CSC2414 - Metric Embeddings*
Lecture 8: Sparsest Cut and Embedding to ℓ_1

Notes taken by Nilesh Bansal
revised by Hamed Hatami

Summary: Sparsest Cut (SC) is an important problem with various applications, including those in VLSI layout design, packet routing in distributed networking, and clustering. But since sparsest cut is NP-hard, we need to find approximate algorithms. Solution to uniform Multi Commodity Flow (MCF) problem using Linear Programming (LP) can be used to approximate SC by $O(\log n)$ in polynomial time.

We then discuss, Poincaré inequalities for ℓ_1 metrics, which can be used to find lower bounds for distortion for embedding a metric to ℓ_1. This discussion is further continued, and we define k-gonal inequalities and hypermetrics.

1 Sparsest Cut

Definition 1.1. Flux of a graph $G = (V, E)$ is defined as,

$$\alpha_G = \min_{S \subset V, |S| \leq |V|/2} \frac{|E(S, \overline{S})|}{|S|}, \text{ where } \overline{S} = V \setminus S.$$

The cut S which minimizes the flux is known as the minimum quotient separator. Computing minimum quotient separator is NP-complete.

Definition 1.2. Sparsity of a graph $G = (V, E)$ is defined as,

$$\beta_G = \min_{S \subset V} \frac{|E(S, \overline{S})|}{|S| \cdot |\overline{S}|}.$$

The cut S which minimizes the sparsity is known as the sparsest cut (SC), which is NP-hard to compute.

Remark 1.3. Sparsity and flux of a graph are closely related.

$$\alpha_G \leq n\beta_G \leq 2\alpha_G$$

* Lecture Notes for a course given by Avner Magen, Dept. of Computer Sciecne, University of Toronto.

1.1 Approximate Solutions to Sparsest Cut

Lemma 1.4. *Solving sparsest cut is equivalent to solving*

$$\begin{aligned} \textit{minimize} \quad & \sum_{ij \in E} d(i,j) \\ \textit{subject to} \quad & \sum_{i,j \in V} d(i,j) = 1 \\ & d \in \ell_1 \end{aligned}$$

Proof. If δ_S represents the metric corresponding to the cut S, we can write,

$$\frac{|E(S,\overline{S})|}{|S| \cdot |\overline{S}|} = \frac{\sum_{i,j \in E} \delta_S(i,j)}{\sum_{\forall i,j} \delta_S(i,j)},$$

and therefore,

$$\min_S \frac{|E(S,\overline{S})|}{|S| \cdot |\overline{S}|} = \min_S \frac{\sum_{i,j \in E} \delta_S(i,j)}{\sum_{\forall i,j} \delta_S(i,j)}.$$

Recall that ℓ_1 metrics are linear combinations of cut metrics, and therefore cut metrics are extreme rays of ℓ_1. From the lemma proved in the last lecture, ratio in the equation above is minimized at one of the extreme rays of the cone. Therefore,

$$\min_S \frac{|E(S,\overline{S})|}{|S| \cdot |\overline{S}|} = \min_{d \in \ell_1} \frac{\sum_{i,j \in E} d_{ij}}{\sum_{\forall i,j} d_{ij}}.$$

Since this is invariant to scaling, without loss of generality, we can assume that the sum $\sum_{\forall i,j} d_{ij} = 1$. □

If we relax our requirement from $d \in \ell_1$ to d is a metric by adding $3\binom{n}{3}$ triangle inequalities, we can solve this problem in polynomial time using Linear Programming (LP). The relaxed LP to solve is.

$$\begin{aligned} \text{minimize} \quad & \sum_{ij \in E} d(i,j) \\ \text{subject to} \quad & \sum_{\forall i,j \in V} d(i,j) = 1 \\ & d(i,j) \geq 0, \text{ and } d(i,j) = d(j,i) \\ & d(i,j) \leq d(i,k) + d(j,k). \end{aligned} \tag{1}$$

Theorem 1.5. *There exists an $O(\log n)$ approximate algorithm for the sparsest cut problem.*

Theorem 1.5 is due to [LLR94] but it originally appeared in [LR88].

