On Discrete Isoperimetry Problems II (Higher Order Cheeger Inequalities and Metric Embedding)

Amir Daneshgar
Sharif University of Technology

Ramin Javadi
Isfahan University of Technology

http://www.sharif.ir/~daneshgar
http://rjavadi.iut.ac.ir

4 September 2014 (13 Shahrivar 1393)

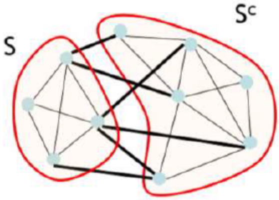
Part I.

Higher Cheeger Inequalities

Isoperimetric numbers

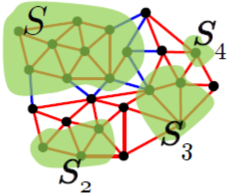
Given a graph $G = (V, E)$
For every $S \subset V$ define

$$\Phi(S) = \frac{|E(S, S^c)|}{|S|}.$$



For every $1 \leq k \leq n$,
define kth isoperimetric number as

$$\iota_k(G) = \min_{\substack{S_1,\ldots,S_k \subset V \\ \text{disjoint}}} \max_{1 \leq i \leq k} \{\Phi(S_i)\}.$$



All arguments could be generalized to weighted graphs.

Laplacian

A. Daneshgar
R. Javadi

For graph G, define normalized Laplacian of G as

$$L := I - D^{-1}A,$$

where A is the adjacency matrix of G and D is the diagonal matrix of the degrees.
A is positive semidefinite with spectrum

$$0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n.$$

For every function $f : V \to \mathbb{R}$, define Rayleigh quotient of f as

$$\mathcal{R}(f) = \frac{\langle Lf, f\rangle}{\langle f, f\rangle} = \frac{\sum_{u\sim v} |f(u) - f(v)|^2}{\sum_v |f(v)|^2}.$$

$$\lambda_2 = \min_{f \perp \mathbf{1}} \mathcal{R}(f).$$

Cheeger's Inequality

Theorem (Alon-Millman 85)

$$\frac{\lambda_2}{2} \leq \iota_2(G) \leq \sqrt{2\lambda_2}.$$

- G is disconnected if and only if $\lambda_2 = 0$.
- λ_2 is called algebraic connectivity.

Higher Cheeger's inequality

In general,

$$\lambda_k = \min_{\substack{f_1,\dots,f_k \\ \text{orthogonal}}} \max_{\substack{f \in \text{span}(f_1,\dots,f_k) \\ f \neq 0}} \mathcal{R}(f).$$

Conjecture (Daneshgar, J. 10, Daneshgar, J., Miclo 12)

$$\frac{\lambda_k}{2} \leq \iota_k(G) \leq c(k)\sqrt{\lambda_k},$$

where $c(k)$ depends only on k.

If f is the characteristic function of $S \subset V$, then

$$\mathcal{R}(f) = \Phi(S).$$

Thus, the l.h.s. is clearly true.

Dirichlet Cheeger inequality

$$\mathrm{supp}(f) := \{v \in V \mid f(v) \neq 0\}.$$

Theorem (Chung 96)

For every function $f : V \to \mathbb{R}$, there exists a set $S \subset \mathrm{supp}(f)$ such that

$$\Phi(S) \leq \sqrt{2\mathcal{R}(f)}.$$

If $f_1, \ldots, f_k$ are orthonormal eigenfunctions corresponding to $\lambda_1, \ldots, \lambda_k$, one may find subsets $S_1, \ldots, S_k \subset V$ such that

$$\max_i \Phi(S_i) \leq \sqrt{2\lambda_k}.$$

But the problem is that S_i's are not necessarily disjoint.

Miclo's conjecture

Miclo's conjecture (Miclo 08)

For any graph G and any number k, there exist disjointly supported functions $\psi_1, \ldots, \psi_k : V \to \mathbb{R}$ such that

$$\max_{1 \le i \le k} \mathcal{R}(\psi_i) \le c'(k)\lambda_k,$$

where $c'(k)$ depends only on k.

