

## Lecture 5: Cheeger-type Inequalities for $\lambda_n$

*In which we prove an analog of Cheeger's inequalities for the largest Laplacian eigenvalue and we show how to use it to develop a spectral approximation algorithm for Max Cut.*

### 1 Cheeger-type Inequalities for $\lambda_n$

Let  $G = (V, E)$  be an undirected graph (not necessarily regular),  $D$  its diagonal matrix of degrees,  $A$  its adjacency matrix,  $L = I - D^{-1/2}AD^{-1/2}$  its normalized Laplacian matrix, and  $0 = \lambda_1 \leq \dots \leq \lambda_n \leq 2$  be the eigenvalues of  $L$ , counted with multiplicities and listed in non-decreasing order.

In Handout 2, we proved that  $\lambda_k = 0$  if and only if  $G$  has at least  $k$  connected component and  $\lambda_n = 2$  if and only if there is a connected component of  $G$  (possibly, all of  $G$ ) that is bipartite.

A special case of the former fact is that  $\lambda_2 = 0$  if and only if the graph is disconnected, and the Cheeger inequalities give a “robust” version of this fact, showing that  $\lambda_2$  can be small if and only if the expansion of the graph is small. In these notes we will see a robust version of the latter fact; we will identify a combinatorial parameter that is zero if and only if the graph has a bipartite connected component, and it is small if and only if the graph is “close” (in an appropriate sense) to having a bipartite connected components, and we will show that this parameter is small if and only if  $2 - \lambda_n$  is small.

Recall that

$$2 - \lambda_n = \min_{\mathbf{x} \in \mathbb{R}^n - \{\mathbf{0}\}} \frac{\sum_{\{u,v\} \in E} (x_u + x_v)^2}{\sum_{v \in V} d_v x_v^2}$$

We will study the following combinatorial problem, which formalizes the task of finding an “almost bipartite connected component:” we are looking for a non-empty subset of vertices  $S \subseteq V$  (we allow  $S = V$ ) and a bipartition  $(A, B)$  of  $S$  such that there is a small number of “violating edges” compared to the number of edges incident

on  $S$ , where an edge  $\{u, v\}$  is *violating* if it is in the cut  $(S, V - S)$ , if it has both endpoints in  $A$ , or if it has both endpoints in  $B$ . (Note that if there are no violating edges, then  $S$  is a bipartite connected component of  $G$ .)

It will be convenient to package the information about  $A, B, S$  as a vector  $\mathbf{y} \in \{-1, 0, 1\}^n$ , where the non-zero coordinates correspond to  $S$ , and the partition of  $S$  is given by the positive versus negative coordinates. We define the “bipartiteness ratio” of  $\mathbf{y}$  as

$$\beta(\mathbf{y}) := \frac{\sum_{\{u,v\} \in E} |y_u + y_v|}{\sum_{v \in V} d_v |y_v|}$$

Note that in the numerator we have the number of violating edges, with edges contained in  $A$  or in  $B$  counted with a weight of 2, and edges from  $S$  to  $V - S$  counted with a weight of 1. In the denominator we have the sum of the degrees of the vertices of  $S$  (also called the *volume* of  $S$ , and written  $\text{vol}(S)$ ) which is, up to a factor of 2, the number of edges incident on  $S$ .

(Other definitions would have been reasonable, for example in the numerator we could just count the number of violating edges without weights, or we could have the expression  $\sum_{\{u,v\} \in E} (y_u + y_v)^2$ . Those choices would give similar bounds to the ones we will prove, with different multiplicative constants.)

We define the bipartiteness ratio of  $G$  as

$$\beta(G) = \min_{\mathbf{y} \in \{-1, 0, 1\}^n - \{\mathbf{0}\}} \beta(\mathbf{y})$$

We will prove the following analog of Cheeger’s inequalities:

$$\frac{2 - \lambda_n}{2} \leq \beta(G) \leq \sqrt{2 \cdot (2 - \lambda_n)}$$

The first inequality is the easy direction

$$\begin{aligned} 2 - \lambda_n &= \min_{\mathbf{x} \in \mathbb{R}^n - \{\mathbf{0}\}} \frac{\sum_{\{u,v\} \in E} (x_u + x_v)^2}{\sum_{v \in V} d_v x_v^2} \\ &\leq \min_{\mathbf{y} \in \{-1, 0, 1\}^n - \{\mathbf{0}\}} \frac{\sum_{\{u,v\} \in E} |y_u + y_v|^2}{\sum_{v \in V} d_v |y_v|^2} \\ &\leq \min_{\mathbf{y} \in \{-1, 0, 1\}^n - \{\mathbf{0}\}} \frac{\sum_{\{u,v\} \in E} 2 \cdot |y_u + y_v|}{\sum_{v \in V} d_v |y_v|} \end{aligned}$$

The other direction follows by applying the following lemma to the case in which  $\mathbf{x}$  is the eigenvector of  $\lambda_n$ .

