ORIE 6334 Spectral Graph Theory	November 29, 2016
Lecture 26	
Lecturer: David P. Williamson	*Scribe: Venus Lo*

1 Recap of Previous Lecture

Last time we showed that an algorithm due to Leighton and Rao (1988) is an $O(\log n)$-approximation algorithm for the sparsest cut problem. Starting in this lecture, we will see an $O(\sqrt{\log n})$-approximation algorithm due to Arora, Rao, and Vazirani (2004).

Recall the definition of a sparsest cut.

Definition 1 *The sparsity of S is:*

$$\rho(S) = \frac{|\delta(S)|}{|S||V-S|}$$

Then sparsest cut of a graph G is

$$\rho(G) = \min_{S \subset V} \rho(S)$$

The relaxation that we will use for the sparsest cut problem is:

$$\min \sum_e x(e)$$

$$\sum_{i,j} d_x(i,j) \geq 1$$

$$d_x \text{ is a negative type metric}$$

Here, $d_x(i,j)$ is the distance of the shortest i-j path using $x(e)$ as the length of edge e. Last time we proved that this is indeed a relaxation of the problem; in that proof, we could have replaced the inequality with $\sum_{i,j} d_x(i,j) = 1$, and this time it will be useful to do so. Last time, we also showed a connection between this relaxation and $\lambda_2(L_F)$ where L_F is the weighted Laplacian with weights equal to the flows of a flow packing.

For what follows, let us scale x up by n^2 and the objective value down by n^2:

$$\frac{1}{n^2} \min \sum_e x(e)$$

$$\sum_{i,j} d_x(i,j) = n^2$$

$$d_x \text{ is a negative type metric}$$

Recall the following equivalent definition of negative-type metrics we mentioned last time:

[0]Parts of this lecture are taken from W and Shmoys, Section 15.4 and from lecture notes of Rothvoss `https://arxiv.org/abs/1607.00854`.

Theorem 1 *d is a negative type metric if and only if there exists $f : V \to \mathbb{R}^n$ such that $d(i,j) = ||f(i) - f(j)||^2$.*

Thus we can rewrite the relaxation in terms of vectors $v_i \in \mathbb{R}^n$ such that $x(e) = ||v_i - v_j||^2$ for each edge $e = (i,j)$. We obtain the following:

$$\begin{aligned}
&\frac{1}{n^2} \min \sum_{(i,j)\in E} ||v_i - v_j||^2 && \\
&\sum_{i,j} ||v_i - v_j||^2 = n^2 && \\
&||v_i - v_j||^2 + ||v_j - v_k||^2 \geq ||v_i - v_k||^2 && \forall i,j,k \\
&v_i \in \mathbb{R}^n && \forall i
\end{aligned}$$

This vector program can be written and solved as a semidefinite program in polynomial time to within ϵ error. We want to use the optimal solution to this SDP to find a sparse cut.

Recall that we broke the Leighton-Rao algorithm in two different cases, one of which was a case in which there were many points that were relatively close to each other. In particular, we saw that if there is a cluster C such that $|C| \geq 2/3n$ and C had diameter $\leq 1/(4n^2)$, then we could find S such that $\rho(S) \leq 6 \sum_{e\in E} x(e) \leq 6 \cdot OPT$.

There is a similar case analysis for the Arora-Rao-Vazirani result; once again, if there are many points that are close to each other, we can easily find an S of sparsity within a constant of optimal. This time, we will use the vector relaxation for our bound.

Claim 2 *If there is some $i \in V$ such that:*

$$|\{j \in V : ||v_i - v_j||^2 \leq 1/4\}| \geq n/4$$

(i.e. a large subset with a small radius), then we can find S such that:

$$\rho(S) \leq \frac{16}{n^2} \sum_{(i,j)\in E} ||v_i - v_j||^2 \leq 16 \cdot OPT$$

The proof is analogous to the 6-approximation that we showed in Leighton-Rao, and is not repeated here.

For simpler notation, let $d(i,j) \equiv ||v_i - v_j||^2$ and

Definition 2 $B(i,r) = \{j \in V : d(i,j) \leq r\}$.

