Low-Distortion Embeddings of Metric Spaces

Yoshio Okamoto

Dept. Information and Computer Sciences, Toyohashi University of Technology

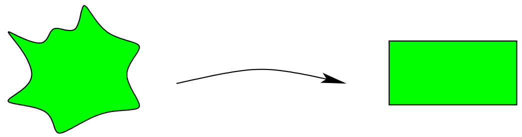
October, 2007

NHC Autumn School on Computational Geometry and Integer Programming

What we want to do

We want to do

Transform one metric space into another



Why we want to do that

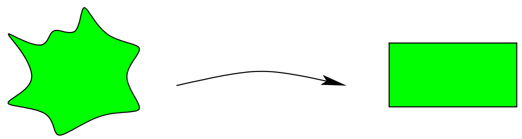
Algorithmic applications

- Geometric approximation algorithms
  Closest pairs, Nearest neighbors, ...
- Combinatorial approximation algorithms
  Sparsest cuts, Multi-commodity flows, Bandwidths, ...
  (cf. J. Lee's NHC Workshop Talk)
- Inapproximability (integrality gap of SDP relaxation)
  Vertex covers, Unique game conjectures, ...
- On-line algorithms
  Metrical task systems, ...
- Streaming algorithms
- ...

How we want to transform the space

Criteria

- efficiently
- into a space as low-dimensional as possible
- into a space as simple as possible
- without much distortion



Contents

References

Books

- "Lectures on Discrete Geometry" by Matoušek, Springer, 2002
- "Geometry of Cuts and Metrics" by Deza and Laurent, Springer, 1997

Surveys

- "Low-distortion embeddings of finite metric spaces" by Indyk and Matoušek, in "Handbook on Discrete and Computational Geometry," 2004
- "Algorithmic applications of low-distortion embeddings" by Indyk, FOCS 2001
- "Finite metric spaces—combinatorics, geometry and algorithms" by Linial, ICM 2002

Definition: Metric Spaces

Def.: metric space

A pair (X, μ) of a set X and a map $\mu\colon X \times X \to \mathbb{R}_+$ is called a metric space if

- $\mu(x, y) = 0 \Leftrightarrow x = y$,
- $\mu(x, y) = \mu(y, x)$ for all $x, y \in X$,
- $\mu(x, y) + \mu(y, z) \geq \mu(x, z)$ for all $x, y, z \in X$.

Def.: finite metric space

A finite metric space is a metric space (X, μ) with X finite.

- Note: This is *different* from a discrete metric space.

Representing a finite metric space

By a matrix

	x_1	x_2	x_3	x_4	x_5
x_1	0	2	3	4	3
x_2	2	0	4	2	1
x_3	3	4	0	2	5
x_4	4	2	2	0	3
x_5	3	1	5	3	0

The i, j-component represents $\mu(x_i, x_j)$.

Definition: Normed Spaces

Def.: normed spaces

A pair $(X, \|\cdot\|)$ of a vector space X on $\mathbb{R}$ and a map $x \in X \mapsto \|x\| \in \mathbb{R}_+$ is called a normed space if

- $\|x\| = 0 \Leftrightarrow x = 0$,
- $\|\alpha x\| = |\alpha|\|x\|$ for all $\alpha \in \mathbb{R}$, $x \in X$,
- $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

Observation (or Exercise)

Given a normed space $(X, \|\cdot\|)$, we define $\mu\colon X \times X \to \mathbb{R}_+$ as

$$\mu(x, y) = \|x - y\|.$$

Then, (X, μ) is a metric space.

$\therefore$ we may think of normed spaces as metric spaces.

Typical Norms: ℓ_p-Norms

Def.: ℓ_p-norms

Define $\|x\|_p \in \mathbb{R}_+$ for every $x \in \mathbb{R}^d$ as

$$\|x\|_p = \left(\sum_{i=1}^{d} |x_i|^p\right)^{1/p}.$$

Then $(\mathbb{R}^d, \|\cdot\|_p)$ is a normed space, denoted by ℓ_p^d.
This norm is called the ℓ_p-norm.

FANs (frequently asked norms)

- $p = 1$: $\sum_{i=1}^{n} |x_i|$ (the Manhattan norm)
- $p = 2$: $\sqrt{\sum_{i=1}^{n} |x_i|^2}$ (the Euclidean norm)
- $p = \infty$: $\max_{i=1}^{n} |x_i|$ (the maximum norm)

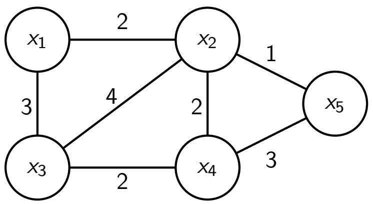
Definition: Metrics on Graphs

$G = (V, E)$ a finite undirected graph, connected
$w\colon E \to \mathbb{R}_+$ an edge-weight function

