Seminar on Algorithms and Geometry

Lecture 2

March 26, 2009

Lecturer: Robert Krauthgamer Scribe by: Liah Kor Updated: April 1, 2009

1 Embedding into ℓ_∞

Theorem 1 (Matousek 96', based on Bourgain 85') *Let $q \geq 1$ be an integer, then any n-point metric space,$(X;d)$ embeds with distortion $2q-1$ into ℓ_∞^k, where $k = O(qn^{\frac{1}{q}} \log n)$.*

Proof Let $p = n^{-\frac{1}{q}}$. For each $j \in \{1 \dots q\}$let $p_j = \min\{p^j, \frac{1}{2}\}$. Let $\{A_{j,1} \dots A_{j,m}\} \subseteq X$ where $m = 24 \cdot n^{\frac{1}{q}} \ln n$, be chosen at random by including each element with probability p_j. Consider f, the corresponding Frechet embedding:

$$\text{f:X}\hookrightarrow \ell_\infty^{q\cdot m} \text{ such that } \ f_{j,i} = d(x, A_{j,i})$$

We'll show using the following lemmas, that f embeds $(X;d)$ into ℓ_∞^k where $k = q \cdot m = qn^{\frac{1}{q}} \log n$ with distortion $2q-1$.
Note that: $\forall x, y \in X, \|f(x) - f(y\| \leq \max_{i\in[1\dots m], j\in[1\dots q]} |f_{j,i}(x) - f_{j,i}(y)| \leq d(x,y)$ (the last inequality is implied by Fact 5 in lecture 1).
Define $\Delta \stackrel{def}{=} \frac{1}{2q-1} d(x,y)$

Lemma 2 *For every $x, y \in X$ exists $j \in [1 \dots q]$ s.t $\forall i \in [1 \dots m]$ with probability $\frac{P}{12}$*

$$|d(x, A_{j,i}) - d(y, A_{j,i})| \geq \Delta$$

Assuming the lemma, for every $x, y \in X$ exist i, j s.t with probability $1 - \frac{1}{n^2}$:

$$\|f(x) - f(y)\|_\infty \geq |f_{j,i}(x) - f_{j,i}(y)| \geq \Delta$$

Applying union bound over all $\binom{n}{2}$ pairs (x, y), we get that with probability $\geq \frac{1}{2}$ for all $x, y \in X$, $\|f(x) - f(y)\|_\infty \geq \Delta = \frac{1}{2q-1} d(x,y)$. ■

Proof of Lemma 2 Define $B(x, r)$ the ball of radius r around $x \in X$. Consider the sequence $B_0 = B(x, 0), B_1 = B(y, \Delta), B_2 = B(x, 2\Delta), B_3 = B(y, 3\Delta) \dots B_q = B(x \text{ or } y, q\Delta)$.

Claim 3 *Exist $j \in [1 \dots m]$ and $t \in [0 \dots q-1]$ such that $|B_t| \geq n^{\frac{j-1}{q}}$ and $|B_{t+1}| \leq n^{\frac{j}{q}}$*

Proof Idea Note that $1 \le |B_i| \le n$. Consider the partitioning of $[1,n]$ into q intervals $1, n^{\frac{1}{q}}, n^{\frac{2}{q}}, \dots n^{\frac{q-1}{q}}, n$. Use counting argument for the case $|B_i|$ is increasing. Otherwise, use the existence of t s.t $|B_t| > |B_{t+1}|$. ■

Let B_t and B_{t+1} as in the claim. For $|d(x, A_{j,i}) - d(y, A_{j,i})|$ to be at least Δ it suffices that $A_{j,i}$ contains at least one point from B_t and no points from the open ball $\mathring{B}_{t+1}$. By using the claim, we can bound the probabilities of the events to be at least $\frac{P}{3}$ and $\frac{1}{4}$ respectively. Using the fact the events are independent we get:

$$Pr[|d(x, A_{j,i}) - d(y, A_{j,i})| \ge \Delta] \ge \frac{P}{12}$$

■

Corollary 4 *Every n-point metric $(X; d)$ embeds into l_2 (and l_1) with distortion $O(\log^2 n)$*

Proof Idea Consider the same embedding f as embedding into l_2^k with $q = \log n$ which implies $k = O(\log^2 n)$ ■

Remark In general $l_2 \subseteq l_1$, i.e every n-point metric that embeds isometrically into l_2 also embeds isometrically into l_1.

