Asymptotic Geometric Analysis: An Overview

V.D. Milman

Tel Aviv University

Rome, 2007

I. The Subject of Asymptotic Geometric Analysis

II. Geometrization of Probability

III. Concentration Phenomenon: Isomorphic Form of Isoperimetric Problems

I. The Subject of Asymptotic Geometric Analysis

Some History

The framework of the subject we will discuss involves very high dimensional spaces (normed spaces, convex bodies) and accompanying asymptotic (by increasing dimension) phenomena.

The starting point was open problems of Geometric FA (from the 60s and 70s). This development naturally led to the Asymptotic Theory of Finite Dim. spaces (in 80s and 90s).

During this period, the problems and methods of Classical Convexity were absorbed by Asymptotic Theory (including geometric inequalites and many geometric, i.e. "isometric", as opposed to "isomorphic" problems).

As an outcome, we derived a new theory:

Asymptotic Geometric Analysis.

One of the most important points of already the first stage is a change in intuition about the behavior of high-dimensional spaces. Instead of the diversity expected in high dimensions and chaotic behavior, we observe a unified behavior with very little diversity.

Unusual intuition of high dimension:

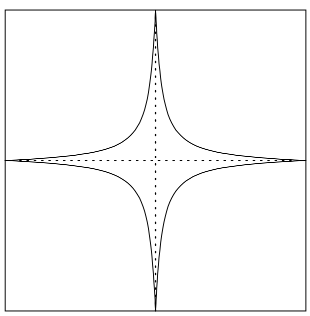
First simple examples

$$\begin{aligned} \mathrm{Vol}\, D_n \quad : \quad n &= 1 \qquad \mathrm{Vol} = 2 \\ &= 2 \qquad\qquad\ \ = \pi \\ &= 3 \qquad\qquad\ \ = \tfrac{4}{3}\pi \end{aligned}$$

$$\mathrm{Vol}\, D_6 > \mathrm{Vol}\, D_5 > \mathrm{Vol}\, D_4 > \mathrm{Vol}\, D_3$$

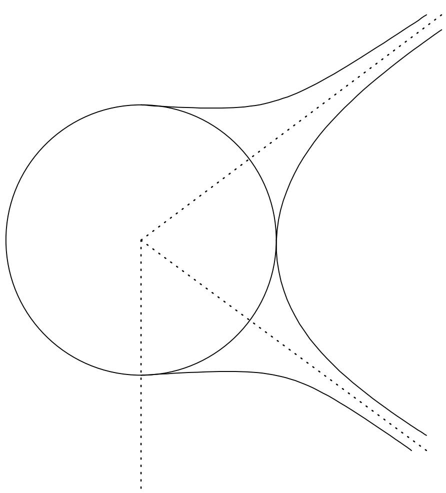
But $\mathrm{Vol}\, D_n = \left(\frac{c_n}{\sqrt{n}}\right)^n$ where $c_n \to \sqrt{2\pi e}$.

So, it is "very difficult" to find points of D_n inside $C^n = [-1, 1]^n$ (i.e. probability of random point of C^n to be in D_n is exponentially small).

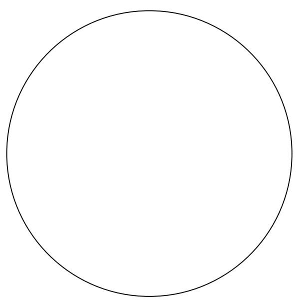


$(1, 1, \ldots, 1)\,, \quad \|(1, \ldots, 1)\| = \sqrt{n}$

$D_n \subset C_n$, emphasizing the geometry of C_n.



or again $D_n \subset C_n$ but emphasizing geometry of D_n



or even more reflecting that C_n has 2^n such direction–mustaches

So n-dim. cube looks "similar" to $\sqrt{n}D_n$ from the volume distribution point of view.

Some precise facts, which sounded very surprizing at the end of the '70s became natural and intuitive from this (correct) picture.

Now a precise fact (Figiel-Lindenstrauss-Milman, '76/'77):

Fix small $\varepsilon > 0$; a "typical" projection $P_E C^n$ on a subspace E, $\dim E = [\varepsilon n] := k$,
is $\sim \sqrt{\varepsilon}$-isometric to a euclidean ball $R_n D_k$ for radius $R_n \sim \sqrt{\frac{2}{\pi}}\sqrt{n}$:

$$(1 - \sqrt{\varepsilon}) R_n D_k \subset P_E C^n \subset (1 + \sqrt{\varepsilon}) R_n D_k$$

[if $aT \subset K \subset bT$, we say that distance $d(K,T) \leq b/a$].

Also, even almost full dimensional projections look euclidean:

Kashin ('77). For any $\frac{1}{2} < \lambda < 1$, $\exists\, C(\lambda)$, s.t. for a typical projection P_E, for $k = \dim E = [\lambda n]$

$$\frac{1}{C(\lambda)} R_n D_k \subset P_E C^n \subset C(\lambda) R_n D_k \ .$$

Unified behaviour

First example: (from: M., 71; M.-Schechtman, 97; M., 2000)

$K \subset (\mathbb{R}^n, |\cdot|)$ – a subset; D – the unit eucl. ball

$D(E) = D \cap E$, for a subspace E.

Let $d(K)$ – diameter of K, $P_E K$ – orthoproj. of K onto subspace $E \hookrightarrow \mathbb{R}^n$.

Let $D_\ell(K) := \mathbb{E}(d(P_E K) \mid \dim E = \ell)$
and mean width

$$D_1(K) := w(K) = \int_{S^{n-1}} w(K,u) d\sigma(u)$$

$w(K,u)$ – width in the direction $u \in S^{n-1}$:

$$w(K,u) = \sup\Big\{(u,x) \mid x \in K\Big\} - \inf\Big\{(u,x) \mid x \in K\Big\}$$

Define

$$k^* = n\left(\frac{w(K)}{d(K)}\right)^2$$

(Note that $w(D)/d(K) \geq 1/\sqrt{n}$.)

Then: $\exists c > 0$ and C s.t. $\forall n \quad \forall K \subset \mathbb{R}^n$

for $k^* \leq \ell \leq n$,

$$c\sqrt{\frac{\ell}{n}}\, d(K) \leq D_\ell(K) \leq C\sqrt{\frac{\ell}{n}}\, d(K)$$

and

$$cw(K) \leq D_\ell(K) \leq Cw(K)$$

for $1 \leq \ell \leq k^*$ (stabilization). Actually, the same is true for a "typical" projection $P_E K$.

k^* – critical value

We observe a phase transition type behaviour (it is a typical behaviour).

Additional information:

The **only** reason for stabilization is that the "typical" projection $P_E K$ for $\dim E \lesssim \varepsilon^2 k^*$ is an ε-net of a euclidean ball (of radius $w(K)/2$).

If K convex, $K = -K$, $\|x\|_K$ – K-norm, then the "typical" projection is an almost euclidean ball

$$P_E K \sim \frac{w(K)}{2} D(E) \ .$$

Next example; Corresponding **global** problem. (from: Bourgain-Lindenstrauss-M., 87; M.-Schechtman, 97; M., 2000).

Problem: Approximate euclidean ball rD by averaging rotations of K: $K_t = \frac{1}{t}\sum_{i=1}^{t} u_i K$, $u_i \in O(n)$.

We will discuss the problem in the language of Funct. Analysis instead of the language of Geometry.

Let $X = (\mathbb{R}^n, \|\cdot\|, |\cdot|)$, denote

$$b := \|Id : \ell_2^n \to X\|, \text{ and } M = \int_{S^{n-1}} \|x\| d\sigma(x) \ .$$

Consider a new averaging norm

$$|||x|||_t = \frac{1}{t}\sum_1^t \|u_i x\|, \quad u_i \in O(n)$$

and the space $X_t = (\mathbb{R}^n, |||\cdot|||_t, |\cdot|)$.

