March 8, 2022 CPSC 536F

HW 1: Due March 15

Today:

- Expander Mixing

- Min-max Rayleigh quotient

- For fixed d, $n \to \infty$,
  a d-regular graph on n-vertices has:

$$\lambda_2 \geq 2\sqrt{d-1}\left(1 - O\left(\frac{1}{\log_{d-1} n}\right)\right)$$

Last week:

G d-regular graph on n vertices,

then

$$A_G = \frac{d}{n}\begin{bmatrix} 1 & \cdots & 1 \\ \vdots & & \\ 1 & \cdots & 1 \end{bmatrix} + \mathcal{E}$$

if d fixed, $n \to \infty$, then either $-\frac{d}{n}$ in most places, and $1-\frac{d}{n}$ in nd entries

where

$$\left(\begin{array}{l} \mathcal{E} = \sum_{i=2}^{n} \lambda_i \vec{v}_i \vec{v}_i^T, \\ \lambda_1 = d \geq \lambda_2 \geq \cdots \geq \lambda_n \\ A\vec{v_i} = \lambda_i \vec{v}_i, \quad \vec{v}_1, \ldots, \vec{v}_n \text{ orthonormal} \end{array} \right)$$

Really: $\vec{v}_1 = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} / \sqrt{n}$

and

$$\mathcal{E} = \sum_{i=2}^{n} \lambda_i v_i \vec{v_i}$$

view as

$$\mathcal{E} v = \quad \left(\text{proj}_{\mathbb{1}^\perp} \vec{v}\right)$$

=

Examples!

$$K_n, \quad A_{K_n} = \begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & \ddots & & \vdots \\ \vdots & & & 1 \\ 1 & \cdots & & 0 \end{bmatrix}$$

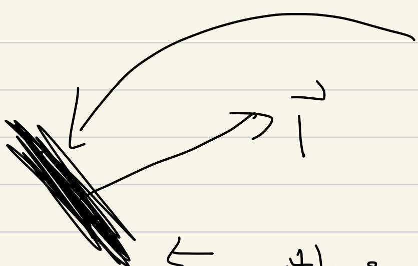
$$= \begin{bmatrix} 1 & \cdots & 1 \\ & \ddots & \\ 1 & & 1 \end{bmatrix} - I$$

Given $\vec{v}$!

$$\text{proj}_{\vec{1}} \vec{w} = \vec{v}_1 \frac{\vec{w}\cdot\vec{v}_1}{\vec{v}_1\cdot\vec{v}_1}$$

$$= \begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}\frac{1}{\sqrt{n}} \; \frac{\begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}\frac{1}{\sqrt{n}}\cdot\vec{w}}{1}$$

$$= \begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}\frac{1}{n}\left(w_1+w_2+\dots+w_n\right)$$



$$\text{proj}_{\vec{1}^{\perp}}\vec{w} = \vec{w} - \text{proj}_{\vec{1}}(w)$$

$\vec{1}^{\perp}$ ← orthog comp to $\vec{1}$

If $A, B \subset \overline{V}_G$

$e(A,B) =$ # edges from A to B

$$= \underbrace{[1\text{'s}, 0\text{'s}]} A_G \begin{bmatrix} \\ \\ \\ \end{bmatrix} \chi_B$$

1's are
the vertices
corres to A

$$\underbrace{\qquad\qquad\qquad}$$

χ_A or $\mathbb{1}_A$

$$\chi_A^T \, \mathrm{Adj}_G \, \chi_B$$

$$=$$

$$v^T A \omega$$

$$= \sum_{i,j} v_i (A)_{ij} \omega_j$$

$$=$$

$$\sum_{i,j} (\chi_A)_i A_{ij} (\chi_B)_j$$

$$= \sum_{i \in A,\, j \in B} 1 \cdot A_{ij} \cdot 1 = e(A,B)$$

$$e(A,B) = \chi_A^T A_G \chi_B$$

$$= \chi_A^T \frac{d}{n} \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix} \chi_B$$

$$+ \chi_A^T \mathcal{E} \chi_B$$

$$= \frac{d}{n} |A| \cdot |B|$$

expected # edges from A to B if A, B fixed, G is chosen "randomly"

$$+ \underbrace{\left(\text{proj}_{\vec{1}^\perp} \chi(A)\right)^T \mathcal{E} \left(\text{proj}_{\vec{1}^\perp} \chi(B)\right)}$$

$$\leq \underbrace{\left(\max_{i \geq 2} |\lambda_i|\right)}_{\rho} \|\quad\| \; \|\quad\|$$

