Massachusetts Institute of Technology Lecturer: Michel X. Goemans
18.409: Topics in TCS: Embeddings of Finite Metric Spaces
October 16, 2006 Scribe: Dan Stratila

Lecture 10

We first provide a clarification of the discussion in Lectures 8 and 9 for the case of weighted graphs, and then discuss the sparsest cut problem and negative type metrics.

1 Clarification on lectures 8 and 9

Recall that in the previous lectures we considered a planar graph metric (X, d) and, for any Δ, gave an embedding $f_\Delta : X \to \ell_2$ such that

$$\frac{\Delta}{32} \leq \|f_\Delta(u) - f_\Delta(v)\| \leq d(u,v), \quad \text{for} \quad 34\Delta \leq d(u,v) \leq 68\Delta. \tag{1}$$

In the unweighted case, we subdivided the edges to make sure that all original distances were at least $4 \times 34 = 136$, took[1] $\Delta = 4, 8, \ldots, 2^k$ with $k = \lceil \log \operatorname{diam}(G) \rceil + 1$, let the embedding be given by $f : x \to (f_4(x), f_8(x), \ldots, f_{2^k}(x))$, and showed that $\|f(u) - f(v)\| \geq \frac{d(u,v)}{2176}$. On the other hand, $\operatorname{diam}(G) < n$ (recall that n is the number of vertices in the original graph), hence

$$\|f(u) - f(v)\| = \sqrt{\sum_{\Delta=4,8,\ldots,2^k} \|f_\Delta(u) - f_\Delta(v)\|^2} \leq \sqrt{\log \operatorname{diam}(G)} \cdot d(u,v) < \sqrt{\log n} \cdot d(u,v). \tag{2}$$

This implies that f is an embedding into ℓ_2 with distortion $O(\sqrt{\log n})$. However, this does not work for weighted graphs, since $\operatorname{diam}(G)$ is no longer bounded by n.

We next show how to proceed in the weighted case. Let each edge e have a weight $w(e)$, and consider $\Delta = 4, \ldots, 2^k$. For each Δ, we introduce a graph G_Δ obtained by contracting all edges with $w(e) \leq \frac{\Delta}{n}$. Note that

$$0 \leq d_G(u,v) - d_{G_\Delta}(u,v) \leq \Delta, \tag{3}$$

and that d_{G_Δ} is still a planar graph metric. We then insert $w(e) - 1$ surrogate nodes on each edge e to obtain an unweighted graph. For each Δ, let f_Δ be a mapping from G_Δ satisfying property (1), and let $f : x \to (f_4(x), f_8(x), \ldots, f_{2^k}(x))$.

Observation 1. If $d_G(u,v) \leq \frac{\Delta}{n}$, then all edges on the shortest path from u to v in G have been contracted in G_Δ. Hence $f_\Delta(u) = f_\Delta(v)$. □

Observation 2. Assuming the notation of the previous lecture, with $u \in s_i$ and $v \in s_j$,

$$\begin{aligned}\|f_\Delta(u) - f_\Delta(v)\|^2 &\leq \sum_{A,\sigma} \gamma(A,\sigma)(\sigma(s_i) d_{G_\Delta}(u,A) - \sigma(s_j) d_{G_\Delta}(v,A))^2 \\ &\leq \sum_{A,\sigma} \gamma(A,\sigma)(d_{G_\Delta}(u,A) + d_{G_\Delta}(v,A))^2 \overset{(*)}{\leq} \sum_{A,\sigma} \gamma(A,\sigma)(2\Delta)^2 = (2\Delta)^2. \end{aligned} \tag{4}$$

Here, $(*)$ is due to the construction of A by "slicing" the vertex set of G_Δ in components of "width" at most Δ.

[1]The reason to start at $\Delta = 4$ is because the probabilistic analysis required Δ to be a multiple of 4.

We now upper bound the squared distance for $u, v \in V(G)$:

$$\|f(u) - f(v)\|^2 = \sum_{\Delta=4,\ldots,2^k} \|f_\Delta(u) - f_\Delta(v)\|^2 \stackrel{(*)}{=} \sum_{\substack{\Delta=4,\ldots,2^k \\ \Delta \le n d_G(u,v)}} \|f_\Delta(u) - f_\Delta(v)\|^2$$

$$\stackrel{(**)}{\le} \sum_{\substack{\Delta=4,\ldots,2^k \\ \Delta \le n d_G(u,v)}} \min\{d_G^2(u,v), (2\Delta)^2\} \stackrel{(***)}{\le} O(\log(n)) d_G^2(u,v). \quad (5)$$

Here, $(*)$ is due to Observation 1, and $(**)$ is due to Observation 2 and f_Δ being non-expanding. To see $(***)$ note that when $d_G^2(u,v)$ is smaller, we have $\frac{\Delta^2}{n^2} \le d_G^2(u,v) \le (2\Delta)^2$, which can occur for at most $O(\log n)$ terms. The terms when $(2\Delta)^2$ is smaller can be viewed as a decreasing geometric progression with the base term at most $d_G^2(u,v)$.