Given the claim above, we will assume from now on that $|B(i,1/4)| < n/4$ for all $i \in V$ (condition $\star$), otherwise we can find a cut with sparsity within a constant factor of optimal.

Furthermore, given the assumption, we can prove the following, which will allow us to prove that the ARV algorithm works.

Claim 3 *There exists $o \in V$ such that $|B(o,4)| \geq 3n/4$. Furthermore, let $U = B(o,4) - B(o,1/4)$. Then $|U| \geq n/2$ and for all $i \in U$, there exists at leas $n/4$ vertices $j \in U$ such that $d(i,j) > 1/4$.*

Proof: Assume by contradiction that no such o exists. Then for all $i \in V$, we have more than $n/4$ vertices at least distance 4 away. Summing over the distances, this gives us

$$\sum_{i,j \in V} d(i,j) = \sum_{i \in V} \left(\sum_{j \in V} d(i,j) \right) > n \cdot \left(\frac{n}{4} \cdot 4 \right) = n^2.$$

This contradicts the feasibility of v to our vector program. So there must exist some $o \in V$ such that $|B(o,4)| \geq 3n/4$. Then because $|B(o,1/4)| \leq n/4$ by condition $\star$, for $U \equiv B(o,4) - B(o,1/4)$,

$$|U| \geq |B(o,4)| - |B(o,1/4)| \geq \frac{3}{4}n - \frac{1}{4}n = \frac{1}{2}n$$

Finally, pick any $i \in U$. Since $|B(i,1/4)| < n/4$ but $|U| \geq n/2$, there must be at least $n/4$ remaining $j \in U$ such that $d(i,j) > 1/4$. □

Next, we know that solutions to the vector program are invariant under translation, so we can move o to the origin. This allows us to simplify notation and say that for all $i \in U$:

$$||v_i||^2 = ||v_i - v_o||^2 \leq 4$$
$$||v_i||^2 = ||v_i - v_o||^2 \geq 1/4.$$

2 ARV Algorithm

As with the Goemans-W max cut algorithm, the ARV algorithm takes a random vector r. The GW algorithm lets S be the set of all vertices i such that $v_i \cdot r \geq 0$. The problem with doing something similar here is that the ratio between the probability an edge (i,j) ends up in the cut and the contribution of that edge to the objective function can get unboundedly large as $\|v_i - v_j\|^2$ becomes small, which makes a max cut style analysis not work in this case. So the ARV algorithm is going to use some similar ideas, but try to avoid looking at edges with $\|v_i - v_j\|^2$ small.

Let $(v_1, \ldots, v_n)$ be the vectors we get from solving the SDP. Our algorithm proceeds as follows. In the algorithm below, Δ is a parameter we will set later.

Algorithm 1: ARV Algorithm

Pick a random vector r such that $r(i) \sim N(0,1)$
Let $L = \{i \in V : v_i \cdot r \leq -1\}$ and $R = \{i \in V : v_i \cdot r \geq 1\}$
Find a maximal matching $M \subseteq \{(i,j) \in L \times R : d(i,j) \leq \Delta\}$
Let L', R' be the vertices in L, R respectively that remain uncovered
Sort $i \in V$ by increasing distance to L' (i.e. $d(i, L')$) to get $i_1, i_2, \ldots, i_n$
Let $S_k = \{i_1, \ldots, i_k\}$ and return $S = \arg\min_{1 \leq k \leq n-1} \rho(S_k)$

Observation 1 *Step 4 above implies that if $i \in L', j \in R'$, then $d(i,j) > \Delta$, otherwise we could have increased the size of M.*

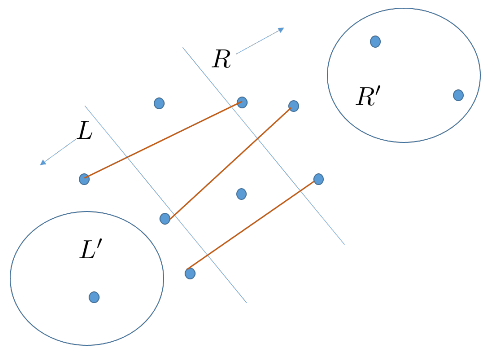

Figure 1: Pictorial depiction of algorithm

3 Starting the Analysis

To prove that ARV is an $O(\sqrt{\log n})$-approximation algorithm, we will need the following two theorems.

