CS 270: Combinatorial Algorithms and Data Structures Spring 2023

Lecture 26 — April 20, 2023

Prof. Jelani Nelson *Scribe: Yang Han*

1 Overview

In the last lecture we introduced Spectral Graph Theory, and started a bit of Cheeger's inequality. In this lecture we will first prove Cheeger's inequality, and then mention other Spectral Graph Theory topics such as spectral sparsification and Laplacian linear system solving. (Note that the second part is actually not covered in lecture)

2 Cheeger's Inequality

2.1 Isoperimetric Ratio and Conductance

First we recap some information that we know about Laplacian from the last lecture.

Definition 2.1. The *Laplacian matrix* $L(G)$ is defined as $D(G)-A(G)$, where $D(G)$ is the diagonal matrix with $D_{u,u}$ the weighted degree of u, which is $d(u) = \sum_{e=(u,\cdot)\in E} w_e$, and $A(G)$ is the adjacency matrix.

We assume $\forall e, w_e \geq 0$.

Definition 2.2. The *isoperimetric ratio* $\theta_G(S)$ is defined to be

$$\theta_G(S) = \frac{w(\partial S)}{\min\{|S|, |V \setminus S|\}}$$

Definition 2.3. and the *conductance* $\Phi_G(S)$ is defined to be

$$\Phi_G(S) = \frac{w(\partial S)}{\min\{\mathrm{vol}(S), \mathrm{vol}(V \setminus S)\}}$$

where the boundary $\partial S = E \cap (S \times (V \setminus S))$ is the set of edges that leaves S, and the volume is $\mathrm{vol}(S) = \sum_{u\in S} d(u)$.

Then we could define

$$\theta(G) = \min_{S\subseteq V} \theta_G(S)$$

and

$$\Phi(G) = \min_{S\subseteq V} \Phi_G(S)$$

Both the isoperimetric ratio and the conductance could be used to capture a sparse cut, where only a small fraction of edges leave a large set of vertices. However, they get slightly different motions since the former uses the size of the sets and the latter uses the volume.

Either could be related to the second smallest eigenvalue $\lambda 2$. The isoperimetric ratio relates to the unnormalized Laplacian, while the conductance is related to the normalized Laplacian.

2.2 Normalized Laplacian

Definition 2.4. Given any PSD matrix M, the *Rayleigh quotients* are:

$$\min_{x\neq 0} \frac{x^\top M x}{x^\top x} = \lambda_1$$

$$\min_{x\neq 0, x\perp v_1} \frac{x^\top M x}{x^\top x} = \lambda_2$$

$$\cdots$$

$$\min_{x\neq 0, x\perp v_1, v_2, \ldots, v_{i-1}} \frac{x^\top M x}{x^\top x} = \lambda_i$$

$$\cdots$$

$$\max_{x\neq 0} \frac{x^\top M x}{x^\top x} = \lambda_n$$

Specifically for Laplacians,

$$\lambda_2 = \min_{x\neq 0, x\perp \mathbb{1}} \frac{x^\top L x}{x^\top x}$$

We choose $x = \mathbb{1}_S - \alpha\mathbb{1}$ with the appropriate α to make it orthogonal with $\mathbb{1}$. Since $\langle x, \mathbb{1}\rangle = |S| - \alpha n$,we let $\alpha = \frac{|S|}{n}$. Then for the Rayleigh quotient

$$R(x) = \frac{x^\top L x}{x^\top x}$$

Since the $\alpha\mathbb{1}$ part of x is in the kernel of L,

$$\begin{aligned} x^\top L x &= \mathbb{1}_S^\top L \mathbb{1}_S \\ &= \sum_{u,v\in V} w(u,v) * (x_u - x_v)^2 \\ &= w(\partial S) \end{aligned}$$

And $x^\top x$ is roughly the size of S. Therefore $R(x) = \frac{x^\top L x}{x^\top x}$ is roughly $\frac{w(\partial S)}{|S|}$, which is the isoperimetric ratio $\theta_G(S)$ with $|S| < |V \setminus S|$.

As for normalized Laplacian, we are actually relating $\Phi(G)$ with $\frac{y^\top L y}{y^\top D y}$, where D is the same diagonal matrix of weighted degree. Let $x = D^{\frac{1}{2}}x$, then

$$\frac{y^\top L y}{y^\top D y} = \frac{x^\top D^{-\frac{1}{2}} L D^{-\frac{1}{2}} x}{x^\top x}$$

which matches the form of Rayleigh quotient.

