CS880: Approximation and Online Algorithms	**Scribe:** Jennifer Cao
Lecture 15: Sparsest Cut Cont.	**Date:** 10/18/2019

Last time we introduced *Sparsest Cut Problem* by showing an efficient LP relaxation based algorithm, which has an integrality gap of $O(logn)$. Such approximation is achieved by embedding the metric returned by LP into l_1 with distortion $O(logn)$.

In today's course, we are going to use SDP Relaxation to get a better approximation, which is $O(\sqrt{logn})$.

15.1 SDP Relaxation

Recall the setting of sparsest cut problem. Given a graph $G = (V, E)$ with positive cost c_e on every edge $e \in E$, and let $n = |V|$ be the number of vertices. The goal of sparsest cut problem is to find a partition $(S, V\backslash S)$ that minimizes $\frac{c(\delta(S))}{|S|\cdot|V\backslash S|}$.

To obtain LP relaxation, we embeded the cut metrics into l_1 metrics. Intuitively, we can embed the cut metrics into a smaller class of metrics that contains l_1 so as to get a tighter SDP relaxation.

It's known that it is possible to optimize over negative type metric in semidefinite programming, and a "squared-Euclidean" metric is negative type metric with a nice property: if $d \in l_1$, then $d \in l_2^2$. Therefore, given param γ, a semidefinite relaxation[1] can be formulated as:

$$\begin{aligned} \min \quad & \sum_e c_e \cdot d_e \\ \text{subject to} \quad & \text{d is a metric} \\ & \sum_{i,j} d_{ij} = \gamma n^2 \\ & d_{ij} = \|x_i - x_j\|^2, \quad \forall i,j \in V \\ & \|x_i\|^2 = 1, \quad \forall i \in V \end{aligned}$$

where the last two constraints are obtained by definition of l_2^2 metric.

The following theorem shows that we can get an α-approximation for sparsest cut as long as there exists an embedding with distortion α.

Theorem 15.1.1 ([2]) *Suppose there exists an embedding f from a negative type metric d into l_1 such that $\forall i, j$,*

1. $|f(i) - f(j)|_1 \leq d_{ij}$
2. $\sum_{i,j} |f(i) - f(j)|_1 \geq \frac{1}{\alpha} \sum_{i,j} d_{ij}$

Then the integrality gap for sparsest cut SDP is at most α.

15.2 Arora Rao Vazirani Theorem

Theorem 15.2.1 (ARV 04'[3]) *Given any l_2^2 metric d over n points, there exists a 1- dimensional embedding f with average distortion $O(\sqrt{logn})$.*

Theorem ARV shows that given any n-point l_2^2 metric d, there exists a set $S \subseteq [n]$ such that the specific embedding $f : X \to S, f(i) = d(S,i)$ achieves a relative low distortion of $O(\sqrt{logn})$. In fact, ARV shows that there exists two sets S and T of size $\Omega(n)$ such that $\forall i \in S$ and $j \in T$, there is always

$$d_{ij} \geq \frac{\gamma}{O(\sqrt{logn})}$$

Hence, to prove this theorem we want

$$\sum_{i,j} |d(S,i) - d(S,j)| \geq O(\frac{1}{\sqrt{logn}}) \sum_{i,j} d_{ij}$$

.

Take a Fréchet embedding, embedding metric $d(S,i) = \min_{j \in S} d_{ij}$, by triangle inequality it has $\sum_{i,j} |d(S,i) - d(S,j)| \geq \sum_{j \notin S} d(S,j) \cdot |S|$. Hence we only need to prove

$$\sum_{j \notin S} d(S,j) \cdot |S| \geq O(\frac{1}{\sqrt{logn}}) \sum_{i,j} d_{ij}$$

Proof: Considering ball $B(u,r) = \{w \in V : d(u,w) \leq r\}$ around $u \in V$ with radius r.

Case 1. There exists a radius $\frac{\gamma}{4}$ ball of size $\geq \frac{n}{4}$, i.e., $\exists u \in V : |B(u, \frac{\gamma}{4})| \geq \frac{n}{4}$.
We can claim that it suffices to pick $S = B(u, \frac{\gamma}{4})$.
From figure 15.2.1, distance of two points inside set S is at most the diameter, which is $\frac{\gamma}{2}$. Thus,

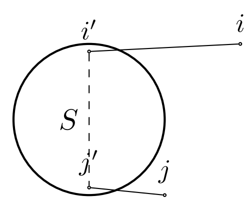

Figure 15.2.1: node i and j are embeded in $S = B(u, \frac{\gamma}{4})$.

