Dimension Reduction in L_p, $0 < p < 2$

Gideon Schechtman*

October 10, 2011

Abstract

Complementing a recent observation of Newman and Rabinovich for $p = 1$ we observe here that for all $0 < p < 2$ any k points in L_p embeds with distortion $(1 + \varepsilon)$ into ℓ_p^n where n is linear in k (and polynomial in ε^{-1}).

1 Introduction

The very well known Johnson–Lindenstrauss Lemma [JL] asserts that, for all k and $\varepsilon > 0$, any k points in a Hilbert space embed with distortion $1+\varepsilon$ into ℓ_2^n for $n = O(\varepsilon^{-2} \log k)$. It is also known that nothing similar to that occurs for the L_1 norm: There are k points in L_1 which if embedded in ℓ_1^n with distortion D forces $n \geq k^{c/D^2}$ for large D ([BC], and [LN] for a simpler proof) and $n \geq k^{1-O(1/\log(1/(D-1)))}$ for D close to 1 ([ACNN]). As for upper bound on n, until recently the best that was known was that any n points in L_p embed isometrically in $\ell_p^{O(k^2)}$ ([B]) and with distortion $1+\varepsilon$ in $\ell_p^{O(\varepsilon^{-2}k \log k)}$ for $0 < p < 2$ and in $\ell_p^{O(\varepsilon^{-2}k^{p/2} \log k)}$ for $2 < p < 4$ ([Sc1]). Recently, Newman and Rabinovich [NR] observed that a recent result of Batson, Spielman and Srivastava [BSS] implies that one can remove the $logk$ for $p = 1$ and get that k points in L_1 $(1+\varepsilon)$-embed into $\ell_p^{O(\varepsilon^{-2}k)}$. Here we show, also using [BSS], a similar result for all $0 < p < 2$. Our dependence on ε is however worse: Any k points in L_p, $0 < p < 2$, $(1+\varepsilon)$-embed into $\ell_p^{O(\varepsilon^{-(2+\frac{2}{p})}k)}$. The O notation hides a constant depending on p.

*Supported by the Israel Science Foundation. AMS subject classification: 46B85

2 The main result

We shall use the following two theorems. The first one is due in this form to Talagrand [T] and is not known to give the best possible dependence on ε. (For $\rho = 1/2$ there is a better result by Kahane [K].)

Theorem 1 *([T]) For each $\varepsilon > 0$ and $0 < \rho < 1$ there is a positive integer $k = k(\varepsilon, \rho)$ and a map $\varphi : \mathbb{R} \to \mathbb{R}^k$ such that*

$$(1+\varepsilon)^{-1}|x-y|^\rho \le \|\varphi(x) - \varphi(y)\|_2 \le (1+\varepsilon)|x-y|^\rho$$

for all $x, y \in \mathbb{R}$. Moreover, $k \le K(\rho)\varepsilon^{-1/\rho}$.

The second theorem we shall use is a relatively new theorem of Batson, Spielman and Srivastava.

Theorem 2 *([BSS]) Suppose $0 < \varepsilon < 1$ and $A = \sum_{i=1}^m v_i v_i^T$ are given, with v_i column vectors in $\mathbb{R}^k$. Then there are nonnegative weights $\{s_i\}_{i=1}^m$, at most $\lceil k/\varepsilon^2 \rceil$ of which are nonzero, such that, putting $\tilde{A} = \sum_{i=1}^m s_i v_i v_i^T$,*

$$(1+\varepsilon)^{-2} x^T A x \le x^T \tilde{A} x \le (1+\varepsilon)^2 x^T A x \qquad (1)$$

for all $x \in \mathbb{R}^k$.

We shall need the following simple corollary of this theorem which in turn is a generalization of Corollary 1 of [Sc2], dealing with the case $s = 1$.

Corollary 1 *Let $X_l, i = 1, 2, \ldots, s$, be s k-dimensional subspaces of ℓ_2^m and let $0 < \varepsilon < 1$. Then there is a set $\sigma \subset \{1, 2, \ldots, m\}$ of cardinality at most $n \le \varepsilon^{-2} ks$ and positive weights $\{s_i\}_{i \in \sigma}$ such that*

$$(1+\varepsilon)^{-1}\|x\|_2 \le \Big(\sum_{i\in\sigma} s_i x^2(i)\Big)^{1/2} \le (1+\varepsilon)\|x\|_2 \qquad (2)$$

for all $l = 1, 2, \ldots, s$ and all $x = (x(1), x(2), \ldots, x(m)) \in X_l$.