Observation 3. If $d_G(u,v) \ge 35\Delta$, then $d_{G_\Delta}(u,v) \ge 34\Delta$, hence $\|f_\Delta(u) - f_\Delta(v)\| \ge \frac{\Delta}{32}$.

To conclude, we find a lower bound, also for $u, v \in V(G)$:

$$\|f(u) - f(v)\|^2 \ge \sum_{\Delta=4,\ldots,2^k} \|f_\Delta(u) - f_\Delta(v)\|^2 \ge \sum_{\substack{\Delta=4,\ldots,2^k \\ d_G(u,v) \ge 35\Delta}} \|f_\Delta(u) - f_\Delta(v)\|^2$$

$$\ge \sum_{\substack{\Delta=4,\ldots,2^k \\ d_G(u,v) \ge 35\Delta}} \left(\frac{\Delta}{32}\right)^2 \ge \left(\frac{d_G(u,v)}{2 \cdot 35 \cdot 32}\right)^2. \quad (6)$$

The last inequality holds because we can assume that $d_G(u,v) \ge 35$ for all u and v, and for the last term in the summation we will have $35\Delta \le d_G(u,v) \le 2 \cdot 35\Delta$. Taking square roots of both bounds, we obtain that f is an embedding into ℓ_2 with distortion $O(\sqrt{\log n})$.

A more general technique to combine embeddings for different scales (different values of Δ) was obtained by Lee [Lee05].

2 Sparsest cut and negative type metrics

Recall the *uniform sparsest cut problem*:

$$\text{Given a graph G, find } \beta(G) = \min_{\emptyset \ne S \subsetneq V(G)} \frac{\delta(S, \overline{S})}{|S| \cdot |\overline{S}|}. \quad (7)$$

Leighton and Rao [LR88] obtained an $O(\log n)$-approximation algorithm for this problem. Linial, London, and Rabinovich [LLR95] obtained an $O(\log n)$-approximation algorithm for the non-uniform version of the problem (with general demands). In the remainder of this lecture, we focus on the uniform problem.

As before, let $\alpha(G)$ be the largest fraction of multicommodity flow that can be sent given unit capacities on edges and one unit-demand commodity between each pair of vertices. Remember that

$$\beta(G) = \min_{(V,d)\ \ell_1\text{-embeddable}} \frac{\sum_{(i,j)\in E} d(i,j)}{\sum_{i<j} d(i,j)}, \quad (8)$$

and that

$$\alpha(G) = \min_{(V,d)\ \ell_\infty\text{-embeddable}} \frac{\sum_{(i,j)\in E} d(i,j)}{\sum_{i<j} d(i,j)}. \tag{9}$$

For more details on these definitions and properties, see Lecture 4.

The quantity $\alpha(G)$ can be computed by linear programming, which suggests that we can approximate $\beta(G)$ by taking the corresponding ℓ_∞ embedding d, and embedding it into ℓ_1 through Bourgain's result. This yields

$$\frac{c}{\log n} d(u,v) \leq l(u,v) \leq d(u,v), \tag{10}$$

and thus $\alpha(G) \leq \beta(G) \leq O(\log n)\alpha(G)$.

We would like to to obtain a better approximation for $\beta(G)$ by considering a class of metrics that would yield a tighter bound than ℓ_∞ but would still allow us to optimize in polynomial time. For this purpose, we consider negative-type metrics.

Definition 1. *A finite metric (X,d) is of negative type if $(X,\sqrt{d})$ is ℓ_2-embeddable.*

For background on negative-type metrics, see [Sch38a, Sch38b]. Note that taking $d(u,v) = \|u-v\|_2^2$ does not necessarily yield a metric. For three points u,v,w the inequality

$$\|u-v\|_2^2 + \|v-w\|_2^2 \geq \|v-w\|_2^2 \tag{11}$$

holds if and only if $\angle uvw \leq \frac{\pi}{2}$. Hence, taking $X \subset \mathbb{R}^k$ with all angles non-obtuse and $d(u,v) = \|u-v\|_2^2$ for any $u,v \in X$ yields a negative-type metric. An example of a set with all angles non-obtuse is $X = \{0,1\}^k$, the set of vertices of the unit hypercube.

