Note on bi-Lipschitz embeddings into normed spaces

Jiří Matoušek

Abstract. Let (X,d), (Y,ρ) be metric spaces and $f : X \to Y$ an injective mapping. We put $\|f\|_{\mathrm{Lip}} = \sup\{\rho(f(x), f(y))/d(x,y);\ x,y \in X, x \neq y\}$, and $\mathrm{dist}(f) = \|f\|_{\mathrm{Lip}} . \|f^{-1}\|_{\mathrm{Lip}}$ (the *distortion* of the mapping f). We investigate the minimum dimension N such that every n-point metric space can be embedded into the space ℓ_∞^N with a prescribed distortion D. We obtain that this is possible for $N \geq C(\log n)^2 n^{3/D}$, where C is a suitable absolute constant. This improves a result of Johnson, Lindenstrauss and Schechtman [JLS87] (with a simpler proof). Related results for embeddability into ℓ_p^N are obtained by a similar method.

Keywords: finite metric space, embedding of metric spaces, distortion, Lipschitz mapping, spaces ℓ_p

Classification: 46B99, 54C25

Let us begin with some notation. The symbol ℓ_p^n denotes the n-dimensional real vector space equipped with the L_p-norm, given by $\|(x_1, \dots, x_n)\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$ (for $1 \leq p < \infty$; for $p = \infty$ it is $\|(x_1, \dots, x_n)\|_\infty = \max\{|x_i|;\ i = 1, \dots, n\}$). Similarly ℓ_p denotes the space of countable sequences of real numbers with the L_p-norm. If P is a finite set equipped by a measure, we will sometimes use the notation $L_p(P)$, meaning the space of $|P|$-tuples of real numbers indexed by members of P, equipped by the L_p-norm (thus ℓ_p^n is just $L_p(\{1, \dots, n\})$, where the set $\{1, \dots, n\}$ is considered with the counting measure).

Every n-point metric space can be isometrically embedded into ℓ_∞^n (this is an old observation due to Fréchet): If $X = \{x_1, \dots, x_n\}$, the embedding $f : X \to \ell_\infty^n$ is defined by $f(x_i)_j = \rho(x_i, x_j)$.

For other ℓ_p spaces, there exist finite metric spaces which cannot be embedded isometrically (a classical work on isometric embeddability into Hilbert space is [Scho38]). One can quantitatively measure the degree of "metric non-embeddability" using so-called Lipschitz distance of metric spaces.

Let (X,d), (Y,ρ) be metric spaces. We let

$$\mathrm{dist}(X, \subseteq Y) = \inf\{\mathrm{dist}(f);\ f : X \to Y \text{ an injective mapping}\}$$

(the distortion of a mapping was defined in the abstract). When $|X| = |Y|$ and the infimum is taken over all *bijective* mappings, this quantity is called the *Lipschitz distance* of X and Y in the literature.

This research was performed while the author was at Department of Computer Science, Charles University

A problem studied in the recent literature is the minimum distortion necessary for embedding of general finite metric spaces into normed spaces (in particular, into ℓ_p^n) and also the minimum dimension, needed for an embedding with a prescribed distortion.

For embedding into Hilbert space, the situation has been essentially cleared out by the works [JL84] and [Bou85]. J. Bourgain proved the following:

Theorem 1 [Bou85]**.**

(i) *For every n-point metric space X,* $\mathrm{dist}(X, \subseteq \ell_2) = O(\log n)$.

(ii) *For every n, there exists a metric space X with* $\mathrm{dist}(X, \subseteq \ell_2) \geq c \log n / \log\log n$, *where $c > 0$ is an absolute constant.*

This gives nearly tight bounds for the embeddability (without a limit on the dimension of the image space). Since every finite subspace of ℓ_2 is isometrically embeddable into any other ℓ_p, the upper bound (i) holds for all p. The lower bound proof in (ii) can be re-formulated using graphs without short cycles, and the same lower bound can be extended to all $p \in [1, 2]$ ([Ma89]). A good lower bound for $p > 2$ remains an open problem; the best known bound follows from [BMW86] and it is $(c \log n)^{1/p}$ for an absolute constant $c > 0$.

Another interesting question is what happens when we limit the dimension of the normed space into which we want to embed. For Euclidean spaces, the following "flattening lemma" was established by Johnson and Lindenstrauss:

Theorem 2 [JL84]**.** *For every $\varepsilon > 0$ there exists a constant $C = C(\varepsilon)$, such that if X is an n-point subset of ℓ_2^n for some $n \geq 2$, then* $\mathrm{dist}(X, \subseteq \ell_2^{C \log n}) \leq 1 + \varepsilon$.

If some analogue of this lemma holds for other values of p is another interesting open problem.

