Lecture Notes

Regularity through convexity in high dimensions

Boaz Klartag

Winter School
Institute Henry Poincaré
Paris, January 7–11, 2013

Notes taken by Silouanos Brazitikos

Contents

Chapter 1

Probability measures in high dimensions

In this first lecture we will present the first protagonist of the mini-course, the *high dimension*. Explanations regarding *convexity*, whose rôle cannot be diminished, will have to wait for the next lectures. We suggest that the students pay attention to our multiple points of view of the subject matter: The geometric perspective, the probabilistic perspective, and the analytic one. We begin by considering in detail some simple examples of probability measures in $\mathbb{R}^n$, trying to imagine how they look like geometrically when the dimension n approaches infinity. This last sentence might sound a bit confusing to some of the students: A probability measure may be defined on $\mathbb{R}^n$ for some n, but then how can the fixed dimension n tend to infinity? One standard answer is that in fact we consider a sequence of Borel probability measures μ_n defined on $\mathbb{R}^n$, for $n = 1, 2, \ldots$ and we investigate this sequence as a whole. Another standard answer, which we find more convenient in practice, is that we are looking after explicit bounds, depending on the dimension n alone, for interesting quantities related to a probability measure μ on $\mathbb{R}^n$. As the dimension gets larger, these bounds may sometimes get sharper. Thus, we obtain good approximations to various features of the measure μ, assuming that n is sufficiently large. A third possible answer could involve infinite-dimensional probability measures, a subject that will not be touched upon in this mini-course.

1.1 The example of the cube

Consider the unit cube $Q^n = [-1/2, 1/2]^n$. In order to warm up a bit, let us compute the diameter of the unit cube, that is

$$diam(Q^n) = \sup_{x,y \in Q^n} |x - y| \tag{1.1}$$

where $|x|^2 = \sum_i x_i^2$ is the square of the standard Euclidean norm of the vector $x \in \mathbb{R}^n$. A moment of reflection reveals that the supremum in (1.1) is attained when x and y are two antipodal vertices of the unit cube Q^n. Therefore

$$diam(Q^n) = \sqrt{n}.$$

In high dimensions, there are two lengthscales that are associated with the unit cube: Its side-length, which is one, and its diameter, which is $\sqrt{n}$. It appears that the latter lengthscale is frequently more dominant. Very roughly, the unit cube Q^n resembles a Euclidean ball of radius $\sqrt{n}$ more than it resembles a Euclidean ball of radius one.

For instance, let us compute the typical Euclidean distance between two random points in Q^n. Thus, let $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$ be two independent random vectors, that are distributed uniformly in the cube Q^n. Then $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ are independent real-valued random variables, distributed uniformly in the interval $[-1/2, 1/2]$. The L^2-average of the distance between X and Y is

$$\sqrt{\mathbb{E}|X-Y|^2} = \sqrt{\sum_{i=1}^n \mathbb{E}|X_i - Y_i|^2} = \sqrt{n \cdot \mathbb{E}|X_1 - Y_1|^2} = \frac{\sqrt{n}}{\sqrt{6}}.$$

Thus the L^2-norm of the distance between the random points X and Y has the order of magnitude of $\sqrt{n}$, for large n. It is also not too hard to show also that the expectation of the distance, the median of the distance and also the L^p-norm of the distance have the order of magnitude of $\sqrt{n}$. This will also follow from a general principle, to be discussed later in this mini-course. We conclude that in any reasonable sense, the typical distance between two random points in the cube has the order of magnitude of $\sqrt{n}$.

An important feature of the uniform distribution on the cube Q^n is related to the classical *central limit theorem*. When X is a random vector that is distributed uniformly in Q^n, its coordinates $X_1, X_2, \ldots, X_n$ are independent, identically-distributed random variables of mean zero and variance $1/12$. Let us dilate the cube Q^n by a factor of $\sqrt{12}$ in order to obtain variance 1. Thus, denote

$$\sqrt{12}Q^n = \left\{ \sqrt{12}x \, ; \, x \in Q^n \right\} = [-\sqrt{3}, \sqrt{3}]^n.$$

Suppose that $Y = (Y_1, \ldots, Y_n)$ is a random vector, distributed uniformly in $\sqrt{12}Q^n$. According to the central limit theorem, the random variable

$$\frac{\sum_{i=1}^n Y_i}{\sqrt{n}} = \sum_{i=1}^n \frac{1}{\sqrt{n}} \cdot Y_i$$

is approximately a standard, normal random variable. Furthermore, it is not essential that all of the weights are precisely $1/\sqrt{n}$. In place of the $1/\sqrt{n}$, we may use coefficients $\theta_1, \ldots, \theta_n \in \mathbb{R}$ with $\sum_i \theta_i^2 = 1$, and consider the random variable

$$\sum_{i=1}^n \theta_i Y_i.$$

The latter random variable has mean zero and variance one, according to our normalization of the θ_i's. By using standard expansions related to the central limit theorem, such as the Berry-Esseen bound, we have

$$\forall t \in \mathbb{R}, \qquad \left| \mathbb{P}\left(\sum_{i=1}^n \theta_i Y_i \leq t \right) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-s^2/2} ds \right| \leq C \sum_{i=1}^n \theta_i^4, \tag{1.2}$$

where $C > 0$ is a universal constant. That is, the constant C does not depend on t or on θ_i, it is simply a numerical constant such as 5 or 10 or $2\pi e$ which in principle can be written explicitly, if one is interested in its value. When substituting $\theta_i = 1/\sqrt{n}$ in (1.2), we see that in the case of the cube, the error in the central limit theorem has the order of magnitude of at most $1/n$. We also see that the Gaussian approximation is pretty good when all of the θ_i's are rather small, since the right-hand side of (1.2) is at most $C \max_i |\theta_i|^2$.

It is perhaps helpful to try and visualize the central limit theorem in a more geometric fashion. One may think about estimates such as (1.2) in the following way: Fix a unit vector of coefficients $\theta = (\theta_1, \ldots, \theta_n) \in \mathbb{R}^n$, and let f_θ be the probability density function of the random variable $\sum_i \theta_i X_i$, where X is distributed uniformly in the unit cube Q^n. Denote

$$H_{\theta,t} = \{x \in \mathbb{R}^n; x \cdot \theta = t\} \qquad\qquad (t \in \mathbb{R}),$$

a hyperplane orthogonal to the direction θ. Here, $x \cdot y = \sum_i x_i y_i$ is the standard scalar product. Consider the intersection of the unit cube with H_t and observe that

$$f_\theta(t) = Vol_{n-1}(Q^n \cap H_{\theta,t}).$$

In the case where $\theta = (1, \ldots, 1)/\sqrt{n}$, and also in many other cases, we can assert that

$$\left| f_\theta(\sqrt{12} \cdot t) - \frac{\exp(-t^2/2)}{\sqrt{2\pi}} \right| \leq \frac{C}{n} \qquad\qquad (t \in \mathbb{R}), \tag{1.3}$$

for a universal constant $C > 0$. The bound (1.3) does not formally follow from (1.2), yet it is very similar to (1.2), and its standard proof is outlined in Exercise 1.1.1 below. We conclude that to a very good approximation, the Gaussian law dictates the volumes of parallel hyperplane slices of the high-dimensional cube.

In the case where $(\theta_1, \ldots, \theta_n) = (1, 0, \ldots, 0)$, the random variable $\sum_i \theta_i X_i$ is not at all close to Gaussian: It is distributed uniformly in an interval. Thus, the Gaussian approximation does not hold for *all* directions θ. Yet, in sense to be made precise, it holds true for *most* directions θ.

Exercise 1.1.1 Let $n \geq 2,\ \theta = (1, \ldots, 1)/\sqrt{n}$, suppose that X is distributed uniformly in Q^n and denote by f_θ the probability density function of the random variable $X \cdot \theta$.

(i) Verify that $\widehat{f_\theta}(s) := \mathbb{E}\exp(2\pi i s \theta \cdot X) = \left(\frac{\sin(\pi s/\sqrt{n})}{\pi s/\sqrt{n}} \right)^n$ for any $s \in \mathbb{R}$.

(ii) Use the Taylor expansion in order to prove that for $|r| < 1/2$,

$$\log \frac{\sin(\pi r)}{\pi r} = -\frac{\pi^2 r^2}{6} + O\left(r^4\right).$$

The notation $O(x)$, for some expression x, is an abbreviation for some complicated quantity y with the property that $|y| \leq Cx$ for a universal constant $C > 0$.

(iii) Conclude that for any $|s| \leq n^{1/4}/2$,

$$\left| \widehat{f_\theta}(s) - \exp\left(-\pi^2 s^2/6\right) \right| \leq C \frac{s^4 \exp\left(-s^2/2\right)}{n}. \tag{1.4}$$

(iv) Write down the Fourier inversion formula for the function f_θ, and recall the formula for the Fourier transform of a Gaussian function. Pretend that (1.4) holds true for all s (this is false, will be corrected in the next step), and deduce the desired estimate (1.3).

(v) In order to prove that the contribution of $|s| > n^{1/4}/2$ is indeed negligible, you may argue as follows: Show that for $|r| \geq n^{-1/4}/2$,

$$\left| \frac{\sin(\pi r)}{\pi r} \right| \leq \min\left\{ 1 - \frac{1}{C\sqrt{n}}, \frac{1}{\pi |r|} \right\}$$

and therefore,

$$\int_{\frac{n^{1/4}}{2}}^{\infty} \left| \widehat{f_\theta}(s) \right| ds \leq \int_0^\infty \min\left\{ \left(1 - \frac{1}{C\sqrt{n}}\right)^n, \left(\frac{1}{\pi s/\sqrt{n}}\right)^n \right\} ds = O\left(\frac{1}{n^{10}}\right).$$

1.2 The example of the Euclidean ball

The next example we will consider in detail is that of the Euclidean unit ball

$$B^n = \{x \in \mathbb{R}^n; |x| \leq 1\}.$$

The volume of B^n is computed in advanced Calculus classes around the globe in various ways, and it always turns out that

$$Vol_n\left(B^n\right) = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)} = \left(\frac{\sqrt{2\pi e}+o(1)}{\sqrt{n}}\right)^n \tag{1.5}$$

where the equality on the right-hand side follows from Stirling's formula. In this mini-course, the notation $o(x)$, for a certain expression x, is an abbreviation for some complicated quantity y with the property that as the dimension n tends to infinity, the ratio y/x tends to zero. We conclude from (1.5) that the volume of the unit ball is rather small, in high dimensions. It is the Euclidean ball whose radius is approximately

$$\frac{\sqrt{n}}{\sqrt{2\pi e}+o(1)}$$

which has volume one. In order not to carry the cumbersome factor "$\sqrt{2\pi e}+o(1)$" throughout the discussion, we choose to consider the Euclidean ball of radius $\sqrt{n}$. Thus, suppose that $X = (X_1, \ldots, X_n)$ is a random vector in $\mathbb{R}^n$ that is distributed uniformly in $\sqrt{n}B_2^n$. Unlike in the case of the cube, this time the random variables $X_1, \ldots, X_n$ are no longer independent. Nevertheless, let us still investigate the distribution of the random variable

$$\sum_{i=1}^n \theta_i X_i$$

for some unit vector $\theta \in \mathbb{R}^n$. Geometrically, this means that we select a unit vector θ, we intersect our Euclidean ball with hyperplanes orthogonal to θ and then we compute their volume. The probability density of $\sum_{i=1}^n \theta_i X_i$ is the function

$$f_\theta(t) = \frac{Vol_{n-1}\left\{\sqrt{n}B_2^n \cap H_{\theta,t}\right\}}{Vol_n\left\{\sqrt{n}B_2^n\right\}} \qquad (t \in \mathbb{R}). \tag{1.6}$$

The denominator in (1.6) is the price we have to pay for our choice to consider a Euclidean ball whose volume is not exactly one. A first observation, is that $f_\theta(t)$ does not depend on the unit vector θ, because of the symmetries of the Euclidean ball. A second observation is that the numerator in (1.6) is the volume of an $(n-1)$-dimensional Euclidean ball of radius $\sqrt{n-t^2}$, by the Pythagoras theorem. Consequently,

$$f_\theta(t) = \frac{Vol_{n-1}\left(B_2^{n-1}\right)}{Vol_n\left(\sqrt{n}B_2^n\right)} \cdot \left(n-t^2\right)^{(n-1)/2} = C_n\left(1-\frac{t^2}{n}\right)^{\frac{n-1}{2}} \qquad \text{for } |t| \leq \sqrt{n}, \tag{1.7}$$

where C_n is some constant depending solely on n. Let us look closely at the right-hand side of (1.7). What is it like for large n? As we all learned in Calculus, we have $f_\theta(t) \approx C_n \exp(-t^2/2)$, which is approximately the density of the standard Gaussian. More precisely, we can estimate the error in the approximation in a standard manner by using Taylor's theorem (Exercise!), and obtain

$$f_\theta(t) = C_n \left[e^{-t^2/2} + O\left(\frac{1}{n}\right) \right].$$

Recalling that f_θ is a probability density with integral one, we may now guess and prove that $C_n = (2\pi)^{-1/2} + O(1/n)$. To conclude,

$$f_\theta(t) = \frac{e^{-t^2/2}}{\sqrt{2\pi}} + O\left(\frac{1}{n}\right) \qquad (t \in \mathbb{R}).$$

Hence also in the case of the high-dimensional Euclidean ball, the Gaussian law approximately governs the behavior of volumes of parallel hyperplane slices. The latter fact is sometimes attributed to Maxwell.

Students who solved Exercise 1.1.1 will surely appreciate the simple, direct argument that we used for the proof of the central limit theorem for the uniform distribution on the Euclidean ball. Unlike in the case of the cube, we did not have to rely on very indirect tools such as the Fourier transform. Therefore, one may argue that the Euclidean ball is very much related to Gaussian approximations and to the central limit theorem, albeit the random variables $X_1, \ldots, X_n$ are dependent.

The boundary of B_2^n is the unit sphere

$$S^{n-1} = \{x \in \mathbb{R}^n\,;\, |x| = 1\}.$$

Next, we observe that in high dimensions, a random point in B_2^n is typically located rather close to the boundary. To that end, suppose that $Y = (Y_1, \ldots, Y_n)$ is a random vector, distributed uniformly in the unit ball B_2^n. Note that $0 \leq |Y| \leq 1$ almost surely, and

$$\mathbb{P}\left(|Y| \leq t\right) = Vol_n(tB_2^n)/Vol(B_2^n) = t^n \qquad (0 \leq t \leq 1).$$

In particular,

$$\mathbb{P}\left(|Y| \geq 1 - \frac{1}{n}\right) = 1 - \left(1 - \frac{1}{n}\right)^n \geq 1 - \frac{1}{e} > \frac{1}{2}.$$

Therefore at least half of the mass of B_2^n lies inside the *thin spherical shell*,

$$B_2^n \setminus \left(1 - \frac{1}{n}\right) B_2^n.$$

Yet another equivalent formulation, is that a random point in B_2^n is typically located at distance at most $1/n$ from the boundary of B_2^n.

In the remainder of this subsection (and also in the next one), we will discuss the measure σ_{n-1}, which is the surface area measure on the unit sphere S^{n-1}, normalized to be a probability measure. We refer to σ_{n-1} as the uniform probability measure on S^{n-1}. Its most important property is that it is invariant under $O(n)$, the group of orthogonal transformation in $\mathbb{R}^n$. Here we have the following exercises related to the above.

Exercise 1.2.1 *Let X and Y be independent random vectors supported on the unit sphere, with Y being distributed uniformly in S^{n-1}. Then the random variables*

$$X \cdot Y \qquad \textit{and} \qquad Y_1$$

have exactly the same distribution.

Exercise 1.2.2 *Suppose that Let $X = (X_1, \cdots, X_n)$ is a random vector uniformly distributed S^{n-1}. Then $(X_1, \cdots X_{n-2})$ is uniformly distributed in B^{n-2}.*

Hint: The density of $(X_1, \cdots X_{n-1})$ is proportional to $\dfrac{1}{\sqrt{1-|x|^2}}$ in B^{n-1}.