Proof: Let $u_1^l, u_2^l, \ldots, u_k^l$ be an orthonormal basis for X_l, $l = 1, 2, \ldots, s$; $u_j^l = (u_j^l(1), u_j^l(2), \ldots, u_j^l(m))$, $j = 1, \ldots, k$. Put $v_i^{lT} = (u_1^l(i), u_2^l(i), \ldots, u_k^l(i))$, $i = 1, \ldots, m$, $l = 1, \ldots, s$. Let also v_i be the concatenation of $v_i^1, v_i^2, \ldots, v_i^s$ forming a column vector in $\mathbb{R}^{ks}$. Then $A = \sum_{i=1}^m v_i v_i^T$, is a $ks \times ks$ matrix whose s $k \times k$ successive central submatrices are the $k \times k$ identity matrix.

Let s_i be the weights given by Theorem 2 with k replaces by ks. Let also $\sigma \subset \{1, \dots, m\}$ denote their support; The cardinality of σ is at most $\varepsilon^{-2}ks$. Let $l = 1, \dots, s$ and $x = \sum_{i=1}^{k} a_i u_i^l = (x(1), x(2), \dots, x(m)) \in X_l$. Apply (1) to the vector $\bar{a} \in \mathbb{R}^{ks}$ where $\bar{a}^T = (\bar{0}, \dots, \bar{0}, a^T, \bar{0}, \dots, \bar{0})^T$ with $\bar{0}$ denotes 0 vector in $\mathbb{R}^k$ and $a^T = (x_1, \dots, x_k)$ stand in the $(l-1)k+1$ to the lk places. Then

$$(1+\varepsilon)^{-2}\|x\|_2^2 = (1+\varepsilon)^{-2} a^T \sum_{i=1}^{m} v_i^l v_i^{lT} a \le a^T \sum_{i=1}^{m} s_i v_i^l v_i^{lT} a \le (1+\varepsilon)^2 \|x\|_2^2.$$

Finally, notice that, for each $i = 1, \dots, m$ and $l = 1, \dots, s$, $a^T v_i^l v_i^{lT} a = x(i)^2$ is the square of the i-th coordinate of x. Thus,

$$a^T \sum_{i=1}^{m} s_i v_i^l v_i^{lT} a = \sum_{i=1}^{m} s_i x(i)^2.$$

■

The main result of this note is:

Theorem 3 *For all $0 < p < 2$ there is a constant $K(p)$ such that for all $\varepsilon > 0$ and all $z_1, z_2, \dots, z_k$ in L_p there are $w_1, w_2, \dots, w_k$ in ℓ_p^n satisfying*

$$\|z_i - z_j\| \le \|w_i - w_j\| \le (1+\varepsilon)\|z_i - z_j\|$$

for all i, j, where $n \le K(p)k/\varepsilon^{2+\frac{2}{p}}$.

Proof: Let $\varphi : \mathbb{R} \to \mathbb{R}^s$ with $s \le K(p)\varepsilon^{-2/p}$ be the function from Theorem 1:

$$(1+\varepsilon)^{-1}|r - s|^{p/2} \le \|\varphi(r) - \varphi(r')\| \le (1+\varepsilon)|r - s|^{p/2} \tag{3}$$

for all $r, r' \in \mathbb{R}$. Assume as we may that $z_1, z_2, \dots, z_k \in \ell_p^m$ for some finite m and consider the map $\phi : \mathbb{R}^m \to \mathbb{R}^{ms}$ given by

$$\phi(r_1, r_2, \dots, r_m) = (\varphi(r_1), \varphi(r_2), \dots, \varphi(r_m)).$$