Now! So $A \subset V_G$

$$\chi_A = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ \vdots \end{bmatrix} \leftarrow \begin{array}{l} |A| \quad 1\text{'s} \\ n-|A| \quad 0\text{'s} \end{array}$$

$$\mathrm{proj}_{\mathbb{1}}(\chi_A) = \frac{|A|}{n} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$\mathrm{proj}_{\mathbb{1}^\perp}(\chi_A) = \chi_A - \frac{A}{n} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} \quad \end{bmatrix} \leftarrow \text{has } |A| \quad 1 - \frac{|A|}{n}$$

$$\text{and } n-|A| \quad -\frac{|A|}{n}$$

Homework:

$$\left\| \mathrm{proj}_{\vec{1}^{\perp}} x_A \right\|_{L^2} = \sqrt{\frac{|A|(n-|A|)}{n}}$$

$=$

$|A| = \frac{n}{2}$

$$\sqrt{\frac{(n/2)(n/2)}{n}} = \frac{\sqrt{n}}{4}$$

$\Rightarrow$

$$\left| e(A,B) - \frac{d}{n}|A|\,|B| \right|$$

$$\leq \rho \sqrt{\frac{|A|(n-|A|)}{n}} \sqrt{\frac{|B|(n-|B|)}{n}}$$

where $\rho = \max\limits_{i \geq 2} |\lambda_i|$

for simplicity we often use:

$$\left. \begin{aligned} &\left| e(A, B) - \frac{d}{n} |A|\,|B| \right| \\ &\qquad \leq \rho \sqrt{|A|\,|B|} \end{aligned} \right\} \text{Expander mixing lemma}$$

reasonable approx for $|A|, |B|$

far from n, say $\leq \frac{n}{2}$,

or $\leq n\left(\frac{2}{3}\right)$

Shows $e(A,B)$ close to $\frac{d}{n}|A||B|$

provided that

$$\rho\sqrt{|A||B|} \quad \text{bounded away from} \quad \frac{d}{n}|A||B|$$

$$\frac{\rho n}{d} \quad \text{" "} \quad \sqrt{|A||B|}$$

Often used to run A_G k times:

$$G^k: \quad A_{G^k} = (A_G)^k$$

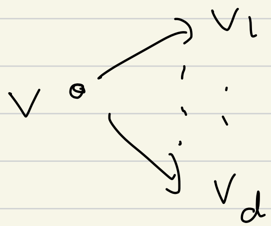
G^k graph with vertices V_G
edges walks of length k in G

So

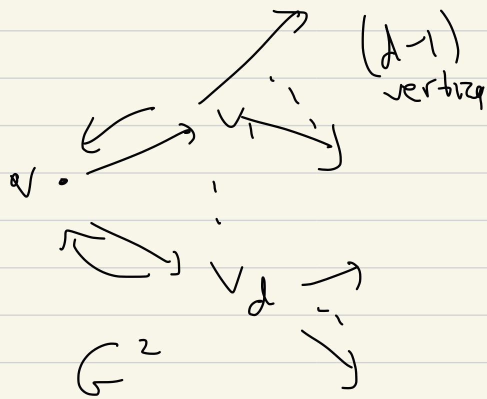
G^k has eigenvalues $(\lambda_1(G))^k, \ldots, (\lambda_n(G))^k$

So $\lambda_1(A_{G^k}) = \lambda_1^k = d^k$

G^k is d^k-regular



G



G^2

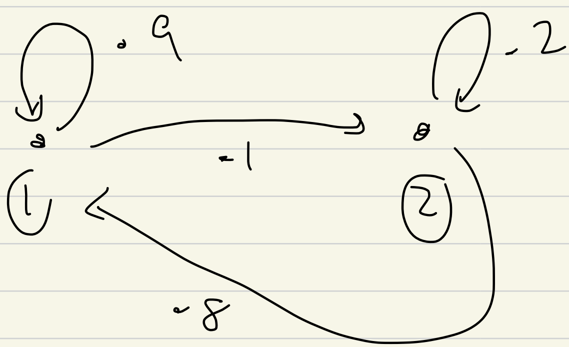
$$\left| e_{G^k}(A,B) - \frac{d^k}{n} |A| \cdot |B| \right|$$

$$\leq \lambda^k \sqrt{|A|\ |B|}$$

Mixing in Markov chains/ matrices!

