

## Lecture 11: Cheeger's Inequality &amp; Spectral Graph Partitioning

Instructor: *Alex Andoni*Scribes: *Patrick Chi, William Wang*

## 1 Definitions and Motivations

Recall from last lecture we define the Laplacian and its normalized form for a graph  $G = (V, E)$  as:

$$L = D - A$$

$$\hat{L} = D^{-1/2} L D^{-1/2}$$

with

$$0 = \mu_1 \leq \mu_2 \leq \dots \leq \mu_n \leq 2$$

where  $\mu_1, \dots, \mu_n$  are the sorted eigenvalues of  $\hat{L}$ .

**Definition 1.** A cut is defined as a separation of a graph  $G$  into two subsections  $S$  and  $\bar{S}$ .

**Definition 2.** A graph is disconnected if there are no edges across a cut. We define a boundary of  $S$  as:

$$\partial S = \{(i, j) \in E, i \in S, j \in \bar{S}\}$$

**Theorem 3.**  $\mu_2 = 0 \iff G$  is disconnected.

*Proof.* This can easily be shown by the fact that we break down the adjacency matrix into two separate parts. Each part is its own distinct graph, thus there will be at least one other eigenvalue equal to 0.  $\square$

### 1.1 Conductance

We want to develop a quantitative statement to show how close to disconnected a graph is. This motivates the idea of conductance.

**Definition 4.** We define the conductance of a cut  $S$  to be:

$$\phi(S) \triangleq \frac{|\partial S|}{\text{vol} S} \quad \text{where} \quad \text{vol}(S) \triangleq \sum_{i \in S} d_i$$

We define the conductance of a graph  $G$  to be

$$\phi(G) = \min_{S \subset V} \phi(S) \text{ s.t. } 0 < \text{vol}(S) \leq \frac{\text{vol}(V)}{2}$$

or in other words

$$\phi(G) = \min_{S: \text{vol}(S) \neq 0} \frac{|\partial S|}{\min\{\text{vol}(S), \text{vol}(\bar{S})\}}$$

Note that  $\phi(S) \neq \phi(\bar{S})$  as we divide by the total number of edges in  $S$ . We generally look at sets with less than  $\frac{1}{2}$  the edge volume for this reason.

Note we can also consider an alternate variation of conductance of  $S$  as the following:

$$\phi(S) = \frac{\# \text{ edges crossing } S \rightarrow \bar{S}}{2(\# \text{ internal edges}) + (\# \text{ edges } S \rightarrow \bar{S})} \leq 1$$

**Remark 5.** *Computing  $\phi(G)$  is NP-hard. Moreover, even computing to a constant factor approximation is NP-hard.*

## 2 Cheeger's Inequality

This motivates **Cheeger's Inequality**, which we define as follows:

**Theorem 6.**

$$\frac{\mu_2}{2} \leq \phi(G) \leq \sqrt{2\mu_2}$$

Cheeger's inequality allows us to bound the connectivity of a graph, and get an idea of how "connected" a graph is just from its Laplacian. The left-hand inequality is somewhat easier and more "intuitive" to prove, while the right-hand inequality involves a much more involved proof.

### 2.1 Proof of $\frac{\mu_2}{2} \leq \phi(G)$

Note that the motivation for this proof comes from the idea that if  $\mu_2$  is small we should be able to find a cut in  $G$  that has small conductance.

*Proof.* This is the "easier" of the two inequalities to prove.

Recall that

$$\mu_2 = \min_{x \neq 0, x \perp v_1} \frac{x^T \hat{L}x}{\|x\|^2} \qquad \phi(G) = \min_{S: 0 < \text{vol}(S) \leq \frac{\text{vol}(G)}{2}} \Phi(S)$$

Thus, it is sufficient to exhibit an  $x$  such that  $R(x) \leq 2\phi(S^*)$ , where  $S^* = \arg \min_S \Phi(S)$

**Fact 7.**

$$y^T Ly = \sum_{(i,j) \in E} (y_i - y_j)^2$$

Using this fact, we try to construct a vector  $x$  to map  $R(x)$  to  $\phi(S^*)$

$$\mu_2 = \min_{x \neq 0, x \perp v_1} R(x) = \min_{x \neq 0, x \perp v_1} \frac{x^T D^{-\frac{1}{2}} L D^{-\frac{1}{2}} x}{\|x\|^2}$$