In effect, the uniform probability distribution on the Euclidean ball B_2^n is not so much different from the uniform probability measure on the unit sphere S^{n-1}, in high dimensions. More precisely, let be a random vector uniformly distributed on S^{n-1}. Then using exercise 1.2.2 we get that

$$\left| \mathbb{P}\left(\sqrt{n}X \cdot \theta \le t\right) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-s^2/2}\, ds \right| \le \frac{c}{\text{"}n-2\text{"}} \le \frac{\tilde{c}}{n}.$$

Additionally, the density of $\sqrt{n}X$ is $c_n \left(1 - \frac{t^2}{n}\right)^{\frac{n-3}{2}}$, so this vector is approximately Gaussian and finally for $|t| \le \sqrt{n}$ we get the following large deviation inequality

$$\begin{aligned}
\mathbb{P}\left(\left|\sqrt{n}X\right| \ge t\right) &= c_n \int_t^{\sqrt{n}} \left(1 - \frac{t^2}{n}\right)^{\frac{n-3}{2}} ds \\
&\le c \int_t^{\infty} e^{-s^2 \frac{n-3}{2n}}\, ds \\
&\le \tilde{c} e^{-t^2 - \frac{3t^2}{2n}} \le \overline{c} e^{-t^2/2}.
\end{aligned}$$

Renormalising, we get that for each $t > 0$ the following holds

$$\mathbb{P}\left(|X_i| \ge t\right) \le Ce^{-nt^2/2}. \tag{1.8}$$

One surprising consequence of (**??**) is that most of the mass of the high-dimensional unit sphere S^{n-1} lies very close to the equator. What is the probability that a random point Y in the sphere will be outside a narrow strip of width $1/10$ around the equator $\{x \in S^{n-1}; x_1 = 0\}$? This is precisely,

$$\mathbb{P}\left(|Y_1| \ge \frac{1}{10}\right) \le Ce^{-cn} \tag{1.9}$$

according to (**??**). Thus, only exponentially-small amount of mass is located at distance $1/10$ from the equator! Furthermore, by the symmetries of the sphere, the same phenomenon is observed for all of the equators. This strange high-dimensional effect is usually referred to as the *concentration of measure phenomenon*. Perhaps the circle is not a very good way to depict the high-dimensional sphere on paper, since such a drawing does not reflect the true distribution of mass. I learned from V. Milman that in order to have a good grasp on the geometry of the high-dimensional sphere, perhaps you might want to imagine it like that:

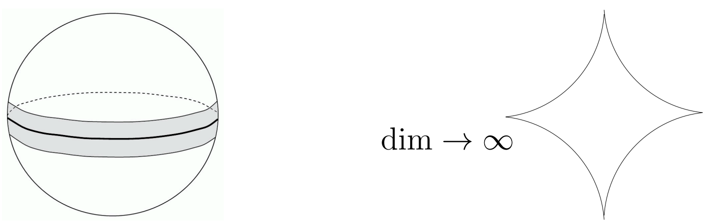

Figure 2.1 – "Concentration of Measure"

1.3 Concentration of measure on the sphere

The concentration of measure phenomenon is one of the most powerful effects in high dimension. The sphere S^{n-1} is a convenient setting for the demonstration of the effect, because the *isoperimetric problem* has an explicit, simple solution here. What is the isoperimetric problem? For $x, y \in S^{n-1}$ we write $d(x, y)$ for the geodesic distance between x and y on the sphere, hence $\cos d(x, y) = x \cdot y$. Note that

$$|x - y| \leq d(x, y) \leq \frac{\pi}{2}|x - y|$$

for any two points $x, y \in S^{n-1}$. Thus, the geodesic distance on the sphere and the Euclidean "tunnel distance" are comparable up to a constant factor. Since we tend to view the sphere "extrinisically", as a subset of $\mathbb{R}^n$, we would stick to the Euclidean distance $|x - y|$. The entire argument works equally well when one instead uses the geodesic distance $d(x, y)$. For a compact subset $A \subset S^{n-1}$ and $\varepsilon > 0$ we denote

$$A_\varepsilon = \left\{x \in S^{n-1}; \exists y \in A, |x - y| \leq \varepsilon\right\},$$

the ε-neighborhood of the set A. For example, when $H = \{x \in S^{n-1}; x_1 \leq 0\}$ is a hemisphere, then,

$$H_\varepsilon = \left\{x \in S^{n-1}\,;\, x_1 \leq \varepsilon\right\} \qquad (0 \leq \varepsilon \leq 1).$$

Clearly, $\sigma_{n-1}(H) = 1/2$. According to (**??**), the measure of the ε-neighborhood of the hemisphere is typically very close to one, that is

$$\sigma_{n-1}(H_\varepsilon) \geq 1 - C\exp(-\tilde{c}\varepsilon^2 n) \qquad (\varepsilon > 0). \tag{1.10}$$

What happens if one consider others subsets $A \subset S^{n-1}$ with measure $1/2$, in place of the hemisphere? The *isoperimetric inequality on* S^{n-1} states that among all subsets of the sphere of measure $1/2$, the hemisphere is the subset that has the minimal ε-neighborhood in terms of measure.

Theorem 1.3.1 *For any Borel set $A \subset S^{n-1}$ and $\varepsilon > 0$,*

$$\sigma_{n-1}(A) \geq 1/2 \qquad \Longrightarrow \qquad \sigma_{n-1}(A_\varepsilon) \geq \sigma_{n-1}(H_\varepsilon),$$

where $H = \{x \in S^{n-1}; x_1 \leq 0\}$ is a hemisphere.

Proof: We fix $\varepsilon > 0$. We will prove first that it is enough to restrict attention to closed subsets $A \subset S^{n-1}$ of measure exactly $1/2$. To prove this we need two claims.

Claim 1.3.2 *The set*

$$\mathcal{A} = \left\{\sigma_{n-1}(A_\varepsilon);\ A \subset S^{n-1} \textit{ Borel and } \sigma_{n-1}(A) \geq \frac{1}{2}\right\}$$

attains minimum.

Explanation: The space of the closed subsets of S^{n-1} is compact with regard to Haussdorff metric

$$d_H(A, B) = \inf\left\{\varepsilon > 0\,;\ A \subset B_\varepsilon \text{ and } B \subset A_\varepsilon\right\}.$$

Now it's easy easy to see that the map $A \mapsto \sigma_{n-1}(A)$ is upper semi-continuous, hence the set

$$\left\{A \subset S^{n-1} \,:\, \sigma_{n-1}(A) \geq \frac{1}{2}\right\}$$

is compact. Finally the map $A \mapsto \sigma_{n-1}(A_\varepsilon)$ is continuous so attains its minimum and we are done.

$\diamondsuit$

Claim 1.3.3 *There exist a minimizer of the set $\mathcal{A}$ with maximal intersection with the southern hemisphere H.*

Explanation: The set of all minimizers of $\mathcal{A}$ is non-empty due to Claim 1.3.2 and compact. Now, the map $A \mapsto \sigma_{n-1}(A \cap H)$ is upper semi-continuous so there exists minimizer with the desired property.

$\diamondsuit$

To summarize, we can take $A \subset S^{n-1}$ with $\sigma_{n-1}(A) = \frac{1}{2}$, minimal $\sigma_{n-1}(A_\varepsilon)$ and maximal intersection with H. The rest of the proof is based on the idea of "polarization" or "two-point-symmetrization" of A. To begin with, for $\theta \in S^{n-1}$ we set

$$\pi_\theta(x) = x - 2(x \cdot \theta)\theta$$

the reflection operator. The main step of the proof is the following claim.

Claim 1.3.4 *We fix $\theta \in S^{n-1}$ with $\theta_1 > 0$ and we take a subset E of the set*

$$A \cap \left\{x \in S^{n-1},\ x_1 > 0,\ x \cdot \theta > 0,\ \pi_\theta(x) \in H\right\}$$

with positive measure. Then the set

$$\pi_\theta(E) \cap A$$

is non-empty.

Explanation: First, we let

$$T_A(x) = \begin{cases} \pi_\theta(x) & \text{if } x \cdot \theta > 0 \text{ and } \pi_\theta(x) \notin A \\ x & \text{otherwise} \end{cases}$$

and we define the 2-point symmetrization as the set

$$S_\theta(A) = \{T_A(x) :\ x \in A\}.$$

Equivalently, if $A^+ = \{x \in A \,:\, x \cdot \theta > 0\}$ and $A^- = \mathbb{R}^n \setminus A^+$, then

$$A = A^- \cup \left(A^+ \cap \pi_\theta(A^-)\right) \cup \left(A^+ \setminus \pi_\theta(A^-)\right)$$

and

$$S_\theta(A) = A^- \cup \left(A^+ \cap \pi_\theta(A^-)\right) \cup \pi_\theta\left(A^+ \setminus \pi_\theta(A^-)\right).$$

We list three properties of the 2-point symmetrization.

- $\sigma_{n-1}\left(S_\theta(A)\right) = \sigma_{n-1}(A)$

- If $A \subset B$ then $S_\theta(A) \subset S_\theta(B)$
- $S_\theta(A)_\varepsilon \subset S_\theta(A_\varepsilon)$
- $\sigma_{n-1}\left(S_\theta(A) \cap H\right) \geq \sigma_{n-1}(A \cap H) + \sigma_{n-1}\left(\pi_\theta(E) \setminus A\right)$

The first two are almost trivial. To prove the third one, we take $x \in S_\theta(A)$. We need to show that $B(x,\varepsilon) \subset S_\theta(A_\varepsilon)$. To prove this we have to consider three cases.

1. If $x \in A$ and $\pi_\theta(x) \in A$ then $B(x,\varepsilon) \cup B(\pi_\theta(x),\varepsilon) \subset A_\varepsilon$, so $B(x,\varepsilon) \cup B(\pi_\theta(x),\varepsilon) \subset S_\theta(A_\varepsilon)$.
2. If $x \in A$ and $\pi_\theta(x) \notin A$ then $x \cdot \theta < 0$ and $B(x,\varepsilon) \subset A_\varepsilon$ so $B(x,\varepsilon) \subset S_\theta(A_\varepsilon)$.
3. If $x \notin A$ and $\pi_\theta(x) \in A$ then $x \cdot \theta < 0$ and $B(\pi_\theta(x),\varepsilon) \subset A_\varepsilon$ so $B(x,\varepsilon) \subset S_\theta(A_\varepsilon)$.

Now to prove the fourth property, we will show that $S_\theta(A) \cap H$ contains the disjoint union of and $\pi_\theta(E) \setminus A$ and $A \cap H$. If $x \in A \cap H$ then $T_A(x) \in H$ and $T_A(x)_1 \leq x_1$ so we get the first inclusion and if $x \in \pi_\theta(E) \setminus A$ then from the definition $\pi_\theta(E) \setminus A \subset S_\theta(A)$ and we get the desired. To summarize, we proved above that $S_\theta(A)$ has measure $\frac{1}{2}$ (equal to measure of A), it has not larger $\varepsilon-$enlargment and not smaller intersection with H. By the minimality of A we conclude that $\sigma_{n-1}\left(\pi_\theta(E) \setminus A\right) = 0$ and so $\pi_\theta(E) \cap A \neq \emptyset$. To finish the proof we will prove that

$$H \subset A.$$

Suppose the opposite, then there exist $\delta > 0$ and x_0, y_0 such that $B(x_0,\delta) \subset H^c \setminus A$ and $B(y_0,\delta) \subset H$. Now, take $\theta \in S^{n-1}$ with $\pi_\theta(x_0) = x_0$ and $\theta_1 > 0$. Then, if

$$E = A \cap B(y_0,\delta) \subset A \cap \left\{x \in S^{n-1},\ x_1 > 0,\ x \cdot \theta > 0,\ \pi_\theta(x) \in H\right\}$$

then E has positive measure and so from Claim 1.3.4, $\pi_\theta(E) \cap A \neq \emptyset$. But $B(x_0,\delta)$ contains $\pi_\theta(E)$ so it is not disjoint from A and this is a contradiction.

■

The proof of Theorem 1.3.1 may be generalized to other types of measure-metric spaces (that is, metric spaces with a measure) with lots of symmetries. By combining Theorem 1.3.1 with the concentration effect (1.10), we see that for any Borel set $A \subset S^{n-1}$ and $0 < \varepsilon < 1$,

$$\sigma_{n-1}(A) \geq 1/2 \qquad \Longrightarrow \qquad \sigma_{n-1}(A_\varepsilon) \geq 1 - C\exp(-c\varepsilon^2 n). \tag{1.11}$$

Thus, for any $A \subset S^{n-1}$ of measure $1/2$, the ε-neighborhood of A captures almost the entire sphere in sense of volume! For a real-valued function f we write $\{f > t\} = \{x; f(x) > t\}$.

Corollary 1.3.5 *Let $f : S^{n-1} \to \mathbb{R}$ be a 1-Lipschitz function (that is, $|f(x) - f(y)| \leq |x - y|$). Let M be the median of f, so that $\sigma_{n-1}\{f \geq M\} \geq 1/2$ and $\sigma_{n-1}\{f \leq M\} \geq 1/2$. Then,*

$$\sigma_{n-1}\left(\{x \in S^{n-1}\,;\ |f(x) - M| \geq t\}\right) \leq C\exp(-ct^2 n) \qquad (t > 0), \tag{1.12}$$

where $C, c > 0$ are universal constants.

Proof: Denote $A = \{f \leq M\}$ and $B = \{f \geq M\}$. Observe that $A_t \subseteq \{f \leq M + t\}$ and $B_t \subseteq \{f \geq M - t\}$. Now apply (1.11) and obtain the bound (1.13). $\square$

Corollary 1.3.5 roughly states that 1-Lipschitz functions on the high-dimensional sphere are effectively constant. Apriori, you would have expected that such a function attains values, say, in the entire interval $[0, 1]$. However, when one evaluates a 1-Lipschitz function at, say, five randomly selected points, the typical answer will be five numbers that are very close to one another.

It is sometimes more convenient to work with the expectation rather than the median. Fortunately, it is possible to replace M in Corollary 1.3.5 by $\int_{S^{n-1}} f d\sigma_{n-1}$.

Proposition 1.3.6 *Let $f : S^{n-1} \to \mathbb{R}$ be an L-Lipschitz function. Let $E = \int_{S^{n-1}} f d\sigma_{n-1}$. Then,*

$$\sigma_{n-1}\left(\{x \in S^{n-1} \,;\, |f(x) - E| \geq t\}\right) \leq C \exp(-ct^2 n/L^2) \qquad (t > 0), \tag{1.13}$$

where $C, c > 0$ are universal constants.

Proof: Replacing f by f/L, we may restrict attention to the case where $L = 1$. We claim that the expectation is rather close to the median. In fact, by Corollary 1.3.5,

$$\begin{aligned} |E - M| \leq \int_{S^{n-1}} |f - M| \, d\sigma_{n-1} &= \int_0^\infty \sigma_{n-1}\left(\{|f - M| \geq t\}\right) dt \\ &\leq \int_0^\infty C e^{-ct^2 n} dt \leq \frac{\tilde{C}}{\sqrt{n}}. \end{aligned}$$

Here, we used the well-known identity, that for any non-negative random variable X,

$$\mathbb{E}X = \mathbb{E}\left(\int_0^\infty 1_{\{t \leq X\}} dt\right) = \int_0^\infty \mathbb{P}(X \geq t) dt,$$

where $1_{\{t \leq X\}}$ is one when $t \leq X$ and vanishes otherwise. The desired inequality (1.13) holds trivially for $t \leq 1/\sqrt{n}$, since the left-hand side of (1.13) is at most one. When $t \geq 1/\sqrt{n}$, we can may our bound for $|E - M|$ and assert that

$$\{x \in S^{n-1} \,;\, |f(x) - E| \geq t\} \subseteq \{x \in S^{n-1} \,;\, |f(x) - M| \geq Ct\}$$

for some universal constant $C > 0$. The bound (1.13) now follows from Corollary 1.3.5. $\square$

Open problem: The analogous isoperimetric problem on $\mathbb{R}P^n$ remains open for all $n \geq 4$.

1.4 The thin-shell theorem

We move on to discuss an application of the concentration of measure phenomena, which is one of the many versions of the "thin-shell theorem". To a certain extent, this theorem explains why the Gaussian distribution appears in the central limit theorem.

Let $X = (X_1, \ldots, X_n)$ be a random vector in $\mathbb{R}^n$ with $\mathbb{E}|X|^2 < \infty$. We assume that X is normalized as follows:

$$\mathbb{E}X_i = 0, \quad \mathbb{E}X_i X_j = \delta_{i,j} \qquad \forall i, j = 1, \ldots, n. \tag{1.14}$$

Equivalently, all of the one-dimensional marginals of X have mean zero and variance one. Here, a one-dimensional marginal of X means a random variable of the form $X \cdot \theta$ for some $\theta \in S^{n-1}$. A random vector that satisfies the normalization condition (1.14) will be called "normalized" or "isotropic".

