CS-67720 Metric Embedding Theory and Its Algorithmic Applications

Lecture 1

Lecturers: Nova Fandina and Yair Bartal *Fall 2018/19*

In this course we present various topics of Metric Embedding theory, ranging from the classical old results to the newest developments of the field. Particularly, we will discuss the following topics:

Metric/Normed Spaces, Embeddings: basic definitions, distortion, average distortion, more notions of quality of embedding.
Dimension Reduction in ℓ_p: known upper/lower bounds and open questions.
Embeddings into normed spaces. Embeddings into trees.
Partial Embeddings. Ramsey theory of metric spaces.
Probabilistic embeddings.
Applications to approximation theory: Distance Oracle, Nearest Neighbor Search, Sparsest Cut and more.

Metric embedding is a widely used technique in computer science and related disciplines. Let us give a very general example of usefulness of a metric embedding. Suppose we have a (big) set of objects, with defined pairwise distances. The distances are given in the (big) table. From such table is quite difficult to "understand" the structure of the set. If for each object we could match a point on the plane, such that for every pair of the resulting points the distance equals the original distance, then we could get a clear graphical representation of the given set. With such representation we can analyze the set using geometric methods. In the above example we embedded the set into Euclidean plain, while keeping the distances exactly equal to the original ones - isometric embedding. Not always isometric embedding is possible. In such cases we wish to keep resulting distances as close to the original as possible, i.e. with minimum distortion. We will see later in the course that such embeddings, low distortion embeddings, are very useful in the theory of approximation algorithms: the general methodology is to translate the problem over a "difficult" metric space into the problem over a "simple" one, with provided guarantees on the loss of precision.
Historical notes. The study of metric spaces takes its roots in the study of dissimilarities of normed spaces that was carried by mathematicians in the first half of the 20th century, such as Banach, Mazur, John, Dvoretzky, etc. In 1985 Bourgain have studied the quantitative question of the minimum distortion needed to embed a metric space into a normed space. In his paper he provided the first such result: any n-point metric space embeds into ℓ_p with distortion $O(\log n)$. The study of the dissimilarities of normed spaces continued through the 80's-90's, again by the mathematicians: Pisier, Talagrand, Schechtman, Milman, etc. In 1995 the seminal paper of Linial, London and Rabinovich has shown the algorithmic importance of the theorem of Bourgain and other results from the theory of normed spaces. Since then, the theory of metric spaces is studied by both communities, by mathematicians as well as by CS community.

The main goal of this lecture is a review of the most intriguing (to the taste of the lecturers) concepts and results in the metric embedding theory.

1.1 Metric Spaces

Definition 1.1. *Metric space is a pair* (X, d)*, where* X *is a set of points, and* $d\colon X^2 \to \mathbb{R}^+$ *is a distance function, such that for all* $u, v, w \in X$*:*

- *reflexivity:* $d(u, v) = 0 \iff u = v$*;*
- *symmetry:* $d(u, v) = d(v, u)$*;*
- *triangle inequality:* $d(u, w) \leq d(u, v) + d(v, w)$*.*

If a distance function satisfies all the properties except $d(u, v) = 0 \Rightarrow u = v$, then the space is called a *semi-metric space* or *pseudo-metric space*. If the distance function satisfies the first and the second properties, it's called a *quasimetric*.

Examples of metric spaces:

1. Let $X = \{x_1, x_2, \ldots, x_n\}$, and put $\forall i \neq j \quad d_X(x_i, x_j) = a$, for a positive $a > 0$, and 0 otherwise. We cal it an **equilateral** metric space.

2. Let $X = \mathbb{R}$, and put $d(x, y) = |x - y|$. This is the usual metric on the real line.

3. Graph Defined Metric Spaces.

   Let $G = (V, E, w)$ be a connected, undirected, weighted graph such that $w\colon E \to \mathbb{R}^+$. We define a metric space on the nodes of G, (V, d_G), where $\forall u, v \in V$, $d_G(u, v)$ is the weight of a lightest path between u and v in G. It is easily checked that d_G is a metric function.

   Note that given any metric space we can define a weighted complete graph of the metric space.

