ISRAEL JOURNAL OF MATHEMATICS, Vol. 63, No. 2, 1988

l_p^n SUPERSPACES OF SPANS OF INDEPENDENT RANDOM VARIABLES

BY

A. ARIAS†
Department of Mathematics, Texas A&M University, College Station, TX 77843, USA

ABSTRACT

We show that for $1 \leqq p < \infty$, $p \neq 2$, if $\varepsilon > 0$ is small enough and $X \leqq L_p$ is the span of n independent Rademacher functions or n independent Gaussian random variables, then any superspace Y of X satisfying $d(Y, L_p^m) \leqq 1 + \varepsilon$ has dimension larger than r^n, where $r = r(\varepsilon, p) > 1$.

1. Introduction

In this paper we investigate a quantitative aspect of the local $\mathscr{L}_p$-structure of L_p. The problem we consider is:

($\mathscr{L}$) Given a subspace X of L_p, $\dim X = n$, and $\varepsilon > 0$, estimate the smallest $m = m_p(X, \varepsilon)$ such that there is a subspace Y of L_p with $X \subseteq Y$ and $d(Y, l_p^m) \leqq 1 + \varepsilon$. In particular, estimate $m_p(n, \varepsilon) = \sup\{m_p(X, \varepsilon) : \dim X = n\}$.

The concept was introduced by Pelczynski and Rosenthal [PR], who proved that $m_p(n, \varepsilon)$ is finite. In the same paper there is an argument, due to Kwapien, that $m_p(n, \varepsilon)$ is of order no larger than $(n/\varepsilon)^{Cn}$ for some constant C. More precise estimates were given by Figiel, Johnson and Schechtman [FJS], who proved that "exponential of n" is the right order of $m_\infty(n, \varepsilon)$ and for "natural" n-dimensional Euclidean subspaces X of L_1, $m_1(X, \varepsilon) \geqq r^n$, where $r = r(\varepsilon) > 1$ is independent of n and $\varepsilon > 0$ is arbitrary.

In the present paper we prove that if $\varepsilon > 0$ is small enough, and X is the

† This forms part of the author's doctoral dissertation prepared at Texas A&M University under the direction of Professor W. B. Johnson.

Supported in part by NSF DMS-85 00764.
Received December 3, 1987 and in revised form May 2, 1988

subspace of L_p spanned by n independent Rademacher functions or by n independent Gaussian random variables, then $m_p(X, \varepsilon) \geqq r^n$ where $r = r(\varepsilon, p) > 1$. More generally.

THEOREM 1.1. *Let $1 \leqq p < \infty$, $p \neq 2$. For every $\eta > 0$ and $0 \leqq \gamma < \frac{1}{2}$ there exists $\varepsilon_p = \varepsilon_p(\eta, \gamma, p) > 0$ such that if $X_n = \operatorname{span}\{g_1, \dots, g_n\}$ where $\| g_i \|_p = 1$, $P[|g_i| < \eta] \leqq \gamma$ for $i = 1, 2, \dots, n$ and the* sgn *g_i's are independent identically distributed* (i.i.d.) *random variables taking values* 1 *and* -1 *with probability $\frac{1}{2}$; and $0 < \varepsilon < \varepsilon_p$, then we can find $r = r(\varepsilon, \eta, \gamma, p) > 1$ and such that $m_p(X_n, \varepsilon) \geqq r^n$.*

REMARKS. (1) Theorem 1.1 is not true for arbitrarily large $\varepsilon > 0$ and $p > 1$. Johnson and Schechtman proved in [JS] that for ε large enough and for g_i independent random variables, $m_p(X_n, \varepsilon)$ satisfies a polynomial upper estimate. It would be interesting to see if $\lim_{p \to 1} \varepsilon_p = \infty$.

(2) We can eliminate the restriction $P[g_i = 0] = 0$ *for* $i = 1, \dots, n$ of Theorem 1.1 by requiring: $P[|g_i| < \eta] \leqq \gamma < \frac{1}{4}$ (instead of $\gamma < \frac{1}{2}$) *for $i = 1, \dots, n$ and g_i's are independent symmetric random variables*. The proof goes essentially the same way.