Def.: shortest-path metrics

A pair (V, μ) is a shortest-path metric on G if
$\mu(u, v)$ is the shortest-path distance between u & v on G w.r.t w.

	x_1	x_2	x_3	x_4	x_5
x_1	0	2	3	4	3
x_2	2	0	4	2	1
x_3	3	4	0	2	5
x_4	4	2	2	0	3
x_5	3	1	5	3	0



Remark (or Exercise)

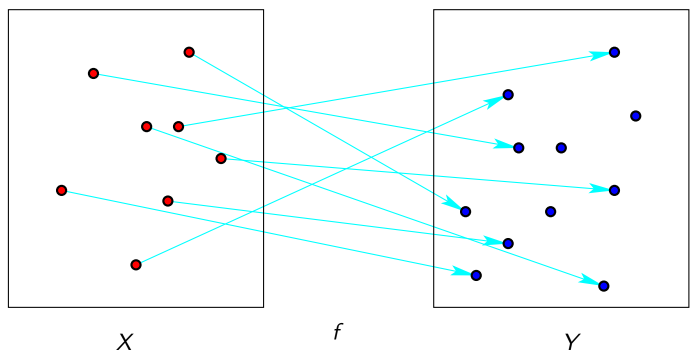
Every finite metric space is a shortest-path metric on a graph.

Definitions: Embedding and Distortion

(X, μ), (Y, ν) two metric spaces

Def.: embeddings

An embedding of (X, μ) into (Y, ν) is a map $f\colon X \to Y$.



Definition: Distortions

(X, μ), (Y, ν) two metric spaces, $D \geq 1$

Def.: D-embeddings

A *D-embedding* is an embedding $f: X \to Y$ s.t.
$\exists\, r > 0, \forall\, x, y \in X$

$$r \cdot \mu(x, y) \leq \nu(f(x), f(y)) \leq D \cdot r \cdot \mu(x, y).$$

Def.: distortion

The *distortion* of an embedding f is $\inf\{D : f \text{ is a } D\text{-embedding}\}$.

Def.: isometry

An embedding is an *isometry* if its distortion is 1.

Embedding into ℓ_p

Def.: embedding into ℓ_p

An *embedding into ℓ_p* of a metric space (X, μ) is an embedding $(X, \mu) \to \ell_p^d$ for some finite d.

Note:

- We may also define the normed space ℓ_p, but it will be a bit subtle. We need to work around the infinity and convergence issues. Thus, we use ℓ_p just as a notational convenience.

Three basic results

We are going to look at the following.

1. Every n-point metric space can be isometrically embedded into ℓ_∞.
2. Every n-point metric space can be embedded into ℓ_2 with distortion $O(\log n)$.
3. There exists an n-point metric space that requires the distortion of $\Omega(\log n)$ when embedded into ℓ_2.

Contents

Isometric embeddability into ℓ_∞

(X, μ) a finite metric space

Theorem

(X, μ) can be embedded into ℓ_∞ with distortion 1.

Proof Outline:

- Explicitly construct a particular embedding
- Prove that it is an isometry

Construction of an isometric embedding

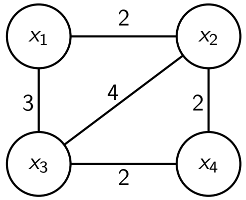
(X, μ) an n-point metric space; $X = \{x_1, \ldots, x_n\}$

Construction

Define a map $f\colon X \to \ell_\infty^n$ as

$$f(x)_i = \mu(x, x_i)$$

for every $x \in X$ and $i \in \{1, \ldots, n\}$



$$\begin{aligned} f(x_1) &= (0, 2, 3, 4) \\ f(x_2) &= (2, 0, 4, 2) \\ f(x_3) &= (3, 4, 0, 2) \\ f(x_4) &= (4, 2, 2, 0) \end{aligned}$$

Proof of the isometry of f

Claim

The constructed f is an isometry.

Proof:

- No contraction

$$\begin{aligned} \|f(x_i) - f(x_j)\|_\infty &= \max\{|f(x_i)_k - f(x_j)_k| \colon k \in \{1, \ldots, n\}\} \\ &\geq |f(x_i)_j - f(x_j)_j| \\ &= |\mu(x_i, x_j) - \mu(x_j, x_j)| \\ &= \mu(x_i, x_j). \end{aligned}$$

- No expansion

$$\begin{aligned} \|f(x_i) - f(x_j)\|_\infty &= |f(x_i)_k - f(x_j)_k| \quad \text{(for some } k\text{)} \\ &= |\mu(x_i, x_k) - \mu(x_j, x_k)| \\ &\leq \mu(x_i, x_j). \end{aligned}$$

Contents

$O(\log n)$-embeddability into ℓ_2

(X, μ) an n-point metric space

Theorem (Bourgain '85)

(X, μ) can be embedded into ℓ_2 with distortion $O(\log n)$.