Definition 1.2. *Normed space is a linear vector space* V *(we will focus on vector spaces defined over* $\mathbb{R}$*) equipped with a norm function* $\|\cdot\| : V \to \mathbb{R}^+$ *such that for all* $u, v \in V, \alpha \in \mathbb{R}$*:*

- *reflexivity:* $\|v\| = 0 \iff v = 0_V$*;*
- *linearity by scalar multiplication:* $\|\alpha v\| = |\alpha| \cdot \|v\|$*;*
- *triangle inequality:* $\|u + v\| \leq \|u\| + \|v\|$*.*

If we remove the first requirement we obtain a *seminorm*.

Claim 1.1. *Let* $(V, \|\cdot\|)$ *be a normed space. Define a function* $d(u, v) = \|u - v\|$ *for all* $u, v \in V$*. Then* (V, d) *is a metric space.*

Following are the typical norms defined on $\mathbb{R}^n$.

ℓ_2-norm: $\|x\|_2 = \left(\sum_{i=1}^n |x_i|^2\right)^{\frac{1}{2}}$; *$\ell_1$-norm* : $\|x\|_1 = \sum_{i=1}^n |x_i|$; *$\ell_\infty$-norm*: $\|x\|_\infty = \max_{i \in [n]}\{|x_i|\}$.

In general, for any $p \geq 1$, define $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$.

It can be shown (Minkowski Inequality) that this function is a norm function. We note that for $0 < p < 1$ the above function is not a norm function, since the triangle inequality does not hold, yet the function $d(u,v) = \sum_{i=1}^{n} (u_i - v_i)^p$ is a proper metric function.

We denote the normed space $(\mathbb{R}^n, \|\cdot\|_p)$ by ℓ_p^n.

Definition 1.3. *The unit ball of the given normed space* $(V, \|\cdot\|)$ *is a set* $B_1 = \{u \in V | \, \|u\| \leq 1\}$. *The unit sphere is the set* $S_1 = \{u \in V | \, \|u\| = 1\}$.

Note that the whole normed space is defined by the unit sphere (since the norm function is linear by scalar multiplication). In Fig. 1.1 the unit balls of several normed spaces are presented.

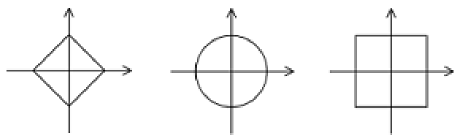

Figure 1.1: Unit balls of normed spaces.

1.1.1 Basic Properties of Normed Spaces

Claim 1.2. *For all* $1 \leq p \leq q \leq \infty$, *and for each* $x \in \mathbb{R}^n$: $\|x\|_p \geq \|x\|_q$.

Proof. For $p, q \neq \infty$, let $P(y) = y^{\frac{q}{p}}$. Since $\frac{q}{p} \geq 1$ the function $P(\cdot)$ is convex. In addition $P(0) = 0$, therefore $P(\cdot)$ is superadditive. Namely, $P\left(\sum_{i=1}^{n} y_i\right) \geq \sum_{i=1}^{n} P(y_i)$. Let $y_i = |x_i|^p$. Then

$$P\left(\sum |x_i|^p\right) \geq \sum P\left(|x_i|^p\right) \Rightarrow \left(\sum |x_i|^p\right)^{\frac{q}{p}} \geq \sum \left(|x_i|^p\right)^{\frac{q}{p}} = \sum |x_i|^q \Rightarrow \left(\sum |x_i|^p\right)^{\frac{1}{p}} \geq \left(\sum |x_i|^q\right)^{\frac{1}{q}}.$$

For $q = \infty$, $(\|x\|_\infty)^p = \max\{|x_i|^p\} \leq \sum_{i=1}^{n} |x_i|^p = \left(\|x\|_p\right)^p$. $\square$

And the following upper bound can be proved similarly, using the Jensens's inequality.

Claim 1.3. *For all* $1 \leq p \leq q \leq \infty$, *and for each* $x \in \mathbb{R}^n$: $\|x\|_p \leq \|x\|_q n^{\frac{1}{p} - \frac{1}{q}}$.

