Non-Uniform Sparsest Cut and Metric Embeddings[1]

- The final cut problem we will look at is a generalization of the *sparsest cut* problem we saw last time: it's called the *non-uniform* sparsest cut problem, or the *generalized* sparsest cut problem. The input, as usual, is an undirected graph $G = (V, E)$ with non-negative costs $c(e)$ on edges. Along with this, we are now given *pairs of terminals* or *demand pairs*: $\{s_i, t_i\}$ with $1 \leq i \leq k$. The sparsity of a subset of vertices $S \subseteq F$ with respect to these demand pairs is defined to be the ***ratio*** $\Phi(S) := \frac{\sum_{e \in \partial S} c(e)}{|\{i\colon \{s_i,t_i\} \cap S = 1\}|}$. That is, it is the ratio of the cut edges and the number of demand pairs separated by this set S. This generalizes the notion from last time which can be seen as the special case of *all* pairs being demand pairs. The objective is to find the cut $S \subseteq V$ of minimum sparsity. We use Φ^*_G to denote this minimum value.

- ***Linear Programming Relaxation.*** The LP relaxation is similar to the uniform sparsest cut problem; the only difference is that the sum of distances between only the terminal pairs is set to 1.

$$\mathsf{lp} := \min \sum_{e=(u,v)\in E} c(e) d_{uv} \qquad \text{(General Sparsest Cut LP)}$$

$$d_{uw} \leq d_{uv} + d_{vw}, \quad \forall i \in F,\ \forall \{u, v, w\} \subseteq V \qquad (1)$$

$$d_{vv} = 0, \qquad \forall v \in V \qquad (2)$$

$$\sum_{i=1}^{k} d_{s_i t_i} = 1$$

- ***Cut Metrics, ℓ_1-metrics, and Metric Embeddings.*** The solution to (General Sparsest Cut LP) returns a distance function d between all pairs of vertices, and thus (V, d) forms a *finite metric space*; it is precisely a finite set of points equipped with a distance function. And this distance function minimizes a certain objective. We now notice that the sparsest cut problem (and in fact many cut problems) are precisely a question of finding a distance function d^* "from a certain special class" which minimizes the same objective. The approximation algorithm follows by taking the "general" metric returned by (General Sparsest Cut LP) and mapping/embedding into this "special class" that "warps" the distance function as little as possible.

  We now discuss this special class of metric/distance function. Fix a subset $S \subseteq V$. The ***elementary cut-metric*** (V, d_S) induced by this subset is simply defined as

$$d_S(u,v) = \begin{cases} 1 & \text{if } |S \cap \{u,v\}| = 1 \\ 0 & \text{otherwise} \end{cases}$$

[1] *Lecture notes by Deeparnab Chakrabarty. Last modified : 18th Mar, 2022*
These have not gone through scrutiny and may contain errors. If you find any, or have any other comments, please email me at deeparnab@dartmouth.edu. Highly appreciated!

Definition 1 (Cut Metric)**.** A metric/distance function $d : V \times V \to \mathbb{R}_{\geq 0}$ is a ***cut metric*** if it can be expressed as a *non-negative, linear* combination of elementary cut metrics. That is, there exists $\lambda_S \geq 0$ for all $S \subseteq V$ such that for any $u, v \in V \times V$, we have

$$d(u,v) = \sum_{S \subseteq V} \lambda_S d_S(u,v)$$

We use CUT to denote the cone[a] of all cut-metrics.

[a] A cone is a subset C of $\mathbb{R}^k$ where $x \in C$ implies $\alpha x \in C$ for any $\alpha \geq 0$, and $x \in C, y \in C$ implies $x + y \in C$. When V is finite, the distance function can be thought of as an $\binom{|V|}{2}$-dimensional vector.

The next observation asserts that the optimum sparsity is obtained by minimizing the "ratio function" over the special class of cut-metrics.

Observation 1. For any graph G,

$$\Phi^*(G) = \min_{d \in \mathsf{CUT}} \frac{\sum_{(u,v) \in E} c(u,v) d(u,v)}{\sum_{i=1}^{k} d(s_i, t_i)} \tag{3}$$

Exercise: ☕ *Prove the above. Note that Φ^* is precisely the minimum when d is an elementary cut metric. Show that taking non-negative linear combinations cannot decrease the RHS.*

The next observation connects cut metrics to more well-known metrics.

Definition 2 (ℓ_1-Metric)**.** A metric/distance function $d : V \times V \to \mathbb{R}_{\geq 0}$ is an ℓ_1 ***metric*** if there exists a map $\phi : V \to \mathbb{R}^h$ for some non-negative integer h such that for every pair of points $u, v \in V \times V$, we have

$$d(u,v) = \|\phi(u) - \phi(v)\|_1 = \sum_{i=1}^{h} \Big|\phi(u)[i] - \phi(v)[i]\Big|$$

We use $\mathcal{L}_1$ to denote the cone of all ℓ_1 metrics on V.