Proof Outline:

- Explicitly construct a particular embedding
  Some probability involved
- Prove that the distortion is $O(\log n)$

Construction of a low-distortion embedding

(X, μ) an n-point metric space; $q = \lfloor \log_2 n \rfloor + 1$

Construction

- Construct $A \subseteq X$ at random as follows:
  - For each $j \in \{1, \ldots, q\}$ construct $A_j \subseteq X$ by sampling every point of X independently with prob. $1/2^j$
  - Choose $j \in \{1, \ldots, q\}$ uniformly at random
  - Set $A = A_j$
- Define a map $f \colon X \to \ell_2^{2^n}$ as

$$f(x)_S = \sqrt{\Pr[S = A]}\ \mu(x, S)$$

for every $x \in X$ and $S \subseteq X$

Notation

$\mu(x, S) = \min\{\mu(x, y) \colon y \in S\}$

No expansion of the constructed embedding

(X, μ) an n-point metric space; $q = \lfloor \log_2 n \rfloor + 1$

Claim 1

The constructed f satisfies

$$\|f(x) - f(y)\|_2 \leq \mu(x, y)$$

for every $x, y \in X$

Proof: exercise

Logarithmically bounded contraction of the constructed embedding

(X, μ) an n-point metric space; $q = \lfloor \log_2 n \rfloor + 1$

Claim 2

The constructed f satisfies

$$\mu(x, y) \leq 32q\|f(x) - f(y)\|_2$$

for every $x, y \in X$

Note

The exact coefficient $32q$ is not important (it could be easily improved). It is only important that the coefficient is $O(\log n)$.

Proof:

- Let's first look at some calculation

Some calculation

Notation

$p_S = \Pr[S = A]$

$$\begin{aligned}
\|f(x) - f(y)\|_2 &= \sqrt{\sum_{S \subseteq X} |f(x)_S - f(y)_S|^2} \\
&= \sqrt{\sum_{S \subseteq X} |\sqrt{p_S}\mu(x,S) - \sqrt{p_S}\mu(y,S)|^2} \\
&= \sqrt{\sum_{S \subseteq X} p_S|\mu(x,S) - \mu(y,S)|^2} \\
&= \sqrt{\sum_{S \subseteq X} p_S|\mu(x,S) - \mu(y,S)|^2}\sqrt{\sum_{S \subseteq X} p_S} \\
&\geq \sum_{S \subseteq X} p_S|\mu(x,S) - \mu(y,S)|
\end{aligned}$$

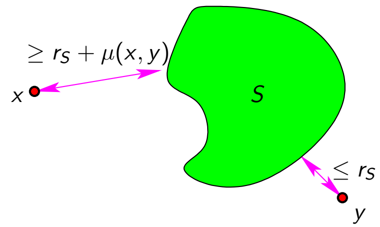
Claim modified

Claim 2'

$$\sum_{S \subseteq X} p_S|\mu(x,S) - \mu(y,S)| \geq \frac{\mu(x,y)}{32q}$$

Proof idea:

- Try to show for "many" sets $S \subseteq X$ there exists r_S s.t. $\mu(x,S) \geq r_S + \mu(x,y)$ and $\mu(y,S) \leq r_S$



A bit of thought

Suppose..

$\forall\ S \subseteq X\ \exists\ r_S > 0$ s.t. $\mu(x,S) \geq r_S + \mu(x,y)$ and $\mu(y,S) \leq r_S$

Then,

$$\begin{aligned}
\sum_{S \subseteq X} p_S|\mu(x,S) - \mu(y,S)| &\geq \sum_{S \subseteq X} p_S((r_S + \mu(x,y)) - r_S) \\
&= \sum_{S \subseteq X} p_S\mu(x,y) = \mu(x,y)
\end{aligned}$$

Thus, we have no expansion!! But, ...

- Not all S satisfy this assumption
- However, if "many" sets satisfy the assumption, the result holds!

Claim modified

Claim 2'

$$\sum_{S \subseteq X} p_S|\mu(x,S) - \mu(y,S)| \geq \frac{\mu(x,y)}{32q}$$

Proof idea:

- Try to show for "many" sets $S \subseteq X$ there exists r_S s.t. $\mu(x,S) \geq r_S + \mu(x,y)$ and $\mu(y,S) \leq r_S$

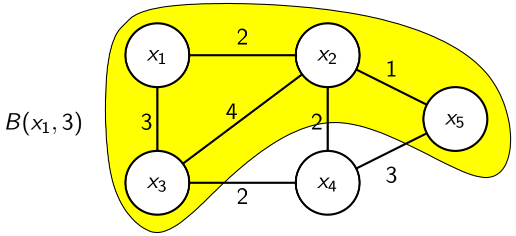
To this end

- Instead of counting, we look at probabilities...
- Quantize $|\mu(x,S) - \mu(y,S)|$ with j
- Throw out some sets S from the summation
- Bound p_S

