

High dimensional convex bodies: phenomena, intuitions and results

Nicole Tomczak-Jaegermann

Fall 2006

– CRM-Fields-PIMS Prize, Fall 2006

Plan of the talk

- further introductory remarks and notation (for non-specialists)
- linear-metric structure and diversity of fin.-dim. normed spaces
- applications:
 - approximation problems in compressed sensing
 - m -neighborly polytopes
- metric entropy and related duality issues

Introductory remarks and notation

Typical setting and objective:

unspecified finite but usually high dimension

study of quantitative invariants, up to *universal constants*

$$cf \leq \text{invariant} \leq Cf$$

where f is an explicit function of the parameters involved
(such as the dimension)

Leads to *isomorphic* rather than isometric properties

Geometric vs. functional-analytic objects

- normed space $X \leftrightarrow$ its unit ball B_X
- convex body $K \subset \mathbb{R}^n$ with $0 \in \text{Int}K \leftrightarrow$ its gauge $\|\cdot\|_K$

$$\text{i.e., } \|x\|_K := \inf\{t > 0 : x \in tK\}$$

In particular, if K is centrally symmetric then

- $K \leftrightarrow$ the normed space $(\mathbb{R}^n, \|\cdot\|_K)$

Fundamental concept : Banach-Mazur distance

$$d(K, B) := \inf\{\lambda > 0 : \exists u \in GL(n) \quad K \subset u(B) \subset \lambda K\}$$

or, in terms of normed spaces,

$$d(X, Y) := \inf\{\|u\| \cdot \|u^{-1}\| : u \in L(X, Y), \text{ isomorphism}\}$$

Linear-metric structure: Subspaces/Quotients

Study: family of subspaces (dually, of quotients) of a given Banach space.

The aim may be two folded:

- to detect some possible regularities in subspaces which might have not existed in the whole space, or oppositely,
- to identify some “irremovable” structures present in *every* subspace (or quotient) of sufficiently large dimension

Dvoretzky's Th. 1961 (strengthened by V. Milman, 1970):

Every normed space X of (large) dimension n has an “almost” Euclidean subspace of dimension $k \geq c \log n$ ($c > 0$ depends on the degree of appr.)

Based on *concentration of measure on sphere* phenomenon.

k optimal, in general: If $E \subset X = \ell_\infty^n$, $d(E, \ell_2^k) \leq 2$ then $k \leq C \log n$.

For large class of spaces, k can be proportional to n , case of e.g. $X = \ell_1^n$.

Large Subspaces of Quotients

Milman (1983): For any $\theta \in (0, 1)$, every n -dim. normed space X admits a subspace of a quotient E , “nearly” Euclidean and of dimension $k \geq \theta n$.

$\exists X \rightarrow X_0$ quotient $\exists E \subset X_0$ subspace s.t. $k \geq \theta n$ and $d(E, \ell_2^k) \leq f(\theta)$.

A byproduct: every n -dimensional normed space admits a “proportional dimensional” quotient of well-bounded *volume ratio*.

A considerable regularity in a global invariant achieved by passing to a quotient of prop. dim.

Milman [ICM 1986]: Does every n -dimensional normed space admit a quotient of dimension $\geq n/2$ whose cotype 2 constant is bounded by a universal numerical constant?

Cotype 2 and Cotype 2 constants

Cotype 2 constant of a space X is the smallest C (if it exists) such that, for every finite sequence (x_j) in X one has

$$\text{Ave}_{\pm} \left\| \sum_j \pm x_j \right\|^2 \geq C^{-2} \sum_j \|x_j\|^2$$

(relaxed parallelogram inequality). If such a constant exists, the space is said to have cotype 2.

Examples: classical and non-commutative L_p -spaces, Schatten classes S_p , for $p \in [1, 2]$.

Saturating spaces I

$X, \dim X = n; \quad V, \dim V = k; \quad k \ll n$

X is *saturated* with V , or V *saturates* X ,

if every subspace (resp. quotient) \tilde{X} of X of sufficiently large dimension (depend. on k) has a subspace (resp. quotient) well-isomorph. to V

By Dvoretzky's theorem, *every* normed space X is saturated with the Euclidean space, i.e., we can take $V = \ell_2^k$. and “large” means $m \geq e^{Ck}$.

Are there any other spaces V that can saturate *some* normed spaces?

Saturating spaces II

S. Szarek/T. [2004]: *Any* space V can saturate. Sample result:
Let n and m_0 with $\sqrt{n} \log n \leq m_0 \leq n$. Then, for every V satisfying

$$k := \dim V \leq c m_0 / \sqrt{n}$$

there exists an n -dimensional normed space X such that every quotient \tilde{X} of X with $\dim \tilde{X} \geq m_0$ contains a 1-complemented subspace isometric to V . (Here $c > 0$ is a universal constant.)