Then with a very high probability (by selection u_i)

$$b_t := \|Id : \ell_2^n \longrightarrow X_t\| \sim \frac{1}{\sqrt{t}} \|Id : \ell_2^n \to X\|$$

for

$$t < (b/M)^2 = (\sup_{|x|=1} \|x\| / \mathbb{E}_{x \in S^{n-1}} \|x\|)^2 := t_0$$

and

$$|||x|||_t \sim M \cdot |x| \quad \text{for} \quad t \gtrsim t_0 \ .$$

Again phase transition,

stabilization after critical value.

Reason for stabilization:

averaging norm becomes

euclidean.

Note: In the "global" process: "the best" possibility is $\approx$ the same as a "typical" selection.

Other very general facts

Dvoretzky-type theorems

We already noted that typical projection $P_E K$ of centrally sym. convex body K on subspace E of dim. k^* is close to a euclidean ball D. Let us put some facts in precise form (and in the dual form, i.e. for "typical" sections).

Theorem (Milman, 1971).

Let $X = (\mathbb{R}^n, \|\cdot\|, |\cdot|)$, $\|\cdot\|$ *– some norm,* $|\cdot|$ *– euclidean norm,* $S^{n-1} = \{x \mid |x| = 1\}$, $b = \max_x \|x\|/|x|$ *and* M $(= \mathbb{E}_{S^{n-1}}(\|x\|))$ *as before.*

Let $k(X) := k = [cn(M/b)^2]$, c *– universal constant. Then,*

$$\text{Prob}\Big\{E \in G_{n,k} \text{ satisfies } \tfrac{1}{2}M|x| \le \|x\| \le 2M|x| \text{ for } \forall\, x \in E\Big\} \ge 1 - e^{-k}.$$

(*i.e. with high probability* k*-dim. subspaces are* $\sim$ *euclidean*).

Remarks. 1. Of course, we may write $1+\epsilon$, $\epsilon > 0$, instead of 2, by reducing k.

2. Always $b \leq \sqrt{2n}M$, and there is $|\cdot|$ s.t. $b \leq c\sqrt{\frac{n}{\log n}}M$ which would imply $k \gtrsim \log n$.

3. However, k (or k^* for random projections) may be very small ($\sim \log n$, say), and it is not useful for applications. At the same time, for some universal constant $c > 0$

$$k(X) \cdot k^*(X) \geq cn$$

(Figiel, Lindenstrauss, Milman – 1977).

Using two "dual" operations consequently changes the picture.

The first example of this nature:

"Quotient of a Subspace Theorem" (Milman, 1985): *Let* $1/2 \le \lambda < 1$ *and* $X - n$*-dim. normed space. Then* $\exists$ *subspaces* $F \hookrightarrow E \hookrightarrow X$ *s.t. for* $Y = E/F$

$$k = \dim Y \ge \lambda n\,,\ d(Y, \ell_2^k) \le c\frac{|\log(1-\lambda)|}{1-\lambda}$$

This is already a structural fact. One may start to feel how we can approach dealing with an **arbitrary** convex body and normed space.

Concept of "Concentration"

Rough Idea (which we will develop later):

A good (say, 1-Lip) function on high-dimensional objects cannot be distinguished from a constant in "polynomial time".

Example: $S^{n+1} \subset \mathbb{R}^{n+2}$; $f(x) \in C(S^{n+1})$ and 1-Lip. Select random points $\{x_i\}_1^N \subset S^{n+1}$. Fix $\varepsilon > 0$. Then

$$\text{Prob.}\{|f(x_i) - f(x_j)| < \varepsilon,\ \forall i, j\} > 1 - Ne^{-\varepsilon^2 n/8}\ .$$

So, say, for $N \sim n^t$ we don't observe osc. of $f(x)$ above level $\varepsilon \gtrsim C\sqrt{\frac{\log n}{n}}$.

For spaces (X_n, ρ_n, μ_n) with metric (to define a "good" class) and probability measures μ_n, of increasing "dim" n, this is a typical observation, called "concentration phenomenon". Technically (and ideologically) this phenomenon is responsible for changing our intuition; it compensates exponential increase in covering (entropy) and "suppresses" an expected diversity within high dim.

This should be explained with some examples and I will return to this later.

Simplicity (versus Complexity)

Fix a family of procedures (= steps) we call **simple** and a family of objects (or systems) which are considered to be **simple** ones.

Then, we measure **simplicity** of (apparently very complex) objects by a (minimal) number of **simple** steps which brings the object we study to a **simple** object.

> It is the opposite philosophy to the standard understanding of **complexity**. But procedures are not reversible and "small simplicity" may co-exist with huge complexity.

Examples

Symmetrizations:

Steiner symmetrization

$\mathbb{R}^n \supset K$ – convex, h – hyperplane

$St_h K$ – one step of Steiner sym.

$$\prod_{h=1}^{N} St_{h_i} K := K_N$$

Fix some $c > 1$.

What is the smallest N s.t. for $\forall K \subset \mathbb{R}^n$ one may find $\{h_i\}_1^N$ such that $\exists$ ellipsoid $\mathcal{E}$ and

$$(*) \qquad \mathcal{E} \subset K_N \subset c \cdot \mathcal{E} \quad ?$$

Hadwiger ($\sim$ 55): $\qquad N \lesssim (c_1 \cdot n)^{n/2}$

Bourgain–Lindenstrauss–Milman ($\sim$1987):

$N \lesssim c_1 n \log n$ (for some $c \sim 3$)

Klartag–Milman (2003):

$$N = \frac{3}{2} n \quad \text{enough .}$$

Actually: $\forall\varepsilon > 0\ \exists c(\varepsilon)$ instead of c in $(*)$ s.t.

$$N \leq (1 + \varepsilon)n$$

(and, for some K, at least $n - c_1 \log n$ is necessary).

If D is the standard euclidean ball then additional $n-1$ sym. turns $\mathcal{E}$ to $r.D$ and altogether $N \leq (2 + \varepsilon)n$ is enough.

(Note: isomorphic answer)

So, the "simplicity" of any convex K (to derive an ellipsoid by using Steiner sym. as elementary steps) is around the same as the simplicity of an ellipsoid with respect to a euclidean ball.

Explicit versus random; derandomization

Simplicity/complexity may be measured also through comparing explicit constructions (of interesting features of high dimensionality) with random ones.

More precisely, we are interested in a number of **explicit** steps and a number of **randomly** selected steps needed to solve a specific problem.

The reason for such an approch is the following:

Most (all?) remarkable structures we discovered in high-dimensional convex bodies are “not visible”.

We prove their existence “with very high probability” but we don’t know explicit constructions which lead to these structures (and what is “explicit”?) or even just algorithms to recover them.

One such test-example is ℓ_1^N (cross-polytope which is the unit ball of ℓ_1^N).

Let $0 < c, C, C_1, \ldots$ be universal constants.

Figiel-Lindenstrauss-Milman (1976 and 1977)

$\exists \epsilon_\circ$ s.t. $\forall \epsilon$, $\epsilon_\circ > \epsilon > 0$, $\forall N$,

$\exists$ subspace $E \hookrightarrow \ell_1^N$, $\dim E = n \geq c\epsilon^2 N$

and $\mathrm{dist}(E, \ell_2^n) \leq 1 + \epsilon$

(i.e. for some $r > 0$, $r \cdot |x| \leq \|x\| \leq (1+\epsilon) r \cdot |x|$)

— . — . —

Kashin (1977)

$$\forall \lambda, \quad \frac{1}{2} < \lambda < 1,$$

$$\exists E \hookrightarrow \ell_1^N, \quad \dim E = n \geq \lambda N$$

$$\text{and } \mathrm{dist}(E, \ell_2^n) \leq C(\lambda).$$

— . — . —

Examples of partially explicitly (and partially randomly) solved problems.

(i) We already discussed that $t \sim (b/M)^2$ rotations $\{u_i\}_1^t$ are needed (M.-Sch.) and also sufficient (BLM) to approximate the euclidean norm

$$|x| \sim \frac{1}{t} \sum_1^t \|u_i x\|$$

(and random $\{u_i\}_1^t$ are OK with high probability).

However: just $4 \log_2 b/M + C$ random $v_i \in O(n)$ is enough and then explicit (short) construction brings the same result. [Artstein-M.2006]

(ii) From Klartag (2002): 5 selections of orthogonal basis enough to approx. eucl. ball using Minkowski symmetrizations. But only 1 selection is random, and $\log^* n$ selections are enough to achieve approximation explicitly (using very simple steps).