Definition: Balls

Definition (a ball of radius r centered at c)

$B(c, r) = \{z \in X : \mu(c, z) \leq r\}$ (closed)
$B^\circ(c, r) = \{z \in X : \mu(c, z) < r\}$ (open)



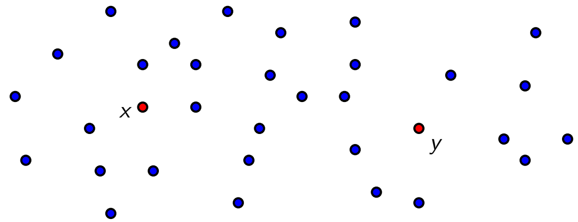
Quantize $|\mu(x, S) - \mu(y, S)|$ (1)

Fix $x, y \in X$

Notation

- For every $j \in \{0, 1, \ldots, q\}$

$$\tilde{r}_j = \min\{r : |B(x, r)| \geq 2^j \text{ and } |B(y, r)| \geq 2^j\}$$



Quantize $|\mu(x, S) - \mu(y, S)|$ (2)

Fix $x, y \in X$

Notation

- For every $j \in \{0, 1, \ldots, q\}$

$$\tilde{r}_j = \min\{r : |B(x, r)| \geq 2^j \text{ and } |B(y, r)| \geq 2^j\}$$

- Let i be an index satisfying

$$\tilde{r}_0 \leq \tilde{r}_1 \leq \cdots \leq \tilde{r}_{i-1} \leq \frac{1}{2}\mu(x, y) \leq \tilde{r}_i \leq \cdots \leq \tilde{r}_q$$

- For every $j \in \{0, 1, \ldots, i\}$

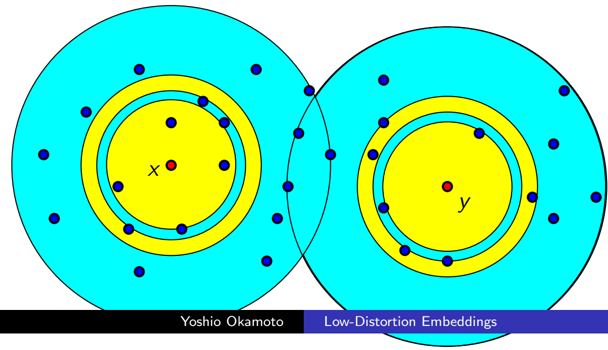
$$r_j = \begin{cases} \tilde{r}_j & (j \in \{0, 1, \ldots, i-1\}) \\ \frac{1}{2}\mu(x, y) & (j = i) \end{cases}$$

Quantize $|\mu(x, S) - \mu(y, S)|$ (3)

$j \in \{1, \ldots, i\}$ fixed

Observations

- $|B^\circ(x, r_j)| < 2^j$ or $|B^\circ(y, r_j)| < 2^j$ (from the def of r_j)
- WLOG $|B^\circ(x, r_j)| < 2^j$
- $|B(y, r_{j-1})| \geq 2^{j-1}$ (from the def of r_{j-1})

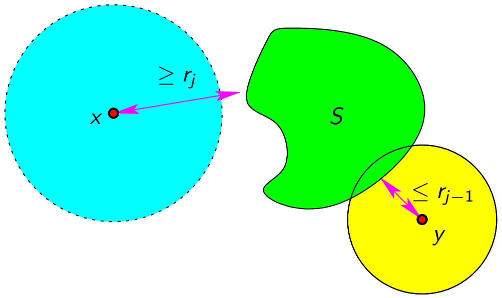


Quantize $|\mu(x,S) - \mu(y,S)|$ (4)

$j \in \{1, \dots, i\}$ fixed

Important observation

$B^\circ(x, r_j) \cap S = \emptyset$ and $B(y, r_{j-1}) \cap S \neq \emptyset$
$\Rightarrow |\mu(x,S) - \mu(y,S)| \geq r_j - r_{j-1}$



Throw out some sets S from the summation (1)

$$
\begin{aligned}
&\sum_{S \subseteq X} \Pr[S = A] |\mu(x,S) - \mu(y,S)| \\
&= \sum_{S \subseteq X} \sum_{j=1}^{q} \Pr[S = A \mid j \text{ chosen}] \Pr[j \text{ chosen}] |\mu(x,S) - \mu(y,S)| \\
&= \sum_{S \subseteq X} \sum_{j=1}^{q} \Pr[S = A \mid j \text{ chosen}] \frac{1}{q} |\mu(x,S) - \mu(y,S)| \\
&= \frac{1}{q} \sum_{j=1}^{q} \sum_{S \subseteq X} \Pr[S = A \mid j \text{ chosen}] |\mu(x,S) - \mu(y,S)| \\
&\geq \frac{1}{q} \sum_{j=1}^{i} \sum_{S \subseteq X} \Pr[S = A \mid j \text{ chosen}] |\mu(x,S) - \mu(y,S)| \\
&\geq \frac{1}{q} \sum_{j=1}^{i} \sum_{\substack{S \subseteq X, \\ B^\circ(x,r_j) \cap S = \emptyset, \\ B(y,r_{j-1}) \cap S \neq \emptyset}} \Pr[S = A \mid j \text{ chosen}] |\mu(x,S) - \mu(y,S)|
\end{aligned}
$$

