MATH 7251: High-Dimensional Probability Lecture Notes

Galyna V. Livshyts

Scribes: Max Dabagia, Ruijia Cao, Dekang Meng, Xuanzhou Chen, Miao Li, Wanying Fu, Rakshit Naidu, Hanyang Jiang, Rahul Sethi, Chengyue Huang, Yuzhou Wang, William Held, Sam van der Poel, Akash Harapanahalli, Jingyi Zhang, Chen Lin, Daqian Bao, Zichong Li, Adam Brown, Oscar Hasburn-Babich, Shikun Liu, José R. Tuirán Rangel, Yinan Huang, Tian-Yi Zhou, Venkatakrishnan Vaidyanathapuram Krishnan, Jonghyeok Lee, Mengyu Yang

Abstract

Disclaimer: these lecture notes are currently under construction, and may not be fully proofread yet. If you spot a typo, please let me know! Also, references are in the process of being matched to the text.

Contents

1 Notation and Preliminaries

1.1 Notation

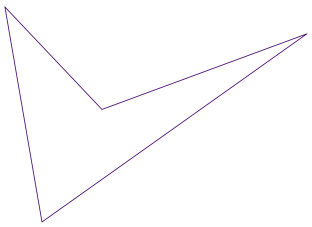
- $\mathbb{R}^n$ is the $n-$dimensional Euclidean space

- Lebesgue measure (volume) in $\mathbb{R}^n$ of a measurable set $A \subset \mathbb{R}^n$ is denoted by $|A|$
- $\mathbb{N}$ is the set of positive integers
- $\|x\|_p = (|x_1|^p + \cdots + |x_n|^p)^{1/p}$ is the p-norm in $\mathbb{R}^n$ for $p \geq 1$
- $\|x\|_\infty = \max_{i=1,\ldots,n} |x_i|$ is the ∞-norm
- $|x| = \|x\|_2$ is a shorthand for Euclidean length
- $\mathbf{B}_p^n = \{x \in \mathbb{R}^n \colon \|x\|_p \leq 1\}$ is p-ball in $\mathbb{R}^n$
- $\mathbb{S}^{n-1} = \partial \mathbf{B}_2^n = \{x \in \mathbb{R}^n \colon |x| = 1\}$ is the n-dimensional hypersphere, or the boundary of B_2^n
- $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$ is the standard inner product
- For $\theta \in \mathbb{S}^{n-1}$, we have the hyperplane
$$\theta^\perp = \{x \in \mathbb{R}^n \colon \langle x, \theta \rangle = 0\}$$
and the affine hyperplane
$$\theta^\perp + t\theta = \{x \in \mathbb{R}^n \colon \langle x, \theta \rangle = t\}$$
for all $t \in \mathbb{R}$.
- A half-space is a set of the form $\{x \in \mathbb{R}^n : \langle x, \theta \rangle \leq t$, for some given $\theta \in \mathbb{S}^{n-1}$ and $t \in \mathbb{R}$.
- A strip is a set of the form $\{x \in \mathbb{R}^n : |\langle x - y, \theta \rangle| \leq t$, for some given $y \in \mathbb{R}^n$, $\theta \in \mathbb{S}^{n-1}$ and $t \geq 0$.
- Fix $x \in \mathbb{R}$. Then $[x]$ is the floor function, the largest integer which is no larger than x.

1.2 Preliminaries from geometry in High dimension and convexity

Definition 1.1 (Convex Set)**.** A set $K \subseteq \mathbb{R}^n$ is called *convex* if for all $x, y \in K$, the line segment $[x, y] = \{\lambda x + (1 - \lambda)y : \lambda \in [0, 1]\}$ is contained in K.

Note that strips, half-spaces and B_p^n for $p \geq 1$ are convex sets. On the other hand, B_p^n for $p < 1$, $\mathbb{S}^{n-1}$, sets which are not 1-connected, are non-convex. Below see an example of a non-convex set:



Any convex set is the intersection of (possibly infinitely many) half-spaces. A convex *polytope* is an intersection of finitely many half-spaces.

1.3 Preliminaries from linear algebra

The operator norm $\|\cdot\|_{op}$ of a matrix A is defined by

$$\|\mathbf{A}\|_{op} = \sup_{x \in \mathbb{R}^n \setminus \{0\}} \frac{|\mathbf{A}x|}{|x|} = \sup_{y \in \mathbb{S}^{n-1}} |Ay|. \tag{1}$$

Definition 1.2 (Hilbert-Schmidt norm)**.** Given a matrix $\mathbf{A} = (a_{ij})$,

$$\|\mathbf{A}\|_{HS} = \sqrt{\sum_{i,j} a_{ij}^2} = \sqrt{\sigma_1^2 + ... + \sigma_n^2},$$

where $\sigma_1 \geq \cdots \geq \sigma_n$ are the singular values of $\mathbf{A}$.

Recall that the smallest singular value is also defined as

$$\sigma_n(A) = \inf_{x \in \mathbb{S}^{n-1}} |Ax|.$$

Remark 1.3. *Note that*

$$\sqrt{n}\|\mathbf{A}\|_{op} \geq \|\mathbf{A}\|_{HS} \geq \|\mathbf{A}\|_{op},$$

since $\sigma_i \leq \sigma_1$*, and* $\sigma_1 \leq \sqrt{\sigma_1^2 + ... + \sigma_n^2} \leq \sqrt{n}\sigma_1$*.*

Theorem 1.4 (Spectral Decomposition)**.** *Let A be a symmetric matrix over $\mathbb{R}$ with n eigenvalues $\lambda_1, \cdots, \lambda_n \in \mathbb{R}$ and corresponding eigenvectors $u_1, \cdots, u_n \in \mathbb{S}^{n-1}$. Then*

$$A = \sum_{i=1}^{n} \lambda_i u_i \otimes u_i \tag{2}$$

Note that for $\forall u \in \mathbb{R}^n, u \otimes u = uu^T = \begin{bmatrix} u_1^2 & u_1u_2 & \cdots & u_1u_n \\ u_1u_2 & u_2^2 & \cdots & u_2u_n \\ \vdots & \vdots & \vdots & \vdots \\ u_1u_n & u_2u_n & \cdots & u_nu_n \end{bmatrix}$ *is a rank-1 matrix. For* $\forall x \in \mathbb{R}^n$*,* $Ax = \sum_{i=1}^n \lambda_i \langle x, u_i \rangle u_i$*.*

Definition 1.5 (Functions on Matrices)**.** For any function $f : \mathbb{R} \to \mathbb{R}$ and $n \times n$ symmetric matrix $X = \sum_{i=1}^n \lambda_i u_i \otimes u_i$, then $f(X) := \sum_{i=1}^n f(\lambda_i) u_i \otimes u_i$.

