$G=(V,E)$ undir graph
$w: E \to \mathbb{R}_+$ edge weights.

Recall $H=(V,F)$ $w': F \to \mathbb{R}_+$ is an ε-approximate spectral sparsifier of G if

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

In previous lecture we saw a random sampling algorithm that shows that for any $\varepsilon > 0$ ε-approximate spectral sparsifiers with $O\left(\frac{n \log n}{\varepsilon^2}\right)$ edges exist.

Recall the algorithm:

We see

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

where $\mathcal{L}_{uv} = (\chi_u - \chi_v)(\chi_u - \chi_v)^T$

is a rank-1 matrix

corresponding to edge uv.

We obtained H by importance sampling:

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

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

$= 0$ otherwise.

And we chose $p_e = \frac{c \log n}{\varepsilon^2} R_{eff}(e)$

Since the "importance" or "influence" of $w_e L_e$ on L_G is $R_{eff}(e)$

captured by the two lemmas.

① $w_e L_e \preceq R_{eff}(e) L_G$

② $\sum_e R_{eff}(e) = n-1$.

the extra $\frac{\log n}{\varepsilon^2}$ in the probability

is standard when using

Chernoff-type concentration ineq.

and is often necessary if one uses

independent random sampling.

Two questions

① Can we derandomize?

② Is $\frac{n \log n}{\varepsilon^2}$ a tight bound?

One standard derandomization strategy when using Chernoff type inequalities is to use pessimistic estimators. In the context of Chernoff bounds this yields a variant of the well-known Multi-plicative

Weight Update method that
(MWU)
relies on exponential potential
function. In the context of
Matrix Chernoff bound one
needs to use the Matrix MWU
which is technical but most
likely yields a deterministic
algorithm but since it essentially
relies on the analysis of the
randomized algorithm and
will also lead to a sparsifier
with $\Omega\left(\frac{n \log n}{\varepsilon^2}\right)$ edges in the

Worst case.

[Batson-Spielman-Srivastava] used another potential function and an elegant analysis to obtain a deterministic algorithm that yields a sparsifier with $O(\frac{n}{\varepsilon^2})$ edges which is optimal.

The initial algorithm was slow (n^4) but subsequent work has yielded faster deterministic and randomized algorithms.

Recall that $\mathcal{L}_H \approx_\varepsilon \mathcal{L}_G$ iff all eigenvalues of $\mathcal{L}_H$ are roughly same as those of $\mathcal{L}_G$.

Even in the sampling algorithm we found it easier to work with $\mathcal{L}_G^{+/2} \mathcal{L}_G \mathcal{L}_G^{+/2} = \Pi$ which is identity on space $\perp$ to $\mathbb{1}$.

The reason for this is that the min and max eigen values of I are 1 and hence requiring that λ_{min} and λ_{max}

are well-approximated ensures that all eigen values are approximated.

To this end this is the main theorem of [BSS]

Main Theorem

Let $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m \in \mathbb{R}^n$ s.t

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

($\bar{v}_1, \ldots, \bar{v}_m$ are in isotropic position).

Then for any $\varepsilon > 0$ $\exists$ a set $S \subseteq [m]$ and c_i, $i \in S$ s.t

① $|S| = O\left(\frac{n}{\varepsilon^2}\right)$

(2) $(1-\varepsilon) I \preceq \sum_{i \in S} c_i v_i v_i^T \preceq (1+\varepsilon)^2 I$

Exercise: Use above theorem to derive $O\left(\frac{n}{\varepsilon^2}\right)$ sized ε-approximate spectral sparsifier of a n-vertex graph.

Hint:

$$L_G^{+/2} L_G L_G^{+/2} = \sum_{e \in E} L_G^{+/2} w_e L_e L_G^{+/2}$$

$$= \sum_{e \in E} v_i v_i^T$$

Proof of Main Theorem

The algorithmic idea is the following.

- Start with matrix $A = 0$ (zeros)
- For some # of steps do
    - Find an index $i \in [m]$ and a step size c and update $A \leftarrow A + c v_i v_i^T$

.

Basically it follows an iterative scheme where in each step

We add a small amount of some $\bar{v}_i v_i^T$ to current matrix A to get it close to I.

