Dimensional Brunn–Minkowski inequalities for the relative entropy

Gautam Aishwarya
(joint work with Liran Rotem)

The Technion- Israel Institute of Technology

October 19, 2023

Overview

- Entropy and geometric inequalities
- The dimensional Brunn–Minkowski problem for measures (Main question) and our approach
- Log-concave measures with homogeneous potential
- the Gaussian measure

The Brunn–Minkowski inequality

Theorem (Brunn–Minkowski)

Let $K_0, K_1 \subseteq \mathbb{R}^n$ be compact sets, then

$$\mathrm{Vol}_n((1-t)K_0 + tK_1)^{1/n} \geq (1-t)\mathrm{Vol}_n(K_0)^{1/n} + t\mathrm{Vol}_n(K_1)^{1/n}$$

Here $(1-t)K_0 + tK_1 = \{(1-t)x + ty : \ x \in K_0, \ y \in K_1\}$.

- Says that the Lebesgue measure is "$1/n$-concave".
- By scaling properties of volume, this is equivalent to the a priori weaker 0-concavity:
  $\mathrm{Vol}_n\left((1-t)K_0 + tK_1\right) \geq \mathrm{Vol}_n(K_0)^{1-t}\mathrm{Vol}_n(K_1)^t$.
- How to prove this using "entropy"?

Step 1: find a suitable notion of entropy to encode volume as maximum entropy

Definition

Let μ be a probability measure on $\mathbb{R}^n$ having density f w.r.t. the Lebesgue measure. Then, the *Shannon-Boltzmann entropy* is defined by

$$h(\mu) = -\int f \log f \, dx.$$

- If μ does not have density, can be meaningfully set to be $-\infty$.
- Sometimes we write $h(f)$ for $h(\mu)$, as well as $h(X)$, if X is a random vector with distribution μ.
- Thought of as a measure of how spread out μ is.

Lemma

Let $K \subseteq \mathbb{R}^n$ be compact, then

$$\sup_{\mu \in \mathcal{P}(K)} h(\mu) = \log \mathrm{Vol}_n(K).$$

Proof:

- If $X \sim \mu$ has density f,

$$h(\mu) = -\int f \log f \; dx = \mathbb{E} \log \frac{1}{f(X)} \leq \log \mathbb{E} \frac{1}{f(X)} \leq \log \mathrm{Vol}_n(K).$$

- Equality if $X \sim \mathrm{Uniform}(K)$.

Therefore, if we want to prove

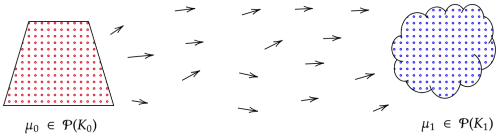
$$\boxed{\log \mathrm{Vol}_n\left((1-t)K_0 + tK_1\right) \geq (1-t)\log \mathrm{Vol}_n(K_0) + t\log \mathrm{Vol}_n(K_1),}$$

we can do the following:

For every $\mu_0 \in \mathcal{P}(K_0)$ and every $\mu_1 \in \mathcal{P}(K_1)$ find an interpolation $\mu_t \in \mathcal{P}\left((1-t)K_0 + tK_1\right)$, such that

$$\boxed{h(\mu_t) \geq (1-t)h(\mu_0) + th(\mu_1).}$$

Step 2: find a suitable interpolation to show concavity of entropy chosen earlier



- We think of probability distributions as distributions (or configurations) of the molecules of a gas in space.
- Look at all time-dependent vector fields v_t which displace μ_0 to μ_1 between $t = 0$ and $t = 1$. Each choice of v_t corresponds to a family μ_t representing the intermediate distribution at time t along the flow.
- Pick the one which minimises the action $\int_0^1 \int |v_t|^2 \, d\mu_t \, dt$, where μ_t denotes the intermediate distribution at time t along the flow induced by v_t.

- The interpolation $\{\mu_t\}$ corresponding to the action-minimising v_t is called the *displacement interpolation*.
- It turns out that the optimal vector field is of the form $\nabla\theta_t$, and displaces every molecule with constant speed along a straight line, leading to $\mu_t \in \mathcal{P}((1-t)K_0 + tK_1)$.