Exercise 1.4.1 Suppose that X is a random vector in $\mathbb{R}^n$ with $\mathbb{E}|X|^2 < \infty$. Assume that X is not supported in a hyperplane. Prove that there exists a vector $b \in \mathbb{R}^n$ and a positive-definite matrix A such that $A(X) + b$ is isotropic.

Exercise 1.4.1 demonstrates that isotropicity is just a matter of normalization. Beginning with any random vector X satisfying certain mild assumptions, we only need to translate it and stretch or shrink it in certain directions in order to transform X into an isotropic random vector.

It turns out that the crucial property of X in the context of Gaussian marginals is a certain *thin spherical shell* bound:

Theorem 1.4.2 *Let X be an isotropic random vector in $\mathbb{R}^n$ and let $\varepsilon > 0$. Assume that*

$$\mathbb{E}\left(\frac{|X|}{\sqrt{n}} - 1\right)^2 \leq \varepsilon^2. \tag{1.15}$$

Then, there exists a subset $\Theta \subseteq S^{n-1}$ with $\sigma_{n-1}(\Theta) \geq 1 - C\exp(-c\sqrt{n})$, such that for any $\theta \in \Theta$ and $t \in \mathbb{R}$,

$$|\mathbb{P}(X \cdot \theta \leq t) - \Phi(t)| \leq C\left(\varepsilon^{1/2} + \frac{1}{n^{1/8}}\right) \tag{1.16}$$

where $\Phi(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^t \exp\left(-s^2/2\right) ds$ and $C, c > 0$ are universal constants.

What is the meaning of condition (1.15)? Via the Chebyshev-Markov inequality this condition implies that

$$\mathbb{P}\left(1 - \sqrt{\varepsilon} \leq \frac{|X|}{\sqrt{n}} \leq 1 + \sqrt{\varepsilon}\right) \geq 1 - \varepsilon.$$

Thus, when $\varepsilon \ll 1$, the condition (1.15) implies that X is concentrated in a *thin spherical shell*.

Theorem 1.4.2 tells us that in order to have many approximately Gaussian marginals, it suffices to verify that most of the mass of the random vector X is contained in a thin spherical shell, whose width is much smaller than its radius. The fact that the radius must be $\sqrt{n}$ is dictated by our isotropic normalization of X.

From the proof of Theorem 1.4.2 one may learn that the thin-shell condition (1.15) is also necessary for the phenomenon of Gaussian approximation. Consider the case where $X = (X_1, \ldots, X_n)$ with $X_1, \ldots, X_n$ being independent random variables with, say, $\mathbb{E}X_i^4 \leq 100$ for all i. The thin-shell condition (1.15) holds true with a rather small ε. Indeed, we may compute that

$$\begin{aligned}\mathbb{E}\left(\frac{|X|}{\sqrt{n}} - 1\right)^2 &\leq \mathbb{E}\left(\frac{|X|^2}{n} - 1\right)^2 = Var\left(\frac{|X|^2}{n}\right) = \sum_{i=1}^n Var\left(\frac{X_i^2}{n}\right) \\ &= \frac{1}{n^2}\sum_{i=1}^n \left[\mathbb{E}X_i^4 - 1\right] \leq \frac{100}{n}.\end{aligned}$$

Thus the standard deviation of $|X|/\sqrt{n}$ is at most $10/\sqrt{n}$, and (1.15) holds true with $\varepsilon = O(n^{-1/2})$. Theorem 1.4.2 thus implies that many of the marginals of X are approximately Gaussian. Yet, our thin-shell theorem has two drawbacks in the context of independent random variables: First, it does not provide any information regarding one specific marginal, it only asserts that most of them are close to Gaussian. Second, the quantitative estimates we obtain do not quite match the sharp Berry-Esseen bound. While it is possible to squeeze a bit more from the proof given below and slightly improve the exponent "$1/8$" in (1.16), it is not entirely clear to us how to go below $1/4$. Nevertheless, Theorem 1.4.2 serves to demonstrate the general principle that approximately Gaussian marginals are related to the more geometric *thin shell* property.

The proof of Theorem 1.4.2 is a beautiful manifestation of the concentration of measure phenomenon. The main idea required for the proof of Theorem 1.4.2 appears in the following lemma:

Lemma 1.4.3 *Let X and ε be as in Theorem 1.4.2. Let $f : \mathbb{R} \to \mathbb{R}$ be an L-Lipschitz function and let Z be a standard, normal random variable. Then there exists a subset $\Theta \subseteq S^{n-1}$ with $\sigma_{n-1}(\Theta) \geq 1 - C\exp(-c\sqrt{n})$ such that*

$$\forall \theta \in \Theta, \qquad |\mathbb{E}f(X \cdot \theta) - \mathbb{E}f(Z)| \leq CL\left(\frac{1}{n^{1/4}} + \varepsilon\right). \tag{1.17}$$

Proof: Denote $F(\theta) = \mathbb{E}f(X \cdot \theta)$ for $\theta \in S^{n-1}$. First, we observe that F is an L-Lipschitz function on the sphere. Indeed, for any $\theta_1, \theta_2 \in S^{n-1}$,

$$\begin{aligned} |F(\theta_1) - F(\theta_2)| &\leq \mathbb{E}\,|f(X \cdot \theta_1) - f(X \cdot \theta_2)| \leq L\mathbb{E}\,|X \cdot (\theta_1 - \theta_2)| \\ &\leq L\sqrt{\mathbb{E}|X \cdot (\theta_1 - \theta_2)|^2} = L|\theta_1 - \theta_2|, \end{aligned}$$

since X is isotropic, and hence the random variable $X \cdot (\theta_1 - \theta_2)$ has variance $|\theta_1 - \theta_2|^2$. The function F is L-Lipschitz, hence it deviates very little from its average on the sphere. In particular, by using Proposition 1.3.6 with $t = L/n^{1/4}$, we deduce the existence of a subset $\Theta \subseteq S^{n-1}$ with $\sigma_{n-1}(\Theta) \geq 1 - C\exp(-c\sqrt{n})$ such that

$$\forall \theta \in \Theta, \qquad \left| F(\theta) - \int_{S^{n-1}} F(\theta) d\sigma_{n-1}(\theta) \right| \leq \frac{L}{n^{1/4}}. \tag{1.18}$$

The next step is to estimate the average of $F(\theta)$ on the sphere, and connect it with $\mathbb{E}f(Z)$. To that end, we introduce a random vector Y, independent of X, that is distributed uniformly on the sphere S^{n-1}. The main observation here is that the random variables $X \cdot Y$ and $|X|Y_1$ have exactly the same distribution. Therefore,

$$\int_{S^{n-1}} F(\theta) d\sigma_{n-1}(\theta) = \mathbb{E}F(Y) = \mathbb{E}f(X \cdot Y) = \mathbb{E}f(|X|Y_1). \tag{1.19}$$

Note that according to (1.15), the random variable $|X|$ is typically very close to $\sqrt{n}$. According to the Maxwell principle, the random variable Y_1 is approximately a Gaussian of mean zero and variance $1/n$. Thus, in high dimensions, the random variable $|X|Y_1$ should be approximately a standard normal random variable. Mathematically, we use Exercise **??**(iv) and argue as follows:

$$\begin{aligned} |\mathbb{E}f(|X|Y_1) - \mathbb{E}f(Z)| &\leq \left|\mathbb{E}f(\sqrt{n}Y_1) - \mathbb{E}f(Z)\right| + \left|\mathbb{E}f(\sqrt{n}Y_1) - \mathbb{E}f(|X|Y_1)\right| \\ &\leq \frac{CL}{n} + L\mathbb{E}\left|\left(\sqrt{n} - |X|\right) Y_1\right| \\ &\leq \frac{CL}{n} + L\sqrt{\mathbb{E}nY_1^2} \cdot \sqrt{\mathbb{E}\left(|X|/\sqrt{n} - 1\right)^2} \leq \frac{CL}{n} + L\varepsilon. \end{aligned}$$

By combining the last computation with (1.18) and (1.19) we deduce the conclusion of the lemma. □

Of course, there is nothing holy about the $n^{-1/4}$ factor in (1.17). Actually, this factor may be improved up to $1/\sqrt{n}$, at the expense of deteriorating our lower bound for the measure of the subset Θ.

Proof of Theorem 1.4.2: Set $\delta = \max\{\sqrt{\varepsilon}, n^{-1/8}\}$. For $t \in \mathbb{R}$ consider the function

$$I_t(x) = \begin{cases} 1 & x < t \\ 1 - (x-t)/\delta & t \leq x \leq t+\delta \\ 0 & x > t+\delta \end{cases}$$

Then I_t is a $(1/\delta)$-Lipschitz function, and

$$\mathbb{P}(X \cdot \theta \leq t) \;\leq\; \mathbb{E} I_t(X \cdot \theta) \;\leq\; \mathbb{P}(X \cdot \theta \leq t + \delta). \tag{1.20}$$

Denote $t_i = \Phi^{-1}(j\delta)$ for $j = 1, \ldots, k$ where $k = \lceil 1/\delta - 1 \rceil \leq n^{1/8}$. Then

$$\mathbb{P}(Z \leq t_j) = j\delta \qquad\qquad \text{for } j = 1, \ldots, k.$$

Set also $t_0 = -\infty$ and $t_{k+1} = +\infty$. Since Φ is a 1-Lipschitz function (why?), then $t_{j+1} \geq t_j + \delta$ for any j, and also

$$|\mathbb{P}(Z \leq t) \,-\, j\delta| \leq 2\delta \qquad \text{for } t \in [t_{j-1} - \delta, t_j + \delta]. \tag{1.21}$$

We will apply Lemma 1.4.3 simultaneously for the functions $I_{t_1}, \ldots, I_{t_k}$. We conclude that there exists a subset $\Theta \subseteq S^{n-1}$ with $\sigma_{n-1}(\Theta) \geq 1 - Ck\exp(-c\sqrt{n})$, such that for $\theta \in \Theta$,

$$\left| \mathbb{E} I_{t_j}(X \cdot \theta) \,-\, \mathbb{E}_{t_j}(Z) \right| \leq \frac{C}{\delta}\left(\frac{1}{n^{1/4}} + \varepsilon\right) \qquad \text{for } j = 1, \ldots, k.$$

Since $k \leq n^{1/8}$ then $\sigma_{n-1}(\Theta) \geq 1 - C\exp(-c\sqrt{n})$. It remains to show that (1.16) holds true for any $\theta \in \Theta$. To that end, let us fix $\theta \in \Theta$ and $t \in \mathbb{R}$. Then there exists $j = 1, \ldots, k+1$ such that $t_{j-1} \leq t \leq t_j$. Hence,

$$\begin{aligned} \mathbb{P}(X \cdot \theta \leq t) &\leq \mathbb{P}(X \cdot \theta \leq t_j) \leq \mathbb{E} I_{t_j}(X \cdot \theta) \leq \mathbb{E}_{t_j}(Z) + \frac{C}{\delta}\left(\frac{1}{n^{1/4}} + \varepsilon\right) \\ &\leq \mathbb{P}(Z \leq t_j + \delta) + \frac{C}{\delta}\left(\frac{1}{n^{1/4}} + \varepsilon\right) \leq \mathbb{P}(Z \leq t) + 2\delta + \frac{C}{\delta}\left(\frac{1}{n^{1/4}} + \varepsilon\right). \end{aligned}$$

Similarly, we may argue that

$$\mathbb{P}(X \cdot \theta \leq t) \geq \mathbb{E} I_{t_{j-1}}(X \cdot \theta) \geq \ldots \geq \mathbb{P}(Z \leq t) - 2\delta - \frac{C}{\delta}\left(\frac{1}{n^{1/4}} + \varepsilon\right).$$

Since $2\delta + \frac{C}{\delta}\left(\frac{1}{n^{1/4}} + \varepsilon\right) \leq \tilde{C}\left(\sqrt{\varepsilon} + n^{-1/8}\right)$, the desired bound (1.16) follows. □

The above discussion demonstrates that the Gaussian approximation property of the marginals is not necessarily associated with independent random variables. The geometry of the high-dimensional sphere plays a central rôle in the context of Gaussian approximation principles.

Exercise 1.4.4 *Let $(\Omega, \mathbb{P})$ be a probability space, and let $f_1, \ldots, f_n \in L^2(\Omega)$ be an orthonormal system such that $\sum_{i=1}^n f_i^2 \equiv 1$, Prove that there exist coefficients $(\theta_1, \ldots, \theta_n) \in S^{n-1}$ such that $f = \sum_{i=1}^n \theta_i f_i$ satisfies*

$$\left| \mathbb{P}(f \leq t) \,-\, \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-s^2/2} ds \right| \leq \frac{C}{n^{1/8}} \qquad\qquad (t \in \mathbb{R}).$$

where m is the Lebesgue measure. (There are many non-trivial examples of such orthonormal systems. For instance, any orthonormal basis of the space of spherical harmonics of a certain degree and dimension).

In the antipode, there are random vectors without thin-shell. For instance we take Y a random vector according to the uniform distribution on the sphere and let δ be such that

$$\mathbb{P}(\delta = 1) = \frac{1}{2} \text{ and } \mathbb{P}(\delta = 2) = \frac{1}{2}$$

and think of the marginals of $X = \delta Y$. Without thin-shell, most marginals are approximate $|X|Y_1$, so a mixture of Gaussians. Finally, without thin-shell, multidimensional marginals are typically sphererically symmetric.

Chapter 2

Probability measures with convexity properties

2.1 Brunn-Minkowski and related inequalities

Recall that a subset $K \subset \mathbb{R}^n$ is convex when

$$x, y \in K,\ 0 < \lambda < 1 \qquad \Longrightarrow \qquad \lambda x + (1-\lambda) y \in K.$$

A convex body is a bounded, open convex set. One of the first things that one needs to know about convex sets is the *Brunn-Minkowski inequality*: For any non-empty, Borel-measurable sets $A, B \subset \mathbb{R}^n$,

$$Vol_n(A+B)^{1/n} \geq Vol_n(A)^{1/n} + Vol_n(B)^{1/n}. \tag{2.1}$$

where $A + B = \{x + y\,;\, x \in A, y \in B\}$ is the Minkowski sum of A and B. We said that (2.1) is about convex sets, yet convexity is not mentioned at all. A hint regarding the connection of Brunn-Minkowski to convex sets is provided by the equality case: When A and B are closed sets of positive measure, there is equality in (2.1) if and only if A is convex and B is homothetic to a translate of A.

A quick argument leads from the Brunn-Minkowski inequality to the isoperimetric inequality in $\mathbb{R}^n$: Suppose that $A \subset \mathbb{R}^n$ is a closed set with a smooth boundary such that $Vol_n(A) = Vol_n(B_2^n)$. By Brunn-Minkowski, for any $\varepsilon > 0$,

$$Vol_n\left(A + \varepsilon B_2^n\right) \geq \left(Vol_n(A)^{1/n} + Vol_n(\varepsilon B_2^n)^{1/n}\right)^n = Vol_n\left(B_2^n + \varepsilon B_2^n\right). \tag{2.2}$$

Here, $\lambda A = \{\lambda x\,;\, x \in A\}$ for a subset $A \subset \mathbb{R}^n$ and $\lambda \in \mathbb{R}$. Therefore,

$$\begin{aligned} Vol_{n-1}(\partial A) &= \lim_{\varepsilon \to 0^+} \frac{Vol_n(A + \varepsilon B_2^n) - Vol_n(A)}{\varepsilon} \\ &\geq \lim_{\varepsilon \to 0^+} \frac{Vol_n(B_2^n + \varepsilon B_2^n) - Vol_n(B_2^n)}{\varepsilon} = Vol_{n-1}(\partial B_2^n). \end{aligned}$$

The latter inequality is usually referred to as the isoperimetric inequality in $\mathbb{R}^n$. As for the proof of (2.1), The original approach of Brunn utilizes the *Steiner symmetrization*. Suppose $A \subset \mathbb{R}^n$ is a Borel set. Let $\theta \in \mathbb{R}^n$ be a unit vector, and denote $H = \theta^\perp$. The "Steiner symmetrization of K with respect to the hyperplane H" is the set

$$S_H(A) = \left\{ y + t\theta\,;\, A \cap (y + \mathbb{R}h) \neq \emptyset,\ y \in H,\ |t| \leq \frac{1}{2} \mathrm{Meas}\left\{A \cap (y + \mathbb{R}h)\right\} \right\}$$

where Meas is the one dimensional Lebesgue measure in the line $x + \mathbb{R}h$.