**Lemma 1 (Main)** For every  $\mathbf{x} \in \mathbb{R}^n - \{\mathbf{0}\}$  there is a threshold  $t$ ,  $0 < t \leq \max_v |x_v|$ , such that, if we define  $\mathbf{y}^{(t)} \in \{-1, 0, 1\}^n$  as

$$y_v^{(t)} = \begin{cases} -1 & \text{if } x_v \leq -t \\ 0 & \text{if } -t < x_v < t \\ 1 & \text{if } x_v \geq t \end{cases}$$

we have

$$\beta(\mathbf{y}^{(t)}) \leq \sqrt{2 \cdot \frac{\sum_{\{u,v\} \in E} (x_u + x_v)^2}{\sum_{v \in V} d_v x_v^2}}$$

Note that the Lemma is giving the analysis of an algorithm that is the ‘‘bipartite analog’’ of Fiedler’s algorithm. We sort vertices according to  $|x_v|$ , and then we consider all sets  $S$  which are suffixes of the sorted order and cut  $S$  into  $(A, B)$  according to sign. We pick the solution, among those, with smallest bipartiteness ratio. Given  $\mathbf{x}$ , such a solution can be found in time  $O(|E| + |V| \log |V|)$  as in the case of Fiedler’s algorithm.

## 1.1 Proof of Main Lemma

We will assume without loss of generality that  $\max_v |x_v| = 1$ . (Scaling  $\mathbf{x}$  by a multiplicative constant does not change the Rayleigh quotient and does not change the set of  $\mathbf{y}$  that can be obtained from  $\mathbf{x}$  over the possible choices of thresholds.)

Consider the following probabilistic experiment: we pick  $t$  at random in  $[0, 1]$  such that  $t^2$  is uniformly distributed in  $[0, 1]$ , and we define the vector  $\mathbf{y}^{(t)}$  as in the statement of the lemma. We claim that

$$\frac{\mathbb{E} \sum_{\{u,v\} \in E} |y_u^{(t)} + y_v^{(t)}|}{\mathbb{E} \sum_{v \in V} d_v |y_v^{(t)}|} \leq \sqrt{2 \cdot \frac{\sum_{\{u,v\} \in E} (x_u + x_v)^2}{\sum_{v \in V} d_v x_v^2}} \quad (1)$$

and we note that the Main Lemma follows from the above claim and from the fact, which we have used before, that if  $X$  and  $Y$  are random variables such that  $\mathbb{P}[Y > 0] = 1$ , then there is a positive probability that  $\frac{X}{Y} \leq \frac{\mathbb{E}X}{\mathbb{E}Y}$ .

We immediately see that

$$\mathbb{E} \sum_{v \in V} d_v |y_v^{(t)}| = \sum_v d_v \mathbb{P}[|x_v| \geq t] = \sum_v d_v x_v^2$$

To analyze the numerator, we distinguish two cases

1. If  $x_u$  and  $x_v$  have the same sign, and, let's say,  $x_u^2 \leq x_v^2$  then there is a probability  $x_u^2$  that both  $y_u^{(t)}$  and  $y_v^{(t)}$  are non-zero (and have the same sign), meaning that  $|y_u^{(t)} + y_v^{(t)}| = 2$ ; and there is an additional probability  $x_v^2 - x_u^2$  that  $y_u^{(t)} = 0$  and  $y_v^{(t)} = \pm 1$ , so that  $|y_u^{(t)} + y_v^{(t)}| = 1$ . Overall we have

$$\mathbb{E} |y_u^{(t)} + y_v^{(t)}| = 2x_u^2 + x_v^2 - x_u^2 = x_u^2 + x_v^2$$

since the last expression is symmetric with respect to  $u$  and  $v$ , the equation

$$\mathbb{E} |y_u^{(t)} + y_v^{(t)}| = x_u^2 + x_v^2$$

holds also if  $x_u^2 \geq x_v^2$ ;

