College of Computer & Information Science
Northeastern University
CS G399: Algorithmic Power Tools I

Fall 2007
22 November 2007
Scribe: Tao Jin

Lecture Outline:

- Review of Sparsest-Cut problem
- l_∞-Embedding
- Bourgain's Theorem
  - Construction of Bourgain's Embedding
  - Proof of Bourgain's Theorem

In this lecture we will first review the sparsest cut problem and its relation with metric embeddings. Then we will show that for any metric space, there is an isometric Frechet's embedding. Last, we will prove the Bourgain's Theorem, which guarantee us a $O(logn)$ approximation factor to the optimal solution of the sparsest-cut problem.

1 Sparsest-Cut Review

In the last lecture we've defined the sparsest-cut problem. Recall that:

Sparsest-Cut Problem:

Given a graph $G = (V, E)$ and $w : E \to Z^+$. Determine cut $(S, \bar{S})$ such that $\frac{E(S,\bar{S})}{|S||\bar{S}|}$ is minimized where $E(S, \bar{S})$ is number of edges crossing S and $\bar{S}$.

For any cut $(S, \bar{S})$, we defined the distance between two vertices in the graph as follow,

$$d_S(i,j) = \begin{cases} 1 & (i,j) \textit{ is cross cut edge} \\ 0 & \textit{otherwise} \end{cases}$$

We claim that any cut induces a so-called cut metric d_S. Also, we have proved that we could relax any metric cut to l_1 metric. Then we have,

$$\begin{aligned} SparsestCut &= \min_S \frac{E(S,\bar{S})}{|S||\bar{S}|} \\ &= \min_{d_S} \frac{\sum_{i,j\in E} d_S(i,j)}{\sum_{i,j} d_S(i,j)} \\ &= \min_{d_{l_1}} \frac{\sum_{i,j\in E} d_{l_1}(i,j)}{\sum_{i,j} d_{l_1}(i,j)} \end{aligned} \tag{1}$$

Suppose there is a metric d which embeds in l_1 with distortion D. According to the definition of distortion,

$$\frac{d(i,j)}{D} \le d_{l_1}(i,j) \le d(i,j)$$

$d(i,j) \geq d_{l_1}(i,j)$

$\sum_{i,j \in E} d(i,j) \geq \sum_{i,j \in E} d_{l_1}(i,j)$

$d(i,j) \leq Dd(i,j)$

$\sum_{i,j} d(i,j) \leq \sum_{i,j} d_{l_1}(i,j)$

So for any metric d

$$\frac{\sum_{i,j \in E} d(i,j)}{\sum_{i,j} d(i,j)} \geq \frac{1}{D} \frac{\sum_{i,j \in E} d_{l_1}(i,j)}{\sum_{i,j} d_{l_1}(i,j)}$$

Since we have proved the equality (1), it is easy to see that for any metric d which embeds in l_1 with distortion D, the optimal solution to minimize $\frac{\sum_{i,j \in E} d(i,j)}{\sum_{i,j} d(i,j)}$ has approximation factor of D to the optimal solution of the sparsest-cut problem.

By Bourgain's Theorem, we can choose d to have $\text{dist}(d) = O(logn)$, thus we could obtain a $O(logn)$-approximate solution by solving the linear program (2).

2 l_∞-Embedding

Frechet's Theorem: Let(V,d) be an arbitrary n-point metric space. Then there is an isometric embedding Φ:$V \rightarrow l_\infty^n$.

Proof: Let's consider the following Frechet's embedding. Given a metric (V, d), $|V| = n$. Define $\Phi : v_i \rightarrow < d(v_i, v_1), d(v_i, v_2), ..., d(v_i, v_n) >$.

By triangle inequality,

$$\begin{aligned} \|\Phi(v_i) - \Phi(v_j)\|_\infty &= \max_t |d(v_i, v_t) - d(v_j, v_t)| \\ &\leq d(v_i, v_j) \end{aligned}$$

so the embedding is nonexpanding. On the other hand,

for all $u \in V$:

$$\|\Phi(v_i) - \Phi(v_j)\|_\infty \geq |\Phi_j(v_i) - \Phi_j(v_j)| = d(v_i, v_j)$$

where $\Phi_j(v_i)$ denotes the *j-th* coordinate of v in the new metric space. By the above two inequalities, it is easy to see that the embedding is isometric.

Generally, it is rare to find isometric embeddings between two spaces of interst, so we relax and allow the embedding to alter the distances. By Bourgain's Theorem, any metric embeds in l_1 with $O(logn)$ distortion. In the rest of the lecture, we will prove Bourgain's Theorem.

