Size of line intersections in high dimensional L_p-balls and product measures

Bo'az Klartag
Weizmann Institute of Science

Joint work with Dor Elboim

Calculus of Variations in Probability and Geometry,
Institute for Pure and Applied Math. (IPAM), February 2022.

Segments in convex bodies in high dimensions

What's the first thing you note in high dim. Euclidean geometry? Maybe:

The unit cube $[0,1]^n \subseteq \mathbb{R}^n$ has volume one, yet it contains long segments "in diagonal", of length $\sqrt{n}$.

- The lengthscale of $\sqrt{n}$ is typical in high dim., it is roughly the radius of the Euclidean ball of volume one.

Theorem (Classical isodiametric inequality)

Any convex $K \subseteq \mathbb{R}^n$ of volume one, contains a segment of length at least

$$\left(\sqrt{\frac{2}{\pi e}} + o(1) \right) \cdot \sqrt{n},$$

the minimum is attained for the Euclidean ball.

Removing a little bit of mass

Question

Can one remove 1% of the mass of the convex body, and avoid these long segments?

Trivial answer: remove points with a rational coordinate or so.

Better formulation

Let $K \subseteq \mathbb{R}^n$ be convex, volume one. Does there exist a subset $A \subseteq K$ with $|A| = 1/2$ such that

$$|A \cap \ell| < C$$

for all lines in $\mathbb{R}^n$? Here $C > 0$ means a universal constant.

- The constant $1/2$ can be replaced by any $c \in (0, 1)$.
- If K is the Euclidean ball, the answer is YES.
  We select $A = K \setminus (1 - \frac{c}{n})K$, a thin spherical sphell.

The curvature of the Euclidean sphere

The case of the centered Euclidean ball $K \subseteq \mathbb{R}^n$ of volume one

Set $A = K \setminus (1 - \frac{1}{n})K$, so that $|A| = 1 - (1 - \frac{1}{n})^n \approx 1 - 1/e$. Fix $x \in \mathbb{R}^n$ and a unit vector θ. We need all $t \in \mathbb{R}$ such that

$$\left(1 - \frac{1}{n}\right)^2 r_n^2 \leq |x + t\theta|^2 \leq r_n^2$$

with $r_n = Vol_n(B^n)^{-1/n} \sim \sqrt{n}$. This has a bounded measure.

- For this K and for $A \subseteq K$ with $|A| = 1/2$, there is always a line ℓ with $|A \cap \ell| > c$, by Fubini's theorem.
- Slightly less trivial: The same is true for any convex body $K \subseteq \mathbb{R}^n$ of volume one.
- Consider $B_p^n = \left\{x \in \mathbb{R}^n \,;\, \|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} \leq \kappa_{p,n}\right\}$ for $1 \leq p \leq \infty$, where $\kappa_{p,n} = \frac{\Gamma(1+n/p)^{1/n}}{2\Gamma(1+1/p)} = \Theta(n^{1/p})$.

The quantity $L(\mu, a)$

Definition

For a probability measure μ on $\mathbb{R}^n$ and $0 < a < 1$ define

$$L(\mu, a) := \inf_{\mu(A) \geq a} \sup_{\ell \text{ line}} \left| \ell \cap A \right|,$$

where the inf runs over all Borel sets $A \subseteq \mathbb{R}^n$ with $\mu(A) = a$.

- For $K \subseteq \mathbb{R}^n$ of volume one, abbreviate $L(K, a) = L(\lambda|_K, a)$, where λ is the Lebesgue measure.

Theorem 1

Let μ be the uniform measure on the unit cube, or the regular simplex, or the standard Gaussian measure γ_n in $\mathbb{R}^n$. Then,

$$L(\mu, 1/2) = \Theta(n^{1/4}).$$

- Here, the extremal set A is a thin spherical shell.

What's going on with these L_p-balls?

- Let X be a random vector, $X \sim Unif(B_p^n)$. For $p \neq 2$,

$$Var(\|X\|_2) = \Theta_p(1)$$

while for $p = 2$ we have $Var(\|X\|_2) = \Theta(1/n)$.
The thin spherical shell does not explain everything:

Theorem 2

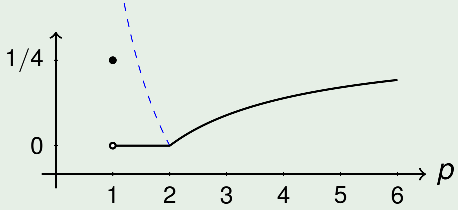
Fix $1 \leq p \leq \infty$. Then as $n \to \infty$,

$$L(B_p^n, 1/2) = \begin{cases} \Theta\left(n^{1/4}\right) & p = 1, \infty \\ \Theta_p\left((\log n)^{\frac{2-p}{2p}}\right) & 1 < p \leq 2 \\ \Theta_p\left(n^{\frac{p-2}{4p+2}}\right) & 2 \leq p < \infty \end{cases}$$

- The constant $1/2$ can be replaced by any fixed $a \in (0, 1)$.

Phase transitions of different types

The exponent of n as a function of p (and the exponent of $\log n$ in blue)



- Phase transitions of different types... But do we have universality in some class? Isn't $n^{1/4}$ somewhat universal?

Definition

A product measure $\frac{d\mu}{dx} = \prod_{i=1}^n \rho_i(x_i)$ in $\mathbb{R}^n$ is "admissible" if

(i) Each ρ_i is smooth and positive, at least in the interval $(-1/2, 1/2)$, with uniform estimates.

(ii) Each ρ_i has a sub-Gaussian tail, with uniform estimates.