Figure

By Fubini's theorem, $Vol_n(S_H(A)) = Vol_n(A)$. It may be verified directly that for any Borel sets $A, B \subset \mathbb{R}^n$,

$$S_H(A+B) \supseteq S_H(A) + S_H(B). \tag{2.3}$$

Intuitively, by applying consecutive Steiner symmetrizations to A, we may get closer and closer to a Euclidean ball of the same volume as A. The Brunn-Minkowski inequality is in fact an equality when both sets are Euclidean balls (why?). Therefore, the inclusion (2.3) seems to indicate that for any $r, s > 0$, the infimum of

$$\inf \{Vol_n(A+B)\,;\, Vol_n(A) = r, Vol_n(B) = s\}$$

is attained when A and B are Euclidean balls. The actual formal proof involves compactness and continuity with respect to the Hausdorff metric, quite similar to the argument presented in the proof of Theorem 1.3.1 in the previous lecture.

Exercise 2.1.1 Fill in the missing details in the above heuristic argument for the Brunn-Minkowski inequality. (e.g., consider a minimizing pair (A, B) with maximal intersections with the Euclidean balls of the same volumes. Prove that both A and B must be invariant under all Steiner symmetrizations, hence they are Euclidean balls).

The Brunn-Minkowski inequality has quite a few equivalent formulations and consequences. For instance, it follows from (2.1) and the Arithmetic/Geometric means inequality that for any Borel sets $A, B \subset \mathbb{R}^n$ and $0 < \lambda < 1$,

$$Vol_n(\lambda A + (1-\lambda)B) \geq Vol_n(A)^\lambda Vol_n(B)^{1-\lambda}. \tag{2.4}$$

The formulation (2.4) is usually referred to as the multiplicative form of the Brunn-Minkowski inequality. One little advantage of (2.4) is that one does not need to assume anymore that A and B are non-empty.Still, for the applications that we have in mind, it will be much more convenient to have an inequality involving *functions* rather than *sets*. The following theorem presents such a functional variant of the classical Brunn-Minkowski inequality.

Theorem 2.1.2 (the Prékopa-Leindler inequality) *Suppose $f, g, h : \mathbb{R}^n \to [0, \infty)$ are integrable functions and $0 < \lambda < 1$. Assume that for any $x, y \in \mathbb{R}^n$,*

$$h(\lambda x + (1-\lambda)y) \geq f^\lambda(x) g^{1-\lambda}(y). \tag{2.5}$$

Then,

$$\int_{\mathbb{R}^n} h \geq \left(\int_{\mathbb{R}^n} f\right)^\lambda \left(\int_{\mathbb{R}^n} g\right)^{1-\lambda}. \tag{2.6}$$

In some sense, the Prékopa-Leindler inequality is "opposite" to the Hölder inequality, which states that for any integrable, non-negative functions f, g and $0 < \lambda < 1$,

$$\int f^\lambda g^{1-\lambda} \leq \left(\int f\right)^\lambda \left(\int g\right)^{1-\lambda}.$$

Observe that when we plug in $f = 1_A, g = 1_B$ and $h = 1_{\lambda A + (1-\lambda)B}$ in the Prékopa-Leindler inequality, we recover the Brunn-Minkowski inequality in the form (2.4).

2.2 Introduction to log-concave measures

We would like to introduce a convenient family of probability measures in $\mathbb{R}^n$, which are the *logarithmically-concave measures*. A non-negative function $\rho : \mathbb{R}^n \to [0,\infty)$ is *log-concave* if it takes the form

$$\rho = \exp(-H)$$

for a convex function $H : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. Here, of course, $\exp(-\infty) = 0$. Students who still sometimes confuse between convex and concave should recall that a function $H : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is convex when

$$H(\lambda x + (1-\lambda)y) \leq \lambda H(x) + (1-\lambda)H(y) \qquad \text{for } x, y \in \mathbb{R}^n, \lambda \in (0,1). \tag{2.7}$$

Thus, for instance, the Euclidean norm is a convex function in $\mathbb{R}^n$, as well as its square. Note that convex functions here may attain the value $+\infty$. Consequently, whenever $K \subset \mathbb{R}^n$ is a convex set, the function

$$H(x) = \begin{cases} 0 & x \in K \\ +\infty & x \notin K \end{cases}$$

is convex. A convex function H is necessarily continuous and locally-Lipschitz in the interior of $\{x \in \mathbb{R}^n; H(x) < +\infty\}$. The Hessian matrix of a smooth, convex function is positive semi-definite,

$$\nabla^2 H \geq 0$$

in the sense of symmetric matrices.

Here are some examples of log-concave functions. The characteristic function 1_K of a convex set $K \subset \mathbb{R}^n$, which equals one on the convex set and vanishes otherwise, is a log-concave function. The Gaussian function $t \mapsto \exp(-t^2/2)$ is log-concave on the real line. Whenever A is an $n \times n$, positive-definite matrix, the probability density

$$\mathbb{R}^n \ni x \mapsto \left(\frac{\det A}{2\pi}\right)^{n/2} \cdot \exp(-Ax \cdot x/2)$$

is a log-concave function. This is the probability density of a *multi-dimensional Gaussian* random vector, with mean zero and covariance matrix A^{-1}.

Exercise 2.2.1 *Suppose ρ is an integrable, log-concave function in $\mathbb{R}^n$.*

(i) Prove that the set $K = \{\rho > \varepsilon\}$ is open, convex and bounded for any $\varepsilon > 0$.

(ii) In the case where $\rho(0) > 0$, prove that there exists $R > 0$ such that

$$\rho(x) \leq \rho(0)\exp(-|x|/R) \qquad \textit{for all } |x| \geq R.$$

(iii) Conclude that any integrable, log-concave function decays exponentially at infinity. That is, there exists $A, B > 0$ such that

$$\rho(x) \leq A\exp(-B|x|) \qquad \textit{for all } x \in \mathbb{R}^n.$$

We say that a probability measure on $\mathbb{R}^n$ is *log-concave* if it is supported on some affine subspace $E \subset \mathbb{R}^n$ (usually $E = \mathbb{R}^n$), and it has a log-concave density in the affine subspace E. The standard Gaussian measure in $\mathbb{R}^n$, the uniform probability measure on a bounded convex set, and Dirac's delta measure are all examples for log-concave probability measure. A random vector X is said to be log-concave when it is distributed according to a log-concave measure. The pointwise product of log-concave densities is certainly log-concave. The following proposition shows that the class of log-concave random vectors is stable under more operations.

Proposition 2.2.2 *Suppose X and Y are independent, log-concave random vectors. Then,*

(a) For a subspace $E \subset \mathbb{R}^n$, denote by $Proj_E : \mathbb{R}^n \to E$ the orthogonal projection operator onto E in $\mathbb{R}^n$. Then,

$$Proj_E(X)$$

is log-concave.

(b) For any affine map T, the random vector $T(X)$ is log-concave.

(c) The random vector $X + Y$ is log-concave.

Proof: Property (a) is deduced from the Prékopa-Leindler inequality, as follows. We may assume that the support of X is not contained in a hyperplane (why?). Denote by $\rho : \mathbb{R}^n \to [0, \infty)$ the log-concave density of X and set

$$\rho_E(x) = \int_{E^\perp} \rho(x + y)dy \qquad (x \in E),$$

the density of $Proj_E(X)$. We need to show that f_E is log-concave. Pick $x_1, x_2 \in E$ and $0 < \lambda < 1$ and denote $x = \lambda x_1 + (1 - \lambda)x_2$. We need to establish the inequality

$$\int_{E^\perp} \rho(x + y)dy \geq \left(\int_{E^\perp} \rho(x_1 + y)dy \right)^\lambda \left(\int_{E^\perp} \rho(x_2 + y)dy \right)^{1-\lambda}. \tag{2.8}$$

However, the three functions $f(y) = \rho(x_1 + y)$, $g(y) = \rho(x_2 + y)$ and $h(y) = \rho(x + y)$ satisfy the requirements of the Prékopa-Leindler inequality, thanks to the log-concavity of ρ. Hence (2.8) follows, and (a) is proven.

We move on to the proof of (b). Whenever $T : \mathbb{R}^n \to \mathbb{R}^n$ is an *invertible*, affine map, then the density of $T(X)$ is proportional to the the functions $x \mapsto \rho(T(x))$ where ρ is the log-concave density of X. Since $x \mapsto \rho(T(x))$ is also log-concave, then we proved (b) in the case of an *invertible*, affine map T. All that remains is to not that any affine map T is the composition of invertible affine maps and an orthogonal projection operator. Thus (b) holds in view of (a).

Regarding (c), we need to note that the random vector $(X, Y) \in \mathbb{R}^{2n}$ is log-concave, since the product of log-concave functions is log-concave. We use (b) with the linear map $T(u, v) = u + v$ for $u, v \in \mathbb{R}^n$, and deduce (c). □

Now, we will prove a very useful lemma which refers to log-concavity of the moments.

Lemma 2.2.3 *Let $f : \mathbb{R}_+ \mapsto [0, \infty)$ be a log-concave integrable function. We define*

$$M_f(p) = \frac{\int_0^\infty t^p f(t)\, dt}{\Gamma(p + 1)},$$

with $p \geq 0$. Then $M_f(p)$ is log-concave.

Example: If we take $f(t) = Ae^{-Bt}$, then $M_f(p) = B^{-p-1}A$ whence in some sense we have equality. In other words, in our problem the exponential is extremal along log-concave functions.
Proof of the Lemma: From the exercise 2.2.1 we deduce that f has all moments, so $M_f(p)$ is finite and continuous. Let $p < q$. We need to show that

$$M_f\left(\frac{p+q}{2}\right) \geq \sqrt{M_f(p)M_f(q)}.$$

Now we find $A, B > 0$ such that

- $M_f(p) = AB^{-p-1} = M_g(p)$,
- $M_f(q) = AB^{-q-1} = M_g(q)$,

where $g(t) = Ae^{-Bt}$. Then the following claim holds.

Claim 2.2.4 *Functions f and g intersect at least twice.*

Explanation: First from our choice of A, B we have that

1. $\int_0^\infty t^p g(t)\,dt = \int_0^\infty t^p f(t)\,dt$,
2. $\int_0^\infty t^{q-p}\left(t^p g(t)\right)\,dt = \int_0^\infty t^{q-p}\left(t^p f(t)\right)$.

From the above it is clear that it is impossible to have always $f \geq g$ or $g \geq f$. So there is an intersection of f and g and suppose that is the only one. Then from the first relation we have that either

$$\int_x^\infty t^p g(t)\,dt > \int_x^\infty t^p f(t)\,dt,$$

or

$$\int_x^\infty t^p g(t)\,dt < \int_x^\infty t^p f(t)\,dt,$$

for $x \in \text{supp}(f)^\circ$. However, this contradicts the second relation since we write

$$\int_0^\infty t^{q-p}\left(t^p f(t)\right)\,dt = \int_0^\infty (q-p)x^{q-p-1}\left(\int_x^\infty t^p f(t)\,dt\right)\,dx.$$

$\diamondsuit$

Claim 2.2.5 *We can find $a, b > 0$ such that*

- *$f(t) \geq g(t)$ for all $t \in [a, b]$ and*
- *$f(t) \leq g(t)$ for all $t \notin [a, b]$.*

Explanation: This holds due to the log-concavity of the difference $f - g$.

$\diamondsuit$

Now we are able to finish the proof. Set $r = \frac{q-p}{2}$ and its easy to see that

$$\int_0^\infty (t^r - a^r)(t^r - b^r)t^p(g(t) - f(t))\,dt \geq 0$$

and after the manipulations we get

$$\int_0^\infty t^{2r+p}(g(t) - f(t))\,dt + (ab)^r \int_0^\infty t^p(g(t) - f(t))\,dt - 2(ab)^r \int_0^\infty t^{r+p}(g(t) - f(t))\,dt \geq 0$$

Note that the first two integrals are equal to zero due to (1) and (2). So finally we get

$$\int_0^\infty t^{\frac{p+q}{2}} f(t)\,dt \geq \int_0^\infty t^{\frac{p+q}{2}} g(t)\,dt$$

which gives

$$M_f\left(\frac{p+q}{2}\right) \geq M_g\left(\frac{p+q}{2}\right) = \sqrt{M_g(p)M_g(q)} = \sqrt{M_f(p)M_f(q)},$$

which is the desired inequality.

■

The following exercise shows that in fact the above result can easily be extended.

Exercise 2.2.6 *We set $M_{-1}(f) = f(0)$. Prove that we can extend the log-concavity of moments to the range $p \in [-1, \infty)$.*

To see an immediate example of using the lemma above, take $X \geq 0$ to be a random variable with log-concave density f. Then

$$M_f(2) \geq \sqrt{M_f(0)M_f(4)},$$

which, after some easy algebra gives

$$\mathbb{E}X^4 \leq 6\left(\mathbb{E}X^2\right)^2,$$

which is a reverse Hölder type inequality. Using such type inequalities we can find a first connection between log-concavity and thin-shell. This is described by the following corollary.

Corollary 2.2.7 *Let X be a random vector in $\mathbb{R}^n$ with density f,such that f is log-concave and f is radial, which means $f(x) = f(|x|)$. If we denote by $\sigma^2 = \mathbb{E}|X|^2$, then*

$$\mathbb{E}\left(\frac{|X|}{\sigma} - 1\right)^2 \leq \frac{c}{n}.$$

Proof. For any $\theta \in S^{n-1}$ we set $g(r) = \mathrm{Vol}_{n-1}(S^{n-1})f(r\theta)$. Integrating in polar coordinates we get:

$$\mathbb{P}(|X| \leq t) = \int_{tB^n} f = \int_{S^{n-1}} \int_0^t f(r\theta)r^{n-1}\,dr d\theta = \int_0^t r^{n-1}g(r)\,dr.$$

Hence the density of $|X|$ is the function $r^{n-1}g(r)$. So we write:

$$\mathbb{E}|X|^4 = \int_0^\infty r^4 \cdot r^{n-1}g(r)\,dr = (n+3)! \cdot M(n+3)$$

and
$$\mathbb{E}|X|^2 = \int_0^\infty r^2 \cdot r^{n-1} g(r)\, dr = (n+1)! \cdot M(n+1).$$
Finally note that
$$1 = \mathbb{E}|X|^0 = (n-1)! \cdot M(n-1).$$
The lemma of log-concavity of moments gives
$$M(n+1) \geq \sqrt{M(n+3)M(n-1)}.$$
Inserting the above results:
$$\mathbb{E}|X|^4 \leq \sigma^2 \frac{(n+3)(n+2)}{n(n+1)} \leq \sigma^2 \left(1 + \frac{c}{n}\right).$$
This gives the desired result because
$$\mathbb{E}\left(\frac{|X|}{\sigma} - 1\right)^2 \leq \mathbb{E}\left(\frac{|X|^2}{\sigma^2} - 1\right)^2 = \frac{\mathbb{E}|X|^4 - \sigma^4}{\sigma^4} \leq \frac{c}{n}$$

Next we present a theorem which is due to Brascamp and Lieb.

Theorem 2.2.8 *Let Ψ be a convex function on $\mathbb{R}^n$ which is C^2 smooth and $\nabla^2\Psi(x) > 0$ for all $x \in \mathbb{R}^n$. Assume that μ is a finite measure such that $\frac{d\mu}{dx} = e^{-\Psi}$. Then for each C^1 smooth $f \in L^2(\mu)$ we have:*
$$\mathrm{Var}_\mu(f) \leq \int_{\mathbb{R}^n} \left(\nabla^2\Psi\right)^{-1} \nabla f \cdot \nabla f \, d\mu(x),$$
where $\mathrm{Var}_\mu(f) = \int_{\mathbb{R}^n} (f-E)^2\, d\mu(x)$, with $E = \frac{\int_{\mathbb{R}^n} f\, d\mu}{\mu(\mathbb{R}^n)}$.

