Lecture 8: Sparsest Cut and the Cheeger bound

Recall: sparsity of $S \subseteq V$ $\quad \phi(S) = \dfrac{cap(S,\bar{S})}{(|S|\,|\bar{S}|)} = \dfrac{|E(S,\bar{S})|}{(|S|\,|\bar{S}|)}$ if capacities $=1$ for simplicity. (today).

want sparsest cut S^* st minimizes $\phi(S)$.

Imp. $\boxed{\text{For today: focus on capacities} = 1 \text{ and regular graphs (all vertices have degree} = d)}$

one way to write sparsity is (very similar to max-cut lecture

$$\min_{x \in \{-1,1\}^n} \frac{\sum'_{(i,j)\in E} (x_i - x_j)^2}{\sum'_{i<j \in V} (x_i - x_j)^2} = \frac{2 \cdot \sum'_{i,j \in E} (x_i - x_j)^2}{\sum'_{i,j \in V} (x_i - x_j)^2}$$

$i<j \in V$ ← ordered pairs $\qquad$ ↖ unordered pairs

suppose we relax this to be all $x \in \mathbb{R}^n$, then value can only fall. Now,

if $x \in \mathbb{R}^n$ then ① shift so that $x \perp \mathbb{1}$ (ie. $\sum' x_i = 0$). Since both num. & den. are differences, does not change. then

② Observe that $\displaystyle\sum'_{i,j \in V} (x_i - x_j)^2 = \Big(\sum'_{i \in V} x_i^2\Big)(2n) - \underbrace{\Big(\sum'_{i \in V} x_i\Big)\Big(\sum'_{j \in V} x_j\Big)}_{=0}$

$$= \Big(\sum'_{i \in V} x_i^2\Big)(2n).$$

$$\Rightarrow \text{relaxation} = \min_{x \perp \mathbb{1}} \left(\frac{2\sum'_{(i,j)\in E} (x_i - x_j)^2}{\sum'_{i \in V} x_i^2} \right) \frac{1}{(2n)} \le \phi(G).$$

$$= \min_{x \perp \mathbb{1}} \underbrace{\frac{x^T L_G x}{x^T x}} \cdot \frac{1}{(n)} .$$

where L_G = Laplacian matrix of a graph
$= dI -$ Adjacency
for d-regular graphs.

Interestingly, $\mathbb{1}$ is the top eigenvector of L_G, so this is the second eigenvalue this means we get that $\boxed{\phi(S) \geq \lambda_2(L_G) \cdot \frac{1}{2n}}$.

The Cheeger inequality bounds $\phi(S)$ above by a function of λ_2 as well, giving a good characterization of $\phi(S)$ for some families of graphs (esp. expanders). (But it can be poor for other families as we'll soon see).

— x —

N.b: there are many variants of sparsity.

- Sparsity,
- edge expansion $\approx$ (sparsity) $\times n$.
- conductance $= \dfrac{E(S,\bar{S})}{\min\left(\sum_{v \in S} \text{degree}(v), \sum_{v \in \bar{S}} \text{degree}(v)\right)}$ etc.

And for irregular graphs we may need to look at the normalized Laplacian which is $I - D^{-1/2} A D^{-1/2}$.

There are connections, but when you look up "Cheeger" you should make sure the version you have is relating the quantities you have in hand. Else you may need to do translations.

— x —

Today: Cheeger bounds, for d-regular graphs

$$\boxed{\frac{\lambda_2(L_G)}{2n} \leq \phi(S) \leq \frac{\sqrt{\lambda_2(L_G) \cdot d}}{\theta(n)}}$$

Already saw the lower bound, see the upper bound soon!

(Also $\exists$ graphs where bounds are tight.)

Upper bound (Algorithmic)

Find an eigenvector corresponding to λ_2. Call it $x \in \mathbb{R}^n$

Sort ~~vectors~~ vertices so that $x_1 \leq x_2 \leq x_3 \cdots \leq x_n$.

Let $S_i = \{1, 2, \ldots, i\}$ in this order.

Output the set S_i with smallest sparsity $\phi(S_i)$.

Notation: $R(x) = \dfrac{\sum_{ij \in E}(x_i - x_j)^2}{\sum_i x_i^2} = \dfrac{x^T L_G x}{x^T x}$ (Rayleigh quotient)

Proof in 2 lemmas.

First take $x \rightarrow$ break into its positive and negative parts (almost).

get vectors y and z.

Lemma 1: Given x, get y, z st both non-negative and.

$$R(y), R(z) \leq \cancel{8} \cdot R(x).$$

and have disjoint supports.

Say y has $\leq n/2$ sized support (disjointness $\Rightarrow$ at least one of y, z does.)

Lemma 2: For non-negative vector x (obtained from Lemma 1), $\exists S \subseteq \text{support}(x)$ st.