Recall we want $\lambda_{min}(A) \geq (1-\varepsilon)$

and $\lambda_{max}(A) \leq 1+\varepsilon$

at the end.

Of course we can simply add all of the v_i's with step size 1 but then we won't sparsify.

Idea is to use $\boxed{O\left(\frac{n}{\varepsilon^2}\right)}$ iterations so we get a sparsifier

Some mathematical background

$$\text{Trace}(A) = \sum_{i=1}^{n} A_{ii}$$ sum of diagonal entries of a $n \times n$ matrix.

① $Tr(A) = \sum_i \lambda_i$ of eigen values

Why? λ's are roots of characteristic poly $\det(A - \lambda I) = 0$.

② $Tr(A+B) = Tr(A+B)$ clear.

③ Less clear: cyclic property
$Tr(AB) = Tr(BA)$ whenever
AB is square matrix
(A, B may not be square!).

Easy to verify.

<u>**Lemma:**</u> $\displaystyle\sum_{i=1}^{m} v_i^T M \bar{v}_i = Tr\left(\left(\sum_{i=1}^{m} v_i v_i^T\right) M\right)$

and hence if $\displaystyle\sum_{i=1}^{m} v_i v_i^T = I$ then

$$\sum_{i=1}^{m} \bar{v}_i^T M \bar{v}_i = Tr(M).$$

<u>**Proof:**</u> $\displaystyle\sum_{i=1}^{m} v_i^T M \bar{v}_i = \sum_{i=1}^{m} Tr\left(v_i^T M \bar{v}_i\right)$

\#   \#

$$= \sum_{i=1}^{m} Tr\left(v_i v_i^T M\right) \quad \text{by cyclic prop of Trace}$$

$$= Tr\left(\left(\sum_{i=1} v_i v_i^T\right) M\right)$$

$$= \mathrm{Tr}(IM) = \mathrm{Tr}(M).$$

D.

Sherman-Morrison-Woodbury Formula

Theorem: Let A be a symmetric non-singular matrix and let c be a real # and $\bar{v}$ be a vector.

Then

$$(A - c v v^T)^{-1} = A^{-1} + c \frac{A^{-1} \bar{v} \bar{v}^T A^{-1}}{1 - c \bar{v}^T A^{-1} \bar{v}}.$$

Proof: check by multiplying. D.

or view A as $\sum_i \lambda_i \bar{u}_i \bar{u}_i^T$

n eigenvalu...

Barrier functions

We define two barrier functions. "upper" and "lower".

Let A be a $n \times n$ symmetric matrix with eigen values

$$\lambda_1 \leq \lambda_2 \cdots \leq \lambda_n .$$

For $u > \lambda_n$ we let

$$\Phi_u(A) = \sum_{i=1}^{n} \frac{1}{u-\lambda_i}$$

$$= \operatorname{Tr}\left((uI - A)^{-1}\right).$$

For $l < \lambda_1$ we let

$$\Phi_l(A) = \sum_{i=1}^{n} \frac{1}{\lambda_i - l}$$

$$= \mathrm{Tr}\left((A - lI)^{-1}\right).$$

As $u \to \infty$ $\Phi_u(A) \to 0$

and as $u \searrow \lambda_n$ $\Phi_u(A) \to \infty$.

As $l \to -\infty$ $\Phi_l(A) \to 0$

$l \to \lambda_1$ $\Phi_l(A) \to \infty$.

We need to understand how the barrier functions changes as u and l change.

Claim: $\Phi_{u+\delta}(A) < \Phi_u(A)$ for $\delta > 0$

Claim: $\lambda_n < u - \frac{1}{\Phi_u(A)}$

Claim: $l + \frac{1}{\Phi_l(A)} \leq \lambda_1$.

More technical

Claim: Let $l < \lambda_1$ and $\delta < \frac{1}{\Phi_l(A)}$

$$\Phi_{l+\delta}(A) \leq \frac{1}{\frac{1}{\Phi_l(A)} - \delta}$$

—

Need basic algebra. See Spielman notes.

How do we use barrier function?

Initially $A = 0$. $\lambda_1 = \lambda_2 \cdots = \lambda_n = 0$.