Miclo's Conjecture $\Rightarrow$ Higher Cheeger inequality

Known results (Lee, Gharan, Trevisan 14)

- Miclo's conjecture is true with $c'(k) = O(k^6)$, i.e. there exist disjointly supported functions $\psi_1, \dots, \psi_k$ such that

$$\max_{1 \le i \le k} \mathcal{R}(\psi_i) \le O(k^6)\lambda_k.$$

- Higher Cheeger inequality holds with $c(k) = O(k^3)$ (it is improved to $O(k^2)$).

$$\iota_k(G) \le O(k^2)\sqrt{\lambda_k}.$$

- For every graph G and every number k,

$$\iota_k(G) \le O(\sqrt{\lambda_{2k} \log k}).$$

- If G is planar, then

$$\iota_k(G) \le O(\sqrt{\lambda_{2k}}).$$

Proof of Miclo's conjecture

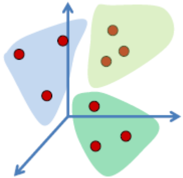
Let $f_1, \ldots, f_k$ be the orthonormal system of eigenfunctions with eigenvalues $\lambda_1, \ldots, \lambda_k$.
Define $F : V \to \mathbb{R}^k$, by

$$F(v) = (f_1(v), \ldots, f_k(v)).$$

Then

$$\mathcal{R}(F) = \frac{\sum_{u \sim v} \|F(u) - F(v)\|^2}{\sum_v \|F(v)\|^2} \leq \lambda_k.$$

The idea is to localize F on k disjoint regions to produce disjointly supported functions $\psi_1, \ldots, \psi_k$, whose Rayleigh quotients are not much far from $\mathcal{R}(F)$.



Spectral clustering

Let $f_1, \ldots, f_k$ be the orthonormal system of eigenfunctions with eigenvalues $\lambda_1, \ldots, \lambda_k$.

Spectral clustering consists of

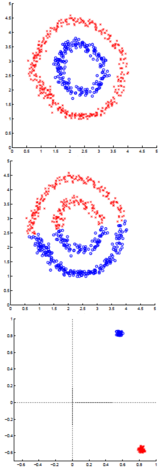
1. Computing the matrix

$$M = (f_1 \cdots f_k).$$

2. Considering each row of M as a vector in $\mathbb{R}^k$.
3. Use a clustering algorithm e.g. k-means to cluster these vectors.

There is no quantitative bound for approximation of $k-$means.

We use random partitioning instead of $k-$means.



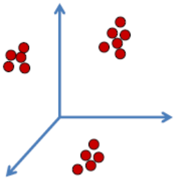
Isotropy

Let $f_1, \ldots, f_k$ be orthonormal functions.

$$M = (f_1 \cdots f_k), \quad F(v) = (f_1(v), \ldots, f_k(v)), \forall\ v \in V$$

Total Mass of F:

$$\sum_{v \in V} \|F(v)\|^2 = k.$$



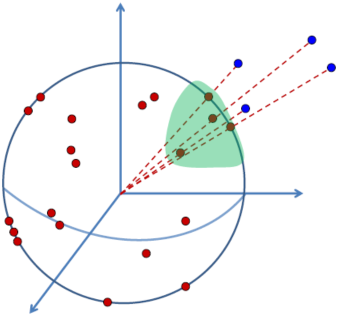
For every unit vector $x \in \mathbb{R}^k$,

$$\sum_{v \in V} \langle x, F(v) \rangle^2 = x^T M^T M x = \|x\|^2 = 1.$$

Radial projection metric

Define radial projection metric on V as follows,

$$d_F(u,v) = \left\| \frac{F(u)}{\|F(u)\|} - \frac{F(v)}{\|F(v)\|} \right\|$$



Radial projection metric

Define radial projection metric on V as follows,

$$d_F(u,v) = \left\| \frac{F(u)}{\|F(u)\|} - \frac{F(v)}{\|F(v)\|} \right\|$$

1) $\forall\ u, v \in V \quad \|F(u)\| \leq \frac{2}{d_F(u,v)} \|F(u) - F(v)\|.$

2) Spreading property: For every $S \subset V$,

$$\operatorname{diam}(S) \leq \Delta \ \Rightarrow \sum_{v \in S} \|F(v)\|^2 \leq \frac{1}{k(1-\Delta^2)} \sum_{v \in V} \|F(v)\|^2.$$

Localizations

A. Daneshgar
R. Javadi

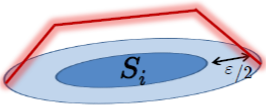
The goal is to find disjoint subsets $S_1, \ldots, S_k \subset V$ such that

1. Mass: $\sum_{v \in S_i} \|F(v)\|^2 \geq \delta \sum_{v \in V} \|F(v)\|^2, \quad \forall\ i,$
2. Separation: $d_F(S_i, S_j) \geq \epsilon, \quad \forall\ i \neq j.$

Then we can define $\psi_i : V \to \mathbb{R}^k$ by

$$\psi_i(v) = F(v).\max\left(0, 1 - \frac{2d_F(v, S_i)}{\epsilon}\right).$$

$\psi_1, \ldots, \psi_k$ are disjointly supported and $\psi_i|_{S_i} = F|_{S_i}$.