for any i, j,

$$d(i,j) \leq d(S,i) + d(S,j) + \frac{\gamma}{2}$$

We know that $\sum_{i,j} d_{ij} = \gamma n^2$. By summing up both sides, we get

$$\begin{aligned}\gamma n^2 &= \sum_{i,j} d(i,j) \\ &\leq |V| \textstyle\sum_i d(S,i) + |V| \sum_j d(S,j) + \frac{\gamma}{2} \cdot n^2 \\ \Rightarrow \quad & \tfrac{\gamma n^2}{2} \leq 2n \textstyle\sum_i d(S,i) \\ \Rightarrow \quad & \textstyle\sum_i d(S,i) \geq \frac{\gamma n}{4} \\ \Rightarrow \quad & \sum_{i \in S, j} d(S,j) = |S| \textstyle\sum_j d(S,j) \geq \frac{n}{4} \cdot \frac{\gamma n}{4} \geq \frac{\gamma n^2}{16}\end{aligned}$$

Case 2. There's no ball of radius $\frac{\gamma}{4}$ containing at least $\frac{n}{4}$ elements, i.e., $\forall u \in V, |B(u, \frac{\gamma}{4})| < \frac{n}{4}$. Notice that there exists $u \in V$ such that for $S = B(u, 2\gamma)$, $|S| \geq \frac{n}{2}$ and $\sum_{i,j \in S} d_{ij} \geq \frac{1}{32}\gamma n^2$.

Proof of Case 2: From constraint $\sum_{i,j} d_{ij} = \gamma n^2$, we know that average distance overall is γ. Hence $\exists u : \frac{1}{n} \sum_j d(u,j) \leq \gamma$.
According to Markov's inequality, then for at least $\frac{n}{2}$ j, $d(u,j) \leq 2\gamma$. By observation of Figure

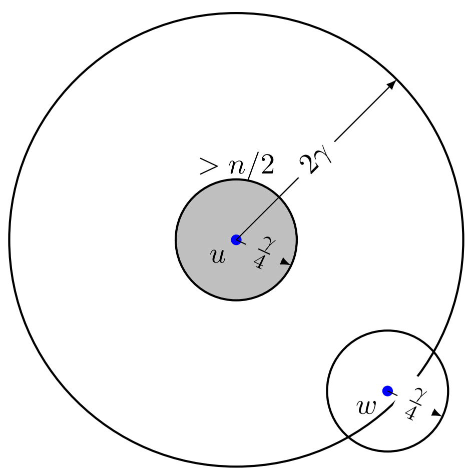

Figure 15.2.2: There are at most $\frac{n}{4}$ points inside of the ball w, and at least $\frac{n}{2} - \frac{n}{4} = \frac{n}{4}$ outside of ball w.

15.2.2,

$$\begin{aligned}\sum_{w \in S, j \in S} d(w,j) &\geq \sum_{w \in S} \sum_{j \in s, j \notin B(w, \frac{\gamma}{4})} d(w,j) \\ &\geq \sum_{w \in S} \tfrac{n}{4} \cdot \tfrac{\gamma}{4} \\ &\geq \tfrac{n}{2} \cdot \tfrac{n}{4} \cdot \tfrac{\gamma}{4} = \tfrac{\gamma n^2}{32}\end{aligned}$$

■

Thus, $\sum_{i,j\in S} d_{ij}$ is bounded by a constant factor, so the expected value of SDP is at most $O(\frac{1}{O(logn)})$.
■

15.3 Hyperplane Rounding Algorithm

15.3.1 Master Theorem

Theorem 15.3.1 (Master Theorem[3]) *Given n points in the unit ball in $\mathbb{R}^m$ with the l_2^2 metric, suppose $\sum_{ij} d_{ij} \geq c \cdot n^2$ for some constant $c > 0$, then there exists sets S and T of size $\Omega(n)$ with $\min_{i\in S, j\in T} d_{ij} \geq \frac{1}{\sqrt{logn}}$.*

Note that the above theorem is not hold if d is an arbitrary metric or square euclidean, as the triangle inequality does not hold.

See Figure 15.3.3, the algorithm contains of two phases: projection and prunning.