Perform the substitution  $y = D^{-\frac{1}{2}} x$ , then we can minimise with respect to  $y$

$$\mu_2 = \min_y \frac{y^T Ly}{y^T Dy}$$

We now set  $\mathbf{y} = \mathbb{1}_{S^*}$ , where we define  $\mathbf{y}$  as:

$$y_i = \begin{cases} 1 & \text{if } i \in S^* \\ 0 & \text{otherwise} \end{cases}$$

Note now that the numerator in our expression for  $\mu_2$  becomes:

$$\mathbf{y}^T L \mathbf{y} = \sum_{(i,j) \in E} (y_i - y_j)^2 = |\partial S^*|$$

The second equality follows from the definition of  $y_i$ , where only an edge crossing from  $S$  to  $\bar{S}$  will contribute to the summation. Now that we have constructed a  $\mathbf{y}$  that relates the  $R(x)$  to  $|\partial S^*|$ , we need to prove the corresponding vector  $x$  is orthogonal to  $v_1$  (since clearly  $x \neq 0$ ).

We will modify  $\mathbf{y}$  such that  $x$  becomes perpendicular to  $v_1$ . Consider  $\mathbf{y}' \triangleq \mathbf{y} - \sigma \mathbb{1}_n$ , where  $\sigma$  is some constant. Note that we still have

$$(\mathbf{y}')^T L \mathbf{y}' = |\partial S^*|$$

Also, recall that  $v_1 = (\sqrt{d_1} \cdots \sqrt{d_n})^T$ . Our objective is now to set  $\sigma$  such that

$$\begin{aligned} v_1 \cdot D^{-\frac{1}{2}} \cdot \mathbf{y}' = 0 &\iff \mathbb{1} \cdot D \cdot \mathbf{y}' = 0 \\ &\iff \sum_{i \in [n]} d_i (y_i - \sigma) = 0 \\ &\iff \sum_{i \in [n]} d_i y_i = \sigma \sum_{i \in [n]} d_i \\ &\iff \text{vol}(S) = \sigma \cdot \text{vol}(V) \end{aligned}$$

Thus, if we set  $\sigma = \frac{\text{vol}(S)}{\text{vol}(V)}$ , we have  $x \perp v_1$  as desired. Also note that we always choose the smaller half of the graph for  $S$ , so we have  $\sigma \leq \frac{1}{2}$ .

Finally, we need to calculate  $(\mathbf{y}')^T D \mathbf{y}'$  for the denominator:

$$\begin{aligned} (\mathbf{y}')^T D \mathbf{y}' &= \sum_{i \in [n]} y'_i d_i y'_i \\ &= \sum_{i \in S^*} (1 - \sigma)^2 d_i + \sum_{i \notin S^*} (-\sigma)^2 d_i \\ &= (1 - \sigma)^2 \text{vol}(S^*) + \sigma^2 (\text{vol}(V) - \text{vol}(S^*)) \\ &= (1 - \sigma) \text{vol}(S^*) \end{aligned}$$

Putting everything together, we get:

$$\begin{aligned}
 R(y') &= \frac{(y')^T Ly'}{(y')^T Dy'} \\
 &= \frac{|\partial S^*|}{(1 - \sigma) \text{vol}(S^*)} \\
 &= \frac{\phi(S^*)}{1 - \sigma} \\
 &\leq 2\phi(S^*)
 \end{aligned}$$

□

## 2.2 Proof of $\phi(G) \leq \sqrt{2\mu_2}$

The proof is very involved, but we will outline a general strategy to approach the proof. Our motivation is to show that if  $\mu_2$  is small, the conductance should be small as well and thus we should be able to construct a cut that is small.

*Proof.* Suppose that  $v_2$  is the second eigenvector, or in other words  $R(v_2) = \mu_2$ . We want to construct a set  $S$  such that:

$$\frac{|\partial S|}{\min\{\text{vol}(S), \text{vol}(\bar{S})\}} \leq \sqrt{2R(v_2)}$$

The question is now how do we construct a set given this eigenvector. As this is a proof sketch, we claim that we can do this with the following algorithm.

- Sort vertices by their value in  $y \triangleq v_2$ , so we have  $y_1 \leq y_2 \cdots \leq y_n$
- Consider every cut  $S_t = \{i; y_i \leq t\}$
- Output  $\arg \min_t \phi(S_t)$

The objective of this algorithm is to essentially find the “cut threshold” that will give us the graph with the least conductance, which we can then show to be less than the RHS of our inequality. As a side remark, it’s easy to verify that the algorithm described above runs in  $\mathcal{O}(m)$  time after computing the second eigenvector  $v_2$ . □

More regarding Cheeger’s Inequality will be described in later lectures.