CSC2414 - Metric Embeddings*
Lecture 1: A brief introduction to metric embeddings, examples and motivation

Notes taken by Costis Georgiou
Revised by Hamed Hatami

Summary: We are interested in representations (embeddings) of one metric space into another metric space that preserve or approximately preserve the distances. In this lecture we provide the basic definitions as well as many examples and a brief overview of the history and main theorems in the area.

1 Introduction

In recent years, the study of distance-preserving embeddings has introduced a powerful tool to algorithm designers due to the connection between combinatorics and geometry. The "embedding method" is considered to be one of the important methods in the design of approximation algorithms, and has gained a lot of popularity in the computer scientists community in the past decade. Its application in the algorithm design is usually in the following framework: One often takes a problem defined over a "difficult" metric and reduce it to a problem over an "easier" metric. Since the solution of many problems is strongly connected to the geometric properties of their input, embeddings are advisable for solving problems over metric spaces.

A finite metric space is simply a set of points with distances between them that satisfy triangle inequality and that two distinct points have nonzero distance. A Euclidean metric, on the other hand, is a metric that is obtained when points are placed in some Euclidean space and distances are inherited from the Euclidean norm. We will discuss such families of metric spaces and their differences, as well as develop algorithms to "embed" one metric space into the other. We will introduce techniques for obtaining negative results saying that some spaces cannot be well embedded into certain normed spaces and discuss the issue of dimension of normed spaces, i.e. when can we reduce dimension and at what cost. We will also draw a link between combinatorial properties of graphs (e.g. expansion, girth, planarity) to the quality in which their metric embeds into certain normed spaces.

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

Much of the discussion ties these questions to the art of algorithmic design. One aspect of this connection is the following: many problems deal explicitly or implicitly with a distance measure between items. When these distances come from a simple geometrical space, say low dimensional Euclidean space, an efficient algorithm is typically available or easy to find. For a general class of metric spaces it is therefore useful to embed the items in such geometrical spaces so that the new distances approximate the original ones. Then, applying the simple algorithm in the geometrical space gives an approximation algorithm for the original problem. We will also discuss less obvious connections to approximation algorithms. For example, we will link certain LP-relaxations or Semi-Definite relaxations to the quality of embeddings of particular metric spaces.

2 Metric spaces and embeddings

Many practical problems that relate metric spaces arise from many different disciplines. To illustrate the need for metric embeddings, it is good to start with such a problem from bioinformatics.

Example 2.1. Biological data, such as DNA or proteins are usually represented as sequences of elements taken from an alphabet. After many years of research, more than half a million different proteins have been discovered with known sequences. Finding the similarity of two different proteins is a fundamental concern and is related to a notion of distance. Thus a set of biological data can been thought as a finite metric space.

Assume that we are given a set of biological $0-1$ sequences X,

0110011000110011101
0100110010010011000
0110111011010001001
1001100011100011110
0100110001110001111

We can observe that the last two sequences seem more similar. We would ask the question that "is there a function f that maps X to a plane with Euclidean distance, keeping the distances the same?"

A mapping of a metric space to a Euclidean space might be useful but not always possible to achieve, as this is the case for the following example:

Example 2.2. Consider the metric space X defined in Figure 1.

Our claim is that it is not possible to realize X in Euclidean space of any dimension. We Prove the claim by contradiction: assume that there is an integer k, such that $f : X = \{a, b, c, d\} \to \mathbb{R}^k$ and moreover f preserves the distances. Since the $\triangle$-inequality is tight for the elements d, a, b we conclude that $f(d), f(a), f(b)$ are collinear in the space $\mathbb{R}^k$. Using the same argument $f(d), f(a), f(c)$ are collinear too. But then $\|f(b) - f(c)\|_2 = 0$ contradicting the fact that the distance between b and c is 2.

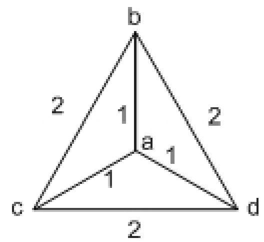

Figure 1: The metric space X

A nice property is the ability to embed sequences into the Euclidian space so that distances are preserved. Example 2.2 indicates that we can only hope to approximate this property. To discuss the relation between metric spaces, first we have to formally define the notion of a metric space.

Definition 2.3. A *metric space* is a pair (X, ρ), where X is a set of *points* and a function $\rho : X \times X \to \mathbb{R}^{\geq 0}$ so that

1. $\rho(x, y) = 0 \Leftrightarrow x = y$
2. $\rho(x, y) = 0 = \rho(y, x)$ (symmetry)
3. $\rho(x, y) + \rho(y, z) \geq \rho(x, z)$ ($\triangle$-inequality)

A mapping ρ that satisfies (1) and (2) is called a *distance*, while a distance that satisfies property (3) is called a *metric*. Moreover, if we allow $\rho(x, y) = 0$ then the metric is called *semi-metric* or *pseudometric*. Let (X, ρ) and (Y, μ) be two metric spaces. Any one-to-one map $f : X \to Y$ is called an *embedding*.