We set $l_0 = -n$

and $u_0 = n$

So that $\Phi_{u_0}(A) = \Phi_{l_0}(A) = 1$.

Our goal is eventually to make A have the property that

$$\frac{\lambda_{max}(A)}{\lambda_{min}(A)} \approx O(1).$$

ideally $(1-\varepsilon)$.

<u>Plan</u>: Show that in each iteration we can find scalar c and v_i and δ_u and δ_ℓ s.t.

$$\Phi^{u+\delta_u}(A + c\,\bar{v}_i\bar{v}_i^T) \le \Phi^u(A) \le 1$$

$$\Phi_{\ell+\delta_\ell}(A + k\,\bar{v}_i\bar{v}_i^T) \le \Phi_\ell(A) \le 1$$

We want δ_u and δ_ℓ to be fixed constants. Suppose we can do that. We run algorithm for $\Omega(n)$ steps.

For concreteness say $\delta_u = 2$ $\delta_\ell = \frac{1}{3}$.

and we run for $6n$ steps.

At end $u = 13n$ and $\Phi_{13n}(A) \le 1$.

$$\Rightarrow \lambda_n \le u - \frac{1}{\Phi_u(A)} \le 13n.$$

$$l \geq -n + 6n \times \frac{1}{3} \geq n.$$

$$\lambda_1 \geq l + \frac{1}{\Phi_l(A)} \geq n.$$

$$\Rightarrow \quad \frac{\lambda_{max}(A)}{\lambda_{min}(A)} \le 13.$$

We need to scale A by $6n$ to normalize.

How to show existence of c
δ_u, δ_ℓ. Note that for given
c, v_i, δ_u, δ_ℓ we can
compute all desired quantities
and check if conditions hold.
So issue is proving existence.

<u>Updating barrier function</u>

We want to change
A to $A + c v_i v_i^T$
rank-1 update.

How does Φ_u change?

$$\phi_u(A) = \mathrm{Tr}\left((uI - A)^{-1}\right)$$

$$\phi_u(A + c v v^T) = \mathrm{Tr}\left[(uI - A - c v v^T)^{-1}\right]$$

We use SMW formula

$$= \mathrm{Tr}\left[(uI - A)^{-1} + c\,\frac{(uI-A)^{-1} v v^T (uI-A)^{-1}}{1 - c v^T (uI-A)^{-1} v}\right]$$

$$= \phi_u(A) + \mathrm{Tr}(\quad)$$

$$= \phi_u(A) + \frac{c\, v^T (uI-A)^{-2} v}{1 - c v^T (uI-A)^{-1} v}$$

because $\mathrm{Tr}(XY) = \mathrm{Tr}(YX)$.

$$= \Phi_u(A) + \frac{c\, \bar{v}^T (uI - A)^{-2} \bar{v}}{1 - c\, \bar{v}^T (uI - A)^{-1} \bar{v}}$$

+ve since $c \geq 0$
and PSD of $(uI - A)^{-1}$.

So adding $c v v^T$ increases
Φ_u. (intuitively λ's
increase $\Rightarrow \sum \frac{1}{u - \lambda_i} \nearrow$.)

To keep $\Phi_u(A + c\bar{v}\bar{v}^T) \leq 1$
we increase u to $u' = u + \delta_u$.

<u>Q:</u> how much should we
increase u s.t

$$\Phi_{u+\delta_u}(A + c\bar{v}\bar{v}^T) \le \Phi_u(A).$$

From above derivation

$$\Phi_{u'}(A + \bar{v}\bar{v}^T)$$

$$= \Phi_{u'}(A) + \frac{c\bar{v}^T(u'I - A)^{-2}\bar{v}}{1 - c\bar{v}^T(u'I - A)^{-1}\bar{v}}$$

Thus to ensure

$$\Phi_{u'}(A + c\bar{v}\bar{v}^T) \le \Phi_u(A)$$

it suffices to have

$$\Phi_u(A) - \Phi_{u'}(A) \geq \frac{c\, v^T (u' I - A)^{-2} \bar{v}^T}{1 - c\, \bar{v}^T (u' I - A)^{-1} \bar{v}}$$

$$\Rightarrow \frac{1}{c} \geq \frac{v^T (u' I - A)^{-2} \bar{v}}{\Phi_u(A) - \Phi_{u'}(A)} + \bar{v}^T (u' I - A)^{-1} \bar{v}$$