2. If  $x_u$  and  $x_v$  have opposite signs, and, let's say,  $x_u^2 \leq x_v^2$ , there is probability  $x_v^2 - x_u^2$  that  $y_u^{(t)} = 0$  and  $y_v^{(t)} = \pm 1$ , in which case  $|y_u^{(t)} + y_v^{(t)}| = 1$ , and otherwise we have  $|y_u^{(t)} + y_v^{(t)}| = 0$ . If  $x_u^2 \geq x_v^2$ , then  $|y_u^{(t)} + y_v^{(t)}|$  equals 1 with probability  $x_u^2 - x_v^2$  and it equals zero otherwise. In either case, we have

$$\mathbb{E} |y_u^{(t)} + y_v^{(t)}| = |x_u^2 - x_v^2|$$

In both cases, the inequality

$$\mathbb{E} |y_u^{(t)} + y_v^{(t)}| \leq |x_u + x_v| \cdot (|x_u| + |x_v|)$$

is satisfied.

Applying Cauchy-Schwarz as in the proof of Cheeger's inequalities we have

$$\begin{aligned} \mathbb{E} \sum_{\{u,v\} \in E} |y_u^{(t)} + y_v^{(t)}| &\leq \sum_{\{u,v\} \in E} |x_u + x_v| \cdot (|x_u| + |x_v|) \\ &\leq \sqrt{\sum_{\{u,v\} \in E} (x_u + x_v)^2} \cdot \sqrt{\sum_{\{u,v\} \in E} (|x_u| + |x_v|)^2} \end{aligned}$$

and

$$\sum_{\{u,v\} \in E} (|x_u| + |x_v|)^2 \leq \sum_{\{u,v\} \in E} 2x_u^2 + x_v^2 = 2 \sum_v d_v x_v^2$$

and, combining all the bounds, we get (1).

## 2 Application to Max Cut

In the Max Cut problem, given an undirected graph  $G = (V, E)$  we want to find a cut  $(C, V - C)$  maximizing the number of cut edges  $E(C, V - C)$ . We call  $maxcut(G)$  the

fraction of edges of  $G$  cut by the optimal solution. We see that if  $\text{maxcut}(G) = 1 - \epsilon$ , then  $2 - \lambda_n \leq 2\epsilon$ , as seen by taking the boolean vector  $\mathbf{x} \in \{-1, 1\}^n$  such that  $x_v = 1$  iff  $v \in C$ .

This means that, in polynomial time, we can find a  $\mathbf{y} \in \{-1, 0, 1\}$  such that  $\beta(\mathbf{y}) \leq 2\sqrt{\epsilon}$ . Unfortunately, the number of non-zero vertices in  $\mathbf{y}$  could be very small, and so  $\mathbf{y}$  would not help find a way to partition all vertices.

We can, however, apply the algorithm of the previous section recursively: after we find  $\mathbf{y}$ , we remove the non-zero vertices from  $G$  (because  $\mathbf{y}$  gives us a way to partition them with few violating edges) and then we can recurse on the residual graph. If the (volume of) the residual graph is very small, then we already almost have a global cut, and if the volume of the residual graph is large, then the optimal cut of  $G$  is still a good cut in the residual graph, and it implies that  $\lambda_n$  is close to 2 and that we can find a  $\mathbf{y}$  of small bipartiteness ratio, and so on.

The overall algorithm is as follows:

**RecursiveSpectralCut** ( $G = (V, E)$ )

- Use algorithm of previous section to find disjoint sets of vertices  $A, B$  such that

$$2E(A) + 2E(B) + E(A \cup B, V - A \cup B) \leq \sqrt{2 - 2\lambda_n} \cdot \text{vol}(A \cup B)$$

- If  $V = A \cup B$ , then return  $(A, B)$
- Else
  - Let  $V' := V - (A \cup B)$ , and let  $G' = (V', E')$  be the subgraph of  $G$  induced by  $V'$
  - $(C, V' - C) := \text{RecursiveSpectralCut}(G' = (V', E'))$
  - Return best cut between  $(C \cup A, (V' - C) \cup B)$  and  $(C \cup B, (V' - C) \cup A)$

We prove the following bound.