It is easy to see that alternatively, having the elements of X ordered, we can think of a symmetric $|X| \times |X|$ matrix to describe the metric space.

On the other hand a normed space is defined as in the following:

Definition 2.4. A normed space is a pair $(V, \|\cdot\|)$, where V is a vector space over $\mathbb{R}$ or $\mathbb{C}$ and $\|\cdot\|$ is a function from V to $\mathbb{R}^{\geq 0}$ satisfying

1. $\|x\| = 0$ iff $x = 0$;
2. $\|\lambda x\| = |\lambda|\|x\|$ for all $x \in V$ and scalar λ;
3. $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in V$.

It is easy to see that every normed space is a metric space with the distance function defined as $d(x, y) = \|x - y\|$, for every $x, y \in V$.

Example 2.5. Some examples of the finite metric spaces are stated in the following.

1. Every graph induces a metric on its vertices that is characterized by the shortest path between vertices. Consider a path with 5 nodes and edges of length 1. A matrix that describes this metric space is

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 0 & 1 & 2 & 3 \\ 2 & 1 & 0 & 1 & 2 \\ 3 & 2 & 1 & 0 & 1 \\ 4 & 3 & 2 & 1 & 0 \end{pmatrix}$$

2. We can represent a cycle with 5 nodes and edges of length 1 with the matrix

$$\begin{pmatrix} 0 & 1 & 2 & 2 & 1 \\ 1 & 0 & 1 & 2 & 2 \\ 2 & 1 & 0 & 1 & 2 \\ 2 & 2 & 1 & 0 & 1 \\ 1 & 2 & 2 & 1 & 0 \end{pmatrix}$$

3. Given a set X and its subset S, we can think of the partition $S, \overline{S}$ as a semi-metric space in the following way. Elements of different sets have distance 1 while elements of the same set have distance 0. This is known as a cut metric (see Figure 2).

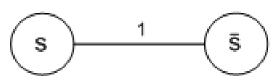

Figure 2: A cut metric

If we denote by $\mathbf{0}_{n\times m}$, the $n \times m$ zero matrix and with $J_{n\times m}$ the all one $n \times m$ matrix, then the previous semi-metric can be represented as a matrix in the following way

$$\begin{pmatrix} \mathbf{0}_{|S|\times|S|} & J_{|S|\times|\overline{S}|} \\ J_{|\overline{S}|\times|S|} & \mathbf{0}_{|\overline{S}|\times|\overline{S}|} \end{pmatrix}.$$

4. In the Euclidean space consider a 1×1 square. Clearly this metric space can be described as

$$\begin{pmatrix} 0 & 1 & \sqrt{2} & 1 \\ 1 & 0 & 1 & \sqrt{2} \\ \sqrt{2} & 1 & 0 & 1 \\ 1 & \sqrt{2} & 1 & 0 \end{pmatrix}.$$

Having defined an input or output space as a metric space, we would hope that analyzing the corresponding metrics will provide us useful information. Therefore it is desirable to look for simple metrics such as:

- A metric which has a short description;
- A low dimensional normed space;
- A metric that comes from "simple" graphs such as a cycle, a tree, a sparse or planar graph.

The infinite spaces $\mathbb{R}^k$ equipped with the so called Minkowski norm $\|\cdot\|_p$ (for $p \in [1, \infty]$) gives rise to the most commonly used metric spaces. Recall that the norm ℓ_p on $\mathbb{R}^k$ is defined as

$$\|x\|_p = \left(\sum_{i=1}^{k} |x_i|^p \right)^{\frac{1}{p}},$$

where $1 \leq p < \infty$, while $\|x\|_\infty$ is defined as $\max_i |x_i|$.

Definition 2.6. We say that a finite metric space (X, ρ) is *realized* in ℓ_p^k if there is a function $f : X \to \mathbb{R}^k$ so that $\rho(x, y) = \|f(x) - f(y)\|_p$. We also call a finite metric space (X, ρ) an ℓ_p*-metric* if it can be realized in ℓ_p^k for some k.

The following simple theorem shows that it is possible to realize every metric space in ℓ_∞.

