Lecture 3: Cheeger Inequality Continued

In which we finish the proof of Cheeger's inequalities and we discuss some generalizations.

1 Completing the Proof Cheeger's Inequality

It remains to prove the following statement.

Lemma 1 *Let* $\mathbf{y} \in \mathbb{R}^V_{\geq 0}$ *be a vector with non-negative entries. Then there is a* $0 < t \leq \max_v \{y_v\}$ *such that*

$$\phi(\{v : y_v \geq t\}) \leq \sqrt{2R_L(\mathbf{y})}$$

We will provide a probabilistic proof. Without loss of generality (multiplication by a scalar does not affect the Rayleigh quotient of a vector) we may assume that $\max_v y_v = 1$. We consider the probabilistic process in which we pick $t > 0$ in such a way that t^2 is uniformly distributed in $[0, 1]$ and then define the non-empty set $S_t := \{v : y_v \geq t\}$.

We claim that

$$\frac{\mathbb{E}\, E(S_t, V - S_t)}{\mathbb{E}\, d|S_t|} \leq \sqrt{2R_L(\mathbf{y})} \tag{1}$$

Notice that Lemma 1 follows from such a claim, because of the following useful fact.

Fact 2 *Let* X *and* Y *be random variables such that* $\mathbb{P}[Y > 0] = 1$*. Then*

$$\mathbb{P}\left[\frac{X}{Y} \leq \frac{\mathbb{E}\, X}{\mathbb{E}\, Y}\right] > 0$$

Proof: Call $r := \frac{\mathbb{E} X}{\mathbb{E} Y}$. Then, using linearity of expectation, we have $\mathbb{E} X - rY = 0$, from which it follows $\mathbb{P}[X - rY \leq 0] > 0$, but, whenever $Y > 0$, which we assumed to happen with probability 1, the conditions $X - rY \leq 0$ and $\frac{X}{Y} \leq r$ are equivalent. $\square$

It remains to prove (1).

To bound the denominator, we see that

$$\mathbb{E}\, d|S_t| = d \cdot \sum_{v \in V} \mathbb{P}[v \in S_t] = d \sum_v y_v^2$$

because

$$\mathbb{P}[v \in S_t] = \mathbb{P}[y_v \geq t] = \mathbb{P}[y_v^2 \geq t^2] = y_v^2$$

To bound the numerator, we say that an edge is *cut* by S_t if one endpoint is in S_t and another is not. We have

$$\mathbb{E}\, E(S_t, V - S_t) = \sum_{\{u,v\} \in E} \mathbb{P}[\{u,v\} \text{ is cut}]$$

$$= \sum_{\{u,v\} \in E} |y_v^2 - y_u^2| = \sum_{\{u,v\} \in E} |y_v - y_u| \cdot (y_u + y_v)$$

Applying Cauchy-Schwarz, we have

$$\mathbb{E}\, E(S_t, V - S_t) \leq \sqrt{\sum_{\{u,v\} \in E} (y_v - y_u)^2} \cdot \sqrt{\sum_{\{u,v\} \in E} (y_v + y_u)^2}$$

and applying Cauchy-Schwarz again (in the form $(a+b)^2 \leq 2a^2 + 2b^2$) we get

$$\sum_{\{u,v\} \in E} (y_v + y_u)^2 \leq \sum_{\{u,v\} \in E} 2y_v + 2y_u^2 = 2d \sum_v y_v^2$$

Putting everything together gives

$$\frac{\mathbb{E}\, E(S_t, V - S_t)}{\mathbb{E}\, d|S_t|} \leq \sqrt{2 \frac{\sum_{\{u,v\} \in E} (y_v - y_u)^2}{d \sum_v y_v^2}}$$

which is (1).

2 Cheeger-type Inequalities for λ_n

Let $G = (V, E)$ be an undirected graph (not necessarily regular), D its diagonal matrix of degrees, A its adjacency matrix, $L = I - D^{-1/2} A D^{-1/2}$ its normalized

Laplacian matrix, and $0 = \lambda_1 \leq \cdots \leq \lambda_n \leq 2$ 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 λ_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 $vol(S)$) 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 λ_n.

Lemma 3 (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*

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

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.

2.1 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

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

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 (2).