Particular case: Given V , if $k \leq c\sqrt{n}$ then there is X such that every $n/2$ -dim. quotient of X contains a 1-complemented isometric copy of V .

Here $m_0 \sim n/2$ and $k \sim \sqrt{n}$ is allowed.

We may decrease m_0 a little, paying the price of smaller k allowed.

Relation to Milman's problem

Mysterious “threshold” \sqrt{n} :

upper bound for k and if $m_0 \sim n/2$ then $k \sim \sqrt{n}$ is allowed.

lower bound for m_0 ($\geq \sqrt{n} \log n$)

Setting $V = \ell_\infty^k$ implies that every quotient \tilde{X} of X with $\dim \tilde{X} \geq n/2$ contains ℓ_∞^k ($k \sim \sqrt{n}$) and so its cotype 2 constant is $\sqrt{k} \sim n^{1/4}$

Complementability of copies of V imply that “every quotient \tilde{X} of X ” can be replaced by “every subspace \tilde{X} of X ,” thus implying the “subspace” variant of the Theorem.

Thus, in general, passing to large subspaces *or* large quotients can not erase k -dimensional features of a space if k is below certain threshold value.

Reconstruction from random linear measurements

Problem: given $T \subset \mathbb{R}^n$, approximate any $v \in T$ using $k \ll n$ *random linear measurements*.

Given $X_1, \dots, X_k \in \mathbb{R}^n$ i.i.d. random vectors, $(\langle X_j, v \rangle)_{j=1}^k$ and T , find $t \in T$, such that $\langle X_j, v \rangle = \langle X_j, t \rangle$ and $|t - v| \leq \varepsilon(k)$ for $\varepsilon(k)$ as small as possible.

Γ has X_1, \dots, X_k as rows

S. Mendelson/A. Pajor/T. (2005, 06)

Our initial motivation: results by E. Candes and T. Tao ('05)
they considered $T =$ the unit ball in ℓ_1^n or *weak*- ℓ_p^n ($0 < p < 1$)
uniform proof in terms of *spectral* properties of Γ .

Γ determined by the Gaussian or Bernoulli or Fourier ensemble.

Linear approximate reconstruction

Given $X_1, \dots, X_k \in \mathbb{R}^n$ i.i.d. random vectors, and $(\langle X_j, v \rangle)_{j=1}^k$,
find $t \in T$, such that $\langle X_j, v \rangle = \langle X_j, t \rangle$
and $|t - v| \leq \varepsilon(k)$ for $\varepsilon(k)$ as small as possible.

Γ has X_1, \dots, X_k as rows, then $t - v \in \ker \Gamma \cap aT$ if T quasi-convex;
thus $\varepsilon(k) = \text{diam}(\ker \Gamma \cap aT)$ works.

Question: Describe $r(T)$, depending on T , such that

$$\text{diam}(\ker \Gamma \cap T) < r(T)$$

with probability close to 1.

For Γ Gaussian: techniques developed in AGA in mid-80's,
using concentration (Milman, Pajor/T., Milman/Pisier,)

Back to concentration phenomena

$T \subset \mathbb{R}^n$ sym. (quasi-)convex; for $\rho > 0$, let $T_\rho = \rho T \cap S^{n-1}$.

$\forall F \subset \mathbb{R}^n, \rho > 0 \quad \text{diam}(F \cap T) < 1/\rho \quad \text{equivalent} \quad F \cap T_\rho = \emptyset$

For $F = \ker \Gamma$, stronger: $(*) \quad |\Gamma x| \sim \text{constant for } x \in T_\rho$

When ρ increases, T_ρ become richer and the condition eventually fails.
Complicated formula for critical ρ , right measure of complexity of T is

$$\ell_*(T) := \mathbb{E} \sup_{t \in T} \left| \sum_{i=1}^n g_i t_i \right| \quad \text{for } T \subset \mathbb{R}^n; \quad g_i \text{'s are i.i.d. } N(0, 1).$$

MPT: for *subgaussian* measurements. Prime examples: coordinates of X_i are *Gaussian* or *Bernoulli* (or any bounded) i.i.d. random variables
all examples of T studied earlier follow from our formula

Exact reconstruction

Problem from signal processing: reconstruct *exactly* sparse vector $z \in \mathbb{R}^n$ by performing $k \ll n$ random linear measurements

Sparse: supported on at most r coordinates

We want k small, but how large does it have to be?

Surprise: possible with $k \geq Cr \log(n/r)$

For Gaussian results: Candes/Tao and M. Rudelson/R. Vershynin.