Localizations

The goal is to find disjoint subsets $S_1, \ldots, S_k \subset V$ such that

1. Mass: $\sum_{v \in S_i} \|F(v)\|^2 \geq \delta \sum_{v \in V} \|F(v)\|^2, \quad \forall\ i,$
2. Separation: $d_F(S_i, S_j) \geq \epsilon, \quad \forall\ i \neq j.$

Then we can define $\psi_i : V \to \mathbb{R}^k$ by

$$\psi_i(v) = F(v).\max\left(0, 1 - \frac{2d_F(v, S_i)}{\epsilon}\right).$$

$$\forall\ uv \in E,\ \|\psi_i(u) - \psi_i(v)\| \leq \left(1 + \frac{4}{\epsilon}\right) \|F(u) - F(v)\|$$

Localizations

The goal is to find disjoint subsets $S_1, \dots, S_k \subset V$ such that

1. Mass: $\sum_{v \in S_i} \|F(v)\|^2 \geq \delta \sum_{v \in V} \|F(v)\|^2, \quad \forall\, i,$
2. Separation: $d_F(S_i, S_j) \geq \epsilon, \quad \forall\, i \neq j.$

Then we can define $\psi_i : V \to \mathbb{R}^k$ by

$$\psi_i(v) = F(v).\max\left(0, 1 - \frac{2d_F(v, S_i)}{\epsilon}\right).$$

$$\forall\, uv \in E,\ \|\psi_i(u) - \psi_i(v)\| \leq \left(1 + \frac{4}{\epsilon}\right) \|F(u) - F(v)\|$$

Thus $\quad \mathcal{R}(\psi_i) \leq \dfrac{1}{\delta}\left(1 + \dfrac{4}{\epsilon}\right)^2 \mathcal{R}(F).$

Random partitioning of metric spaces

A. Daneshgar
R. Javadi

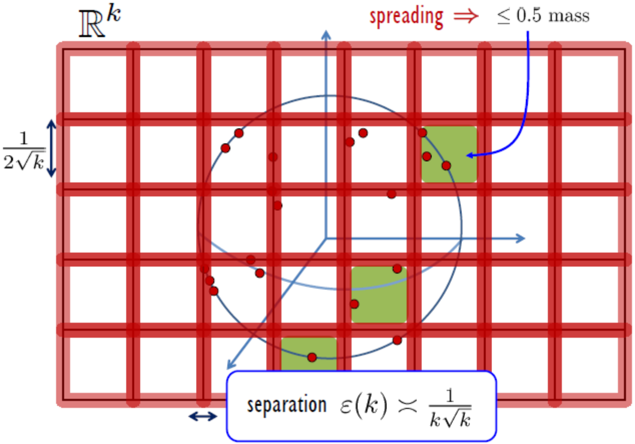
To find such S_i's, we use random partitioning theory.

Theorem (Gupta, Krauthgamer, Lee 03)

If $X \subset \mathbb{R}^k$, then for every $\Delta, \delta > 0$, X admits a random partition $\mathcal{P}$ satisfying

- $\forall\ S \in \mathcal{P},\ \operatorname{diam}(S) \leq \Delta$,
- $\forall\ x \in X,\ \mathbb{P}[B_{d_F}(x, \Delta\delta/k) \subseteq \mathcal{P}(x)] \geq 1 - \delta$.

$\mathcal{P}(x)$ is the partition containing x.


$\mathbb{R}^k$
spreading $\Rightarrow$ ≤ 0.5 mass
$\frac{1}{2\sqrt{k}}$
separation $\varepsilon(k) \asymp \frac{1}{k\sqrt{k}}$

For every $S \subset V$, let

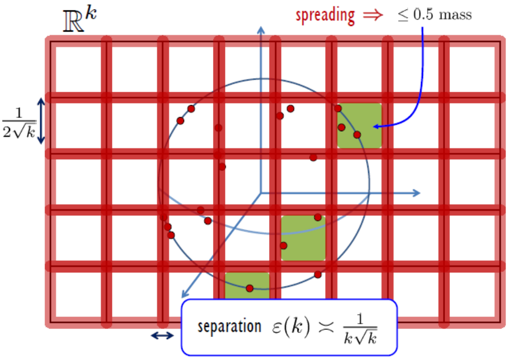
$$\tilde{S} = \{x \in S \mid B_{d_F}(x, \Delta\delta/k) \subseteq S\}.$$

Since

$$\forall\ x \in X,\ \mathbb{P}[B_{d_F}(x, \Delta\delta/k) \subseteq \mathcal{P}(x)] \geq 1 - \delta,$$

there exists partition P, where

$$\sum_{S \in P} \sum_{v \in \tilde{S}} \|F(v)\|^2 \geq (1 - \delta) \sum_{v \in V} \|F(v)\|^2.$$



