ADVANCED ALGORITHMS (VI)

CHIHAO ZHANG

Last week we introduced the notion of edge expansion, and its relation with the eigenvalues of the Laplacian. For general graphs (not necessarily d-regular), we define for every $S \subseteq V$,

$$\phi(S) = \frac{\left|E(S, \bar{S})\right|}{\sum_{i \in S} \deg(i)},$$

and

$$\phi(G) = \min_{S \subseteq V} \max\left\{\phi(S), \phi(\bar{S})\right\}.$$

The Cheeger's inequality provides both an upper bound and a lower bound for $\phi(G)$, in terms of the second smallest eigenvalue of the normalized Laplacian N:

$$\frac{\lambda_2}{2} \le \phi(G) \le \sqrt{2\lambda_2}.$$

Today we will prove the inequality.

1. Proof of the Lower Bound

In this section, we prove $\lambda_2 \le 2\phi(G)$. We use the characterization

$$\lambda_2 = \min_{\text{2-dim subspace } X \subseteq \mathbb{R}^n} \max_{\mathbf{x} \in X \setminus \{0\}} R_N(\mathbf{x}).$$

Therefore, in order to prove that λ_2 is small, we only need to show that for some subspace $X \subseteq \mathbb{R}^n$, it holds that for every $x \in X \setminus \{0\}$, $R_N(\mathbf{x}) \le 2\phi(G)$.

Recall that $\phi(G) = \min_{S \subseteq V} \max\left\{\phi(S), \phi(\bar{S})\right\}$. We let S be the set of vertices achieving the minimum, namely $\phi(S) = \phi(G)$. Let $\mathbf{1}_S \in \mathbb{R}^n$ be the vector that

$$\mathbf{1}_S(i) = \begin{cases} 1, & i \in S \\ 0, & i \notin S. \end{cases}$$

Define $\mathbf{1}_{\bar{S}}$ similarly. We let X be the space $\mathrm{span}(D^{\frac{1}{2}}\mathbf{1}_S, D^{\frac{1}{2}}\mathbf{1}_{\bar{S}})$ where $D \triangleq \mathrm{diag}(\deg(1), \deg(2), \dots, \deg(n))$. Then every $\mathbf{x} \in X$ can be written as $\mathbf{x} = aD^{\frac{1}{2}}\mathbf{1}_S + bD^{\frac{1}{2}}\mathbf{1}_{\bar{S}}$ for some $a, b \in \mathbb{R}$. First note that

$$R_N(aD^{\frac{1}{2}}\mathbf{1}_S) = R_N(D^{\frac{1}{2}}\mathbf{1}_S) = \frac{\langle D^{\frac{1}{2}}\mathbf{1}_S, ND^{\frac{1}{2}}\mathbf{1}_S\rangle}{\langle D^{\frac{1}{2}}\mathbf{1}_S, D^{\frac{1}{2}}\mathbf{1}_S\rangle} = \frac{\langle \mathbf{1}_S, L\mathbf{1}_S\rangle}{\langle \mathbf{1}_S, D\mathbf{1}_S\rangle} = \phi(S),$$

and similarly

$$R_N(bD^{\frac{1}{2}}\mathbf{1}_{\bar{S}}) = \phi(\bar{S}) \le \phi(S).$$

If one of a or b is zero, the inequality obviously follows. Therefore, as long as we can show for every $a, b \neq 0$, it holds taht

$$R_N(\mathbf{x}) = R_N(aD^{\frac{1}{2}}\mathbf{1}_S + bD^{\frac{1}{2}}\mathbf{1}_{\bar{S}}) \le R_N(aD^{\frac{1}{2}}\mathbf{1}_S) + R_N(bD^{\frac{1}{2}}\mathbf{1}_{\bar{S}}),$$

the inequality is proved.

In fact, we prove the following stronger statement: For every symmetric M, every pair of nonzero vectors $\mathbf{x}, \mathbf{y}$ such that $\langle \mathbf{x}, \mathbf{y}\rangle = 0$, it holds that $R_M(\mathbf{x} + \mathbf{y}) \le 2 \cdot \max\left\{R_M(\mathbf{x}), R_M(\mathbf{y})\right\}$.

