$G = (V, E)$ $\quad w : E \to \mathbb{R}_+$

We saw cut sparsifier.

$H = (V, F)$ with $w' : F \to \mathbb{R}_+$ is an ε-cut sparsifier if

$\forall\ S \subseteq V$

$$w(\delta_E(S)) \approx (1 \pm \varepsilon)\, w'(\delta_F(S)).$$

H is an ε-approximate spectral sparsifier if $\forall\ \bar{x} \in \mathbb{R}^n$

$$\bar{x}^T \mathcal{L}_G \bar{x} \approx (1 \pm \varepsilon)\, \bar{x}^T \mathcal{L}_H \bar{x}$$

Since $\bar{x}^T \mathcal{L}_h \bar{x} = w(\delta(S))$
when $\bar{x} = \chi_S$
for $S \subseteq V$

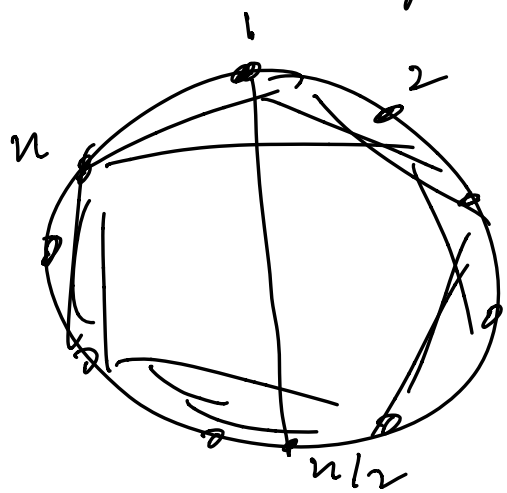
an ε-spectral sparsifier is an ε-cut sparsifier.

However examples show that a cut sparsifier is not always a spectral sparsifier.

See below.

Example which shows cut sparsifier does not imply spectral sparsifier.

Fix n and some parameter k.



Edges $(i-j) \bmod n \le k$ and $(1, \frac{n}{2})$

G' is graph G - edge $(1, \frac{n}{2})$.

Claim: min cut of G' is $\Theta(k)$

$\Rightarrow G'$ is ε-cut sparsifier for G with $\varepsilon = O(\frac{1}{k})$.

Is G' a good spectral sparsifier for G?

Consider $\bar{x} = (0, 1, 2, \ldots, \frac{n}{2}-1, \frac{n}{2}-1, \ldots, 1, 0)$

$$\bar{x}^T \mathcal{L}_{G'} \bar{x} = \Theta(\underline{\underline{n}} k^3).$$

$$\bar{x}^T \mathcal{L}_{G} \bar{x} = \Theta(n k^3) + \left(\frac{n}{2}-1\right)^2$$

If k is constant very different.

So edge $(1, \frac{n}{2})$ important to retain in spectral sparsifier.

$\square$.

Random Sampling based Scheme

[Spielman-Srivastava].

Inspired/motivated by cut sparsification

Recall $\mathcal{L}_G = \sum_{e \in E} w_e \mathcal{L}_e$

Want sparse H s.t.

$$(1-\epsilon)\mathcal{L}_G \preceq \mathcal{L}_H \preceq (1+\epsilon)\mathcal{L}_G.$$

- Pick p_e for each edge e a probability.
- Sample each edge e independently with probability p_e.
  If e chosen set $w'_e = \frac{w_e}{p_e}$

- Output H as the chosen edge weights.

Proof relies on <u>matrix</u> <u>Chernoff</u> inequality.

Let $X_e = \frac{w_e}{p_e} \mathcal{L}_e$ with prob p_e

$= 0$ otherwise.

$$\mathcal{L}_H = \sum_e X_e.$$

$$E[\mathcal{L}_H] = \mathcal{L}_G.$$

Want to use matrix Chernoff inequality to argue that

$\mathcal{L}_H$ is an independent sum of random rank-1 matrices and hence by concentration "behaves" like $\mathcal{L}_G$ whp.

