Advanced Approximation Algorithms **(CMU 18-854B, Spring 2008)**

Lecture 27: Algorithms for Sparsest Cut

Apr 24, 2008

Lecturer: Anupam Gupta *Scribe: Varun Gupta*

In lecture 19, we saw an LP relaxation based algorithm to solve the sparsest cut problem with an approximation guarantee of $O(\log n)$. In this lecture, we will show that the integrality gap of the LP relaxation is $O(\log n)$ and hence this is the best approximation factor one can get via the LP relaxation. We will also start developing an SDP relaxation based algorithm which provides an $O(\sqrt{\log n})$ approximation for the uniform sparsest cut problem (where demands between all pairs of vertices is $D_{ij} = 1$), and an $O(\sqrt{\log n}\log\log n)$ algorithm for the sparsest cut problem with general demands.

1 Problem Definition and LP relaxation review

Recall that the the sparsest cut problem is defined as follows. We are given an undirected graph $G = (V, E)$ with

- non-negative edge costs (or capacities) $c_e = c_{ij}$ for all $e = \{i, j\} \in \binom{V}{2}$,
- non-negative demands D_{ij} between every pair of vertices $\{i, j\} \in \binom{V}{2}$.

With the edge capacities and the demands, we can associate vectors $\bar{c}, \bar{D} \in \Re^{\binom{n}{2}}$ ($n = |V|$). The sparsest cut problem seeks to find

$$\Phi^* = \min_{S \subseteq V} \frac{\bar{c} \cdot \bar{\delta}_S}{\bar{D} \cdot \bar{\delta}_S} = \frac{\mathsf{cap}(S, \bar{S})}{\mathsf{dem}(S, \bar{S})} \tag{1}$$

where

$$\mathsf{cap}(S, \bar{S}) = \textstyle\sum_{i \in S, j \in \bar{S}} c_{ij}$$
$$\mathsf{dem}(S, \bar{S}) = \textstyle\sum_{i \in S, j \in \bar{S}} D_{ij}$$

and $\bar{\delta}_S \in \Re^{\binom{n}{2}}$ is the cut metric associated with S:

$$\delta_{Sij} = \begin{cases} 0 & \text{if } i, j \in S \text{ or } i, j \in \bar{S} \\ 1 & \text{otherwise} \end{cases}$$

To form the LP relaxation, we relax the requirement of minimizing over the cut metrics to minimizing over *all* metrics. That is,

$$\lambda^* = \min_{\text{metrics } d} \frac{\bar{c} \cdot \bar{d}}{\bar{D} \cdot \bar{d}}$$

The above relaxation is solved by the following linear program:

$$\begin{array}{rlll} \min & \sum_{i,j} c_{ij} d_{ij} & & \\ \text{subject to} & \sum_{i,j} D_{ij} d_{ij} & = 1 & \\ & d_{ij} + d_{jk} & \geq d_{ik} & \forall i,j,k \\ & d_{ij} & \geq 0 & \forall i,j \end{array} \tag{2}$$

Clearly, $\lambda^* \leq \Phi^*$. In lecture 19, we proved $\Phi^* \leq \lambda^* \cdot O(\log n)$ by embedding the metric returned by the LP into ℓ_1 with distortion $O(\log n)$.

2 Integrality gap for sparsest cut LP relaxation

A natural question to ask is, can we get a better approximation ratio than $O(\log n)$ using the LP relaxation? In this section we will see that the answer is *no*, since the LP relaxation has an integrality gap of $O(\log n)$.

Claim 2.1. *The integrality gap between* Φ^* *and* λ^* *is* $\Omega(\log n)$.

To prove the above claim, we will first take a small digression and introduce the *maximum concurrent flow* problem which takes the same input as the sparsest cut problem. We will show that the optimal value τ^* of the maximum concurrent flow problem is equal to the optimal value λ^* of the sparsest cut LP relaxation on the same input graph. Finally we will prove that the integrality gap of Φ^* and $\tau* = \lambda^*$ is $\Omega(\log n)$.

2.1 The maximum concurrent flow problem

Definition 2.2. Given an undirected graph $G = (V, E)$ with

- non-negative edge capacities $c_e = c_{ij}$ for all $e = \{i, j\} \in \binom{V}{2}$,
- non-negative demands D_{ij} between every pair of vertices $\{i, j\} \in \binom{V}{2}$.

the maximum concurrent problem seeks to maximize τ, such that we can send $\tau \cdot D_{ij}$ flow between vertices i and j *simultaneously* for all $\{i, j\} \in \binom{V}{2}$ while satisfying the edge capacity constraints.