If the g_i's are independent symmetric random variables, we can use Kanter's inequality (see [AG] p. 112),

$$(*) \qquad P\left[\left|\sum_{i=1}^{k} g_i\right| < \eta\right] \leqq \frac{\frac{3}{2}}{\left(1 + \sum_{i=1}^{k} P[|g_i| \geqq \eta]\right)^{1/2}},$$

to prove:

COROLLARY 1.2. *Let $1 \leqq p < \infty$, $p \neq 2$. For every $\eta > 0$ and $0 \leqq \gamma < 1$ there exists $\varepsilon_p = \varepsilon_p(\eta, \gamma, p) > 0$ such that if $X_n = \operatorname{span}\{f_1, \dots, f_n\}$ where the f_i's are normalized independent symmetric random variables satisfying $P[|f_i| < \eta] \leqq \gamma$ for $i = 1, 2, \dots, n$; and $0 < \varepsilon < \varepsilon_p$, then we can find $r = r(\varepsilon, \eta, \gamma, p) > 1$ such that $m_p(X_n, \varepsilon) \geqq r^n$.*

PROOF. Take k (independent of n) so that the right hand side of $(*)$ is less than $\frac{1}{4}$. Let

$$g_i = \sum_{j=(i-1)k-1}^{ik} f_j \Big/ \left\| \sum_{j=(i-1)k-1}^{ik} f_j \right\|, \qquad i = 1, \dots, [n/k] \quad \text{and} \quad \eta' = \eta/k.$$

Then span $\{g_i\}_{i=1}^{[n/k]}$ satisfies the hypothesis of Remark 2 and the proof follows. ∎

Theorem 1.1 answers a question raised in [FJS], and gives a variety of subspaces of L_p for which $m_p(X, \varepsilon)$ satisfies an "exponential lower bound". The remaining main problem is to see it there is a similar exponential behavior for the uniform approximation property (or uniform projection approximation property) for L_p, $1 \leqq p \neq 2 < \infty$. It was shown in [FJS] that this is the case for $p = 1$.

It also remains open whether $m_p(n, \varepsilon)$ admits an exponential upper estimate; or, at least, when X_n is the span of n i.i.d. symmetric random variables. Figiel [F] proved that this is correct if X_n is the span of n i.i.d. Gaussian random variables and $p = 1$.

2. Proof of Theorem 1.1

The main tool for the proof is:

THEOREM 2.1 (Dor–Schechtman) [D], [S]. *Let* $1 \leqq p < \infty$, $p \neq 2$. *There exists a function* $d(\varepsilon)$ *such that* $d(\varepsilon) \to 0$ *as* $\varepsilon \to 0$ *and, if* $f_1, \ldots, f_m$ *are functions in* $L_p[0, 1]$ *which satisfy*

$$(1) \qquad (1-\varepsilon)\left(\sum_{i=1}^{m} |a_i|^p\right)^{1/p} \leqq \left\| \sum_{i=1}^{m} a_i f_i \right\|_p \leqq (1+\varepsilon)\left(\sum_{i=1}^{m} |a_i|^p\right)^{1/p}$$

for all scalars $a_1, \ldots, a_m$; *then there exists a partition* $A_1, \ldots, A_m$ *of* $[0, 1]$ *such that*

$$(2) \qquad \left\| \sum_{i=1}^{m} a_i f_{i|A_i^c} \right\|_p \leqq d(\varepsilon)\left(\sum_{i=1}^{m} |a_i|^p\right)^{1/p}.$$

PROOF OF THEOREM 1.1. Take $Y \subseteq L_p[0, 1]$ such that $X_n \subset Y$ and $d(Y, l_p^m) \leqq 1 + \varepsilon$, where $m = m_p(X_n, \varepsilon)$. Then we can find:

(i) $f_1, \ldots, f_m$ in Y satisfying (1),

(ii) A partition $A_1, \ldots, A_m$ of $[0, 1]$ satisfying (2), and

(iii) some constants a_{ik} such that

$$g_k = \sum_{i=1}^{m} a_{ik} f_i \qquad \text{for } k = 1, 2, \ldots, n.$$

Using (2) and the fact that $\| g_k \|_p = 1$, we get

$$\left\| \sum_{i=1}^{m} a_{ik} f_{i|A_i^c} \right\|_p \leqq \frac{d(\varepsilon)}{1-\varepsilon} \qquad \text{for } k = 1, \ldots, n.$$