II. Geometrization of Probability

The goal of these talks is directed more to the next stage of the theory (as I see it).

Extension of the category of Convex Bodies to the category of log-concave functions (measures)

Definitions. Consider $d\mu = f dx$, $f \geq 0$ μ is log-concave iff $\forall\, A, B \subset \mathbb{R}^n$ and $0 < \lambda < 1$

$$\mu\Big(\lambda A + (1-\lambda)B\Big) \geq \mu(A)^{\lambda}\mu(B)^{1-\lambda}$$

Very important example: (i) $\mu(K) = \operatorname{Vol} K$ (by Brunn-Minkowski theorem)

(ii) $\mu_K(A) = \frac{\operatorname{Vol}(K \cap A)}{\operatorname{Vol} K}$, K-convex.

(iii) Marginals of Volume; $\operatorname{Proj}_E \mu$, $E \hookrightarrow \mathbb{R}^n$, with the density $\operatorname{Proj}_E f = \int_{x+E^{\perp}} f(y)dy$.

Function $f(x) \geq 0$ is called log-concave if $\log f$ is concave, i.e. $f(x) = e^{-\varphi(x)}$ and φ is convex.

Connection (C. Borell): Let $\operatorname{Supp}\mu$ not belong to any affine hyperplane. Then μ is log-concave iff μ is abs. continuous on $\operatorname{Supp}\mu$ and the density f is a log-concave function.

More examples: $e^{-|x|}$; $\frac{1}{(\sqrt{2}\pi)^n}e^{-|x|^2/2}$; $e^{-\|x\|^p/p}$, for any norm and $1 \leq p < \infty$.

Log-concavity was used in Convexity Theory already from the '50s (Henstock–MacBeath) and later, say, Prékopa–Leindler's extension of Brunn–Minkowsky inequality, or the use of log-concave functions to study volume of sections of ℓ_p^n by Meyer–Pajor. But a purely geometric study of log-concavity waited until the end of the '80s, and was initiated by K. Ball, with the study of isotropicity of such measures and its connection with isotropicity of convex bodies.

Recently, it was observed that asymptotic theory of high-dim. convexity is extended to the much larger category of log-concave measures. In this extension we identify K with the measure

$$\mu_K := \mathsf{Vol}_{|_K} \text{ (i.e. } \mu(A) = \mathsf{Vol}(A \cap K)).$$

Three features characterize this extension.

(i) On the one hand, important geometric inequalities (and other kinds of geometric statements) are interpreted, extended and proved for log-concave measures.

(ii) On the other hand, some typical probabilistic results (and thinking) are interpreted and proved in a geometric framework.

(iii) And most importantly, an extension of the geometric approach to the log-concave category is needed to solve some central problems of a purely geometric nature.

Examples of results to confirm this picture.

(i) Functional form of some geometric inequalities.

Prékopa–Leindler inequality (functional version of Brunn–Minkowski inequality)

We introduce first sup-convolution which we call
Asplund product:

$$(f \star g)(x) = \operatorname{Sup}_{x_1+x_2=x} f(x_1)g(x_2).$$

Example: $1_K \star 1_T = 1_{K+T}$.
λ-homothety for function is

$$(\lambda \cdot f)(x) := f^{\lambda}\left(\frac{x}{\lambda}\right), \quad \lambda > 0$$

(So $f \star f = 2 \cdot f$)

Let us start with the very important and beautiful Brunn-Minkowski inequality: For $A, B \subset \mathbb{R}^n$ (and all involved sets are measurable)

$$\mathrm{Vol}(A+B)^{1/n} \geq (\mathrm{Vol}\, A)^{1/n} + (\mathrm{Vol}\, B)^{1/n}$$

In multiplicative, dimension-free form

$$\mathrm{Vol}\left(\lambda A + (1-\lambda)B\right) \geq (\mathrm{Vol}\, A)^{\lambda} \cdot (\mathrm{Vol}\, B)^{1-\lambda}$$

for $0 < \lambda < 1$.

Then the functional analogue of Brunn–Minkowski is exactly Prékopa–Leindler inequality:
For $f, g : \mathbb{R}^n \to [0, \infty)$, $0 < \lambda < 1$

$$\int (\lambda \cdot f) \star \left((1-\lambda) \cdot g\right) \geq \left(\int f\right)^{\lambda} \cdot \left(\int g\right)^{1-\lambda}$$

Also "isomorphic" inequalities have their functional form. E.g. geometric statement:

Reverse Brunn–Minkowski (Milman, '85): $\exists C$ s.t. for any convex, sym., $K, P \subset \mathbb{R}^n$, there are $T_K, T_P \in SL_n$ such that if $\tilde{K} = T_K(K)$, $\tilde{P} = T_P(P)$, then

$$\mathrm{Vol}(\tilde{K} + \tilde{P})^{\frac{1}{n}} < C\Big[\mathrm{Vol}(\tilde{K})^{\frac{1}{n}} + \mathrm{Vol}(\tilde{P})^{\frac{1}{n}}\Big]$$

where T_K depends solely on K and T_P solely on P.

Its functional analogue (Klartag-Milman, '05): For any even log-concave $f, g : \mathbb{R}^n \to (0, \infty)$ there are $T_f, T_g \in SL_n$, s.t. $\tilde{f} = f \circ T_f$, $\tilde{g} = g \circ T_g$ satisfy

$$\Big[\int \tilde{f} \star \tilde{g}\Big]^{\frac{1}{n}} < C\Big[\Big(\int \tilde{f}\Big)^{\frac{1}{n}} + \Big(\int \tilde{g}\Big)^{\frac{1}{n}}\Big]$$

where T_f depends solely on f and T_g solely on g (and C is, as before, a universal constant).

Notion of Polarity;

functional version of Santaló inequality

Let $K \subset \mathbb{R}^n$, convex, $0 \in K$

$$K^\circ := \left\{x \in \mathbb{R}^n : (x,y) \le 1 \ \ \forall\, y \in K\right\}$$

[FA interpretation: If $K = -K$, $\|x\|_K$ – Minkowski functional, i.e. K is the unit ball of $X = (\mathbb{R}^n, \|\cdot\|_K)$. Then $X^* = (\mathbb{R}^n, \|\cdot\|_K^*)$ has K° its unit ball.]

Let $\mathcal{D}$ be the unit euclidean ball.

Blaschke–Santaló inequality

Let $K = -K$, then

$$|K| \cdot |K^\circ| \le |\mathcal{D}|^2$$

(max. achieved on $K := \mathcal{D}$).

Problem: What is $\min |K| \cdot |K^\circ|$ (Mahler, ∼'39)?

Asymptotic answer [Bourgain-Milman, 1985]:

$\exists c > 0$ universal s.t.

$$c \leq \left(\frac{|K| \cdot |K^\circ|}{|\mathcal{D}|^2} \right)^{1/n}.$$

For general convex K: $\exists x_0$ s.t. for $\widehat{K} = K - x_0$

$$|\widehat{K}| \cdot |\widehat{K}^\circ| \leq |\mathcal{D}|^2$$

($\min_x |K| \cdot |(K - x)^\circ|$ achieved for x_0 called Santaló point; then 0 is the barycenter of $(K - x_0)^\circ$.)

Now the functional version. We start with

Legendre transform

$$\mathcal{L}\varphi(x) = \sup_{y\in\mathbb{R}^n} \Big[(x,y) - \varphi(y)\Big].$$

If φ – convex and low semi-continuous $\to$ $\mathcal{L}\mathcal{L}\varphi = \varphi$.

Define polarity:

$f^\circ = e^{-\mathcal{L}(-\log f)}$, i.e. $-\log f^\circ = \mathcal{L}(-\log f)$,

or

$$f^\circ(x) = \inf_{y\in\mathbb{R}^n} \frac{e^{-(x,y)}}{f(y)}.$$

Let $LC(\mathbb{R}^n)$ be the class of (all) upper semicontinuous non-negative functions s.t. their logs are concave (in short, "log-concave" functions.) Then $f \in LC(\mathbb{R}^n)$ implies $(f^\circ)^\circ = f$.