1.4 Preliminaries from Functional Analysis and Probability

Recall that for any non-negative random variable

$$\mathbb{E}\, X = \int_0^\infty \mathbb{P}(X > t) dt. \tag{3}$$

Indeed, $X = \int_0^\infty 1_{\{X>t\}} dt$, and therefore

$$\mathbb{E}\, X = \mathbb{E} \int_0^\infty 1_{\{X>t\}} dt = \int_0^\infty \mathbb{E}\, 1_{\{X>t\}} dt = \int_0^\infty \mathbb{P}(X > t) dt.$$

Lemma 1.6 (Markov)**.** *Let $X \geq 0$ be a random variable defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then, for $t > 0$, we have*

$$\mathbb{P}\left(X > t\right) \leq \frac{\mathbb{E}\, X}{t}.$$

Proof. For any $t > 0$, note that

$$\mathbb{E}\, X = \int_0^\infty \mathbb{P}(X > s)\ \mathrm{d}s \geq \int_0^t \mathbb{P}(X > s)\ \mathrm{d}s \geq t \cdot \mathbb{P}(X > t).$$

□

Definition 1.7. A function $F\colon \mathbb{R}^n \to \mathbb{R}$ is called **convex** if for all $x, y \in \mathbb{R}^n$

$$F(\lambda x + (1-\lambda)y) \leq \lambda F(x) + (1-\lambda)F(y).$$

Remark 1.8. *Note that if F is convex, then we may conclude inductively that*

$$F\left(\sum_{i=1}^m \lambda_i x_i\right) \leq \sum_{1 \leq i \leq m} \lambda_i F(x_i),$$

where $\lambda_1 + \cdots + \lambda_m = 1$.

Theorem 1.9 (Jensen Inequality)**.** *Let μ be any probability measure and $g \in L^1(\mathbb{R}^n)$. Let $F\colon \mathbb{R} \to \mathbb{R}$ be a convex function. Then,*

$$F\left(\int_{\mathbb{R}^n} g\ \mathrm{d}\mu\right) \leq \int_{\mathbb{R}^n} F(g)\ \mathrm{d}\mu.$$

Remark 1.10. *Note that Jensen inequality implies Remark 1.8 by taking μ to be the discrete measure on $\mathbb{R}^n$ with $\operatorname{supp}\mu = \{1, \ldots, m\}$ and $\mu(\{i\}) = \lambda_i$ for $1 \leq i \leq m$.*

Definition 1.11. Let X be a random variable. For the $p \geq 1$, the p-norm of X is defined by

$$\left\|X\right\|_p = \left(\mathbb{E}\left|X\right|^p\right)^{1/p}.$$

Theorem 1.12 (Minkowski)**.** *Let X, Y be two random variables and $p \geq 1$. Then,*

$$\|X + Y\|_p \leq \|X\|_p + \|Y\|_p .$$

Theorem 1.13 (Cauchy-Schwarz)**.** *Let $a, b \in V$, where $(V, \langle \cdot, \cdot \rangle)$ is an inner product space. Then,*

$$|\langle a, b \rangle| \leq \|a\| \cdot \|b\| ,$$

where $\|\cdot\|$ denotes the norm induced by $\langle \cdot, \cdot \rangle$.

Theorem 1.14 (Hölder's Inequality)**.** *Suppose that $p, q \geq 1$ and $1/p + 1/q = 1$. Then,*

$$\int |fg| \, \mathrm{d}\mu \leq \left(\int |f|^p \, \mathrm{d}\mu \right)^{1/p} \cdot \left(\int |g|^q \, \mathrm{d}\mu \right)^{1/q}$$

2 The Probabilistic Method and the concept of the high-dimensional phenomena

2.1 A cool fact about the cube

We will mostly study $\mathbb{R}^n$ – the n-dimensional Euclidean space – where n is a large, positive integer (or, in other words, the dimension n of our space is high). We will often think that $n \to \infty$ and analyze things asymptotically. We know intuitively that functions on $\mathbb{R}^2$ are more complex than those on $\mathbb{R}$, and functions on $\mathbb{R}^3$ are more complex than on $\mathbb{R}^2$. However, in some sense, objects in high-dimensional spaces actually become simpler, and more predictable. We will soon see some examples of this phenomenon, but for the time being we concentrate on the so-called probabilistic method in High-Dimensional Geometry.

We now state the following cool fact:

Fact 2.1. *Consider the cube $[0,1]^n$ (with side-length 1) and pick $x \in [0,1]^n$. Take any (arbitrary!) $\theta \in \mathbb{S}^{n-1}$, and consider the strip centered at x orthogonal to θ of width 1, i.e.*

$$S = \left\{ y \in \mathbb{R}^n \colon |\langle y - x, \theta \rangle| \leq \frac{1}{2} \right\} .$$

Then at least one vertex of $[0,1]^n$ belongs to S.

The proof will be based on the *probabilistic method*. Imagine that someone shows you a non-transparent box full of balls and tells you that if you draw a ball from it then with probability 0.3 you get a red ball from it. Then you can conclude that *there exists* at least one red ball in the box.

In our case, the box will be the vertices of the cube, and the red ball will be the vertex with the desired property of falling into the specific strip.

2.2 Random rounding

We will first define the concept of randomized rounding, which was introduced in one dimension by Raghavan, Thompson [31], and later extended and studied by many authors including Kannan, Vempala [16], Alon, Klartag [2], Klartag, Livshyts [18], see a survey by Srinivasan [42]. This object is very useful in Computer Science, as well as in High-Dimensional Probability and related areas.

Definition 2.2 (Random Rounding)**.** We outline the definition in two steps.