Johnson, Lindenstrauss and Schechtman proved the following result:

Theorem 3 [JLS87]**.** *For every n point metric space X and a number D, there exists a k-dimensional normed space Z with* $\mathrm{dist}(X, \subseteq Z) \leq D$, *where* $k = O((\log n)^3 D^2 n^{K/D})$, *for some absolute constant K.*

They combine the technique of [Bou85] with some other methods from normed space theory. We will show a strengthening of their result (namely the embedding will always be into ℓ_∞^k), using a much simpler method. A similar method yields also some estimates for embedding into ℓ_p^k.

Theorem 4. *Let X be an n-point metric space.*

(i) *If $N \geq C(\log n)^2 n^{3/D}$, where C is a suitable absolute constant, then* $\mathrm{dist}(X, \subseteq \ell_\infty^N) \leq D$.

(ii) *For every p, $1 \leq p < \ln n/3$,* $\mathrm{dist}(X, \subseteq \ell_p^N) = O((\log n)^{1+1/p}/p)$, *provided that $N \geq C^p (\log n)^2$, where C is a suitable absolute constant.*

(iii) *Let $p \in [1, \infty]$ and $N \geq C(\log n)^2$, where C is a suitable absolute constant. Then* $\mathrm{dist}(X, \subseteq \ell_p^N) = O(\log n)$ *(the constant of proportionality can be chosen independently of p).*

Let us remark that (iii) without a bound on the dimension follows immediately from [Bou85]. The part (ii) shows that the necessary distortion really decreases with growing p, and for p of the order $\log n$ we get an embedding with distortion bounded by a constant.

The technique we will use for embedding of finite metric spaces into normed spaces is due to J. Bourgain ([Bou85]). Let (X, ρ) be an n-point metric space, $m = \lceil \log_2 n \rceil + 1$ and let M_k denote the set of all subsets of X of size 2^k, $k = 0, 1, \ldots, m-1$. Let us put $M = M_0 \cup \ldots \cup M_{m-1}$. On every M_k, we introduce a probabilistic measure μ_k, which assigns the same probability to every element of the set M_k, and a probabilistic measure μ on M is defined by $\mu(\{A\}) = \mu_k(\{A\})/m$ for every $A \in M_k$.

Let $x, y \in X$ be two points and let $A \in M$; we denote $d_A(x, y) = |\rho(A, x) - \rho(A, y)|$. Obviously $d_A(x, y) \le \rho(x, y)$. The following lemma contains two versions of the same idea and its proof is not too difficult:

Lemma 5. *Let x, y be two points of a metric space X.*

(i) [JLS87] *For every $\alpha \in (0, 1/3)$ there exists k, such that*

$$\mu_k(\{A \in M_k;\ d_A(x, y) \ge \alpha\rho(x, y)\}) \ge cn^{-3\alpha},$$

where c is a positive constant.

(ii) [Bou85] *There exist nonnegative numbers $\rho_0, \ldots, \rho_{m-1}$ and pairwise distinct indices $k_0, \ldots, k_{m-1}$, such that $\rho_0 + \rho_1 + \cdots + \rho_{m-1} \ge \rho(x, y)/3$ and*

$$\mu_{k_i}(\{A \in M_{k_i};\ d_A(x, y) \ge \rho_i\}) \ge c\,,$$

where c is a positive constant.

Proof of Theorem 4: (i) Let a set P_k $(k = 0, 1, ..., m-1)$ arise by r independent random draws from the set M_k, where $r = N/m$. Let us put $P = P_0 \cup \ldots \cup P_{m-1}$ (so $|P| \le N$). An embedding $f : X \to L_\infty(P)$ is defined by $f(x)_A = \rho(x, A)$ for $A \in P$. Clearly $\|f\|_{\mathrm{Lip}} \le 1$.

Let x, y be a pair of distinct points of X, and let $\alpha = 1/D$. Let k be an index as in Lemma 5 (i). Then

$$\begin{aligned}\mathrm{Prob}(\forall A \in P_k;\ d_A(x, y) < \alpha\rho(x, y)) &= \mu_k(\{A \in M_k;\ d_A(x, y) < \alpha\rho(x, y)\})^r \le \\ &\le (1 - cn^{-3/D})^{N/m} \le \exp(-cn^{-3/D}N/m) < \exp(-c.C.\log n) < n^{-2}\,,\end{aligned}$$

hence

$$\begin{aligned}\mathrm{Prob}(\|f^{-1}\|_{\mathrm{Lip}} > D) &\le \mathrm{Prob}(\exists\, x, y \in X;\ \forall\, A \in P;\ d_A(x, y) < \alpha\rho(x, y)) \le \\ &\le \binom{n}{2} \mathrm{Prob}(\forall\, A \in P_k;\ d_A(x, y) < \alpha\rho(x, y)) < 1\,.\end{aligned}$$

This means that there exists some embedding $f : X \to \ell_\infty^N$ with distortion at most D.

(ii) Similarly as in (i), we select P_k from M_k using $r = N/m$ independent random draws. We define $f : X \to L_p(P)$ (where we take the uniform probability measure on $P = P_0 \cup \ldots \cup P_{m-1}$) by $f(x)_A = \rho(x, A)$. Similarly as in the previous, $\|f\|_{\mathrm{Lip}} \le 1$.