Spreading property: For every $S \subset V$,

$$\operatorname{diam}(S) \leq \Delta \;\Rightarrow\; \sum_{v \in S} \|F(v)\|^2 \leq \frac{1}{k(1-\Delta^2)} \sum_{v \in V} \|F(v)\|^2.$$

Some open questions

- For noisy hypercube graph H, we have $\iota_k(H) \geq \Omega(\sqrt{\lambda_k \log k})$. But in general it is proved that $\iota_k(G) \leq O(k^2)\sqrt{\lambda_k}$. Is it true that $\iota_k(G) \leq O(\sqrt{\lambda_k \log k})$?
- Can we provide a quantitative justification for $k-$means approximation instead of random partitioning?
- Can we proved a higher Cheeger inequality for the mean isoperimetric numbers?

Part II.

Metric Embedding and Application to Sparsest Cut

Metric Space

A. Daneshgar
R. Javadi

A metric space (shortly metric) is a pair (X, d), where
X is a set of points and
$d : X \times X \to \mathbb{R}^+$ is a distance function satisfying:

1. $d(x, y) = 0$ if and only if $x = y$.
2. $d(x, y) = d(y, x)$.
3. $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality).

We are concerned with the finite metrics.

The size of X is denoted by n.

We will be lax in the first property of metrics.

Metrics and Graphs

A weighted graph is a pair (G, w), where G is a simple graph and $w : E(G) \to \mathbb{R}^+$ is a weight function.
Any weighted graph (G, w) on n vertices induces a metric d_G on n points defined as

$$\forall x, y \in V(G):$$
$$d_G(x, y) := \text{weight of shortest path between } x \text{ and } y.$$

Also for every metric (X, d), one may define a complete graph (G, w) whose vertex set is X and

$$w(x, y) := d(x, y).$$

We will use natural correspondence between graphs and metrics.

$$(G, w) \Longleftrightarrow d_G$$

A critical graph for a metric d is a minimal graph G which induces d, i.e.

$$d_G = d \text{ but } d_{G \setminus e} \neq d \text{ for each } e \in E(G).$$

Example.

	a	b	c	d	e
a	0	3	8	6	1
b		0	9	7	2
c			0	2	7
d				0	5
e					0

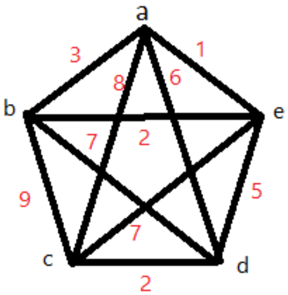
A critical graph for a metric d is a minimal graph G which induces d, i.e.

$$d_G = d \text{ but } d_{G\setminus e} \neq d \text{ for each } e \in E(G).$$

Example.

	a	b	c	d	e
a	0	3	8	6	1
b		0	9	7	2
c			0	2	7
d				0	5
e					0

$\longrightarrow$



A critical graph for a metric d is a minimal graph G which induces d, i.e.

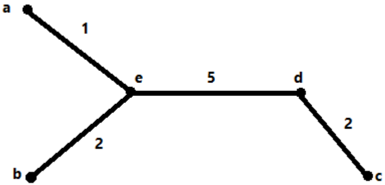
$$d_G = d \text{ but } d_{G \setminus e} \neq d \text{ for each } e \in E(G).$$

Example.

	a	b	c	d	e
a	0	3	8	6	1
b		0	9	7	2
c			0	2	7
d				0	5
e					0

$\longrightarrow$

Critical graph

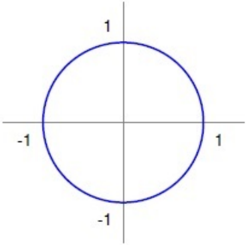


Minkowski Norms

For every $1 \leq p \leq \infty$, Minkowski norm ℓ_p is defined as follows,

$$\forall x \in \mathbb{R}^k, \quad \|x\|_p := \left(\sum_{i=1}^{k} |x_i|^p \right)^{\frac{1}{p}}$$

$$\|x\|_\infty := \max_{i=1}^{k} |x_i|$$



Unit ball

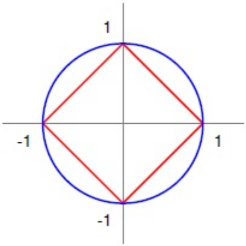
- ℓ_2: Euclidean norm $\|x\|_2 := \sqrt{\sum_{i=1}^{k} |x_i|^2}$