Two useful properties are:

- *Continuity equation*: For every test function ϕ,

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\int \phi \, \mathrm{d}\mu_t = \int \langle \nabla\phi, \nabla\theta_t\rangle \, \mathrm{d}\mu_t.}$$

- *The Hamilton–Jacobi equation*,

$$\boxed{\frac{\partial\theta_t}{\partial t} + \frac{|\nabla\theta_t|^2}{2} = 0.}$$

An equivalent way to describe things

Let $\mu_0, \mu_1 \in \mathcal{P}_{2,ac}(\mathbb{R}^n)$, and consider the *Monge problem* of minimising

$$\boxed{I = \int |x - T(x)|^2 \, \mathrm{d}\mu_0(x),}$$

over all maps T such that $T_{\#}\mu_0 = \mu_1$.

- (Brenier '91, McCann '97) The minimum is uniquely attained by a map $T = \nabla\psi$, where ψ is convex. The displacement interpolation is $\mu_t = T_{t\#}\mu_0$, where $T_t = (1-t)I + tT$.
- The time-dependent velocity field from before solves,

$$\nabla\theta_t(T_t(x)) = \frac{\mathrm{d}}{\mathrm{d}t} T_t(x).$$

Displacement concavity of entropy

Theorem (McCann '97)

The Shannon–Boltzmann entropy is displacement concave:

$$h(\mu_t) \geq (1-t)h(\mu_0) + th(\mu_1),$$

for all displacement interpolations.

Proof: Change of variables obtained by the description of μ_t as $T_{t\#}\mu_0$.

Corollary

For all compact $K_0, K_1 \subseteq \mathbb{R}^n$, we have

$$\mathrm{Vol}_n\left((1-t)K_0 + tK_1\right) \geq \mathrm{Vol}_n(K_0)^{1-t}\mathrm{Vol}_n(K_1)^t.$$

How to obtain the dimensional form directly?

- (McCann '97) The functional $\mu \mapsto \int \frac{d\mu}{dx}^{1-\frac{1}{n}}\, dx$ is displacement concave. Or,

Theorem (Erbar-Kuwada-Sturm, '15)

The functional $e^{h(\cdot)/n}$ is displacement concave:

$$e^{h(\mu_t)/n} \geq (1-t)e^{h(\mu_0)/n} + te^{h(\mu_1)/n}.$$

- While they prove something much^k more general, the argument specialised to obtain the above is very simple.
- In the language of random vectors, after some change of variables, this implies $e^{h(X+Y)/n} \geq e^{h(X)/n} + e^{h(Y)/n}$, when (X, Y) are coupled so that it minimise $\mathbb{E}|X-Y|^2$.
- Compare with the Entropy Power Inequality: $e^{2h(X+Y)/n} \geq e^{2h(X)/n} + e^{2h(Y)/n}$, when X, Y are independent.

Which problem do we want approach using these ideas?

The question of a-concavity of Borel measures ν on $\mathbb{R}^n$:

$$\nu\left((1-t)K_0 + tK_1\right) \geq \left((1-t)\nu(K_0)^a + t\nu(K_1)^a\right)^{1/a},$$

is relatively well understood.

- (Borell, '75) We know this is equivalent to b-concavity of the density of μ, where $\frac{1}{a} = \frac{1}{b} + n$.
- (EKS, '15) We know this is equivalent to $(0, 1/a)$-convexity of the relative entropy w.r.t. ν.

Broad question: Do concavity properties of a measure improve when K_0, K_1 are restricted to some sub-class of compact sets?
In particular: Can 0-concavity (i.e., log-concavity) for an even measure improve when restricted to sets with some convexity and symmetry properties?

Some developments on this theme

- (Livshyts-Marsiglietti-Nayar-Zvavitch, '17) The influential Log-Brunn–Minkowski conjecture of Böröczky, Lutwak, Yang and Zhang, implies that

$$\nu((1-t)K_0 + tK_1)^{1/n} \geq (1-t)\nu(K_0)^{1/n} + t\nu(K_1)^{1/n},$$

  holds for every even log-concave measure ν and symmetric convex bodies K_0, K_1.
- The $1/n$-concavity question for the Gaussian measure specifically was asked by Gardner and Zvavitch ('10) when $0 \in K_0 \cap K_1$. Nayar and Tkocz ('12) showed that this is not true, and suggested central symmetry may be the correct condition.