Proof. We will use ideas related to Ricci curvature, Bochner formula and computation of derivatives. Denote by $\mathcal{S}$ the space of C^∞ smooth compactly supported functions. We will find first the adjoint representation of ∂_i in $L^2(\mu)$. For $u \in \mathcal{S}$ and v smooth we have:
$$\int_{\mathbb{R}^n} (\partial_i u)\, v e^{-\Psi(x)}\, dx = -\int_{\mathbb{R}^n} u \left(\partial_i v - v \partial_i \Psi\right) e^{-\Psi(x)}\, dx.$$
For $u \in \mathcal{S}$ we define:
$$\partial^* u = \partial_i u - (\partial_i \Psi) u.$$
We observe the identity:
$$\partial_j \left(\partial_i^* u\right) = \partial_i^* \left(\partial_j u\right) - \left(\partial_{ij}\Psi\right) u.$$
For every $u \in \mathcal{S}$ we define
$$Lu = \sum_{i=1}^n \partial_i^* \partial_i u = \Delta u - \nabla\Psi \cdot \nabla u.$$
Then for every smooth function f we have that
$$\int (Lu) f\, d\mu = \sum_{i=1}^n \int \partial_i^* (\partial_i u) f\, d\mu = -\sum_{i=1}^n \int \partial_i u\, \partial_i f\, d\mu = -\int \nabla u \cdot \nabla f\, d\mu$$

and for each $u \in \mathcal{S}$ we get $\int Lu\, d\mu = 0$. We also observe the identity

$$\partial_k(Lu) = \sum_{i=1}^n \partial_k \partial_i^*(\partial_i u) = \sum_{i=1}^n \partial_i^* \partial_{ki} u - \sum_{i=1}^n \partial_{ik}\Psi\, \partial_i u$$
$$= L(\partial_k u) - \sum_{i=1}^n \partial_{ik}\Psi\, \partial_i u.$$

Now are now ready to prove Bochner's formula.

Proposition 2.2.9 *For each $u \in \mathcal{S}$ we have that*

$$\int_{\mathbb{R}^n} (Lu)^2\, d\mu = \int_{\mathbb{R}^n} \left(\nabla^2\Psi\right) \nabla u \cdot \nabla u\, d\mu + \sum_{i=1}^n \int_{\mathbb{R}^n} |\nabla \partial_i u|^2\, d\mu.$$

Proof. Integration by parts gives

$$\int_{\mathbb{R}^n} (Lu)^2\, d\mu = -\int_{\mathbb{R}^n} \nabla(Lu) \cdot \nabla u\, d\mu = -\sum_{i=1}^n \int_{\mathbb{R}^n} \partial_i(Lu)\partial_i u\, d\mu$$
$$= -\sum_{i=1}^n \int_{\mathbb{R}^n} L(\partial_i u)\partial_i u\, d\mu + \sum_{i,j=1}^n \int_{\mathbb{R}^n} \partial_{ij}\Psi \partial_j u \partial_i u\, d\mu$$
$$= \sum_{i=1}^n \int_{\mathbb{R}^n} |\nabla \partial_i u|^2\, d\mu + \int_{\mathbb{R}^n} \left(\nabla^2\Psi\right) \nabla u \cdot \nabla u\, d\mu.$$

■

We will prove now a technical lemma.

Lemma 2.2.10 *The image of $\mathcal{S}$ under the operator L is dense in*

$$H = \left\{ f \in L^2(\mu) : \int f\, d\mu = 0 \right\} \subset L^2(\mu).$$

Proof. For each $u \in \mathcal{S}$ we set

$$Lu = \Delta u - \nabla\Psi \cdot \nabla u.$$

Then suppose that $f \in L^2(\mu)$ is such that for each $u \in \mathcal{S}$ we have that $Lu \perp f$. We will show that f is constant.

Observe first that L is elliptic and symmetric of second order. From our assumption, we classically get that f is a $C^\infty-$smooth function and $Lf = 0$. This gives that for each $\theta \in \mathcal{S}$ we have that

$$L(f^2) = 2fLf + 2\nabla f \cdot \nabla f = 2|\nabla f|^2.$$

Using this we make the following computation:

$$\int |\nabla(\theta f)|^2\, d\mu = \int \left(|\nabla\theta|^2 f^2 + 2\theta f \nabla\theta \cdot \nabla f + |\nabla f|^2 \theta^2 \right)\, d\mu$$
$$= \int |\nabla\theta|^2 f^2\, d\mu + \frac{1}{2}\int \nabla(\theta^2) \cdot \nabla(f^2)\, d\mu + \int |\nabla f|^2 \theta^2\, d\mu$$
$$= \int |\nabla\theta|^2 f^2\, d\mu - \frac{1}{2}\int \theta^2 L(f^2)\, d\mu + \int |\nabla f|^2 \theta^2\, d\mu.$$

Now take

$$\theta(x) = \begin{cases} 1 & \text{if } |x| \leq 0 \\ 0 & \text{if } |x| \geq 2. \end{cases}$$

This is a C^∞ function with $|\nabla\theta(x)| \leq M$ for each x. Setting $\partial_k(x) = \partial(x/k)$ we get that $|\nabla\theta_k(x)| \leq \frac{M}{k}$. In addition $\partial_k f \longrightarrow f$ pointwise and $|\nabla(\theta_k f)|^2 \longrightarrow f$ pointwise. So by Fatou's lemma

$$\int |\nabla f|^2\, d\mu \leq \liminf_{k\to\infty} \int |\nabla(\theta_k f)|^2\, d\mu = \liminf_{k\to\infty} \int |\nabla\theta_k f|^2 f^2\, d\mu$$
$$\leq \liminf_{k\to\infty} \int f^2 \frac{M}{k} = 0,$$

which means that $\int |\nabla f|^2\, d\mu = 0$ and so $\nabla f = 0$ and f is constant.

■

We are now ready to give a proof of Theorem 2.2.8.

Proof. We take $f \in L^2(\mu)$ which is C^1-smooth and we assume that $\int f\, d\mu = 0$.

Fix $\varepsilon > 0$. From lemma 2.2.10 we can find $u \in \mathcal{S}$ such that

$$\|Lu - f\|_{L^2(\mu)} < \varepsilon.$$

We write

$$\begin{aligned}\mathrm{Var}_\mu(f) = \|f\|^2_{L^2(\mu)} &= \|Lu - f\|^2_{L^2(\mu)} + 2\int fLu\, d\mu - \int (Lu)^2\, d\mu \\ &\leq \varepsilon^2 - 2\int \nabla f \cdot \nabla u\, d\mu - \int (\nabla^2\Psi)\nabla u \cdot \nabla u\, d\mu \\ &\leq \varepsilon^2 + \int (\nabla^2\Psi)^{-1}\nabla f \cdot \nabla f\, d\mu,\end{aligned}$$

where we used that

$$\int (Lu)^2\, d\mu \geq \int (\nabla^2\Psi)\nabla u \cdot \nabla u\, d\mu,$$

which follows from Bochner's formula and

$$-2x \cdot y - Ax \cdot x \leq A^{-1}y \cdot y \Longleftrightarrow |\sqrt{A}x + \sqrt{A^{-1}}y|^2 \geq 0.$$

If we take $\varepsilon \to 0$ we are done.

■

Now let $u : \mathbb{R}^n \mapsto \mathbb{R} \cup \{+\infty\}$ convex or not and define

$$u^*(x) = \sup_{y\in\mathbb{R}^n} [x \cdot y - u(y)].$$

Then u^* is convex. Define

$$I(u) = \log \int e^{-u^*}.$$

Exercise 2.2.11 *Let u and u^* be as above. Then we have that:*

(i) If u is smooth

$$u^*\left(\nabla u(x)\right) = x \cdot \nabla u(x) - u(x).$$

(ii) If $\Psi = u^*$ is smooth, $\nabla^2\Psi(x) > 0$ and f is a C^1-smooth function then

$$\frac{\partial^2}{\partial t^2} I(u+tf)\mid_{t=0} = \mathrm{Var}_{e^{-\Psi}}(g) - \int_{\mathbb{R}^n} \left(\nabla^2\Psi\right)^{-1} \nabla g \cdot \nabla g \, e^{-\Psi},$$

(iii) Applying Précopa's inequality, prove directly that I is concave without smoothness and strict convexity assumptions.

2.3 First application of Brascamp-Lieb to thin shell bounds

We begin with a proposition.

Proposition 2.3.1 *Let μ be a probability measure on $\mathbb{R}^n$ with density $e^{-\Psi}$. Suppose that for each $x \in \mathbb{R}^n$ we have that*

$$\nabla^2\Psi(x) \geq \delta > 0.$$

Then for each $f \in L^2(\mu)$ which is also C^1-smooth we have that

$$\mathrm{Var}_\mu(f) \leq \frac{1}{\delta}\int |\nabla f|^2 \, d\mu.$$

Proof. From the assumption we easily get that $\left(\nabla^2\Psi\right)^{-1} \leq \frac{1}{\delta}$. Using this, it is immediate that

$$\left(\nabla^2\Psi\right)^{-1} \nabla f \cdot \nabla f \leq \frac{1}{\delta}|\nabla f|^2.$$

Now the result comes from an application of Brascamp-Lieb inequality.

■

In order to use this for thin-shell bounds we take $f(x) = \frac{|x|^2}{n}$ and μ isotropic. Then,

$$|\nabla f(x)|^2 = \frac{4x_1^2 + \cdots + 4x_n^2}{n^2} = \frac{4|x|^2}{n^2}.$$

If $\nabla^2\Psi(x) > \delta >> \frac{1}{n}$ then

$$\mathbb{E}\left(\frac{|X|}{\sqrt{n}} - 1\right)^2 \leq \mathrm{Var}\left(\frac{|X|^2}{n}\right) \leq \frac{1}{\delta}\int \frac{4|X|^2}{n^2} = \frac{4}{\delta n} << 1.$$

2.4 Second application of Brascamp-Lieb to $\frac{1}{2}-$convexity

Definition 2.4.1 *Let $0 < p \leq 1$. A set $K \subset \mathbb{R}^n_+$ is called $p-$convex if the set*

$$\{(x_1^p, \cdots, x_n^p) : \; x \in K\}$$

is convex.

Note first that $1-$convexity is the known convexity.
For simplicity we deal with the case $p = \frac{1}{2}$. As a first example in that case we observe that the set
$$B_p^n = \left\{ x \in \mathbb{R}_+^n : \sum_{i=1}^n x_i^p \le 1 \right\}$$
is $\frac{1}{2}-$convex if $p \ge \frac{1}{2}$.

Definition 2.4.2 *A set $K \subset \mathbb{R}_+^n$ is monotone or oriented downwards, if*
$$(x_1, \cdots, x_n) \in K,\ g_i \le x_i \Longrightarrow g \in K.$$

The following exercise shows that the definition of $p-$convexity is monotone in p only for monotone sets.

Exercise 2.4.3 *If $K \subset \mathbb{R}_+^n$ is a $p-$convex and monotone set, then is $q-$convex for $q < p$.*

(i) In particular if $K \subset \mathbb{R}_+^n$ is a convex and monotone set, then is $\frac{1}{2}-$convex.

(ii) If $K \subset \mathbb{R}_+^n$ is unconditional and convex then $K \cap \mathbb{R}_+^n$ is monotone and hence $\frac{1}{2}-$convex.

Definition 2.4.4 *We say that a function $\Psi : \mathbb{R}_+^n \mapsto \mathbb{R} \cup \{+\infty\}$ is $\frac{1}{2}-$convex when the function*
$$x = (x_1, \cdots, x_n) \mapsto \Psi(x_1^2, \cdots, x_n^2)$$
is convex.
Also for a $\frac{1}{2}-$convex function H, the function e^{-H} is $\frac{1}{2}-$log concave.
The connection with the definition of $\frac{1}{2}-$convex sets is obvious now. For, if $K \subset \mathbb{R}_+^n$ is $\frac{1}{2}-$convex if and only if $\mathbf{1}_K$ is $\frac{1}{2}-$log concave.

We can state now the following proposition.

Proposition 2.4.5 *Let μ be a $\frac{1}{2}-$log concave finite measure in $\mathbb{R}_+^n$. Then for each function f which is C^1-smooth*
$$\mathrm{Var}_\mu(f) \le 4 \int_{\mathbb{R}^n} \sum_{i=1}^n x_i^2 |\partial_i f|^2 \, d\mu(x).$$

Proof. The main idea for the proof is to change variables first and then to use Brascamp-Lieb.
Denote $\frac{d\mu}{dx} = e^{-\Psi}$. Then if
$$\pi(x_1, \cdots, x_n) = (x_1^2, \cdots, x_n^2),$$
the function $\Psi(\pi(x))$ is convex. Set
$$\varphi(x) = \Psi(\pi(x)) - \sum_{i=1}^n \log(2x_i).$$

Then,
$$\nabla^2 \varphi(x) \ge \nabla^2 \left(- \sum_{i=1}^n \log(2x_i) \right) = \begin{pmatrix} \frac{1}{x_1^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{x_2^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{x_n^2} \end{pmatrix} > 0,$$

and so we get that

$$\left(\nabla^2\varphi(x)\right)^{-1} \leq \begin{pmatrix} x_1^2 & 0 & \cdots & 0 \\ 0 & x_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_n^2 \end{pmatrix}.$$

Set $g(x) = f(\pi(x))$. By the Brascamp-Lieb inequality we have that

$$\mathrm{Var}_{e^{-\varphi}}(g) \leq \int_{\mathbb{R}^n_+} \left(\nabla^2 f\right)^{-1} \nabla g \cdot \nabla g e^{-\varphi(x)}\, dx \leq \int_{\mathbb{R}^n_+} \sum_{i=1}^n x_i^2 |\partial_i g(x)| 2 e^{-\varphi(x)}\, dx.$$

When $y = \pi(x) = (x_1^2, \cdots, x_n^2)$ then the Jacobian of π equals to

$$\det \frac{dy}{dx} = \prod_{i=1}^n (2x_i).$$

Also note that

$$e^{-\varphi(x)} = e^{-\Psi(\pi(x))} \prod_{i=1}^n (2x_i).$$

From the last two relations we have that π pushes forward the one measure to another

$$e^{-\varphi(x)} dx \overset{\pi_*}{\curvearrowright} e^{-\Psi(y)} dy.$$

So

$$\mathrm{Var}_{e^{-\varphi}}(g) = \mathrm{Var}_{e^{-\Psi}}(f).$$

When $y = \pi(x) = (x_1^2, \cdots, x_n^2)$ we have

$$x_i^2 |\partial_i g(x)|^2 = 4 y_i^2 |\partial_i f(y)|^2,$$

so

$$\mathrm{Var}_{e^{-\Psi}}(f) \leq \int_{\mathbb{R}^n} \sum_{i=1}^n x_i^2 |\partial_i g(x)|^2 e^{-\varphi(x)} dx = 4 \int_{\mathbb{R}^n} \sum_{i=1}^n y_i^2 |\partial_i f(x)|^2 e^{-\Psi(x)} dy.$$

■

Comment. When $Y \in \mathbb{R}^n_+$ has a $\frac{1}{2}-$log concave density then $(\sqrt{Y_1}, \cdots, \sqrt{Y_n})$ is log-concave but a bit more.

Now we ready to prove the main theorem.