Minkowski Norms

For every $1 \leq p \leq \infty$, Minkowski norm ℓ_p is defined as follows,

$$\forall x \in \mathbb{R}^k, \quad \|x\|_p := \left(\sum_{i=1}^{k} |x_i|^p \right)^{\frac{1}{p}}$$

$$\|x\|_\infty := \max_{i=1}^{k} |x_i|$$



Unit ball

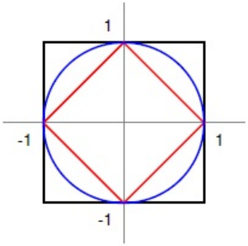
- ℓ_2: Euclidean norm $\|x\|_2 := \sqrt{\sum_{i=1}^{k} |x_i|^2}$
- ℓ_1: Manhattan norm $\|x\|_1 := \sum_{i=1}^{k} |x_i|$

Minkowski Norms

For every $1 \leq p \leq \infty$, Minkowski norm ℓ_p is defined as follows,

$$\forall x \in \mathbb{R}^k, \quad \|x\|_p := \left(\sum_{i=1}^{k} |x_i|^p \right)^{\frac{1}{p}}$$

$$\|x\|_\infty := \max_{i=1}^{k} |x_i|$$



Unit ball

- ℓ_2: Euclidean norm $\|x\|_2 := \sqrt{\sum_{i=1}^{k} |x_i|^2}$
- ℓ_1: Manhattan norm $\|x\|_1 := \sum_{i=1}^{k} |x_i|$
- ℓ_∞ : $\|x\|_\infty := \max_{i=1}^{k} |x_i|$

Isometry

$\mathbb{R}^k$ endowed with ℓ_p-norm is simply denoted by ℓ_p^k.

Isometry

$\mathbb{R}^k$ endowed with ℓ_p-norm is simply denoted by ℓ_p^k.

Given two metrics $(X, d), (X', d')$,
An embedding (map) $f : X \to X'$ is called isometry if it is distance-preserving, i.e.

$$\forall x, y \in X,\ \ d(x, y) = d'(f(x), f(y)).$$

Why embedding?

The input data for the problem at hand is (or can be interpreted as) a metric.

The idea is to embed the metric into simpler metrics (low-dimensional or well-known e.g. ℓ_p) on which the problem can be solved easier.

An example: Furthest points in ℓ_1.

Given: A set X of n points in ℓ_1^k $(n \gg 2^k)$.
Query: Find the furthest pair of points in X.

$$\max_{x,y \in X} \|x - y\|_1.$$

A brute-force algorithm can do this in $O(kn^2)$ time (not great when $n \gg k$).
But in ℓ_∞ we can do better:

$$\max_{x,y \in X} \|x - y\|_\infty = \max_{x,y \in X} \max_i |x_i - y_i|$$
$$= \max_i \text{ furthest pair along the i-th coordinate.}$$

Thus, in ℓ_∞^k, furthest pair can be found in $O(kn)$.

An example: Furthest points in ℓ_1.

Now it remains to find an isometry from ℓ_1 to ℓ_∞.

Define the isometry $f : \ell_1^k \to \ell_\infty^{2^k}$ as follows:

For every $x \in \mathbb{R}^k$ and $s \in \{-1,1\}^k$, define

$$f_s(x) = < s, x > .$$

Then

$$\|f(x)\|_\infty = \max_{s \in \{-1,1\}^k} < s, x > = \sum_{i=1}^{k} |x_i| = \|x\|_1 .$$

Therefore, $\|f(x) - f(y)\|_\infty = \|f(x-y)\|_\infty = \|x-y\|_1$.

Algorithm for furthest points in ℓ_1^k.

(1) For every $x \in X$, compute $f(x)$.

(2) Find furthest pair in $f(X) \subset \ell_\infty^{2^k}$.

Item (1) can be done in $O(nk2^k)$ and Item (2) can be done in $O(n2^k)$.

So it can be done in $O(nk2^k)$ (compare with brute-force $O(kn^2)$).

The Universal Space: ℓ_∞

Theorem (Fréchet)

Any n point metric (X, d) can be isometrically embedded into ℓ_∞^n.

Define embedding $f : X \to \ell_\infty^n$ as follows,

For every $x, y \in X$, define $f_y(x) := d(x, y)$.

f is isometry:

$$\begin{aligned} \|f(x) - f(y)\|_\infty &= \max_z |f_z(x) - f_z(y)| \\ &= \max_z |d(z, x) - d(z, y)| \\ &= d(x, y). \end{aligned}$$

We are not always so lucky to find an isometry.

For instance, 3-star $K_{1,3}$ cannot be embedded into ℓ_2 isometrically, regardless of the number of dimensions.

Try to prove it!

Even when there is an isometry, it is not practical due to curse of high dimension.