Examples:

$$\mathbb{1}_K^\circ = e^{-\|x\|_{K^\circ}} \qquad \left(e^{-\|x\|_K^2/2}\right)^\circ = e^{-\|x\|_{K^\circ}^2/2}.$$

Some elementary properties:

1. $(f^\circ)^\circ = f\,,$

2. if $f \leq g \Rightarrow f^\circ \geq g^\circ,$

3. $(f \star g)^\circ = f^\circ \cdot g^\circ$

(for log-concave functions $(f \cdot g)^\circ = f^\circ \star g^\circ$)

4. $(\lambda \cdot f)^\circ = (f^\circ)^\lambda$

Theorem (Artstein, Klartag, Milman; 2005):

Let $f : \mathbb{R}^n \to \mathbb{R}^+$, $\int f < \infty$. *Then*

(i) *for some* x_0 *and* $\widetilde{f}(x) = f(x - x_0)$,

$$\int \widetilde{f} \cdot \int \widetilde{f}^\circ \leq (2\pi)^n \ . \qquad (*)$$

For log-concave f, *we may take* $x_0 = \int xf \Big/ \int f$.

[If f-even, $x_0 = 0$ and $(*)$ was proved by K. Ball.]

(ii) $\min_{x_0} \int \widetilde{f} \cdot \int \widetilde{f}^\circ = (2\pi)^n$ *iff* f *is a gaussian.*

The standard geometric Santaló inquality for convex bodies follows from $(*)$: apply $(*)$ to $f = e^{-\|x\|_K^2/2}$.

Then $\int_{\mathbb{R}^n} f\,dx = c_n |K|$ where $c_n = (2\pi)^{n/2}/|\mathcal{D}|$.

The generalization of Santaló's inequality for a log-concave class had as its main goal to justify the notion of polarity. However, many more inequalities of this kind follow. They are interesting in themselves but demonstrate that generalization of Santaló's inequality does not justify in itself that we deal with the correct notion of duality.

What should we call "duality"?

We start with the class $Cvx(\mathbb{R}^n)$ of all *lower-semi-continuous* convex functions $\{f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$.

Theorem (Arstein-Avidan, Milman). *Let* $T : Cvx(\mathbb{R}^n) \to Cvx(\mathbb{R}^n)$ (*1–1 and onto*) *satisfying*

1. $T \cdot T\varphi = \varphi$ (*for any* $\varphi \in Cvx(\mathbb{R}^n)$);

2. $\varphi \leq \psi$ *implies* $T\varphi \geq T\psi$.

Then T *is essentially the Legendre transform* $\mathcal{L}$*:* $\exists c_0 \in \mathbb{R}$, $v_0 \in \mathbb{R}^n$, *symmetric linear* $B \in \mathrm{GL}_n$ *s.t.*

$$(T\varphi)(x) = (\mathcal{L}\varphi)(Bx + v_0) + (x, v_0) + C_0.$$

Turn to the class $LC(\mathbb{R}^n)$.

What should be a duality transform for this class (and does it exist in some natural sense)?

The corollary of the previous theorem is

Corollary (Artstein-Avidan; Milman)**.** *Let* $T : LC(\mathbb{R}^n) \to LC(\mathbb{R}^n)$ *satisfy,* $\forall f \in LC(\mathbb{R}^n)$

1. $T \cdot Tf = f$;
2. $f \leq g$ *implies* $Tf \geq Tg$.

Then $\exists\, 0 < c_0 \in \mathbb{R}^n$, $v_0 \in \mathbb{R}^n$ *and the symmetric operator* $B \in \mathrm{GL}_n$ *s.t.* T *is defined by*

$$(Tf)(x) = c_0 e^{-(v_0,x)} \inf_y \frac{e^{-(Bx+v_0,y)}}{f(y)},$$

i.e. $(Tf)(x) = c_0 e^{-(v_0,x)} f^\circ(Bx + v_0)$.

The Concept of Duality

Definition. Let S be a set of functions defined on $\mathbb{R}^n$. We say that a transform $T : S \to S$ (onto, 1–1) generates a duality if

1. $\forall f \in S$, $T \cdot Tf = f$;

2. $\forall f, g \in S$, $f \leq g$ implies $Tf \geq Tg$.

The above theorem and corollary describe all existing dualities on classes $Cvx(\mathbb{R}^n)$ and $LC(\mathbb{R}^n)$.

A recent result by Böröczky–Schneider implies the same kind of characterization for the case of the family of all compact convex bodies with 0 in their interiors (i.e. the standard duality is, essentially, the only existing duality for the class of compact convex bodies).

Our method gives another proof of this and also provides the same result for the class of convex bodies containing 0 and not necessarily bounded. The duality for the case of the family of norms on $\mathbb{R}^n$ is also characterized in the same way, which follows from joining Gruber's (1992) and Böröczky–Schnieder's results (the whole weight of the proof lies within Gruber's paper, but it was written and developed for a different problem).

Another example which demonstrates the level of non-triviality of formulas which may be derived from a simple starting point.

Let $s > 0$:

$Conc_s^+(\mathbb{R}^n)$ is the family of all bounded s-concave functions $f : \mathbb{R}^n \to \mathbb{R}^+$ (i.e. $f^{1/s}$ is concave on its convex support) and s.t. $f(0) > 0$.

The importance of these classes in Convex Geometry is in the fact that, for s an integer, they are marginals of uniform distributions (Volumes) on convex bodies.

We define, in [Artstein-Klartag-Milman], a duality for this class by

$$\mathcal{L}_s f = \inf_{y: f(y)>0} \frac{(1-(x,y))_+^s}{f(y)}.$$

Theorem (Artstein–Milman)**.** *Let*

$T : Conc_s^+(\mathbb{R}^n) \to Conc_s^+(\mathbb{R}^n)$, *1–1, onto*

1. $T \cdot Tf = f$;
2. $f \leq g$ *implies* $Tf \geq TG$.

Then, there exists a constant $C_0 \in \mathbb{R}$ *and an invertible symmetric linear transformation* $B \in \mathrm{GL}_n$ *such that*

$$(Tf)(x) = C_0 \inf_{(y: f(By)>0)} \frac{(1-\langle x,y\rangle)_+^s}{f(By)}.$$

Return to the Functional version of Santaló's inequality.

Also the reverse inequality is true in the functional form (similar to Bourgain-Milman's reverse of Santaló's original geometric inequality).

Theorem (Klartag-Milman; 2005). $\exists c > 0$, C *s.t.* $\forall$ *log-concave* $f : \mathbb{R}^n \to \mathbb{R}^+$, $\int f < \infty$, *we have*

$$c < \left(\int_{\mathbb{R}^n} f \cdot \int_{\mathbb{R}^n} f^\circ \right)^{1/n}.$$

Functional form of Urysohn inequality

(Klartag–Milman)

(Urysohn inequality for log-concave functions)

Recall the classical Urysohn inequality:

$$\left(\frac{\operatorname{Vol} K}{\operatorname{Vol} \mathcal{D}}\right)^{1/n} \leq M^\star(K) := \int_{S^{n-1}} \sup_{y\in K}(x,y)d\sigma(x)$$

and

$$\operatorname{Vol}(\mathcal{D}+\varepsilon K) = \operatorname{Vol}\mathcal{D}+\varepsilon n M^\star(K)\operatorname{Vol}\mathcal{D}+O(\varepsilon^2).$$

So, we may define the analogous quantity. Let $G(x) = e^{-|x|^2/2}$. Then define

$$V_G(f) = \lim_{\varepsilon\to O^+} \frac{\int G\star[\varepsilon\cdot f] - \int G}{\varepsilon}$$

(one may show that lim exists).

Denote $M^\star(f) = 2\frac{V_G(f)}{n\int G} = \frac{V_G(f)}{\frac{n}{2}(2\pi)^{n/2}}$.

Then $M^\star(G) = 1$.

If $f = 1_K$ then (calculation)

$$V_G(1_K) = \frac{(2\pi)^{\frac{n-1}{2}} n\kappa_n}{\kappa_{n-1}} M^\star(K)$$

($\kappa_n = \mathrm{Vol}\,\mathcal{D}_n$).

