Contemporary Mathematics
Volume 26, 1984

EXTENSIONS OF LIPSCHITZ MAPPINGS INTO A HILBERT SPACE

William B. Johnson[1] and Joram Lindenstrauss[2]

INTRODUCTION

In this note we consider the following extension problem for Lipschitz functions: Given a metric space X and $n = 2, 3, 4, \ldots$, estimate the smallest constant $L = L(X, n)$ so that every mapping f from every n-element subset of X into ℓ_2 extends to a mapping $\tilde{f}$ from X into ℓ_2 with

$$\|\tilde{f}\|_{\ell ip} \leq L \|f\|_{\ell ip} .$$

(Here $\|g\|_{\ell ip}$ is the Lipschitz constant of the function g.) A classical result of Kirszbraun's [14, p. 48] states that $L(\ell_2, n) = 1$ for all n, but it is easy to see that $L(X, n) \to \infty$ as $n \to \infty$ for many metric spaces X.

Marcus and Pisier [10] initiated the study of $L(X, n)$ for $X = L_p$. (For brevity, we will use hereafter the notation $L(p, n)$ for $L(L_p(0,1), n)$.) They prove that for each $1 < p < 2$ there is a constant $C(p)$ so that for $n = 2, 3, 4, \ldots$,

$$L(p, n) \leq C(p) \; (\mathrm{Log}\; n)^{1/p - 1/2} .$$

The main result of this note is a verification of their conjecture that for some constant C and all $n = 2, 3, 4, \ldots$,

$$L(X, n) \leq C(\mathrm{Log}\; n)^{1/2}$$

for all metric spaces X. While our proof is completely different from that of Marcus and Pisier, there is a common theme: Probabilistic techniques developed for linear theory are combined with Kirszbraun's theorem to yield extension theorems.

The main tool for proving Theorem 1 is a simply stated elementary geometric lemma, which we now describe: Given n points in Euclidean space, what

[1]Supported in part by NSF MCS-7903042.

[2]Supported in part by NSF MCS-8102714.

© 1984 American Mathematical Society
0271-4132/84 $1.00 + $.25 per page

is the smallest $k = k(n)$ so that these points can be moved into k-dimensional Euclidean space via a transformation which expands or contracts all pairwise distances by a factor of at most $1 + \varepsilon$? The answer, that $k \leq C(\varepsilon)$ Log n, is a simple consequence of the isoperimetric inequality for the n-sphere in the form studied in [2].

It seems likely that the Marcus-Pisier result and Theorem 1 give the right order of growth for $L(p, n)$. While we cannot verify this, in Theorem 3 we get the estimate

$$L(p, n) \geq \delta \left(\frac{\text{Log } n}{\text{Log Log } n}\right)^{1/p - 1/2} \quad (1 \leq p < 2)$$

for some absolute constant $\delta > 0$. (Throughout this paper we use the convention that Log x denotes the maximum of 1 and the natural logarithm of x.) This of course gives a lower estimate of

$$\delta \left(\frac{\text{Log } n}{\text{Log Log } n}\right)^{1/2}$$

for $L(\infty, n)$. That our approach cannot give a lower bound of $\delta(\text{Log } n)^{1/p - 1/2}$ for $L(p, n)$ is shown by Theorem 2, which is an extension theorem for mappings into ℓ_2 whose domains are ε-separated.

The minimal notation we use is introduced as needed. Here we note only that $B_Y(y, \varepsilon)$ (respectively, $b_Y(y, \varepsilon)$) is the closed (respectively, open) ball in Y about y of radius ε. If $y = 0$, we use $B_Y(\varepsilon)$ and $b_Y(\varepsilon)$, and we drop the subscript Y when there is no ambiguity. S(Y) is the unit sphere of the normed space Y. For isomorphic normed spaces X and Y, we let

$$d(X,Y) = \inf \|T\| \, \|T^{-1}\|,$$

where the inf is over all invertible linear operators from X onto Y. Given a bounded Banach space valued function f on a set K, we set

$$\|f\|_\infty = \sup_{x \in K} \|f(x)\|.$$

1. THE EXTENSION THEOREMS

We begin with the geometrical lemma mentioned in the introduction.

LEMMA 1. *For each* $1 > \tau > 0$ *there is a constant* $K = K(\tau) > 0$ *so that if* $A \subset \ell_2^n$, $\overline{\overline{A}} = n$ *for some* $n = 2, 3, \ldots$, *then there is a mapping* f *from* A *onto a subset of* ℓ_2^k $(k \equiv [K \log n])$ *which satisfies*

$$\|\tilde{f}\|_{\ell ip}\ \|\tilde{f}^{-1}\|_{\ell ip} \le \frac{1+\tau}{1-\tau} .$$

PROOF. The proof will show that if one chooses at random a rank k orthogonal projection on ℓ_2^n, then, with positive probability (which can be made arbitrarily close to one by adjusting k), the projection restricted to A will satisfy the condition on $\tilde{f}$. To make this precise, we let Q be the projection onto the first k coordinates of ℓ_2^n and let σ be normalized Haar measure on $O(n)$, the orthogonal group on ℓ_2^n. Then the random variable

$$f : (O(n), \sigma) \to L(\ell_2^n)$$

defined by

$$f(u) = U^* QU$$

determines the notion of "random rank k projection." The applications of Levy's inequality in the first few self-contained pages of [2] make it easy to check that $f(u)$ has the desired property. For the convenience of the reader, we follow the notation of [2].