Theorem 4 *There is some constant* c' *such that* $\Pr[|L|, |R| > c'n] \geq c'$.

Theorem 5 (Structure Theorem) *For* $\Delta = \Omega(1/\sqrt{\log n})$, $E[|M|] \leq (c'/2)^2 n$.

Using Theorem 5 and Markov's Inequality, we can conclude that

$$\Pr\left[|M| \geq \frac{c'}{2}n\right] \leq \frac{E[|M|]}{c'/2 \cdot n} \leq \frac{c'}{2}.$$

If the size of the matching is at most $\frac{c'}{2}n$ with probability at least $1-c'/2$ while the size of $|L|$ and $|R|$ are at least $c'n$ with probability at least c', we can then conclude that $|L'|, |R'| \geq \frac{c'}{2}n$ with constant probability.

Using a similar analysis to last class and using the theorems above, we get our $O(\sqrt{\log n})$-approximation:

$$\begin{aligned}
\min_{1\leq k\leq n-1} \rho(S_k) &= \min_{1\leq k\leq n-1} \frac{|\delta(S_k)|}{|S_k||V-S_k|} \\
&\leq \frac{\sum_{k=1}^{n-1} |d(i_{k+1}, L') - d(i_k, L')||\delta(S_k)|}{\sum_{k=1}^{n-1} |d(i_{k+1}, L') - d(i_k, L')||S_k||V-S_k|} \\
&= \frac{\sum_{(i,j)\in E} |d(i, L') - d(j, L')|}{\sum_{i,j\in V} |d(i, L') - d(j, L')|} \\
&\leq \frac{\sum_{(i,j)\in E} ||v_i - v_j||^2}{\sum_{i\in L'} \sum_{j\in R'} |d(j, L')|} \\
&\leq \frac{\sum_{(i,j)\in E} ||v_i - v_j||^2}{|L'||R'|\Delta} \\
&= O(\sqrt{\log n}) \cdot \frac{1}{n^2} \sum_{(i,j)\in E} ||v_i - v_j||^2 \\
&\leq O(\sqrt{\log n}) \cdot OPT
\end{aligned}$$

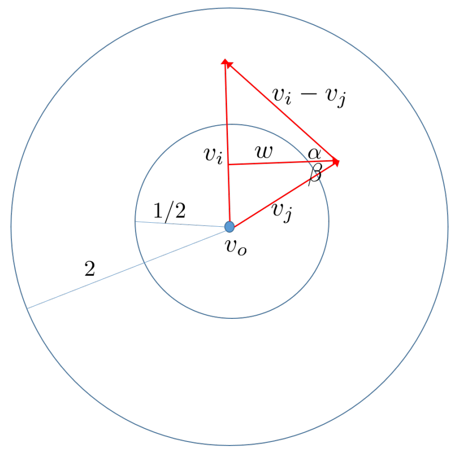

Figure 2: Fact **??**

Note that the second to the last inequality follows in the denominator since $d(j, L') \geq \Delta$ for any $j \in R'$.

We now turn to the proof of Theorem 4.

Proof of Theorem 4]: First, pick $i, j \in U$ such that $d(i, j) = ||v_i - v_j||^2 > 1/4$. We have:

$$1/4 \leq ||v_i||^2 \leq 4$$
$$1/4 \leq ||v_j||^2 \leq 4$$

Without loss of generality, assume $||v_i|| \geq ||v_j||$. Let w be the projection of v_j onto v_i. From the relaxation, we know that the following inequalities are obeyed.

$$||v_i - v_o||^2 \leq ||v_i - v_j||^2 + ||v_j - v_o||^2$$
$$||v_i - v_j||^2 \leq ||v_i - v_o||^2 + ||v_o - v_j||^2$$

These two inequalities imply that the angle between v_j and $v_i - v_j$ is not obtuse, and similarly the angle between v_i and v_j is not obtuse by the Law of Cosines. So let α be the angle between $v_i - v_j$ and w, and β be the angle between v_j and w (See Figure 2). Then it must be the case that $\alpha + \beta \leq \pi/2$ and $v_i \cdot v_j \geq 0$. We can now prove a bound on $\|w\|$ by applying some case analysis.