Let τ^* denote the optimal value of the maximum concurrent flow problem. Consider a partition $(S, \bar{S})$ of V. The total flow crossing this partition is $\tau^* \cdot \mathsf{dem}(S, \bar{S})$ whereas the capacity of the partition is $\mathsf{cap}(S, \bar{S})$. Since we can't have more flow than the capacity,

$$\tau^* \cdot \mathsf{dem}(S, \bar{S}) \leq \mathsf{cap}(S, \bar{S}) \qquad\qquad \forall S \subseteq V$$

and hence,

$$\tau^* \leq \min_{S \subseteq V} \frac{\mathsf{cap}(S, \bar{S})}{\mathsf{dem}(S, \bar{S})} = \Phi^*$$

In fact, the maximum concurrent flow value τ^* is exactly the same as the optimum value Λ^* for the LP relaxation.

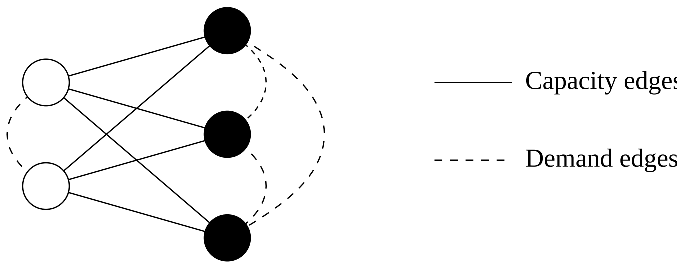


Figure 1: An instance of the sparsest cut/maximum concurrent flow problem.

Fact 2.3. *Given an instance of the sparsest cut problem, and maximum concurrent flow problem on the same input graph,* $\lambda^* = \tau^*$.

Proof. The proof follows by observing that the LP relaxation of sparsest cut problem and the LP for the maximum concurrent flow problem are duals of each other. The LP for the maximum concurrent flow has a variable $f(P)$ for each simple path in the input graph.

$$\begin{array}{rlll} \text{min} & \tau & & \\ \text{subject to} & \sum_{\substack{\text{paths } P \\ \text{between } i,j}} f(P) & \geq \tau \cdot D_{ij} & \forall i,j \\ & \sum_{\substack{\text{paths } P \\ \text{containing} \\ e=\{u,v\}}} f(P) & \leq C_{uv} & \forall u,v \\ & f(P) & \geq 0 & \forall P \end{array}$$

The dual has variables α_{ij} and β_{uv} corresponding to the two sets of contraints above:

$$\begin{array}{rlll} \text{max} & \sum_{ij} C_{uv}\beta_{uv} & & \\ \text{subject to} & \sum_{\{u,v\}\in P} \beta_{uv} & \geq \alpha_{ij} & \forall \text{ paths } P \text{ between } i,j \\ & \sum_{ij} D_{ij}\alpha_{ij} & \geq 1 & \\ & \alpha,\beta & \geq 0 & \end{array}$$

Now if we consider α_{ij} to be the distance between i, j and β_{uv} to be the edge length of the edge $\{u, v\}$, then this LP can be checked to be the same as λ^*. $\square$

2.2 Integrality Gaps

For example, consider the graph in Figure 1. The solid edges represent edges with capacity 1. The dotted edges represent pairs $\{i, j\}$ with $D_{ij} = 1$. Remaining capacities and demands are 0. Note that the value of the sparsest cut in the graph in Figure 1 is $\Phi^* = 1$ (choose any solid vertex as the set S). Furthermore, $\tau^* \leq \frac{3}{4}$. If send τ units of flow between every demand pair, the total volume of the flow is 8τ, since each path between a demand pair has two edges and there are four demand pairs. The total capacity is 6, and hence $\tau^* \leq \frac{3}{4}$. In fact $\tau^* = \frac{3}{4}$, by sending $\frac{1}{4}$ units of flow on each of the three paths between the white vertices and $\frac{3}{8}$ units on each of the two paths between every pair of solid vertices.

Theorem 2.4. *[7] The integrality gap between Φ^* and $\tau*$ ($= \lambda^*$) is $\Omega(\log n)$.*

Proof. Let $G = (V, E)$ be a constant degree expander with unit capacity on each edge $c_{i,j} = 1\ \forall\{i,j\} \in E$, and unit demand between every pair of vertices $D_{ij} = 1\ \forall i \neq j$. We will show that G has a large Φ^* but small τ^*.