Lemma 1.4. Jensens's inequality
Let $P(\cdot)$ *be any real convex function. Then for any* $n \geq 1$ *and any* $y_i, 1 \leq i \leq n$ *from the domain of* $P(\cdot)$ *it holds that:*

$$P\left(\frac{\sum_{i=1}^{n} y_i}{n}\right) \leq \frac{\sum_{i=1}^{n} P(y_i)}{n}.$$

Inequality 1.5. Cauchy-Schwarz inequality
$\forall x, y \in \mathbb{R}^n \quad |\langle x, y\rangle| \leq \|x\|_2 \cdot \|y\|_2$.

Following is the generalization of the Cauchy-Schwarz inequality.

Inequality 1.6. Holder inequality
$\forall x, y \in \mathbb{R}^n$ *and* $\forall 1 \leq p, q \leq \infty$ *so that* $\frac{1}{p} + \frac{1}{q} = 1$ *the following holds:* $|\langle x, y\rangle| \leq \|x\|_p \cdot \|y\|_q$.

1.2 Embeddings: Basic Definitions

Definition 1.4. Embedding
Let (X, d_x) and (Y, d_y) be any metric spaces. A map $f: X \to Y$ is called an embedding.

Definition 1.5. Isometry
An embedding $f : X \to Y$ is an isometry, if $\forall u, v \in X$, $d_x(u, v) = d_y(f(u), f(v))$.

Examples. An equilateral space on 3 points isometrically embeds into ℓ_2^2 - equilateral triangle. (In fact, any 3-point metric space isometrically embeds into ℓ_2^2). An equilateral space on n points isometrically embeds into ℓ_2^{n-1} – the vertices of the $(n-1)$-dimensional simplex. It can be shown that the number of dimensions needed for isometric embedding is at least $n - 1$.(Exercise.)

Even more, there exist metric spaces that do not admit isometric embedding into ℓ_2 of any dimension.

Claim 1.7. *There exists a 4-point metric space that cannot be isometrically embedded into ℓ_2.*

Proof. Let $X = \{A, B, C, D\}$,with $d_x(A, B) = d_x(B, C) = d_x(C, D) = d_x(D, A) = 1$, and $d_x(A, C) = d_x(B, D) = 2$. Clearly, (X, d_x) is a metric space. Note that this is the metric space that is defined on the weighted cycle graph on 4 vertices with the unit weighs on the edges. Such graph is denoted by C_4.

Assume by contradiction that $f: X \to l_2$ is an isometry. Denote $f(A) = A'$, $f(B) = B'$, $f(C) = C'$, $f(D) = D'$. Then, we obtain that $\|A' - B'\|_2 + \|B' - C'\|_2 = \|A' - C'\|_2$. Therefore, A', B' and C' are located on the same line (since in strictly convex normed spaces $\|x + y\| = \|x\| + \|y\|$ iff x and y are on the same ray. Note that ℓ_p, for all $1 < p < \infty$ is strictly convex). Moreover, B' is a middle point of the $[A', C']$. The same observation is about D'. Namely, $B' = D'$. And therefore $\|B' - D'\|_2 = 0 \neq d_x(B, D) = 2$, contradiction.

□

In fact, the only unweighted graph metrics that are isometrically embeddable into ℓ_2 are equilateral metric space or a path metric.

Claim 1.8. *(Exercise) A finite connected unweighted graph G admits an isometric embedding into ℓ_2 if and only if it is either a complete graph K_n or a path P_n.*

Remark 1.9. *There is an algorithm that in polynomial time decides weather a given finite metric space is isometrically embeddable into ℓ_2. The same question for the ℓ_1 norm is shown to be NP-hard.*

1.2.1 Approximate Embedding: Measuring The Quality

We might want to relax the demand on the exactness of an embedding, and discuss an "approximate" embedding – a map that preserves the original distances within some *error* (distortion), which we want to measure and control. There are various natural ways to define an error of an embedding. Perhaps the classical one is a worst case multiplicative error, which we define in a sequel.

Definition 1.6. Expansion
Let $f\colon X \to Y$ be an embedding, and let $u \neq v \in X$ be any pair of points.