3 Bourgain's Theorem

3.1 Construction of Bourgain's Embedding

Given a metric (V, d) on n points, we will randomly pick k sets of nodes: $S_1, ..., S_k \subseteq V$ and define the embedding as

$\Phi : v_i \rightarrow < d(v_i, S_1), d(v_i, S_2), ..., d(v_i, S_k) >$, where

The aim in Sparsest Cut is to break the graph into two pieces such that the pieces are large but the number edges crossing the cut is small. In other words, we want small number edges connecting two large pieces.

Problem:

Given a graph $G = (V, E)$ and $w : E \rightarrow Z^+$. Determine cut $(S, \bar{S})$ such that $\frac{E(S,\bar{S})}{|S||\bar{S}|}$ is minimized where $E(S, \bar{S})$ is number edges crossing S and $\bar{S}$.

Motivation:

Sparsest-Cut Problem is much better motivated as compared to Max-Cut Problem. One motivation for Sparsest-Cut is Balanced-Cut (where we have two pieces with small number of edges crossing and each piece is constant fraction of n). Example: (1/3,2/3) balanced cuts, $(\alpha, (1 - \alpha))$ balanced cuts.

Balanced Cuts are widely studied because it is very fundamental to Divide and Conquer paradigm. It is because if we can break our problem instance into balanced pieces, then the algorithm would be better in terms of running time and quality. From efficiency standpoint if both the pieces are some constant fraction of total size, it is guaranteed that we are making progress. From quality of algorithm standpoint, if the instance is broken evenly, then when the solutions of each pieces are stitched together, the loss of stitching the constant fraction pieces is not too much. As Divide and Conquer is a basic paradigm in algorithm design, people care about Balanced-Cut. It can be shown that if we can get a approximation for Sparsest-Cut, then we can get a pseudo-approximation to Balanced-Cut.

Another motivation for Sparsest-Cut is counting substructures in graph. For this random walks on graphs is required. And to do random walks on graph, it is important to show that we get good counts if the random walks converge to a stationary distribution. Intuition is to randomly walk on a graph for a long time, and after sometime get lost in the graph (i.e. after getting lost it is equally likely to be present anywhere in the graph). And getting lost is necessary requirement for getting substructures (i.e. this means that you count substructures in different parts of the graphs). Sparsest-Cut is a motivation for random walks.

3.2 Algorithm

We discussed Leighton-Rao(LR) algorithm couple of weeks back which is a combinatorial algorithm for uniform case. We will present $O(logn)$ due to Linial-London-Rabinovich (LLR) and Auman-Rabani(AR) which trivially handles non-uniform case.

As per LLR, we take $\frac{E(S,\bar{S})}{|S||\bar{S}|}$ as optimization over a *metric.*

Definition 1. If S is a set of points and $R^{\geq 0}$ is a set of positive reals then **metric** is a function $d : (S * S) \rightarrow R^{\geq 0}$ satisfying:

i. $StrictMetrics : d(S_i, S_j) = 0 \Leftrightarrow S_i = S_j$

ii. $Symmetry : d(S_i, S_j) = d(S_j, S_i)$

iii. $TriangleInequality : d(S_i, S_j) + d(S_i, S_j) \geq d(S_i, S_k)$

Definition 2. Suppose we have d-dimensional vector $\vec{v_i}$ i.e. $\vec{v_i} \in R^d$, then **norm** (represented by $\|\|$) can be defined as a function $\|\| : R^d \rightarrow R^{\geq 0}$. We can think of a norm as the length of the vector. It satisfies the following:

i. There is only vector whose norm is zero and that is zero vector i.e. $\|\vec{v}\| = 0 \Leftrightarrow \vec{v} = \vec{0}$

ii. Linearity of a norm under scalar multiplication (if we scale a vector by some amount, its norm also scales by the same amount) i.e. $\|\lambda\vec{v}\| = |\lambda|\|\vec{v}\|$

iii. Triangle Inequality i.e. $\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$

Examples of norm:

- $L_1 = \sum |v_i|$
- $L_2 = \sqrt{\sum |v_i|^2}$
- $L_n = \sqrt[n]{\sum |v_i|^n}$ where n is a real number ≥ 1
- $L_{\text{inf}} = \max_i |v_i|$

Consider R^d. Take any centrally symmetric convex (CSC) body. Centrally Symmetric means if $\vec{u} \in body$ then $-\vec{u} \in body$. *Convex* means if $\vec{u}, \vec{v} \in body$ then $\alpha\vec{u} + (1-\alpha)\vec{v} \in body$.

