum_{S' \subset \Omega} \hat{K}^n(\{x\},S')\Lambda(S',y) = \sum_{S' \subset \Omega} \hat{K}^n(\{x\},S')\pi_{S'}(y) \Rightarrow (3)$

Relating L^2 distance to evolving sets process

Theorem

If $x \in \Omega$, $S_0 = \{x\}$, then $\|P^n(x,\cdot),\pi\|_2 \leq \hat{\mathbb{E}}^n_{\{x\}} \sqrt{\frac{1-\pi(S_n)}{\pi(S_n)}}$

Proof. $\|P^n(x,\cdot),\pi\|_2 = \left\|\hat{\mathbb{E}}^n_{\{x\}} \pi_{S_n},\pi\right\|_2 \leq \hat{\mathbb{E}}^n_{\{x\}} \|\pi_{S_n},\pi\|_2$

Now, $\|\pi_{S_n},\pi\|_2 = \sqrt{\sum_{y\in\Omega} \pi(y)\left(\frac{\pi_{S_n}(y)}{\pi(y)} - 1\right)^2} =$
$\sqrt{\sum_{y\in\Omega}\left(\frac{\pi^2_{S_n}(y)}{\pi(y)}\right) - 1} = \sqrt{\sum_{y\in S_n}\left(\frac{\pi(y)}{\pi^2(S_n)}\right) - 1} = \sqrt{\frac{1-\pi(S_n)}{\pi(S_n)}}$

Conductance vs. Congestion

Symmetric Conductance

$\tilde{\Phi}(r) := \min_{S\subset\Omega}\left\{\tilde{\Phi}(S) : 0 < \pi(S) \leq r\right\}$ with $\tilde{\Phi}(S) = \frac{Q(S,\bar{S})}{\pi(S)\pi(\bar{S})}$ and $\tilde{\Phi}^* = \min_{S\subset\Omega}\tilde{\Phi}(S)$.

$\tilde{\Phi}^*$ and Φ^* are equal up to a factor of $1/2$.

Congestion

$\mathcal{C}(r) := \max_{S\subset\Omega}\{\mathcal{C}(S) : 0 < \pi(S) \leq r\}$ with $\mathcal{C}(S) = \mathbb{E}_S\left[\sqrt{\frac{(1-\pi(S'))\pi(S')}{(1-\pi(S))\pi(S)}}\right]$ and $\mathcal{C}^* = \max_{S\subset\Omega}\mathcal{C}(S)$.

For lazy MC, $\mathcal{C}(S) \leq \sqrt{1-\tilde{\Phi}(S)^2}$.

Simple Result

Lemma

$\hat{\mathbb{E}}\{x\}^{n+1} f(\pi(S_{n+1})) - \hat{\mathbb{E}}^n_{\{x\}} f(\pi(S_n)) = -\hat{\mathbb{E}}^n_{\{x\}} f(\pi(S_n))(1 - C(S_n))$
with $f(z) = \sqrt{\frac{1-z}{z}}$.

Proof.
$\hat{\mathbb{E}}^{n+1}_{\{x\}} f(\pi(S_{n+1})) = \hat{\mathbb{E}}^n_{\{x\}} \sum_{S\subset\Omega} \hat{K}(S_n, S) f(\pi(S))$
$= \hat{\mathbb{E}}^n_{\{x\}} f(\pi(S_n)) K(S_n, S) \sqrt{\frac{\pi(S)(1-\pi(S))}{\pi(S_n)(1-\pi(S_n))}}$
$= \hat{\mathbb{E}}^n_{\{x\}} f(\pi(S_n)) \mathcal{C}(S_n)$
Since $\mathcal{C}^* \geq \mathcal{C}(S)\ \forall S \subset \Omega$,
$\hat{\mathbb{E}}^{n+1}_{\{x\}} f(\pi(S_{n+1})) \leq \mathcal{C}^* \hat{\mathbb{E}}^n_{\{x\}} f(\pi(S_n))$.

Simple bound on mixing time

- $\|P^n(x,\cdot),\pi\|_2 \leq \hat{\mathbb{E}}^n_{\{x\}}\sqrt{\frac{1-\pi(S_n)}{\pi(S_n)}} \leq (\mathcal{C}^*)^n f(\pi(S_0)) = (\mathcal{C}^*)^n f(\pi(x))$.
- Solving for when RHS drops to ϵ and using $\log \mathcal{C}^* \leq -(1-\mathcal{C}^*)$ gives

$$\begin{aligned}\tau^2(\epsilon) &\leq \left\lceil \frac{1}{1-\mathcal{C}^*}\log\frac{f(\pi_*)}{\epsilon}\right\rceil \\ &\leq \left\lceil \frac{1}{1-\mathcal{C}^*}\log\frac{1}{\epsilon\sqrt{\pi_*}}\right\rceil \\ &\leq \left\lceil \frac{2}{(\tilde{\Phi}^*)^2}\log\frac{1}{\epsilon\sqrt{\pi_*}}\right\rceil\end{aligned}$$

using $\mathcal{C}^* \leq \sqrt{1-(\tilde{\Phi}^*)^2}$

Improved bound on mixing time

- Before we used $\mathcal{C}^* \geq \mathcal{C}(S), \forall S \subset \Omega$
- Infact, we can do better by using $\mathcal{C}(\pi(S)) \geq \mathcal{C}(S), \forall S \subset \Omega$
- Exploiting this fact and using *a bit of analysis*,

$$\tau^2(\epsilon) \leq \left[\int_{4\pi_*}^{4/\epsilon^2} \frac{du}{u(1 - \mathcal{C}(u))} \right]$$
$$\leq \left[\int_{4\pi_*}^{4/\epsilon^2} \frac{du}{u\tilde{\Phi}^2(u)} \right]$$

using $\mathcal{C}(u) \leq \sqrt{1 - \tilde{\Phi}(u)^2}$

Remarks

- Evolving set L^2 bounds are at least as good as bounds discussed in class.
- Conductance bounds apply to non reversible chains.
- Conductance is inappropriate for non-lazy chains. In particular, conductance does not distinguish between periodic and aperiodic chains.