- Expansion of (u,v) under embedding f is given by: $expans_f(u,v) = \frac{d_y(f(u),f(v))}{d_x(u,v)}$.

- Expansion of the embedding f is given by: $expans(f) = \sup_{u\neq v\in X}\{expans_f(u,v)\}$.

Definition 1.7. Contraction
Let $f\colon X \to Y$ be an embedding, and let $u \neq v \in X$ be any pair of points.

- Contraction of (u,v) under embedding f is givenby: $contr_f(u,v) = \frac{d_x(u,v)}{d_y(f(u),f(v))}$.

- Contraction of embedding f is given by: $contr(f) = \sup_{u\neq v\in X}\{contr_f(u,v)\}$.

Definition 1.8. *Distortion of an embedding f is defined by: $dist(f) = expans(f)\cdot contr(f)$.*

We say that f is non-contractive if $contr(f) \leq 1$, and we say that f is non-expansive if $expans(f) \leq 1$.

Let us enumerate few remarks regarding the definition.

1. Distortion is a measure of quality of the worst case. Indeed, for all $x \neq y \in X$, we have: $\frac{1}{contr(f)} \leq \frac{d_y(f(x),f(y))}{d_x(x,y)} \leq expans(f)$, which implies, $d_x(x,y) \leq contr(f)\cdot d_y(f(x),f(y)) \leq dist(f)\cdot d_x(x,y)$.
2. The definition captures the scaling of the metric as embedding with distortion 1. (Another direction is also true, i.e., if embedding has distortion 1 it is a scaling map).

   Moreover, the definition allows to rescale the source and destination metrics without affecting the distortion of the embedding $f : X \to Y$. Namely, if distortion of f is α, and all the distances of X are multiplied by some constant $a > 0$, and all the distances of Y are multiplied by some constant $b > 0$, the distortion of the embedding f is still α.

Following are few fundamental results with worst case distortion guarantees, which we will cover in the course.

- **Embedding into normed spaces.** Given any finite metric space X, embed it into ℓ_p with low distortion. Bourgain has shown the folowing result: Any n-point metric space can be embedded into ℓ_p^d with $O(\log n)$ distortion, for any $p \geq 1$. Moreover, in [1] the authors have proved that $d = O(\log n)$ (alongside with a stronger guarantees, which we will discuss in a moment). In addition, the distortion is proved to be tight by Matausek (expander graphs).
- **Dimension reduction.** Given a finite $X \subset V$ in a normed space of dimension d, find an embedding of X into the subspace $W \subset V$, of smaller dimension d', while keeping low distortion. The central result is the Johnson-Lindenstrauss lemma for Euclidean sets: Given any n-point set $X \in \ell_2^d$ and any $\epsilon > 0$, there is an embedding $f : S \to \ell_2^{O\left(\frac{\log n}{\epsilon^2}\right)}$, with distortion $(1+\epsilon)$. Recently, the dimension/distortion tradeoff is shown to be tight.

- **Embedding into trees.** Tree metrics are often preferable destination metric spaces, due to their algorithmic simplicity. Yet, the following negative result can be shown (we will prove it later in the course): Every embedding of C_n into a tree metric, requires distortion at least $\Omega(n)$. To overcome this barrier, the new type of embeddings was suggested.

- **Probabilistic Embedding** The idea is to introduce some randomness into the process: Instead of embedding into a metric space Y, embed into a distribution D over a collection of metric spaces (Y_i, d_{Y_i}), i.e. for all i there is a (non-contractive) embedding $f_{Y_i} : X \to Y_i$, and there is a distribution D over the set of embeddings. The quality of a probabilistic embedding is then measured by an *expectation* of the worst case distortion. More formally, the distortion is α if for all $u \neq v \in X$, $E_D[d_Y(f_Y(u), f_Y(v))] \leq \alpha d_X(u, v)$. In the course, we will study probabilistic embeddings into tress. The central result is due to [3], [5]: Any n-point metric space X can be embedded into a distribution of tree metrics, with expected distortion of $\Theta(\log n)$.

