CSC2414 - Metric Embeddings*
Lecture 1: Embnedding trees into Euclidian space

Notes taken by

Summary: In this tutorial we discuss a result of Bourgain which says that every binary tree can be embedded with distortion $O(\sqrt{\log \log n})$ into ℓ_2.

1 Introduction

In this tutorial we discuss a result of Bourgain regarding the distortion of the Euclidian embeddings of a binary tree. First let us notice that every tree metric is ℓ_1.

Theorem 1.1. *Let T be a tree. There exists an isometric embedding of T into ℓ_1^n.*

Proof. The proof is by induction on n, the number of vertices. The base of induction is trivial. Let T be a tree on n vertices and v be a leaf of T. By induction hypothesis there exists an isometric embedding f from $T - v$ to ℓ_1^{n-1}. Define the map g from f to ℓ_1^n by $g(u) := f(u) \oplus 0$ for every $u \in T - v$ and $g(v) := f(w) \oplus 1$ where w is the unique neighbor of u in T. Now it is easy to see that g is an isometric embedding of T into ℓ_1^n. □

Now let us consider the Euclidian space, ℓ_2. Let T be a complete binary tree and f be its embedding into ℓ_1^n as it is defined in the proof of Theorem 1.1. What is the distortion of the embedding f when we consider it as a mapping into ℓ_2^n? It is easy to see that the new distance of two vertices u and v in ℓ_2 is $\sqrt{d_T(u,v)}$. Since the maximum value of $d_T(u,v)$ is $2\log_2 n$, we conclude that the distortion of f is $\sqrt{2\log_2 n}$. But this is not optimal. Bourgain in a short paper showed that it is possible to improve this bound to $O(\sqrt{\log \log n})$, and this is sharp.

Theorem 1.2. *Let T be the complete binary tree of height h. It is possible to embed T into ℓ_2 with distortion $O(\sqrt{h})$.*

Proof. Label the vertices of T by $1, \dots, n$, where $n = 2^{h+1} - 1$ is the number of vertices, and denote by $h(i)$ the distance of i from the root of the tree. Consider the embedding $f : T \to \ell_2^n$ defined as

$$f(i)_j = \begin{cases} \sqrt{1 + h(i) - h(j)} & j \text{ is an ancestor of } i \\ 0 & \text{otherwise} \end{cases}$$

* Lecture Notes for a course given by Avner Magen, Dept. of Computer Sciecne, University of Toronto.

For example if T is a complete binary tree with three vertices labelled as $1, 2, 3$ (where 1 is the root). Then

$$\begin{aligned} f(1) &= (1, 0, 0) \\ f(2) &= (\sqrt{2}, 1, 0) \\ f(3) &= (\sqrt{2}, 0, 1). \end{aligned}$$

First let us notice that the contraction of f is a constant. Consider two vertices i and j of distance d, and the shortest path between them $i_0(= i), i_1, \dots, i_{d-1}, i_d(= j)$. Suppose that i_r is the common ancestor of i and j. Then for $0 \leq k < r$ we have $f(i)_{i_k} = \sqrt{k+1}$ while $f(j)_{i_k} = 0$. Also for $r < k \leq d$, we have $f(i)_{i_k} = 0$ while $f(j)_{i_k} = \sqrt{d-k+1}$. This shows that

$$\|f(i) - f(j)\|_2 \geq \sqrt{\sum_{k=0}^{r-1}(k+1) + \sum_{k=r+1}^{d}(d-k+1)} \geq \sqrt{\sum_{k=1}^{d/2} k} \geq \frac{\sqrt{d}}{2}.$$

So it remains to compute the expansion of f. Lemma 1.3 below shows that we only need to compute the expansion on the edges of the tree. Let ij be an edge of the tree where i is the parent of j. The two vectors $f(i)$ and $f(j)$ are the same except on the coordinates corresponding to the vertices of the path from i to the root. For any vertex t on this path $f(j)_t = \sqrt{1 + h(j) - h(t)}$ while $f(i)_t = \sqrt{1 + h(i) - h(t)} = \sqrt{h(j) - h(t)}$. So

$$\begin{aligned} \|f(i) - f(j)\|_2 &= \sqrt{\sum_{k=0}^{h(j)}(\sqrt{k+1} - \sqrt{k})^2} = \sqrt{\sum_{k=0}^{h(j)}(\frac{1}{\sqrt{k+1} + \sqrt{k}})^2} \\ &\leq \sqrt{\sum_{k=0}^{h(j)}(\frac{1}{2\sqrt{k+1}})^2} \approx \frac{\sqrt{\ln h(j)}}{2} \leq \sqrt{h}/2. \end{aligned}$$

□

Lemma 1.3. *Let G be a weighted graph and d, its corresponding metric. Let (X, ρ) be a metric space and $f : (G, d) \to (X, \rho)$. Then the expansion of f is*

$$A := \max_{uv \in G} \frac{\rho(f(u), f(v))}{d(u, v)}$$

Proof. Consider two arbitrary vertices v and w and the shortest path between them $v_1(= v), \dots, v_k(= w)$. Then

$$\begin{aligned} \frac{\rho(f(v), f(w))}{d(v, w)} &\leq \frac{\sum_{i=1}^{k-1} \rho(v_i, v_{i+1})}{\sum_{i=1}^{k-1} d(v_i, v_{i+1})} \leq \\ & \frac{\sum_{i=1}^{k-1} d(v_i, v_{i+1}) \left(\max_{uv \in G} \frac{\rho(f(u), f(v))}{d(u,v)}\right)}{\sum_{i=1}^{k-1} d(v_i, v_{i+1})} = \max_{uv \in G} \frac{\rho(f(u), f(v))}{d(u, v)}. \end{aligned}$$

□

N. Linial, A. Magen and M. Saks showed that every tree can be embedded into the Euclidian space with distortion $O(\log\log n)$.

Theorem 1.4. *Let T be a tree on n vertices. It is possible to embed T into ℓ_2 with distortion $O(\log\log n)$.*

The result is improved by Matoušek.

Theorem 1.5. *Let T be a tree on n vertices. It is possible to embed T into ℓ_2 with distortion $O(\sqrt{\log\log n})$.*