Let $P_l : \mathbb{R}^{ms} \to \mathbb{R}^m$, $l = 1, \dots, s$, be the restriction operator to the coordinates $\{l, s+l, s+2l, \dots, s+(m-1)l\}$. Consider the s subspaces of $\mathbb{R}^m$ given by

$$X_l = \text{span}\{P_l\phi(z_1), \dots, P_l\phi(z_k)\}$$

$l = 1, \dots, s$. Apply Corollary 1 to these s subspaces to get a set $\sigma \subset \{1, 2, \dots, m\}$ of cardinality at most $n \le \varepsilon^{-2} ks$ and positive weights $\{s_i\}_{i \in \sigma}$ such that

$$(1+\varepsilon)^{-1}\|x\|_2 \le (\sum_{i\in\sigma} s_i x^2(i))^{1/2} \le (1+\varepsilon)\|x\|_2 \tag{4}$$

for all $l = 1, 2, \dots, s$ and all $x = (x(1), x(2), \dots, x(m)) \in X_l$. Applying (4) to $x = P_l\phi(z_u) - P_l\phi(z_v)$ we get

$$\begin{aligned}(1+\varepsilon)^{-2}\|P_l\phi(z_u) - P_l\phi(z_v)\|_2^2 \le &\sum_{i\in\sigma} s_i(P_l\phi(z_u) - P_l\phi(z_v))^2(i) \\ \le &(1+\varepsilon)^2\|P_l\phi(z_u) - P_l\phi(z_v)\|_2^2.\end{aligned} \tag{5}$$

Adding up these s inequalities, we get

$$(1+\varepsilon)^{-2}\|\phi(z_u)-\phi(z_v)\|_2^2 \le \sum_{i\in\sigma} s_i\|(\phi(z_u)-\phi(z_v))(i)\|_2^2 \le (1+\varepsilon)^2\|\phi(z_u)-\phi(z_v)\|_2^2$$

where by $\phi(\bar{r})(i)$ we mean the restriction of $\phi(\bar{r})$ to the s coordinates $(i-1)s+1$ to is. Applying (3) we now get

$$(1+\varepsilon)^{-6}\|z_u - z_v\|_p^p \le \sum_{i\in\sigma} s_i\|(z_u - z_v)(i)\|_p^p \le (1+\varepsilon)^6\|z_u - z_v\|_p^p$$

for all u and v. ∎

References

[ACNN] A. Andoni, M. Charikar, O. Neiman, and H.L. Nguyen, Near Linear Lower Bounds for Dimension Reduction in ℓ_1, 52th Annual IEEE Symposium on Foundation of Computer Science, October 2011.

[B] K. Ball, Markov chains, Riesz transforms and Lipschitz maps, Geom. Funct. Anal., vol. 2, no. 2, pp. 137172, 1992.

[BSS] J.D. Batson, D.A. Spielman, and N. Srivastava, Twice-Ramanujan sparsifiers, STOC 09: Proceedings of the 41st annual ACM symposium on Theory of computing (New York, NY, USA, 2009), ACM, pp. 255262.

[BC] B. Brinkman and M. Charikar, On the impossibility of dimension reduction in ℓ_1, J. ACM, vol. 52, pp. 766788, September 2005.

[JL] W. B. Johnson and J. Lindenstrauss, Extensions of Lipschitz mappings into a Hilbert space, in Conference in modern analysis and probability (New Haven, Conn., 1982). Providence, RI: Amer. Math. Soc., 1984, pp. 189206.

[K] J-P. Kahane, Hélices et quasi-hélices, (French) [Helices and quasihelices] Mathematical analysis and applications, Part B, pp. 417433, Adv. in Math. Suppl. Stud., 7b, Academic Press, New York-London, 1981.

[LN] J. R. Lee and A. Naor, Embedding the diamond graph in ℓ_p and dimension reduction in ℓ_1, Geometric and Functional Analysis (GAFA) 14(4), pp. 745747, 2004.

[NR] Ilan Newman, Yuri Rabinovich, Finite Volume Spaces and Sparsification, http://arxiv.org/abs/1002.3541

[Sc1] G. Schechtman,More on embedding subspaces of L_p in ℓ_r^n, Compositio Math. 61 (1987), no. 2, 159169.

[Sc2] G. Schechtman, Tight embedding of subspaces of L_p in ℓ_p^n for even p, Proc. Amer. Math. Soc. 139 (2011), 4419-4421.

[T] M. Talagrand, Approximating a helix in finitely many dimensions. Ann. Inst. H. Poincar Probab. Statist. 28 (1992), no. 3, 355363.

Gideon Schechtman
Department of Mathematics
Weizmann Institute of Science
Rehovot, Israel
E-mail: gideon.schechtman@weizmann.ac.il