So $M^\star(K) = c_n M^\star(1_K)$ for $c_n \sim \sqrt{n}$.

Properties: $M^\star(f \star g) = M^\star(f) + M^\star(g)$

$$M^\star(\lambda \cdot f) = \lambda M^\star(f),\ \lambda > 0.$$

Theorem (Klartag–M., 2005). Let $f : \mathbb{R}^n \to [0, \infty]$ be an even log-concave function s.t. $\int f = \int G$ $(= (2\pi)^{n/2}$. Then

$$M^\star(f) \geq M^\star(G) = 1.$$

(ii) **A Central Limit Theorem (CLT) for Convex Sets**

In the classical approach, we study a geometric shape of projections (or sections) of convex body K (and we know that they are, with high probability, close to euclidean balls for small enough rank of projections).

But what about measure projections (marginals) in place of geometric projections?
(Gromov '87)

Normalize the convex body $K \subset \mathbb{R}^n$ s.t.

$$\mathrm{Vol}\, K = 1, \quad \int_K \vec{x}dx = 0, \quad \int_K \langle x, \theta \rangle^2 dx = |\theta|^2 L_K^2,$$

for any $\theta \in \mathbb{R}^n$. We say that K is in "isotropic" position and the constant L_K is called the isotropic constant of K.

Then most marginals are approximately gaussian! Progress towards this goal was obtained earlier by Brehn–Voigt (2000) and Antilla–Ball–Perissinaki (2003) and others. But the complete solution was achieved recently by Klartag:

Theorem (Klartag, '06):

Suppose $K \subset \mathbb{R}^n$ is convex and isotropic, and X is distributed uniformly in K.

Then $\exists \Theta \subset S^{n-1}$ with $\sigma_{n-1}(\Theta) \geq 1 - \delta_n$, such that for $\theta \in \Theta$,

$$\sup_{A \subset \mathbb{R}} \left| \text{Prob}\left\{ \langle X, \theta \rangle \in A \right\} - \frac{1}{L_K \sqrt{2\pi}} \int_A e^{-\frac{t^2}{2L_K^2}} dt \right| \leq \epsilon_n.$$

Here, say, $\delta_n < \exp(-c\sqrt{n}), \epsilon_n < C(n^{-1/30})$.

There is an analogue multi-dimensional version.

The proofs of these results use, very essentially, the extension of the whole theory to the log-concave category.

A Sketch of the Proof

It will consist of a few steps. Let μ be the uniform distribution over K, i.e. $\mu = 1_K dx$.

Step 1. Change the problem and consider another measure (convolution with a gaussian $N(0; \epsilon^2)$)

$$\nu = \mu * \gamma_\epsilon$$

where γ_ϵ has the density $\frac{1}{(\sqrt{2\pi})^n \epsilon^n} e^{-|x|^2/2\epsilon^2}$.

ν is log-concave measure.

Step 2. Consider marginal $P_E \nu$ of ν on k-dim. subspace $E \in G_{n,k}$ and its density $(dP_E\nu/dx)(x)$ at $x \in E$.

Important observation: log of this density

$$g(E; x) = \log \left(\frac{dP_E \nu}{dx} \right)(x)$$

has the Lip. constant (with respect to E and x) strongly improved (and the role of gaussian γ_ϵ is crucial here). Actually it could even be ∞ before convolution.

Step 3. As a consequence, on a "random" subspace $E \in G_{n,k}$ (for $k \sim n^{\kappa}$, for some $\kappa > 0$) g is almost spherically invariant.

Step 4. The measure $P_E\nu$ on E (now a fixed subspace) is log-concave and $\sim$ spherically symmetric. This implies that it is strongly concentrated around ONE sphere (of, say, radius R). Also, this radius R may be taken independently of E. (This means that for some $\alpha > 0$ and $\beta > 0$

$$P\Big\{\Big||x| - R\Big| \geq t\Big\} \leq ce^{-c_1 n^{\alpha} t^{\beta}}. \quad)$$

Step 5. Because Step 4 is true for random E, it follows that also ν itself is strongly concentrated around the sphere of radius $R_1 \sim \sqrt{\frac{n}{k}}\, R$.

Step 6. Return to the orginal measure μ. It is also strongly concentrated around ONE sphere (because gaussian γ_ϵ is taken negligibly small).

Step 7. Then its marginal is gaussian because it is true for a sphere.

(iii) Isotropic position and isotropic constant (Slicing problem)

We repeat that a convex body $K \subset \mathbb{R}^n$, with the barycenter of K at 0, is in isotropic position iff $\operatorname{Vol} K = 1$, and $\forall i, j = 1, \ldots, n$

$$\int_K x_i x_i dx = \delta_{ij} L_K^2$$

($x = (x_1, \ldots, x_n)$). We call L_K the isotropic constant of K. It is an old and famous problem of Bourgain if isotropic constants $\{L_K\}$ are uniformly bounded (by dim. n and convex bodies in $\mathbb{R}^n$). Well-known 20 year old estimate of Bourgain states that $L_K \leq Cn^{1/4} \log n$.

However, recently Klartag proved

Theorem (Klartag, '05). *For any convex body $K \subset \mathbb{R}^n$ and $\epsilon > 0$ there exists a convex body $T \subset \mathbb{R}^n$, s.t.*

$$(1 - \epsilon)T \subset K - x_0 \subset (1 + \epsilon)T$$

and $L_T < c/\sqrt{\epsilon}$.

Corollary (Klartag '05, relies on Paouris' recent theorem).

$$L_K < Cn^{1/4} \text{ when } K \subset \mathbb{R}^n.$$

Important to note the proof of the last theorem requires the extension of Asymptotic Theory of Convexity to the category of log-concave measures.

To give some details we need to establish a connection between log-concavity and convex bodies.

For any even log-concave $f : \mathbb{R}^n \to \mathbb{R}^+$ we associate a norm (K. Ball, 1990)

$$\|x\|_f = \left(\int_0^\infty n f(rx) r^{n-1} dr \right)^{-1/n}.$$

Denote K_f the unit ball of $\| \cdot \|_f$.

A few properties of this correspondence:

1. $\mathrm{Vol}\, K_f = \int f$

2. Define $\overline{\overline{K}}_f = \{x \in \mathbb{R}^n : f(x) > e^{-n}\}$. Then, for a universal $c > 0$,

$$K_f \subset \overline{\overline{K}}_f \subset c K_f.$$

3. Let f and g – log-concave and $f(0) = g(0) = 1$. Then, for some universal constants c_1 and c_2

$$c_1 K_{f \star g} \subset K_f + K_g \subset c_2 K_{f \star g}$$

and

$$c_1 n K_f^{\circ} \subset K_{f^{\circ}} \subset c_2 n K_f^{\circ}.$$

Let us define the isotropic constant of log-concave measure.

We say that f is in *isotropic position* if

$$\operatorname*{Sup}_{x\in\mathbb{R}^n} f(x) = 1 = \int f(x)dx \quad \text{and}$$
$$\int_{x\in\mathbb{R}^n} x_i x_j \, f dx = \delta_{ij} L_f^2$$

and the constant L_f is called the *isotropic constant* of the measure fdx.

One may write a formula for L_f without "putting" fdx in the isotropic position,

$$L_f = \left(\frac{\operatorname*{Sup}_{x\in\mathbb{R}^n} f(x)}{\int_{\mathbb{R}^n} f dx} \right)^{1/n} \quad (\det \operatorname{Cov} f)^{1/2n}$$

where covariance matrix

$$\operatorname{Cov} f = \Big(\operatorname{Cov}_f(x_i, x_j) \Big),$$
$$\operatorname{Cov}_f(x_i, x_j) = \frac{\int_{\mathbb{R}^n} x_i x_j f dx}{\int f dx} - \frac{\int x_i f}{\int f} \cdot \frac{\int x_j f}{\int f}.$$

Then, for any K convex, $L_K = L_{1_K}$.

A sketch of Klartag's proof of a solution of the "isomorphic" slicing problem.

