Tight Embeddings and Weighted Samplings for subspaces of L_p

Gideon Schechtman

Online Workshop on Discretization and related questions

March 2021

Tight embeddings for subspaces of L_p

- L_p will denote any L_p space over any measure space but for the problems I'll consider one can think of $L_p = L_p[0,1]$ with respect to Lebesgue measure.
- $0 < p < \infty$.
- Given a finite dimensional subspace $X \subseteq L_p$ and $\varepsilon > 0$ denote

$$\boxed{N_p(X,\varepsilon) = \inf\{n;\ X \overset{1+\varepsilon}{\hookrightarrow} \ell_p^n\}.}$$

Here $X \overset{K}{\hookrightarrow} Y$ means

$\exists$ linear $T: X \to Y$ with

$$\|x\| \le \|Tx\| \le K\|x\|.$$

$$\boxed{N_p(X,\varepsilon) = \inf\{n;\ X \overset{1+\varepsilon}{\hookrightarrow} \ell_p^n\}} .$$

Denote also

$$\boxed{N_p(k,\varepsilon) = \sup\{N_p(X,\varepsilon);\ \dim X = k\}}$$

- The investigation of these quantities was a subject of intensive investigation mostly in the 1980-s 1990-s.

- We shall see later the relation with sampling/discretization

- Except for log factors and the exact dependence on ε, p the dependence of $N_p(k,\varepsilon)$ on k is now known:

$$\boxed{N_p(k,\varepsilon) \underset{\log k}{\overset{c(p,\varepsilon)}{\approx}} \begin{cases} k & 0 < p \leq 2 \\ k^{p/2} & 2 < p < \infty \end{cases}}$$

History w/personal bias

Dvoretzky (1960): $\ell_2^k \overset{1+\varepsilon}{\hookrightarrow} X$ for any X with $\dim(X)$ large enough $\left(\geq e^{c(\varepsilon)k^2}\right)$

Milman (1971): same with $\dim X \geq e^{c(\varepsilon)k}$

The proof gives much more, in particular

FLM (1977) $\ell_2^k \overset{1+\varepsilon}{\hookrightarrow} \ell_p^n$ for

$$\boxed{n \leq C(p,\varepsilon) \begin{cases} k^{p/2} & p > 2 \\ k & 1 \leq p < 2 \end{cases}}$$

$(0 < p < 1$ o.k. too$)$

Also, Kashin (77): $\forall\ \lambda > 1$

$\ell_2^k \overset{C(\lambda)}{\hookrightarrow} \ell_p^{\lambda k}$, $1 \leq p < 2$

Dvoretzky, Milman, FLM use "concentration" related to the exponential decay of Gaussian variables.

- For $1 \leq p < q < 2$ $\ell_q \overset{1}{\hookrightarrow} L_p$, even $L_q \overset{1}{\hookrightarrow} L_p$ using q stable r.v. (Levy, Kadec)

• q stables have only polynomial tail. So the following was surprizing at the time:

JS (1982): $\ell_q^k \overset{1+\varepsilon}{\hookrightarrow} \ell_p^{ck}$, $1 \le p < q < 2$, $c = c(p, q, \varepsilon)$.

(using discretization of q-stables).

(1980's) I found some polynomyal estimates for $N_p(k, \varepsilon)$ Using complicated methods based on JS

(1987) A much simpler proof basically for all that was known using

$\boxed{\text{Change of density + sampling}}$

(89- 90's) The real powers (including Bourgain & Talagrand) took over -

added techniques of entropy estimates

and estimates of size of Gaussian processes.

- Best known results by ~2000:

$$\boxed{N_p(k,\varepsilon) \le C(p,\varepsilon) \begin{cases} k^{p/2} \log k, & 2<p, \quad [BLM] \\ k \log k (\log\log k)^2, & 1<p<2 \quad [T] \\ & 0<p<1 \quad [Z] \\ k \log k, & p=1 \quad [T] \end{cases}}$$

- Except for the log factors these can't be improved [BGDJN] (1977).

- Also, $0<p<q\le 2$, $X \subseteq L_p$, $X \overset{2}{\hookrightarrow} L_q$,

$$\boxed{N_p(X,\varepsilon) \le C(p,q,\varepsilon) \dim X}$$

- In the 21 century:

[JS] $0<p<q\le 2$, $\lambda>1$, $X \subset L_q$

$$\boxed{N_p(X, C(p,q,\lambda)) \le \lambda \dim X}$$

[S] p - even integer:

$$\boxed{N_p(k,\varepsilon) \le C(p,\varepsilon) k^{p/2}}$$

(using Batson, Spielman, Srivastava (2009)).