What should p_e be?

Based on the setting of regular Chernoff bound for real values

p_e should correspond to the "importance" of $\mathcal{L}_e$ in the sum $\mathcal{L}_G = \sum_{e \in E} \mathcal{L}_e$

Is there such a measure?

Turns out the $p_e \approx R_{eff}(e)$ the effective resistance is quite useful here. Why?

<u>Lemma 1</u>: $\mathcal{L}_e \preceq R_{eff}(e)\, \mathcal{L}_G$

<u>Lemma 2</u>: $\sum_e R_{eff}(e) = n-1.$

To get concentration we will need to set $p_e = \frac{c \log n}{\varepsilon^2} R_{eff}(e)$

From Lemma 2 it is easy to see that # of edges chosen in expectation is $O\left(\frac{n \log n}{\varepsilon^2}\right)$.

The main part is to understand <u>Lemma 1</u>.

For this we need some background on what it means that

$$\bar{x}^t A \bar{x} \leq \bar{x}^t B \bar{x} \quad \forall \bar{x} \in \mathbb{R}^n.$$

Background on PSD matrices

Recall PSD matrices satisfy

$\bar{x}^T A \bar{x} \geq 0 \quad \forall \bar{x} \in \mathbb{R}^n$. where $A \in \mathbb{R}^{n \times n}$ is symmetric

$A \succeq 0$ is notation to indicate A is psd.

all eigen values are real and non-negative

symmetric

because of psd.

The psd property gives rise to a partial order on $n \times n$ symmetric matrices.

We say $A \succeq B$ if

$$\bar{x}^T A \bar{x} \geq \bar{x}^T B \bar{x} \quad \forall \bar{x} \in \mathbb{R}^n$$

$\Rightarrow A - B \succeq 0$.

$A \succeq B$ and $B \succeq C$

$\Rightarrow A \succeq C$.

also $A \succeq B \Rightarrow A + C \succeq B + C$

for all symmetric A, B, C.

One can define for graphs

$G \succeq H$ if

$L_G \succeq L_H$.

<u>Observation:</u> A, B, C symmetric matrices $(n \times n)$.

$$A \succeq B \Rightarrow CAC^T \succeq CBC^T$$

If C is non-singular then

$$CAC^T \succeq CBC^T \Rightarrow A \succeq B.$$

<u>Proof:</u> Check $\bar{x}^T A \bar{x} \geq \bar{x}^T B \bar{x}$

via $\bar{x}^T C A C^T \bar{x} \geq \bar{x}^T C B C^T \bar{x}$

$\square$

Note that $\bar{x}^T \mathcal{L}_G x^T = \sum_{i \sim j} w_{ij} (x_i - x_j)^2$

Therefore if we <u>reduce</u> weights of edges then we obtain a new graph H s.t

$$G \succeq H.$$

<u>Lemma:</u> Suppose $\mathcal{L}_G \succeq c\, \mathcal{L}_H$ for some constant $c \geq 0$. Then

$$\lambda_k(\mathcal{L}_G) \geq \lambda_k(\mathcal{L}_H) \quad \forall k.$$

<u>Proof:</u> From Courant-Fischer theorem

$$\lambda_k(\mathcal{L}_G) = \min_{\substack{S \subseteq \mathbb{R}^n \\ \dim(S) = k}} \max_{\bar{x} \in S} \frac{\bar{x}^T \mathcal{L}_G \bar{x}}{x^T x}$$

$$\geq c \min_{\substack{S \subseteq \mathbb{R}^n \\ \dim(S)=k}} \max_{\bar{x} \in S} \frac{x^T \mathcal{L}_H \bar{x}}{x^T \bar{x}}$$

$\square$.

Matrix Norms

Given $n \times n$ matrix A one can define several norms.

① Frobenius norm $\|A\|_2$

$$= \sqrt{\sum_{ij} A_{ij}^2}$$ which is same

as l_2 norm of A viewed as a n^2 sized vector.