Throw out some sets S from the summation (2)

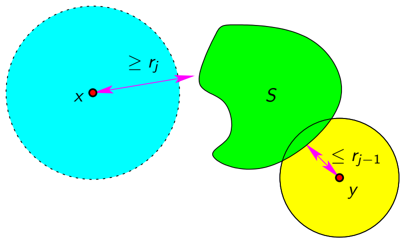
$$
\begin{aligned}
&\sum_{S \subseteq X} \Pr[S = A] |\mu(x,S) - \mu(y,S)| \\
&\geq \frac{1}{q} \sum_{j=1}^{i} \sum_{\substack{S \subseteq X, \\ B^\circ(x,r_j) \cap S = \emptyset, \\ B(y,r_{j-1}) \cap S \neq \emptyset}} \Pr[S = A \mid j \text{ chosen}] |\mu(x,S) - \mu(y,S)| \\
&\geq \frac{1}{q} \sum_{j=1}^{i} \sum_{\substack{S \subseteq X, \\ B^\circ(x,r_j) \cap S = \emptyset, \\ B(y,r_{j-1}) \cap S \neq \emptyset}} \Pr[S = A \mid j \text{ chosen}] (r_j - r_{j-1}) \\
&= \frac{1}{q} \sum_{j=1}^{i} (r_j - r_{j-1}) \underbrace{\sum_{\substack{S \subseteq X, \\ B^\circ(x,r_j) \cap S = \emptyset, \\ B(y,r_{j-1}) \cap S \neq \emptyset}} \Pr[S = A \mid j \text{ chosen}]}_{\text{need to bound}}
\end{aligned}
$$

Bound the probability (1)

$$
\begin{aligned}
&\sum_{\substack{S \subseteq X, \\ B^\circ(x,r_j) \cap S = \emptyset, \\ B(y,r_{j-1}) \cap S \neq \emptyset}} \underbrace{\Pr[S = A \mid j \text{ chosen}]}_{=\Pr[S = A_j]} \\
&= \Pr[B^\circ(x, r_j) \cap A_j = \emptyset \text{ and } B(y, r_{j-1}) \cap A_j \neq \emptyset] \\
&= \Pr[B^\circ(x, r_j) \cap A_j = \emptyset] \cdot \Pr[B(y, r_{j-1}) \cap A_j \neq \emptyset]
\end{aligned}
$$

Note

$j \leq i$ implies $B^\circ(x, r_j) \cap B(y, r_{j-1}) = \emptyset$,
$\therefore$ the events $B^\circ(x, r_j) \cap S = \emptyset$ and $B(y, r_{j-1}) \cap S \neq \emptyset$ independent.



Bound the probability (2)

$$\begin{aligned}\Pr[B^\circ(x,r_j)\cap A_j=\emptyset] &= \Pr[z\notin A_j \text{ for all } z\in B^\circ(x,r_j)]\\ &= \left(1-\frac{1}{2^j}\right)^{|B^\circ(x,r_j)|} > \left(1-\frac{1}{2^j}\right)^{2^j}\\ &\geq \left(1-\frac{1}{2^1}\right)^{2^1} = \frac{1}{4}\end{aligned}$$

Reminder (construction of A_j)

For each $j\in\{1,\dots,q\}$ construct $A_j\subseteq X$ by sampling every point of X independently with prob. $1/2^j$

Reminder (assumptions and a fact)

$|B^\circ(x,r_j)|<2^j$, $j\geq 1$ and $(1-1/2^j)^{2^j}$ monotonically increasing

Bound the probability (3)

$$\begin{aligned}\Pr[B(y,r_{j-1})\cap A_j\neq\emptyset] &= 1-\Pr[B(y,r_{j-1})\cap A_j=\emptyset]\\ &= 1-\left(1-\frac{1}{2^j}\right)^{|B(y,r_{j-1})|}\\ &\geq 1-\left(1-\frac{1}{2^j}\right)^{2^{j-1}}\\ &\geq 1-\exp\left(\frac{1}{2^j}2^{j-1}\right)\\ &= 1-\frac{1}{\sqrt{e}}\geq\frac{1}{4}\end{aligned}$$

Reminder (assumption and a well-known fact)

$|B(y,r_{j-1})|\geq 2^{j-1}$, $1+x\leq\exp(x)$ for all $x\in\mathbb{R}$

Summing up...