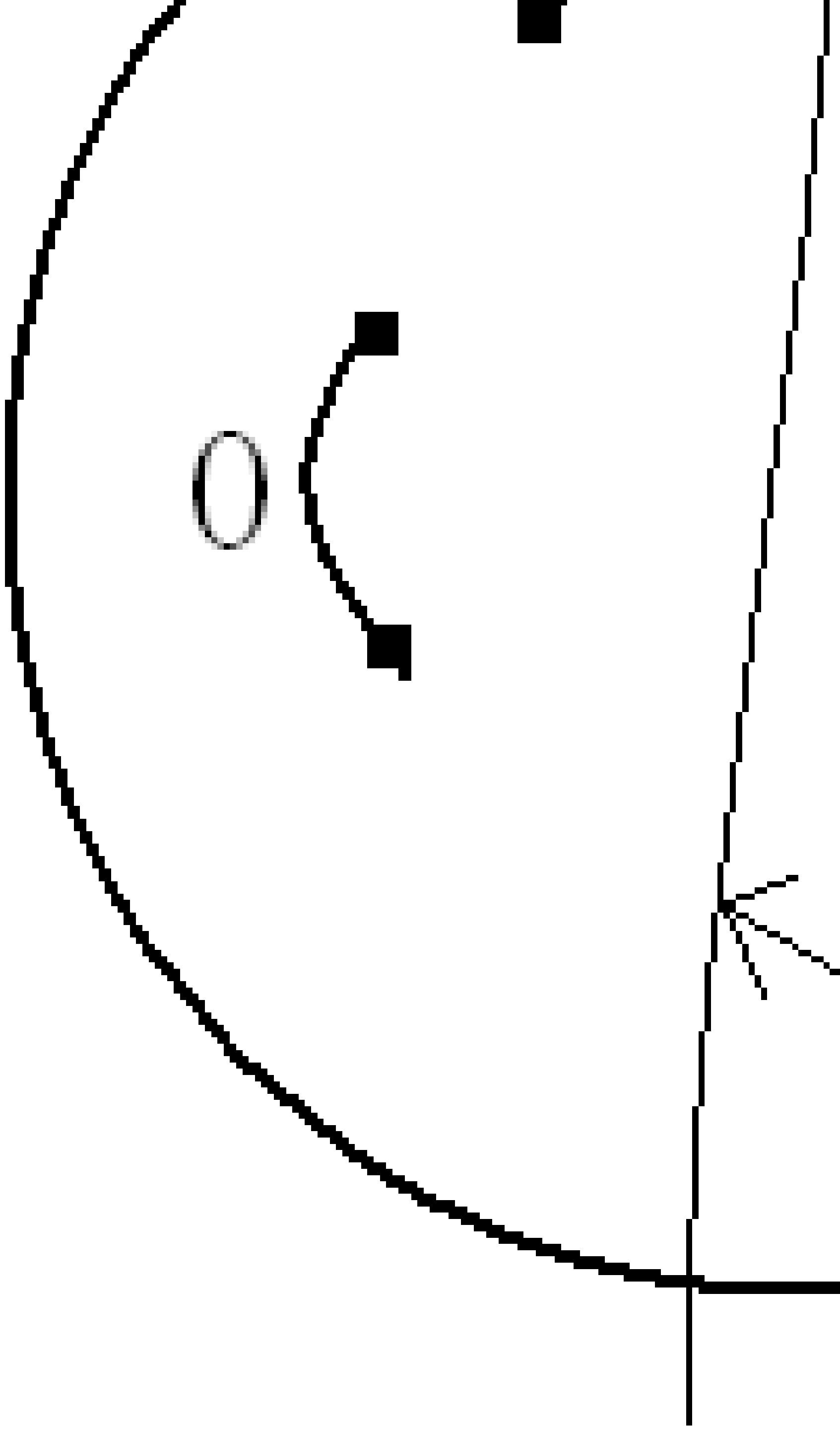
Now it is to be noted that every norm corresponds one-to-one to CSC. In other words, CSC induces a norm. To reduce a CSC to a norm, for any vector, we will compute the λ by which, the vector need to be scaled to bring it to the boundary of the ball. Since CSC is convex and centrally symmetric, there is exactly one point at which the vector can be scaled to the boundary. Then, λ is the norm of the vector. This norm satisfies *linearity* trivially. The *triangle inequality* comes from the *convexity* of CSC. The following are CSC corresponding to some norms.

