CSCI5160 Approximation Algorithms Lecturer: Siu On Chan

Notes 11: Cheeger–Alon–Milman inequality

1. Local sweep cut

We now prove the hard direction of Cheeger–Alon–Milman inequality from the previous lecture.

Theorem 1.1 (Cheeger–Alon–Milman). $\varphi(G) \leq \sqrt{2\lambda_2}$.

The proof is a "rounding algorithm" that converts any $y \in \mathbb{R}^V$ with small Rayleigh quotient $R(y) = \frac{y^\top L y}{y^\top D y} = \frac{\sum_{(i,j)\in E} w_{ij}(y_i - y_j)^2}{\sum_{i\in V} d(i)y_i^2}$ into a subset S with small conductance.

Lemma 1.2. Given any $y \in \mathbb{R}^V$, there is an algorithm to find $S \subseteq \operatorname{supp}(y)$ with $\varphi(S) \leq \sqrt{2R(y)}$.

Here $\operatorname{supp}(y) = \{i \in V \mid y_i \neq 0\}$ denotes the support of y. How to turn $y \in \mathbb{R}^V$ into a subset? We saw from last lecture that if y is the indicator $\mathbb{1}_T$ of some subset $T \subseteq V$, then $R(y) = \varphi(T)$. It is natural to consider rounding by thresholding: Choose threshold $t \in \mathbb{R}$ and output $S_t = \{i \in V \mid y_i > t\}.$

The algorithm instead output $S_t = \{i \in V \mid y_i^2 > t\}$. The squaring allows us to relate conductance to Rayleigh quotient, which involves squared terms $(y_i - y_j)^2$ and y_i^2 in the numerator and denominator, respectively.

Proof of Lemma 1.2. Imagine threshold t increases from zero to infinity, and $S_t = \{i \in V \mid y_i^2 > t\}$ shrinks from $\operatorname{supp}(y)$ to \emptyset . The cut weight $w(S_t, \overline{S}_t)$ and total degree $d(S_t)$ also changes as t grows.

We will assume all $|y_i| \leq 1$, as scaling y by a constant does not affect R(y). We will also pick $t \in [0,1]$ uniformly at random. We now analyze the expected cut weight $\mathbb{E}_t[w(S_t, \overline{S}_t)]$ and expected total degree $\mathbb{E}_t[d(S_t)]$.

$$\begin{split} \mathbb{E}_{t}[w(S_{t},\overline{S}_{t})] &= \sum_{(i,j)\in E} w_{ij} \mathbb{E}_{t}[\mathbb{1}((i,j) \text{ is cut by } S_{t})] \\ &= \sum_{(i,j)\in E} w_{ij}(y_{j}^{2} - y_{i}^{2}) \quad \text{assuming } y_{i}^{2} \leqslant y_{j}^{2} \\ &= \sum_{(i,j)\in E} w_{ij}(y_{j} - y_{i})(y_{j} + y_{i}) \leqslant \sqrt{\sum_{(i,j)\in E} w_{ij}(y_{j} - y_{i})^{2}} \sqrt{\sum_{(i,j)\in E} w_{ij}(y_{j} + y_{i})^{2}} \end{split}$$

The inequality is Cauchy–Schwarz. The first term under square-root is the numerator of the Rayleigh quotient. For the second term under square-root,

$$\sum_{(i,j)\in E} w_{ij}(y_i + y_j)^2 \leqslant \sum_{(i,j)\in E} w_{ij}2(y_i^2 + y_j^2) = 2\sum_{i\in V} d(i)y_i^2 .$$

Altogether,

$$\mathbb{E}_{t}[w(S_{t},\overline{S}_{t})] \leqslant \sqrt{\sum_{(i,j)\in E} w_{ij}(y_{j}-y_{i})^{2}} \sqrt{2\sum_{i\in V} d(i)y_{i}^{2}} .$$

Now for the expected total degree,

$$\mathbb{E}_t[d(S_t)] = \sum_{i \in V} d(i) \mathbb{E}_t[\mathbb{1}(i \in S_t)] = \sum_{i \in V} d(i) y_i^2.$$

So their ratio satisfies

$$\frac{\mathbb{E}_t[w(S_t, \overline{S}_t)]}{\mathbb{E}_t[d(S_t)]} \leqslant \sqrt{2R(y)} \ .$$

By the following proposition, there must be some choice of $t = t_*$ such that

$$\varphi(S_{t_*}) = \frac{w(S_{t_*}, S_{t_*})}{d(S_{t_*})} \leqslant \sqrt{2R(y)} . \qquad \square$$

Proposition 1.3. Let f and g be arbitrary real-valued integrable functions. There must be some choice of t_* such that

$$\frac{f(t_*)}{g(t_*)} \leqslant \frac{\mathbb{E}_t[f(t)]}{\mathbb{E}_t[g(t)]} \; .$$

Proof. Let $C = \mathbb{E}_t[f(t)] / \mathbb{E}_t[g(t)]$, so that

$$0 = \mathop{\mathbb{E}}_{t}[f(t)] - C\mathop{\mathbb{E}}_{t}[g(t)] = \mathop{\mathbb{E}}_{t}[f(t) - Cg(t)]$$

There must be some choice of $t = t_*$ such that the term in the expectation is nonpositive:

$$f(t_*) - Cg(t_*) \leq 0 \qquad \Longleftrightarrow \qquad \frac{f(t_*)}{g(t_*)} \leq C = \frac{\mathbb{E}_t[f(t)]}{\mathbb{E}_t[g(t)]}.$$

The algorithm in Lemma 1.2 can find small conductance S_{t_*} deterministically: Simply try all thresholds t that lead to different $S_t = \{i \in V \mid y_i^2 > t\}$, and output the one with the smallest conductance. There are at most n choices for t once vertices are sorted according to y_i^2 .