Let $$U_A = \frac{(u' I - A)^{-2}}{\Phi_u(A) - \Phi_{u'}(A)} + (u' I - A)^{-1}$$

<u>Lemma:</u> Fix u, δu and $\bar{v}$

Then $\max c$ s.t. $\Phi_{u+\delta u}(A + c\, v\, v^T) \leq \Phi_u(A)$ is given by.

$$\frac{1}{c} \geq \bar{v}^T U_A \bar{v} .$$

Very clean dependence on A u and δu and v is "outside".

What happens when we increase l to δ_l? $l' = l + \delta_l$.

$\Phi_{l'}(A)$ increases. but when we add $c\bar{v}\bar{v}^T$ to A

$\Phi_l(A + c v \bar{v}^T)$ <u>decreases</u>. (since λ's intuitively increase).

How much? Etc.

Want $\phi_{l'}(A + t v \bar{v}^T) \leq \phi_l(A)$.

Lemma:

Let $L_A = \dfrac{(A - l'I)^{-2}}{\phi_{l'}(A) - \phi_l(A)} - (A - l'I)^{-1}$

If $\frac{1}{c} \leq \bar{v}^T L_A \bar{v}$ then

$\phi_{l'}(A + c v \bar{v}^T) \leq \phi_l(A)$.

Inductive argument.

Given u, A, δ_u, δ_ℓ and $\bar{v}$ we have seen that

$$\Phi_{u+\delta_u}(A + c\, v \bar{v}^t) \leq \Phi_u(A)$$

if $\frac{1}{c} \geq \bar{v}^T U_A \bar{v}$

and $$\Phi_{\ell+\delta_\ell}(A + c\, v \bar{v}^T) \leq \Phi_\ell(A)$$

if $\frac{1}{c} \leq \bar{v}^T L_A \bar{v}$.

Recall $\sum_{i=1}^{m} v_i v_i^T = \underline{I}$

Fix δ_u and δ_ℓ.

Want to find scalar $c>0$ and $\bar{v}_i$ s.t. we can update A to $A + c\, v_i v_i^T$ without violating barriers.

<u>Lemma</u>: $\displaystyle\sum_{i=1}^{m} v_i^T U_A \bar{v}_i \leq \frac{1}{\delta_u} + \Phi_u(A).$

<u>Lemma</u>: $\displaystyle\sum v_i^T L_A \bar{v}_i$

$$\geq \frac{1}{\delta_\ell} - \frac{1}{\frac{1}{\Phi_\ell(A)} - \delta_\ell}$$

$$\geq \frac{1}{\delta_\ell} - \Phi_\ell(A).$$

Thus if we have that

$$\frac{1}{\delta_u} + \Phi_u(A) \leq \frac{1}{\delta_l} - \Phi_l(A).$$

$\Rightarrow \exists\ v_i$ s.t $\bar{v}_i^T U_A \bar{v}_i \leq \bar{v}_i^T L_A \bar{v}_i$

and

Then we can find c and $\bar{v}_i$ to add to A..

Parameter settings

d fixed constant $d = \frac{1}{\varepsilon^2}$.

$l_0 = -\sqrt{d}\, n.$ $\quad \phi_{l_0}(0) = \frac{1}{\sqrt{d}}$

$u_0 = \frac{d+\sqrt{d}}{\sqrt{d}-1} n \quad \phi_{u_0}(0) = \frac{\sqrt{d}-1}{d+\sqrt{d}}.$

$\delta_l = 1$, $\delta_u = \frac{\sqrt{d}+1}{\sqrt{d}-1}.$

Initially.

$$\frac{1}{\delta_u} + \phi_u(A) = \frac{\sqrt{d}-1}{\sqrt{d}+1} + \frac{\sqrt{d}-1}{d+\sqrt{d}}$$

$$= \frac{\sqrt{d}-1}{\sqrt{d}+1}\left(1+\frac{1}{\sqrt{d}}\right)$$

$$= \frac{\sqrt{d}-1}{\sqrt{d}} = 1 - \frac{1}{\sqrt{d}}$$

$$\frac{1}{\delta_\ell} - \frac{1}{\phi_\ell(0) - \delta_\ell} = 1 - \frac{1}{\sqrt{d} - 1}$$

Hence we satisfy the condition.