(A)

$$\begin{aligned}\Phi^* &= \min_S \frac{\mathsf{cap}(S,\bar{S})}{\mathsf{dem}(S,\bar{S})} = \min_{|S|\leq\frac{n}{2}} \frac{\mathsf{cap}(S,\bar{S})}{|S||\bar{S}|} \\ &= \min_{|S|\leq\frac{n}{2}} \frac{\mathsf{cap}(S,\bar{S})}{|S|} \left(q \in \left[\frac{1}{n}, \frac{2}{n}\right]\right) \\ &= \Omega(1)\Omega\left(\frac{1}{n}\right)\end{aligned}$$

where the last step follows since $\min_{|S|\leq\frac{n}{2}} \frac{\mathsf{cap}(S,\bar{S})}{|S|}$ is the edge expansion which is $\Omega(1)$ for a constant degree expander.

(B) Recall the following claim, which follows from problem 7 in homework 5:

Claim 2.5. *In a constant degree expander (say degree=10), $\Omega(n^2)$ pairs of vertices are at a distance greater than $\frac{1}{10}\log n$.*

Since all D_{ij} are 1, at least $\tau \cdot \Omega n^2 \log n$ volume of flow is needed to send τ units of flow between these $\Omega(n^2)$ pairs. However, since the graph is a constant degree expander, total edge capacity is $O(n)$. Therefore,

$$\tau^* \leq O\left(\frac{1}{n\log n}\right).$$

This completes the proof of $\Omega(\log n)$ gap between Φ^* and τ^*, and hence of Claim 2.1. $\square$

3 SDP relaxation for sparsest cut

To obtain the LP relaxation, we had relaxed the requirement of minimizing over all cut metrics to minimizing over all metrics. To obtain the SDP relaxation we consider the following tighter relaxation:

$$\beta^* = \min_{d\,\in\,\text{metric}\cap\ell_2^2} \frac{\bar{c}\cdot\bar{d}}{\bar{D}\cdot\bar{d}}$$

Recall that an n-point metric d is in ℓ_2^2 (it is a "squared-Euclidean" metric) if there exist points $v_1, v_2, \cdots, v_n \in \Re^k$ such that the distances are

$$d_{ij} = \|v_i - v_j\|_2^2.$$

Note that the condition that the squared distances form a metric (i.e., satisfy the triangle inequality) is equivalent to saying that in the space $\Re^k$, none of the triangles between these n points have obtuse angles. Note that $\ell_2^2\cap$metric forms a convex cone. Some more properties of $\ell_2^2\cap$metric:

1. If $d \in \ell_1$, then $d \in \ell_2^2$. (Why?) This is what we require since we need to optimize over ℓ_1 and hence the feasible set of the SDP relaxation should be a superset of ℓ_1.

2. In $\Re^k$, we can have at most 2^k points with ℓ_2^2 metric (in fact, any negative type metric). This is achieved by the hypercube.

3. Given n points on the real line $\Re^1$ with the ℓ_1 metric, the ℓ_2^2 embedding of these points requires n dimensions—a new dimension for each point. (Use this to infer that $d \in \ell_1 \Rightarrow d \in \ell_2^2$.)

The SDP to compute β^* is given by:

$$\begin{array}{lrll} \min & \sum_{i,j} c_{ij} \|x_i - x_j\|^2 & & \\ \text{subject to} & \sum_{i,j} D_{ij} \|x_i - x_j\|^2 & = 1 & \\ & \|x_i - x_j\|^2 + \|x_j - x_k\|^2 & \geq \|x_i - x_k\|^2 & \forall i, j, k \\ & x_i & \in \Re^t & \forall i \end{array} \tag{3}$$

The approximation ratio of the SDP relaxation naturally depends on how well (low distortion) one can embed an ℓ_2^2 metric into ℓ_1. The following theorems give upper bounds on the integrality gap for the SDP relaxation (3).

Theorem 3.1 (Goemans, unpublished)**.** *If the SDP returns a solution in $\Re^k$, then the integrality gap is $O(\sqrt{k})$.*

Theorem 3.2 ([2])**.** *For the uniform sparsest cut problem ($D_{ij} = 1 \ \forall i \neq j$), the SDP integrality gap is $O(\sqrt{\log n})$.*

Theorem 3.3. *[1] For general sparsest cut, the SDP integrality gap is $O(\sqrt{\log n} \log \log n)$.*

The techniques used in proving above theorems are useful as tools to round SDP relaxations in minimizations problems (earlier we have seen rounding techniques for maximization problems).