2. FROM ORTHOGONALITY TO SMALL SUPPORT

Does Lemma 1.2 prove Theorem 1.1? Not yet, the subset S_{t_*} produced need not contain at most half of the total degree. It may even be the case that $S_{t_*} = V$.

But we also did not exploit the orthogonality condition: that $\sum_{i \in V} d(i)y_i = 0$. In this section, given $y \in \mathbb{R}^V$ with small Rayleigh quotient and satisfying the orthogonality condition, we will produce two vectors z_- and z_+ both with "small support", and apply the algorithm in previous section to z_- or z_+ .

Note that the numerator of the Rayleigh quotient does not change if all entries of y are shifted by the same $c \in \mathbb{R}$. Among all shifts $z = y + c\mathbb{1}$, the denominator of the Rayleigh quotient is minimized when $\sum_{i \in V} d(i)z_i = 0$, because the quadratic form

$$z^{\top}Dz = \sum_{i \in V} d(i)z_i^2 = \sum_{i \in V} d(i)(y_i + c)^2$$

has derivative (with respect to c) $2\sum_{i \in V} d(i)y_i$.

Assume without loss of generality that y is sorted, so that $y_1 \leq \ldots \leq y_n$. Find the smallest j such that $\sum_{1 \leq i \leq j} d(i) \geq d(V)/2$. We will then shift y by $c = -y_j$ to obtain $z = y - y_j \mathbb{1}$. The previous paragraph implies that $R(z) \leq R(y)$, because the numerator stays the same but the denominator can only increase after the shift.

Note that $z_j = 0$. The above choice of j ensures both sets $S_- = \{i \in V \mid y_i < y_j\} = \{i \in V \mid z_i < 0\}$ and $S_+ = \{i \in V \mid y_i > y_j\} = \{i \in V \mid z_i > 0\}$ contain at most half of the total degree of V. We will take the positive and negative part of z to get z_+ and z_- :

$$z_{-} = \begin{cases} z_i & z_i < 0\\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad z_{+} = \begin{cases} z_i & z_i > 0\\ 0 & \text{otherwise} \end{cases}$$

We now show z_{-} or z_{+} has Rayleigh quotient at most that of z.

Lemma 2.1. $\min\{R(z_{-}), R(z_{+})\} \leq R(z)$.

Proof. $z^{\top}Dz = z_{+}^{\top}Dz_{+} + z_{-}^{\top}Dz_{-}$, because left-hand-side is a weighted sum of z_{i}^{2} , and each nonzero z_{i}^{2} is counted in z_{+} or z_{-} .

 $z^{\top}Lz \ge z_{+}^{\top}Lz_{+} + z_{-}^{\top}Lz_{-}$, because left-hand-side is a weighted sum of $(z_i - z_j)^2$ over edges, and every edge that contribute to left-hand-side, it either get dropped if z_i and z_j have opposite signs, or is retained otherwise.

We have therefore shown $\frac{z_{-}^{\top}Lz_{-} + z_{+}^{\top}Lz_{+}}{z_{-}^{\top}Dz_{-} + z_{+}^{\top}Dz_{+}} \leq R(z)$. The result follows once we can show

$$\min\left\{\frac{A}{C}, \frac{B}{D}\right\} \leqslant \frac{A+B}{C+D}. \text{ And it is implied by Proposition 1.3.}$$

3. DISCUSSION

The task of finding subset of smallest conductance is known as Sparsest Cut. This problem is NP-hard, so we settle for an approximation algorithm.

By the above arguments, an algorithm to find a set S of small conductance is as follows:

(1) Compute an eigenvector y to the second smallest eigenvalue of \mathcal{L}

(2) Sort all entries of y so that $y_{i_1} \leq \ldots \leq y_{i_n}$, i.e. vertex i_1 has the smallest value, i_n the largest (3) Try all cut of the form $S = \{i_1, \ldots, i_j\}$ (or \overline{S} , whichever has smaller total degree)

By both sides of Cheeger–Alon–Milmon, this algorithm is guaranteed to find a subset S with $\varphi(S) \leq 2\sqrt{\varphi(G)}$.

The approximation guarantee is quite bad if $\varphi(G)$ is very small, say order of 1/n.

There are other approximation algorithm with better guarantee. There is an SDP-based approximation algorithm by Arora–Rao–Vazirani with approximation ratio $O(\sqrt{\log n})$.