Step 1 (dimension 1). For $x \in \mathbb{R}$, define η_x to be a random variable such that

$$\eta_x = \begin{cases} [x] & \text{w.p. } 1-p \\ [x]+1 & \text{w.p. } p, \end{cases}$$

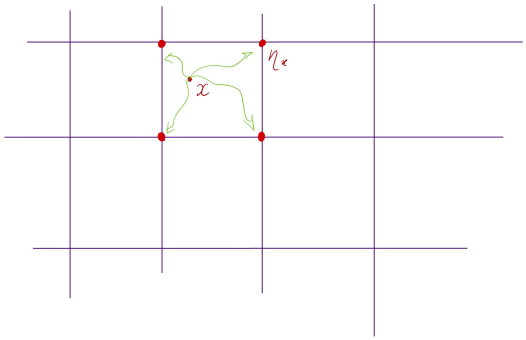
where $p = p(x) = x - [x]$. Note that p is chosen so that η_x is centered at x:

$$\mathbb{E}\,\eta_x = [x](1-(x-[x])) + ([x]+1)(x-[x]) = [x] + x - [x] = x.$$

Step 2 (dimension n). For $x \in \mathbb{R}^n$, define η_x to be a random vector taking values in the vertex set of the lattice cube which x falls into, so that the coordinates of η_x are independent and $\mathbb{E}\,\eta_x = x$. In other words, each coordinate of x is independently randomly rounded using the 1-dimensional definition. Namely, we let $\eta_x = ((\eta_x)_1, ..., (\eta_x)_n)$, where the random variables $(\eta_x)_i$ are independent, and distributed as follows:

$$(\eta_x)_i = \begin{cases} [x_i] & \text{w.p. 1 - p} \\ [x_i]+1 & \text{w.p. p,} \end{cases}$$

where $p = p(i,x) = x_i - [x_i]$,



2.3 Proof of the cool fact about the cube

First, recall the following basic fact:

Lemma 2.3. *For independent random variables $X_1, \dots, X_n$,*

$$\mathrm{Var}(X_1 + \cdots + X_n) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n).$$

Proof. First, consider

$$\mathbb{E}(X_1 + ... + X_n)^2 = \sum_{i=1}^n \mathbb{E}\, X_i^2 + \sum_{i \neq j} 2\, \mathbb{E}\, X_i X_j = \sum_{i=1}^n \mathbb{E}\, X_i^2 + \sum_{i \neq j} 2\, \mathbb{E}\, X_i\, \mathbb{E}\, X_j.$$

where in the last step we used that by independence, $\mathbb{E}\, X_i X_j = \mathbb{E}\, X_i\, \mathbb{E}\, X_j$. On the other hand,

$$(\mathbb{E}\, X_1 + ... + \mathbb{E}\, X_n)^2 = \sum_{i=1}^n (\mathbb{E}\, X_i)^2 + \sum_{i \neq j} 2\, \mathbb{E}\, X_i\, \mathbb{E}\, X_j$$

Therefore,

$$\mathrm{Var}(X_1 + \cdots + X_n) = \mathbb{E}(X_1 + ... + X_n)^2 - (\mathbb{E}(X_1 + ... + X_n))^2 = \sum_{i=1}^n \mathbb{E}\, X_i^2 - \sum_{i=1}^n (\mathbb{E}\, X_i)^2 =$$

$$\mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n).$$

□

We may now proceed with the proof of the cool fact.

Proof of Fact 2.1. For the vector $x \in [0,1]^n$ consider the random rounding η_x. Note that $\mathbb{E}\langle \eta_x - x, \theta \rangle = 0$, so

$$\mathbb{E}\langle \eta_x - x, \theta \rangle^2 = \mathrm{Var}\langle \eta_x - x, \theta \rangle = \sum_{i=1}^n \mathbb{E}((\eta_x)_i - x_i)^2 \theta_i^2,$$

by Lemma 2.3. Hence,

$$\mathbb{E}\langle \eta_x - x, \theta\rangle^2 = \sum_{i=1}^{n} \theta_i^2 (\mathbb{E}(\eta_x)_i^2 - x_i^2).$$

Note that (by definition of random rounding),

$$\mathbb{E}(\eta_x)_i^2 = 0 \cdot (1 - x_i) + 1 \cdot x_i = x_i,$$

so

$$\mathbb{E}(\eta_x)_i^2 - x_i^2 = x_i - x_i^2 \leq \frac{1}{4},$$

where we use the fact that for any $a \in [0,1]$ one has $a - a^2 \leq \frac{1}{4}$. We conclude that

$$\mathbb{E}\langle \eta_x - x, \theta\rangle^2 = \sum_{i=1}^{n} \theta_i^2 x_i(1 - x_i) \leq \frac{|\theta|^2}{4} = \frac{1}{4}. \tag{4}$$

Therefore, there exists some realization of η_x, a vertex $y \in \{0,1\}^n$ such that $\langle y - x, \theta\rangle^2 \leq 1/4$. Hence, $|\langle y - x, \theta\rangle| \leq 1/2$. □

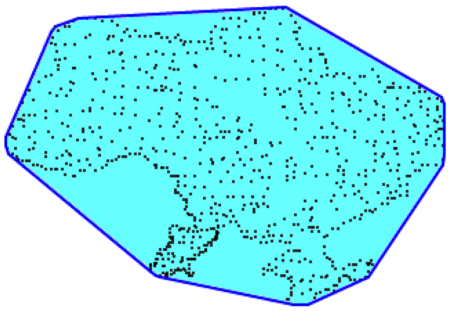
2.4 The Randomized Carathéodory theorem

We will now state a second cool fact, which is an approximate version of Carathéodory's theorem. We follow Vershynin [53] in this subsection.

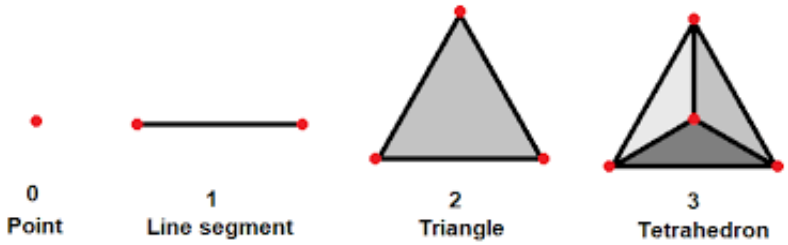
Definition 2.4 (Convex hull)**.** Let $A \subseteq \mathbb{R}^n$. The convex hull of A is

$$\mathrm{conv}(A) = \left\{ \sum_{i=1}^{m} \lambda_i x_i : \; m \in \mathbb{N},\, x_1, \ldots, x_m \in A,\, \lambda_i \geq 0,\, \sum_{i=1}^{m} \lambda_i = 1 \right\}.$$

The expression $\sum_{i=1}^{m} \lambda_i x_i$ is called a *convex combination* of $x_1, ..., x_m$.



For example, the convex hull of $n+1$ linearly independent points in $\mathbb{R}^n$ is called a *simplex*.