The following lemma was conjectured by Erdös in 1948 and proven by Danzer and Grünbaum [DG62].

Lemma 1. *If $X \subset \mathbb{R}^k$ and no three points in X form an obtuse angle, then $|X| \leq 2^k$.*

An elegant proof of this lemma is presented in [AZ04], but the reader is encouraged to prove it independently first (hint: try a volume argument). Here is another characterization of negative type metrics.

Lemma 2. *A finite metric (X,d) is of negative type if and only if for any set $\{b_i \in \mathbb{R} : i \in X\}$ with $\sum_{i\in X} b_i = 0$ we have*

$$\sum_{i\in X}\sum_{j\in X} b_i b_j d(i,j) \leq 0. \tag{12}$$

Proof sketch of $\Leftarrow$. Use the fact that $(X,\sqrt{d})$ is ℓ_2-embeddable to obtain that $(d(i,n) + d(j,n) - d(i,j))_{1\leq i,j\leq n-1} \succeq 0$ (see Lecture 2) and pre and post-multiply to obtain the characterization. □

Note that if (X,d) is a negative-type metric and $\alpha > 0$, then $(X,\alpha d)$ is also of negative type.

Lemma 3. *If (X,d_1) and (X,d_2) are of negative type, then so is (X,d_1+d_2).*

Proof sketch. Take the ℓ_2 embeddings of $(X,\sqrt{d_1})$ and $(X,\sqrt{d_2})$ and combine the dimensions. □

This implies that negative-type metrics form a (convex) cone. On the other hand, recall that ℓ_1 metrics also form a cone with cut metrics as its extreme rays.

Lemma 4. *Cut metrics are of negative type.*

Proof. Let the cut metric be (X, d) and the cut defining it given by $S \subset X$. For any b_i-s, we have

$$\sum_{i \in X} \sum_{j \in X} b_i b_j d(i,j) = 2 \sum_{i \in S} \sum_{j \in \overline{S}} b_i b_j \leq 0, \tag{13}$$

because $\sum_{i \in S} b_i + \sum_{i \in \overline{S}} b_i = 0$ implies one of the sums is less than or equal to zero, and the other is greater than or equal to zero. □

Therefore, the cone of ℓ_1 metrics is contained in the cone of negative-type metrics. Note that the negative-type metric induced by squared l_2 distance in the hypercube is ℓ_1 embeddable. There are, however, negative type metrics which are not l_1-embeddable. Here is one such construction from lattices.

Let

$$v_1, \ldots, v_k \in \mathbb{R}^k \quad \text{and} \quad L = \left\{ \sum_{i=1}^{k} \lambda_i v_i : \lambda_i \in \mathbb{Z} \right\}. \tag{14}$$

Consider a sphere $B(a, r)$ such that its interior $\dot{B}(a, r) = \{x : \|x - a\| < r\}$ does not intersect the lattice (a so-called empty sphere), and let $X = B(a, r) \cap L$. Note that no three points in X form an obtuse angle. Indeed, if $\angle uvw > \frac{\pi}{2}$, then $u - v + w$ is in both the lattice and in $L \cap \hat{B}(a, r)$, which is a contradiction. For some lattices (and some empty spheres), the squared ℓ_2 metric induced by X is not ℓ_1 embeddable.

Additional examples of negative type metrics not embeddable in ℓ_1 include shortest path metrics on some graphs (e.g. K_9 minus two adjacent edges). To learn more about negative type metrics, see the book by Deza and Laurent [DL97].

3 Computing $\gamma(G)$ and a lower bound

Let

$$\gamma(G) = \min_{(X,d) \text{ is of neg. type}} \frac{\sum_{(i,j) \in E} d(i,j)}{\sum_{i<j} d(i,j)}. \tag{15}$$

Since the cone of ℓ_1 metrics is contained in the cone of negative type metrics, $\beta(G) \geq \gamma(G)$.

Next, we show how to compute $\gamma(G)$ using semidefinite programming. Recall that for a symmetric $n \times n$ matrix X, the following are equivalent:

1. X is positive semidefinite, denoted $X \succeq 0$.

2. The eigenvalues λ_i are nonnegative.

3. $a^T X a \geq 0$ for any $a \in \mathbb{R}^n$.

A semidefinite program has the form

$$\min\ C \circ X, \tag{16a}$$
$$\text{s.t. } A_i \circ X = b_i, \quad i = \overline{1,p}, \tag{16b}$$
$$X \succeq 0. \tag{16c}$$

Here C and A_i are symmetric $n \times n$ matrices, and $A \circ B = \sum_{i=1}^n \sum_{j=1}^n a_{ij} b_{ij} = \operatorname{tr}(A^T B)$ is the Frobenius product. (Although this is not specified in the constraints, X can be assumed symmetric due A and B being symmetric. Also note that linear programming can be obtained as a special case by restricting non-diagonal entries of X to be zero.)

