Seminar on Algorithms and Geometry

Lecture 1

March 18, 2009

Lecturer: Robert Krauthgamer   Scribe by: Noam Arkind   Updated: March 22, 2009

1 Introduction

1.1 Basic Definitions

Definition 1 *A **Metric Space** $(X; d)$, is defined by the set X and a metric $d : X \times X \to \mathbb{R}^+$ i.e. $d(x, y)$ $(x, y \in X)$ is non-negative, symmetric and such that the triangle inequality holds*

$$d(x, y) \le d(x, z) + d(z, y) \quad \forall x, y, z \in X.$$

Definition 2 *Let $(X; d_X)$ and $(Y; d_Y)$ be two metric spaces. An **Embedding** of X into Y is a map $f : X \to Y$. It is said to have **distortion** $D \ge 1$ if $\exists r \ge 0$ such that*

$$d_X(x, y) \le r d_Y(f(x), f(y)) \le D d_X(x, y)) \quad \forall x, y \in X.$$

*We sometimes write X embeds into Y as $X \hookrightarrow Y$. If $D = 1$ then f is called **Isometry**.*

Remark *In this definition r is a scaling factor and most of the times it will be absorbed in f.*

Remark *In this course we will usually consider the case when the metric space is finite, i.e. $|X| = n < \infty$.*

Remark Potentially two problems may occur when searching for an embedding, low dimension of the image (for example 3-D lattice embedding into a 2-D plane will have large distortion), or if the target is not rich enough and too constrained.

Definition 3 *The metric space l_p^k is the set $\mathbb{R}^k$ with the metric $d(x, y) = \|x - y\|_p$, where*

$$\|z\|_p = \left(\sum_{i=1}^{k} |z_i|^p\right)^{\frac{1}{p}} \quad \forall 1 \le p < \infty$$
$$\|z\|_\infty = \max_{1 \le i \le k} |z_i|.$$

The metric space l_p is the same, but for infinite p-summable sequences x, (i.e. $\sum_{i=1}^{\infty} |x_i|^p < \infty$).

1.2 Motivation

The motivation for these definitions, is that in many cases we want to simplify our problem by mapping our space $(X; d)$, that might be very complicated, into a space with simpler geometry, such that the original structure is preserved. An example of an application, is to embed a network into l_p^2, this will enable us to keep only two coordinates for each node, and still find the distance between two nodes easily by direct calculation instead of computing the shortest path. For example, it is of interest to answer questions of the form

- Can we embed a n-cycle with $d(x, y) =$[length of shortest path] into a tree with the same metric?
- $\{0, 1\}^k$ with l_1 norm $\hookrightarrow l_2$?
- $\{0, 1\}^k$ with l_2 norm $\hookrightarrow l_1$?
- n-path with $d(i, j) = \sqrt{|i - j|} \hookrightarrow l_1$ or l_2?

2 Embeddings into l_∞

Definition 4 *A* ***Frechet Embedding*** $f : X \to \mathbb{R}^k$ *is defined for some collection of sets* $A_1, \ldots, A_k \subseteq X$, *as* $f_i(x) \stackrel{\text{def}}{=} \{f(x)\}_i = d_X(x, A_i) \stackrel{\text{def}}{=} \min_{a \in A_i} d_X(x, a)$.

Fact 5 *(non-expansive)* $|f_i(x) - f_i(y)| \leq d(x, y), \forall x, y \in X$.

Proof Let $x' = \arg\min_{a \in A_i} d(x, a)$, then $f_i(y) \leq d(y, x')$ by definition, and $d(y, x') \leq d(y, x) + d(x, x')$ (where $d(x, x') = f_i(x)$) by triangle inequality. ■

Proposition 6 *Every n-point metric embeds isometrically into* l_∞^n.

Proof Let $X = \{1, \ldots, n\}$, we use Frechet Embedding for $A_j = \{j\}, \forall 1 \leq j \leq n$. Fix $j, k \in X$, then on the one hand by the previous fact $\|f(j) - f(k)\|_\infty = \max_i |f_i(j) - f_i(k)| \leq d_X(j, k)$. But on the other hand, $f_j(j) = 0$, hence $\|f(j) - f(k)\|_\infty = \max_i |f_i(j) - f_i(k)| \geq |f_j(k)| = d(j, k)$. ■

Theorem 7 (Matousek 96', based on Bourgain 85') *Let* $q \geq 1$ *be an integer, then any n-point metric space embeds with distortion* $2q - 1$ *into* l_∞^k, *where* $k = O(qn^{\frac{1}{q}} \log n)$.