Markov matrix is an $n \times n$ matrix, P, with non-neg entries, each of whose row sums $= 1$

$$\underbrace{\begin{array}{cc} & \begin{array}{cc} 1 & 2 \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{bmatrix} .9 & .1 \\ .8 & .2 \end{bmatrix} \end{array}}_{P}$$



So

$$[1\ 0]\,P^k = \text{the row vector of probabilities } \cancel{\text{that}} \text{ where you are}$$

1 starting in state 1, 2 after k steps

If P is irreducible

(i.e. for each $i, j \in \{1, \dots, n\}$

there is a walk of (some length)

from i to j) then there

is a unique stationary distribution,

$\pi = \vec{\pi}$ (row vector) s.t.

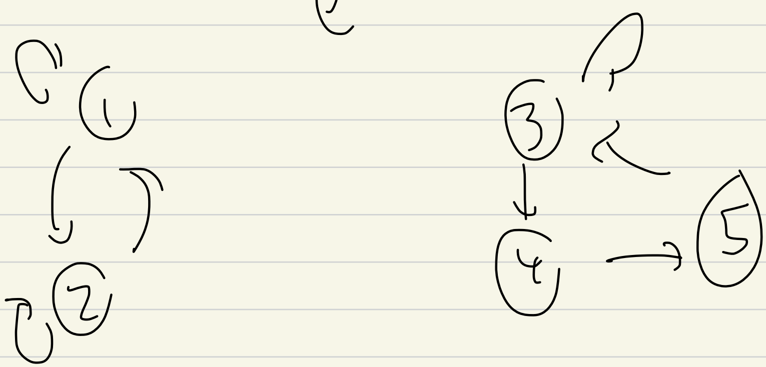
① π is stochastic: $\pi_1 + \dots + \pi_n = 1$

and $\pi_1, \dots, \pi_n \geq 0$

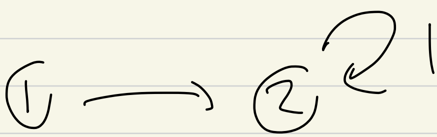
② and $\pi P = \pi$.

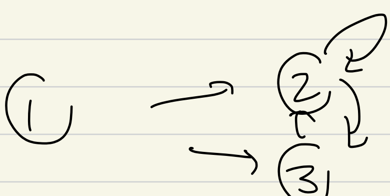
no transitions

If, say

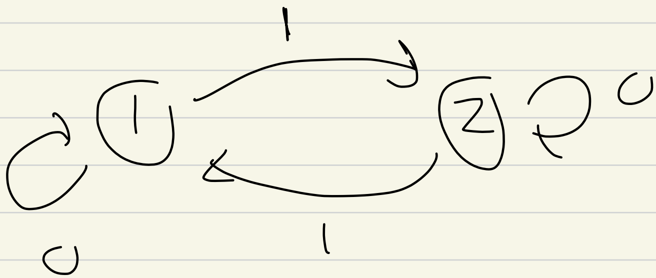


then get a stationary distribution
for each connected comp.





So, for example:



$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

then P represents a bipartite graph, and

$$\overbrace{\begin{bmatrix} 1 & 0 \end{bmatrix}}^{e_1} P^k$$

has no limit as $k \to \infty$.

Really want: pick $i = 1, \ldots, n$

$\vec{e}_i P^k$ = what happens when you start purely at state i, "run" the Markov chain k-times

For each $\varepsilon > 0$, we define the ε-mixing time, $t = t_{mix}(\varepsilon)$, s.t. for all $i \in \{1, \ldots, n\}$

$$\text{distance}\left(\vec{e}_i P^t, \pi\right) \le \varepsilon.$$

For reasons ~~that~~ we'll summarize,

we take

$$\text{distance}(\mu, \nu)$$

$$= \|\mu - \nu\|_{\text{Total Var}}$$

$$= \frac{1}{2} \|\mu - \nu\|_{L^1}$$

$$= \frac{1}{2} \sum_{i=1}^{n} |\mu_i - \nu_i|$$

$=$

It turns out that $t_{mix} = t_{mix}\left(\frac{1}{4}\right)$

Then $t_{mix}\left(\frac{1}{4}\right) \cdot k$ is an upper bound for $t_{mix}\left(1/2^{k+1}\right)$