An optimal bound is obtained in the following theorem:

Theorem 5 (Bourgain 85') *Every n-point metric $(X; d)$ embeds into l_2 with distortion $O(\log n)$*

2 Sparsest Cut

The input of the *Sparsest Cut* problem is a graph $G = (V, E)$ s.t $|V| = n$, and k pairs of vertices called demand pairs $\{s_1, t_1\}, \dots, \{s_k, t_k\} \subseteq V$.
Given such a graph and $S \subseteq V$ we can define the notion of *sparsity* for the cut $(S, \overline{S})$.

Definition 6

$$\text{sparsity} \stackrel{\text{def}}{=} \frac{|\{(\mathrm{u},\mathrm{v}) \in \mathrm{E} : |(\mathrm{u},\mathrm{v}) \cap \mathrm{S}| = 1\}|}{|\{\mathrm{i} : |(\mathrm{s_i},\mathrm{t_i}) \cap \mathrm{S}| = 1\}} = \frac{\textit{number of edges crossing the cut}}{\textit{number of demand pairs separated by the cut}}$$

The objective of the *Sparsest Cut* problem is to find a cut S that minimizes the sparsity.
Remark A special case of the sparsest cut problem is *uniform demand* in which there are $\binom{n}{2}$ demand pairs (demand pair for each pair of vertices).
Denote by $e(S, \overline{S})$ the number of edges crossing a cut S.
$sparsity_{\text{ uniform demand}} = \frac{e(S,\overline{S})}{|S| \cdot |\overline{S}|}$
Without loss of generality $|S| \le \frac{|V|}{2} = \frac{n}{2}$ and $\frac{n}{2} \le |\overline{S}| \le n$ thus after scaling the problem is equivalent to minimizing $\frac{e(S,\overline{S})}{|S|}$ (up to a factor of 2).

2.1 Formulation as a Linear Program

Definition 7 *A given cut $(S, \overline{S})$ may be thought of as a Cut Metric $d(\cdot,\cdot)$ defined by:*

$$f(x) = \begin{cases} 1 & x \in S \\ 0 & x \notin S \end{cases}$$

$$\forall u, v \in V, d(u,v) = |f(u) - f(v)|$$

The sparsest cut problem can be formulated as the following optimization problem:

$$OPT = \min \frac{\sum_{(u,v)\in E} d(u,v)}{\sum_{i=1}^{k} d(s_i, t_i)}$$

$$\text{subject to: } d(\cdot,\cdot) \text{ is a cut metric}$$

2.1.1 Relaxation to a linear program

$$LP = \min \sum_{(u,v)\in E} d(u,v) \text{ s.t:} \begin{cases} \sum_{i=1}^{k} d(s_i, t_i) = 1 \\ \text{d}(\cdot,\cdot) \text{ is a metric} \end{cases}$$

Since the relaxation allows also solutions which are not a cut-metric LP $\leq$ OPT.
The relaxation of the sparsest set problem is a linear program,hence it can be solved in polynomial time.

2.1.2 Approximation of Sparsest Cut using LP relaxation

An interesting question is can we "round" a solution of the LP to a solution of the sparsest cut problem (i.e find the corresponding cut)?
A first step of the rounding will be given by Bourgain's theorem which establishes an embedding of the metric d found by the LP into l_1 with distortion $O(\log n)$.
So exists $f : V \to l_1$ s.t $\forall u, v \in V,\ \ d(u,v) \leq \|f(u) - f(v)\|_1 \leq O(\log n) \cdot d(u,v)$
We get: $\sum_{(u,v)\in E} \|f(u) - f(v)\|_1 \leq O(\log n) \cdot LP$ where LP is the solution found by the linear program relaxation. By normalization we can also meet the condition
$\sum_{i=1}^{k} \|f(s_i) - f(t_i)\|_1 = 1$

Theorem 8 *For every n-point metric (X, d) that embeds isometrically to l_1, there exist $\alpha_i \geq 0$ and cut metrics τ_i s.t $\forall x, y \in X\ d(x,y) = \sum_i \alpha_i \tau_i(x,y)$.*