Distortion

Given two metrics $(X, d), (X', d')$ and a map $f : X \to X'$, expansion of f is the maximum factor by which distances are stretched, i.e.

$$\text{expansion}(f) := \max_{x,y \in X} \frac{d'(f(x), f(y))}{d(x, y)},$$

Also, contraction of f is the maximum factor by which distances are shrunk, i.e.

$$\text{contraction}(f) := \max_{x,y \in X} \frac{d(x, y)}{d'(f(x), f(y))},$$

The distortion of f is the product of the contraction and the expansion.

$$\|f\|_{\text{dist}} := \text{contraction}(f) \cdot \text{expansion}(f).$$

Distortion

If there exists an embedding $f : X \to X'$ with distortion D, we say that (X, d) can be embedded into (X', d') with distortion D and we write

$$(X, d) \overset{D}{\hookrightarrow} (X', d').$$

Metric Embedding Algorithms

An isometry leads to an exact algorithm, however, a non-isometric embedding leads to an approximation algorithm.

The distortion of the embedding determines the approximation factor of the algorithm.

Hence, we are interested in embeddings with low distortion into low dimensional spaces.

Férchet-type embeddings

Let (X, d) be a metric. For every $S \subseteq X$, define the mapping $f_S : X \to \mathbb{R}$ as

$$f_S(x) := d(x, S) = \min_{s \in S} d(x, s), \quad \forall\, x \in X.$$

A Férchet-type embedding (Subset embedding) is a map $f : X \to \mathbb{R}^k$, defined as

$$f(x) = \bigoplus_{S \subseteq X} \beta_S f_S,$$

for $\beta_S \in \mathbb{R}$.

Universal mapping into ℓ_∞ was obtained by setting

$$\beta_S = \begin{cases} 1 & |S| = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Embedding into ℓ_1

Theorem (Bourgain (1985))

For every metric (X, d) with n points,

$$(X, d) \overset{O(\log n)}{\hookrightarrow} \ell_1^{2^n}.$$

The embedding is the Férchet-type embedding, by setting

$$\beta_S := \frac{1}{|S|\binom{n}{|S|}}.$$

Embedding into ℓ_p

Using Férchet-type embedding with suitable choice of β_S, Matousek generalized Bourgain's result.

Theorem (Matousek (1995))

For every $1 \leq p < \infty$ and for every metric (X, d) with n points,

$$(X, d) \overset{O(\log n)}{\hookrightarrow} \ell_p^{2^n}.$$

Embedding into ℓ_p

Random sampling ideas can be applied to randomly select some suitable subsets in Férchet-type embedding and reduce the dimensions.

Theorem (Linial, London, Rabinovich (1995))

For every $1 \leq p \leq \infty$ and for every metric (X, d) with n points,

$$(X, d) \overset{O(\log n)}{\hookrightarrow} \ell_p^{\log^2 n}.$$

Go Back

Go Back

Dimensionality Reduction in $\mathbb{R}^k$

For $X \subset \mathbb{R}^k$, if distances are allowed to be slightly distorted, one can substantially reduce the dimensions.

Theorem (Johnson and Lindenstrauss (1984))

Given any $X \subset \mathbb{R}^k$ of n points and any $\epsilon > 0$,

$$X \overset{1+\epsilon}{\hookrightarrow} \mathbb{R}^{\log n/\epsilon^2}.$$

- The embedding is obtained by projecting the point set on a random $k' = O(\log n/\epsilon^2)$-dimensional subspace and then scaling up the resulting distances by $\sqrt{k/k'}$.
- This is an example of the concentration of measure phenomenon (random variables in high dimensions are concentrated around their means).

Application: Shortest Path

An easy application: computation of shortest paths.

If $G \overset{D}{\hookrightarrow} \ell_p^k$, then

we can store the graph in space $O(nk)$ and D-approximate the shortest path between any two vertices in $O(k)$ time.

e.g. imagine $G \overset{O(\log n)}{\hookrightarrow} \ell_p^{\log^2 n}$ we get space $O(n \log^2 n)$ and time $O(\log n)$.

Application: Sparsest Cut

Given a complete graph $G = (V, V \times V)$,
an edge-capacity function $\mathrm{cap} : V \times V \to \mathbb{R}^+$ and
a pairwise demand function $\mathrm{dem} : V \times V \to \mathbb{R}^+$.

The sparsity of a cut $S \subset V$ is given by

$$\alpha_S := \frac{\mathrm{cap}(S, V \setminus S)}{\mathrm{dem}(S, V \setminus S)}.$$

The sparsity of G is defined as

$$\alpha(G) = \min_{S \subset V} \alpha_S.$$

When cap and dem are constant, it is just mean isoperimetric number.

