Inapproximability for metric embeddings into $\mathbb{R}^d$

Jiri Matousek (Charles University)
Anastasios Sidiropoulos (U. of Toronto)
Metric spaces

Metric space $M = (X, D)$

- Positive definiteness
  $D(p, q) = 0$ iff $p = q$
- Symmetry
  $D(p, q) = D(q, p)$
- Triangle inequality
  $D(p, q) \leq D(p, r) + D(r, q)$
Metric spaces

Metric Spaces

Euclidean Spaces
Metric embeddings

[Bourgain ’85]

Finite Metric Spaces
\( n \)-point

\( O(\log n) \)

Euclidean Spaces
Metric embeddings

- Given spaces $M=(X,D)$, $M'=(X',D')$
- Mapping $f:X \rightarrow X'$
- Distortion $c$ if:

$$D(x_1,x_2) \leq D'(f(x_1),f(x_2)) \leq c \cdot D(x_1,x_2)$$
Motivation

• Geometric interpretation
• Succinct data representation
  – Embedding into low-dimensional spaces
• Visualization
  – Embedding into the plane
  – Multi-dimensional scaling
• Optimization
  – Embedding into “easy” spaces
## Known results

<table>
<thead>
<tr>
<th>Host space</th>
<th>Distortion</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O(\log n)$ –dimensional $L_2$ (also true for $L_p$)</td>
<td>$O(\log n)$</td>
<td>[Bourgain ’85], [Johnson-Lindenstrauss], [Alon], [Linial, London, Rabinovich ’94], [Abraham, Bartal, Neiman ‘06]</td>
</tr>
</tbody>
</table>
| $d$-dimensional $L_2$                            | $\tilde{O} \left( n^{\text{const}/d} \right)$ | [Matousek ’90]  
Also: [Gupta ‘99], [Babilon, Matousek, Maxova, Valtr 2003], [Badoiu, Chuzhoy, Indyk, S ‘06], [Bateni, Demaine, Hajiaghayi, Moharrami 2007] |

Random projection is optimal in the worst case!
Random projection
Absolute vs. Relative embeddings

• Small dimension $\rightarrow$ high distortion $\left(n^{\Omega(1/d)}\right)$
  - E.g. embedding a cycle into the line
• What if a particular metric embeds with small distortion?
• Computational problem:
  Approximate best possible distortion
Known results on approximation

- Into $\mathbb{R}^1$
  - Unweighted graphs: $n^{1/2}$-approx, $1.01$-hard [BDGRRRS ‘05]
  - Trees: $n^{1-a}$-approx, $n^b$-hard [Badoiu, Chuzhoy, Indyk, S ‘05]
  - General metrics: $(\text{OPT} \cdot \log n)^{O(\sqrt{\log \Delta})}$ [Badoiu, Indyk, S‘07]

- Into $\mathbb{R}^d$
  - Ultrametrics: $\log^6 \Delta$-approx, $\text{NP}$-hard [Badoiu, Chuzhoy, Indyk, S ‘06], [Onak, S ‘08]
  - General metrics: $\tilde{O} (n^{2/d})$ worst case [Matousek ‘90]
    $\Omega(n^{1/22d})$-hard [Matousek, S ‘08]

Random projection is a near-optimal approximation algorithm for general metrics (unless P=NP)!
Reduction outline

Embedding into $\mathbb{R}^1 \rightarrow$ Embedding into $\mathbb{R}^d$

**Theorem** [Badoiu, Chuzhoy, Indyk, S ‘05]
Embedding into $\mathbb{R}^1$ is NP-hard to approximate within $n^{1/12}$

\[\downarrow\]

**Theorem**
For $d \geq 2$, embedding into $\mathbb{R}^d$ is NP-hard to approximate within $n^{1/22d}$
Reduction outline

Reduction from embedding into $\mathbb{R}^1$

Product with $S^{d-1}$

$M \xrightarrow{c} \mathbb{R}^1$ iff $M' \xrightarrow{O(c)} \mathbb{R}^d$
Reduction outline (easy direction)

\[ R^1 \rightarrow \text{product} \rightarrow R^d \]
Reduction outline (hard direction)
Nesting lemma

\[ f_1, f_2 : S^{d-1} \to \mathbb{R}^d \text{ continuous} \]

- Non-intersecting
- \( |f_i(x) - f_i(y)| > |x - y| - \varepsilon \)
- \( |f_1(x) - f_2(x)| < \varepsilon \)

One sphere is "inside" the other!

Ideas from [Vaisala ‘08]
Proof of nesting lemma: Techniques

- discrete spaces
  - topological spaces
    - (d-1)-dimensional cohomology groups
      - Alexander duality
        - 0-dimensional homology groups
  - discrete embeddings
    - continuous functions
      - homomorphisms

Extension

Cohomology
What if OPT=O(1)?

• It is NP-hard to distinguish between metrics that embed into $\mathbb{R}^d$ with distortion $n^{a/d}$ vs $n^{b/d}$ ($a<b$)

• Can we distinguish between $O(1)$ vs $n^{b/d}$?

**NO! (for $d \geq 3$)**
Improved reduction for $d \geq 3$
Further directions

• Intriguing open problem:
  Embedding into $\mathbb{R}^d$, $d \leq 2$.
  Is there an algorithm achieving distortion $\text{OPT}^{O(1)}$?

• Minimize the dimension.