The next classical result tells us that every point in a convex hull of some set in $\mathbb{R}^n$ can be represented as a convex combination of at most $n+1$ points from this set, i.e. m in the definition of convex hull need not be larger than $n+1$.

Theorem 2.5 (Carathéodory's theorem)**.** *Let $A \subseteq \mathbb{R}^n$ and $x \in \text{conv}(A)$. Then there exist $x_1, \ldots, x_{n+1} \in A$ and $\lambda_1, \ldots, \lambda_{n+1} \geq 0$, $\sum_{i=1}^{n+1} \lambda_i = 1$ such that $x = \sum_{i=1}^{n+1} \lambda_i x_i$. In other words, x belongs to a simplex with vertices spanned by $x_1, \ldots, x_{n+1} \in A$.*

Proof. Home work! □

Carathéodory's Theorem could be useful for constructing various algorithms. However, having to operate with $n+1$ points could still be too difficult if $n = 100000000$, say. Would it be possible to represent x with less than $n+1$ points from A? In general, of course, not. However, if we were willing to represent x *approximately*, with some small error, then we could get away with using a lot less points, potentially:

Fact 2.6 (Randomized/Approximate Carathéodory)**.** *Suppose $A \subseteq \mathbb{R}^n$ with diameter*

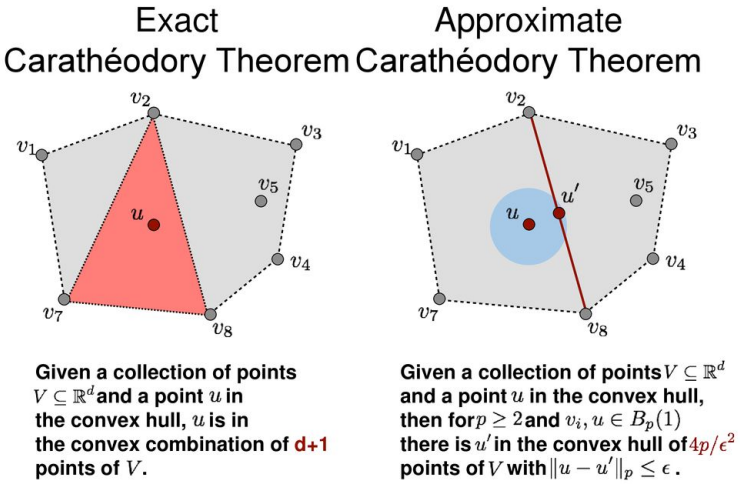
$$\text{diam}(A) = \sup_{x,y \in A} |x - y| \leq 1.$$

Then for all $x \in \text{conv}(A), k \in \mathbb{N}$, there exist $x_1, \ldots, x_k \in A$, such that

$$\left\| x - \frac{1}{k} \sum_{i=1}^{k} x_i \right\| \leq \frac{1}{\sqrt{k}}.$$

Note that we allow repetitions among the points.

While $\frac{1}{k} \sum_{i=1}^{k} x_i$ does not equal to x, but rather approximates it with error $\frac{1}{\sqrt{k}}$, the advantage is that we only need k points, and potentially $k \ll n+1$. Another interesting feature here is that we can take $\lambda_1 = \lambda_2 = \ldots = \lambda_k = \frac{1}{k}$, rather than deal with different weights.



Proof of Fact 2.6. Choose any $x \in \mathrm{conv}(A)$. By Carathéodory's theorem, there exist $z_1, \ldots, z_{n+1} \in A$ and $\lambda_1, \ldots, \lambda_{n+1} \geq 0$ such that $\sum_{i=1}^{n+1} \lambda_i = 1$ and

$$x = \sum_{1 \leq i \leq n+1} \lambda_i z_i.$$

Consider a random vector Z which takes value z_1 with probability λ_1, value z_2 with probability λ_2, and so on (overall, this random vector takes $n+1$ values $z_1, \ldots, z_n \in A$). Note that

$$\mathbb{E}\, Z = \sum_{1 \leq i \leq n+1} \lambda_i z_i = x.$$

Now, given an integer k, consider k i.i.d. copies of Z denoted by $Z_1, \ldots, Z_k$ (the existence of these random vectors is guaranteed by Kolmogorov's extension theorem, see Theorem 1 in section 2.9 of [57]). Note that by Lemma 2.3,

$$\mathbb{E} \left| x - \frac{1}{k} \sum_{1 \leq i \leq k} Z_i \right|^2 = \frac{1}{k^2} \sum_{1 \leq i \leq k} \mathbb{E} \left| x - Z_i \right|^2 \leq \frac{1}{k^2} \cdot k = \frac{1}{k},$$

where in the last passage we used the fact that $x \in \mathrm{conv}(A)$, and $\frac{1}{k} \sum_{1 \leq i \leq k} Z_i \in A \subset \mathrm{conv}(A)$ with probability 1, and the diameter of the convex hull of A is bounded from above by the diameter of A (homework!), which in turn is bounded by 1.

We conclude that

$$\mathbb{E} \left| x - \frac{1}{k} \sum_{1 \leq i \leq k} Z_i \right|^2 \leq \frac{1}{k}.$$

Hence, there exist $y_1, \ldots, y_k \in A$, the realizations of $Z_1, ..., Z_n$, such that

$$\left| x - \frac{1}{k} \sum_{1 \le j \le k} y_i \right| \le \frac{1}{\sqrt{k}},$$

which proves our Fact. □

The Approximate Carathéodory Theorem has, for example, the following useful consequence:

Corollary 2.7. *Let P be a polytope in $\mathbb{R}^n$ with at most N vertices, i.e.,*

$$P = \text{conv}\{x_1, \ldots, x_N\},$$

where each $x_i \in \mathbb{R}^n$. Suppose that $\text{diam}(P) \le 1$. Fix any $\varepsilon > 0$. Then, P can be covered by at most $N^{\lfloor 1/\varepsilon^2 \rfloor}$ euclidean balls of radius ε.

Proof. Home work! □

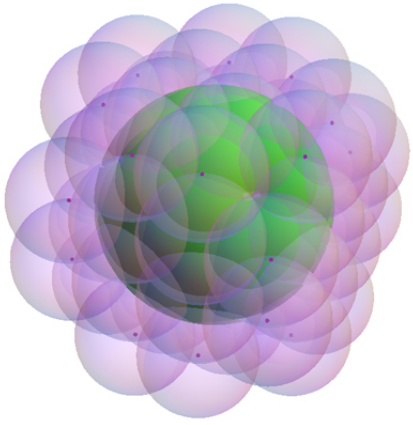
While we leave this fact as a home work, in the next sub-section we prove a related and very useful result.