Sparsest Cut

For a subset $S \subset V$, define $d : V \times V \to \mathbb{R}$ to be the characteristic function of the cut $(S, V \setminus S)$.

$$d(x,y) := \begin{cases} 1 & x \in S, y \in V \setminus S \\ 0 & \text{otherwise.} \end{cases}$$

Actually d is a metric and is called a cut metric. Then

$$\alpha(G) = \min_{d \text{ is a cut metric}} \frac{\sum_{x,y} \operatorname{cap}(x,y) d(x,y)}{\sum_{x,y} \operatorname{dem}(x,y) d(x,y)}.$$

IP for sparsest cut

$\alpha(G)$ is the solution of the following integer programming,

$$\begin{aligned} \text{minimize} \quad & \sum_{x,y} \text{cap}(x,y)\ d(x,y) \\ \text{subject to} \quad & \sum_{x,y} \text{dem}(x,y)\ d(x,y) = 1 \\ & d \text{ is a cut metric.} \end{aligned}$$

LP relaxation of sparsest cut

A. Daneshgar
R. Javadi

Since we have no way to write being cut metric as a linear condition, we relax it to be a metric.

$$\begin{aligned} \text{minimize} \quad & \sum_{x,y} \mathrm{cap}(x,y)\ d(x,y) \\ \text{subject to} \quad & \sum_{x,y} \mathrm{dem}(x,y)\ d(x,y) = 1 \\ & d(x,y) \leq d(x,z) + d(z,y), \ \forall x,y,z \\ & d(x,y) \geq 0 \end{aligned}$$

Let LP^* and d^* denote respectively the optimal value and the metric acheiving the optimal value of this LP.
Since this LP is a relaxation, we have $LP^* \leq \alpha(G)$.
But unfortunately the inequality can be strict.

ℓ_1 Metrics and Cut Metrics

A. Daneshgar
R. Javadi

Lemma

Any ℓ_1 metric on the set $V \subset \mathbb{R}^k$ can be written as a linear combination of cut metrics, i.e. for every $x, y \in V$

$$\|x - y\|_1 = \sum_{S \subset V} \lambda_S d_S,$$

where λ_S are constants and d_S is characteristic function of the cut $(S, V \setminus S)$. Moreover, at most kn of λ_S are nonzero. (The converse is also true.)

Proof. Fix coordinate i and order the points $v_1(i) \leq v_2(i) \leq \dots v_n(i)$. For every $1 \leq t \leq n$, set $S_{t,i} := \{v_1, \dots, v_t\}$ and $\lambda_{S_{t,i}} := v_{t+1}(i) - v_t(i)$. Put the other λ_S be equal to zero. □

Integrality Gap and Distortion into ℓ_1

Theorem

Given an instance G of the sparsest cut problem, let d^* be the metric achieving the optimal value of the above LP.

If $(V, d^*) \overset{D}{\hookrightarrow} \ell_1$, then $LP^* \leq \alpha(G) \leq D \cdot LP^*$.

Proof. Let μ be the ℓ_1 metric satisfying $\mu \leq d^* \leq D \cdot \mu$ and write $\mu = \sum_S \lambda_S d_S$. Then

$$\begin{aligned} LP^* &= \frac{\sum_{x,y} \mathrm{cap}(x,y) d^*(x,y)}{\sum_{x,y} \mathrm{dem}(x,y) d^*(x,y)} \geq \frac{\sum_{x,y} \mathrm{cap}(x,y) \mu(x,y)}{D \sum_{x,y} \mathrm{dem}(x,y) \mu(x,y)} \\ &= \frac{\sum_S \lambda_S \,\mathrm{cap}(S, V \setminus S)}{D \cdot \sum_S \lambda_S \,\mathrm{dem}(S, V \setminus S)} \geq \min_{S : \lambda_S \neq 0} \frac{\mathrm{cap}(S, V \setminus S)}{D \cdot \mathrm{dem}(S, V \setminus S)} \\ &\geq \frac{\alpha(G)}{D}. \end{aligned}$$

$O(\log n)-$Approximation Algorithm for Sparsest Cut Problem

The above theorem along with a low-distortion embedding into ℓ_1 give rise to an $O(\log n)-$approximation algorithm for sparsest cut problem.

Given an instance G of sparsest cut problem,

1. Solve LP to find optimal metric d^*.
2. Using embedding $(V, d^*) \overset{O(\log n)}{\hookrightarrow} \ell_1^{O(\log^2 n)}$, find ℓ_1-metric μ.

Embedding

3. Write $\mu = \sum_{S \subset V} \lambda_S d_S$ and find the cut S^* with the minimum sparsity among all S whose λ_S is nonzero. (There are at most $n \log^2 n$ of such S.)