We define $N = D^{-\frac{1}{2}} L D^{-\frac{1}{2}}$, and let $\gamma_1, \gamma_2, \cdots, \gamma_n$ to be the eigenvalues of N. Since

$$\begin{aligned} D^{-\frac{1}{2}} L D^{-\frac{1}{2}} \sqrt{\vec{d}} &= D^{-\frac{1}{2}} L \mathbb{1} \\ &= D^{-\frac{1}{2}} \times 0 \\ &= 0 \\ &= 0 * \sqrt{\vec{d}} \end{aligned}$$

we know that $\gamma_1 = 0$ and $v_1 = \sqrt{\vec{d}}$. From $x \perp \sqrt{\vec{d}}$ we also know that $y \perp \vec{d}$.

2.3 Proof of Cheeger's Inequality

Cheeger's Inequality for N is

$$\frac{\gamma_2}{2} \leq \Phi \leq \sqrt{2\gamma_2}$$

which is what we will prove today.

Before we start with the easier side $\Phi \geq \frac{\gamma_2}{2}$, we start with proving a lemma:

Lemma 2.5. $\forall S \subsetneq V, S \neq \emptyset, \frac{w(\partial S)}{vol(S) vol(V \setminus S)} w(v) \geq \gamma_2$.

Proof. Define $y = \mathbb{1}_S - \sigma \mathbb{1}$ with $\sigma = \frac{\text{vol}(S)}{\text{vol}(V)}$.

- First we claim that our $y \perp \vec{d}$.

  Claim. $\vec{d}^\top y = 0$

  Proof.

$$\begin{aligned} \vec{d}^\top y &= \vec{d}^\top \mathbb{1}_S - \frac{\text{vol}(S)}{\text{vol}(V)} \vec{d}^\top \mathbb{1} \\ &= \text{vol}(S) - \text{vol}(S) \\ &= 0 \end{aligned}$$

□

- $y \top Ly = w(\partial S)$.
- As for $y^\top Dy$,

$$\begin{aligned} y^\top Dy &= \sum_{u \in S} d(u)(1-\sigma)^2 + \sum_{u \notin S} d(u)\sigma^2 \\ &= \text{vol}(S) - 2\sigma \text{vol}(S) + \sigma^2 \text{vol}(S) + \sigma^2 \text{vol}(V \setminus S) \\ &= \text{vol}(S) - 2\sigma \text{vol}(S) + \sigma^2 \text{vol}(V) \\ &= (1-\sigma)\text{vol}(S) \\ &= \frac{\text{vol}(S)\text{vol}(V \setminus S)}{\text{vol}(V)} \end{aligned}$$

- Therefore

$$\gamma_2 \leq \frac{y \top Ly}{y^\top Dy} = \frac{w(\partial S)}{\text{vol}(S)\text{vol}(V \setminus S)} w(v)$$

□

With this lemma we can prove the easy side of Cheeger.

Proof.

$$\begin{aligned} \gamma_2 &\leq \frac{w(\partial S)\text{vol}(V)}{\text{vol}(S)\text{vol}(V \setminus S)} \\ &= \frac{w(\partial S)\text{vol}(V)}{\min(A,B)\max(A,B)} \\ &\leq \frac{2w(\partial S)}{\min\{\text{vol}(S), \text{vol}(V \setminus S)\}} \end{aligned}$$

□

Now we come to the other side $\Phi \leq \sqrt{2\gamma_2}$.

Proof. We need to show that $\exists S \subsetneq V$ such that $\Phi_G(S) \leq \sqrt{2\gamma_2}$.

- Define $S_\tau = \{u : y_u \leq \tau\}$, where y is the minimizer of $\frac{y^\top Ly}{y^\top Dy}$. We view the vector as embedding of the vertices on the line, and we sort V so that $y_1 \leq y_2 \leq \cdots \leq y_n$. Then we define a family of cuts that $y_1, \ldots, y_i$ are in the cut and $y_{i+1}, \ldots, y_n$ are not. Note that since $\vec{d}$ is a nonnegative vector and y is orthogonal with it, there must be some y_is to be negative and others to be positive.

- For purpose of this proof, we want

$$\sum_{u:y-u<0} d(u) \leq \frac{1}{2}\mathrm{vol}(V)$$
$$\sum_{u:y-u>0} d(u) \leq \frac{1}{2}\mathrm{vol}(V)$$

  However this is not always possible, so we instead let j be the smallest coordinate that $\sum_{u=1}^{j} d(u) \geq \frac{1}{2}\mathrm{vol}(V)$, and define $z = y - y_j\mathbb{1}$. (Note that $z_j = 0$)

- **Claim.**
$$\frac{z^\top Lz}{z^\top Dz} \leq \frac{y^\top Ly}{y^\top Dy}$$

  Proof. Define $v_s = y + s\mathbb{1}$. We claim that $f(s) = v_s^\top Dv_s$ is minimized with $s = 0$ and prove by calculus (taking derivative). Then we know $z^\top Dz \geq y^\top Dy$, with $z^\top Lz = y^\top Ly$,we get $\frac{z^\top Lz}{z^\top Dz} \leq \frac{y^\top Ly}{y^\top Dy}$. □

- **Claim.** $\exists \tau$ so that $\Phi(S_\tau) \leq \sqrt{2\frac{z^\top Lz}{z^\top Dz}}$.