- **Ramsey Type Embeddings.** The quest is to find an order in disorder. Given a finite metric space X, find a largest subspace $Y \in X$ that can be embedded into a simple metric Z with low distortion. We will show the following result: In any n-point metric space X there exists a subspace $Y \subseteq X$ of size at last $n^{1-\epsilon}$, which admits an embedding into a tree with distortion $O(1/\epsilon)$.

- **Average Distortion.** The straightforward generalization to the worst case distortion is the *average distortion*: How well an embedding performs on average over the set of all pairs of the metric space. For an embedding $f : X \to Y$ we define the following notions.

  Definition 1.9. *(* **Distortion of a pair)** *For any $u \neq v \in X$, the distortion of the pair with respect to f is $dist_f(u, v) = \max\{contr_f(u, v), exp_f(u, v)\}$.*

  Informally, this is the multiplicative change the pair (u, v) experiences under f.

  Definition 1.10. *(* **Average Distortion / ℓ_q-distortion)** *For all $1 \leq q < \infty$, we define*

$$\ell_q\text{-}dist(f) = \left(\frac{\sum_{u \neq v \in X} (dist_f(u, v))^q}{|\binom{X}{2}|} \right)^{\frac{1}{q}},$$

  and

$$\ell_\infty\text{-}dist(f) = \sup_{u \neq v \in X} \{dist_f(u, v)\}.$$

  The above definition of ℓ_q-distortion can be generalized to work with any distribution over the pairs of X, for any finite or infinite metric space. (In the above definition we have used a uniform distribution).

  One of the central results on the ℓ_q-distortion measure is the following variation on Bourgain's embedding [1]: For any finite X there is an embedding $f : X \to \ell_p^{O(\log n)}$ with $\ell_q\text{-}dist(f) = O(q)$, and with worst case distortion of $O(\log n)$. Another interesting result is about dimension reduction: It can be shown that any finite Euclidean set embeds into ℓ_2^k, for any given $k \geq 1$, with ℓ_q-distortion of $1 + O(\alpha(q, k))$, $q < k$. For the embedding into trees the following result is known: Deterministic algorithm of any finite metric space into a tree with constant average distortion; and probabilistic embedding with ℓ_q-dist of $O(q)$ and worst case distortion of $O(\log n)$.

- **Terminal/Prioritized Distortion.** In the setting of the embedding with terminal distortion, we are given a metric space X and a set of *terminal* points $T \subseteq X$, $|T| = t$. The aim is to embed X such that all the pairs (x, z), $x \in X, z \in T$ have distortion of $\alpha(t)$. Further generalization of this notion is *prioritized* distortion: given a priority ranking $\pi = (x_1, x_2, \ldots, x_n)$ of the points of X, we seek an embedding such that for all pair x_i, x_j, with $j < i$ the distortion of this pair is $\alpha(j)$. There are various results in this setting, let us state a few that were proved in [4]: Given any finite space X and any priority ranking π, there is an embedding of X into a tree metric with prioritized distortion $O(j \log j (\log\log j)^2)$. And of course, the variation on Bourgain's embedding: there is an embedding of X into $\ell_p^{O(\log n)}$ with prioritized distortion $O(\log j)$.

Now, we are ready to see an examples of some "real" embeddings!

1.3 Universality of ℓ_∞

We start with the following interesting and basic fact in the theory of metric mbeddings.

Theorem 1.10. **[Fréchet]** *Any n-point metric space is isometrically embeddable into ℓ_∞^{n-1}.*

Proof. Let $X = (\{z_1, z_2, \ldots z_n\}, d_X)$. We define an embedding $f\colon X \to l_\infty^{n-1}$ as follows:

$$\forall z_i \in X, \quad f(z_i) = (d_X(z_1, z_i), d_X(z_2, z_i), \ldots, d_X(z_i, z_i), \ldots, d_X(z_{n-1}, z_i)).$$