Consider the spectral decompositions of the two vectors $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{v}_i$ and $\mathbf{y} = \sum_{i=1}^n b_i \mathbf{v}_i$. We have

$$\begin{aligned}
R_M(\mathbf{x}+\mathbf{y}) &= \frac{\langle \sum_{i=1}^n (a_i+b_i)\mathbf{v}_i, M(\sum_{i=1}^n (a_i+b_i)\mathbf{v}_i)\rangle}{\langle \sum_{i=1}^n (a_i+b_i)\mathbf{v}_i, \sum_{i=1}^n (a_i+b_i)\mathbf{v}_i\rangle} \\
&= \frac{\sum_{i=1}^n \lambda_i (a_i+b_i)^2}{\sum_{i=1}^n (a_i+b_i)^2} \\
&\overset{①}{\leq} \frac{\sum_{i=1}^n \lambda_i 2(a_i^2+b_i^2)}{\sum_{i=1}^n a_i^2 + \sum_{i=1}^n b_i^2} \\
&\overset{②}{\leq} 2\cdot \max\left\{ \frac{\sum_{i=1}^n \lambda_i a_i^2}{\sum_{i=1}^n a_i^2}, \frac{\sum_{i=1}^n \lambda_i b_i^2}{\sum_{i=1}^n b_i^2} \right\} \\
&= 2\cdot \max\left\{ R_M(\mathbf{x}), R_M(\mathbf{y}) \right\}.
\end{aligned}$$

In the above calculation, the denominator of ① is due to $\mathbf{x} \perp \mathbf{y}$ and the numerator follows from $(a+b)^2 \leq 2(a^2+b^2)$; ② is due to the inequality $\frac{a_1+a_2}{b_1+b_2} \leq \max_{i=1,2} \frac{a_i}{b_i}$ for nonnegative a_i and b_i.

2. Proof of the Upper Bound

The proof of the upper bound $\phi(G) \leq \sqrt{2\lambda_2}$ is more involved. The proof we are going to introduce today is in fact an analysis of the following approximation algorithm for edge expansion $\phi(G)$.

> Fiedler's Algorithm
> Input: A graph $G = (V, E)$ and a vector $\mathbf{x} \in \mathbb{R}^n$.
> 1. Number the vertex set $V = \{v_1, \dots, v_n\}$ according to $\mathbf{y} \triangleq D^{-\frac{1}{2}}\mathbf{x}$ so that $\mathbf{y}(i) \leq \mathbf{y}(i+1)$ for every $i = 1, \dots, n-1$.
> 2. For every $i = 1, 2, \dots, \lfloor \frac{n}{2} \rfloor$, define $S_i = \{1, 2, \dots, i\}$.
> 3. Return $\min_{1 \leq i \leq \lfloor \frac{n}{2} \rfloor} \phi(S_i)$.

The performance of Fiedler's algorithm depends on the input vector $\mathbf{x}$. We now prove

Theorem 1. *For every $\mathbf{x} \perp D^{\frac{1}{2}}\mathbf{1}$, Fiedler's algorithm finds a set S such that*

$$\phi(S) \leq \sqrt{2R_N(\mathbf{x})}.$$

Then Cheeger's inequality follows by taking $\mathbf{x}$ to $\mathbf{v}_2$.

Now we start to prove Theorem 1. Let $\mathbf{x} \perp D^{\frac{1}{2}}\mathbf{1}$ be a vector. Fiedler's algorithm defines n sets $S_1, \dots, S_n$ and returns the one with minimum expansion. We now use probabilistic method to show that one of S_i has expansion at most $\sqrt{2R_N(\mathbf{x})}$.

We already know from the last lecture that if we let $\mathbf{y} = D^{-\frac{1}{2}}\mathbf{x}$, then $R_N(\mathbf{x}) = \frac{\langle \mathbf{y}, L\mathbf{y}\rangle}{\langle \mathbf{y}, D\mathbf{y}\rangle}$. Moreover, $\mathbf{x} \perp D^{\frac{1}{2}}\mathbf{1}$ if and only if $\mathbf{y} \perp D\mathbf{1}$. Assume without loss of generality that $\mathbf{y}(1) \leq \mathbf{y}(2) \leq \cdots \leq \mathbf{y}(n)$. Let ℓ be the smallest index such that

