Lecture Outline:

- Max-Cut Recap
- Sparsest Cut
  - Metrics and Embeddings
  - LP based $O(\log n)$
  - LP based $O(\sqrt{\log n})$

In this lecture we will review the Max-Cut which we did in the last class. Then we will study Sparsest-Cut along with notions of metrics, norms and embedding.

1 Max-Cut Review

In the last class we saw an algorithm for Max-Cut using SDP as a black box. In Max-Cut, we break the graph into two pieces, to maximize the number of edges going across. The applications of Max-cut are in fields of VLSI design where we may desire high bandwidth/data-rates across two pieces.

We wrote vector program [P] as follows:

$$\max \frac{1}{2} \sum_{(i,j)\in E} (1 - v_i.v_j)$$
$$\text{such that } v_i \cdot v_i = 1$$

The optimal value of [P] $\geq$ Max-Cut. Here, we are replacing the naive upper bound by Z_P^* and hence getting a better approximation ratio. Cut is discrete and binary in nature i.e. either the vertex will be in S or in $\bar{S}$. So the approach of the algorithm is to make it continuous, solve it and then use that solution to guide the discrete. For further analysis of this algorithm, refer to last lecture.

Lovasz realized that algorithmic ideas from Linear Programming can be extended to Semi-Definite Programming. But SDP had a kind of continuous relaxation and getting something discrete out of it was challenging. The paper by Goemans-Williamson was the first demonstration in which we can write the SDP and which can then be rounded cleverly to get a good approximation ratio.

2 Sparsest-Cut Problem

The aim in Sparsest Cut is to break the graph into two pieces such that the pieces are large but the number of 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 an 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 and the number of edges crossing between the pieces is small, 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 graphs. For several counting applications, random walks on graphs are 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. The 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 a necessary requirement for getting substructures (i.e. this means that you count substructures in different parts of the graphs). The size of the sparsest cut is related to the time taken for random walks to converge.

2.1 Algorithm

We discussed Leighton-Rao(LR) algorithm couple of weeks back which is a combinatorial algorithm for uniform case. We will present $O(\log n)$ 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 m-dimensional vector $\vec{v_i}$ i.e. $\vec{v_i} \in R^m$, then **norm** (represented by $\|\|$) can be defined as a function $\|\| : R^m \to R^{\geq 0}$. 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. $\|\lambda\vec{v}\| = |\lambda|\|\vec{v}\|$

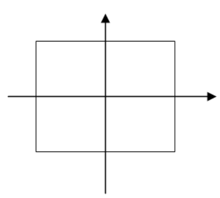
iii. Triangle Inequality i.e. $\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$

Examples of norm:

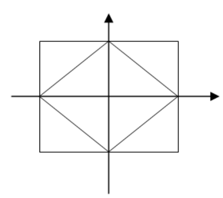
- $L_1 = \sum |v_i|$
- $L_2 = \sqrt{\sum |v_i|^2}$
- $L_n = \sqrt[n]{\sum |v_i|^n}$ where n is a real number ≥ 1
- $L_{\text{inf}} = \max_i |v_i|$

Consider R^m. 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 $\alpha\vec{u} + (1-\alpha)\vec{v} \in body$.

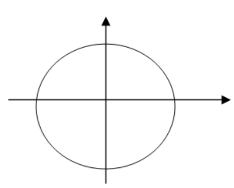
Now it is to be noted that every norm corresponds one-to-one to CSC. In other words, CSC induces a norm. To reduce a CSC to a norm, for any vector, we will compute the λ by which, the vector need to be scaled to bring it to the boundary of the ball. Since CSC is convex and centrally symmetric, there 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.



L_∞



L_1



L_2

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

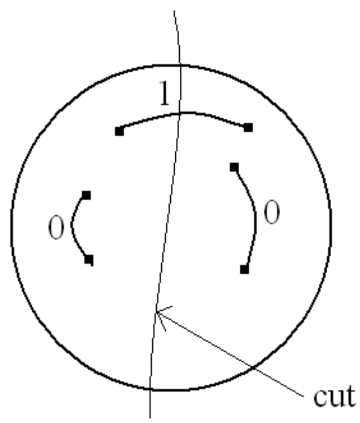
l_1 is used to denote family of metrics. In particular l_1^m denotes that there are n points belonging to m-dimensional space and the distance between them is induced by the L_1 norm.

Definition 3. We take a metric with distance function d on it and embed into another metric with distance function d' over the same set of points. *Embedding* (E) can be formally defined as

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

If i and j are the two points, then ratio of $d(i,j)$ and $d'(E(i), E(j))$ is captured by *distortion* ($D_{i,j}$) for the edge (i,j). We will focus our attention on non-expansive embeddings, that is, embeddings for which $D_{i,j} \leq 1$ for all i and j. The *Distortion* D for the entire *embedding* is $\max_{(i,j)} 1/D_{i,j}$. The notion of *Average Distortion* can be defined as $\frac{\sum_{i,j} D_{i,j}}{\text{number of pairs}}$.

Consider the following graph. Take a subset S. Distance between any pair of vertices on the same side is 0, else if they are on opposite side then 1 (irrespective of if there exists an edge between them). We claim that any *cut* induces a *metric* called *cut metric* d_S (note that distance defined satisfies non negativity, strict metrics, symmetry and triangular inequality).



Recall, we said Sparsest Cut as optimization over a *metric*.

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

Note the equality (1). It says, we can relax from *cut metrics* to l_1 *metrics*. We now argue the last step of the above equation. First, it can be easily seen that any cut metric is an l_1 metric.

We now consider the other direction. 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). Thus, there exists scalar λ_S, $S \subseteq V$ such that

$$d_{l_1} = \sum_{S \subseteq V} (\lambda_S d_S)$$

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

Roadmap to Solution:

We wanted to minimize over cut metrics. Then we said that minimizing over cut metrics 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 $O(\log n)$ 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 cut metrics. Then minimizing over cuts is same as minimizing over l_1. Then we relax it and minimize over l_2^2 metric (this metric is captured by an SDP). By ARV, l_2^2 embeds in l_1 with $O(\sqrt{\log n})$ average distortion. From l_1, we can obtain the cut.