If ε is small enough, we have that

$$g_k \approx \sum_{i=1}^{m} a_{ik} f_{i|A_i}.$$

Then, for "most" of the $j = 1, \ldots, m$ and $k = 1, \ldots, n$; $a_{jk} f_{j|A_j}$ is a "good" approximation of $g_{k|A_j}$. Since $|g| \geqq \eta$ in a set of big measure,

$$P[\omega \in A_j : \operatorname{sgn} g_k(\omega) \neq \operatorname{sgn} a_{jk} \operatorname{sgn} f_j(\omega)] \qquad (\operatorname{sgn} 0 = 1)$$

is a "reasonable" estimate of the measure of the set of all $\omega \in A_j$ such that $g_k(\omega)$ is "not close" to $a_{jk} f_j(\omega)$. (Notice that if $g_k = r_k$, the kth Rademacher function, then $\operatorname{sgn} g_k(\omega) \neq \operatorname{sgn} a_{jk} \operatorname{sgn} f_j(\omega)$ implies $|g_k(\omega) - a_{jk} f_j(\omega)| \geqq 1$.)

The quantity we want to estimate is

$$q(n) = \frac{1}{n} \sum_{k=1}^{n} \left[\sum_{j=1}^{m} P[w \in A_j : \operatorname{sgn} g_k(\omega) \neq \operatorname{sgn} a_{jk} \operatorname{sgn} f_j(\omega)] \right].$$

It represents the average of the measure of the sets where g_k is "not close" to $\sum_{i=1}^{m} a_{ik} f_{i|A_i}$.

The idea of the proof is that, on the one hand, this quantity must be small since $g_k \approx \sum_{i=1}^{n} a_{ik} f_{i|A_i}$, and, on the other hand, if m is small relative to n, the independence of $\operatorname{sgn} g_k$ forces it to be large.

CLAIM 1.

$$q(n) \leqq \left[\frac{d(\varepsilon)}{\eta(1-\varepsilon)} \right]^p + \gamma.$$

Set

$$b(\varepsilon) = \left[\frac{d(\varepsilon)}{\eta(1-\varepsilon)} \right]^p + \gamma + \varepsilon$$

and choose $\varepsilon_p > 0$ so that $b(\varepsilon_p) < \frac{1}{2}$.

CLAIM 2. Let $h = h(n)$ be the smallest number satisfying

$$\frac{2m}{2^n} \left[\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{h} \right] \geqq 1. \tag{3}$$

Then there exists $N \in \mathbf{Z}^+$ such that for $n \geqq N$, we have

$$h/n \leqq b(\varepsilon_p) - \varepsilon_p/2.$$

We postpone the proof of Claims 1 and 2 and finish the proof of Theorem 1.1. Taking a crude estimate of (3), we obtain

$$1 \leqq n \frac{\binom{n}{h}}{2^n} m. \tag{4}$$

By Stirling's formula, there is a constant $A > 0$ such that

$$\frac{\binom{n}{h}}{2^n} \leqq A \left[\frac{f(h/n)}{2}\right]^n,$$

where the function

$$f(x) = \frac{(1-x)^{x-1}}{x^x}$$

satisfies $f(0) = 1, f(\frac{1}{2}) = 2$ and $f'(x) > 0$ for $0 < x < 1/2$.

Since f is increasing, for $n \geqq N$ we have that

$$\frac{\binom{n}{h}}{2^n} \leqq A \left[\frac{f(b(\varepsilon_p) - \varepsilon_p/2)}{2}\right]^n. \tag{5}$$

Set $r = 2/f(b(\varepsilon_p))$ (notice that $r > 1$).

Substituting (5) in (4), we see that m has to be at least of order r^n to compensate $\binom{n}{h}/2^n$. That is, we can find $N_1 \in \mathbf{Z}^+$ s.t.

$$m \geqq r^n \qquad \text{for every } n \geqq N_1. \qquad \blacksquare$$

REMARK. If $g_k = r_k$ for every k, and we take $\eta = 1$ and $\gamma = 0$; we have that $\lim_{\varepsilon \to 0} b(\varepsilon) = 0$. Then we get that

$$\liminf_n m^{1/n} \geqq 2.$$

This follows from the fact that $\lim_{\varepsilon \to 0} r(\varepsilon) = 2$. Thus we recover the "right order" for the Rademacher case.