**Lemma 2** *If  $\text{maxcut}(G) = 1 - \epsilon$ , then **RecursiveSpectralCut**  $G$  finds a cut crossed by at least  $(1 - 4\sqrt{\epsilon})|E|$  edges.*

PROOF: We proceed by induction on the number of recursive steps.

If there is no recursive call, then  $A \cup B = V$ , and so  $(A, B)$  is already a cut. The number of edges of  $G$  that *do not* cross the cut is

$$E(A) + E(B) \leq \frac{1}{2} \sqrt{2 \cdot (2 - \lambda_n)} \cdot \text{vol}(V) \leq 2\sqrt{\epsilon} \cdot |E| \leq 4\sqrt{\epsilon}|E|$$

because  $2 - \lambda_n \leq 2\epsilon$  and  $\text{vol}(V) = 2|E|$ . This settles the base case.

For the inductive step, the number of edges not cut by the algorithm is at most

$$E(A) + E(B) + \frac{1}{2}E(A \cup B, V') + (|E'| - E'(C, V' - C))$$

because we should count all the edges with both endpoints in  $A$  and both endpoints in  $B$ , all the edges of  $G'$  not cut in the recursive step, and half of the edges from  $A \cup B$  to  $V'$ , because the best way of combining the cuts loses at most half of those edges.

By using the fact that  $2 - \lambda_n \leq 2\epsilon$  and the inductive hypothesis we have

$$\begin{aligned} E(A) + E(B) + \frac{1}{2}E(A \cup B, V') &\leq \sqrt{\epsilon} \cdot \text{vol}(A \cup B) \\ |E'| - E'(C, V' - C) &\leq 4\sqrt{\epsilon'} \cdot |E'| \end{aligned}$$

where  $\epsilon'$  is defined so that  $1 - \epsilon' = \text{maxcut}(G')$ .

Let us call  $\rho := \frac{|E| - |E'|}{|E|}$  the fraction of edges of  $G$  that is not in  $G'$ . Then we have

$$\text{vol}(A \cup B) \leq 2|E - E'| = 2\rho|E|$$

and

$$|E'| = |E| \cdot (1 - \rho)$$

We also have

$$\epsilon'|E'| \leq \epsilon|E|$$

because the number of edges not cut by the optimal cut of  $G'$  is at most the number of edges not cut by the optimal cut of  $G$ , given that  $G'$  is a subgraph of  $G$ . So we have

$$\epsilon' \leq \epsilon \cdot \frac{|E|}{|E'|} = \epsilon \cdot \frac{1}{1 - \rho}$$

Putting everything together, the number of uncut edges is at most

$$\begin{aligned} &\sqrt{\epsilon} \cdot \text{vol}(A \cup B) + 4\sqrt{\epsilon'} \cdot |E'| \\ &\leq \sqrt{\epsilon} \cdot 2\rho|E| + 4\sqrt{\epsilon \cdot (1 - \rho)}|E| \\ &\leq 4\sqrt{\epsilon}|E| \end{aligned}$$

where the last line is equivalent to

$$2\rho + 4\sqrt{1 - \rho} \leq 4$$

which follows from the fact that

$$\sqrt{1 - \rho} \leq \sqrt{\left(1 - \frac{\rho}{2}\right)^2} = 1 - \frac{\rho}{2}$$

□

For small  $\epsilon$ , the bound of the previous lemma is close to the bound achieved by the Goemans-Williamson semidefinite programming-based algorithm, that, under the assumption that  $\text{maxcut}(G) = 1 - \epsilon$ , finds a cut crossed by about a  $1 - \frac{2}{\pi}\sqrt{\epsilon}$  fraction of edges, which is best possible for polynomial time algorithms assuming the unique games conjecture.

The bound of the lemma is not very good for large  $\epsilon$ , however: if  $\epsilon > \frac{1}{64}$ , the lemma guarantees a number of cut edges that is less than half the number of edges, which is worse than the performance of a simple greedy algorithm, and if  $\epsilon > \frac{1}{16}$  the lemma does not guarantee that any edge at all is cut.

If one always chooses the best between the cut returned by the recursion and a greedy cut, it is possible to derive an analysis that works well even in graphs in which the value of  $\text{maxcut}(G)$  is small, and show that the algorithm finds a cut crossed by at least a  $.531 \cdot \text{maxcut}(G)$  fraction of edges.