2.5 The standard ε-net argument

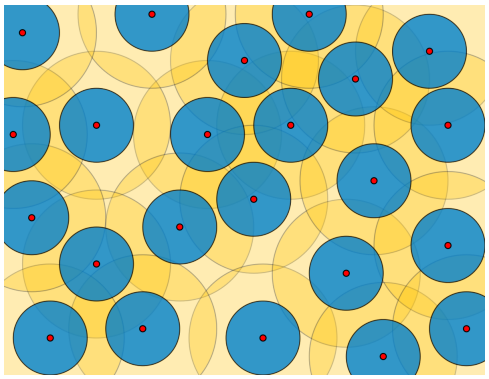
Suppose we would like to view the sphere or the unit ball as a discrete set, in which each point is represented approximately. What would be a good way to do that? We start by presenting the most classical result.

Theorem 2.8 (Classical ε-net Construction)**.** *Let $n \in \mathbb{N}$ and $\varepsilon > 0$. Then, there exists a positive integer $m \le \left(\frac{2+\varepsilon}{\varepsilon}\right)^n$, and $y_1, \ldots, y_m \in \mathbf{B}_2^n$, such that the unit ball $\mathbf{B}_2^n$ is covered by balls of radius ε with the centers at y_i, that is,*

$$\mathbf{B}_2^n \subseteq \bigcup_{1 \le i \le m} (y_i + \varepsilon \mathbf{B}_2^n).$$



Proof. Without loss of generality, we may assume that $\varepsilon < 1$ (otherwise, the statement is straightforward). Consider a **maximal** packing of balls of radius $\varepsilon/2$ in $(1+\varepsilon/2)\mathbf{B}_2^n$, i.e., pick $y_1, \ldots, y_m$ such that $y_i + (\varepsilon/2)\mathbf{B}_2^n$ are disjoint, $y_i + (\varepsilon/2)\mathbf{B}_2^n \subseteq (1+\varepsilon/2)\mathbf{B}_2^n$, and m is the largest possible number of balls with such properties (see Remark below for a justification for the existence of such a packing).



First, we show that

$$B_2^n \subseteq \bigcup_{1 \le i \le m} (y_i + \varepsilon B_2^n). \tag{5}$$

Suppose not. Then, there exists $x \in B_2^n$ such that $|x - y_i| > \varepsilon/2$ for all $1 \le i \le m$, and $x + \frac{\varepsilon}{2}\mathbf{B}_2^n \subset (1+\frac{\varepsilon}{2})\mathbf{B}_2^n$. But then $x + (\varepsilon/2)\mathbf{B}_2^n \cap y_i + (\varepsilon/2)\mathbf{B}_2^n = \emptyset$, contradicting the maximality of m. Hence (5) is confirmed. That is, we found a construction of some covering of B_2^n with balls of radius ε; our next task is to estimate its size.

Observe that

$$(1+\varepsilon/2)\mathbf{B}_2^n \supseteq \bigsqcup_{1 \le i \le m} (y_i + \varepsilon/2\mathbf{B}_2^n)$$

Therefore, by the additive property of the Lebesgue measure, we have

$$\sum_{1 \le i \le m} \left| y_i + \frac{\varepsilon}{2}\mathbf{B}_2^n \right| \le \left| (1+\varepsilon/2)\mathbf{B}_2^n \right|,$$

where $|\cdot|$ denotes the n-dimensional Lebesgue measure. Therefore, we have

$$m \cdot \left(\frac{\varepsilon}{2}\right)^n \le \left(1 + \frac{\varepsilon}{2}\right)^n$$

Hence, $m \le (1 + 2/\varepsilon)^n$, as desired. □

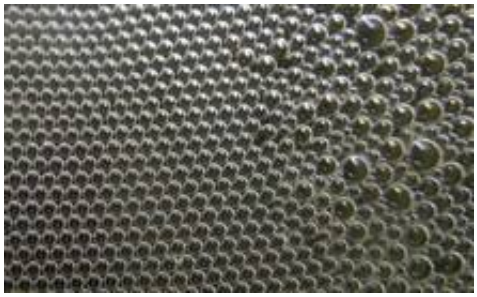
Remark 2.9 (Ideas by Ruijia Cao, not something Galyna suggested)**.** *To show that a maximal packing used in the proof above exists, we can use the following argument. Note that for $t > 0$ and any packing of m disjoint balls, we have*

$$\bigsqcup_{i=1}^{m} (y_i + t\mathbf{B}_2^n) \subseteq \mathbf{B}_2^n,$$

Then, $m \leq 1/t^n$. Let $\varepsilon > 0$. We will use the following algorithm to explicitly construct a maximal packing of $\mathbf{B}_2^n$ using balls with radii $\varepsilon/2$:

1. *First, pick any $x_1 \in \mathbf{B}_2^n$.*
2. *In the second step, choose $x_2 \in \mathbf{B}_2^n$ such that $\|x_2 - x_1\| \geq \varepsilon$.*
3. *At the k-th step, suppose that $x_1, \ldots, x_{k-1}$ have been chosen such that $x_i + (\varepsilon/2)\mathbf{B}_2^n$ are pairwise disjoint. Choose x_k such that $\|x_k - x_i\| \geq \varepsilon$ for all $1 \leq i \leq k-1$.*
4. *If no such x_k exists, then the algorithm terminates.*

Since the number of balls in any packing of disjoint ball is finite, the above algorithm will terminate in finitely many steps.



Remark 2.10 (a small addition by Galyna)**.** *Arguing along the lines of the previous Remark (by Ruijia), one can consider the collection of all such "locally optimal" packings. Since the sizes of all of them are bounded by t^{-n} (as was explained using the volumetric argument), and since the supremum of a bounded from above sequence is attained, we conclude that at least one of such packings is maximal. For details, see e.g. Vershynin [53].*

Remark 2.11. *Note that the size m of the covering of the unit ball by balls of radius ε is necessarily at least $\frac{1}{\varepsilon^n}$, since $m \cdot |\varepsilon \mathbf{B}_2^n| \geq |\mathbf{B}_2^n|$. Therefore, the bound in Theorem 2.8 is sharp up to 3^n, when $\varepsilon < 1$.*

Therefore, one may conclude that the optimal ε-covering is of the size $\left(\frac{c}{\varepsilon}\right)^n$ where $c \in [1, 3]$ (technically, the sharp c might depend on ε but we could discuss the "limiting value" in some sense). There is vast literature on the subject and tighter bounds are known, however the only dimensions in which the limiting value of c is known (when $\varepsilon \to 0$) are $2, 8$ and 24. Dimension 2 is assigned as a (difficult!) home work, while in dimensions 8 and 24 this result was established by a Fields medalist Viazovska [55], and Cohn, Kumar, Miller, Radchenko, Viazovska [56].