Let K – convex compact, $O \in K$, $\operatorname{Vol} K = 1$.

1. Let $f : K \to [0, \infty)$ and log-concave.

   Assume
   $$\left(\frac{\operatorname{Sup}_{x \in K} f}{\inf_{x \in K} f} \right)^{1/n} < C.$$

   Then K_f isomorphic to K, i.e. $\exists\, c_1 := c_1(c)$ s.t.
   $$\frac{1}{c_1} K_f \subset K \subset c_1 K_f$$

   (here, as before,
   $$K_f = \left\{ x \in \mathbb{R}^n;\ \int_0^\infty n f(rx) r^{n-1} dr \geq 1 \right\}$$

   and is a convex set by K. Ball).

2. (K. Ball) $L_f \simeq L_{K_f}$.

   So, we should find such an f that $L_f <$ const.

3. Consider a (convex) function $F_K(x) = F(x)$

$$F(x) = \log \int_K e^{\langle x,y\rangle} dy.$$

   (a) This function produces a *transportation of measure*

$$\nabla F := \psi : \mathbb{R}^n \longrightarrow \overset{\circ}{K}$$

   (similar to the so called 'momentum' map) which means that

$$\forall A \subset \mathbb{R}^n, \quad \int_A \det \operatorname{Hess} F = \operatorname{Vol}((\nabla F)A) \leq 1.$$

Let μ_1 and μ_2 be two Borel measures in $\mathbb{R}^n$ and $T : \mathbb{R}^n \to \mathbb{R}^n$ s.t. $\forall A \subset \mathbb{R}^n$ (measurable)

$$\mu_2(A) = \mu_1(T^{-1}A).$$

Then T transports μ_1 to μ_2.

Equivalently, $\forall \varphi \in C^+(\mathbb{R}^n)$

$$\int_{\mathbb{R}^n} \varphi(x) d\mu_2(x) = \int_{\mathbb{R}^n} \varphi(Tx) d\mu_1(x).$$

Fact. *Let $F : \mathbb{R}^n \to \mathbb{R}$ be C^2-smooth strictly convex and $K = \mathrm{Im}(\nabla F)$. Let measure μ have density* $\frac{d\mu}{dx} = \det \mathrm{Hess}\, F(x)$.
Then $\nabla F : \mathbb{R}^n \to \mathbb{R}^n$ transports μ to $\mathrm{Vol}_{|K}$.

Applying to our situation, we see that ∇F_K transports the measure μ_K to the uniform measure on K.

(b) Note that $\nabla F(x) = \int y \, d\mu_{K,x}(y)$ and the density of $\mu_{K,x}$ is

$$\frac{e^{\langle x,y\rangle} 1_K(y)}{\int_K e^{\langle x,z\rangle} dz}.$$

Also Hess$(F)(x) =$ Cov$(\mu_{K,x})$.

Therefore

$$\det \text{Hess}\, F(x) = \Big(\int f_x / \,\text{Sup}\, f_x\Big)^2 \cdot L_{f_x}^{2n}$$

where $f_x(y) = e^{\langle x,y\rangle} 1_K(y)$.

So, we consider the family of log-concave functions and we search for a function as in 1. and 2. inside this family.

4. Let $x \in nK^\circ$. (Note: $|K| = 1 \Rightarrow |nK^\circ|^{1/n} \sim 1$.)

(a) Then

$$\left(\frac{\operatorname{Sup}_{y \in K} f_x(y)}{\inf_{y \in K} f_x(y)} \right)^{1/n} < C.$$

Indeed, $\operatorname{Sup}_{y \in K} f_x(y) = \operatorname{Sup}_{y \in K} e^{\langle x, y \rangle} \leq e^{\|x\|^*} \leq e^n$. (Similarly for $\inf \geq e^{-n}$.)

So we know that $K_{f_x} \sim K$ for $\forall x \in nK^\circ$.

(b) We want to find $x \in nK^\circ$ s.t. $L_{f_x} <$ Const., i.e. to estimate

$$\Big(\det \mathrm{Hess}\, F(x)\Big)^{1/2n} \left(\frac{\mathrm{Sup}\, f_x}{\int f_x}\right)^{1/n} <?$$

Actually, it is enough to find $x \in nK^\circ$, s.t.

$$\det \mathrm{Hess}\, F(x) < \mathrm{Const.}^n.$$

We prove this "in average":

$$\frac{1}{|nK^\circ|}\int_{nK^\circ} \det \mathrm{Hess}\, F(x) \leq \frac{1}{|nK^\circ|}\mathrm{Vol}(\mathrm{Im}(\nabla F)) \leq C^n$$

(reverse Santaló).

What is next?

Does the family of log-concave measures represent the largest class of probability measures where Geometry is extended so naturally?

It is not clear; but let us consider a much larger class of "convex measures" (I also like the terminology "hyperbolic measures").

Definition (C. Borell, '74): Fix $-\infty \leq s \leq 1$; a measure μ on $\mathbb{R}^n$ is s-concave iff $\forall\, A, B \subset \mathbb{R}^n$ non-empty and measurable, $t \in (0,1)$,

$$\mu\Big(tA + (1-t)B\Big) \geq \Big[t\mu(A)^s + (1-t)\mu(B)^s\Big]^{1/s}.$$

Note, for $s = 0$,

$$\mu\Big(tA{+}(1{-}t)B\Big) \geq \mu(A)^t \mu(B)^{1-t} \quad \text{(log-concavity)}$$

and, for $s = -\infty$,

$$\mu\Big(tA + (1-t)B\Big) \geq \min\Big(\mu(A), \mu(B)\Big).$$

Denote $\mathcal{M}(s)$ the class of all finite s-concave measures. Clearly $\mathcal{M}(s_i) \supseteq \mathcal{M}(s_2)$ for $s_1 < s_2$.

New example: Cauchy distribution with density

$$p(x) = \frac{c_n}{(1+|x|^2)^{\frac{n+1}{2}}};$$

in this case $s = -1$ ("heavy tails" distributions).

C. Borell: (i) $\forall\, \mu \in \mathcal{M}(-\infty)$ has a convex supp$K \subset \mathbb{R}^n$ and μ is abs. continuous (w.r.t. Lebesgue measure on K);

(ii) If μ is s-concave, then $s \leq 1/\dim K$;

(iii) If $\dim K = n$, the density p of μ satisfies $\forall\, x, y \in K$

$$p\Big(tx + (1-t)y\Big) \geq \Big(tp(x)^{s_n} + (1-t)p(y)^{s_n}\Big)^{1/s_n}$$

for $s_n = \frac{s}{1-ns}$.
(So, if μ is log-concave then also its density is a log-concave function; however,
if $s = -\infty$ then its density is $-\frac{1}{n}$ concave, and we call this class *convex measures*).

Also, levels of density of convex measures are boundaries of convex sets.

Connection with Classical Convexity

The definition of convex measures corresponds to the unified principle behind most (all) geometric inequalites, a principle of minimization:

$$f(A;B) \geq \min\Big\{f(A;A), f(B;B)\Big\}$$

[“the minimum is achieved on equal objects”].

I call inequalities satisfying this principle geometric inequalities of *hyperbolic type*.

To continue this study we need to recall a few facts from *Convex Classic*.

Next we use a short notation $|K|$ for $\mathrm{Vol}\, K$.

A. From Convex Classic:

1. Brunn-Minkowski ineq.: $K, T \subset \mathbb{R}^n$,

$$|K+T|^{1/n} \geq |K|^{1/n} + |T|^{1/n}.$$

2. Notion of mixed volumes:

K_i – convex $\subset \mathbb{R}^n$, $\lambda_i \geq 0$, $m \geq n$

$$\left|\sum_1^m \lambda_i K_i\right| = \sum_{1\leq i_1 \leq \cdots \leq i_n \leq m} \lambda_{i_1}\lambda_{i_2}\cdot\ldots\cdot\lambda_{i_n} V(K_{i_1}, \ldots, K_{i_n})$$

coefficients (invariant w.r.t. permutations) – called mixed volumes.

