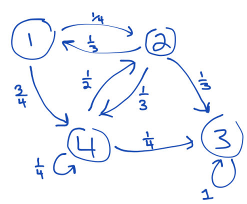
Today more Markov chains
- mixing time
- random walks on graphs

Sequence of r.v.'s $(X_t)_{t=0}^{\infty}$ on discrete state space

$$\Pr(X_t = j \mid X_{t-1} = i, X_{t-2}, \dots, X_0) = \Pr(X_t = j \mid X_{t-1} = i) = p_{ij}$$

Finite Markov Chains

- random walk on directed graph
- each vertex is a "state" of MC.
- each arc describes corresponding transition probability



$$\begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 0 & \frac14 & 0 & \frac34 \\ 2 & \frac13 & 0 & \frac13 & \frac13 \\ 3 & 0 & 0 & 1 & 0 \\ 4 & 0 & \frac12 & \frac14 & \frac14 \end{array}$$

Use X_t to denote state at time t

$\Pr(X_{t+1} = j \mid X_t = i) = p_{ij}$ $P = (p_{ij})$ transition prob matrix

$\vec{p}^{\,t} = (p_1^t, p_2^t, \dots, p_n^t)$ describes prob distn over states at time t

$p_i^t = \Pr(X_t = i)$. starting from some state p^0

$\vec{p}^{\,0} = (1, 0, \dots 0)$ means start in state i

$\vec{p}^{\,0} = (\frac1n, \frac1n, \dots, \frac1n)$ means start in uniformly random state

$$\vec{p}^{\,t+1} = \vec{p}^{\,t} P \implies \vec{p}^{\,t+m} = \vec{p}^{\,t} P^m$$

called a Markov chain because it has "Markov property" = next state depends on current state but not on history

Irreducible Markov chain

corresponding graph strongly connected $\forall i,j$ $\exists t$ s.t. $p_{ij}^t > 0$

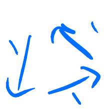
<u>period</u> of state i $= \gcd\{n \geq 1 \mid p_{ii}^n > 0\}$

Markov chain is <u>aperiodic</u> if period of every state is 1

All Markov chains we will consider will be finite, irreducible & aperiodic

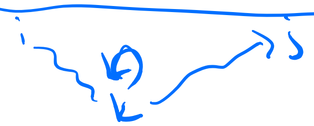
$\implies \exists N > 0$ s.t. P^n strictly positive <u>$\forall n \geq N$</u>

$(P^n)_{ij} = \Pr(X_{t+n} = j \mid X_t = i) > 0$



S.C. ✓

if $\exists$ state k s.t. $p_{kk} > 0$



$(P^n)_{ij} > 0$

$(i,j$ $i = j$

$P' = \frac{1}{2}(I + P)$ $\quad p'_{ij} = \begin{cases} \ldots \\ \frac{1}{2} p_{ij} & i \neq j \end{cases}$

A **stationary** distn $\vec{\pi} = (\pi_1, \pi_2, \ldots, \pi_n)$ of a M.C. is a prob distn s.t.

$$\vec{\pi} = \vec{\pi} P \quad \text{"fixed point"}$$

$$\forall j \quad \pi_j = \sum_i \pi_i p_{ij}$$

$\Rightarrow \vec{\pi} = \vec{\pi} P^t$

Fundamental Thm of Markov Chains

For any finite, irreducible, aperiodic MC

1. $\exists$ stationary distn $\vec{\pi}$ (with $\pi_i > 0 \ \forall i$)
2. $\vec{\pi}$ is unique (unique left eigenvector w/ eigenvalue 1)
3. $\pi_i = \frac{1}{h_{ii}}$
4. $\forall i,j \quad \lim_{n \to \infty} P^t_{ij} = \pi_j$

Notation

$H_{ij} = \min_{t \geq 1} \{ X_t = j \mid X_0 = i \}$

$h_{ij} = E(H_{ij})$

h_{ii} = expected first return time.