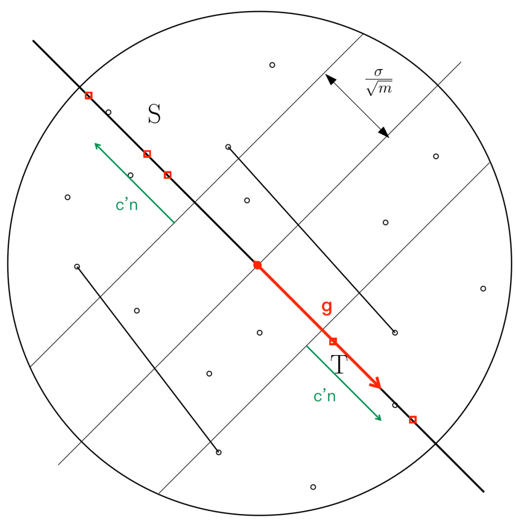

Figure 15.3.3: Separating a unit ball by a hyperplane with a margin. Black circle denotes the original vertices, and red squares are their projections on direction g. Sets S and T are found by the *hyperplane rounding algorithm*. At the *projection step*, the algorithm starts with a "fat" random hyperplane cut, and S and T are chosen as vertices that project far apart; at the *pruning step*, pairs of points that are too close to each other are discarded.

1. Project. Pick a random Gaussian of variance 1 at each dimention, and project all points in V on the line in this direction. Formally, we pick a random unit vector g, and let $Y_i := g \cdot x_i$ be the projection of x_i on g.

Then we define the following sets

$$\tilde{S} = \{i : Y_i \text{ is among the smallest c'n values}\}$$

$$\tilde{T} = \{i : Y_i \text{ is among the largest c'n values}\}$$

2. Prune. While there exists pairs $i \in \tilde{S}$ and $j \in \tilde{T}$ with $d_{ij} < \frac{c''}{\sqrt{logn}}$, pick any such pair and discard it. Finally, all what is remained in $\tilde{S}$ and $\tilde{T}$ the required sets $S \leftarrow \tilde{S}$ and $T \leftarrow \tilde{T}$, and return S and T.

15.3.2 The Projection Step

Claim 15.3.2 *There exists $\delta > 0$ such that, at the end of projection step,*

$$\mathbf{Pr}\left[\min_{i \in S, j \in T} |Y_i - Y_j| > \delta\right] = 1 - O(1)$$

Proof: Fix some $i \in \tilde{S}$ and $j \in \tilde{T}$. Consider the normalized projection Y_i and Y_j of x_i, x_j on a random direction g, and note that g is distributed as a Gaussian random variable $N(0, 1)$.

When projections of x_i and x_j lie in the same side, or in the adjoining sides, the separation of projection fails. Putting these cases together and by Markov's inequality,

$$\mathbf{Pr}\left[|\langle g, x_i - x_j\rangle| < \delta' \sqrt{d_{ij}}\right] < \text{constant} \cdot \delta'$$

Hence there exists $\Omega(Cn^2)$ pairs i, j with $d_{ij} > \frac{c}{2}$, and the probability

$$\mathbf{Pr}[|Y_i - Y_j| < \delta"] < \mathbf{Pr}\left[|Y_i - Y_j| < \delta' \sqrt{d_{ij}}\right] < \text{constant}$$

■

15.3.3 The Pruning Step

In this part we will give an intuitive explanation of pruning step. In this step, we need to show that the number of points discarded from $\tilde{S}$ and $\tilde{T}$ is small, i.e., that no more than $c'n$ pairs of points are deleted from $\tilde{S}$ and $\tilde{T}$.

Consider some (i, j), $d_{ij} < \frac{c''}{\sqrt{logn}}$ being small, and $|Y_i - Y_j| \geq \delta$.

As probability that a gaussian variable is streched by t is

$$\mathbf{Pr}[\text{Gaussian variable} \geq t] \leq \exp^{-t^2}$$

The factor that (i, j)'s projection is larger from expected length is precisely the stretching probability, hence the probability that (i, j) is discarded is

$$\mathbf{Pr}[(i, j) \text{ is discarded}] = \exp^{-\Omega(logn)} = o(1/n)$$

when t is $O(\frac{1}{\sqrt{logn}})$.

But with (i, j) getting "stretched" by a factor of $(logn)^{1/4}$,

$$\mathbf{Pr}[(i, j) \text{ is discarded}] = \exp^{-\Omega(\sqrt{logn})} = \Omega(1)$$

This means that Euclidean distance between the first and the last point is $\Omega(1)$ whereas their projection is $\Omega(\sqrt{logn})$, which is large enough to get large separated sets with high probability.

References

[1] F. Rendl, "Semidefinite programming and combinatorial optimization," *Applied Numerical Mathematics*, vol. 29, no. 3, pp. 255–281, 1999.

[2] Y. Rabinovich, "On average distortion of embedding metrics into the line and into l 1," in *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing.* ACM, 2003, pp. 456–462.

[3] S. Arora, S. Rao, and U. Vazirani, "(2004). expander flows geometric embeddings and graph partitionings," in *STOC*, 2004.