Almost-Euclidean subspaces of ℓ_1^N via tensor products: a low-tech approach to randomness reduction

Stanislaw Szarek

Paris 6/Case Western Reserve

MLV, May 19, 2010

Joint work with Piotr Indyk

arxiv:1001.0041,
http://www.case.edu/artsci/math/szarek/

Background

It has been known since 1970's that ℓ_1^N contains nearly Euclidean subspaces of dimension $\Omega(N)$.

Proofs were probabilistic, hence non-constructive.

Consequence: results not-quite-suitable for subsequently discovered applications to high-dimensional nearest neighbor search, error-correcting codes over the reals, compressive sensing and other computational problems.

In this talk

A "low-tech" scheme which, for any $\delta > 0$, allows to exhibit nearly Euclidean $\Omega(N)$-dimensional subspaces of ℓ_1^N while using only N^δ random bits.

This extends and complements (particularly) recent work by Guruswami-Lee-Wigderson.

Characteristic features:

(1) simplicity (we use only tensor products) and

(2) yielding arbitrarily small distortions, or "almost Euclidean" subspaces.

From the 1970's

There exists a subspace $E \subset \mathbb{R}^N$ of dimension $m \geq \alpha N$ and a scaling constant S such that for all $x \in E$

$$D^{-1}\|x\|_2 \leq S\|x\|_1 \leq \|x\|_2$$

Theorem 1 (FLM, 1977) $\forall\, D > 1\ \exists\, \alpha = \alpha(D)$

Theorem 2 (Kashin, 1977) $\forall\, \alpha < 1\ \exists\, D = D(\alpha)$

Explicit S, usually $S = \Theta(N^{-1/2})$
FLM usually written with $D = 1 + \epsilon$
For the Kashin regime $D^{-1}\|x\|_2 \leq N^{-1/2}\|x\|_1 \ \leq\ \|x\|_2$

Highly non-explicit subspaces

- E − a generic element of $G_{N,m}$

- $E = \text{range}\, M$, where M is a random $N \times m$ matrix with i.i.d. entries

- $E = \ker M$, where M is a random $(N-m) \times N$ matrix with i.i.d. entries

- $E \subset \mathbb{R}^{2m}$, the graph of a generic isometry on R^m

In all cases need (at least) $\Omega(N^2)$ random bits

Until a few years ago: explicit subspaces only for $m = O(\sqrt{N})$ (Rudin 1960)

Randomness reduction

Goal: produce, "in polynomial time," αN-dimensional nearly Euclidean subspaces of ℓ_1^N while using $d \ll N^2$ random bits

Holy Grail: $d = O(\log N)$

Indyk (2000) $d = O(N \log^2 N)$, small α, $D = 1 + \varepsilon$

AM (2006), LS (2007) $d = O(N)$, all α, $D = D(\alpha)$

Methods: expander graphs etc.

State of the art before this work

Indyk (2007) no randomness, all $D > 1$, $m = N^{1-o(1)}$

GLR (2008) no randomness, all $\alpha \in (0, 1)$,
D less than any power of N
More precisely, $D = (\log N)^{O(\log \log \log N)}$

GLW (2008) $d = O(N^{\gamma})$, all $\alpha \in (0, 1)$, $D = D(\alpha, \gamma)$
That is, sublinear randomness in the Kashin regime

Methods: expander graphs and much more.

This work

Sublinear randomness in both Kashin and FLM regimes + simplicity

The result (for the FLM regime)

Let $\epsilon, \gamma \in (0,1)$. Given $N \in \mathbb{N}$, assume that we have at our disposal a sequence of random bits of length N^γ. Then, in deterministic polynomial (in N) time, we can generate numbers $M > 0$, $m \geq c(\epsilon,\gamma)N$ and E, an m-dimensional subspace of ℓ_1^N, for which

$$\forall x \in E \quad (1-\epsilon)M\|x\|_2 \leq \|x\|_1 \leq (1+\epsilon)M\|x\|_2$$

with probability greater than 98%.

[slight cheating]

The result (for the Kashin regime)

Let $\gamma, \alpha \in (0,1)$. Given $N \in \mathbb{N}$, assume that we have at our disposal a sequence of random bits of length N^γ. Then, in deterministic polynomial (in N) time, we can generate numbers $M > 0$, $m \geq \alpha N$ and E, an m-dimensional subspace of ℓ_1^N, for which

$$\forall x \in E \quad D(\gamma, \alpha)^{-1} \|x\|_2 \leq N^{-1/2} \|x\|_1 \leq \|x\|_2$$

with probability greater than 98%.