Let $1 \le i < j \le n$, then, $\|f(z_i) - f(z_j)\|_\infty = \max_{1 \le k \le n-1}\{|d_X(z_i, z_k) - d_X(z_k, z_j)|\}$. For all $1 \le k \le n-1$, and for all $1 \le i, j \le n$, by triangle inequality of $d_X(\cdot)$ it holds that

$$|d_X(z_i, z_k) - d_X(z_k, z_j)| \le d_X(z_i, z_j).$$

Therefore,

$$\max_{1 \le k \le n-1}\{|d_X(z_i, z_k) - d_X(z_k, z_j)|\} \le d_X(z_i, z_j).$$

On the other hand, since $i < j$, it holds that $i \neq n$ and

$$\max_{1 \le k \le n-1}\{|d_X(z_i, z_k) - d_X(z_k, z_j)|\} \ge |d_X(z_i, z_i) - d_X(z_i, z_j)| = |d_X(z_i, z_j)| = d_X(z_i, z_j).$$

Therefore, $\|f(z_i) - f(z_j)\|_\infty = d_X(z_i, z_j)$ as required. □

In fact, Fréchet proved that any *separable* metric space admits an isometric embedding into ℓ_∞ (of infinite dimension). The construction is the same as above, where the destinations points are the points of the countable dense set. Note that we can recognize the above embedding as embedding into ℓ_p^{n-1}, for any $p \neq \infty$ with distortion of $(n-1)^{\frac{1}{p}}$. An exact characterization of the dimension $d(n)$ needed to isometrically embed all n-point metric spaces into $\ell_\infty^{d(n)}$ is an open question. In [7] Wolf has shown that $d(n) \le n - 2$, for $n \ge 4$. The lower bound was studied in [2], where the author proved that $d(n) \ge n - g(n)$, $g(n) = O(n^{\frac{2}{3}} \log n)$. Recently, in [6] the authors showed that for any integer $c > 0$, $d(n) \le n - c$, for n large enough.

What follows next was not formally covered in the class. We did mention the construction and gave an intuition behind the proofs, so we decided to include this material into the lecture scribes.

1.4 Isometric Embeding of ℓ_1^d into $\ell_\infty^{2^{d-1}}$

Let us consider the following worm up example, for $d = 2$.

Claim 1.11. *ℓ_1^2 isometrically embeds into ℓ_∞^2.*

The "proof" is a fact that the unit ball of the ℓ_1^2 space is a rotation and scaling of the unit ball of the ℓ_∞^2 space. This kind of transformation, (linear and preserving the unit ball up to some constant) preserves the norm of all the vectors. And this in turn, implies that transformation preserves the inter-point distances. Let us formalize this observation.

Lemma 1.12. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces, and let $f : X \to Y$ be a linear transformation. Then f is isometry iff for all $x \in X$, with $\|x\|_X = 1$, it holds that $\|f(x)\|_Y = 1$.*

Note: The same is true for the linear embedding with distortion. Namely, if $f\colon X \to Y$ is a linear mapping, then f has distortion α iff there exists a constant $c > 0$ s.t. $\forall u \in X,\ \ \|x\|_X \leq c \cdot \|f(u)\|_Y \leq \alpha \|x\|_X$.

Next, we extend the the above result.

Theorem 1.13. *ℓ_1^d is isometrically embeddable into $\ell_\infty^{2^{d-1}}$.*

Before we proof this theorem, let us examine the actual isometric embedding between ℓ_1^2 to ℓ_∞^2. Define $f : \ell_1^2 \to \ell_\infty^2$ as follows: $\forall (x, y) \in \mathbb{R}^2 \ \ f(x, y) = (x + y, y - x)$. We show that $\|(x, y)\|_1 = \|f(x, y)\|_\infty$. This is true, since, $\|(x, y)\|_1 = |x| + |y| \geq |x + y|$, and also $\|(x, y)\|_1 = |x| + |y| = |-x| + |y| \geq |y - x|$. On the other hand, we have $\|(x, y)\|_1 \geq \max\{|x + y|, |y - x|\} = \|f(x, y)\|_\infty$, and$\max\{|x + y|, |y - x|\} = \max\{|x + y|, |y - x|, |x - y|, |-x - y|\} \geq |sign(x)x + sign(y)y| = ||x| + |y|| = |x| + |y|$. Now we can extend this idea to a higher dimension.