The next lemma shows that CUT and $\mathcal{L}_1$ are the same.

Lemma 1 (ℓ_1-Metric)**.** For any (V, d) with $d \in \mathsf{CUT}$, there is a mapping $\phi : V \to \mathbb{R}^h$ for some h, such that $\|\phi(u) - \phi(v)\|_1 = d(u,v)$ for all pairs u, v. Conversely, given a mapping $\phi : V \to R^h$, there exists $d \in \mathsf{CUT}$ such that $d(u,v) = \|\phi(u) - \phi(v)\|_1$ for all pairs u and v. The second mapping is polynomial time computable and has $\lambda_S > 0$ for at most nh sets.

Proof. Suppose $d \in \mathsf{CUT}$ and let $d = \sum_S \lambda_S d_S$. That is, for any pair u, v, $d(u,v)$ equals the sum of λ_S for all subsets S separating u and v. Now define a mapping on $h := |\{S : \lambda_S > 0\}|$ coordinates, as follows: $\phi(u)[S] := \lambda_S \cdot \mathbf{1}_{u \in S}$. That is, the Sth coordinate of $\phi(u)$ is λ_S if $u \in S$, otherwise it is

0. Note that for any pair u, v, $||\phi(u) - \phi(v)||_1$ equals the sum of λ_S for all S separating u and v. That is, $\|\phi(u) - \phi(v)\|_1 = d(u, v)$.

For the converse, we have a mapping $\phi : V \to \mathbb{R}^h$. For every coordinate, we associate h subsets as follows. Fix a coordinate i. Order the vertices as $\phi(u_1)[i] \leq \phi(u_2)[i] \leq \cdots$. The sets with positive S are precisely $\{u_1, \ldots, u_t\}$ for $t = 1, \ldots, h$, and define $\lambda_S = \phi(u_t)[i] - \phi(u_{t-1})[i] \geq 0$ for $S = \{u_1, \ldots, u_t\}$. Also define $\phi(u_0)[i] = 0$. It is easy to check that $d(u, v) = ||\phi(u) - \phi(v)||_1$. $\square$

Therefore, the above lemma and (3) implies $\Phi^* = \min_{d \in \mathcal{L}_1} \frac{\sum_{(u,v)\in E} c(u,v)d(u,v)}{\sum_{i=1}^k d(s_i,t_i)}$. The final definition we need is the following.

Definition 3 (Metric Embedding)**.** Given two metric spaces (V, d) and (U, ℓ), a mapping $\phi : V \to U$ is an ***embedding*** if it is injective, that is, $x \neq y \Rightarrow \phi(x) \neq \phi(y)$. This metric has *dilation* at most $\alpha \geq 1$ and *contraction* at most $\beta \geq 1$ if for any two vertices $u, v \in V \times V$, we satisfy

$$\forall u, v \in V \times V : \quad \frac{d(u, v)}{\alpha} \leq \ell(\phi(u), \phi(v)) \leq \beta \cdot d(u, v) \tag{4}$$

The parameter $\rho = \alpha \cdot \beta$ is the *distortion* of this embedding.

At times, we only care about the dilation/contraction only for a subset of vertices; that is, (4) holds only for $u, v \in S \times S$. We then say the metric embedding has distortion ρ when restricted to the subset S. Another thing before we move on. One could think of ϕ as an algorithm mapping V to U. Many embeddings are described as a *randomized* map/algorithm, and one has (4) holding only with high probability. First, if the probability of error $\ll \frac{1}{|V|^2}$, then union bounding and the probabilistic method implies there exists one embedding with distortion ρ. Second, for most algorithmic purposes, this randomized mapping suffices.

- Finally, we connect to generalized sparsest cut.

Theorem 1. Let (V, d) be a general metric on the vertices of the graph G. Let $S := \{s_i, t_i : 1 \leq i \leq k\}$. Suppose there is a metric embedding $\phi : V \to \mathbb{R}^h$ for $h = \text{poly}(n)$ into $\mathcal{L}_1$, with dilation[a] α wrt S, and contraction 1 with respect to V Furthermore, assume this embedding can be obtained in polynomial time. Then, there is an α-approximation to the non-uniform sparsest cut problem.