$$\sum_{k \leq \ell} \deg(v_k) \geq \sum_{k > \ell} \deg(v_k).$$

We shift the vector $\mathbf{y}$ by letting $\mathbf{y}' = \mathbf{y} - \mathbf{y}(j)\mathbf{1}$. It is not hard to see that $\frac{\langle \mathbf{y}', L\mathbf{y}'\rangle}{\langle \mathbf{y}', D\mathbf{y}'\rangle} \leq \frac{\langle \mathbf{y}, L\mathbf{y}\rangle}{\langle \mathbf{y}, D\mathbf{y}\rangle}$, since shifting in the direction of $\mathbf{1}$ does not change the numerator but increasing the denominator due to $\mathbf{y} \perp D\mathbf{1}$ (this can be verified by considering $\langle \mathbf{y} + z\mathbf{1}, D(\mathbf{y} + z\mathbf{1})\rangle$ as a function on z and looking at its derivative). Moreover, if for every $t \in \mathbb{R}$, we let $S_t \triangleq \{v_i : \mathbf{y}'(i) \leq t\}$, then every S_t is among the separators considered by Fiedler's algorithm with input $\mathbf{x}$. Therefore, we can sample separators considered by Fiedler's algorithm by sampling a number t in $\mathbb{R}$. To define a suitable distribution on $\mathbb{R}$, we can further assume $\mathbf{y}'(1)^2 + \mathbf{y}'(n)^2 = 1$ without loss of generality. Then we can sample t in $[\mathbf{y}'(1), \mathbf{y}'(n)]$ with probability density $f(t) = 2|t|$ (Figure 1).

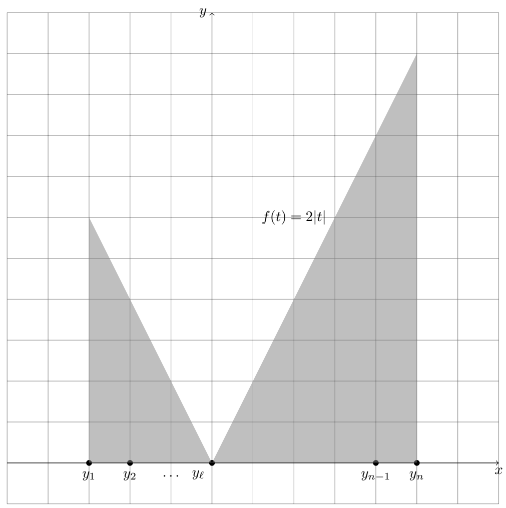

FIGURE 1. The probability desnity $f(t) = 2\,|t|$.

Everything is going to be in a very nice form with this mysterious distribution. Recall that

$$\phi(G) = \min_{S \subseteq V} \max\left\{\phi(S), \phi(\bar{S})\right\} = \min_{S \subseteq V} \frac{\left|E(S, \bar{S})\right|}{\min\left\{\sum_{i \in S} \deg(i), \sum_{i \in \bar{S}} \deg(i)\right\}},$$

we have

$$\begin{aligned}
\mathbf{E}\left[\left|E(S_t, \bar{S}_t)\right|\right] &= \sum_{\substack{\{i,j\} \in E \\ i \le j}} \mathbf{Pr}\left[i \in S_t, j \in \bar{S}_t\right] \\
&= \sum_{\substack{\{i,j\} \in E \\ i \le j}} \int_{\mathbf{y}'(i)}^{\mathbf{y}'(j)} f(t)\ \mathrm{d}t \\
&= \sum_{\substack{\{i,j\} \in E \\ i \le j}} \operatorname{sgn}(\mathbf{y}'(j)) \cdot \mathbf{y}'(j)^2 - \operatorname{sgn}(\mathbf{y}'(i)) \cdot \mathbf{y}'(i)^2 \\
&\le \sum_{\substack{\{i,j\} \in E \\ i \le j}} \left(\left|\mathbf{y}'(j)\right| + \left|\mathbf{y}'(i)\right|\right)\left(\mathbf{y}'(j) - \mathbf{y}'(i)\right) \\
&\overset{①}{\le} \sqrt{\sum_{\substack{\{i,j\} \in E \\ i \le j}} \left(\left|\mathbf{y}'(j)\right| + \left|\mathbf{y}'(i)\right|\right)^2} \cdot \sqrt{\sum_{\substack{\{i,j\} \in E \\ i \le j}} \left(\mathbf{y}'(j) - \mathbf{y}'(i)\right)^2} \\
&\overset{②}{\le} \sqrt{\sum_{\substack{\{i,j\} \in E \\ i \le j}} 2(\mathbf{y}'(i)^2 + \mathbf{y}'(j)^2)} \cdot \sqrt{\sum_{\substack{\{i,j\} \in E \\ i \le j}} \left(\mathbf{y}'(j) - \mathbf{y}'(i)\right)^2} \\
&= \sqrt{2 \sum_{i \in V} \deg(i) \cdot \mathbf{y}'(i)^2} \cdot \sqrt{\sum_{\substack{\{i,j\} \in E \\ i \le j}} \left(\mathbf{y}'(j) - \mathbf{y}'(i)\right)^2} \\
&= \sqrt{2\langle \mathbf{y}', D\mathbf{y}' \rangle} \cdot \sqrt{\sum_{\substack{\{i,j\} \in E \\ i \le j}} \left(\mathbf{y}'(j) - \mathbf{y}'(i)\right)^2}
\end{aligned}$$