General strategy: (initiated by Kolesnikov-Milman, '18, '22)

- Try to prove $\frac{\mathrm{d}^2}{\mathrm{d}t^2}\left(\mu\left((1-t)K+tL\right)^a\right) \leq 0$ by directly computing this second derivative.
- Enough to check at $t = 0$, which gives a functional inequality for $f : \partial K \to \mathbb{R}$.
- Transform this inequality to an inequality for functions $u : K \to \mathbb{R}$ by taking u to be the solution of a certain elliptic PDE with f as its Neumann boundary condition.
- As far as we can tell, the convexity of K, L is crucial.

More developments

- (Kolesnikov-Livshyts, '21) Obtained $1/2n$-concavity, $\gamma_n((1-t)K_0 + tK_1)^{\frac{1}{2n}} \geq (1-t)\gamma_n(K_0)^{\frac{1}{2n}} + t\gamma_n(K_1)^{\frac{1}{2n}}$, for all convex sets $K_0, K_1 \subseteq \mathbb{R}^n$ containing 0.
- (Eskenazis-Moschidis, '21) Using a sufficiency criteria obtained in the above work, obtained $1/n$-concavity of the Gaussian measure when restricted to origin symmetric convex sets.
- (Cordero-Erasquin and Rotem, '23) $1/n$-concavity for Rotationally-invariant log-concave measures and symmetric convex sets
- (Livshyts, '21) $\frac{1}{n^{4+o(1)}}$-concavity for all even log-concave measures and symmetric convex sets.

Our Approach

Let ν be a log-concave measure on $\mathbb{R}^n$.

Step 1: Find the right notion of entropy to encode the ν-measure as maximum entropy.

Step 2: Prove the right displacement concavity property for this entropy.

Main difference: Differentiate entropy instead of differentiating the ν-measure of interpolating convex bodies.

Step 1: Relative entropy

Definition (Relative entropy)

Let ν be a σ-additive Borel measure on $\mathbb{R}^n$. We define the *relative entropy of μ with respect to ν* by,

$$D(\mu\|\nu) = \begin{cases} \int f \log f \, \mathrm{d}\nu, & \text{if } \mu \text{ has density } f \text{ w.r.t. } \nu, \\ +\infty, & \text{otherwise.} \end{cases}$$

- $D(\mu\|\nu)$ quantifies how much μ is "spread out" from the viewpoint of ν.
- We get an absolute measure of spread by looking at negative the amount μ is spread out from the most spread out measure Vol_n, $h(\mu) = -D(\mu\|\mathrm{Vol}_n)$.

Lemma

Let $K \subseteq \mathbb{R}^n$ be a compact set, then

$$\sup_{\mu \in \mathcal{P}(K)} e^{-D(\mu \| \nu)} = \nu(K)$$

.

Proof:

- If $X \sim \mu \in \mathcal{P}(K)$ has density f w.r.t. ν,

$$-D(\mu\|\nu) = -\int f \log f \, \mathrm{d}\nu = \mathbb{E} \log \frac{1}{f(X)} \leq \log \mathbb{E} \frac{1}{f(X)} \leq \log \nu(K).$$

- Equality if $\frac{\mathrm{d}\mu}{\mathrm{d}\nu}$ is constant, that is, if $\mu(\cdot) = \nu_K(\cdot) = \frac{\nu(\cdot \cap K)}{\nu(K)}$.

General method

- Suppose $S \subseteq \mathcal{P}_2(R^d)$ is a displacement convex set, that is, $\mu_0, \mu_1 \in S$ implies $\{\mu_t\} \subseteq S$.
- Let ν be a σ-finite Borel measure on $\mathbb{R}^n$. Assume that $\{\nu_K : K \in \mathcal{K}\} \subseteq S$, for some class $\mathcal{K}$ of compact sets.
- If $e^{-aD(\cdot\|\nu)}$ is displacement concave, $a > 0$, on S:

$$e^{-aD(\mu_t\|\nu)} \geq (1-t)e^{-aD(\mu_0\|\nu)} + te^{-aD(\mu_1\|\nu)};$$

then,

$$\nu((1-t)K_0 + tK_1)^a \geq (1-t)\nu(K_0)^a + t\nu(K_1)^a,$$

for all $K_0, K_1 \in \mathcal{K}$.