Goemans, and independently, Linial made the following conjecture on the integrality gap of the SDP relaxation:

Conjecture 3.4. *[4, 8] The integrality gap between Φ^* and β^* is $\theta(1)$.*

The Goemans-Linial conjecture was first disproved by Khot and Vishnoi [5] who proved an $\Omega\left((\log \log n)^{1/6-\epsilon}\right)$ integrality gap for the non-uniform case. This was then improved to $\Omega(\log \log n)$ by Krauthgamer and Rabani [6]. For uniform sparsest cut, $\Omega(\log \log n)$ integrality gap was shown by Devanur, Khot, Saket and Vishnoi [3].

3.1 From SDP relaxation to sparse cuts

In this section we will see an important *structure theorem* and some intuition of how this structure theorem can lead to a $O(\sqrt{\log n})$ approximation for the sparsest cut problem; the proof will be given in the next lecture.

Lemma 3.5. ***Structure Lemma [2]:*** *Let $v_1, v_2, \cdots, v_n$ be points in the unit ball in $\Re^k$ satisfying $d_{ij} = \|v_i - v_j\|^2$ is a metric. Suppose the points satisfy the following "well-spread-out property":*

$$\tfrac{1}{n^2} \textstyle\sum_{i,j} d_{ij} \geq \delta = \Omega(1)$$

Then there exist disjoint sets S and T such that $|S|, |T| \geq \Omega(n)$ and

$$\min_{i \in S, j \in T} d_{ij} \geq \Omega\left(\tfrac{1}{\sqrt{\log n}}\right)$$

Intuition for the $O(\sqrt{\log n})$ approximation in the uniform case.

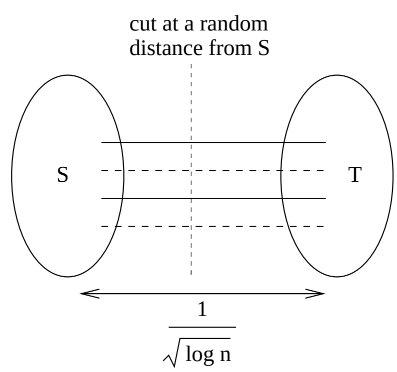
Since all demands are the same, we can scale the demands and set them to $D_{ij} = \frac{1}{n^2}\ \forall i \neq j$. Now the SDP ensures that

$$\sum D_{ij} d_{ij} = \sum_{ij} d_{ij} = \frac{1}{n^2} \sum d_{ij} = 1.$$

Now suppose we got lucky, and the SDP embedding lies on the unit ball, so that we can use the Structure lemma. (This will not happen in general: we'll give a rigorous proof next time.) If now we pick the S and T satisfying the Structure Lemma and cut at a random distance from S, the probability an edge e is cut is:

$$\mathbf{Pr}[\text{edge } e \text{ is cut}] = \frac{d_e}{\frac{1}{\sqrt{\log n}}} = d_e \sqrt{\log n}$$

Thus the expected total capacity crossing the cut is $\beta^* \cdot O(\sqrt{\log n})$. Furthermore, since S and T are both $\Omega(n)$ and the demands are equal, we lose a constant in the demand crossing the cut.



3.2 The Structure Lemma is Tight

A natural question is, can we tighten the Structure Lemma to obtain a better approximation? The answer to this is *no*: an example where the Structure Lemma is tight is the hypercube $\{-\frac{1}{\sqrt{K}}, \frac{1}{\sqrt{K}}\}^K$

where $K = \log_2 n$. The ℓ_2^2 distance between two vertices v_i and v_j is given by:

$$\begin{aligned} d_{ij} &= \|v_i - v_j\|^2 \\ &= \frac{4 \text{ Hamming dist}(i,j)}{K} \end{aligned}$$

To prove this, consider two sets S, T with $|S|, |T| = s$ for some parameter s to be specified soon. The vertex-isoperimetric inequality for the hypercube says that for all sets with size s, the set X that has the fewest neighbors outside X (i.e., the smallest $|N(X) \backslash X|$) is a ball around some vertex. Therefore, one such set X with the smallest vertex-expansion is the set $|X^*|$ containing exactly the s points closest to $\frac{1}{\sqrt{K}}\{-1, -1, \cdots, -1\}$. And hence $|S \cup N(S)| \geq |X^* \cup N(X^*)|$ for all $|S| = s = |X^*|$. Now suppose we choose:

$$s = \sum_{i \leq \frac{K}{2} - \sqrt{K \log\left(\frac{1}{\alpha}\right)}} \binom{K}{i}$$

then $s \approx \alpha n$ via tail bounds on the binomial distribution. Let S be any set of this size, then:

$$|S \cup N(S)| \leq |X^* \cup N(X^*)| = \sum_{i \leq \frac{K}{2} - \sqrt{K \log\left(\frac{1}{\alpha}\right)} + 1} \binom{K}{i}$$