This time we will bound the probability that the difference $|f(x)_A - f(y)_A|$ (for given x, y) is large only for a small fraction of A's from P_k. Let us put $\alpha = p/\log n$. Let $x, y \in X$ and k be index as in Lemma 5 (i). Let τ denote the probability that for a random $A \in M_k$ it is $d_A(x, y) \ge \alpha\rho(x, y)$; we have $\tau \ge cn^{-3\alpha} \ge C_1^{-p}$ for some absolute constant C_1.

We bound the probability $\theta = \mathrm{Prob}(|\{A \in P_k;\ d_A(x, y) \ge \alpha\rho(x, y)\}| < \tau r/2)$. This probability is bounded by the probability that we achieve less than $\tau r/2$ successes in a series of r independent (Bernoulli) trials with success probability τ. By Chernoff inequality (see e.g. [Spe]) we get

$$\theta \le \exp(-\frac{\tau r}{8}) = \exp(-(C/C_1)^p \log n/8) < n^{-2},$$

provided that C is large enough compared to C_1.

Hence for a certain choice of the set P we may assume that for every pair $x, y \in X$ there exists k such that $|\{A \in P_k;\ d_A(x, y) \ge \alpha\rho(x, y)\}| \ge \tau r/2$. Let f be a mapping defined as above for such a set P. Then for every x, y we have

$$\|f(x) - f(y)\|_p = \left(\sum_{A\in P} \frac{d_A(x, y)^p}{|P|}\right)^{1/p} \ge \left(\sum_{A\in P_k} \frac{d_A(x, y)^p}{N}\right)^{1/p} \ge$$
$$\ge \left(\sum_{A\in P_k;\ d_A(x,y)\ge\alpha\rho(x,y)} \frac{\alpha^p\rho(x, y)^p}{N}\right)^{1/p} \ge \left(\frac{\tau r\alpha^p\rho(x, y)^p}{2N}\right)^{1/p} \ge$$
$$\ge (\tau/2m)^{1/p}\alpha\rho(x, y) \ge \left(\frac{C_1^{-p}}{2\log n}\right)^{1/p} \frac{p}{\log n}\rho(x, y) \ge \frac{\rho(x, y)}{O((\log n)^{1+1/p}/p)}\,.$$

(iii) The proof is quite analogous to (ii), only we use Lemma 5 (ii) instead of (i). Again we put $r = N/m$, and the sets $P_0, \ldots, P_{m-1}$ will be as in the previous. Let for a given pair x, y the numbers $\rho_0, \ldots, \rho_{m-1}$ and indices $k_0, \ldots, k_{m-1}$ be as in Lemma 5 (ii). One proves that for every $i = 0, 1, \ldots, m-1$ it is ($c > 0$ is the constant from Lemma 5 (ii))

$$\mathrm{Prob}(|\{A \in P_{k_i};\ d_A(x, y) \ge \rho_i\}| < cr/2) < n^{-2}m^{-1},$$

so there exists a set P such that for the corresponding mapping $f : X \to \ell_p(P)$ we have (for every $x, y \in X$)

$$\|f(x) - f(y)\|_1 \ge \frac{1}{N} \sum_{i=0}^{m-1} \frac{cr\rho_i}{2} \ge \frac{1}{m} \cdot \frac{c}{2} \cdot \frac{\rho(x, y)}{3} = \frac{\rho(x, y)}{O(\log n)}\,,$$

and finally it is $\|f(x) - f(y)\|_1 \le \|f(x) - f(y)\|_p \le \|f(x) - f(y)\|_\infty \le \rho(x, y)$, hence $\|f\|_{\mathrm{Lip}} = O(\log n)$ — we even use the same mapping for each p.

References

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

[BMW86] Bourgain J., Milman V., Wolfson H., *On type of metric spaces*, Trans. Amer. Math. Soc. **294** (1986), 295–317.

[JL84] Johnson W., Lindenstrauss J., *Extensions of Lipschitz maps into a Hilbert space*, Contemporary Math. **26** (Conference in modern analysis and probability), 189–206, Amer. Math. Soc., 1984.

[JLS87] Johnson W., Lindenstrauss J., Schechtman G., *On Lipschitz embedding of finite metric spaces in low dimensional normed spaces*, in: *Geometrical aspects of functional analysis* (J. Lindenstrauss, V.D. Milman eds.), Lecture Notes in Mathematics **1267**, Springer-Verlag, 1987.

[Ma89] Matoušek J., *Lipschitz distance of metric spaces* (in Czech), CSc. degree thesis, Charles University, 1990.

[Scho38] Schoenberg I.J., *Metric spaces and positive definite functions*, Trans. Amer. Math. Soc. **44** (1938), 522–536.

[Spe87] Spencer J., *Ten Lectures on the Probabilistic Method*, CBMS-NSF, SIAM 1987.

Department of Applied Mathematics, Charles University, Malostranské nám. 25, 118 00 Praha 1, Czechoslovakia

(*Received* July 29, 1991)