Proof. Equation 1 can be solved using LP to get a solution d^* (which is a metric). Using the Bourgain's theorem [Bou85], we can find an embedding of d^* to $d \in {\ell_1}^{O(\log^2 n)}$ with distortion $O(\log n)$. Now d can be expressed as a linear combination of $O(n \log^2 n)$ cut metrics.

$$d = \sum_{S \in \mathcal{S}} \lambda_S \delta_S, \text{ where } \mathcal{S} \text{ is a collection of cuts.}$$

Since d is in the cone of cut metrics,

$$\min_{S \in \mathcal{S}} \frac{\sum_{i,j \in E} \delta_S(i,j)}{\sum_{\forall i,j} \delta_S(i,j)} \leq \frac{\sum_{i,j \in E} d_{ij}}{\sum_{\forall i,j} d_{ij}}.$$

From Bourgain's theorem,

$$\frac{\sum_{i,j \in E} d_{ij}}{\sum_{\forall i,j} d_{ij}} \leq O(\log n) \frac{\sum_{i,j \in E} d^*_{ij}}{\sum_{\forall i,j} d^*_{ij}}.$$

But,

$$\frac{\sum_{i,j \in E} d^*_{ij}}{\sum_{\forall i,j} d^*_{ij}} = \min_{d' \text{ is metric}} \frac{\sum_{i,j \in E} d'_{ij}}{\sum_{\forall i,j} d'_{ij}} \leq \min_{\forall S} \frac{\sum_{i,j \in E} \delta_S(i,j)}{\sum_{\forall i,j} \delta_S(i,j)}.$$

Therefore,

$$\min_{S \in \mathcal{S}} \frac{\sum_{i,j \in E} \delta_S(i,j)}{\sum_{\forall i,j} \delta_S(i,j)} \leq O(\log n) \frac{\sum_{i,j \in E} d^*_{ij}}{\sum_{\forall i,j} d^*_{ij}} \leq O(\log n) \min_{\forall S} \frac{\sum_{i,j \in E} \delta_S(i,j)}{\sum_{\forall i,j} \delta_S(i,j)}.$$

□

1.2 Non-Uniform Sparsest Cut

What we have discussed till now can be generalized to the case of non-uniform sparsest cut, where we have to minimize

$$\frac{\sum_{\forall i,j} \gamma_{ij} \delta_S(i,j)}{\sum_{\forall i,j} \eta_{ij} \delta_S(i,j)}.$$

For the problem of uniform sparsest cut, $\gamma_{ij} = 1$ if $i, j \in E$, and 0 otherwise; and $\eta_{ij} = 1$ always.

2 Multi Commodity Flow Problem

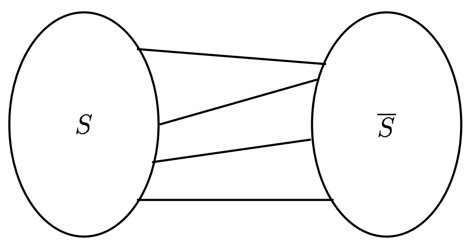
In a Multi Commodity Flow (MCF) problem, there are $k \geq 1$ commodities, each with its own source s_i, sink t_i and demand D_i. The aim is to simultaneously route all the commodities from their source to sink in a way that total amount of commodity passing through an edge is not more than the capacity of the edge. In our analysis we will only discuss a special kind of MCF that we call a uniform multi-commodity flow problem. In this special case, all edges have capacity 1, and demand D_i is same for all the commodities. Hence the problem statement in the uniform multi-commodity flow problem is to ship simultaneously maximum amount λ of commodity between each pair of vertices.

Remark 2.1. Uniform multi-commodity flow problem forms the dual to the approximate sparsest cut problem presented in Equation (1).