Sampling & change of density

Let (Ω, μ) be a probability space and $X \subseteq L_p(\Omega, \mu)$ of dimension k.

A naive way to try and embed X in ℓ_p^n is to find $t_1, \dots, t_n \in \Omega$ s.t.

$$\boxed{(1-\varepsilon)\|f\|_p \leq \left(\frac{1}{n}\sum |f(t_i)|^p\right)^{1/p} \leq (1+\varepsilon)\|f\|_p}$$

for all $f \in X$.

For example if $X = \operatorname{span}\{\chi_{(0,\frac{1}{2})}, \chi_{[\frac{1}{2},1]}\}$ in $L_p(0,1)$. Then taking $t_1 \in (0, \frac{1}{2})$, $t_2 \in [\frac{1}{2}, 1)$ gives $\frac{1}{2}\left(|f(t_1)|^p + |f(t_2)|^p\right) = \|f\|_{L_p}^p$.

<u>BUT!</u> if $Y = \text{span}\{\chi_{(0,\delta)}, \chi_{(\delta,1)}\}$

and $f = a\chi_{(0,\delta)} + b\chi_{(\delta,1)}$

$t_1 \in (0,\delta)$, $t_2 \in (\delta,1)$ then

$$\frac{1}{2}\left(|f(t_1)|^p + |f(t_2)|^p\right) = \frac{1}{2}\left(|a|^p + |b|^p\right)$$

where $\|f\|_p^p = \delta|a|^p + (1-\delta)|b|^p$

and the two quatities are very different if δ is small.

Same holds if we sample more points.

Also, if the sampling is random we are likely to miss the interval $(0,\delta)$ all together!

Note: X and Y are isometric.
So for the purposes of tight embeddings
if we were given a "bad" subspace
like Y we would try and first
change it to a "good" isometric
copy and then do some sampling.

This was the basic idea in my
paper from 1987.

The isometry I used was a certain
"change of density".

change of density

Given $g \in L_1^+(\Omega, \mu)$, $\|g\|_1 = 1$ define

$$T_g : L_p(\Omega, \mu) \to L_p(\Omega, g\,d\mu) \quad \text{by}$$

$$T_g f = \frac{f}{g^{1/p}}$$

$$\|T_g f\|_p = \left(\int \frac{|f|^p}{g} \, g \, d\mu \right)^{1/p} = \left(\int |f|^p d\mu \right)^{1/p} = \|f\|_p$$

---

g or T_g are called change of density.

The point is, given an $X \subseteq L_p$, to find a change of density that makes the function in $T_g X$ as "flat" as possible.

The one that seems to work best is Lewis' change of density.

Lewis (1978): $X \subseteq L_p$ of dim k

Then $\exists$ change of density g s.t. $T_g X$ admits an orthonormal basis (in $L_2(g\,d\mu)$) $f_1, \dots, f_k$ with $\sum_{i=1}^{k} f_i^2 \equiv k$.

Note, if $f \in T_g X$, $0 < p < 2$, say $f = \sum_{i=1}^{k} a_i f_i$,

$$\|f\|_\infty \leq \left(\sum a_i^2\right)^{1/2} k^{1/2} = \|\sum a_i f_i\|_{L_2} k^{1/2}.$$

$$\leq \|\sum a_i f_i\|_p^{p/2} \|\sum a_i f_i\|_\infty^{1-p/2} k^{1/2}.$$

So $\|f\|_\infty^{p/2} \leq \|f\|_p^{p/2} k^{1/2}$ and $\boxed{\|f\|_\infty \leq k^{1/p} \|f\|_p}$

So the functions in $T_g X$ are no longer arbitrary "picky".

One can now sample and hope for good.

In the actual proofs more is used.

Note: If

$$(1-\varepsilon)\|f\|_{L_p(g\,d\mu)} \le \left(\frac{1}{n}\sum_{1}^{n} |f(t_i)|^p\right)^{1/p} \le (1+\varepsilon)\|f\|_{L_p(g\,d\mu)}$$

for all $f \in T_g X$

then

$$\boxed{(1-\varepsilon)\|f\|_{L_p(\mu)} \le \left(\frac{1}{n}\sum_{1}^{n} g(t_i)|f(t_i)|^p\right)^{1/p} \le (1+\varepsilon)\|f\|_{L_p(\mu)}}$$

for all $f \in X$.