Remark 2.12. *Since $\mathbb{S}^{n-1} \subset \mathbf{B}_2^n$, Theorem 2.8 implies that there exists a covering of size $\left(\frac{3}{\varepsilon}\right)^n$ of the unit sphere by $\varepsilon-$balls. In fact, a stronger result is true: there is such a cover of size at most $\left(\frac{c}{\varepsilon}\right)^{n-1}$. This important fact is left as a home work.*

Theorem 2.8 is very classical and it has many powerful applications, some of which we will see in this course. However, there are situations when this result is, in fact, insufficiently sharp. Another disadvantage of the argument we presented is that it is not constructive (as one of the students cleverly pointed out), because an optimal packing is not constructed explicitly.

2.6 A lattice net

In this subsection, we discuss a refinement of Theorem 2.8 which is furthermore constructive (although of course one cannot only gain and not lose, so some features of Theorem 2.8 might not be kept). First, we note the following classical

Lemma 2.13. *Let $n, N \in \mathbb{Z}$ be positive integers. Then, there are $\binom{N+n-1}{n-1}$ number of solutions to the following equation*

$$x_1 + \cdots + x_n = N,$$

where $x_i \in \mathbb{Z}$ and $x_i \geq 0$.

Proof. Note that the number of ways to represent N as an ordered sum of n non-negative integers is the same as the number of ways to distribute N identical balls into n different boxes (any box can contain between 0 and n balls). The answer to this classical combinatiorial riddle is $\binom{N+n-1}{n-1}$. Indeed, this corresponds to the number of ways to arrange a sequence of N zeroes (corresponding to the balls) and $n-1$ ones (corresponding to the sides of the boxes, except the two on the edges). And this is the same as selecting $n-1$ spots to place ones in the collection of $N+n-1$ spots. See Figure 1. □

Theorem 2.14. *Let $R > 0$. Then, there exists a constant $C > 0$ such that*

$$\#\{x \in \mathbb{Z}^n : \|x\|_1 \leq R\} \leq \left(\frac{CR}{n}\right)^n.$$

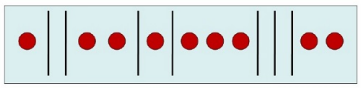


Figure 1: $(n-1)$ dividers and N balls

Proof. Applying Lemma 2.13, we have

$$\begin{aligned}\#\left\{x \in \mathbb{Z}^n : \|x\|_1 \leq R\right\} &\leq 2^n \cdot \#\left\{x \in \mathbb{Z}^n : x_i \geq 0, \sum_{1\leq i\leq n} x_i \leq R\right\} \\ &= 2^n \cdot \sum_{N=1}^{\lfloor R\rfloor} \binom{N+n-1}{n-1}.\end{aligned}$$

Now, Stirling's formula states that for any $k \geq 1$

$$k! = (1+o(1)) \cdot \sqrt{2\pi k}\left(\frac{k}{e}\right)^k.$$

Therefore, with a bit of arithmetic, one can show (*Home work!*) that there exists a constant $C > 0$ such that

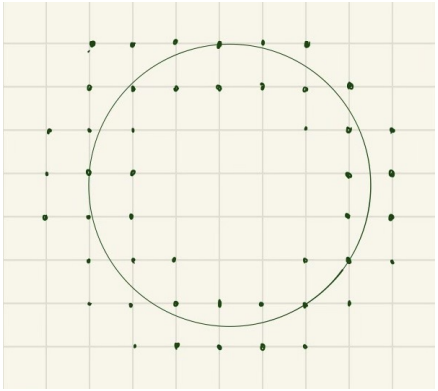
$$2^n \cdot \sum_{N=1}^{\lfloor R\rfloor} \binom{N+n-1}{n-1} \leq \left(\frac{CR}{n}\right)^n.$$

□

Corollary 2.15. *Fix any $\varepsilon > 0$. Then, there exist $y_1, \ldots, y_m \in \mathbf{B}_2^n$ and a constant $c > 0$ such that*

$$\bigcup_{1\leq i\leq m}\left(y_i + \frac{\varepsilon}{\sqrt{n}}\mathbf{B}_\infty^n\right) \supseteq \mathbf{B}_2^n,$$

where $m \leq (C/\varepsilon)^n$.



Proof. Let $R = n/\varepsilon$. Then, by Theorem 2.14, there exist $c > 0$ and $y_1, \ldots, y_m \in \mathbb{Z}^n$ with $\|y_i\|_1 \le R$, where $m \le (C/\varepsilon)^n$. Therefore,

$$\frac{R}{\sqrt{n}} \cdot \mathbf{B}_2^n \subseteq \bigcup_{1 \le i \le m} (y_i + \mathbf{B}_\infty^n).$$

Now, by excluding those y_i that lie outside a ball of radius $\frac{3}{2}\frac{R}{\sqrt{n}}$ and by re-indexing if necessary, we may choose $y_1, \ldots, y_{m'} \in \frac{3}{2}\frac{R}{\sqrt{n}}\mathbf{B}_2^n \cap \{y_1, \ldots, y_m\}$ such that

$$\frac{R}{\sqrt{n}} \cdot \mathbf{B}_2^n \subseteq \bigcup_{1 \le i \le m'} (y_i + \mathbf{B}_\infty^n),$$

where $m' \le m$. Therefore,

$$\mathbf{B}_2^n \subseteq \bigcup_{1 \le i \le m} \left(\frac{\varepsilon}{\sqrt{n}} y_i + \frac{\varepsilon}{\sqrt{n}} \mathbf{B}_\infty^n \right).$$

Note that each $\frac{\varepsilon}{\sqrt{n}} y_i \in \frac{3}{2}\mathbf{B}_2^n$ for $1 \le i \le m'$ because $\|y_i\|_1 \le R$, and because

$$\sqrt{n}|x| \ge \|x\|_1 \ge |x|,$$

for any $x \in \mathbb{R}^n$. □

Remark 2.16. *Here, we give a comparison between Theorem 2.8 and Corollary 2.15. By Theorem 2.14 and Corollary 2.15,*

$$\mathbf{B}_2^n \subseteq \bigcup_{1 \le i \le m} \left(y_i + \frac{\varepsilon}{\sqrt{n}} \mathbf{B}_\infty^n \right), \tag{6}$$

where $m \le (\frac{c}{\varepsilon})^n$. On the other hand, Theorem 2.8 implies that

$$\mathbf{B}_2^n \subseteq \bigcup_{1 \le i \le m} (y_i + \varepsilon \mathbf{B}_2^n), \tag{7}$$

where $m \le \left(\frac{3}{\varepsilon}\right)^n$.