Idea 1: Decompose relative entropy, work on pieces

- Suppose $\nu = e^{-V}\,\mathrm{d}x$, V convex. (that is, ν is log-concave with potential V)
- $\boxed{D(\mu\|\nu) = -h(\mu) + \mathcal{V}(\mu)}$, where $\mathcal{V}(\mu) = \int V\,\mathrm{d}\mu$.
- Thus,

$$e^{-aD(\cdot\|\nu)} = \underbrace{e^{ah(\cdot)}}_{\text{know}}\underbrace{e^{-a\mathcal{V}(\cdot)}}_{\text{want}}.$$

- Concavity of $e^{-a\mathcal{V}(\cdot)}$ is equivalent to

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{V}(\mu_t) \geq a\left(\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}(\mu_t)\right)^2.$$

- These derivatives can be calculated in terms of the velocity-field $\nabla\theta_t$ associated with the displacement interpolation $\{\mu_t\}_{t\in[0,1]}$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}(\mu_t) = \int \langle \nabla V, \nabla\theta_t \rangle \,\mathrm{d}\mu_t,$$
$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{V}(\mu_t) = \int \langle \nabla^2 V \cdot \nabla\theta_t, \nabla\theta_t \rangle \,\mathrm{d}\mu_t.$$

- These quantities can be related by a Hölder-type inequality,

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}(\mu_t)\right)^2 = \left(\int \langle \nabla\theta_t, \nabla V \rangle \,\mathrm{d}\mu_t\right)^2$$
$$\leq \int \langle \nabla^2 V \cdot \nabla\theta_t, \nabla\theta_t \rangle \,\mathrm{d}\mu_t \cdot \int \langle \left(\nabla^2 V\right)^{-1} \cdot \nabla V, \nabla V \rangle \,\mathrm{d}\mu_t$$
$$= \left(\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{V}(\mu_t)\right) \cdot \int \langle \left(\nabla^2 V\right)^{-1} \cdot \nabla V, \nabla V \rangle \,\mathrm{d}\mu_t$$

- Therefore, displacement concavity of $e^{-a\mathcal{V}}$ on S can be obtained by showing $\int \langle \left(\nabla^2 V\right)^{-1} \cdot \nabla V, \nabla V \rangle \, \mathrm{d}\mu \leq \frac{1}{a}$ on S.
- When V is *p-homogeneous*, that is, when $V(\lambda x) = \lambda^p V(x)$ for all $\lambda > 0$, the quantity we need to bound becomes simpler allowing several computations.
- In this case, set $S = \{\mu : \mathcal{V}(\mu) \leq \frac{n}{p}\}$.
- This set is displacement concave and contains all measures whose density with respect to ν is radially decreasing.

Theorem (A-Rotem, '23+)

Let $V : \mathbb{R}^n \to [0, \infty)$ be a p-homogeneous convex function, $p \in (1, \infty)$. Let $\mathrm{d}\nu = e^{-V+c} \, \mathrm{d}x \in \mathcal{P}(\mathbb{R}^n)$, for some constant c.. Then, the functional $e^{-\frac{p-1}{n}\mathcal{V}(\cdot)}$, where $\mathcal{V}(\mu) = \int V \, \mathrm{d}\mu$, is displacement concave on S.

Putting things together

- So now we know $e^{\frac{h(\cdot)}{n}}$ and $e^{-\frac{p-1}{n}\mathcal{V}(\cdot)}$ are both displacement concave on S.
- $e^{-\frac{p-1}{pn}D(\cdot\|\nu)} = e^{\frac{p-1}{pn}c}\left(e^{\frac{h(\cdot)}{n}}\right)^{1-\frac{1}{p}}\left(e^{-\frac{p-1}{n}\mathcal{V}(\cdot)}\right)^{\frac{1}{p}}$ is a geometric-mean of concave functions, hence is itself concave.

Theorem (A-Rotem, '23+)

Let $V : \mathbb{R}^n \to [0,\infty)$ be a p-homogeneous convex function, $p \in (1,\infty)$. Let $\mathrm{d}\nu = e^{-V+c}\,\mathrm{d}x \in \mathcal{P}(\mathbb{R}^n)$, for some constant c. Then, the functional $e^{-\frac{p-1}{pn}D(\cdot\|\nu)}$ is displacement concave on the set S.