  Proof. Without lost of generality, $z_1^2 + z_n^2 = 1$.

  - We will define a distribution over τ so that

$$\mathop{\mathbb{E}}_{\tau}[w(\partial S_\tau)] \leq \sqrt{2\gamma_2}\mathop{\mathbb{E}}_{\tau}[\min\{\mathrm{vol}(S_\tau), \mathrm{vol}(V \setminus S_\tau)\}]$$

    With the linearity of expetation,

$$\mathbb{E}[w(\partial S_\tau) - \sqrt{2\gamma_2}\min\{\mathrm{vol}(S_\tau), \mathrm{vol}(V \setminus S_\tau)\}] \leq 0$$

    which indicates that there exists a τ such that

$$w(\partial S_\tau) - \sqrt{2\gamma_2}\min\{\mathrm{vol}(S_\tau), \mathrm{vol}(V \setminus S_\tau)\} \leq 0$$

    and this is the τ that we need.

- So how do we define the distribution? The distribution is supported on $[z_1, z_n]$, and we define a probability density function such that $\forall t \in [z_1, z_n], \phi(t) = 2|t|$. First we show that this is a valid PDF.

  Proof.

$$\int_{z_1}^{z_n} 2|t|dt = -\int_{z_1}^{0} 2tdt + \int_{0}^{z_n} dt = \int_{0}^{|z_1|} 2tdt + \int_{0}^{z_n} 2tdt = z_1^2 + z_n^2 = 1$$

  □

  Also we can show that

$$\mathbb{P}_{\tau}(\tau \in [a,b]) = |sgn(b)b^2 - sgn(a)a^2|$$

  where

$$sgn(x) = \begin{cases} 1, x > 0 \\ 0, x = 0 \\ -1, x < 0 \end{cases}$$

- Next we deal with the left side with a claim.

  Claim.

$$\mathbb{E}[w(\partial S)] \le \sum_{e \in E} w_e |z(a) - z(b)|(|z(a)| + |z(b)|)$$

  Proof.

$$\begin{aligned}
\mathbb{E}[w(\partial S)] &= \sum_{e=(a,b)\in E} w_e \, \mathbb{P}[e \in \partial S_\tau] \\
&= \sum_{e=(a,b)\in E} w_e \, \mathbb{P}(z(a) \le \tau \le z(b)) \\
&= \sum_{e=(a,b)\in E} w_e |sgn(z(b))z(b)^2 - sgn(z(a))z(a)^2| \\
&= \begin{cases} \sum_e w_e |z(b)^2 - z(a)^2|, sgn(a) = sgn(b) \\ \sum_e w_e z(b)^2 + z(a)^2, sgn(a) \ne sgn(b) \end{cases} \\
&\le \begin{cases} \sum_e w_e |(z(a) - z(b))(z(a) + z(b))|, sgn(a) = sgn(b) \\ \sum_e w_e (z(b) - z(a))^2, sgn(a) \ne sgn(b) \end{cases} \\
&\le \sum_e w_e |z(a) - z(b)|(|z(a)| + |z(b)|)
\end{aligned}$$

  □

- Finally we deal with the right side with another claim, this time without proof.

  Claim.

$$\mathbb{E}_{\tau}[\min\{\text{vol}(S_\tau), \text{vol}(V \setminus S_\tau)\}] = z^\top D z$$

Putting the two claims together,

$$\begin{aligned}
\mathbb{E}[w(\partial S)] &\leq \sum_e w_e |z(a) - z(b)|(|z(a)| + |z(b)|) \\
&\leq \sqrt{\sum_e w_e (z(a) - z(b))^2} \sqrt{\sum_e w_e (|z(a)| + |z(b)|)^2} \\
&\leq \sqrt{z^\top L z} \sqrt{2 \sum_e w_e (z(a)^2 + z(b)^2)} \\
&= \sqrt{z^\top L z} \sqrt{2 \sum_{u \in V} d(u) z(u)^2} \\
&= \sqrt{z^\top L z} \sqrt{2 z^\top D z} \\
&= \sqrt{\frac{z^\top L z}{z^\top D z}} \sqrt{2} z^\top D z \\
&\leq \sqrt{2\gamma_2} \min\{\mathrm{vol}(S_\tau), \mathrm{vol}(V \setminus S_\tau)\}
\end{aligned}$$

□

□