Note that in some sense, Corollary 2.15 is stronger than Theorem 2.8 since $\mathbf{B}_\infty^n \subseteq \sqrt{n}\mathbf{B}_2^n$ (see Figure 2), and therefore (6) implies (7). However, Corollary 2.15 does have a minor disadvantage comparing to Theorem 2.8: the constant $C > 3$ is unspecified in our computation. But this usually doesn't play a role in applications.

Since $\mathbb{S}^{n-1} \subset \mathbf{B}_2^n$, and by removing some unused points of small Euclidean norm, we get the following result about sphere covering from Corollary 2.15:

Corollary 2.17 (Home Work)**.** *For any $\varepsilon > 0$, there there exist $y_1, \ldots, y_m \in B_2^n$ such that for any $x \in \mathbb{S}^{n-1}$ there exists $i \in [m]$ with $\|x - y_i\|_\infty < \varepsilon/\sqrt{n}$ and $m \le (C/\varepsilon)^{n-1}$.*

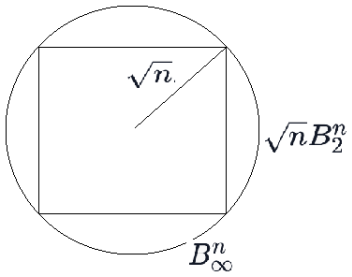


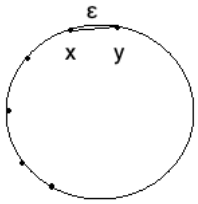
Figure 2: L_2 and L_∞ norms comparison

Remark 2.18. *Note that not all $y_i \in \mathbb{S}^{n-1}$, as in our previous construction of the standard ϵ-net argument.*

We can also restate Theorem 2.14 and Corollary 2.17 in the following manner:

Theorem 2.19. *1. For all $\varepsilon > 0$, there exists a collection of points $y_1, ..., y_m \in \mathbb{S}^{n-1}$ with $m \leq (\frac{C}{\varepsilon})^n$, such that for all $x \in \mathbb{S}^{n-1}$, there exists $i \in \{1, ..., m\}$ such that $|x - y_i| < \varepsilon$.*

2. For all $\varepsilon > 0$, there exists $y_1, ..., y_m \subset \frac{3}{2}\mathbf{B}_2^n \backslash \frac{1}{2}\mathbf{B}_2^n$ with $m \leq (\frac{C'}{\varepsilon})^n$, such that for all $x \in \mathbb{S}^{n-1}$, there exists $i \in \{1, ..., m\}$ such that $\|x - y_i\|_\infty \leq \frac{\varepsilon}{\sqrt{n}}$.



Proof. The first part follows from By Theorem 2.8 and the standard net argument, while the second part follows from Corollary 2.17. □

Remark 2.20. *Note that $\frac{1}{\sqrt{n}}|a| \leq \|a\|_\infty \leq |a|$, where $|\cdot| = \|\cdot\|_2$, and therefore the second part of Theorem 2.19 is stronger than the first part, in the essential sense.*

Remark 2.21. *However note that in the second part of Theorem 2.19, unlike in the first part, one needs to assume that the points of the net are in $\frac{3}{2}\mathbf{B}_2^n \backslash \frac{1}{2}\mathbf{B}_2^n$, rather than simply in B_2^n. This is because the collection of our approximating points in the second part consists of the vertices of the lattice cubes, rather than centers! This will not make a difference for our applications, however.*

2.7 An application: an efficient net for matrix multiplication

From now on, we will use the word *net* to mean a collection of points which has some good approximating properties for a given set (usually, the sphere). We will use notation $\mathcal{N}$ for a finite set of points, in place of just listing the points $\{y_1, \ldots, y_m\}$ as before, to save time. First, we outline the following

Theorem 2.22. *For all $\varepsilon > 0$, there exists a set $\mathcal{N} \subseteq \mathbb{S}^{n-1}$ such that $\#\mathcal{N} \leq (\frac{c}{\varepsilon})^n$, and for all $x \in \mathbb{S}^{n-1}$, there exists $y \in \mathcal{N}$ such that for any integer $N \geq 0$, and for any $N \times n$ matrix $\mathbf{A}$,*

$$\|\mathbf{A}(x-y)\| \leq \|\mathbf{A}\|_{op} \cdot \varepsilon,$$

where we recall that the operator norm of a matrix was defined in (1).

Proof. By the usual net argument, for all $x \in \mathbb{S}^{n-1}$, there exists some $y \in \mathcal{N}$ such that $|x-y| < \varepsilon$. Take any matrix $\mathbf{A}$, we have $|\mathbf{A}(x-y)| \leq \|\mathbf{A}\|_{op} \cdot |x-y| \leq \|\mathbf{A}\|_{op} \cdot \varepsilon$. □

Next, we now explain a more precise way to approximate points on the sphere in order to compare the functional $|\mathbf{A}x|$ at those points. This net construction combines the ideas we already explained in Theorem 2.19 (part 2), with the idea of random rounding which we discussed in subsection 2.2. This net construction first appears in Klartag, Livshyts [18], and was later further developed in Livshyts [25].

Theorem 2.23. *Fix $n \in \mathbb{N}$. For all $\varepsilon > 0$, there exists $\mathcal{N} \subset \frac{3}{2}\mathbf{B}_2^n \backslash \frac{1}{2}\mathbf{B}_2^n$ such that $\#\mathcal{N} \leq (\frac{C}{\varepsilon})^n$, and so that for all $x \in \mathbb{S}^{n-1}$ and for all $N \in \mathbb{N}$, for any $N \times n$ matrix $\mathbf{A}$, there exists $y \in \mathcal{N}$ such that*

$$|\mathbf{A}(x-y)| \leq \varepsilon \cdot \frac{\|\mathbf{A}\|_{HS}}{\sqrt{n}},$$

where the Hilbert-Schmidt norm of a matrix was defined in Definition 1.2.