Special case (D_n is the euclidean ball in $\mathbb{R}^n$):

$$|K + tD_n| = \sum_0^n \binom{n}{i} W_i(K) t^i,$$

where $W_i(K) = V(\underbrace{K, \ldots, K}_{n-i}, \underbrace{D, \ldots, D}_{i})$

– i-th Quermassintegral

(other notation: $W_i(K) := V_{n-i}(K)$). Then

$$W_i(K+T)^{1/n-i} \geq W_i(K)^{1/n-i} + W_i(T)^{1/n-i}.$$

3. Alexandrov-Fenchel inequalities:

$$V(K, T, A_3, \ldots, A_n)^2$$
$$\geq V(K, K, A_3, \ldots, A_n) \cdot V(T, T, A_3, \ldots, A_n)$$

and particular cases, called Alexandrov ineq.,

$$\left(\frac{V_i(K)}{|D_n|}\right)^{1/i} \geq \left(\frac{V_j(K)}{|D_n|}\right)^{1/j}, \quad 1 \leq i < j \leq n.$$

(The case $i = 1$ and $j = n$ is Urysohn's inequality.)

B. Going back in time to locate very similar-looking **numerical inequalities.**

Newton inequalities:

Let $\overline{x} = \{x_i\}_{i=1}^n$ – positive numbers

Elementary symmetric (and normalized) functions $1 \leq i \leq n$

$$E_i(x_1, \ldots, x_n) = \frac{1}{\binom{n}{i}} \sum_{1 \leq j_1 < \cdots < j_i \leq n} x_{j_1} \cdot x_{j_2} \cdot \ldots \cdot x_{j_i}$$

(and $E_0 \equiv 1$).

So $E_1(\overline{x})$ – arithmetic means,

$E_n(\overline{x})$ – geometric means.

For $k = 1, \ldots, n-1$

$E_k^2(\overline{x}) \geq E_{k-1}(x_1, \ldots, x_n) \cdot E_{k+1}(x_1, \ldots, x_n)$. (N)

$\Longrightarrow$ Corollary (Maclaurin)

$$E_1(\overline{x}) \geq E_2^{1/2}(\overline{x}) \geq \cdots \geq E_n^{1/n}(\overline{x}).$$

(Recall Alexandrov ineq.; and arithmetic-geom. mean ineq. is analogous to Urysohn inequality).

Indeed:

(N) for $k = 1$: $E_1^2 \geq E_2 \Rightarrow E_1 \geq E_2^{1/2}$

$$k = 2\text{: } E_2^2 \geq E_1 \cdot E_3 \Rightarrow \geq E_2^{1/2} E_3 \Rightarrow$$

$$\Rightarrow E_2^{1/2} \geq E_3^{1/3},$$

and so on.

Proof of (N): Use polynomials with real roots. Direction of ineq. in (N) is consequence of real roots (and "hyperbolic" is named because of it).

Proof of (N): Consider

$$P(x) = \prod_1^n (x - x_i) = \sum_0^n (-1)^j \binom{n}{j} E_j x^{n-j}$$

or, in homogeneous form

$$Q(t, \tau) = \tau^n P\left(\frac{t}{\tau}\right) = \sum_{j=0}^n (-1)^j \binom{n}{j} E_j(\vec{x}) t^{n-j} \tau^j$$

(Note: for a fixed τ, only real roots by t, and for fixed t, real roots by τ.)

Take $(n-k-1)$ times derivative of P by t and then $(k-1)$ times by τ; we have

$$\frac{n!}{2} E_{k-1} t^2 - n! E_k t\tau + \frac{n!}{2} E_{k+1} \tau^2,$$

and for, say, $\tau = 1$, this polynom has real roots in t. This means

$$E_k^2 \geq E_{k-1} \cdot E_{k+1} \text{ !}$$

To feel analogy even more, few "modern" numeric inequalites (Markus-Lopes, 1956):

$$E_k(x+y)^{1/k} \geq E_k(x)^{1/k}+E_k(y)^{1/k}, \quad k = 1, \dots n$$

($k = 1$ – equality; $k = n$ follows from geom. arithm. means ineq.)

C. Multi-dimensional (but still numerical) case.

Let S_n be the space of real sym. $n \times n$ matrices. If $A_i \in S_n$, $t_i > 0$, then

$$\det(t_1 A_1 + \cdots + t_m A_m)$$
$$= \sum_{1 \le i_1 \le \cdots \le i_n \le m} n! D(A_{i_1}, \ldots, A_{i_n}) t_{i_1} \cdot t_{i_2} \cdots t_{i_n}.$$

Coefficient $D(A_1, \ldots, A_n)$ is called the mixed discriminant for $A_1, \ldots, A_n$.

That $P(t) = \det(A + tI)$ has only real roots for any $A \in S_n$ leads to many interesting inequalities (similar to Newton ineq.) (Alexandrov, and later more general theory was built).

Examples: Let $A, B, C_i \geq 0$.

Alexandrov inequality (similar to Alexandrov-Fenchel for convex bodies)

$$D(A, B, C_3, \ldots C_n)^2 \geq$$
$$\geq D(A, A, C_3, \ldots, C_n) \cdot D(B, B, C_3, \ldots, C_n)$$

or its consequence

$$D(A_1, \ldots, A_n) \geq \prod_1^n (\det A_i)^{1/n} \quad (A_i \geq 0)$$

(and it is true that, for mixed volumes,

$$V(K_1, \ldots, K_n) \geq \left(\prod_1^n |K| \right)^{1/n}).$$

Again, the reverse triangle inequality, i.e. concavity of $[\det A]^{1/n}$ is true (Minkowski inequality):

for $A_i \geq 0$

$$[\det(A_1 + A_2)]^{1/n} \geq [\det A_1]^{1/n} + [\det A_2]^{1/n}$$

(proof is immediate from geom.-arithm. means ineq.)

Before we discuss "technical" connections between the convex theory (A) and the numerical $[(B)$ and $(C)]$, let us see the unified principle behind all of these inequalities.

All of them have a form of a principle of minimization

$$f(A;B) \geq \min\Big\{f(A;A), f(B;B)\Big\}$$

["The minimum is achieved on equal objects"], which is, in fact, equivalent to the original inequalities.

Examples: (i) Alexandrov–Fenchel inequality is equivalent to

$$V(A;B;C_1,\dots) \geq \min\Big(V(A;A;C_1,\dots);V(B;B;C_1,\dots)\Big)$$

(ii) Brunn–Minkowski ineq.: $\forall\, t,\tau > 0$ and A,B convex:

$$|tA+\tau B|^{1/n} \geq t|A|^{1/n} + \tau|B|^{1/n}$$

is equivalent to

$$|tA+\tau B| \geq \min\Big(|(t+\tau)A|,|(t+\tau)B|\Big).$$

And so on, and so on, and so on!

Is this an incidental similarity to "convex measure", i.e. s-concave measures for $s = -\infty$?

Or does a deeper meaning lie behind it?

How are numerical (multi-dimensional) and convexity theories connected?

First, we describe a couple of remarkable maps.

Let $|T|/|K| = \lambda$, K and T convex bodies.

Knothe map. Fix euclidean structure and orthogonal coordinate system.

$$\text{build } \varphi : \overset{\circ}{K} \longrightarrow \overset{\circ}{T},$$

measure preserving

$$\varphi = \Big(\varphi_1(x_1), \varphi_2(x_1, x_2), \varphi_3(x_1, x_2, x_3), \ldots, \varphi_n(x)\Big)$$

$$\frac{\partial \varphi_i}{\partial x_i} := \lambda_i(x) \text{ and } \prod \lambda_i = \lambda.$$

Example of use:

$$\begin{aligned}|K+T| \geq \Big|\{x+\varphi(x) \mid x \in \mathring{K}\}\Big| = \\ = \int_K \mathrm{det}(I + \mathrm{Jac}\,\varphi) ds = \\ = \int_K \prod(1+\lambda_i) \geq \int_K \Big(1 + \Big(\prod \lambda_i\Big)^{1/n}\Big)^n dx = \\ = (1+\lambda^{1/n})^n \cdot |K|\end{aligned}$$

$$\Longrightarrow |K+T|^{1/n} \geq |K|^{1/n} + |T|^{1/n}.$$

Brenier map.