② Spectral norm based on how A operates.

$$\|A\|_{spec} = \max_{\substack{\|x\|=1 \\ \text{(unit vector)}}} \|A\bar{x}\|_2$$

↑ length of Ax

It is the largest singular value of A.

For PSD matrices it is λ_{max}.

Matrix Chernoff Inequality

Recall Chernoff inequality

$X_1, X_2, \ldots, X_n$ independent random variables in $[0, R]$ (typically $R = 1$)

Let $X = \sum_{i=1}^{n} X_i$.

$$E[X] = \sum_{i=1}^{n} E[X_i] = \mu \text{ say.}$$

Let $\mu_{min} \leq \mu \leq \mu_{max}$.

Then for all $\delta \geq 0$.

① $$\Pr\left[X \geq (1+\delta)\mu_{max}\right] \leq \left[\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right]^{\frac{\mu_{max}}{R}}$$

② $$\Pr\left[X \leq (1-\delta)\mu_{min}\right] \leq \left[\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right]^{\frac{\mu_{min}}{R}}.$$

For $\underline{\underline{0 \le \delta \le 1}}$ we can simplify.

(1) $\le e^{-\delta^2 \mu_{max}/3R}$

(2) $\le e^{-\delta^2 \mu_{min}/2R}$.

Typically we assume $R = 1$.
We can scale and get the results for general R from the one for $R = 1$.

<u>Matrix version</u>

Can we obtain concentration bounds for <u>matrix</u> sums?

<u>Theorem</u> [Tropp] Let $X_1, \ldots, X_k$ be <u>independent</u>, <u>random</u> symmetric real matrices of size $d \times d$ with $0 \leq X_i \leq R\,I$. $X = \sum_{i=1}^{n} X_i$

Let $\mu_{min} \cdot I \leq E[X] \leq \mu_{max} I$.

Then for all $\delta \geq 0$

① $$\Pr\left[\lambda_{max}(X) \geq (1+\delta)\mu_{max}\right] \leq d \cdot \left[\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right]^{\mu_{max}} \leq d\, e^{-\delta^2 \mu_{max}/3R} \quad (\text{if } \delta \leq 1)$$

② $$\Pr\left[\lambda_{min}(X) \leq (1-\delta)\mu_{min}\right] \leq d\left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\frac{\mu_{min}}{R}}$$

$$\leq d \cdot e^{-\delta^2 \mu_{min}/2R}$$

<u>Useful corollary</u>

$$\mu_{min} = \mu_{max} = 1$$

$$\Rightarrow \quad E[X] = I.$$

Then $\Pr[\lambda_{min}(X) \leq (1-\delta)]$

$$\leq e^{-\delta^2/2R}$$

$$\Pr[\lambda_{max}(X) \geq (1+\delta)] \leq e^{-\delta^2/3R}$$

$$\Rightarrow \quad \Pr[(1-\delta)I \preceq X \preceq (1+\delta)I]$$

$$\geq 1 - 2ne^{-\delta^2/3R}$$

$$=.$$

We would like to apply Matrix Chernoff inequality to prove $\mathcal{L}_H \approx \mathcal{L}_G$.

Recall that $\mathcal{L}_H \approx \mathcal{L}_G$

$\iff$ all eigenvalues are similar in both.

However Matrix Chernoff only quantifies λ_{min} and λ_{max}.

For identity matrix I we have

$\lambda_{min} = \lambda_{max} = 1$ and

approximately preserving λ_{min} & λ_{max} $\not\Longrightarrow$ all eigenvalues preserved closely.

Transformation

Let A, B be positive definite matrices.

$$A \leq (1+\varepsilon) B \quad \text{iff}$$

$$B^{-\frac{1}{2}} A B^{-\frac{1}{2}} \leq (1+\varepsilon) I$$

We would like to apply above for psd matrices but a psd matrix may not have an inverse but pseudo inverse will work.

