Wednesday, September 22, 2021 2:47 PM

MAT 1841 Prof. Yun William Yu

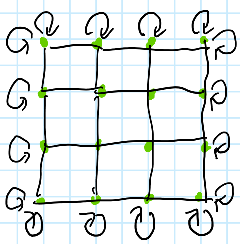
郁 昀

Previously: Conductance + ε-mixing time

Today:

- Minor correction to proofs
- Finish proving lemma
- Convergence of random walks on a grid
- Metropolis-Hasting + Gibbs Sampling
- Volume via MCMC
- Page Rank

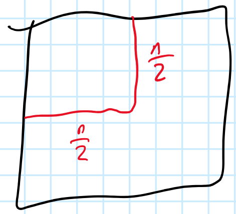
Random walk on 2D grid



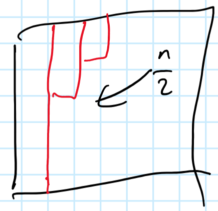
$n \times n$ grid

$p_{ij} = \frac{1}{4}$ if i and j are adjacent

$\pi_i = \frac{1}{n^2}$ is the stationary dist.



Consider any subset S $|S| \le \frac{n^2}{2}$

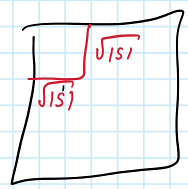


If $|S| \ge \frac{n^2}{4}$, then at least n edges leave S

Thus $$\Phi(S) = \sum_{i \in S} \frac{\pi_i}{\pi(S)} \sum_{j \in \bar{S}} p_{ij} = \frac{1}{|S|} \sum_{\substack{i \in S \\ j \in \bar{S}}} p_{ij} \ge \frac{n}{4|S|} \ge \frac{1}{2n} = \Omega\left(\frac{1}{n}\right).$$

If $|S| < \frac{n^2}{4}$: Case 1: at least a col/row

Case 2: rectangular shape in corner



at least $2\sqrt{|S|}$ pts adjacent to S

Thus $$\Phi(S) = \sum_{i \in S} \frac{\pi_i}{\pi(S)} \sum_{j \in \bar{S}} p_{ij} = \frac{1}{|S|} \sum_{\substack{i \in S \\ j \in \bar{S}}} p_{ij} = \frac{2\sqrt{|S|}}{4|S|} = \frac{1}{2} \cdot \frac{1}{\sqrt{|S|}} \ge \frac{1}{2} \cdot \frac{2}{n} = \frac{1}{n} = \Omega\left(\frac{1}{n}\right)$$

$$\Rightarrow \Phi(S) = \Omega\left(\frac{1}{n}\right). \Rightarrow \varepsilon\text{-mixing time} = O\left(\frac{n^2 \ln \frac{1}{\varepsilon}}{\varepsilon^3}\right).$$

Higher-dim grids

$\underbrace{n \times n \times n \times \cdots \times n}_{d \cdot \text{dim}}$ lattice with same self-loop at boundaries construction.

$\pi_i = \frac{1}{n^d}$ because every node has same degree

Similar reasoning as above gives $\Phi = \Omega\left(\frac{1}{dn}\right)$

$\Rightarrow$ ... $(d^3 n^2 \ldots)$ ← poly in n, d

Similar reasoning as above gives $\Phi = \Omega\left(\frac{1}{dn}\right)$

$\Rightarrow$ ε-mixing time is $O\left(\frac{d^3 n^2 \ln n}{\varepsilon^3}\right)$ ← poly in n, d rather than exponential

much smaller than the cover time of $\Omega(n^d)$

Metropolis-Hasting

Let $\vec{\pi} = (\pi_1, \ldots, \pi_n)$ be our desired stationary dist. on states $V = \{1, 2, \ldots, n\}$.

Let $G = (V, E)$ be any connected undirected graph with max degree r.

At node i, "choose" each adjacent node w.p. $\frac{1}{r}$ $\left(\text{prob. } 1 - \frac{\deg(i)}{r} \text{ of choosing none}\right)$

If an adjacent node j is chosen, if $\pi_j \geq \pi_i$, go to j.

if $\pi_j < \pi_i$, go to j w.p. $\frac{\pi_j}{\pi_i}$.

Otherwise, stay in place.

i.e. $p_{ij} = 0$, if edge (i, j) does not exist and $i \neq j$

$p_{ij} = \frac{1}{r} \min\left(1, \frac{\pi_j}{\pi_i}\right)$ if $\exists (i, j)$ and $i \neq j$.

$p_{ii} = 1 - \sum_j p_{ij}$.

Thus, if $\nexists (i, j)$, $\pi_i p_{ij} = 0 = \pi_j p_{ji}$

if $\exists (i, j)$, $\pi_i p_{ij} = \frac{\pi_i}{r} \min\left(1, \frac{\pi_j}{\pi_i}\right) = \frac{1}{r} \min(\pi_i, \pi_j) = \pi_j p_{ji}$.

$\Rightarrow$ $\vec{\pi}$ is the stationary dist. of the Markov chain.

Gibbs sampling

Let $p(\vec{x})$ be the target distribution, where $\vec{x} = (x_1, \ldots, x_d)$

Let $G = (V, E)$ be an undirected graph where V is the set of all states of $\vec{x}$

and $(\vec{x}, \vec{y}) \in E$ iff $|\vec{x} - \vec{y}|_0 = 1$ (i.e. two states are connected by an edge if they differ in only one coordinate)

Each step choose 1 coordinate to update and keep all others fixed

Each step, choose 1 coordinate to update, and keep all others fixed.

Suppose $|\vec{x}-\vec{y}|_0 = 1$, WLOG suppose they differ in first coordinate.

Then set $P_{\vec{x}\vec{y}} = \frac{1}{d} p(y_1 | x_2, \ldots, x_d)$

$$\Rightarrow p(\vec{x}) P_{\vec{x}\vec{y}} = \frac{p(\vec{x})}{d} p(y_1 | x_2, \ldots, x_d) = \frac{p(x_1, \ldots, x_d)\, p(y_1, \ldots, y_d)}{d \cdot p(x_2, \ldots x_d)} = p(\vec{y}) \cdot P_{\vec{y}\vec{x}}.$$