[a]it is allowed to be randomized and succeed with high probability

Proof. Given an instance of the non-uniform sparsest cut problem, solve (General Sparsest Cut LP) to obtain a metric space (V, d) with $\sum_{i=1}^k d(s_i, t_i) = 1$. Next, use the embedding algorithm which promises the metric embedding $\phi : V \to \mathbb{R}^h$ with the property that

$$\forall u, v \in S \times S \quad : \quad d(u, v) \leq \alpha \cdot \|\phi(u) - \phi(v)\|_1 \quad \text{and} \quad \forall u, v \in V \times V \quad : \quad d(u, v) \geq \|\phi(u) - \phi(v)\|_1$$

For the time being, let $\|\phi(u) - \phi(v)\|_1$ be called $\ell(u, v)$. Then observe that

$$\Phi(\ell) := \frac{\sum_{(u,v)\in E} c(u, v)\ell(u, v)}{\sum_{i=1}^k \ell(s_i, t_i)} \leq \alpha \cdot \frac{\sum_{(u,v)\in E} c(u, v)d(u, v)}{\sum_{i=1}^k d(s_i, t_i)} = \alpha \cdot \mathsf{lp}$$

Finally, we use Lemma 1 (conversely part) to obtain a cut metric $d \in \mathsf{CUT}$ with $\ell(u,v) = d(u,v) = \sum_{S \subseteq V} \lambda_S d_S(u,v)$ with S ranging over at most $nh = \mathrm{poly}(n)$ subsets. Choosing the S among them with smallest $\Phi(S)$ would have sparsity at most $\Phi(\ell)$ which is at most $\alpha \cdot \mathsf{lp}$. Note that if the embedding was randomized, then the set S returned would have sparsity $\leq \alpha \mathsf{lp}$ with probability $\geq 1 - \frac{1}{\mathrm{poly}(k)}$. □

- ***Metric Embedding Results.*** The above discussion, and in particular Theorem 1 delegates all the "hard work" to finding metric embeddings of a general metric into $\mathcal{L}_1$ with low distortion. But such questions have been studied by mathematicians for almost a century, and therefore one can "piggy-back" on such results. The following theorem of Bourgain (stylized to capture distortion with respect to S) immediately implies a $O(\log k)$-approximation for the general sparsest cut problem.

  Theorem 2 (Bourgain's Theorem, the Terminal Version)**.** Given any metric space (V, d) and a set $S \subseteq V$ of size at most k, there is a mapping $\phi : V \to \mathbb{R}^{O(\log^2 k)}$ such that with high probability, we have that for any pair of vertices u and v, $||\phi(u) - \phi(v)||_1 \leq d(u,v)$ and for any pair $u, v \in S$, $d(u,v) \leq O(\log k)||\phi(u) - \phi(v)||_1$.

  In the next lecture note, we see a proof of the above theorem using in fact one of the techniques we saw for the multicut problem.

Notes

After the seminal paper [4] of Leighton and Rao, there were many works on the generalized sparsest cut problem. The first was the paper [3] which gave an $O(\log C \log k)$ approximation where C is the sum of edge capacities. Any ρ-approximate algorithm for the multicut problem can be used to give an $O(\rho \log k)$-approximation for the generalized sparsest cut (it is a nice exercise to figure this out). This was observed in the paperr [2] by Garg, Vazirani and Yannakakis which also gave a $O(\log k)$-approximation for multicut. One could thus obtain an $O(\log^2 k)$-approximate algorithm for the generalized sparsest cut problem. See also the paper [6] for a more general $O(\log^2 k)$-approximation.

The idea of using metric embeddings to solve these style of cut problems are from the seminar paper [5] by Linial, London, and Rabinovich. This paper and also the paper [1] by Aumann and Rabani describe the $O(\log k)$-approximation to the generalized sparsest cut problem using the version of Bourgain's theorem stated above.

References

[1] Y. Aumann and Y. Rabani. An $O(\log k)$ approximate min-cut max-flow theorem and approximation algorithm. *SIAM Journal on Computing (SICOMP)*, 27(1):291–301, 1998. Prelim. Version, in IPCO 1995.

[2] N. Garg, V. V. Vazirani, and M. Yannakakis. Approximate max-flow min-(multi)cut theorems and their applications. *SIAM J. Comput.*, 25(2):235–251, 1996. Prelim. Version in STOC 1993.

[3] P. Klein, A. Agrawal, R. Ravi, and S. Rao. An approximate max-flow min-cut relation for undirected multicommodity flow, with applications. *Combinatorica*, 15:187–202, 1995. Prelim. Ver. in FOCS 1990.

[4] Leighton and Rao. Multicommodity Max-Flow Min-Cut Theorems and Their Use in Designing Approximation Algorithms. *Journal of the ACM*, 46, 1999. Prelim. Version. in FOCS 1988.

[5] N. Linial, E. London, and Y. Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–246, 1995. Prelim. Version. in FOCS 1994.

[6] S. Plotkin and E. Tardos. Improved bounds on the max-flow min-cut ratio for multicommodity flows. *Combinatorica*, 15:425–434, 1995. Prelim. Version in STOC 1993.