where ① uses Cauchy-Schwartz and ② is due to the inequality $(a+b)^2 \leq 2(a^2+b^2)$. Also by the definition of $\mathbf{y}'$, it holds that

$$\begin{aligned}
\mathbf{E}\left[\min\left\{\sum_{i\in S_t}\deg(i), \sum_{i\in \bar{S}_t}\deg(i)\right\}\right] &= \mathbf{Pr}\left[t\leq 0\right]\cdot \mathbf{E}\left[\sum_{i\in S_t}\deg(i)\;\middle|\; t\leq 0\right] + \mathbf{Pr}\left[t>0\right]\cdot \mathbf{E}\left[\sum_{i\in \bar{S}_t}\deg(i)\;\middle|\; t>0\right]\\
&= \sum_{i\in V}\deg(i)\cdot \mathbf{Pr}\left[t\leq 0, i\in S_t\right] + \sum_{i\in V}\deg(i)\cdot \mathbf{Pr}\left[t>0, i\in \bar{S}_t\right]\\
&= \sum_{i\leq \ell}\deg(i)\cdot \mathbf{Pr}\left[\mathbf{y}'(i)\leq t\leq 0\right] + \sum_{i>\ell}\deg(i)\cdot \mathbf{Pr}\left[0\leq t\leq \mathbf{y}'(i)\right]\\
&= \sum_{i\in V}\deg(i)\cdot \mathbf{y}'(i)^2 = \langle \mathbf{y}', D\mathbf{y}'\rangle.
\end{aligned}$$

Therefore, putting above together yields

$$\frac{\mathbf{E}\left[\left|E(S_t,\bar{S}_t)\right|\right]}{\mathbf{E}\left[\min\left\{\sum_{i\in S_t}\deg(i), \sum_{i\in\bar{S}_t}\deg(i)\right\}\right]} \leq \frac{\sqrt{2\langle \mathbf{y}', D\mathbf{y}'\rangle}\cdot\sqrt{\sum_{\{i,j\}\in E}\left(\mathbf{y}'(j)-\mathbf{y}'(i)\right)^2}}{\langle \mathbf{y}', D\mathbf{y}'\rangle} = \sqrt{\frac{2\langle \mathbf{y}', L\mathbf{y}'\rangle}{\langle \mathbf{y}', D\mathbf{y}'\rangle}} \leq \sqrt{\frac{2\langle \mathbf{y}, L\mathbf{y}\rangle}{\langle \mathbf{y}, D\mathbf{y}\rangle}} = \sqrt{2R_N(\mathbf{x})}.$$

It remains to verify that for two random variables $X\geq 0$ and $Y>0$, $\frac{\mathbf{E}[X]}{\mathbf{E}[Y]}\leq r$ implies $\mathbf{Pr}\left[\frac{X}{Y}\leq r\right]>0$. To see this, notice that

$$\frac{\mathbf{E}\left[X\right]}{\mathbf{E}\left[Y\right]}\leq r \iff \mathbf{E}\left[X-rY\right]\leq 0 \implies \mathbf{Pr}\left[X-rY\leq 0\right]>0 \implies \mathbf{Pr}\left[\frac{X}{Y}\leq r\right]>0.$$

3. Remark

In the class I proved Cheeger's inequality for regular graphs. Please carefully read the proof for general graphs here. The proofs are adapted from two wonderful lecture notes [Spi15, Tre16].

References

[Spi15] Dan Spielman. Lecture notes on spectral graph theory. 2015. Available at http://www.cs.yale.edu/homes/spielman/561/lect06-15.pdf. 4

[Tre16] Luca Trevisan. Lecture notes on graph partitioning, expanders and spectral methods. 2016. Available at https://people.eecs.berkeley.edu/~luca/books/expanders-2016.pdf. 4