$$\begin{aligned}&\sum_{S\subseteq X} p_S|\mu(x,S)-\mu(y,S)|\\ &\geq \frac{1}{q}\sum_{j=1}^{i}(r_j-r_{j-1})\Pr[B^\circ(x,r_j)\cap A_j=\emptyset]\cdot\Pr[B(y,r_{j-1})\cap A_j\neq\emptyset]\\ &\geq \frac{1}{q}\sum_{j=1}^{i}(r_j-r_{j-1})\cdot\frac{1}{4}\cdot\frac{1}{4}\\ &= \frac{1}{16q}\sum_{j=1}^{i}(r_j-r_{j-1})\\ &= \frac{1}{16q}(r_i-r_0) \quad \text{(telescopic sum)}\\ &\geq \frac{1}{16q}\left(\frac{1}{2}\mu(x,y)-0\right)\\ &= \frac{1}{32q}\mu(x,y)\end{aligned}$$

Contents

Lower bound for the minimum distortion into ℓ_2

Theorem (Linial, London, and Rabinovich '95)

An embedding of a shortest-path metric on an n-vertex constant-degree expander (with unit weight) into ℓ_2 needs $\Omega(\log n)$ distortion.

Proof: exercise (with guides)

Instead, we now prove the following

Theorem (Enflo '69)

An embedding of a shortest-path metric on an n-vertex Hamming cube (with unit weight) into ℓ_2 needs $\Omega(\sqrt{\log n})$ distortion.

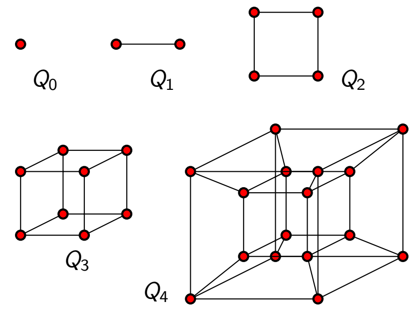
A proof (below) should be a hint to the exercise above.

Hamming cubes

d a positive integer

Definition

A d-dimensional Hamming cube Q_d is a graph defined as
Vertex set $V(Q_d) = \{0,1\}^d$
Edge set $E(Q_d) = \{\{u,v\} : u \text{ \& } v$ differ at exactly one coord.$\}$



Lower bound for Hamming cubes

See Q_d itself as a shortest-path metric space on Q_d w/ unit weight

Theorem

$d \geq 2$ a natural number
$f : V(Q_d) \to \ell_2$ a D-embedding $\Rightarrow D \geq \sqrt{d}$

Note: $n = |V(Q_d)| = 2^d$, $\therefore \sqrt{d} = \sqrt{\log_2 n}$

Theorem (in words)

There exists an n-point metric space that cannot be embedded into ℓ_2 with distortion better than $\sqrt{\log_2 n}$

Proof: Notation

Set-up

- (X,μ), (X,ν) two metric spaces
- $E, F \subseteq \binom{X}{2}$ non-empty sets of 2-element subsets of X

Notation

- $\mathrm{ave}_2(\mu, E) = \sqrt{\frac{1}{|E|} \sum_{\{x,y\} \in E} \mu(x,y)^2}$ (root mean square)
- $R_{E,F}(\mu) = \frac{\mathrm{ave}_2(\mu, F)}{\mathrm{ave}_2(\mu, E)}$

Proof: Observation

For a D-embedding $f\colon X \to \ell_2^k$, let $\nu(x,y) = \|f(x) - f(y)\|_2$

Observation

With the notation above, it holds that

$$R_{E,F}(\mu) \le D \cdot R_{E,F}(\nu)$$

Proof:

- Since f a D-embedding, by the def of D-embeddings, $\exists\, r$: $r \cdot \mu(x,y) \le \nu(x,y) \le D \cdot r \cdot \mu(x,y)$
- $\therefore\ \mathrm{ave}_2(\mu, F) \le \frac{1}{r}\mathrm{ave}_2(\nu, F)$ and $\mathrm{ave}_2(\mu, E) \ge \frac{1}{Dr}\mathrm{ave}_2(\nu, E)$
- $\therefore\ \dfrac{\mathrm{ave}_2(\mu,F)}{\mathrm{ave}_2(\mu,E)} \le \dfrac{\frac{1}{r}\mathrm{ave}_2(\nu,F)}{\frac{1}{Dr}\mathrm{ave}_2(\nu,E)} = D\dfrac{\mathrm{ave}_2(\nu,F)}{\mathrm{ave}_2(\nu,E)}$

Proof: Claim

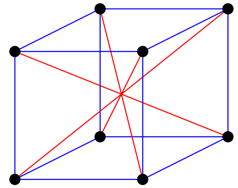
To prove the theorem it is enough to show the following

Claim

$$\frac{R_{E,F}(\mu)}{R_{E,F}(\nu)} \ge \sqrt{d} \text{ for some } E, F \subseteq \binom{V(Q_d)}{2}$$