PROOF OF CLAIM 1. For any $k = 1, \ldots, n$ we have

$$\begin{aligned}\left\|\sum_{i=1}^m a_{ik} f_{i|A_i^c}\right\|_p^p &\geqq \eta^p P\left[\left|\sum_{i=1}^m a_{ik} f_{i|A_i^c}\right| \geqq \eta\right]\\ &\geqq \eta^p P\left[\omega \in B_k : \left|\sum_{i=1}^m a_{ik} f_{i|A_i^c}\right| \geqq \eta\right] \qquad \text{where } B_k = [\,|g_k| \geqq \eta]\\ &= \eta^p \sum_{j=1}^m P[\omega \in B_k \cap A_j : |g_k(\omega) - a_{jk} f_j(\omega)| \geqq \eta]\\ &\geqq \eta^p \sum_{j=1}^m P[\omega \in B_k \cap A_j : \operatorname{sgn} g_k(\omega) \neq \operatorname{sgn} a_{jk} \operatorname{sgn} f_j(\omega)]\\ &\geqq \eta^p \left[\sum_{j=1}^m P[\omega \in A_j : \operatorname{sgn} g_k(\omega) \neq \operatorname{sgn} a_{jk} \operatorname{sgn} f_j(\omega)] - P[\,|g_k| < \eta]\right].\end{aligned}$$

Therefore,

$$\begin{aligned}\sum_{j=1}^m P[\omega \in A_j : \operatorname{sgn} g_k(\omega) \neq \operatorname{sgn} a_{jk} \operatorname{sgn} f_j(\omega)] &\leqq \frac{\left\|\sum_1^m a_{ik} f_{i|A_i^c}\right\|_p^p}{\eta^p} + P[\,|g_k| < \eta]\\ &\leqq \left[\frac{d(\varepsilon)}{\eta(1-\varepsilon)}\right]^p + \gamma,\end{aligned}$$

and taking the average on k, we finish the proof. ■

Proof of Claim 2. We will identify $[0,1]$ with $\{-1, 1\}^{\mathbf{N}}$ in such a way that $\operatorname{sgn} g_k = r_k$ for $k = 1, \ldots, n$. Since we will be interested in counting the number of 1's of elements of $\{-1, 1\}^{\mathbf{N}}$ we will use the Bernoulli functions $d_k = (r_k + 1)/2$.

We want to define a funtion $\varphi : \{-1, 1\}^{\mathbf{N}} \to \{-1, 1\}^{\mathbf{N}}$ that will let us handle $q(n)$ in an easier way (remember $\operatorname{sgn} 0 = 1$). For $\omega \in A_j$, let

$$\begin{aligned} r_k(\varphi(\omega)) &= -\operatorname{sgn} g_k(\omega) \operatorname{sgn} a_{jk} \operatorname{sgn} f_j(\omega) \qquad &&\text{for } k = 1, \ldots, n, \qquad \text{and}\\ r_k(\varphi(\omega)) &= r_k(\omega) &&\text{for } k = n+1, n+2, \ldots .\end{aligned}$$

We can easily check that

$$P[\omega \in A_j : \operatorname{sgn} g_k(\omega) \neq \operatorname{sgn} a_{jk} \operatorname{sgn} f_j(\omega)] = \int_{A_j} d_k(\varphi(\omega)) d\omega.$$

Moreover, given ω_1, ω_2 in A_j such that $\varphi(\omega_1) = \varphi(\omega_2)$, then

$$r_k(\omega_1) = r_k(\omega_2) \qquad \text{for } k = n+1, n+2, \ldots$$

and (remember, $\operatorname{sgn} g_k = r_k$ for $k = 1, \dots, n$) either

$$r_k(\omega_1) = r_k(\omega_2) \qquad \text{for } k = 1, \dots, n,$$

or

$$r_k(\omega_1) = -r_k(\omega_2) \qquad \text{for } k = 1, \dots, n.$$

Therefore

$$\operatorname{card}\{\omega \in A_j : \varphi(\omega) = \omega_0\} \leqq 2 \qquad \text{for all } \omega_0 \in \{-1, 1\}^{\mathbf{N}}.$$

Adding all the pieces together, we get

$$\sum_{j=1}^{m} P[\omega \in A_j : \operatorname{sgn} g_k(\omega) \neq \operatorname{sgn} a_{jk} \operatorname{sgn} f_j(\omega)] = \int_0^1 d_k(\varphi(\omega)) d\omega$$

and $\operatorname{card}\{\omega \in \{-1, 1\}^{\mathbf{N}} : \varphi(\omega) = \omega_0\} \leqq 2m$ for every $\omega_0 \in \{-1, 1\}^{\mathbf{N}}$.