Claim: $\mathcal{L}_H \preceq (1+\varepsilon) \mathcal{L}_G$

$$\iff \mathcal{L}_G^{+/2} \mathcal{L}_H \mathcal{L}_G^{+/2} \preceq \mathcal{L}_G^{+/2} \mathcal{L}_G \mathcal{L}_G^{+/2}$$

$\mathcal{L}_G^{+/2}$ is square root of pseudoinverse

Let $\Pi = \mathcal{L}_G^{+/2} \mathcal{L}_G \mathcal{L}_G^{+/2}$

Π is <u>not the identity matrix</u>

but is the identity operator in the space orthogonal to the null space of $\mathcal{L}_G$ which is $\bar{1}$.

$$\mathcal{L}_G^{+/2} = \sum_{i \geq 2} \frac{1}{\sqrt{\lambda_i}} \bar{u}_i \bar{u}_i^T$$

$$\mathcal{L}_G = \sum_{i=1}^{n} \frac{1}{\lambda_i} \bar{u}_i \bar{u}_i^T$$

Hence for any $\bar{v} \perp \bar{1}$

$$\left(\mathcal{L}_G^{+/2} \mathcal{L}_G \mathcal{L}^{+/2}\right) \bar{v} = \bar{v}.$$

$\Rightarrow$ λ_{max} and λ_{min} of operator $= 1$.

Wanto to prove Lemma 1.

$$\mathcal{L}_H \approx (1 \pm \varepsilon)\mathcal{L}_G.$$

Assume $w_e = 1$ $\forall e \in E$ for simplicity.

$$\mathcal{L}_H = \sum_e X_e \qquad E[X_e] = \mathcal{L}_e$$

$$\mathcal{L}_G = \sum_e \mathcal{L}_e$$

Consider $\mathcal{L}_G^{+/2} \mathcal{L}_H \mathcal{L}_G^{+/2}$

$$= \sum_e \mathcal{L}_G^{+/2} X_e \mathcal{L}_G^{+/2}$$

$$E\left[\mathcal{L}_G^{+/2} \mathcal{L}_H \mathcal{L}_G^{+/2}\right] = \Pi = \mathcal{L}_G^{+/2} \mathcal{L}_G \mathcal{L}_G^{+/2}$$

We want to use Tropp's inequality

We set $p_e = \dfrac{R_{eff}(e)}{R}$

where $R = \dfrac{c\,\varepsilon^2}{\log n}$. for some

Let $Y_e = \mathcal{L}_G^{+/2} X_e \cdot \mathcal{L}_G^{+/2}$

We saw that $\mathcal{L}_e \preceq R_{eff}(e)\,\mathcal{L}_G$

$$\Rightarrow \quad \mathcal{L}_G^{+/2} \mathcal{L}_e \mathcal{L}_G^{+/2} \preceq R_{eff}(e)\, \mathcal{L}_G^{+/2} \mathcal{L}_G \mathcal{L}^{+/2}$$

Recall $Y_e = \dfrac{1}{p_e} \mathcal{L}_G^{+/2} \mathcal{L}_e \mathcal{L}_G^{+/2}$

with prob p_e

$= 0$ other

Hence $Y_e = \frac{R}{R_{eff}(e)} \cdot \mathcal{L}_G^{+/2} \mathcal{L}_e \mathcal{L}_G^{+/2}$

$\succeq 0$.

Hence $\|Y_e\| \leq \frac{R}{R_{eff}(e)} \cdot R_{eff}(e)$

$\leq R$.

We apply Matrix Chernoff to the sum $\sum_e \mathcal{L}_G^{+/2} Y_e \mathcal{L}_G^{+/2}$

$\mathbb{E}[\quad] = \Pi$.

$\lambda_{min}(\Pi) = \lambda_{max}(\Pi) = 1$

viewed as a symmetric operator over $n-1$ dim space.

orthogonal to I.

$$\| L_G^{+/2} Y_e L_G^{+/2} \| \leq R$$

$\Rightarrow$ with prob: $1 - 2n e^{-\frac{\varepsilon^2}{3R}}$

we have

$$(1-\varepsilon) L_G^{+/2} L_G L_G^{+/2} \preceq L_G^{+/2} L_H L^{+/2} \preceq (1+\varepsilon) L_G^{+/2} L_G L^{+/2}$$

If $R \geq c \frac{\log n}{\varepsilon^2}$ for sufficiently large c then done. whp.