A dimensional Brunn–Minkowski for star bodies

Since every ν_K, for a star-shaped body K (i.e., $[0,1]K \subseteq K$), has a radially decreasing density with respect to ν, we have:

Theorem (A-Rotem, 23+)

Let $V : \mathbb{R}^n \to [0,\infty)$ be a p-homogeneous convex function for $1 < p < \infty$, and let $\mathrm{d}\nu = e^{-V}\,\mathrm{d}x$. Suppose $K_0, K_1 \subseteq \mathbb{R}^n$ are star bodies. Then for all $0 \le t \le 1$ we have

$$\nu\left((1-t)K_0 + tK_1\right)^{\frac{p-1}{pn}} \ge (1-t)\nu(K_0)^{\frac{p-1}{pn}} + t\nu\left(K_1\right)^{\frac{p-1}{pn}}.$$

Corollary

$\gamma_n\left((1-t)K_0 + tK_1\right)^{\frac{1}{2n}} \ge (1-t)\gamma_n(K_0)^{\frac{1}{2n}} + t\gamma_n(K_1)^{\frac{1}{2n}}$, *for all star-shaped compact sets $K_0, K_1 \subseteq \mathbb{R}^n$.*

Idea 2: work directly with the relative entropy

Formally, we can directly differentiate the relative entropy itself.

- Consider a measure ν having density e^{-V} with respect to the Lebesgue measure, and the Markov semigroup generated by $L = \Delta - \langle \nabla V, \nabla \rangle$.
- Denote by Γ, Γ_2 the carré du champ operator and its iteration, respectively, of this semigroup.

Then,

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} D(\mu_t \| \nu) &= - \int L\theta_t \, \mathrm{d}\mu_t, \\
\frac{\mathrm{d}^2}{\mathrm{d}t^2} D(\mu_t \| \nu) &= \int \Gamma_2(\theta_t) \, \mathrm{d}\mu_t,
\end{aligned}
$$

where $\nabla \theta_t$ is the velocity-field of $\{\mu_t\}$.

To obtain the displacement concavity $e^{-aD(\cdot\|\nu)}$ on a set S, we need to prove

$$\int \Gamma_2(\theta_t)\ \mathrm{d}\mu_t \geq a\left(\int L\theta_t\ \mathrm{d}\mu_t\right)^2.$$

- Using exponentiated relative entropy (as we do) is perhaps better than using Rényi entropy when working with a proper subset S.
- Remarkably similar to the Livshyts-Kolesnikov criteria for $\nu = \gamma_n$.

The Gaussian case

Let $S_n \subseteq \mathcal{P}_2(\mathbb{R}^n)$ be the collection of all even measures μ satisfying the Poincáre inequality

$$\int f^2 \, d\mu \leq \int |\nabla f|^2 \, d\mu,$$

for all **odd** test functions f.

Theorem (A-Rotem, '23+)

Suppose the displacment interpolation $\{\mu_t\}_{t\in[0,1]}$ *completely lies in* S_n*, then* $e^{-D(\mu_t\|\gamma_n)}$ *is concave in* t*. Moreover,* $S_1 \subseteq \mathcal{P}_2(\mathbb{R}^1)$ *is itself displacement convex, thus* $e^{-D(\cdot\|\gamma)}$ *is displacement concave on* S_1*.*

- The proof of the first part uses an idea very similar to the one used by Eskenazis and Moschidis.

What do we know about the contents of S_n

We know that S_n contains:

1. All even measures whose density is log-concave with respect to the standard Gaussian (that is, even strongly log-concave measures). This class corresponds to origin-symmetric convex bodies.
2. The displacement convex set consisting of all Gaussians with covariance dominated by the identity matrix.
3. All displacemet interpolation $\{\mu_t\}_{t\in[0,\infty]}$ such that one endpoint is the Gaussian, and the other endpoint is even and strong log-concave.

Open Questions

- Is $S_n \subseteq \mathcal{P}_2$ displacement convex in general?
- Does S_n contain a displacement convex set containing all even strongly log-concave probability measures?

THANK YOU!