We want to estimate from below the quantity

$$\frac{1}{n} \int_0^1 \left[\sum_{k=1}^{n} d_k(\varphi(\omega)) \right] d\omega.$$

We have that

$$\int_0^1 \left[\sum_{k=1}^{n} d_k(\varphi(\omega)) \right] d\omega = \sum_{i=1}^{n} iP\left[\omega : \sum_{k=1}^{n} d_k(\varphi(\omega)) = i\right].$$

Let $S_i = [\omega : \Sigma_{k=1}^{n} d_k(\omega) = i]$. Since $\operatorname{card} \varphi^{-1}(\omega) \leqq 2m$ for every ω,

$$S_i = \bigcup_{j=0}^{2m} [\omega \in S_i : \operatorname{card} \varphi^{-1}(\omega) = j].$$

Since φ only changes the first n coordinates of ω, it is 1–1 and measure preserving in each of "the 2^n basic dyadic intervals". Therefore

$$\begin{aligned} P\left[\omega : \sum_{k=1}^{n} d_k(\varphi(\omega)) = i\right] &= \sum_{j=1}^{2m} jP[\omega \in S_i : \operatorname{card} \varphi^{-1}(\omega) = j] \\ &\leqq 2m \sum_{j=1}^{2m} P[\omega \in S_i : \operatorname{card} \varphi^{-1}(\omega) = j] \\ &\leqq 2mP(S_i) = 2m \frac{\binom{n}{i}}{2^n}, \end{aligned}$$

and we have equality only if card $\varphi^{-1}(\omega)=2m$ a.e. in S_i. The worst case occurs when the image of φ is "concentrated" where there are few 1's. That is

$$\frac{1}{n}\int_0^1\left[\sum_{k=1}^n d_k(\varphi(\omega))\right]d\omega$$

(6)

$$\geqq \frac{0}{n}\left[2m\frac{\binom{n}{0}}{2^n}\right]+\cdots+\frac{(h-1)}{n}\left[2m\frac{\binom{n}{h-1}}{2^n}\right]+\frac{h}{n}\left[2m\,p_n\frac{\binom{n}{h}}{2^n}\right],$$

where h is the mininum number satisfying

$$\frac{2m}{2^n}\left[\binom{n}{0}+\cdots+\binom{n}{0}\right]\geqq 1$$

and p_n is chosen so that

$$\frac{2m}{2^n}\left[\binom{n}{0}+\cdots+\binom{n}{h-1}+p_n\binom{n}{h}\right]=1.$$

We will show that the right hand side of (6) is essentially h/n, and this together with Claim 1 will finish the proof.

For any $0<c<\frac{1}{2}$ such that $cn\in\mathbf{Z}^+$ we can easily check that

$$\binom{n}{cn-k}\leqq\left(\frac{c}{1-c}\right)^k\binom{n}{cn}\qquad\text{for } k=1,\ldots,cn.$$

Take $0<a<1$ and let $c=[ah]/n$, then

$$\frac{\binom{n}{0}+\cdots+\binom{n}{[a^2h]}}{\binom{n}{[ah]}}\leqq\left(\frac{c}{1-c}\right)^{(1-a)ah}\left[1+\frac{c}{1-c}+\cdots\right]$$

which goes to zero as $n\to\infty$ ($[x]$ is the integer part of x). Therefore

$$\frac{2m}{2^n}\left[\binom{n}{[a^2h]+1}+\cdots+\binom{n}{h-1}+p_n\binom{n}{h}\right]=1-e_a(n)$$

where $e_a(n)\to 0$ as $n\to\infty$. Hence, (6) implies that

$$\frac{1}{n}\int_0^1\left[\sum_{i=1}^n d_i(\varphi(\omega))\right]d\omega \geqq \frac{[a^2h]}{n}(1-e_a(n)).$$