P - stochastic row sums = 1

Simple cases:

1. P symmetric $\quad \Pi = (\frac{1}{n}, \ldots, \frac{1}{n})$

$$\underline{\underline{\Pi_j}} \overset{?}{=} \sum_i \Pi_i p_{ij} \underset{\Pi_i = \frac{1}{n}}{=} \frac{1}{n} \sum_i p_{ij} = \frac{1}{n} \underbrace{\sum_i p_{ji}}_{=1} = \frac{1}{n}$$

2. P doubly stochastic (col sums are also = 1)

$$\Pi = (\frac{1}{n}, \ldots, \frac{1}{n})$$

3. P "reversible": If $\exists \Pi$ s.t. $\forall i,j$

$$\Pi_i p_{ij} = \Pi_j p_{ji}$$

$$\Pi_j \overset{?}{=} \sum_i \Pi_i p_{ij} \underset{\text{rev}}{=} \sum_i \Pi_j p_{ji} = \Pi_j \sum_i p_{ji} = \Pi_j$$

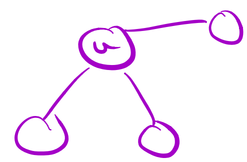
Random walks on graphs.

$G=(V,E)$

$$p_{uv} = \begin{cases} \frac{1}{d_u} & (u,v)\in E \\ 0 & \text{o.w.} \end{cases}$$

$P = D^{-1} A$

$D^{-1} = \begin{pmatrix} \frac{1}{d_1} & & 0 \\ & \frac{1}{d_2} \ddots & \\ 0 & & \frac{1}{d_n} \end{pmatrix}$ where d_i is degree of vertex i; A = adjacency matrix of graph



graph connected, aperiodic $\neq$ bipartite.

Some interesting questions:

1) What is limiting dist'n of random walk? (i.e. stationary dist'n)
2) How long does it take before the walk approaches the limiting dist'n? mixing time
3) Starting from vertex s, what is the exp # of steps to first reach t? E hitting times
4) How long does it take to reach every vertex at least once? E cover time

Claim 1: $\pi_u = \frac{d_u}{2m}$ m total # edges in graph.

$(u,v)\in E$

$$\pi_u p_{uv} = \frac{d_u}{2m}\cdot\frac{1}{d_u} = \frac{1}{2m} = \pi_v p_{vu}$$

$$\sum_u \frac{d_u}{2m} = 1$$

$\forall u,v \quad \pi_u p_{uv} = \pi_v p_{vu}$

Want to bound mixing time for random walks on graphs. start at $\vec{p}^0$

find for what t, $\sup_{p^0} \| p^0 P^t - \pi \|_1 \leq \varepsilon$

($p^0 P^t$: dist'n over states at time t)

~~Proof of thm~~ for d-regular graphs using spectral approach

$\pi = (\frac{1}{n}, \dots, \frac{1}{n})$ $\qquad \pi_u = \frac{d}{2m} = \frac{d}{dn} = \frac{1}{n}$

Spectral Thm

If $M \in \mathbb{R}^{n \times n}$ symmetric, then

- all its eigenvalues are real
- $\exists$ full basis of orthonormal eigenvectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ where $\vec{v}_i \cdot \vec{v}_j = 1_{i=j}$

$M\vec{v}_i = \lambda_i \vec{v}_i$

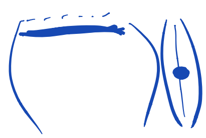
$$M = \sum_{i=1}^{n} \lambda_i v_i v_i^T = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ & & \ddots \\ 0 & & \lambda_n \end{pmatrix} \begin{pmatrix} - & v_1^T & - \\ - & v_2^T & - \\ & \vdots & \\ - & v_n^T & - \end{pmatrix}$$

