A simple bound on L

Lemma: $L_G \preceq 2D$

Pf: $x^T L_G x = \sum_{i,j \in E} (x_i - x_j)^2$

tight when $\quad \leq \sum_{i,j} 2(x_i^2 + x_j^2)$

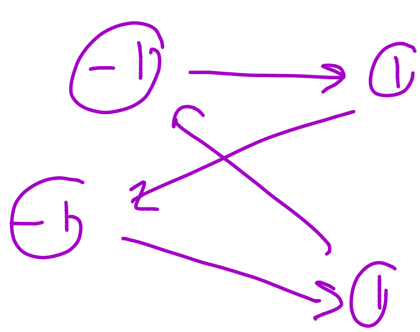
$x_i = -x_j$

$= 2 \sum_{i \in E} d_i x_i^2$

$= 2x^T D x$

Lemma: If there exists $x \neq 0$

$x^T L_G x = 2x^T D x \iff x_i = -x_j$



Bipartide graph.

For $\lambda_2 = 0$ then G is disconnected

For $\lambda_2 = 2 \iff G$ is bipartide.

Next time we will generalize for

(i) Robust clustering (ii) Robust approx to Maxcut

Sparsification

$G = (V, E)$ $e = \{v, u\}$

$$L_G = \sum_{e \in E} \underbrace{\chi_e \chi_e^T}_{\text{rank } 1.} \qquad \chi_e = e_u - e_v$$

Sparsification: Construct graph on samevertex set $H = (V, E_H, \vec{w}_H)$

with

(a) $\| L_G - L_H \| \leq \varepsilon$

(b) H sparse $|E_H| = \tilde{O}(n)$

$= O(n \operatorname{polylog} n)$

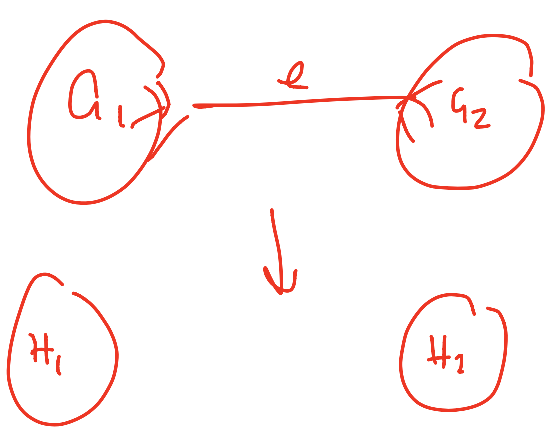
Motivation: faster result on most graph algorithms.

<u>Notion of Error</u>

i) Absolute error

$$|x^T L_G x - x^T L_H x| \leq \varepsilon x^T D x$$

not useful in application..



Then

$$x^T L_e x = (x_i - x_j)^2 = 2(x_i^2 + x_j^2)$$

$$x^T D x = \sum d_i x_i^2 \geq d(x_i^2 + x_j)^2$$

taking $\varepsilon = 2/d$ then these graphs are very similar but very different in terms of random walks.

ii) Relative error.

$$|x^T L_G x - x^T L_H x| \leq \varepsilon x^T L_G x$$

Assigning G_1 $x = 1$ gives
G_2 $x = -1$

$$x^T L_e x \leq \varepsilon x^T L_g x \Rightarrow \varepsilon = 1 \text{ bad}$$

Check: Preserves the degree:

test $x = (0\ 0\ 0 \cdots 1 \cdots 0\ 0\ 0)^T$

K_n $D = (n-1) I$

$|E_G| = O(n^2)$

H

$|E_H| = O\left(\frac{n}{\varepsilon^2}\right)$

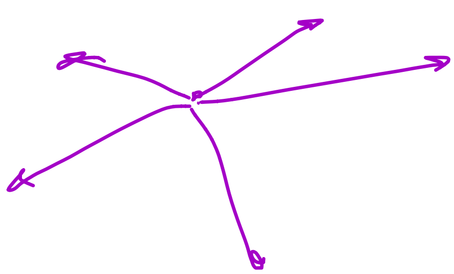
$w \approx \Omega(n)$ $\Big\}$ boost the weight of the edges.

<u>Rmk</u>: $E_H \subseteq E_G$ with the weights changing dramatically.

<u>PSD Sparsification</u>

$A \succeq 0 \qquad A \in \mathbb{R}^{n \times n}$

$$A = \sum_{i=1}^{m} u_i u_i^T \qquad m \gg n$$



dist of vectors.

$$x^T A x = \sum_{i=1} (u_i^T x)^2$$

captures the variance in the u_i direction.

$$\tilde{A} = \sum_{i=1}^{m} \underbrace{s_i}_{\text{R.W. } E(s_i)=1} u_i u_i^T$$

$$\text{e.g. } s_i = \begin{cases} 100 & \text{w.p. } 1/100 \\ 0 & \text{o.w.} \end{cases}$$

Q: How do we construct s_i?

First we want to consider relative error

$$\forall x \; |x^T A x - x^T \tilde{A} x| \le \varepsilon x^T A x$$

$$\Downarrow$$

$\forall y \mid y \perp y - y\, A \; A A \; y \,\| = \varepsilon \|y\|$

$$\tilde{A} = \sum_{i=1}^{T} s_i\, A^{-1/2} u_i u_i^T A^{-1/2}$$

So we look to

compare $I - A^{-1/2} \tilde{A} A^{1/2}$

So we can consider and map

$$I = \sum_{i=1}^{m} u_i u_i^T$$

the problem back. (Partition of unity)

<u>Ex</u>: $A = \sum \vec{a}_i \qquad \tilde{A} = \sum s_i \vec{a}_i$

by defining $\quad p_i \propto \|a_i\|^2$

$$S_i = \begin{cases} \dfrac{1}{p_i} & \text{w.p. } p_i \end{cases}$$

Chosen to minimize variance from

Chernoff Bounds.

Matrix Chernoff Bounds

$X = \sum X_i$ to prove chernoff bounds

$$\mathbb{E}\left(e^{\theta \sum X_i}\right) = \prod_{i=1} \underbrace{\mathbb{E}\left(e^{\theta X_i}\right)}_{\text{MGF}} = \underbrace{\mathbb{E}\left(e^{\theta X}\right)}_{\text{MGF sum.}}$$

$$\mathbb{P}(X \geq \lambda) = \mathbb{P}\left(e^{\theta X} \geq e^{\theta \lambda}\right) \leq \frac{\mathbb{E}\left(e^{\theta X}\right)}{e^{\theta \lambda}}$$

Matrix MGF

$Y \in S^{n \times n}$ symm. random variable.

$$g(\theta) = \underbrace{\mathbb{E}\left(e^{\theta Y}\right)}_{\text{matrix}}$$

exponential.

We want to say:

$$g(\theta) = \mathbb{E}\left(e^{\theta Y}\right) \neq \prod_{i=1}^{n} \mathbb{E}\left(e^{\theta Y_i}\right)$$

not true but

we'll fix it next time

$$\mathbb{P}\left(\lambda_{max}(Y) > \lambda\right) = \mathbb{P}\left(e^{\theta \lambda_{max}} > e^{\theta \lambda}\right)$$

$$\leq \mathbb{P}\left(\mathrm{Tr}\, e^{\theta Y} \geq e^{\theta \lambda}\right)$$

$$\leq \frac{\mathbb{E}\left(\mathrm{Tr}\, e^{\theta Y}\right)}{e^{\theta \lambda}}$$