Let K and T be convex compact bodies. Fix euclidean structure. Then there is

$$\psi : \overset{\circ}{K} \overset{\text{onto}}{\longrightarrow} \overset{\circ}{T}$$

which is "measure preserving", i.e.
$\det \operatorname{Jac} \psi = \text{Const.} := a$ and $\exists\, f$, $\operatorname{Dom} f = \overset{\circ}{K}$, convex and

$$\psi = \nabla f.$$

Example of use:

$$|K+T| = \int_K \det(I + \operatorname{Hess} f) dx \geq \int_K (1 + a^{1/n})^n dx = \\ = |K|(1 + a^{1/n})^n$$

(note $a \cdot |K| = |T|$), which implies the Brunn-Minkowsky inequality.

Warning: $\mathring{K}+\mathring{T}$ may be $\neq \{x+\psi(x) \mid x \in \mathring{K}\}$.

However (Alesker-Dar-Milman, 1999)

$$\exists\, u : \mathring{K} \xrightarrow{\text{onto}} \mathring{T}$$

"measure preserving", s.t.

$$\mathring{K}+\mathring{T} = \left\{x + u(x) \mid x \in \mathring{K}\right\}.$$

We use construction differently.

Let γ be Gaussian standard measure on $\mathbb{R}^n$, K_i – convex bodies. Let $|K_i| = 1$.

Step 1. Use the Brenier (-McCann) maps:

ψ_i – Brenier maps

$$\psi_i : (\mathbb{R}^n, \gamma) \longrightarrow (\mathring{K}_i, \text{Vol}),$$

$\psi_i = \nabla f_i$, f_i – convex functions on $\mathbb{R}^n$ and smooth (by a deep result of Caffarelli).

Step 2 (Gromov). $\forall\, t_i > 0$

$$\text{Im} \sum t_i \psi_i = \sum \text{Im}\, t_i \psi_i \quad :$$

$$(\mathbb{R}^n, \gamma) \quad \xrightarrow[\text{onto}]{\sum t_i \psi_i} \quad \sum t_i \,\text{Im}\, \nabla_i = \sum t_i \mathring{K}_i.$$

Therefore:

$$\Big|\sum t_i K_i\Big| = \int_{\mathbb{R}^n} \det\Big(\sum t_i A_i\Big)dx$$

$$\left\{\text{here } A_i = \text{Hess}\, f_i = \left(\frac{\partial^2 f_i}{\partial x_k \partial x_\ell}\right)\right\}$$

$$= \sum_{1\le i_1 \cdots \le i_n \le n} n!\, t_{i_1}\cdot \ldots \cdot t_{i_n} \int D(A_{i_1},\ldots,A_{i_n})dx,$$

So we have proved polynomiality and

$$V(K_{i_1},\ldots,K_{i_n}) = \int_{\mathbb{R}^n} D(A_{i_1},\ldots,A_{i_n})dx.$$

As ψ_i – measure preserving, $\det\left(\frac{\partial^2 f_i}{\partial x_k \partial x_\ell}\right)(x) = \gamma(x)$ (assuming $|K_i| = 1$), and (e.g.) from determinant inequality

$$D(A_1,\ldots,A_n) \ge \Big(\prod \det A_i\Big)^{1/n}$$

follows

$$V(K_1,\ldots,K_n) \ge \int_{\mathbb{R}^n} \left(\prod_1^n \text{Hess}\, f_i\right)^{1/n} dx$$

$$= \int_{\mathbb{R}^n} \gamma\, dx = 1 \quad \left(= \Big(\prod |K_i|\Big)^{1/n}\right)$$

III. Concentration Phenomenon

Isomorphic Form of Isoperimetric Problems

The standard form

Let (X, ρ, μ) be a metric probability space.

Let A be a measurable subset of X
and $\mu(A \subset X) \geq 1/2$.

Define $A_\epsilon = \{x \in X, \rho(x, A) \leq \epsilon\}$.

Define the function $\alpha(X; \epsilon) = 1 - \inf_A \mu(A_\epsilon)$
called the *concentration function* of X.

Consider a family $\mathcal{X} = \{(X_n, \rho_n, \mu_n)\}$.

For "natural" families, two estimates are typical:

$$\alpha(X_n, \epsilon) \leq c_1 \exp(-c_2 \epsilon^2 n) \qquad (*)$$

or

$$\leq c_1 \exp\Big(-c_2 \epsilon \sqrt{n}\Big) \qquad (**)$$

If $(*)$ is satisfied, $\mathcal{X}$ is called *normal* Lévy family;

these are "elliptic type" families.

Examples: S^n; $W_{n,k}$; SO(n); tori $\mathbb{T}_n$,

$F_2^n = \{\pm 1\}^n$, Π_n – permutation group

(metrics and measures should be specified).

Examples of families satisfying $(**)$ (I call them "hyperbolic type"):

$\mathrm{SL}_2(\mathbb{Z}_p)$ or, for any fixed $k \geq 2$, $\mathrm{SL}_k(\mathbb{Z}_p)$ where p plays the role of "n" above.

The concept of a Lévy family (and especially a normal Lévy family) generalizes the concept behind the law of large numbers in two directions:

a) the measures are not necessarily the product of measures (that is, we have no condition of “independence”), and

b) Lipschitz functions on the space are considered instead of linear functionals only.

To explain the reason for terminology "*concentration phenomenon*" and also outline why a bound of the form $(*)$ is so crucial, let us consider a 1-Lip function $f(x)$ defined on (X, ρ, μ), i.e.

$$\Big|f(x) - f(y)\Big| \leq \rho(x, y)\,.$$

Denote L_f the median of $f(x)$, which is defined by

$$\mu\Big\{x \in X \;\Big|\; f(x) \geq L_f\Big\} \geq \tfrac{1}{2}$$

and

$$\mu\Big\{x \in Z \;\Big|\; f(x) \leq L_f\Big\} \geq \tfrac{1}{2}\,.$$

Then

$$\mu\Big\{x \in X \;\Big|\; |f(x) - L_f| < \epsilon\Big\} \geq 1 - 2\alpha(X, \epsilon)\,.$$

If $\alpha(X, \epsilon)$ is very small, then the values of Lipschitz function "concentrate" in the measure around one value, meaning it is almost constant with high probability.

This is the case in all examples we mentioned above. It is a general property of high-dimensional metric probability spaces which is called "*concentration phenomenon*".

Such a "concentration" of measure balances the exponentially high entropy of n-dimensional spaces and leads to a "regularity" in high dimension, keeping "diversity" under control. The absolute constants involved in the examples are needed to balance rates of exponential decay (coming from Concentration) and exponential expansion (coming from covering/entropy). Surprisingly, both exponents have "roughly" the same order of decay via expansion by dimension and only a factor is needed to compensate them.

Consider an example:

Let X be S^{n+1}, the euclidean unit sphere in $\mathbb{R}^{n+2}$. Then

$$\alpha(X, \epsilon) \leq c_1 \exp(-\epsilon^2 n/2).$$

Let $\mathcal{N} = \{x_i\}_i^N \subset S^{n+1}$ and $N < c_1 \exp(\epsilon^2 n/2)$. Then $\exists\, u \in O(n)$, s.t.

$$f(ux_i) \underset{\epsilon}{\sim} L_f.$$

(Various functions and configurations provide various geometric consequences.)

Functional point of view. Let $\mathcal{N} = \{\mu_i\}_{i=1}^N$ be probability measures on S^{n+1}. Let $f(x)$ be a 1-Lip. function on S^{n+1}. Let $N \leq c \exp(\epsilon^2 n/8)$. Then

$$\exists\, u \in O(n) \text{ s.t. } \forall\, i = 1, \ldots, N,$$
$$\left| \int_{s^{n+1}} f(ux) d\mu_i(x) - L_f \right| < \epsilon.$$

(If $\mu_i(x)$ are δ-measures we return to the previous example.)

There are two ways extending the concept of concentration:

(i) For metric G-spaces (X, ρ) without any measures involved (Gromov-Milman, $\sim$1985);

(ii) Probability spaces (X, μ) without any metric involved (Giannopoulos-Milman, 2000).