Norms induce metrics. Suppose we have bunch of vectors in R^d, then pick any CSC body, which induces a norm. And define metric as $d(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\|$.

l_1 is used to denote family of metrics. In particular l_1^d denotes that there are n points belonging to d-dimensional space and the distance between them is induced by the L_1 norm.

Definition 3. We take a metric S, with distance function d on it and embed into another metric T with distance function d'. *Embedding* (E) can be formally defined as

$$E : (S, d) \underset{D}{\overset{D}{\rightleftharpoons}} (T, d')$$



If S_1 and S_2 are the two poi
distortion ($D_{(S_1,S_2)}$) for the edg
The notion of *AverageDistorti*

Consider the following graph.
side is 0, else if they are on op
them). We claim that any *cut*
satisfies non negativity, strict n

Recall, we said *SparsestCut* as optimization over a *metric*.

$$
\begin{aligned}
SparsestCut &= \min_{S} \frac{E(S,\bar{S})}{|S||\bar{S}|} \\
&= \min_{d_S} \frac{\sum_{i,j\in E} d_S(i,j)}{\sum_{i,j} d_S(i,j)} \\
&= \min_{d_{l_1}} \frac{\sum_{i,j\in E} d_{l_1}(i,j)}{\sum_{i,j} d_{l_1}(i,j)} \qquad (2)
\end{aligned}
$$

Note the equality (1). It says, we can relax from *cut metrics* to l_1 *metrics*. It can be observed that any *cut metric* is an l_1 *metric*.

One claim is, if we take at any l_1 metric, it is equal to the combination of *cut metrics* for some collection of cuts. In other words, l_1 is *the* cut cone. So, if we take any metric induced by *manhattan distance* (L_1), then that can be represented by sum of *cut metrics*. Also, the reverse direction is true i.e. any point that is represented as some combination of cuts, from that we can get a l_1 embedding. (Note that the reverse direction is not required for our analysis).

$$d_{l_1} = \sum_S (\lambda_S d_S)$$

This claim doesn't imply equality (1) as an l_1 metric we have is sum of *cut metrics* (not one particular cut). We can break the equality (1) as follows:

$$\min_{d_{l_1}} \frac{\sum_{i,j \in E} d_{l_1}(i,j)}{\sum_{i,j} d_{l_1}(i,j)} = \min_{S_1 \ldots S_k} \frac{\sum_{i,j \in E} \sum_{S_t} \lambda_{S_t} d_{S_t}(i,j)}{\sum_{i,j} \sum_{S_t} \lambda_{S_t} d_{S_t}(i,j)}$$

Using $\frac{\sum_i a_i}{\sum_i b_i} \geq min_i \frac{a_i}{b_i}$ and the above equality (just switch the order of summation and move λ outside), we can argue that the equality (1) holds.

Now,we know that we can optimize over l_1. If we remove λ from both numerator and denominator, the ratio remains the same. Hence, we can write LP for sparsest cut as follows:

$$\min_{d_{l_1}} \sum_{i,j \in E} d_{l_1}(i,j)$$

$$\text{such that } \sum_{i,j} d_{l_1}(i,j) = 1$$

Now, from l_1, we can relax it to any metric. After relaxation the value of objective function will reduce. But we can add additional constraint (triangle inequality) to represent it as LP.

$$\min \sum_{i,j \in E} d_l(i,j)$$

$$\text{such that } \sum_{i,j} d_l(i,j) = 1$$

$\forall i, j, k d_l(i,j) + d_l(j,k) \geq d_l(i,k)$

We will solve this LP, get a metric and use the following theorem by Bourgain.

Bourgain:

Any metric $\xrightarrow{O(logn)}$ $l_1^{O(log^2 n)}$ i.e. any metric embeds in l_1 with $O(logn)$ distortion.

Roadmap to Solution:

We wanted to minimize over some cuts. Then we said that minimizing over cuts is same as minimizing over l_1. Then we relaxed it and minimize over any metric. By Bourgain, any metric embeds in l_1 with $O(logn)$ distortion. From l_1, we can induce the cut (l_1 and cut cone are same).

The roadmap for Arora, Rao, Vazirani is similar. We want to minimize over some cuts. Then minimizing over cuts is same as minimizing over l_1. Then we relaxed it and minimize over l_2^2 metric (this metric is capturing SDP). By ARV, l_2^2 embeds in l_1 with $O(\sqrt{logn})$ average distortion. From l_1, we can induce the cut.