Since negative type metrics form a cone, we can normalize by $\sum_{i<j} d(i,j) = 1$. The following semidefinite program yields $\gamma(G)$:

$$\min \sum_{(i,j)\in E} d(i,j), \tag{17a}$$
$$\text{s.t. } \sum_{i<j} d(i,j) = 1, \tag{17b}$$
$$d(i,j) + d(j,k) \geq d(i,k), \qquad i,j,k \in V(G), \tag{17c}$$
$$(d_{ij} + d_{jn} - d_{ij})_{1\leq i,j\leq n-1} \succeq 0. \tag{17d}$$

Since we can compute ϵ-approximate and ϵ-feasible solutions to semidefinite programs in time polynomial in the input data size and $\log \frac{1}{\epsilon}$, the same is true for $\gamma(G)$.

We conclude by comparing $\gamma(G)$ to the classical eignevalue bound for the uniform sparsest cut problem. Since

$$\gamma(G) = \min_{\substack{v_1,\ldots,v_n\in\mathbb{R}^n \\ \text{all angles } \leq \pi/2}} \frac{\sum_{(i,j)\in E} \|v_i - v_j\|^2}{\sum_{i<j} \|v_i - v_j\|^2}, \tag{18}$$

a lower bound is given by

$$\delta(G) = \min_{v_1,\ldots,v_n\in\mathbb{R}^n} \frac{\sum_{(i,j)\in E} \|v_i - v_j\|^2}{\sum_{i<j} \|v_i - v_j\|^2} \overset{(*)}{=} \min_{\substack{x_1,\ldots,x_n\in\mathbb{R} \\ x_1+\cdots+x_n=0}} \frac{\sum_{(i,j)\in E} |x_i - x_j|^2}{\sum_{i<j} |x_i - x_j|^2}$$
$$= \min_{x\in\mathbb{R}^n,\ x^t e=0} \frac{x^T L_G x}{n x^T x} = \frac{1}{n}\mu_2(G). \tag{19}$$

Here $(*)$ holds because we can separate coordinate-by-coordinate and shift points. As usual, L_G denotes the Laplacian with

$$(L_G)_{ij} = \begin{cases} -1, & (i,j) \in E, \\ \deg_G(i), & i = j, \\ 0, & \text{otherwise.} \end{cases} \tag{20}$$

This lower bound is good for an expander graph, but in general can be arbitrarily bad. For example, consider the cycle C_n:

$$\beta(C_n) = \frac{2}{\left(\frac{n}{2}\right)^2} = \frac{8}{n^2}, \tag{21}$$

$$\mu_2(L_{C_n}) = 2 - 2\cos\left(\frac{2\pi}{n}\right) = \Theta\left(\frac{1}{n^2}\right). \tag{22}$$

Therefore, $\delta(C_n) = \Theta\left(\frac{1}{n^3}\right)$.

$\gamma(G)$ is thus no worse than both the linear programming bound $\alpha(G)$ and the eigenvalue bound $\delta(G)$. And these two bounds appear to have bad examples which appear to be good for the other one, and this intuitively explains why $\gamma(G)$ will turn out to be better in the worst case.

References

[AZ04] Martin Aigner and Günter M. Ziegler. *Proofs from The Book*. Springer-Verlag, Berlin, third edition, 2004. Including illustrations by Karl H. Hofmann.

[DG62] L. Danzer and B. Grünbaum. Über zwei Probleme bezüglich konvexer Körper von P. Erdős und von V. L. Klee. *Math. Z.*, 79:95–99, 1962.

[DL97] Michel Marie Deza and Monique Laurent. *Geometry of cuts and metrics*, volume 15 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 1997.

[Lee05] James R. Lee. Distance scales, embeddings, and metrics of negative type. In *Proceedings of SODA*, 2005.

[LR88] T. Leighton and S. Rao. An approximate max-flow min-cut theorem for uniform multicommodity flow problems with applications to approximation algorithms. In *Proceedings of the 29th Annual Symposium on Foundations of Computer Science*, pages 422–431, 1988.

[LLR95] Nathan Linial, Eran London, and Yuri Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–245, 1995.

[Sch38a] I. J. Schoenberg. Metric spaces and completely monotone functions. *Ann. of Math. (2)*, 39(4):811–841, 1938.

[Sch38b] I. J. Schoenberg. Metric spaces and positive definite functions. *Trans. Amer. Math. Soc.*, 44(3):522–536, 1938.