Theorem 2.4.6 *Let X be an isotropic random variable in $\mathbb{R}^n$ and assume that its density is unconditional. Denote by Y the conditioning of X to $\mathbb{R}^n_+$ and assume that Y has $\frac{1}{2}-$log concave density. Then*

$$\mathbb{E}\left(\frac{|X|}{\sqrt{n}} - 1\right)^2 \leq \frac{c}{n}.$$

Proof. We apply Proposition 2.4.5 with $f(x) = \frac{|x|^2}{n}$ and using unconditionality we have

$$\mathbb{E}\left(\frac{|X|^2}{n} - 1\right)^2 = \mathbb{E}\left(\frac{|Y|^2}{n} - 1\right)^2 = \mathrm{Var} f(Y) \le 4\mathbb{E}\sum_{i=1}^n Y_i^2 |\partial_i f(Y)|^2 = \frac{16}{n^2}\sum_{i=1}^n \mathbb{E}Y_i^4.$$

Now by Lemma 2.2.3 for the positive and log concave $\sqrt{Y_1}, \cdots, \sqrt{Y_n}$ we have that

$$\left(\frac{\mathbb{E}(\sqrt{Y_i})^8}{8!}\cdot\frac{1}{0!}\right)^{\frac{1}{2}} \le \frac{\mathbb{E}(\sqrt{Y_i})^4}{4!}.$$

Hence,

$$\mathbb{E}Y_i^4 \le \binom{8}{4}\left(\mathbb{E}Y_i^2\right)^2 = \binom{8}{4}\left(\mathbb{E}X_i^2\right)^2 = \binom{8}{4}.$$

Therefore,

$$\mathrm{Var}\left(\frac{|X|^2}{n}\right) \le \frac{16}{n^2}\sum_{i=1}^n \binom{8}{4} = \frac{16\cdot\binom{8}{4}}{n}.$$

■

Using the work of Sudakov, Diaconis, Freedman we get then following Corollary:

Corollary 2.4.7 *Let X be a log-concave, isotropic and unconditional random variable in $\mathbb{R}^n$. Then for each $\theta = (\theta_1, \cdots, \theta_n) \in S^{n-1}$ and for each $t \in \mathbb{R}$,*

$$|\mathbb{P}(X\cdot\theta \le t) - \mathbb{P}(Z \le t)| \le C\sum_{i=1}^n \theta_i^4,$$

where $Z \backsim N(0,1)$. In our case we can compute that with probability greater than $1 - e^{-\sqrt{n}}$ on S^{n-1},

$$\sum_{i=1}^n \theta_i^4 \le \frac{c}{n}.$$

In the general case, without unconditionality, we have the following theorem:

Theorem 2.4.8 *Let X be an isotropic log-concave random vector in $\mathbb{R}^n$. Then*

$$\mathbb{E}\left(\frac{|X|}{\sqrt{n}} - 1\right)^2 \le \varepsilon_n,$$

where $\varepsilon_n \to 0$ as $n \to \infty$.

The best known result is due to Guédon and E.Milman. They showed that

$$\varepsilon_n \le \frac{c}{n^{\frac{1}{3}}}$$

and the famous conjecture states that

$$\varepsilon_n \le \frac{c}{n}.$$

In a similar direction, there is a theorem due to Klartag and Eldan which compares the density of a random projection with the density of the uniform distribution. More precisely we have the following:

Theorem 2.4.9 *Let be a log-concave and isotropic random variable in $\mathbb{R}^n$ and $l \leq n^\alpha$. Then, there exists $\mathcal{A} \subset G_{n,l}$ with $\sigma_{n,l}(\mathcal{A}) \geq 1 - e^{-\sqrt{n}}$, such that for each $E \in \mathcal{A}$*

$$\int_{\mathbb{R}^n} |f_E(x) - g_E(x)| dx \leq \frac{c}{n^\alpha}$$

and

$$\left| \frac{f_E(x)}{g_E(x)} - 1 \right| \leq \frac{c}{n^\alpha}$$

for each $x \in E$ and $|x| \leq cn^\alpha$. Here f_E denotes the log-concave density of $\mathrm{Proj}_E(x)$ and $g_E(x) = (2\pi)^{-l/2} \exp\left(-|x|^2/2\right)$ and c, α are universal constants.

The common idea for all bounds for the thin-shell is to project X to a random subspace and obtain a log-concave and approximately radial distribution. To see why it is enough to look at the projections is more clear by the following lemma.

Lemma 2.4.10 *Let X be a log-concave, isotropic random vector in $\mathbb{R}^n$ and $E \in G_{n,l}$ be a random subspace of X. Then*

$$\mathbb{E}\left(\frac{|X|}{\sqrt{n}} - 1\right)^2 \leq \mathbb{E}_E \mathbb{E}_X \left(\frac{|\mathrm{Proj}_E(X)|}{\sqrt{l}} - 1\right)^2.$$

Proof. For each $z \in \mathbb{R}^n$ we have that

$$\mathbb{E}|\mathrm{Proj}_E z|^2 = \frac{l}{n}|z|^2. \tag{2.9}$$

We need to show that

$$\mathbb{E}\frac{|X|^2}{n} - 2\frac{\mathbb{E}|X|}{\sqrt{n}} + 1 \leq \mathbb{E}_E \mathbb{E}_X \frac{|\mathrm{Proj}_E(X)|^2}{l} - 2\mathbb{E}_E \mathbb{E}_X \frac{|\mathrm{Proj}_E(X)|}{\sqrt{l}} + 1.$$

Using (2.9), it is enough to show that

$$\mathbb{E}_E \mathbb{E}_X |\mathrm{Proj}_E(X)| \leq \sqrt{\frac{l}{n}} \mathbb{E}_X |X|.$$

The last one is true by Jensen's inequality, since

$$\mathbb{E}_E |\mathrm{Proj}_E z| \leq \sqrt{\mathbb{E}_E |\mathrm{Proj}_E z|^2} = \sqrt{\frac{l}{n}} |z|.$$

■

The above shows that it is enough to project X to a random subspace E and prove thin-shell bound for $\mathrm{Proj}_E X$ which is also log-concave and isotropic.
In order to do this, first use Sudakov type phenomenon. By concentration, for a typical $E \in G_{n,l}$, we have that $\mathrm{Proj}_E X$ is approximately spherically symmetric. Also as we proved in Theorem 2.4.6, if X is log-concave, isotropic and radial random vector in $\mathbb{R}^l$, then

$$\mathbb{E}\left(\frac{|X|}{l} - 1\right)^2 \leq \frac{C}{l}.$$

We will see why we care about the best exponent in the thin-shell bound.

Definition 2.4.11 *For a random vector X in $\mathbb{R}^n$, define G_X as the minimal number for which for each function f*

$$\mathrm{Var} f(X) \leq G_X^2 \mathbb{E}|\nabla f(X)|^2.$$

Also denote

$$G_n = \sup_{X \in \mathbb{R}^n} G_X$$

An equivalent description of G_X is given by the next theorem which is due to Buser, Ledoux and E.Milman.

Theorem 2.4.12 *G_X is equivalent, up to a universal constant, to the minimal $R > 0$ such that,*

$$\mathbb{P}(X \in A) = \frac{1}{2} \Longrightarrow \mathbb{P}(X \in A_1) = \frac{1}{2} + \frac{1}{R},$$

for each $A \subset \mathbb{R}^n$, where

$$A_1 = \{x \in \mathbb{R}^n : \ \exists y \in A, \ |x - y| \leq 1\}.$$

Finally, this parameter G_n is connected, by a theorem of Eldan, to the thin-shell bound.

Theorem 2.4.13 *Denote*

$$\sigma_n^2 = \sup_{X \in \mathbb{R}^n} \mathrm{Var}(|X|),$$

where the supremum is taken over all isotropic and log-concave random vectors X in $\mathbb{R}^n$. Then,

$$G_n \leq C\sqrt{\log n}\sqrt{\sum_{k=1}^{n} \frac{G_k^2}{k}}.$$

Since $\sigma_n \leq cn^{1/3}$ it follows that $G_n \leq cn^{1/3}\sqrt{\log n}$.

Chapter 3

The isotropic constant and the Bourgain-Milman inequality

Definition 3.0.14 *Suppose that X is a log-concave vector in $\mathbb{R}^n$ with density f. Set*

$$L_X = L_f = (\det \mathrm{Cov}(X))^{\frac{1}{2n}} \cdot f(\mathbb{E}X)^{\frac{1}{n}},$$

the isotropic constant of X.

It's easy to see that if $T : \mathbb{R}^n \mapsto \mathbb{R}^n$ is affine invertible then $L_X = L_{T(X)}$. In order to see some equivalent definitions of the isotropic constant, we need first a lemma.

Lemma 3.0.15 *Let X be a a log-concave vector in $\mathbb{R}^n$ with density f and set $x_0 = \mathbb{E}X$. Then*

$$e^{-n} \sup f \leq f(x_0) \leq \sup f$$

and

$$\log \frac{1}{f(x_0)} \leq -\int_{\mathbb{R}^n} f(x) \log f(x)\, dx \leq \log \frac{1}{f(x_0)} + n$$

Proof. Set $f(x) = e^{-\Psi(x)}$ for all x, where Ψ is a convex function. Note also that

$$-\int_{\mathbb{R}^n} f(x) \log f(x)\, dx = \mathrm{Ent}(X).$$

We will prove that

$$\Psi(x_0) \leq -\int_{\mathbb{R}^n} f(x) \log f(x)\, dx \leq \Psi(0) + n.$$

For the left hand side inequality we use Jensen's inequality:

$$\Psi(\mathbb{E}X) \leq \mathbb{E}\Psi(X) = -\int_{\mathbb{R}^n} f(x) \log f(x)\, dx.$$

For the right hand side inequality note first that for all $x \in \{\Psi < +\infty\}$ in which Ψ is differentiable

$$\Psi(0) \geq \Psi(X) + \nabla\Psi(X) \cdot (0 - x).$$

Since Ψ is differentiable almost everywhere we have

$$\begin{aligned} \mathrm{Ent}(X) = \int_{\mathbb{R}^n} \Psi(x) e^{-\Psi(x)}\, dx &\leq \int_{\mathbb{R}^n} (\Psi(0) + \nabla\Psi(x) \cdot x)\, e^{-\Psi(x)}\, dx \\ &= \Psi(0) - \sum_{i=1}^{n} \int_{\mathbb{R}^n} x_i \partial_i \left(e^{-\Psi(x)}\right)\, dx \\ &= \Psi(0) + n. \end{aligned}$$

■

From the above lemma we can give the following equivalent versions of the isotropic constant. Up to a universal constant the following is true:

$$L_X \sim [\det \mathrm{Cov}(X)]^{\frac{1}{2n}} (\sup f)^{\frac{1}{n}} \sim [\det \mathrm{Cov}(X)]^{\frac{1}{2n}} \exp\left(-\frac{\mathrm{Ent}(X)}{n}\right).$$

Also, it could be interesting to define

$$n \log L_X = \frac{1}{2} \log \det \mathrm{Cov}(X) - \mathrm{Ent}(X).$$

Notice that $\mathrm{Ent}(X)$ is related to the volume in $\mathbb{R}^n$ that X occupies and $\det \mathrm{Cov}(X)$ is related to the volume of the inertia ellipsoid of X.
To see one more equivalent definition we need the second lemma in the proof of Paouris theorem. Namely, for given $f : \mathbb{R}^n \mapsto [0, \infty)$ which is log concave, set

$$K_f = \left\{ x \in \mathbb{R}^n : \ f(x) \geq (\sup f) \frac{1}{25^n} \right\}.$$

Then we have that

$$\mathbb{P}\left(X \in K_f\right) \geq 1 - e^{-cn}$$

and the following lemma:

Lemma 3.0.16 *Let X be a log-concave vector in $\mathbb{R}^n$ with density $e^{-\Psi}$. Then for each $a \geq 2$,*

$$\mathbb{P}\left(\Psi(X) \geq \inf \Psi + an\right) \leq e^{-can}.$$

From the above lemma, we have the following equivalent version of the isotropic constant. Up to a universal constant the following is true:

$$L_X \sim [\det \mathrm{Cov}(X)]^{\frac{1}{2n}} \int_{\mathbb{R}^n} f^{1+\frac{1}{n}} \sim [\det \mathrm{Cov}(X)]^{\frac{1}{n}} \left(\int_{\mathbb{R}^n} f^2 \right)^{\frac{1}{n}}.$$

The next lemma shows that the isotropic constant is greater than a universal constant.

Lemma 3.0.17 *There exists a universal constant c, such that $L_X > c$ for all log-concave $X \in \mathbb{R}^n$.*

Proof. Applying an affine transformation, assume that $\mathbb{E}X = 0$, $\mathrm{Cov}(X)$ is scalar and $f(0) = 1$. Then,

$$L_X^2 = [\det \mathrm{Cov}(X)]^{\frac{1}{n}} = \frac{\mathrm{Tr}(\mathrm{Cov}(X))}{n} = \frac{\mathbb{E}|X|^2}{n}.$$

Also,

$$\mathbb{P}\left(X \in \frac{\sqrt{n}}{100} B^n \right) = \int_{\frac{\sqrt{n}}{100} B^n} f \leq (\sup f) \mathrm{Vol}\left(\frac{\sqrt{n}}{100} B^n \right) \leq e^n \cdot \frac{e^{-n}}{2} = \frac{1}{2}.$$

Using this we obtain that:

$$\mathbb{E}|X|^2 \geq \mathbb{E}|X|^2 \cdot \mathbf{1}_{|X| \geq \frac{\sqrt{n}}{100}} \geq \frac{cn}{2}.$$

■

The hyperplane conjecture or the slicing problem asks whether L_X is bounded also above by a universal constant.
If we define

$$L_n = \sup_{X \in \mathbb{R}^n} L_X$$

The currently best upper bound is due to Klartag and is

$$L_n \leq c n^{\frac{1}{4}}.$$

Also Klartag and Eldan found a relation to the thin shell, which is

$$L_n \leq c\sigma_n.$$

To give an explanation of the name of the slicing problem, we have the following proposition:

Proposition 3.0.18 *Let K be a convex body on $\mathbb{R}^n$ and let $X \sim \mathrm{Unif}(K)$ with $\mathbb{E}X = 0$ and $\mathrm{Cov}(X)$ to be scalar. Then for each two hyperplanes $H_1, H_2 \subset \mathbb{R}^n$ through zero,*

$$\frac{\mathrm{Vol}_{n-1}(K \cap H_1)}{\mathrm{Vol}_{n-1}(K \cap H_2)} \leq C \leq \sqrt{6}.$$

This one follows from the following lemma:

Lemma 3.0.19 *Let K be a convex body on $\mathbb{R}^n$ and let $X \sim \mathrm{Unif}(K)$ with $\mathbb{E}X = 0$. Then for each $\theta \in S^{n-1}$*

$$c_1 \mathrm{Vol}_n(K) \leq \mathrm{Vol}_{n-1}(K \cap \theta^{\perp})\sqrt{\mathbb{E}(X \cdot \theta)^2} \leq c_2 \mathrm{Vol}_n(K).$$

Proof. Denote by f the density of $X \cdot \theta$. Then,

$$f(t) = \frac{\mathrm{Vol}_{n-1}\left(K \cap \{x \in \mathbb{R}^n : \ X \cdot \theta = t\}\right)}{\mathrm{Vol}_n(K)}.$$

It is easy to see that f is a log-concave probability density and

$$\int_{-\infty}^{+\infty} t f(t)\, dt = 0.$$

Now, what we want to prove is equivalent to

$$c_1 \leq f(0)\sqrt{\int_{-\infty}^{+\infty} t^2 f(t)\, dt} \leq c_2. \tag{3.1}$$

To prove the right-hand side of (3.1) we use Lemma 2.2.3 for $p = -1, 0, 2$ to get

$$M_f^3(0) \geq M_f(2)^1 M_f(-1)^2.$$

This is equivalent to

$$f^2(0)\int_0^{\infty} t^2 f(t)\, dt \leq 2\left(\int_0^{\infty} f(t)\, dt\right)^3 \leq 2\int_0^{\infty} f(t)\, dt.$$

Repeating the argument for $(-\infty, 0)$ and then adding the inequalities

$$f^2(0)\int_{-\infty}^{+\infty} t^2 f(t)\, dt \leq\leq 2\int_{-\infty}^{+\infty} f(t)\, dt = 2.$$

Now to prove the left hand side inequality, use first Lemma 3.0.15

$$\sup f \leq e f(0)$$

Then choose a half line with measure greater or equal to $\frac{1}{2}$, say

$$\mathbb{P}\left(X\cdot\theta \geq 0\right) \geq \frac{1}{2}.$$

Since

$$\mathbb{P}\left(0 \leq X\cdot\theta \leq \frac{1}{4\sup f}\right) = \int_0^{\frac{1}{4\sup f}} f(t)\, dt \leq \frac{1}{4},$$

we obtain

$$\mathbb{P}\left(X\cdot\theta \geq \frac{1}{4\sup f}\right) \geq \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

Using this, we have

$$\mathbb{E}(X\cdot\theta)^2 \geq \frac{1}{4}\left(\frac{1}{4\sup f}\right)^2 \geq \frac{c}{f^2(0)}.$$

■

The following is an immediate consequence.