$\qquad\qquad \Phi^T \qquad \Lambda \qquad \Phi$

$\Phi^T \Phi = \Phi \Phi^T = I$

$P = D^{-1}A$ symmetric, $\frac{1}{d}$ along diag $\Rightarrow$ $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ real; $\vec{v}_1 \quad \vec{v}_2 \quad \vec{v}_n$ orthonormal basis

row sums $=1$

P 

$\lambda_1 = 1 \qquad \vec{e} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$

$\vec{v}_1 = \begin{pmatrix} \frac{1}{\sqrt{n}} \\ \vdots \\ \frac{1}{\sqrt{n}} \end{pmatrix}$ $\quad \|v_1\|_2 = 1$

2 other facts

① G is connected iff $\lambda_2 < 1$

② G is nonbipartite iff $\lambda_n > -1$

$\Rightarrow$ $|\lambda_i| < 1$, $2 \leq i \leq n$

$\lambda = \max_{2 \leq i \leq n} |\lambda_i|$

$1-\lambda$ governs rate of convergence to stationarity.

we will show that $\forall p^0$

$\|p^0 P^t - \pi\|_2 \leq \frac{1}{10}$ $\Leftarrow$ $\left(\frac{1}{n^c}\right)$

if $t = \Omega\left(\frac{\ln n}{1-\lambda}\right)$

$$P = \sum_{i=1}^{n} \lambda_i v_i v_i^T$$

$$P^2 = \sum_{i=1}^{n} \lambda_i^2 v_i v_i^T$$

$$P^t = \sum_{i=1}^{n} \lambda_i^t v_i v_i^T$$

$$= \underbrace{v_1 v_1^T} + \sum_{i=2}^{n} \lambda_i^t v_i v_i^T$$

$$= \frac{1}{n} J + \sum_{i=2}^{n} \lambda_i^t v_i v_i^T$$

↑ all 1's matrix

Write $\underset{\text{row}}{p^0} = \sum_{j=1}^{n} c_j v_j^T$

$c_j = (p^0, v_j)$

$$\underset{\text{row}}{p^0} P^t = p^0 v_1 v_1^T + \sum_{j=1}^{n} c_j v_j^T \left(\sum_{i=2}^{n} \lambda_i^t v_i v_i^T \right)$$

$(\frac{1}{\sqrt{n}} \dots \frac{1}{\sqrt{n}})$

$$= \frac{1}{\sqrt{n}} v_1^T$$

$$= \left(\frac{1}{n}, \dots \frac{1}{n}\right) + \sum_{i=2}^{n} c_i \lambda_i^t v_i^T$$

π

$$p^0 v_1 = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} p_i^0 = \frac{1}{\sqrt{n}}$$

$$p^0 P^t = \pi + \sum_{i=2}^{n} c_i \lambda_i^t v_i^T$$

$$\| p^0 P^t - \pi \|_1 = \left\| \sum_{i=2}^{n} c_i \lambda_i^t v_i^T \right\|_1$$

$$\leq \sqrt{n} \left\| \sum_{i=2}^{n} c_i \lambda_i^t v_i^T \right\|_2$$

Cauchy Schwarz

$$\left\| \sum_{i \geq 2} c_i \lambda_i^t v_i \right\|_2^2 = \sum_{i \geq 2} c_i^2 \lambda_i^{2t}$$

$$\leq \lambda^{2t} \sum_{i \geq 2} c_i^2$$

$$\leq \sum_{i \geq 1} c_i^2 = \|p_0\|_2^2 \leq 1$$

$$\leq \sqrt{n} \, \lambda^t$$

$$[1 - (1-\lambda)]^t$$

$$\leq \sqrt{n} \, e^{-t(1-\lambda)}$$

If take $t = O\left(\frac{\ln n}{1-\lambda}\right)$

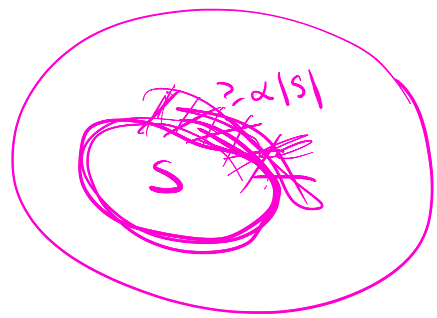
$t = \frac{2 \ln n}{1-\lambda}$.