Proof of Theorem 1.13. Let $S = \{-1, 1\}^d$, be the set of size 2^d, and consider the subset $P \subseteq S$, which is obtained from S by taking off a negation of every string in S. Clearly, the size of P is $2^d/2$ (exactly half of string will survive). Define an embedding as follows:

$$\forall x \in \mathbb{R}^d, \quad f(x) = (< s_1, x >, < s_2, x >, \ldots < s_{2^{d-1}}, x >),$$

for all $s_i \in P$, where $< \cdot >$ is a standard inner product. Note that by the linearity of the inner product, f is a linear. We show that f is a norm preserving. For any $x = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$, we have

$$\|x\|_1 = \sum_{1 \leq i \leq d} |x_i| =^{\forall\, s_j \in S}= \sum_{1 \leq i \leq d} |x_i||s_j^i| \geq \left| \sum_{1 \leq i \leq d} x_i \cdot s_j^i \right|.$$

Therefore, $\|x\|_1 \geq \|f(x)\|_\infty$. On the other hand,

$$\|f(x)\|_\infty = \max_{s_j \in S}\{| < x, s_j > |\} = \max_{s_j \in S \cup \bar{S}}\{| < x, s_j > |\} = \max_{s_j \in S \cup \bar{S}} \left\{ \left| \sum_{1 \leq i \leq d} x_i \cdot s_j^i \right| \right\} \geq,$$

$$\geq \left| \sum_{1 \leq i \leq d} sign(x_i) \cdot x_i \right| = \sum_{1 \leq i \leq d} |x_i| = \|x\|_1.$$

$\square$

Diameter Problem: given a set X of n points in ℓ_1^d. The goal is to find the maximal distance between any two points in X.
Trivial Solution: Compute all the pairwise distances and choose the maximal one. $O(n^2 d)$ time.
Another Solution: Embed all the points into $\ell_\infty^{2^{d-1}}$. The distance between any two embedded points (in terms of ℓ_∞ norm) is the maximal difference between any two appropriate coordinates. Thus, we can compute the diameter of the embedded set as follows. For each coordinate $1 \leq i \leq 2^{d-1}$, compute the maximal difference of n values by finding the minimal and the maximal values. From 2^{d-1}values, compute the maximum. The time complexity is $O(n2^d)$ + the time needed to apply the embedding. To embed one point takes $O(2^d d)$ time. Thus, the the total time to solve the problem is $O(nd2^{d-1})$ which is better than a naive solution, for $d = o(\log n)$.

Note:

We can consider better implementation of the computing the above embedding, the one similar to rolling hash. For each point $x \in \ell_1^d$ we compute the first coordinate of embedding directly, investing $O(d)$ time, and all of the rest coordinates of embedding we can efficiently compute from the previous result, by changing it slightly with $O(1)$ (amortized) operations, thus reducing the time for embedding one point to $O(2^{d-1})$ amortized time.

References

[1] Ittai Abraham, Yair Bartal, and Ofer Neiman. Advances in metric embedding theory. STOC '06, pages 271–286, 2006.

[2] K. Ball. Isometric embedding in ℓ_p-spaces. *European J. Combin.*, 11(4):305–311, 1990.

[3] Yair Bartal. Probabilistic approximations of metric spaces and its algorithmic applications. In *37th Annual Symposium on Foundations of Computer Science, FOCS '96, Burlington, Vermont, USA, 14-16 October, 1996*, pages 184–193, 1996.

[4] Michael Elkin, Arnold Filtser, and Ofer Neiman. Terminal Embeddings. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2015)*, volume 40 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 242–264, 2015.

[5] Jittat Fakcharoenphol, Satish Rao, and Kunal Talwar. A tight bound on approximating arbitrary metrics by tree metrics. *Journal of Computer and System Sciences*, 69(3):485 – 497, 2004. Special Issue on {STOC} 2003.

[6] F. Petrov, D. Stolyarov, and P. Zatitskiy. On embeddings of finite metric spaces in ℓ_∞^n. *ArXiv e-prints*, March 2009.

[7] D. Wolfe. Imbedding a finite metric set in an n-dimensional minkowski space. *Nederl. Akad. Wetensch., Proc. Ser.*, A 70, 1967.