Corollary 3.0.20 *Let K be a convex body on $\mathbb{R}^n$ with $\mathrm{Vol}_n(K) = 1$. Let also $X \sim \mathrm{Unif}(K)$, with $\mathbb{E}X = 0$ and $\mathrm{Cov}(X)$ to be a scalar matrix. Then for each $H \subset \mathbb{R}^n$ hyperplane through 0,*

$$\frac{c_1}{L_K} \leq \mathrm{vol}_{n-1}(K\cap H) \leq \frac{c_2}{L_K}.$$

Exercise 3.0.21 *Let K be a convex body on $\mathbb{R}^n$ with $\mathrm{Vol}_n(K) = 1$. Then we have that:*

(i)

$$L_K \leq c\sqrt{n}.$$

A hint for this one, is to observe that in some direction the width is $\leq c\sqrt{n}$.

(ii)

$$L_n \leq c \sup \frac{1}{\sup \mathrm{Vol}_{n-1}(K\cap H)},$$

where the first supremum is taken over all K convex body on $\mathbb{R}^n$ with $\mathrm{Vol}_n(K) = 1$ and the second supremum is taken over all hyperplanes $H \subset \mathbb{R}^n$.

(iii) When $X \sim \mathrm{Unif}(K)$ is isotropic and the covariance matrix is the identity, then

$$K \subset cnB_2^n.$$

We continue with a lemma.

Lemma 3.0.22 *Let X be a log-concave random vector in $\mathbb{R}^n$ with $\mathbb{E}X = 0$ and $\mathrm{Cov}(X)$ to be scalar. If f is the density of X, then for any subspace $E \subset \mathbb{R}^n$,*

$$\left(\int_E f\right)^{\frac{1}{\mathrm{codim}(E)}} \geq c\frac{f(0)^{\frac{1}{n}}}{L_X}.$$

Proof. Denote

$$Y = \mathrm{Proj}_E X.$$

Then Y is also log-concave and $\mathbb{E}Y = 0$. Also notice that

$$\mathrm{Cov}(Y) = \frac{L_X^2}{f(0)^{\frac{2}{n}}}\mathrm{Id}.$$

So,

$$c \leq L_Y = \frac{L_X}{f(0)^{\frac{1}{n}}}\left(\int_E f\right)^{\frac{1}{\mathrm{codim}(E)}},$$

since $\int_E f$ is the density of Y at 0.

■

In what follows, for an $l-$dimensional convex body K, denote

$$\mathrm{v.rad}(K) = \left(\frac{\mathrm{vol}_l(K)}{\mathrm{vol}_l(B^l)}\right)^{\frac{1}{l}}$$

We can now formulate the following theorem.

Theorem 3.0.23 *Let K be a convex body on $\mathbb{R}^n$ and $1 \leq l \leq n$ with $l = \lambda n$. Choose $E \in G_{n,l}$ randomly. Then with probability grater than $1 - e^{-n}$*

$$\mathrm{diam}(K \cap E)^{1-\lambda}\mathrm{v.rad}(K \cap E)^{\lambda} \leq c\mathrm{v.rad}(K).$$

In order to prove the above theorem, we need a lemma first.

Lemma 3.0.24 *Let K be a convex body on $\mathbb{R}^l$ and $n \geq l$. Then*

$$\mathrm{Vol}_l(K)\mathrm{diam}(K)^{n-l} \leq c^l \int_K |x|^{n-l}\, dx.$$

Proof. Take $x_0 \in \overline{K}$ with

$$|x_0| = \sup_{x \in K} |x|.$$

Then it is clear that $\mathrm{diam}(K) \leq 2|x_0|$. Due to convexity of K we have that

$$\frac{3}{4}x_0 + \frac{1}{4}K \subset K.$$

Notice that for every $x \in \frac{3}{4}x_0 + \frac{1}{4}K$,

$$|x| \geq \frac{3}{4}|x| - \frac{1}{4}\sup_{x \in K}|x| = \frac{1}{2}|x_0|.$$

This is the key point for the proof since

$$\int_K |x|^{n-l}\, dx \geq \int_{\frac{3}{4}x_0 + \frac{1}{4}K} |x|^{n-l}\, dx \geq \left(\frac{1}{2}|x_0|\right)^{n-l} \cdot \mathrm{Vol}_l\left(\frac{3}{4}x_0 + \frac{1}{4}K\right)$$
$$\geq c^n \mathrm{diam}(K)^{n-l}\mathrm{Vol}_l(K).$$

■

Now we are ready to prove the theorem.
Proof of Theorem 3.0.23. Integration in polar coordinates and using Lemma 3.0.24,

$$\begin{aligned}
\mathrm{Vol}_n(K) &= \int_{S^{n-1}} \int_0^\infty \mathbf{1}_K(r\theta) r^{n-1}\, dr d\theta = \mathrm{Vol}_{n-1}(S^{n-1}) \int_{S^{n-1}} \int_0^\infty \mathbf{1}_K(r\theta) r^{n-1}\, dr d\sigma_{n-1}(\theta) \\
&= \mathrm{Vol}_{n-1}(S^{n-1}) \int_{G_{n,l}} \int_{S^{n-1}\cap E} \int_0^\infty \mathbf{1}_K(r\theta) r^{n-1}\, d\sigma_{n,l} dr d\sigma_E(\theta) \\
&= \frac{\mathrm{Vol}_{n-1}(S^{n-1})}{\mathrm{Vol}_{l-1}(S^{l-1})} \int_{G_{n,l}} \left[\int_E \mathbf{1}_K(x)|x|^{n-l}\, dx \right] d\sigma_{n,l}(E) \\
&= \frac{n\mathrm{Vol}_n(B^n)}{l\mathrm{Vol}_l(B^l)} \int_{G_{n,l}} \int_{K\cap E} \mathbf{1}_K(x)|x|^{n-l}\, dx\, d\sigma_{n,l}(E) \\
&\geq c^n \frac{n\mathrm{Vol}_n(B^n)}{l\mathrm{Vol}_l(B^l)} \int_{G_{n,l}} \mathrm{Vol}_l(K\cap E)\mathrm{diam}(K\cap E)^{n-l}\, d\sigma_{n,l}(E).
\end{aligned}$$

After some algebra we arrive at

$$\mathrm{v.rad}(K) \geq c \left(\int_{G_{n,l}} \mathrm{v.rad}(K\cap E)^{\lambda n} \mathrm{diam}(K\cap E)^{(1-\lambda)n}\, d\sigma_{n,l}(E) \right)^{\frac{1}{n}}.$$

By Markov-Chebychev with probability $\geq 1 - e^{-n}$,

$$\mathrm{diam}(K\cap E)^{1-\lambda} \mathrm{v.rad}(K\cap E)^\lambda \leq c\mathrm{v.rad}(K).$$

■

The following theorem is due to Kashin.

Theorem 3.0.25 *Let*

$$B_1^n = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^n |x_i| \leq 1 \right\}.$$

Select a random subspace $E \in G_{n,n/2}$. Then with probability $\geq 1 - e^{-n}$,

$$\frac{1}{\sqrt{n}}(B^n \cap E) \subset K \cap E \subset \frac{c}{\sqrt{n}}(B^n \cap E).$$

Note that if $x \in \dfrac{1}{\sqrt{n}} B_2^n$ then by Cauchy-Schwarz inequality we have that

$$\sum_{i=1}^n |x_i| \leq \sqrt{n} \sqrt{\sum_{i=1}^n x_i^2} \leq 1.$$

So $\dfrac{1}{\sqrt{n}} B_2^n \subset B_1^n$. Also

$$\mathrm{Vol}_n(B_1^n) = \frac{2^n}{n!} = \left(\frac{2e + o(1)}{n} \right)^n$$

so

$$\left(\frac{\mathrm{Vol}_n(B_1^n)}{\mathrm{Vol}_n\left(\frac{1}{\sqrt{n}} B_2^n\right)} \right)^{\frac{1}{n}} \leq C.$$

The following theorem shows that, in a sense, we can obtain the reverse result.

Theorem 3.0.26 *Let K be a convex body on $\mathbb{R}^n$. Suppose that $R > 0$ is such that*

$$B^n \subset K$$

and

$$\left(\frac{\mathrm{Vol}_n(K)}{\mathrm{Vol}_n(B^n)}\right)^{\frac{1}{n}} \leq R.$$

Then for $E \in G_{n,n/2}$ with probability $> 1 - e^{-n}$

$$B^n \cap E \subset K \cap E \subset c(R)\; B^n \cap E.$$

Proof. Since $B^n \subset K$ we have that for any subspace $E \in G_{n,n/2}$

$$\mathrm{v.rad}(K \cap E) \geq 1.$$

Now from Theorem 3.0.23 we have that with probability $> 1 - e^{-n}$

$$\mathrm{v.rad}(K \cap E)^{\frac{1}{2}} \mathrm{diam}(K \cap E)^{\frac{1}{2}} \leq \mathrm{v.rad}(K) \leq R,$$

so

$$\mathrm{diam}(K \cap E) \leq CR^2.$$

We begin first with the definition of the polar body and some easy properties.
Let $K \subseteq \mathbb{R}^n$ convex with K =$-K$. Then we define

$$K^\circ = \{x \;\in\; \mathbb{R}^n : \forall\, y \;\in K \;\; x \cdot y \leq 1\}$$

We will need two basic properties.

$$K \subseteq R\; \cdot B^n \Rightarrow K^\circ \supseteq \frac{1}{R} B^n$$

and

$$(K \cap F)^\circ = \mathrm{Proj}_E K^\circ.$$

Santalo ’ s inequality :
Let $K \;\subseteq\; \mathbb{R}^n$ be a convex body with $K = -K$. Then

$$|K| \cdot |K^\circ| \leq \; |B^n|^2$$

with equality if and only if K is an ellipsoid. Define

$$s(K) = \mathrm{v.rad}(K) \cdot \mathrm{v.rad}(K^\circ)$$

Then we can rewrite Santalo’s inequality as

$$\mathrm{v.rad}(K) \cdot \mathrm{v.rad}(K^\circ) \leq 1$$

We have also a reserve inequality of that type due to Bourgain and Milman.

Theorem 3.0.27 *There exists a universal constant $c > 0$ such that for every $K \subseteq \mathbb{R}^n$ convex with $K = -K$,*

$$s(K) \geq c$$

Lemma 3.0.28 *Let $K \subseteq \mathbb{R}^n$ be convex with $K = -K$. Assume that $f : K \to [0, +\infty)$ is a probability log-concave density, such that*

$$\forall\, x \in K \;\; e^{2n} \leq f(x) \leq e^{2n}.$$

Let X be distributed according to f, and that $\mathrm{cov}(X)$ is a scalar matrix. Let $E \in G_{n,l}$ be random with $\lambda = \frac{l}{n}$. Then, with probability $\geq 1 - e^{-n}$

- $\mathrm{v.rad}(K \cap E) \geq c_\lambda \sqrt{n} L_f$
- $\mathrm{diam}(K \cap E) \geq c_\lambda \sqrt{n} L_f$
- $s(K \cap E) \geq \dfrac{c_\lambda}{(L_f)^{1/2}}$

Proof. Denote $x_0 = \mathbb{E}X$. From the Lemma 3.0.22, with prob $1 - e^{-n}$

$$\left(\int_{E+x_0} f \right)^{\frac{1}{n-l}} \geq c \frac{f(x_0)^{1/n}}{L_X} \geq \frac{\tilde{c}}{L_x}$$

So we have that

$$|K \cap (E + x_0)| \geq e^{-2n} \int_{E+x_0} f \geq c^n \cdot \frac{1}{L_X^{n-1}}$$

which means that

$$\mathrm{v.rad}(K \cap (E \; + x_0)) \geq \frac{c^{\frac{n}{l}} \sqrt{l}}{L_X^{\frac{n-l}{l}}} \geq c^{\frac{1}{\lambda}} \sqrt{\frac{1}{\lambda}} \frac{\sqrt{n}}{L_X^{\frac{1}{\lambda}-1}}.$$

But

$$\frac{K \cap (E + x_0) + K \cap (E - x_0)}{2} \subseteq K \cap E$$

This gives the proof for the first, since

$$\mathrm{v.rad}(K \cap E) \geq \sqrt{\mathrm{v.rad}(K \cap (E + x_0)^2)} = \mathrm{v.rad}(K \cap (E \; + x_0))$$

For the second use Theorem 3.0.23 to get that with probability $1 - e^{-n}$

$$\mathrm{diam}(K \cap E)^{1-\lambda} \mathrm{v.rad}(K \cap \; E)^\lambda \leq c \cdot \sqrt{n} |K|^{\frac{1}{n}} \leq c\sqrt{n} \left(e^{2n} \int_K f \right)^{\frac{1}{n}} \leq \tilde{c}\sqrt{n}$$

So,

$$\mathrm{diam}(K \cap E) \leq c_\lambda \sqrt{n}^{\frac{1}{1-\lambda}} \left(\mathrm{v.rad}(K \cap E)\right)^{-\frac{\lambda}{1+\lambda}} \leq c_\lambda \sqrt{n}^{\frac{1}{1-\lambda}} \left(\sqrt{n}\, L_X^{1-\frac{1}{\lambda}} \right)^{-\frac{\lambda}{1+\lambda}} = c_\lambda \sqrt{n} L_X$$

For the third one we use the first two use the first two to get that with high probability,

$$(K \cap E)^\circ \supseteq \frac{c}{\sqrt{n} L_X} B^n \cap E.$$

This gives that

$$\mathrm{v.rad}((K \cap E)^\circ) \cdot \mathrm{v.rad}(K \cap E) \geq \frac{c}{\sqrt{n} L_X} \cdot c\sqrt{n} L_X^{1-\frac{1}{\lambda}} = \frac{c}{L_X^{\frac{1}{\lambda}}}.$$

■

Corollary 3.0.29 *Let $K \subseteq \mathbb{R}^n$ be a convex body and $f : K \to [0, \infty)$ a log-concave integrable function with*

$$\frac{\sup f}{\inf f} \leq e^{4n}.$$

Then, for all l there exists $E \in G_{n,l}$ with

$$s(K \cap E) \geq \frac{c_\lambda}{L_f^{\frac{1}{\lambda}}},$$

where $\lambda = \frac{l}{n}$.

Corollary shows that we need a log concave density f not too far from uniform measure on K, with control on L_f. To do this, we introduce the log-Laplace transformation. Suppose X is a log-concave random vector in $\mathbb{R}^n$. Set

$$\Lambda_X(\xi) = \log \mathbb{E} e^{X \cdot \xi},$$

the log-Laplace transform. It is easy to see that Λ_X is C^∞-smooth in the open set $\{\Lambda_X < +\infty\}$. Also, given $\xi \in \mathbb{R}^n$, set $f_\xi(x)$ to be the probability density proportional to $x \to e^{\xi \cdot x} f(x)$ where f is the density of X. Finally, the random vector X_ξ, with density f_ξ is log-concave when $\xi \in \{\Lambda_x < +\infty\}$.

Lemma 3.0.30 *For any ξ with $\Lambda(\xi) < +\infty$,*

$$\nabla \Lambda(\xi) = \mathbb{E} X_\xi, \ \ \nabla^2 \Lambda(\xi) = Cov(X_\xi) > 0.$$

In particular, Λ_ξ is strictly convex.

Proof. Computing the partial derivatives we find,

$$\partial_i \Lambda(\xi) = \frac{\displaystyle\int_{\mathbb{R}^n} x_i e^{\xi \cdot x} f(x)\, dx}{\displaystyle\int_{\mathbb{R}^n} e^{\xi \cdot x} f(x)\, dx} = \mathbb{E}\left(X_\xi \cdot e_i\right)$$

and

$$\partial_{ij} \Lambda(\xi) = \mathbb{E}(X_\xi \cdot e_i)(X_\xi \cdot e_i) - \mathbb{E}(X_\xi \cdot e_i) \cdot \mathbb{E}(X_\xi \cdot e_j).$$

■

Corollary 3.0.31 *Let X be a random vector with $X \sim Unif(K)$, where K is convex with $K = -K$. Then,*

$$\int_{\mathbb{R}^n} \det \mathrm{Cov}(X_\xi) \leq \mathrm{Vol}_n(K).$$

Proof. Since Λ is strictly-convex, the function $x \to \nabla \Lambda(x)$ is 1-1. So we can change variables to have

$$\mathrm{Vol}_n(K) = \int_K 1\, dx \geq \int_{\nabla \Lambda(\mathbb{R}^n)} 1\, dx = \int_{\mathbb{R}^n} \det \nabla^2 \Lambda(y)\, dy = \int_{\mathbb{R}^n} \det \mathrm{Cov}(X_\xi) d\xi$$

■

Proposition 3.0.32 *Let $K \subseteq \mathbb{R}^n$ be convex with $K = -K$. Let $\varepsilon > 0$. Then, there exists $\xi \in \varepsilon n K^\circ$ with*

$$L_{X_\xi} \leq \frac{c}{\sqrt{\varepsilon \cdot s(K)}}$$

where $X \sim Unif(K)$.