- If $\alpha \leq \pi/4$, then $||w|| \; ||v_i - v_j|| \cos\alpha \geq (1/\sqrt{2})(1/2)$.
- If $\beta \leq \pi/4$, then $||w|| \; ||v_j|| \cos\beta \geq (1/\sqrt{2})(1/2)$.

Hence we know that $||w|| \geq \frac{1}{2\sqrt{2}}$.

So now we can consider $\Pr[i \in L, j \in R]$. We claim that if $r \cdot v_i \in [-2, -1]$ and $r \cdot w \geq 3$, then $i \in L$ and $j \in R$. We know that $r \cdot v_i \leq -1$ places $i \in L$, so we just need to show that $r \cdot v_j \geq 1$ to place $j \in R$. Notice that

$$v_j = (v_i \cdot v_j)\frac{v_i}{||v_i||^2} + w$$

Taking a dot product with r on both sides

$$\begin{aligned} r \cdot v_j &= (v_i \cdot v_j) \frac{r \cdot v_i}{||v_i||^2} + r \cdot w \\ &\geq (1)(-2) + 3 \\ &= 1, \end{aligned}$$

as desired.

So we can lower bound the probability that $i \in L$ and $j \in R$ as

$$\begin{aligned} \Pr[i \in L, j \in R] &\geq \Pr[-2 \leq r \cdot v_i \leq -1 \text{ and } r \cdot w \geq 3] \\ &= \Pr\left[\frac{-2}{||v_i||} \leq r \cdot \frac{v_i}{||v_i||} \leq \frac{-1}{||v_i||}\right] \cdot \Pr\left[r \cdot \frac{w}{||w||} \geq \frac{3}{||w||}\right] \\ &= \Omega(1) \end{aligned}$$

The first equality is true because v_i and w are orthogonal, and thus the two probabilities are independent by properties of multivariate normal distribution (in particular $r \cdot x$ and $r \cdot y$ are independently distributed for vectors x and y orthogonal to each other). For the last line, given a unit-vector x, then $r \cdot x \sim N(0, 1)$. We have bounds on $||v_i||$ and $||w||$, so we have at least constant-sized interval of a standard normal distribution for each probability and the probabilities must be at least a constant.

Similarly we can show that

$$\Pr[i \in R, j \in L] \geq \Pr[1 \leq r \cdot v_i \leq 2,\ r \cdot w \leq -3] = \Omega(1)$$

Hence we expect $\Omega(n^2)$ pairs of such that $i \in L, j \in R$ or vice versa, and this is sufficient to prove the theorem. $\square$

4 Structure Theorem

In the next lecture we will prove the structure theorem, following a new proof given by Barak and Steurer. To give an idea of the proof, we observed before that

$$\frac{v}{||v||} \cdot r \sim N(0, 1).$$

It is possible to show that $\Pr[v \cdot r \geq \alpha] \leq \exp(-\alpha^2/||v||^2)$. In particular, for $i, j \in U$, we know that $||v_i - v_j||^2 \leq 8$, so that

$$\frac{v_i - v_j}{||v_i - v_j||} \cdot r \sim N(0, 1).$$

It then follows that

$$\Pr\left[(vi - v_j) \cdot r \geq C\sqrt{\ln n}\right] \leq e^{\frac{-C^2 \ln n}{8}} = \frac{1}{n^{C^2/8}}.$$

So for C sufficiently large, we can assume with high probability that $(v_i - v_j) \cdot r \leq C\sqrt{\ln n}$ for all $i, j \in U$. In order to show the Structure Theorem, we will show that:

$$\frac{\Omega(1)}{\Delta}\left(\frac{E[|M|]}{n}\right)^3 \leq E\left[\max_{i,j \in U}(v_i - v_j) \cdot r\right] \leq C\sqrt{\ln n}$$

Then for $\Delta = \Omega(1/\sqrt{\log n})$, we get that $\mathbb{E}[|M|] \leq \left(\frac{c'}{2}\right)^2 n$, as desired.