Iterating this, if we define S_t to be all elements at Hamming distance t from S, we would have

$$|S_t| \leq \sum_{i \leq \frac{K}{2} - \sqrt{K \log\left(\frac{1}{\alpha}\right)} + t} \binom{K}{i}$$

For $t = \sqrt{K \log(1/\alpha)}$ this would be at least $n/2$.

Similarly, $|T_t|$ would be at least $n/2$ for the same value. Since both these sets contain at least half the elements, S_t intersects T_t, and hence S and T are $2t$-close in Hamming distance. But $t = O(\sqrt{K})$, which means that the ℓ_2^2 distance between S and T is $\frac{4 \cdot O(\sqrt{K})}{K} = O(\frac{1}{\sqrt{\log n}})$, which proves the fact that the Structure Lemma is tight up to constants.

3.3 Proving a Small Integrality Gap in the Uniform Case

We end with the following lemma due to Rabinovich [9].

Lemma 3.6. *[9] For the uniform sparsest cut problem $D_{ij} = 1\ \forall i \neq j$, suppose the metric d given by the SDP embeds into $\mu \in \ell_1$ such that,*

1. $\mu \leq d$
2. $\sum_{i,j} \mu_{ij} \geq \frac{\sum_{ij} d_{ij}}{\alpha}$,

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

Proof. The proof is very similar to the proof we saw for the Sparsest Cut problem in Lecture 19 that an embedding into ℓ_1 with distortion α implies an integrality gap of α. Since here we are dealing with the *uniform* case, we show that the average condition above suffices. Indeed,

$$\begin{aligned}\frac{\bar{c}\cdot\bar{\mu}}{\bar{D}\cdot\bar{\mu}} \leq \frac{\bar{c}\cdot\bar{d}}{\sum \mu_{ij}} &\leq \frac{\bar{c}\cdot\bar{d}}{\sum d_{ij}}\alpha \\ &= \frac{\bar{c}\cdot\bar{d}}{\bar{D}\cdot\bar{d}}\cdot\alpha \\ &= \beta^*\alpha\end{aligned}$$

□

In the next lecture, we will see a technique to embed the SDP metric into $\Re^1$ (and hence ℓ_1).

References

[1] S. Arora, J. Lee, and A. Naor. Euclidean distortion and the sparsest cut. In *Proc. 37th Symp. on Theory of Computation (STOC)*, pages 553–562, 2005.

[2] S. Arora, S. Rao, and U. Vazirani. Expander flows, geometric embeddings and graph partitioning. In *Proc. 36th Symp. on Theory of Computation (STOC)*, pages 222–231, 2004.

[3] N. Devanur, S. Khot, R. Saket, and N. Vishnoi. Integrality gaps for sparsest cut and minimum linear arrangement problems. In *Proc. 38th Symp. on Theory of Computation (STOC)*, pages 537–546, 2006.

[4] M. Goemans. Semidefinite programming in combinatorial optimization. *Mathematical Programming*, 79:143–161, 1997.

[5] S. Khot and N. Vishnoi. The unique games conjecture, integrality gap for cut problems and embeddability of negative type metrics into ℓ_1. In *Proc. 46th Symp. on Foundations of Computer Science (FOCS)*, pages 53–62, 2005.

[6] R. Krauthgamer and Y. Rabani. Improved lower bounds for embeddings into l_1. In *Proceedings of the 17th annual ACM-SIAM symposium on Discrete algorithm*, pages 1010–1017. ACM Press, 2006.

[7] F. T. Leighton and S. Rao. An approximate max-flow min-cut theorem for uniform multicommodity flow problems with applications to approximation algorithms. In *Proc. 29th Symp. on Foundations of Computer Science (FOCS)*, pages 422–431, 1988.

[8] N. Linial. Finite metric spaces - combinatorics, geometry and algorithms. In *Proc. International Congress of Mathematicians III*, pages 573–586, 2002.

[9] Y. Rabinovich. On average distortion of embedding metrics into the line and into ℓ_1. In *Proc. 35th Symp. on Theory of Computation (STOC)*, pages 456–462, 2003.