Now run algorithm for dn iterations. Let A be resul. matrix.

$$\frac{\lambda_{max}(A)}{\lambda_{min}(A)} \leq \frac{u_0 + dn\,\delta_u}{l_0 + dn\,\delta_\ell}$$

$$\leq \frac{d + 2\sqrt{d} + 1}{d - 2\sqrt{d} + 1}.$$

for $d = \frac{1}{\varepsilon^2}$ above is $\dfrac{\frac{1}{\varepsilon^2} + 2\frac{1}{\varepsilon} + 1}{\frac{1}{\varepsilon^2} - \frac{2}{\varepsilon} + 1}$

$$\approx \frac{1+2\varepsilon+\varepsilon^2}{1-2\varepsilon+\varepsilon^2}.$$

Proofs of Lemmas

Lemma: $\sum_{i=1}^{m} v_i^T U_A \bar{v}_i \leq \frac{1}{\delta_u} + \Phi_u(A).$

Proof: Recall.

$$\sum_{i=1}^{m} v_i^T U_A \bar{v}_i = \mathrm{Tr}(U_A)$$

$$U_A = \frac{(u'I - A)^{-2}}{\Phi_u(A) - \Phi_{u'}(A)} + \underset{\downarrow}{(u'I - A)^{-1}}$$

$$\mathrm{Tr}\left((u'I - A)^{-1}\right) = \Phi_{u'}(A) \leq \Phi_u(A).$$

$$\mathrm{Tr}\left((u'I - A)^{-2}\right) \ ?$$

$$\phi_u(A) = \sum_{i=1}^{n} \frac{1}{u-\lambda_i}$$

$$\frac{d}{du}\phi_u(A) = -\sum \frac{1}{(u-\lambda_i)^2}$$

$$= -\mathrm{Tr}\left((uI-A)^{-2}\right).$$

Since $\phi_u(A)$ is convex in u.

we have.

$$\phi_u(A) - \phi_{u+\delta_u}(A) \geq -\delta_u \phi'_{u+\delta_u}$$

$$\geq \delta_u \mathrm{Tr}\left((u'I-A)^{-2}\right).$$

$$\Rightarrow \quad \frac{1}{\delta_u} \geq \frac{\mathrm{Tr}\left((u'I-A)^{-2}\right)}{\phi_u(A) - \phi_{u+\delta_u}(A)}.$$

$\square$.

Lemma: $$\sum_{i=1}^{m} v_i^T L_A \bar{v}_i \geq \frac{1}{\delta_l} - \phi_l(A).$$

Proof: Recall $$\sum_{i=1}^{m} v_i^T L_A \bar{v}_i = Tr(L_A).$$

$$L_A = \frac{(A - l'I)^{-2}}{\phi_{l'}(A) - \phi_l(A)} - (A - l'I)^{-1}$$

$$Tr(L_A) = Tr(\quad) + Tr(\quad)$$

$$\mathrm{Tr}\,(A - l'I)^{-1} = \Phi_{l'}(A) \leq \Phi_l(A).$$

$$\mathrm{Tr}\left((A - l'I)^{-2}\right).$$

$$\Phi_l(A) = \sum_{i=1}^{n} \frac{1}{\lambda_i - l}$$

$$\frac{d}{dl}\Phi_l(A) = \sum_{i=1}^{n} \frac{1}{(\lambda_i - l)^2}$$

$\Phi_l(A)$ is convex in l.

$$\Rightarrow \quad \frac{\Phi_{l+\delta_l}(A) - \Phi_l(A)}{\delta_l} \leq \Phi'_{l+\delta_l}(A)$$

$$\Rightarrow \quad \frac{\mathrm{Tr}\left((A - l'I)^{-2}\right)}{\Phi_{l'}(A) - \Phi_l(A)} \geq \frac{1}{\delta_l}.$$