Theorem 2.7. *Every metric space embeds isometrically into ℓ_∞.*

Proof. We will prove this lemma only for finite metric spaces. Consider a metric space (X, d), where $X = (x_1, \ldots, x_n)$. It suffices to find a function $f : X \to \mathbb{R}_n$ such that (X, d) embeds isometrically into $(\mathbb{R}_n, \|\cdot\|)$. For $x_i \in X$ we define

$$f(x_i) = (d(x_1, x_i), d(x_2, x_i), \ldots, d(x_n, x_i))$$

Clearly it suffices to show for every $x_i, x_j \in X$ that $\|f(x_i) - f(x_j)\|_\infty = d(x_i, x_j)$. First we note that since d is a metric, it respects the $\triangle$-inequality, thus $d(x_i, x_k) - d(x_j, x_k) \leq d(x_i, x_j)$ for $k = 1, \ldots n$. It follows that

$$\max_k |d(x_i, x_k) - d(x_j, x_k)| \leq d(x_i, x_j),$$

or in other words

$$\|f(x_i) - f(x_j)\|_\infty \leq d(x_i, x_j). \tag{1}$$

On the other hand, the j-th coordinate of the vector $f(x_i) - f(x_j)$ is $d(x_j, x_i) - d(x_j, x_j) = d(x_i, x_j)$. Therefore the maximum coordinate of $f(x_i) - f(x_j)$ is at least $d(x_i, x_j)$ or in other words

$$\|f(x_i) - f(x_j)\|_\infty \geq d(x_i, x_j). \tag{2}$$

The lemma follows then from (1) and (2). □

3 Similarity of metric spaces

Solving the TSP in a graph is a difficult problem. An embedding of the metric of the graph into a tree that preserves the distances makes the problem trivial. However, as we saw in Example 2.2, we cannot always hope to achieve such embeddings. This gives rise to the following definition that tries to capture the idea of approximately preserving the distances of a metric space.

Definition 3.1. Let f be an embedding from the finite metric space (X, ρ) into another finite metric (Y, μ). We define

$$\text{expansion}(f) = \max_{x,y \in X} \frac{\mu\left(f(x), f(y)\right)}{\rho(x,y)}$$

$$\text{contraction}(f) = \max_{x,y \in X} \frac{\rho(x,y)}{\mu\left(f(x), f(y)\right)}$$

The *distortion* of an embedding f, distortion(f), is defined as the product of expansion(f) and contraction(f). An embedding f with $distortion(f) = 1$ is called *isometric*.

Clearly the best the distortion that we can hope for is 1 (this is achieved when the new distances are just a scaling of the old distances with some number $\beta > 0$. Note that the distortion of an embedding f can also be equivalently defined as the minimum $\alpha \geq 1$ such that there exists $\beta > 0$ such that for every $x, y \in X$:

$$\beta\mu\left(f(x), f(y)\right) \leq \rho(x,y) \leq \alpha\beta\mu\left(f(x), f(y)\right).$$

4 Algorithmic application

Consider a problem that is defined on a set of points in a metric space M. Our aim will be to efficiently "place" the points in a simpler space M' so that

1. The distortion is small;
2. It is easy to solve the problem in M'.

Some possible scenarios include the following cases

1. Embed a metric space into a low dimensional ℓ_p space.
2. Start with high dimensional normed space and embed it into a low dimensional normed space.
3. Embed a metric space into "tree-metrics" and usually solve the problem using a divide $\&$ conquer approach on the tree.

While embedding metric spaces with small distortion into tree-metrics might be useful, this is not always possible. For instance, embedding a metric space into one tree is problematic in the sense that any such embedding could be very poor. For example it is not possible to embed an n-cycle, C_n, into a tree-metric with good distortion.

Example 4.1. The n-cycles cannot be embedded into a tree metric with distortion better than $\Omega(n)$. For instance, the attempt to discard any of the edges creates expansion $n-1$. Furthermore, the contraction is 1 since by discarding an edge the distance between any two nodes cannot be decreased. Therefore, $n-1$ is a lower bound for the distortion of this specific embedding.

On the other hand we can embed C_n into a distribution of trees. Those trees are once again paths that arise by discarding an edge chosen uniformly at random. Below, the expected value is computed with respect to the random choice of the edge e that we discard, or in other words the random tree T_e that is obtained. We denote by f_e the embedding that is defined by discarding the edge e and by $d_e(\cdot,\cdot)$ the cost of the minimum path of any two nodes in T_e. Thus the embedding f_e is defined as the mapping

$$f_e : (C_n, d) \to (T_e, d_e)$$

Now, clearly

$$\begin{aligned}
\mathbf{E}_{T_e}\left[d_e\left(f_e(x), f_e(y)\right)\right] &= \frac{d(x,y)}{n}\left(n - d(x,y)\right) + \frac{n-d(x,y)}{n} d(x,y) \\
&= 2d(x,y)\frac{n-d(x,y)}{n} \\
&\leq 2d(x,y)
\end{aligned}$$

improving the distortion to a constant factor.