Claim (1) says (notice that $q(n)=(1/n)\int_0^1 \Sigma_{i=1}^n d_i(\varphi(\omega))d\omega$)

$$\frac{1}{n}\int_0^1\left[\sum_{i=1}^n d_i(\varphi(\omega))\right]d\omega \leqq b(\varepsilon_p)-\varepsilon_p.$$

Therefore

$$\frac{h}{n}\leqq\frac{b(\varepsilon_p)-\varepsilon_p}{a^2(1-e_a(n))}$$

and an appropriate selection of $0<a<1$ and $N\in\mathbf{Z}^+$ gives us that

$$\frac{h}{n}\leqq b(\varepsilon_p)-\varepsilon_p/2 \qquad \text{whenever } n\geqq N. \qquad \blacksquare$$

3. An application

We finish the paper with an application of Theorem 1.1 which was pointed out to us by G. Schechtman.

PROPOSITION. *Let $1\leqq p<2$. For every $\varepsilon>0$ there exist $\delta=\delta(\varepsilon,p)>0$ and $r=r(\varepsilon,p)>1$ such that for every $n\in\mathbf{Z}^+$ there is an n-dimensional subspace $X_n\subseteq L_p[0,1]$ such that $d(X_n,l_p^n)\leqq\varepsilon$ and $m_p(X_n,\varepsilon)\geqq r^n$.*

PROOF. Let $\varepsilon>0$ and $n\in\mathbf{Z}^+$. Take $f_1,\ldots,f_n$ a sequence of 3-valued symmetric i.i.d. random variables such that $\|f_i\|_p=1$ and $\operatorname{supp} f_i$ small enough to insure that

$$d(\operatorname{span}\{f_1,\ldots,f_n\},l_p^n)\leqq 1+\varepsilon/2.$$

Let $r_1,\ldots,r_n$ be a copy of the Rademacher functions independent of the f_i's. Set $g_k=\eta' f_k+\eta r_k$ where η,η' are chosen so that $\|g_k\|_p=1$ and $\eta=\eta(\varepsilon)>0$ small enough to have

$$d(X_n,l_p^n)\leqq 1+\varepsilon$$

where $X_n=\operatorname{span}\{g_1,\ldots,g_n\}$.

The result follows easily from Theorem 1.1. $\blacksquare$

REMARK. A weaker form is true also for $2<p<\infty$. Let $X_n=\operatorname{span}\{r_k\}_{r=1}^n$. As in Remark 12 of [JS] we can find Y_n such that $X_n\subseteq Y_n\subseteq L_p$ with $d(Y_n,l_p^{\lim Y_n})\leqq K_p$ and $\dim Y_n\leqq K_p n^{p/2}$ (K_p is a constant depending only on p).

Since $m_p(X_n, \varepsilon) \geqq r^n$ then we have that $m_p(Y_n, \varepsilon) \geqq r^n \geqq r_0(\dim Y_n)^{2/p}$ where $r_0 = r_0(\varepsilon, p) > 1$.

Acknowledgement

The author wants to thank Professors E. Giné, W. B. Johnson, G. Schechtman and J. Zinn for helpful discussions concerning the preparations of this paper.

References

[AG] A. Araujo and E. Giné, *The Central Limit Theorem for Real and Banach Valued Random Variables*, John Wiley & Sons, New York, 1980.

[D] L. E. Dor, *On projections in L_1*, Ann. of Math. **102** (1975), 463–474.

[F] T. Figiel, *The exponential estimate for local structure of Gaussian subspaces of L_p*, to appear

[FJS] T. Figiel, W. B. Johnson and G. Schechtman, *Factorization of natural embeddings of l_p^n into L_r, I*, Studia Math., to appear.

[JS] W. B. Johnson and G. Schechtman, *Sums of independent random variables in rearrangement invariant function spaces*, to appear.

[PR] A. Pelczynski and H. P. Rosenthal, *Localization techniques in L^p-spaces*, Studia Math. **52** (1975), 263–289.

[S] G. Schechtman, *Almost isometric L_p subspaces of $L_p(0, 1)$*, J. London Math. Soc. **20** (1979), 516–528.