MPT: results for *subgaussian* measurements

Geometry of random polytopes

A polytope is called m -neighborly if any set of less than m of its vertices is the vertex set of a face.

Random $\{-1, 1\}$ polytopes: $K_n := \text{conv} \{v_1, \dots, v_n\} \subset \mathbb{R}^k$ ($n > k$)
where $v_i \in \mathbb{R}^k$ has coordinates i.i.d. Bernoulli random variables.

Surprise: with probability close to 1, a random $\{-1, 1\}$ -polytope K_n in \mathbb{R}^k is m -neighborly for a relatively large m ,

$$m \leq \frac{ck}{\log(Cn/k)}.$$

Metric entropy

K, B subsets of a vector space, the *covering number* of K by B

$$N(K, B) = \min N \text{ s.t. } \exists x_1, \dots, x_N \quad K \subset \bigcup (x_i + B)$$

The *packing number* $M(K, B) = \max M$ s.t.

$$\exists y_1, \dots, y_M \in K \quad (y_i + B) \cap (y_j + B) = \emptyset \text{ for } i \neq j.$$

Closely related, if B is centrally symmetric:

$$N(K, 2B) \leq M(K, B) \leq N(K, B).$$

If B is a ball in a Banach space X and $K \subset X$, it reduces to smallest ε -nets or the largest ε -separated (or 2ε -separated) subsets of K .

Duality of metric entropy

If $u : Y \rightarrow X$ bounded linear operator (X, Y Banach spaces)
the sequence of *entropy numbers* of u is defined by

$$e_k(u) = \inf \{ \varepsilon : N(u(B_Y), \varepsilon B_X) \leq 2^{k-1} \} \quad \text{for } k \geq 1 \quad (e_k(u)) \downarrow$$

$\lim e_k(u) = 0$ iff u is compact iff u^* is compact
the limiting behaviour of $\{e_k(u)\}$ and $\{e_k(u^*)\}$ is the same.

Duality conjecture [Pietsch, 1972]:

Is it true that for some absolute constants $a, b \geq 1$

$$a^{-1}e_{bk}(u) \leq e_k(u^*) \leq ae_{k/b}(u) ?$$

For symm. convex bodies $K, B \subset \mathbb{R}^n$: do we have

$$b^{-1} \log N(B^0, aK^0) \leq \log N(K, B) \leq b \log N(B^0, a^{-1}K^0),$$

uniformly in K, B and n ? (K^0, B^0 are the polar bodies)

$$K^0 := \{x : |\langle x, y \rangle| \leq 1 \text{ for all } y \in K\}$$

Duality of metric entropy, results

S. Artstein/Milman/Szarek [2004]: The duality holds when one of the spaces X, Y is a Hilbert space; in geometric terms, when either K or B is an ellipsoid.

Artstein/Milman/Szarek/T. [2004]: More generally, the same is true if one of the spaces is K -convex.

K -convexity means the absence of large subspaces resembling f.d. ℓ_1 -spaces equivalently, nontrivial type $p > 1$; also, by deep theorem by Pisier, equiv. boundedness of the Rademacher (or Gaussian) projection on $L_2(X)$.

Examples: all (classical and non-commutative) spaces L_p ($1 < p < \infty$), all uniformly convex/uniformly smooth spaces.

Quantified by the K -convexity constant.

Convexified packing I

Let $K, B \subset \mathbb{R}^n$ sym. convex bodies; the *convexified packing number* $\hat{M}(K, B)$ is the maximal length M of a sequence x_1, \dots, x_M in K ,
 $(x_j + B) \cap \text{conv} \bigcup_{i < j} (x_i + B) = \emptyset$, for $j = 2, \dots, M$.

Unlike for usual packing or covering, the order is important here.

Convexified packing II

- For this modified notion, the duality holds: $\hat{M}(K, B) \leq \hat{M}(B^0, K^0/2)^2$.
- If K or B is K -convex and $K \subset 4B$ then the packing numbers $M(K, B)$ and $\hat{M}(K, B)$ are equivalent.

These ideas were first used in Bourgain/Pajor/Szarek/T. (1987).

The first fact is a direct application of the Hahn-Banach separation theorem.

The second is simple for the Hilbert space case; for uniformly convex/smooth spaces it follows from an elementary convexity argument.

In the K -convex case it is not elementary and follows from so-called Maurey's Lemma

- For a given B , if the duality conjecture holds for all $K \subset \mathbb{R}^n$ s.t. $K \subset 4B$, then it holds for all $K \subset \mathbb{R}^n$.

This was proved by Artstein/Milman/Szarek.