Proof. Using Corollary 3.0.31 we have,

$$\int_{\varepsilon n K^\circ} \det \mathrm{Cov}(X_\xi)\, d\xi \leq \int_{\mathbb{R}^n} \det \mathrm{Cov}(X_\xi)\, d\xi \leq |K|,$$

so

$$\frac{1}{|\varepsilon n K^\circ|} \int_{\varepsilon n K^\circ} \det \mathrm{Cov}(X_\xi)\, d\xi \leq \frac{|K|}{|\varepsilon n K^\circ|}$$

Hence there exists $\xi \in \varepsilon n K^\circ$ with

$$\det \mathrm{Cov}(X_\xi) \leq \frac{|K|}{\varepsilon^n \cdot |nK^\circ|}$$

Now we compute,

$$\begin{aligned} L_{X_\xi} \leq \det \mathrm{Cov}(X_\xi)^{\frac{1}{2n}} \cdot (\sup f_\xi)^{\frac{1}{n}} &\leq \left(\frac{|K|}{\varepsilon^n |nK^\circ|} \right)^{\frac{1}{2n}} \cdot \left(\sup_{x \in K} \frac{e^{\xi \cdot x}/|K|}{e^{\Lambda(\xi)}} \right)^{\frac{1}{n}} \\ &\leq \frac{1}{\sqrt{\varepsilon}} \cdot \frac{|K|^{\frac{1}{2n}}}{|nK^\circ|^{\frac{1}{2n}}} \cdot \frac{e^{\xi}}{|K|^{\frac{1}{n}}} \\ &\leq \frac{c}{\sqrt{\varepsilon}} \cdot \frac{1}{\sqrt{n}|K^\circ|^{\frac{1}{2n}} \cdot |K|^{\frac{1}{2n}}} = \frac{c}{\sqrt{\varepsilon s(K)}} \end{aligned}$$

where we used that since Λ is convex even and $\Lambda(0) = 0$ we have that $\Lambda \geq 0$.

■

So if we take $\varepsilon = \frac{1}{2}$, given $K \subseteq \mathbb{R}^n$ convex with $K = -K$, we found a log-concave random vector $Y = X_\xi$ supported in K, such that for the density of Y, denoted by f, we have

$$\frac{\sup_K f}{\inf_K f} \leq \frac{e^{\frac{n}{2}}}{e^{\frac{-n}{2}}} = e^n$$

and

$$L_y \leq \frac{c}{\sqrt{s(K)}}.$$

This fits with Corollary 3.0.29. Then, for all l there exists $E \in G_{n,l}$ with

$$s(K \cap E) \geq \frac{c_\lambda}{s(K)^{\frac{1}{2\lambda}}},$$

where $\lambda = \frac{l}{n}$. As a main lemma for the proof of Bourgain-Milman inequality, we use a result of Giannopoulos,Vritsiou,Paouris.

Lemma 3.0.33 *Let $K \subseteq \mathbb{R}^n$ be convex body with $K = -K$. Then, there exists $E \in G_{n,l}$ with $l = \lfloor \frac{3}{4} n \rfloor$ such that*

$$s(K \cap E)^\lambda \geq c\sqrt{s(K)}.$$

From the proof, we need a classical result, the Rogers-Shephard's inequality.

Lemma 3.0.34 *Let $K \subseteq \mathbb{R}^n$ be convex body with $0 \in K$ and let $E \in G_{n,l}$. Then,*

$$|K \cap E| \cdot |\mathrm{Proj}_{E^\perp} K| \leq 4^n |K|.$$

Proof. First, observe that

$$\mathrm{Vol}_n(K) = \int_{\mathrm{Proj}_{E^\perp}(K)} Vol_l(K \cap (E + x))\, dx$$

Also:
Claim: If $x \in \frac{1}{2}\mathrm{Proj}_{E^\perp}(x)$, then $K \cap (E + x)$ contains a translation of $\frac{K \cap E}{2}$.
Explanation: Take $x \in \frac{1}{2}\mathrm{Proj}_{E^\perp} K$. Then, there exists $y \in K$ such that $2x = \mathrm{Proj}_{E^\perp} y$. Thus,

$$\frac{y + (K \cap E)}{2} \subseteq K \cap \left(\frac{E + y}{2} \right) = K \cap (E + x).$$

$\lozenge$

Now the proof is immediate since

$$\mathrm{Vol}_n(K) \geq \int_{\frac{1}{2}\mathrm{Proj}_{E^\perp} K} |K \cap (E + x)|\, dx \geq \frac{1}{2^n} \cdot \frac{1}{2^n} |Proj_{E^\perp} K| \cdot |K \cap E|$$

$\blacksquare$

Proof of 3.0.33: Assume first that $4|n$, then $\lambda = \frac{l}{n} = \frac{3}{4}$. By Rogers-Shephard's inequality,

$$|K| \geq 4^{-n} |K \cap E| \cdot |\mathrm{Proj}_{E^\perp} K|$$

and also

$$|K^\circ| \geq 4^{-n} |K^\circ \cap E^\perp| \cdot |\mathrm{Proj}_E K^\circ|.$$

Multiplying we have that:

$$|K| \cdot |K^\circ| \geq |K \cap E| \cdot |\mathrm{Proj}_{E^\perp} K| \cdot |K^\circ \cap E^{perp}| \cdot |\mathrm{Proj}_E K^\circ| \cdot 16^{-n}$$

so

$$|B^n|^2 (s(K))^n \geq |B^l|^2 s(K \cap E)^l |B^{n-l}|^2 \cdot s(K^\circ \cap E^\perp) \cdot 16^{-n}$$

But $|B^n|^{\frac{1}{\sqrt{n}}} \sim \frac{1}{\sqrt{n}}$, $|B^l|^{\frac{1}{l}} \sim \frac{1}{l} \sim \frac{1}{\sqrt{n}}$ and $|B^{n-l}|^{\frac{1}{n-l}} \sim \frac{1}{\sqrt{n}}$. and so

$$s(K) \geq s(K \cap E)^\lambda \cdot s(K^\circ \cap E^\perp)^{1-\lambda}.$$

$\blacksquare$

We left as an exercise to eliminate the assumptions that $4|n$. Define

$$s_N = \inf\{s(T) : \; T \subseteq \mathbb{R}^n \; convex, \; 1 \leq n \leq N, \; T = -T\}$$

This infimum is attained and $s_N > 0$.
Using Lemma 3.0.33

$$s(K) \geq c \cdot \sqrt{s_N},$$

so

$$s_N \geq c \cdot \sqrt{s_N} \Rightarrow s_N \geq c,$$

which proves Bourgain-Milman's inequality.

■

As a consequence of Bourgain-Milman's inequality and 3.0.32 we have the following theorem.

Theorem 3.0.35 *Let $K \subseteq \mathbb{R}^n$ be convex with $K = -K$. Let $\varepsilon > 0$. Then, there exists $\xi \in \varepsilon n K^\circ$ with*

$$L_{X_\xi} \leq \frac{c}{\sqrt{\varepsilon}}$$

where $X \sim Unif(K)$.

We left as an exercise to remove the assumption $K = -K$ by the assumption that K has its barycenter at 0. The first consequence of this theorem is the following corollary.

Corollary 3.0.36 *Let $K \subseteq \mathbb{R}^n$ be a convex body and $0 < \varepsilon < 1$. Then there exists a convex body $T \subseteq \mathbb{R}^n$ with the following properties:*

$$(1 - \varepsilon) \cdot T \subseteq K \subseteq T.$$

and

$$L_T \sim L_{X_\xi} \leq \frac{c}{\sqrt{\varepsilon}}.$$

One more consequence of the theorem is the following Milman's ellipsoid theorem.

Theorem 3.0.37 *Let $K \subseteq \mathbb{R}^n$ be a convex body such that $K = -K$. Then there exists a centrally symmetric ellipsoid $\mathcal{E}$ such that:*

$$(1)\mathrm{Vol}_n(\mathcal{E}) = \mathrm{Vol}_n(K)$$
$$(2)\mathrm{Vol}_n(K) \leq C^n \cdot \mathrm{Vol}_n(\mathcal{E} \cap K),$$

where $C > 0$ is a universal constant. The ellipsoid $\mathcal{E}$ is a Milman ellipsoid with constant C associated with K.

Comment: The hyperplane conjecture for K is equivalent to the statement that the inertia ellipsoid of K is proportional to a Milman ellipsoid. **Proof of Theorem 3.0.37:** Use Theorem 3.0.35 with $\varepsilon = \frac{1}{2}$. There exists a random vector Y supported in K such that,

$$L_Y < c$$

and the log-concave density of Y, which is g, satisfies,

$$e^{-cn} \leq g(x) \leq e^{cn}.$$

Writing the relation for the isotropic constant of Y,

$$1 \sim L_Y = \det \mathrm{Cov}(Y)^{\frac{1}{2n}} \cdot (supg)^{\frac{1}{n}}$$

so

$$\det \mathrm{Cov}(Y)^{\frac{1}{2n}} \sim 1.$$

Let T be an affine map from $\mathbb{R}^n$ to $\mathbb{R}^n$ such that $T(Y)$ is isotropic. Then,

$$\det \mathrm{Cov}(T(Y))^{\frac{1}{n}} = \det(T)^{\frac{2}{n}} \cdot \det \mathrm{Cov}(Y)^{\frac{1}{n}}$$

so

$$\det(T)^{\frac{1}{n}} \sim 1.$$

Using Markov's inequality,

$$\mathbb{E}|T(Y)|^2 = n \Rightarrow \mathbb{P}\left(|T(Y)| \geq 2\sqrt{n}\right) \leq \frac{1}{4}.$$

Hence,

$$\mathrm{Vol}_n\left(T(K) \cap 2\sqrt{n}B^n\right) \geq \frac{1}{\sup g_{T(Y)}} \cdot \int_{T(K)\cap 2\sqrt{n}B^n} g_{T(Y)}(x)\, dx.$$

But

$$g_{T(Y)}(x) = \det(T)^{-1} \cdot g(T^{-1}x)$$

so

$$\mathrm{Vol}_n(T(K) \cap 2\sqrt{n}B^n) \geq c^n \cdot \mathbb{P}\left(|T(Y)| \leq 2\sqrt{n}\right) \geq c^n \cdot \frac{3}{4} \sim \tilde{c}^n$$

Let $\mathcal{E} \subseteq \mathbb{R}^n$ be an ellipsoid, not necessarily symmetric, with

$$T(\mathcal{E}) = 2\sqrt{n}B^n.$$

To summarize what we have till now:

$$(1)\ \mathrm{Vol}_n(T(K) \cap T(E)) \geq c^n$$
$$(2)\ \mathrm{Vol}_n(T(E)) \leq c^n \mathrm{Vol}_n\left(2\sqrt{n}B^n\right) \leq \tilde{c}^n.$$

Hence,

$$(1)\ \mathrm{Vol}_n(K \cap \mathcal{E}) \geq \tilde{c}^n,$$
$$(2)\ \mathrm{Vol}_n(\mathcal{E}) \leq c^n.$$

To finish the proof we need two steps.

Step 1: Rescale $\mathcal{E}_1 = \dfrac{|K|^{\frac{1}{n}}}{|\mathcal{E}|^{\frac{1}{n}}}\mathcal{E}$. Since $|\mathcal{E}|^{\frac{1}{n}} \sim 1$, then $\mathcal{E}_1 \supseteq c\mathcal{E}$ and

$$|K \cap \mathcal{E}_1| \geq |K \cap c\mathcal{E}| \geq |cK \cap c\mathcal{E}| = c^n \cdot |K \cap E| \geq \tilde{c}^n.$$

Step 2:
We have that $1 = |K| = |\mathcal{E}_1|$, and $|K \cap \mathcal{E}_1| \geq c^n$. Let $\mathcal{E}_2 = \mathcal{E}_1 - x_0$ where x_0 is the center of $\mathcal{E}_1$. By Brunn-Minkowski, since

$$K \cap \mathcal{E}_2 \supseteq \frac{K \cap (\mathcal{E}_2 - x_0) + K \cap (\mathcal{E}_2 + x_0)}{2},$$

we have that

$$|K \cap \mathcal{E}_2| \geq |K \cap (\mathcal{E}_2 - x_0)| \geq \tilde{c}^n.$$

So, $\mathcal{E}_2$ is the desired ellipsoid.

■

Exercise 3.0.38 *Use that*

$$\left(\int_E f\right)^{\frac{1}{\mathrm{codim}(E)}} \geq \frac{f(0)^{\frac{1}{n}}}{L_X}$$

and conclude that the ellipsoid $\mathcal{E}_2$ that we constructed satisfies that for all $E \in G_{n,l}$, and $\lambda = \frac{l}{n}$,

$$\mathrm{v.rad}(E \cap K) \geq c_\lambda \cdot \mathrm{v.rad}(E \cap \mathcal{E}).$$

Now we will prove one other form of Rogers-Shepherd inequality.

Proposition 3.0.39 *Let $K, T \subseteq \mathbb{R}^n$ be convex, centrally symmetric. Then,*

$$|K + T| \cdot |K \cap T| \leq 4^n \cdot |K| \cdot |T|. \tag{3.2}$$

Sketch of the proof. Set $A = K \times T$ and define

$$E = \{(x, x) : \ x \in \mathbb{R}^n\} \subseteq \mathbb{R}^{2n}$$

and

$$A \cap E = (x, x) : x \in K \cap T.$$

Thus,

$$\mathrm{Proj}_{E^\perp} A = \{(x, -x) : \ x \in \frac{K + T}{2}\}$$

Now result follows from Lemma 3.0.34.

■

Corollary 3.0.40 *When $\mathcal{E}$ is the Milman ellipsoid of $K \subseteq \mathbb{R}^n$ with a universal constant, then*

$$|K + \mathcal{E}|^{\frac{1}{n}} \leq c|K|^{\frac{1}{n}} = c|\mathcal{E}|^{\frac{1}{n}}. \tag{3.3}$$

Proof. From the above proposition,

$$|K + \mathcal{E}|^{\frac{1}{n}} \cdot |K \cap E|^{\frac{1}{n}} \leq c \cdot |K|^{\frac{1}{n}} \cdot |\mathcal{E}|^{\frac{1}{n}}.$$

Therefore, if $\mathcal{E}$ is Milman ellipsoid of K, then $\mathcal{E}^\circ$ is Milman ellipsoid of K°.
Explanation: From Bourgain-Milman inequality,

$$\mathrm{v.rad}(\mathcal{E}^\circ) \sim \mathrm{v.rad}(K)$$

Thus,

$$|\mathcal{E}^\circ \cap K^\circ|^{\frac{1}{n}} = |conv(K, \mathcal{E})^\circ|^{\frac{1}{n}} \sim |(K + \mathcal{E})^\circ|^{\frac{1}{n}} \geq \frac{1}{n \cdot |K + \mathcal{E}|^{\frac{1}{n}}} \geq \frac{c}{n \cdot |K|^{\frac{1}{n}}} \geq c \cdot |K^\circ|^{\frac{1}{n}}.$$

We conclude with one more corollary.

Corollary 3.0.41 *Let $K \subseteq \mathbb{R}^n$, be a convex body and assume that B_2^n is a Milman ellipsoid for K with a universal constant. Let $E \in G_{n,\frac{n}{2}}$ be random. Then with probability $\geq 1 - e^{-n}$:*

(1) $\mathrm{diam}(K \cap E) \leq c$
(2) $\mathrm{Proj}_E K \supseteq cB^n$
(3) $E \cap B^n$ *is a Milman ellipsoid for* $K \cap E, \mathrm{Proj}_E K,$ *with another universal constant.*

Proof. For the first one,

$$\mathrm{diam}(K \cap E)^{\frac{1}{2}} \cdot \mathrm{v.rad}(K \cap E)^{\frac{1}{2}} \leq c,$$

hence,

$$\mathrm{diam}(K \cap E) \leq c.$$

For the second, just dualize the first: B^n is a Milman ellipsoid for K° and use it for K°. We left the final one as an exercise.