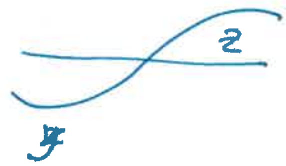
$$\phi(S) \leq \sqrt{}$$

(Moreover this set will be one of the sets considered in our algo above)

Lemma 1:

Try #1: $y = \max(x, 0)$ component wise max
$z = -\min(x, 0)$ —— " —— min

$x = (-0.7, -0.3, 0, 0.5, 1.2$
$y = (0, 0, 0, 0.5, 1.2$
$z = (0.7, 0.3, 0, 0, 0)$



then

(*) $\qquad \sum_{i}(y_i - y_j)^2, \ \sum_{i}(z_i - z_j)^2 \leq \sum_{i}(x_i - x_j)^2.$

if only $\sum_i z_i^2, \sum_i y_i^2 \geqslant \Omega(1)$, but maybe not ?

eg. $x = (1, \frac{-1}{n-1}, \frac{-1}{n-1}, \dots \frac{-1}{n-1})$ so $x \perp \mathbb{1}$
but $\|z\| = \frac{1}{\sqrt{n}}$ ☹

So shift x. Sps $\|z\|^2 < 1/4$ (and $\|y\|^2 \geqslant 3/4$).

Try #2: $x' = x - \frac{1}{2\sqrt{n}}$ $\qquad y' = \max(x', 0)$
$z' = -\min(x', 0)$.

"moves mass to help z!"

(*) still holds for y', z'

but now. $\|x'\|^2 = \|x\|^2 + \|\frac{1}{2n}\mathbb{1}\|^2 = \|x\|^2 + \frac{1}{4} = 5/4.$
↑ $x \perp \mathbb{1}$ (Pythagoras)

- $\|y'\|^2 \leq \|x\|^2 \leq 1 \quad \Rightarrow \quad \|z'\|^2 = \frac{5}{4} - \|y'\|^2 \geqslant \frac{1}{4}.$
↑ because negative shift $\qquad$ ↑ $y' \perp z'$

- $\frac{3}{4} \leq \|y\|^2 = \sum_i y_i^2 \leq \sum_i \left(y_i' + \frac{1}{2\sqrt{n}}\right)^2 \leq 2\left(\sum_i \left(y_i'^2 + \frac{1}{4n}\right)\right) = 2\|y'\|^2 + \frac{1}{2}$

$\Rightarrow \|y'\|^2 \geqslant \frac{1}{8}.$

☺

So $x \to y', z'$ disjoint supports

and $R(y'), R(z') \geqslant 8R(x).$

Also, note that all we've done is shifted and truncated x, so level sets in y, z are both level sets in x.

Now focus on y say. (where ~~to~~ $\text{supp}(y) \le n/2$).

<u>Lemma 2</u>: $\exists$ level set in support of y s.t.

$$\phi(S) \le \sqrt{2d\,R(y)}/n$$

Remember: $y \ge 0$, $\text{Supp}(y) \le n/2$

<u>Pf</u>: Suppose $y_i \in [0,1]$ for all i (scaling does not change $R(y)$).

Choose α uniformly in $[0,1]$.

$$S_\alpha := \{ i \mid y_i^2 \geq \alpha \}.$$ random set.

$$\mathbb{E}[|E(S_\alpha, \bar{S}_\alpha)|] = \sum_{(i,j)\in E} \Pr_\alpha[\text{one of } y_i^2, y_j^2 < \alpha,\ \text{other} \geq \alpha]$$

$$= \sum_{ij\in E} |y_i^2 - y_j^2| = \sum_{ij\in E} |(y_i - y_j)(y_i + y_j)|$$

(Cauchy Schwarz) $$\leq \sqrt{\sum_{ij\in E} (y_i - y_j)^2} \cdot \sqrt{\sum_{ij\in E} \underbrace{(y_i + y_j)^2}_{\le 2(y_i^2 + y_j^2)}}$$

$$\leq \sqrt{\sum_{ij\in E} (y_i - y_j)^2} \cdot \sqrt{2d \sum_i y_i^2}$$ because graph is d-regular

$$= \sqrt{2d \cdot R(y)} \cdot \left(\sum_i y_i^2\right).$$ because $R(y) = \dfrac{\sum_{ij\in E} (y_i - y_j)^2}{\sum_i y_i^2}$

equal

$$\mathbb{E}[|S_\alpha|] = \sum_i \Pr_\alpha[y_i^2 > \alpha] = \sum_i y_i^2$$

$$\Rightarrow \mathbb{E}\left[|E(S_\alpha, \bar{S}_\alpha)| - \sqrt{2d\,R(y)} \cdot |S_\alpha|\right] \leq 0$$

$\Rightarrow$ by the probabilistic method, $\exists \alpha \in [0,1]$ s.t.

$$\frac{|E(S_\alpha, \bar{S}_\alpha)|}{|S_\alpha|} \leq \sqrt{2d\, R(y)}$$

and so

$$\frac{|E(S_\alpha, \bar{S}_\alpha)|}{|S_\alpha||\bar{S}_\alpha|} \leq \frac{|E(S_\alpha, \bar{S}_\alpha)|}{|S_\alpha| \cdot n/2} \leq \frac{2}{n} \cdot \sqrt{2d\, R(y)}.$$

since $|S_\alpha| \leq n/2$
$\Rightarrow |\bar{S}_\alpha| \geq n/2$

from above.