Overview of the proof

Fact: *Tensor products of nearly Euclidean subspaces are nearly Euclidean*

Assume μ_1, μ_2 probability measures and $E_1 \subset L_1(\mu_1)$, $E_2 \subset L_1(\mu_2)$ such that

$$(*) \qquad \forall f \in E_k \quad \lambda_k^{-1}\|x\|_2 \leq \|f\|_1 \leq \|f\|_2$$

Then

$$(\dagger)\ \forall F \in E_1 \otimes E_2 \quad (\lambda_1\lambda_2)^{-1}\|F\|_2 \leq \|F\|_1 \leq \|F\|_2$$

Beckner (1975), Figiel-Johnson (1980)

What is $E_1 \otimes E_2$?

1. $E_1 \otimes E_2 \subset L_2(\mu_1 \otimes \mu_2) \subset L_1(\mu_1 \otimes \mu_2)$

2. If $(\phi_i(s))$, $(\psi_j(t))$ are orthonormal bases in E_1 and E_2, the products $(\phi_i(s)\psi_j(t))$ are an orthonormal basis in $E_1 \otimes E_2$

The proof: apply $(*)$ in one variable, then in the other (modulo some easy extra tricks)

Naive approach

Use $N^\gamma = n^2$ random bits to produce a seed: $\alpha n = \alpha N^{\gamma/2}$-dimensional nearly Euclidean subspace $F \subset \ell_1^n = \ell_1^{N^{\gamma/2}}$ (α close to 1, $n = N^{\gamma/2} \ll N$)

Next, consider the $2/\gamma$-fold tensor product to obtain an $(\alpha N^{\gamma/2})^{2/\gamma} = \alpha^{2/\gamma} N$-dimensional subspace of an ℓ_1 space of dimension $n^{2/\gamma} = (N^{\gamma/2})^{2/\gamma} = N$

The distortion is the original distortion to the power $2/\gamma$.

This actually works in the Kashin regime!

In the Kashin regime

On the αn-dimensional subspace $F \subset \ell_1^n$ we have

$$D^{-1}\|x\|_2 \le n^{-1/2}\|x\|_1 \;\le\; \|x\|_2$$

or, with the normalized counting measure

$$D^{-1}\|f\|_2 \le \|f\|_1 \;\le\; \|f\|_2$$

and after tensorizing

$$D^{-2/\gamma}\|F\|_2 \le \|F\|_1 \le \|F\|_2$$

What about the FLM regime?

Milman's version of Dvoretzky's theorem

Consider the n-dimensional Euclidean space (over $\mathbb{R}$ or $\mathbb{C}$) endowed with the Euclidean norm $|\cdot|$ and some other norm $\|\cdot\|$ such that, for some $b > 0$, $\|\cdot\| \le b|\cdot|$. Denote $M = \mathbb{E}\|X\|$, where X is a random variable uniformly distributed on the unit Euclidean sphere. Let $\varepsilon > 0$ and let $m \le c\varepsilon^2 (M/b)^2 n$, where $c > 0$ is an appropriate (computable) universal constant. Then, for most m-dimensional subspaces F we have

$$\forall x \in F \quad (1-\varepsilon)M|x| \le \|x\| \le (1+\varepsilon)M|x|.$$

The case of ℓ_1^n (with normalized measure)

$$\forall x \in F \quad (1-\varepsilon)M\|x\|_2 \le \|x\|_1 \le (1+\varepsilon)M\|x\|_2$$

with $m = \dim F = \Omega(n)$

The problem: $M \sim \sqrt{2/\pi} \approx 0.8 < 1$

The lower estimate survives tensoring, but the upper does not, so after tensoring we can only have the trivial upper estimate. The gap between the lower and the upper estimates is $(\sqrt{2/\pi}(1-\varepsilon))^{-2/\gamma}$, which can not be close to 1 and is large if γ small

The remedy: replace ℓ_1^n with $\ell_1^{n/B}(\ell_2^B)$

Then, with the normalized counting measure, $\|x\| \leq \|x\|_2$ for all $x \in \mathbb{R}^n$ (the trivial upper bound)

On the other hand, the mean of the norm on the sphere is $\geq \sqrt{1 - \frac{1}{B}}$, so Dvoretzky's theorem yields

$$(1-\varepsilon)\sqrt{1-\frac{1}{B}}\,\|x\|_2 \leq \|x\|$$

on a large subspace.

Enough to choose $B \approx 1/\varepsilon$, then

$$(1-\varepsilon)\sqrt{1-\frac{1}{B}} = (1-\varepsilon)^{3/2}$$

After tensorizing, the gap between the lower and the upper estimates is $(1-\varepsilon)^{3/\gamma}$

It remains to show that a tensorization trick $(*) \Rightarrow (\dagger)$ works also for Hilbert-space-valued functions.

This is less immediate than in the scalar case, but also elementary. Results of this type go back to Marcinkiewicz-Zygmund (1939)