See: Levin, Peres, and Wilmer

reference: Markov Chains and

Mixing Times, Ch 4

Take G: d-regular on n vertices,

$\rho = \max_{i \geq 2} |\lambda_i|$, we know

for $k = 1, 2, \ldots$,

$$A_G^k = \frac{d^k}{n} \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{bmatrix} + \mathcal{E}^k$$

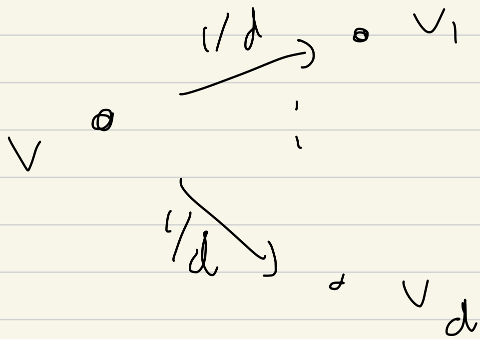
$\|\mathcal{E}^k\|$ is bounded by ρ^k

and $\mathcal{E}$ "lives" on $\mathbb{1}^\perp$.

G is d-regular, so

$$P = \frac{1}{d} A_G = \frac{d}{d}\left(\frac{d}{n}\begin{bmatrix} 1 & \cdots & 1 \\ \vdots & & \\ 1 & & 1 \end{bmatrix} + \mathcal{E} \right)$$

We get a Markov chain



$$\pi = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} \frac{1}{n}$$

Consider the mixing times $t_{mix}(\varepsilon)$:

$$f(t) = e_i^T P^t - \pi$$

$$= \left\| P^t e_i - \begin{bmatrix} 1/n \\ \vdots \\ 1/n \end{bmatrix} \right\|_{L^2}$$

$$\mathrm{Proj}_{\vec{1}}\, e_i = \mathrm{Proj}_{\vec{1}} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ \vdots \end{bmatrix} = \begin{bmatrix} 1/n \\ \vdots \\ 1/n \end{bmatrix}$$

So

$$\left\| P^t e_i - \begin{bmatrix} 1/n \\ \vdots \\ 1/n \end{bmatrix} \right\|_{L^2}$$

$$= \left\| (E/d)^t e_i \right\|_{L^2} \leq (\rho/d)^t \, \|e_i\|_{L^2}$$

$$\leq (\rho/d)^t$$

$$\left\| P^t e_i - \begin{bmatrix} 1/n \\ \vdots \\ 1/n \end{bmatrix} \right\|_{L^1}$$

Fact! If $\vec{x} \in \mathbb{R}^n$, we want

$$\|\vec{x}\|_{L^1} \leq \|\vec{x}\|_{L^2} \cdot (?)$$

$$\downarrow$$

$$|x_1| + \dots + |x_n| \leq \sqrt{x_1^2 + \dots + x_n^2} \quad ?$$

Cauchy Swartz

$$\begin{pmatrix} |x_1| \\ \vdots \\ |x_n| \end{pmatrix} \cdot \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \leq \sqrt{x_1^2 + \dots + x_n^2}\ \underbrace{\sqrt{1^2 + \dots 1^2}}_{n \text{ times}}$$

$$= \|\vec{x}\|_{L^2} \sqrt{n}$$

So

$$\left\| P^t e_i - \begin{bmatrix} 1/n \\ \vdots \\ 1/n \end{bmatrix} \right\|_{L^1} \text{ to be}$$

$$\begin{cases} \leq \varepsilon & t_{mix}(\varepsilon) \\ \leq 1/4 & t_{mix} \end{cases}$$

$$\left\| P^t e_i - \begin{bmatrix} 1/n \\ \vdots \\ 1/n \end{bmatrix} \right\|_{L^1}$$

$$\leq \left(\rho/d\right)^t \sqrt{n} \quad \text{is} \quad \leq \varepsilon$$

if

$$\left(d/\rho\right)^t \geq \sqrt{n}/\varepsilon$$

So

$$t \geq \frac{\log(\sqrt{n}/\varepsilon)}{\log(d/\rho)}$$

$$= \frac{\log\sqrt{n}}{\log(d/\rho)} + \frac{\log(1/\varepsilon)}{\log(d/\rho)}$$

Gives mixing time

$$t_{mix}(\varepsilon) \leq \left\lceil \frac{\log(\sqrt{n}/\varepsilon)}{\log(d/\rho)} \right\rceil$$

(ceiling function), tends to be