Universality for product measures

Theorem 3

Let $n \geq 1, 0 < a < 1$ and let μ be an admissible product probability measure in $\mathbb{R}^n$. Then,

$$L(\mu, a) = \begin{cases} \Theta\left(a \cdot n^{1/4}\right) & e^{-n} \leq a \leq 1/2 \\ \Theta\left(n^{1/4} \cdot |\log(1-a)|^{1/4}\right) & 1/2 \leq a \leq 1 - e^{-n} \end{cases}$$

- Using Chen's recent bound for the thin shell constant, relying on Eldan's stochastic localization:

Proposition 4

Let $K \subseteq \mathbb{R}^n$ be a convex body of volume one in isotropic position (i.e., scalar covariance matrix). Then, uniformly over K,

$$L(K, 1/2) = O\left(n^{1/4+o(1)}\right).$$

Can $L(\mu, 1/2)$ be much larger than $n^{1/4}$?

- Outside the realm of convexity and product measures:

Radially-symmetric example for μ in $\mathbb{R}^n$ with $L(\mu, 1/2) = \Theta(\sqrt{n})$

Take μ to be the law of the random vector

$$X + UY$$

where X, Y, U are independent, with X, Y standard Gaussians in $\mathbb{R}^n$ and U uniform in $[0, 1]$.

Proof of the lower bound: suppose that $\mu(A) \geq 1/2$. Then,

$$\mathbb{P}(X + UY \in A, |Y| \geq \sqrt{n}/2) \geq \mathbb{P}(X + UY \in A) - Ce^{-cn} \geq 1/3.$$

Hence there exist $x, y \in \mathbb{R}^n$ s.t. for the line $\ell = x + \mathbb{R}y$,

$$|\ell \cap A| \geq |y| \cdot \mathbb{P}(x + Uy \in A) \geq \sqrt{n}/6. \quad \square$$

What mathematics does this story remind us of?

A sample of mathematical directions, maybe of a similar spirit:

1. Density **Hales-Jewett Theorem.** Any subset $A \subseteq \{1, \ldots, D\}^n$ of positive density contains a combinatorial line (a row, a column or a "diagonal"). Here D is fixed, $n \to \infty$. This gives "length one", not $n^{1/4}$.
2. **The lower-dimensional Busemann-Petty problem.** Let $K, T \subseteq \mathbb{R}^n$ be convex, volume one, centrally-symmetric. For $2 \leq \ell \leq n-1$, is there an ℓ-dimensional subspace E with
$$|K \cap E| \geq |T \cap E|?$$
3. The **Radon transform** $R(1_A)$ with $A \subseteq [0,1]^n$ of positive density, satisfies $\|R(1_A)\|_\infty \geq cn^{1/4}$ ("reverse Kakeya"?)
4. **Szemerédi-Trotter Theorem.** Given n points and m lines in the plane, there are at least $(nm)^{2/3}$ incidences.

Proofs – Needle decomposition

Our idea for the lower bound: need to approximate μ by a mixture of uniform measures on **long segments.**

Consider the case where $\mu = \gamma_n$ is Gaussian.

A computation (hint: use relative entropy and Pinsker's inequality)

Let $X, Y \sim \gamma_n$ be independent. Then for $r < cn^{-1/4}$,

$$d_{TV}(X, X + rY) \leq 1/10.$$

Hence $d_{TV}(X, X + rUY) \leq 1/10$ for $U \sim Unif([0, 1]), r = cn^{-1/4}$.

- A family of uniform measures on intervals

$$[X, X + rY]$$

  of length $\sim n^{1/4}$ approximates the Gaussian measure well.
- There is a related effect for the uniform measure on the cube.

The exponent in the case $p > 2$

- Given $X \sim Unif(B_p^n)$. Need $d_{TV}(Y, X) \leq 1/4$ with

$$Y_i = X_i + rU\delta_i\psi(X_i)$$

where $\delta \in \{\pm 1\}$ i.i.d symmetric Bernoulli, $U \sim Unif([0, 1])$.

Only small coordinates move, since $(t^p)''$ is small near zero. We take φ to be a bump function in $[1/2, 2]$, $\beta \geq \alpha$ and

$$Y_i = X_i \pm n^{-\beta}\varphi(n^\alpha X_i)$$

Requirements

1. Not too many coordinates enter the interval $[-n^{-\alpha}, n^{-\alpha}]$. This gives the constraint $1 + \alpha - 2\beta \leq (1 - \alpha)/2$.
2. The change in $\|X\|_p^p$ is at most a constant. This gives the constraint $\alpha(p - 1) + 2\beta \geq 1$.

Extremal set is $B_p^n \cap \{|\sum_i h(x_i) - E| \leq C\}$ for certain convex h.

The behaviour in the case $1 < p < 2$

Differences with the previous case:

1. Here $(t^p)''$ is small for large t. So better to move large coordinates.
2. The extremal set, without large line intersections, is
$$A = B_p^n \cap \{\kappa_{p,n}^p - C \leq \|x\|_p^p \leq \kappa_{p,n}^p\} \cap \{\forall i, |x_i| \leq C(\log n)^{1/p}\}.$$
3. The contribution from $(t^p)'$ should be "cancelled". Here the terms are too big, so we move the coordinates in pairs, and in opposite direction, to improve cancellation.
4. We randomly move coordinate X_i with $|X_i| \sim R = (\log n)^{1/p}$ to distance δ so that
$$\#(points) \cdot R^{p-2} \cdot \delta^2 \leq C.$$
The segment length is about $\sqrt{\#(points)} \cdot \delta$.

The end

Thank you!