So the best level set has at most this sparsity

Finishes proof of Lem 2
$\Rightarrow$ of Cheeger!

<u>To recap</u>: Cheeger says

$$\boxed{\Omega\left(\frac{\lambda_2}{n}\right) \leq \phi(G) \leq O\left(\frac{\sqrt{\lambda_2 d}}{n}\right)}$$

~~In fact,~~ So when $d = \text{constant}$, $\lambda_2 = O(1)$ then correct up to constants.
we have a way to certify a constant degree expander.

In general $\lambda_2 \in [0, d]$ so when $\lambda_2 = \Omega(d)$ we're pretty tight!

spectral definition of expander.

<u>Tightness of bounds:</u>

Hypercube $\{0,1\}^d$: $n = 2^d$

- $\phi(G) = \text{sparsity of cut } (\{x_1 = 0\}, \{x_1 = 1\}) = \dfrac{2^{d-1}}{2^{d-1} \cdot 2^{d-1}} = \dfrac{2}{n}$

- $\lambda_2(L_G) = 2$

So the lower bound of Cheeger is tight here.

Path or Cycle:

- $\phi(G) = \phi(\text{any balanced cut}) = \frac{\Theta(1)}{n^2}$
- $\lambda_2(L_G) = \Theta(1/n^2)$ so. $\sqrt{\frac{\lambda_2 \cdot d}{n}} = \Theta\left(\frac{1}{n^2}\right)$.

——— x ———

- Can generalize to non-regular graphs, but there ~~we~~ it is convenient to use "conductance".

$$vol(S) = \sum_{i \in S} degree(i) / 2m \quad = \text{"fraction of edges hitting } S\text{"}$$

$$conductance(G) = \min_{S: vol(S) \le 1/2} \frac{|E(S,\bar{S})|}{vol(S)}$$

And then $\mathcal{L}(G) = I - D^{-1/2} A D^{-1/2}$ is normalized Laplacian

and Cheeger says

(*) $\boxed{\Theta(\lambda_2) \le \text{conductance} \le \Theta(\sqrt{\lambda_2}).}$

λ_2 of Normalized Laplacian

——— x ———

Sanity Check:

for d-regular graphs, $conductance(S) = \frac{|E(S,\bar{S})|}{vol(S)} = \frac{|E(S,\bar{S})|}{d.|S|}$

$$\approx \frac{n}{d} \cdot \phi(S)$$

~~so~~ and $\mathcal{L}(G) = \frac{1}{d} \cdot L_G$ so $\lambda_2(\mathcal{L}) = \frac{1}{d}\lambda_2(L_G)$

$\Rightarrow$ (*) says $\Theta\left(\frac{\lambda_2(L_G)}{d}\right) \le \frac{n}{d}\phi(S) \le \Theta\left(\sqrt{\frac{\lambda_2(L_G)}{d}}\right)$ which is what we proved!

$\lambda_2(L_G)$ as a notion of connectivity

Recall: if Graph has k connected components, then

$$\lambda_1(L_G) = \lambda_2 = \ldots = \lambda_k = 0.$$

So $\lambda_2 = 0$ means graph is disconnected.

- λ_2 "close to 0" means graph has a sparse cut (by cheeger).
- λ_2 "far from 0" means no ~~all~~ sparse cuts (by other side of cheeger).

  - A graph is a "spectral expander" if
    $$\lambda_2(L_G) \geq \Omega(d).$$

    So there's a huge "spectral gap" $\lambda = 0 \longleftrightarrow \lambda_2 = \Omega(d)$. $\ldots \lambda_n \leq \Theta(d)$

—— x ——

Algorithmic Aspects of Cheeger

Upper bound proof Algorithmic.

In fact shows that best level cut has sparsity $\leq \sqrt{\frac{\Theta(d)}{n} R(x)}$

- So we don't really need to find the exact 2nd eigenvalue, as long as we find x with small Rayleigh coefficient, we're good.

- Commonly used in vision applications (this 2nd ev called Fiedler vector)
- Results of Kwok et al: spectral partitioning good even when λ_k large for some later λ_k. $\phi(S) \leq O(k \cdot \lambda_2/\sqrt{\lambda_k}) \cdot 1/n$. See LapChi's notes.

Takeaways:

- Cheeger shows that $\lambda_2(L_G)$ is a good approx to $\phi(G)$ when the expansion of G is large.
- Leighton Rao shows that the LP integrality gaps arose in exactly these expander-like examples.

(Q1) Can we combine the two techniques?
Using SDPs (which capture spectral ideas + ~~[illegible]~~ LPs)?
(YES)

(Q2) This idea of using the λ_2 to get an algorithm for sparsest cut is cool.
Can we use eigenvector based methods for Max-Cut too?
(Yes)

More on these in the next Lectures.

—x—