2.1 Uniform Sparsest Cut

If we want to ship λ units from each vertex in S to $\overline{S}$, the total flow across the cut will be $|S||\overline{S}|$. Since the number of edges carrying this load is $E(S, \overline{S})$, the maximal flow λ between each pair is bounded by

$$\lambda \leq \frac{E(S, \overline{S})}{|S||\overline{S}|}.$$



Any feasible solution to uniform MCF must therefore have $\lambda \leq \beta_G$, where β_G is the solution to sparsest cut problem, $\beta_G = \min_{S \subset V} \frac{|E(S,\overline{S})|}{|S| \cdot |\overline{S}|}$. While $\lambda \leq \beta_G$ is necessary, it is not always sufficient for a uniform MCF to have a flow of size λ. Since MCF forms dual to approximate $beta_G$, from Theorem 1.5, $\beta_G \leq O(\log n)\lambda$. Therefore,

$$\lambda \leq \beta_G \leq O(\log n)\lambda.$$

Now we will prove that $O(\log n)$ is a tight bound by providing an example where $\beta_G \geq \Omega(\log n)\lambda$. Consider a constant degree expander graph G with degree r. We want to ship λ units of commodity between every pair of vertices. The contribution to total load from flow between two vertices x and y is at least $\lambda d_G(x, y)$, where $d_G(x, y)$ is the length of shortest path. Hence total load is at least $\lambda \sum_{\forall i,j} d_G(i, j)$. Since for a constant degree graph, a large fraction of pair of vertices are in distance $O(\log n)$ asymptotically,

$$\lambda \sum_{\forall i,j} d_G(i, j) = \lambda \Omega(n^2 \log n).$$

Since total number of edges is $nr/2$ with capacity 1 each,

$$\lambda\Omega(n^2 \log n) \leq \frac{nr}{2} \cdot 1.$$

Therefore $\lambda = O(\frac{1}{n \log n})$.

Since $\frac{|E(S,\overline{S})|}{\min(|S|,|\overline{S}|)} \geq \epsilon$ for every $S \subseteq V$ in expander graphs,

$$\frac{|E(S,\overline{S})|}{|S| \cdot |\overline{S}|} \geq \frac{1}{n} \cdot \frac{|E(S,\overline{S})|}{\min(|S|,|\overline{S}|)} = \Omega(\frac{1}{n}).$$

The difference in the solution to MCF and SC in this case is $O(\log n)$.

Remark 2.2. Sparsest Cut is NP-hard. MCF is solvable using linear programming. Solution to MCF is within $O(\log n)$ to the solution of the sparsest cut. Hence we can use MCF to find $O(\log n)$ approximation to sparsest cut.

3 Lower Bound for Embedding to ℓ_1

To compute the lower bounds for distortion when embedding to ℓ_1, we will first construct an inequality that holds for ℓ_1, and then use this inequality to say something about distortion while embedding to ℓ_1. The inequality we will construct falls under the general class of Poincaré inequalities, and is of form,

$$\sum_{i,j} \alpha_{ij} \|x_i - x_j\|_1 \geq \sum_{i,j} \beta_{ij} \|x_i - x_j\|_1. \tag{2}$$

We need to determine α, β such that Equation (2) holds true. Distortion for embedding a metric d to ℓ_1 will be at least,

$$\frac{\sum_{i,j} \beta_{ij} d(i,j)}{\sum_{i,j} \alpha_{ij} d(i,j)},$$

if Equation (2) holds true for all ℓ_1 metrics.