More algorithmic extensions of metric embeddings arise from linear and semidefinite programming.

Example 4.2. Problems such as MAX-CUT, SPARSEST-CUT or even VERTEX COVER can be formulated as linear or semidefinite programs (LP/SDP). Their relaxation can be thought as metric spaces. The partition of the nodes that any of these examples requires can be thought as a cut metric, while more rich metric spaces follow by relaxing the LP's and SDP's. An embedding from the rich space to a cut metric space gives rise to an approximation algorithm, where distortion translates to the integrality gap of the relaxation.

5 Historical notes

In this section we review the history of this area and state some of the most structural theorems that we will discuss in the course.

Although the study of the dissimilarity of normed spaces goes back to the first half of the 20th century and works of mathematicians such as Stefan Banach, Stanislaw Mazur, Fritz John[Joh48], Isaac Schoenberg, Dvoretzky [Dvo61], etc, the minimum distortion required to embed a metric space into a normed space remained unstudied until 1985 when Bourgain [Bou85] in a short paper proved an upper-bound for this quantity in a very general case.

Theorem 5.1 (Bourgain). *Every metric space on n points embeds into $\ell_p^{O(\log n)}$ with distortion and dimension $O(\log^2 n)$.*

Meanwhile due to a striking result of the Russian mathematician E. D. Gluskin [Glu81], many great mathematicians including G. Pisier [Pis89], M. Talagrand [LT91], G. Schechtman and V. Millman [MS86], etc, started to restudy the dissimilarity of normed spaces. There is a tight connection between this area and the area of metric embeddings, so that some theorems regarding normed spaces has immediate corollaries or analogues in metric spaces. For example the following theorem of W.B. Johnson and J. Lindenstrauss [JL84] is originally proved as a problem in this area.

Theorem 5.2. *Given n points in ℓ_2^n, they can be embedded into $\ell_2^{O(\log n/\epsilon^2)}$ with distortion $1+\epsilon$.*

Although the research of Bourgain on metric spaces carried further by himself [Bou86, BFM86], J.Matousek [Mat92], M. Deza and H. Maehara [DM90], etc, it was not until 1995 when the seminal paper of N. Linial, E. London, and Y. Rabinovich [LLR95] is published that the algorithmic importance of Bourgain's theorem and other problems in this area is revealed.

References

[BFM86] J. Bourgain, T. Figiel, and V. Milman. On Hilbertian subsets of finite metric spaces. *Israel J. Math.*, 55(2):147–152, 1986.

[Bou85] J. Bourgain. On Lipschitz embedding of finite metric spaces in Hilbert space. *Israel J. Math.*, 52(1-2):46–52, 1985.

[Bou86] J. Bourgain. The metrical interpretation of superreflexivity in Banach spaces. *Israel J. Math.*, 56(2):222–230, 1986.

[DM90] M. Deza and H. Maehara. Metric transforms and Euclidean embeddings. *Trans. Amer. Math. Soc.*, 317(2):661–671, 1990.

[Dvo61] Aryeh Dvoretzky. Some results on convex bodies and Banach spaces. In *Proc. Internat. Sympos. Linear Spaces (Jerusalem, 1960)*, pages 123–160. Jerusalem Academic Press, Jerusalem, 1961.

[Glu81] E. D. Gluskin. The diameter of the Minkowski compactum is roughly equal to n. *Funktsional. Anal. i Prilozhen.*, 15(1):72–73, 1981.

[JL84] William B. Johnson and Joram Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space. In *Conference in modern analysis and probability (New Haven, Conn., 1982)*, volume 26 of *Contemp. Math.*, pages 189–206. Amer. Math. Soc., Providence, RI, 1984.

[Joh48] Fritz John. Extremum problems with inequalities as subsidiary conditions. In *Studies and Essays Presented to R. Courant on his 60th Birthday, January 8, 1948*, pages 187–204. Interscience Publishers, Inc., New York, N. Y., 1948.

[LLR95] Nathan Linial, Eran London, and Yuri Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–245, 1995.

[LT91] Michel Ledoux and Michel Talagrand. *Probability in Banach spaces*, volume 23 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1991. Isoperimetry and processes.

[Mat92] JiříMatoušek. Note on bi-Lipschitz embeddings into normed spaces. *Comment. Math. Univ. Carolin.*, 33(1):51–55, 1992.

[MS86] Vitali D. Milman and Gideon Schechtman. *Asymptotic theory of finite-dimensional normed spaces*, volume 1200 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1986. With an appendix by M. Gromov.

[Pis89] Gilles Pisier. *The volume of convex bodies and Banach space geometry*, volume 94 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1989.