Conclusion: For all $0<p<\infty$, $\varepsilon>0$, $k\in\mathbb{N}$, there is an n,

$$\boxed{n \le C(p,\varepsilon)\begin{cases} k^{p/2}\log k, & 2<p \\ k^{p/2}, & p \text{ even} \\ k\log k(\log\log k)^2, & 1<p<2 \\ k\log k, & p=1 \end{cases}}$$

Such that if $X \subseteq L_p(\Omega)$, $\dim X = k$, then there are $t_1, \dots, t_n \in \Omega$ and weights $w_1, \dots, w_n$ s.t.

$$\boxed{(1-\varepsilon)\|f\|_p \le \Big(\sum_i w_i |f(t_i)|^p\Big)^{1/p} \le (1+\varepsilon)\|f\|_p}$$

for all $f\in X$.

Except for the log factors the estimate on n is best possible.

- If in addition, $0 < p < q \leq 2$ and

$$\boxed{X \overset{2}{\hookrightarrow} L_q}$$

("equivalently":

$$\mathbb{E}\left\|\sum_{i=1}^{m} \varepsilon_i x_i\right\| \leq 2\left(\sum_{1}^{m}\|x_i\|^q\right)^{1/q}$$

$\forall\, x_1, \dots, x_m \in X$)

$$\boxed{n \leq C(p,q,\varepsilon)\, k.}$$

Role of Gaussian processes

$X \subseteq L_1(\Omega, \mu)$ X finite dimensional. For our purposes, by simple discretization we can assume $X \subset \ell_1^m$ (with huge m).

If we can find signs ε_i s.t. for all $x = (x_1, \dots, x_m) \in X$, $\|x\| = 1$,

$$\boxed{\left| \sum_{i=1}^{m} \varepsilon_i |x_i| \right| \leq \delta}$$

Then

$$1 - \delta \leq 2 \sum_{\varepsilon_i = 1} |x_i| \leq 1 + \delta$$

and

$$1 - \delta \leq 2 \sum_{\varepsilon_i = -1} |x_i| \leq 1 + \delta$$

for all $x \in X$.

So $X \overset{\frac{1+\delta}{1-\delta}}{\hookrightarrow} \ell_1^{m/2}$.

Now iterate and hope for good.

So we need to evaluate

$$\sup_{\substack{x \in X \\ \|x\|=1}} \left| \sum_{i=1}^{m} \varepsilon_i |x_i| \right| .$$

We really evaluate

$$\mathbb{E} \sup_{\substack{x \in X \\ \|x\| \le 1}} \left| \sum_{1}^{m} \varepsilon_i |x_i| \right| .$$

where ε_i are independent Bernoulli ± 1 variables.

Easy,

$$\mathbb{E} \sup_{\substack{x \in X \\ \|x\| \le 1}} \left| \sum_{1}^{m} \varepsilon_i |x_i| \right| \le \sqrt{\frac{\pi}{2}}\, \mathbb{E} \sup_{\substack{x \in X \\ \|x\| \le 1}} \left| \sum_{1}^{m} g_i |x_i| \right|$$

where g_i are independent standard Gaussians.

$$\left\{ G_x = \sum_{i=1}^{m} g_i |x_i| \right\}_{\substack{x \in X \\ \|x\|=1}}$$

is a Gaussian process.

The estimates for the expectation of the sup of a Gaussian process is a well study subject. (Dudley, Fernique, Talagrand- majorizing measure).

Open problems

1. Are the log factors necessary?

(not for even p !)

2. Explicite embeddings : especially

for $\ell_2^k \hookrightarrow \ell_1^{ck}$

(known $\ell_2^k \hookrightarrow \ell_1^{ck^2}$)

3. Embedding of subSETS.

$S \subset L_1$, $|S| = k$,

S 2-Lipschitz embed into ℓ_1^n.

what is the smallest n ?

known :

$$k^\alpha \leq \text{best } n \leq ck$$

for some absolute $\alpha > 0$.

Reference : Expository paper

- Johnson & Schechtman, Finite dimensional subspaces of L_p, Handbook of the geometry of Banach spaces Vol 1, Elsevier 2001.

For an elementary proof of Lewis' change of density:

- Schechtman & Zvavitch, Embedding subspaces of L_p into ℓ_p^n, $0 < p < 1$, Math Nachr. 227 (2001), 133-142.