Since linear metrics can be expressed as a linear combination of cut metrics, for every $d_1 \in \ell_1$, $d_1 = \sum_{S \in \mathcal{S}} \lambda_S \delta_S$. Thus Equation (2) will hold true if the equation below holds true for all $S \in \mathcal{S}$.

$$\begin{aligned} \sum_{i,j} \alpha_{ij} \delta_S(i,j) &\geq \sum_{i,j} \beta_{ij} \delta_S(i,j) \\ \Leftrightarrow \sum_{i,j \text{ separated by } S} \alpha_{ij} &\geq \sum_{i,j \text{ separated by } S} \beta_{ij} \end{aligned} \tag{3}$$

Let G be a graph with d_G as the metric induced by it, then one possible attempt to determine α and β can be to set,

$$\begin{aligned} \alpha_{ij} &= 1 \text{ if } i,j \in E \text{ and } 0 \text{ otherwise } (E \text{ is the edge set}), \\ \text{and } \beta_{ij} &= \beta, \text{ a constant.} \end{aligned}$$

In this case, for a cut S,

$$\sum_{i,j \text{ separated by } S} \alpha_{ij} = |E(S,\overline{S})|,$$

$$\text{and} \quad \sum_{i,j \text{ separated by } S} \beta_{ij} = \beta \cdot |S| \cdot |\overline{S}|.$$

Since this must hold true for all S, we can set $\beta = \min_S \frac{|E(S,\overline{S})|}{|S|\cdot|\overline{S}|}$ for the Poincaré inequality in Equation 2 to hold true. Therefore, the minimum distortion for embedding a metric d in ℓ_1 is at least,

$$\frac{\sum_{i,j} \beta_{ij} d(i,j)}{\sum_{i,j} \alpha_{ij} d(i,j)} = \frac{\min_S \frac{|E(S,\overline{S})|}{|S|\cdot|\overline{S}|} \cdot \sum_{i,j} d(i,j)}{|E|}$$

For a constant degree expander G with n nodes and degree r this becomes,

$$\frac{\frac{\epsilon}{n} \cdot \Omega(n^2 \log n)}{nr} = \Omega(\log n),$$

because in a constant degree graph, a constant fraction of n^2 pair of vertices have length $O(\log n)$ asymptotically.

Theorem 3.1. *A constant degree expander graph requires distortion $\Omega(\log n)$ for embedding to ℓ_1.*

3.1 k-Gonal Inequalities

Let b be an n-dimensional integral vector, $b \in \mathbb{Z}^n$, such that $\sum_i b_i = 1$. Equation (2) holds true if we set [1],

$$\alpha_{ij} = (b_i b_j)^- \text{ and } \beta_{ij} = (b_i b_j)^+.$$

[1]For a real number a, $(a)^+ = a$ if $a \geq 0$ and 0 otherwise. $(a)^- = (a)^+ - a$. Examples, $(7)^+ = 7$, $(-7)^+ = 0$, $(-3)^+ = 0$, $(-3)^- = 3$.

This can proved by proving the Equation (3) for all S.

$$\begin{aligned}
\sum_{i,j \text{ separated by } S} -\alpha_{ij} \quad &+ \sum_{i,j \text{ separated by } S} \beta_{ij} \\
&= \sum_{i,j \text{ separated by } S} ((b_i b_j)^- + (b_i b_j)^+) \\
&= \sum_{i,j \text{ separated by } S} b_i b_j \\
&= \left(\sum_{i\in S} b_i\right) \cdot \left(\sum_{j\notin S} b_j\right) \\
&= \left(\sum_{i\in S} b_i\right) \cdot \left(1 - \sum_{j\in S} b_j\right) \quad \text{because } \sum_i b_i = 1 \\
&= M \cdot (1-M) \quad \text{where } M = \sum_{i\in S} b_i \text{ is an integer} \\
&\leq 0
\end{aligned}$$

Remark 3.2. For all $b \in \mathbb{Z}^n$, such that $\sum_i b_i = 1$,

$$\sum_{i,j} (b_i b_j)^- \|x_i - x_j\|_1 \geq \sum_{i,j} (b_i b_j)^+ \|x_i - x_j\|_1, \tag{4}$$

is a valid inequality. This inequality is known as k-gonal inequality, with $k = \sum_i |b_i|$.