$$\leq \frac{1}{n^{3/2}}$$

d-regular graphs

diameter of graph of d const is $\Omega(\log n)$

d-reg graphs with $1-\lambda > \varepsilon$ are called "expander graphs"



- mixing time $O(\log n)$
- sets have const expansion.
- look like random d-regular graphs

random d regular graph

$$\lambda_2 = \Theta\left(\frac{1}{\sqrt{d}}\right)$$

- $\exists$ explicit constructions for expander graphs

Nonregular graphs

$P = D^{-1}A$

$D^{-1}A$ has same eigenvalues as $D^{-1/2}AD^{-1/2}$

$$D^{-1/2}AD^{-1/2}x = \lambda x$$

$$D^{-1/2}D^{-1/2}AD^{-1/2}x = \lambda D^{-1/2}x$$

$$D^{-1}A\left(D^{-1/2}x\right) = \lambda\left(D^{-1/2}x\right)$$

$Q = D^{-1/2}AD^{-1/2}$ $\quad D^{1/2} \quad D^{1/2}$

$$P = D^{-1/2}QD^{1/2}$$

$$P^t = D^{-1/2}Q^tD^{1/2}$$

$P^2 = D^{-1/2}$

$$\sum \lambda_i^t v_i v_i^T$$

. . . . .

random walk on graph $\quad \pi_i = \frac{d_i}{2m}$

h_{ii} = expected time starting at i to return to i $= \frac{2m}{d_i}$

Lemma $\forall$ edge (i,j) $\quad h_{ij} + h_{ji} \le 2m$

h_{ij}: expected time starting at i to reach j

$2m$: # edges in graph

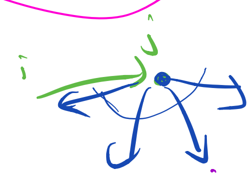
Pf Consider r.w. on directed edges.

$\rightarrow$ 2 states: $i \to j$, $j \to i$

New MC: $2m$

$$q_{(i,j),(j,k)} = \frac{1}{d_j}$$

$$q_{(i,j)(k,\ell)} = 0 \quad k \neq j$$



(j,k)

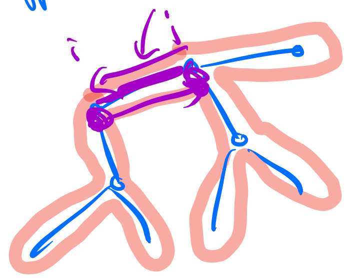
Claim: TPM of random walk on directed edges doubly stochastic

$$\sum_{i \text{ s.t. } (i,j)\in E} q_{(i,j)(j,k)} = d_j \cdot \frac{1}{d_j} = 1$$

$\Rightarrow \pi$ is uniform $\Rightarrow \pi_i = \frac{1}{2m}$

$$\underbrace{h_{ij} + h_{ji}}_{\text{original chain}} \le \underbrace{h'_{i\to j,\, i \to j}}_{\text{in new chain}} = 2m$$

$\Longrightarrow$ Simple upper bound on cover time of any graph

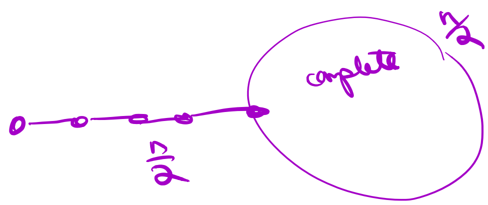


$h_{ij}+h_{ji}$

$$C(G) \le \sum_{(i,j)\in T} (h_{ij}+h_{ji})$$

$$\le (n-1)\,2m$$

$O(nm)$



$\Omega(n^2)$

$C(G) = O(n^3)$