$\square$.

Proofs of Lemmas

Lemma 1: $\mathcal{L}_e \preceq R_{eff}(e) \cdot \mathcal{L}_G$.

Proof: Want to prove

Consider $\bar{v}^T \mathcal{L}_e \bar{v} = (v_i - v_j)^2$

where $e = (i,j)$

Wlg assume $v_i = 1$ and $v_j = 0$.

in which case $\bar{v}^T \mathcal{L}_e \bar{v} = 1$.

What is $\bar{v}^T \mathcal{L}_G \bar{v}$?

$$\bar{v}^T \mathcal{L}_G \bar{v} \geq \min_{v_i = 1,\ v_j = 0} \sum_{a \sim b} (v_a - v_b)^2$$

$$\geq C_{eff}(ij) = \frac{1}{R_{eff}(ij)}$$

(volltäns).

$$\Rightarrow \mathcal{L}_e \preceq R_{eff}(e)\, \mathcal{L}_G$$

$\square$.

Above is a sketch. See notes.

<u>Lemma 2</u>: $\sum_e R_{eff}(e) = n-1$

<u>Proof</u>: $R_{eff}(s,t) = (\chi_s - \chi_t)^T \mathcal{L}_G^{\dagger} (\chi_s - \chi_t)$

Recall.

$$\sum_{u \sim v} R_{eff}(u,v) = \sum_{u \sim v} (\chi_u - \chi_v)^T \mathcal{L}_G^{\dagger} (\chi_u - \chi_v)$$

[Simplicity assume $w_e = 1$, can create parallel edges if necessary.]

Recall $\text{Trace}(A) = \sum_{i=1} A_{ii}$

Sum of diagonal entries.

$\text{Trace}(A) =$ Sum of eigen values of A (even when A may have complex values)

Follows from fact that eigen values are solutions to $\det(A - \lambda I) = 0$

Coefficient of $\sum_i \lambda_i = \sum_{i=1}^{n} A_{ii}$.

$$\sum_{u,v \in E} R_{eff}(u,v) = \sum_{u,v \in E} (\chi_u - \chi_v)^T \mathcal{L}^+ (\chi_u - \chi_v)$$

$$= \sum_{u,v \in E} \text{Trace}\left(\quad \right)$$

(by cyclic property of trace $\text{Trace}(ABC) = \text{Trace}(CAB)$.

$$= \sum_{uv \in E} \text{Trace}\left((x_u - x_v)(x_u - x_v)^T \cdot \mathcal{L}^+ \right)$$

$$\text{Trace}(A(B+C)) = \text{Trace}(AB + AC).$$

$$= \text{Trace}\left(\left(\sum_{uv \in E} (x_u - x_v)(x_u - x_v)^T \right) \mathcal{L}^+ \right)$$

$$= \text{Trace}\left(\mathcal{L} \mathcal{L}^+ \right)$$

Recall $\mathcal{L} = \sum_{i=1}^{n} \lambda_i \bar{u}_i \bar{u}_i^T$

? eigen vectors of $\mathcal{L}$.

$\lambda_1 = 1$. orthonormal

$$\mathcal{L}^+ = \sum_{i=2}^{n} \frac{1}{\lambda_i} \bar{u}_i \bar{u}_i^T$$

$$\mathcal{L}\mathcal{L}^{+} = \sum_{i=2}^{n} u_i u_i^T$$

So essentially identity in the space orthogonal to $\bar{u}_1$.

So eigen values are $0, \underbrace{1, \ldots, 1}_{n-1}$

$$\Rightarrow \text{Trace}(\mathcal{L}\mathcal{L}^{+}) = n-1.$$

$\square$.