Example 3.3. Let $b_i = 1$, $b_j = 1$ and $b_l = -1$ and all other $b_k = 0$. Equation 4 can be written as,

$$\begin{aligned}
b_i b_j \|x_i - x_j\|_1 + b_i b_l \|x_i - x_l\|_1 + b_j b_l \|x_l - x_j\|_1 &\leq 0, \\
\text{i.e., } \|x_i - x_j\|_1 \leq \|x_i - x_l\|_1 + \|x_l - x_j\|_1, &
\end{aligned}$$

which is the well known triangle inequality.

Example 3.4. Consider a vector $b = (1, 1, 1, -1, -1) \in \mathbb{Z}^5$. Thus if a metric d is in ℓ_1, it must satisfy,

$$d_{12} + d_{23} + d_{13} + d_{45} \leq d_{14} + d_{24} + d_{34} + d_{15} + d_{25} + d_{35}. \tag{5}$$

Consider the bipartite graph $K_{2,3}$ with metric $d_{K_{2,3}}$. For this graph metric, LHS of Equation (5) is 8, while RHS is 6. Hence $K_{2,3}$ can not be isometrically embedded to ℓ_1, and the distortion must be at least $\frac{8}{6}$.

Remark 3.5. If a metric d satisfies the all the k-gonal inequalities, then d is a hypermetric. Therefore, ℓ_1 is a hypermetric.

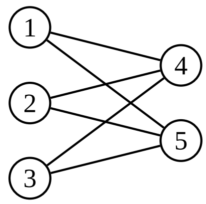


Figure 1: $K_{2,3}$ bipartite graph

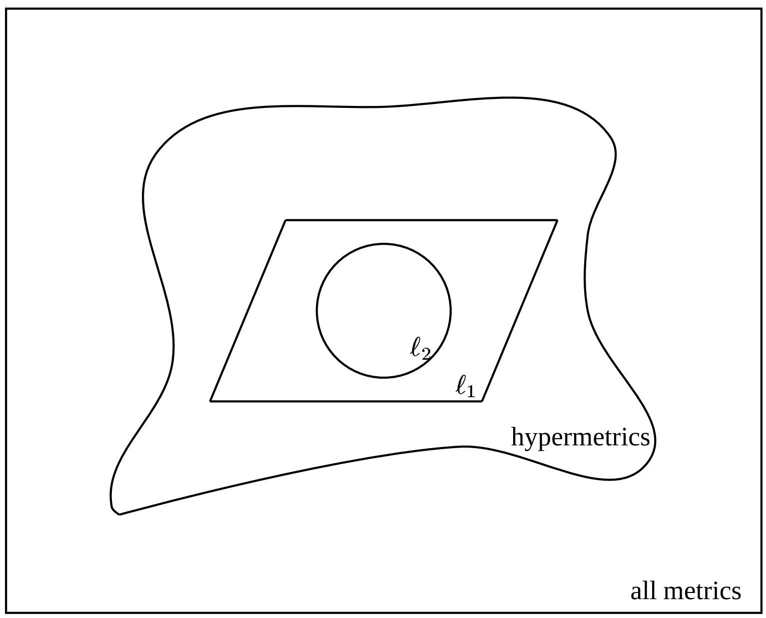


Figure 2: A schematic diagram showing that all ℓ_2 metrics are ℓ_1, which in turn are hypermetrics.

References

[Bou85] J. Bourgain. On lipschitz embedding of finite metric spaces in hilbert space. *Israel Journal of Mathematics*, 1985.

[LLR94] Nathan Linial, Eran London, and Yuri Rabinovich. The geometry of graphs and some of its algorithmic applications. In *35th Annual Symposium on Foundations of Computer Science*, pages 577–591, Santa Fe, New Mexico, 20–22 November 1994. IEEE.

[LR88] Tom Leighton and Satish Rao. An approximate max-flow min-cut theorem for uniform multicommodity flow problems with applications to approximation algorithms. In *29th Annual Symposium on Foundations of Computer Science*, pages 422–431, White Plains, New York, 24–26 October 1988. IEEE.