Let $|||\cdot|||$ denote the usual Euclidean norm on $\mathbb{R}^n$ and for $1 \le k \le n$ and $x \in \mathbb{R}^n$ set

$$r(x) = r_k(x) = \sqrt{n}\left(\sum_{i=1}^{k} x(i)^2\right)^{1/2} ,$$

which is equal to

$$\sqrt{n}\ |||Qx|||$$

for our eventual choice of $k = [K \log n]$. Thus $r(\cdot)$ is a semi-norm on ℓ_2^n which satisfies

$$r(x) \le \sqrt{n}\ |||x||| \qquad (x \in \ell_2^n).$$

(In [2], $r(\cdot)$ is assumed to be a norm, but inasmuch as the left estimate $a|||x||| \le r(x)$ in formula (2.5) of [2] is not needed in the present situation, it is okay that $r(\cdot)$ is only a semi-norm.)

Setting

$$B = \left\{\frac{x-y}{|||x-y|||} : x, y \in A;\ x \ne y\right\} \subset S^{n-1} ,$$

we want to select $U \in O(n)$ so that for some constant M,

$$M(1-\tau) \leq r(Ux) \leq M(1+\tau) \quad (x \in B) .$$

Let M_r be the median of $r(\cdot)$ on S^{n-1}, so that

$$\mu_{n-1}[x \in S^{n-1} : r(x) \geq M_r] \geq 1/2$$

and

$$\mu_{n-1}[x \in S^{n-1} : r(x) \leq M_r] \leq 1/2$$

where μ_{n-1} is normalized rotationally invariant measure on S^{n-1}.

We have from page 58 of [2] that for each $y \in S^{n-1}$ and $\varepsilon > 0$,

$$\sigma[U \in O(n) : M_r - \sqrt{n}\,\varepsilon \leq r(Uy) \leq M_r + \sqrt{n}\,\varepsilon] \geq 1 - 4 \exp\left(\frac{-n\varepsilon^2}{2}\right).$$

Hence

$$\text{(1.1)} \quad \sigma[U \in O(n) : M_r - \sqrt{n}\,\varepsilon \leq r(Uy)) \leq M_r + \sqrt{n}\,\varepsilon \quad \text{for all } y \in B] \geq$$

$$\geq 1 - 2n(n+1) \exp\left(\frac{-n\varepsilon^2}{2}\right).$$

By Lemma 1.7 of [2], there is a constant

$$C \leq 4 \sum_{m=1}^{\infty} (m+1)\, e^{-m^2/2}$$

so that

$$\text{(1.2)} \quad \left|\int_{S_{n-1}} r(x)\, d\mu_{n-1}(x) - M_r\right| < C .$$

We now repeat a known argument for estimating $\int_{S_{n-1}} r(x)\, d\mu_{n-1}(x)$ which uses only Khintchine's inequality.

For $1 \leq k \leq n$ we have:

$$\operatorname*{Av}_{\pm} \int_{S^{n-1}} \left|\sum_{i=1}^{k} \pm x(i)\right| d\mu_{n-1}(x) =$$

$$= \operatorname*{Av}_{\pm} \int_{S^{n-1}} \left|\left< x, \sum_{i=1}^{k} \pm \delta_i \right>\right| d\mu_{n-1}(x)$$

$$= \sqrt{k} \int_{S^{n-1}} |< x, \delta_1 >|\, d\mu_{n-1}(x) \quad \left[\begin{array}{l}\text{by the rotational} \\ \text{invariance of } \mu_{n-1}\end{array}\right] .$$

Setting

$$\alpha_n = \int_{S^{n-1}} |< x, \delta_1 >|\, d\mu_{n-1}(x) ,$$

we have from Khintchine's inequality that for each $1 \le k \le n$,

$$\sqrt{nk}\,\alpha_n \le \int_{S^{n-1}} r_k(x)\, d\mu_{n-1}(x) \le \sqrt{2nk}\,\alpha_n .$$

(We plugged in the exact constant of $\sqrt{2}$ in Khintchine's inequality calculated in [5] and [13], but of course any constant would serve as well.) Since obviously $r_n(x) = \sqrt{n}$, we conclude that for $1 \le k \le n$

(1.3) $$\sqrt{k/} \le \int_{S^{n-1}} r_k(x)\, d\mu_{n-1}(x) \le \sqrt{k} .$$

Specializing now to the case $k = [K \log n]$, we have from (1.2) and (1.3) that

$$\sqrt{k/3} \le M_r$$

at least for $K \log n$ sufficiently large. Thus if we define

$$\varepsilon = \tau \sqrt{k/3n}$$

we get from (1.1) that

$$\sigma\,[U \in O(n) : (1 - \tau)M_r \le r(Uy) \le (1 + \tau)M_r \quad \text{for all} \quad y \in B]$$

$$\ge 1 - 2n(n+1) \exp\left(-\frac{\tau^2 k}{18}\right)$$

$$\ge 1 - 2n(n+1) \exp\left(-\frac{\tau^2 K \log n}{18}\right)$$

which is positive if, say,

$$K \ge (10/\tau)^2 .$$ □

It is easily seen that the estimate $K \log n$ in Lemma 1 cannot be improved. Indeed, in a ball of radius 2 in ℓ_2^k there are at most 4^k vectors $\{x_i\}$ so that $\|x_i - x_j\| \ge 1$ for every $i \ne j$ (see the proof of Lemma 3 below). Hence for τ sufficiently small there is no map F which maps an orthonormal set with more than 4^k vectors into a k-dimensional subspace of ℓ_2 with

$$\|F\|_{\ell ip}\ \|F^{-1}\|_{\ell ip} \le \frac{1+\tau}{1-\tau} .$$

We can now verify the conjecture of Marcus and Pisier [10].