Lower Bound for Distortion

Is Linial et.al embedding Embedding the best low-distortion embedding?

Is there any way to find an embedding into ℓ_1 with less distortion?

Theorem (Leighton and Rao (1988))

For infinitely many values of n, there exists graphs G_n on n vertices such that

$$\alpha(G_n)/LP^* \geq \Omega(\log n).$$

Corollary

If $d^*_{G_n}$ is the metric generated by the LP LP *on G_n, then $(V_n, d^*_{G_n})$ cannot be embedded into ℓ_1 with distortion smaller than $\Omega(\log n)$.*

Multicommodity Flow: LP duality of Sparsest Cut

Given a graph $G = (V, V \times V)$ and
an edge-capacity function $\text{cap} : V \times V \to \mathbb{R}^+$.
For every pair of vertices (x, y), there is a separate commodity and a demand $\text{dem}(x, y) \in \mathbb{R}^+$.

The objective is to maximize number λ such that for each commodity (x, y), at least $\lambda \cdot \text{dem}(x, y)$ units of this commodity can be routed simultaneously, subject to flow conservation and capacity constraints, i.e.,
each commodity must satisfy flow conservation at each vertex other than its own source and sink, and the sum of flows routed through an edge should not exceed the capacity of this edge.

This problem is dual problem of the LP for sparsest cut.

So if λ^* is the optimal value, then $\lambda^* = LP^*$.

Expanders

Edge-expansion of a graph G is defined as

$$\Phi(G) := \min_{S \subset V} \left(\frac{\operatorname{cap}(S, V \setminus S)}{|S|} + \frac{\operatorname{cap}(S, V \setminus S)}{|V \setminus S|} \right).$$

If $\operatorname{dem}(x, y) = 1$, for every x, y, then

$$\Phi(G) = n \cdot \alpha(G).$$

A graph G is an expander if the expansion is a constant independent of n, i.e. $\Phi(G) = \Omega(1)$. Constant-degree expanders are known to exist:

Theorem (Lubotzky, Phillips, and Sarnak (1988))

For infinitely many n, there exist $3-$regular expanders G_n.

Proof of Leighton-Rao Theorem. ▸ Leighton

For $3-$regular expanders G_n, we have $\alpha(G) \geq \Omega(\frac{1}{n})$, but $LP^* \leq O(\frac{1}{n\log n})$.

First note that for every vertex v, the number of vertices at distance at least $\frac{1}{2}\log n$ from v is at least $\frac{1}{2}n$.

This gives at least $(1/2)n \cdot n$ demand pairs, such that sending λ^* units of flow between any one of these pairs uses up at least $\lambda^* \frac{\log n}{2}$ units of volume. Since the total capacity available in the 3-regular graph is only $3n/2$, we must have that

$$\frac{1}{2}n \cdot n \cdot \lambda^* \frac{\log n}{2} \leq \frac{3n}{2}.$$

Thus,

$$\lambda^* \leq \frac{6}{n\log n}.$$

Further Works

A. Daneshgar
R. Javadi

There is no hope to find a better approximation for the sparsest cut problem by rounding the LP.

However, there is a different relaxation through semidefinite programming (SDP).

Arora, Rao and Vazirani used this relaxation to find a better approximation algorithm. Be with us in the next lecture!

Theorem (Arora, Rao and Vazirani (2004))

There is an $O(\sqrt{\log n})-$approximation algorithm which solves the sparsest cut problem.

Some Excersies

Prove that:

1. Any tree with n vertices can be isometrically embedded into ℓ_1^{n-1}.
2. Any tree with n vertices can be isometrically embedded into $\ell_\infty^{O(\log n)}$.
3. The complete graph on n vertices can be isometrically embedded into $\ell_\infty^{\lceil \log n \rceil}$ and this is optimal.
4. $K_{1,3}$ cannot be embedded into ℓ_2 isometrically.
5. A cycle on n vertices cannot be embedded into a tree with distortion lower than $n-1$.
6. A binary tree on n vertices can be embedded into ℓ_2^n with distortion $O(\sqrt{\log \log n})$.

Further Reading

- For preparing these slides, for part I, the slides of James Lee and for Part II, the lecture notes of Räcke has been used, as well as the slides of Montanaro.
- For metric embedding, there are a number lecture courses. Just Google it!
- You may find the following book fruitful: *discrete geometry* by Matousek.
- Also, see: *The geometry of graphs and some of its algorithmic applications* by Linial, London and Rabinovich.

Thank you!

Comments and Criticisms are Welcomed

daneshgar@sharif.ir
rjavadi@cc.iut.ac.ir