Proof outline:

- Let $E = E(Q_d)$ and $F =$ the set of pairs w/ dist. d in Q_d
  More formally
  $F = \{\{v, \overline{v}\} \colon v \in V(Q_d)\}$ where $\overline{v}$ is the comp-wise flip of v
- Show $R_{E,F}(\mu) = d$
- Show $R_{E,F}(\nu) \le \sqrt{d}$

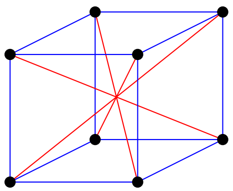


Computing $R_{E,F}(\mu)$

Reminder

$E = E(Q_d)$ and $F = \{\{v, \overline{v}\} \colon v \in V(Q_d)\}$

- $\mu(x,y) = 1$ for every $\{x,y\} \in E$
- $\therefore\ \mathrm{ave}_2(\mu, E) = 1$
- $\mu(x,y) = d$ for every $\{x,y\} \in F$
- $\therefore\ \mathrm{ave}_2(\mu, F) = d$
- $\therefore\ R_{E,F}(\mu) = \dfrac{\mathrm{ave}_2(\mu,F)}{\mathrm{ave}_2(\mu,E)} = d$



For bounding $R_{E,F}(\nu)$

$$\begin{aligned} R_{E,F}(\nu) &= \frac{\mathrm{ave}_2(\nu,F)}{\mathrm{ave}_2(\nu,E)} \\ &= \frac{\sqrt{\frac{1}{|F|}\sum_{\{u,v\}\in F}\nu(u,v)^2}}{\sqrt{\frac{1}{|E|}\sum_{\{u,v\}\in E}\nu(u,v)^2}} \\ &= \sqrt{\frac{\frac{1}{2^{d-1}}\sum_{\{u,v\}\in F}\nu(u,v)^2}{\frac{1}{d2^{d-1}}\sum_{\{u,v\}\in E}\nu(u,v)^2}} \\ &= \sqrt{d\frac{\sum_{\{u,v\}\in F}\nu(u,v)^2}{\sum_{\{u,v\}\in E}\nu(u,v)^2}} \end{aligned}$$

$\therefore$ Enough to show $\sum_{\{u,v\}\in F}\nu(u,v)^2 \le \sum_{\{u,v\}\in E}\nu(u,v)^2$.

Subclaim

Notation

$\nu^2(E) = \sum_{\{u,v\}\in E} \nu(u,v)^2$, $\nu^2(F) = \sum_{\{u,v\}\in F} \nu(u,v)^2$

Subclaim

$\nu^2(F) \le \nu^2(E)$

Proof outline:

- Induction on d
- When $d = 2$, a direct calculation
- When $d > 2$, make use of a product structure of cubes

Proof of Subclaim: $d = 2$

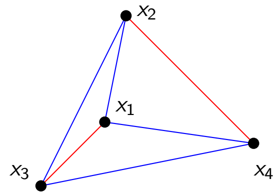
To prove the case $d = 2$ we use the following

Lemma (short diagonals lemma)

For any four points x_1, x_2, x_3, x_4 in a Euclidean space

$$\|x_1 - x_3\|_2^2 + \|x_2 - x_4\|_2^2 \le \|x_1 - x_2\|_2^2 + \|x_2 - x_3\|_2^2 + \|x_3 - x_4\|_2^2 + \|x_4 - x_1\|_2^2$$

Proof: Exercise

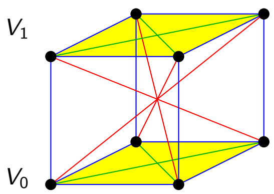


Proof of Subclaim when $d = 2$

- Set $x_i = f(v_i)$ and use the lemma above
- LHS $= \nu^2(F)$ and RHS $= \nu^2(E)$

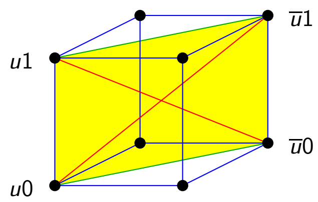
Proof of Subclaim: $d > 2$ (1)

- Partition the vertex set $V = \{0,1\}^d$ into
  - $V_0 = \{0,1\}^{d-1} \times \{0\} = \{(u,0) \colon u \in \{0,1\}^{d-1}\}$
  - $V_1 = \{0,1\}^{d-1} \times \{1\} = \{(u,1) \colon u \in \{0,1\}^{d-1}\}$
- V_i induces a $(d-1)$-dimensional Hamming cube in Q_d
- E_i = the edge set of the Hamming cube induced by V_i
  F_i = the set of pairs w/ dist. $d-1$ in the induced cube
- $\nu^2(F_i) \le \nu^2(E_i)$ by induction hypothesis
- $E_{01} = E \setminus (E_0 \cup E_1)$



Proof of Subclaim: $d > 2$ (2)

- Each pair in F has a form $\{(u,0), (\overline{u},1)\}$
- By the short diagonals lemma
  $\nu(u0, \overline{u}1)^2 + \nu(u1, \overline{u}0)^2 \le$
  $\nu(u0, \overline{u}0)^2 + \nu(\overline{u}0, \overline{u}1)^2 + \nu(\overline{u}1, u1)^2 + \nu(u1, u0)^2$
- By adding up the inequalities over all u
  $\nu^2(F) \le \nu^2(E_{01}) + \nu^2(F_0) + \nu^2(F_1)$
- $\therefore \nu^2(F) \le \nu^2(E_{01}) + \nu^2(E_0) + \nu^2(E_1) = \nu^2(E)$



Contents

- Introduction (10 min)
- Embedding into ℓ_∞ (10 min)
- Embedding into ℓ_2 (30 min)
- Lower bound for ℓ_2 (30 min)
- Remarks (10 min)

Low-distortion embeddings into ℓ_2

Summary

Every n-point metric space can be embedded
- into ℓ_∞ isometrically,
- into ℓ_2 with distortion $O(\log n)$ and this is tight.

We may wonder about
- ℓ_1 and ℓ_p?
- restricted classes of metric spaces?

What about ℓ_p?

Theorem (Bourgain '85)

For any fixed $p \in [1, \infty)$, every n-point metric space can be embedded into ℓ_p with distortion $O(\log n)$

Proof idea:
- Modify the proof for ℓ_2 (exercise)
- Use the following fact

Theorem (a consequence of Dvoretzky's thm)

Every n-point set in ℓ_2 can be isometrically embedded into ℓ_p for any $p \in [1, \infty)$

Tightness?

Theorem (Matoušek '97)

The distortion of $\Omega(\log n)$ is needed (again by expanders)

Restricting the classes of graphs

$\mathcal{G}$ a class of (connected) graphs

Def.: $\mathcal{G}$-metric

A finite metric space (X, μ) is a $\mathcal{G}$-metric if $\exists$ a graph $G \in \mathcal{G}$ and an edge-weight function w s.t. X can be isometrically embedded into the shortest-path metric on G with w

- Tree metric: $\mathcal{G}$ the class of trees
- Outerplanar-graph metric: $\mathcal{G}$ the class of outerplanar graphs
- Planar-graph metric: $\mathcal{G}$ the class of planar graphs
- ...

Tree metrics

Tree: a connected graph with no cycle

Theorem

Every n-point tree metric can be

- isometrically embedded into ℓ_1 (exercise)
- embedded into ℓ_p with distortion $O((\log\log n)^{\min\{1/2,1/p\}})$ (Matoušek '99)
  - this is tight (Bourgain '86)

A bad example is a complete binary tree with unit weight.

Outerplanar-graph metrics

Outerplanar graph: a graph that can be drawn on the plane with no edge crossings and all edges incident to the outer face

Theorem

Every n-point outerplanar-graph metric can be

- isometrically embedded into ℓ_1 (Okamura and Seymour '81)
- embedded into ℓ_2 with distortion $O(\sqrt{\log n})$ (Rao '99)

As far as I surveyed, tightness for ℓ_2 doesn't seem to be known

Planar-graph metrics

Planar graph: a graph that can be drawn on the plane with no edge crossings

Theorem

Every n-point planar-graph metric can be embedded into ℓ_2 with distortion $O(\sqrt{\log n})$ (Rao '99)

- this is tight (Newman and Rabinovich '03)

Conjecture (Gupta, Newman, Rabinovich, and Sinclair '04)

A planar-graph metric can be embedded into ℓ_1 with constant distortion

- They proved it is the case for series-parallel graphs
- They further conjecture that the shortest-path metric on an H-minor-free graph (for H fixed) can be embedded into ℓ_1 with constant distortion

Girths and distortions

The girth of a graph: the length of a shortest cycle

Theorem (Linial, Magen, and Naor '02)

An embedding of the shortest-path metric on an r-regular graph ($r \geq 3$) of girth g with unit weight into ℓ_2 requires the distortion $\Omega(\sqrt{g})$

Conjecture (Linial, London, and Rabinovich '95)

The tight bound is $\Omega(g)$

Note: Upper bound $O(g)$ is easy when r is constant

(a consequence of one exercise)

Lots of missing aspects

No time to mention ...

- Dimension into ℓ_2, ℓ_∞, ℓ_p and trade-offs
- Embeddings into probabilistic tree metrics
- Efficiency of the construction
- Other important special metrics (e.g., edit distance, Hausdorff distance)
- Dimension reduction (e.g., Johnson–Lindenstrauss Lemma)
- Algorithmic applications
- ...

One of the most developing subjects in discrete and computational geometry

Contents

[End of Lecture]