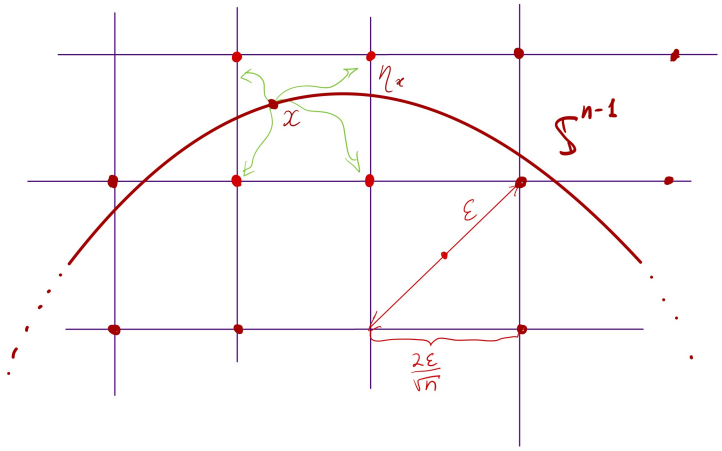
Remark 2.24. *As we noted before, $\frac{\|\mathbf{A}\|_{HS}}{\sqrt{n}} \cdot \varepsilon \leq \|\mathbf{A}\|_{op} \cdot \varepsilon$, and thus the net in Theorem 2.23 is more precise than the net in Theorem 2.22. Apart from some minor differences in the order of quantifiers (which we discuss below), Theorem 2.23 is stronger than Theorem 2.22, in the essential sense (but not literally).*

Proof of 2.23. Consider $\mathcal{N}$ to be the net given in Corollary 2.17. Then indeed $\mathcal{N} \subset \frac{3}{2}\mathbf{B}_2^n \backslash \frac{1}{2}\mathbf{B}_2^n$ such that $\#\mathcal{N} \leq (\frac{C}{\varepsilon})^n$, and for every $x \in \mathbb{S}^{n-1}$ one may find a lattice cube

$$Q = \frac{\epsilon}{\sqrt{n}} \prod_{i=1}^{n} [t_i, t_i + 1],$$

for some integers t_i, such that $x \in Q$, and the vertices of Q all belong to the point set $\mathcal{N}$.

Consider the (scaled) random rounding η_x of $x \in \mathbb{S}^{n-1}$ to the vertices of the cube Q (as defined earlier).



Namely, η_x is a random vector which takes values in the vertices of Q, has independent coordinates, and $\mathbb{E}\,\eta_x = x$. Recall from (4), after the appropriate re-scaling by $\frac{2\epsilon}{\sqrt{n}}$, that

$$\mathbb{E}\langle x - \eta_x, \theta\rangle^2 \le \frac{\epsilon^2|\theta|^2}{n}. \tag{8}$$

Recall that $|\mathbf{A}y|^2 = \sum_{i=1}^N \langle \mathbf{A}^T e_i, y\rangle^2$ (here $A^T e_i$ are the rows of A). Using the inequality (8) with $\theta = A^T e_i$, we get

$$\begin{aligned}
\mathbb{E}\,|\mathbf{A}(x-\eta_x)|^2 &= \mathbb{E}\sum_{i=1}^N \langle \mathbf{A}^T e_i, x-\eta_x\rangle^2 \\
&= \sum_{i=1}^N \mathbb{E}\langle \mathbf{A}^T e_i, x-\eta_x\rangle^2 \\
&\le \frac{\varepsilon^2}{n}\sum_{i=1}^N |\mathbf{A}^T e_i|^2 \\
&= \|\mathbf{A}\|_{HS}^2 \cdot \frac{\varepsilon^2}{n}.
\end{aligned}$$

Since $\mathbb{E}\,|\mathbf{A}(x-\eta_x)|^2 \le \|\mathbf{A}\|_{HS}^2 \cdot \frac{\varepsilon^2}{n}$, there exists a realization of random vector y of η_x such that

$$|\mathbf{A}(x-y)|^2 \le \varepsilon^2 \cdot \frac{\|\mathbf{A}\|_{HS}^2}{n},$$

or equivalently,

$$|\mathbf{A}(x-y)| \le \varepsilon \cdot \frac{\|\mathbf{A}\|_{HS}}{\sqrt{n}}.$$

Note that any realization of η_x is a vertex of Q and in particular, $y \in \mathcal{N}$. □

Remark 2.25. *Note that:*

- *For both Theorem 2.22 and Theorem 2.23, the net does not depend on the matrix* $\mathbf{A}$.
- *For Theorem 2.22, the approximation point* y *depends on both* $\mathbf{A}$ *and* x*, which is not the case for Theorem 2.23. But we will see that in applications that we consider, this does not make a difference.*

Example 2.26. *Let us consider two concrete examples of matrices.*

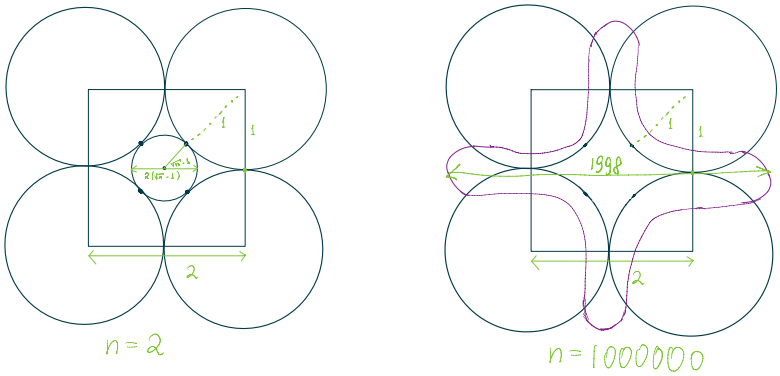
- $\mathbf{A} = \mathrm{Id_n}$. *Then* $\|\mathbf{A}\|_{op} = 1$, *and* $\|\mathbf{A}\|_{HS} = \sqrt{1 + \cdots + 1} = \sqrt{n}$. *Therefore,* $\frac{\|A\|_{HS}}{\sqrt{n}} = \|A\|_{op}$, *and therefore Theorems 2.22 and 2.23 provide the same precision.*
- $\mathbf{A} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$

  Then $\frac{1}{\sqrt{n}}\|\mathbf{A}\|_{HS} = \frac{1}{\sqrt{n}} \ll 1 = \|\mathbf{A}\|_{op}$, *and therefore,* ***Theorem 2.23 is*** $\sqrt{n}$ ***more precise when approximating the functional*** $Ax = x_1$ ***on the sphere, or, equivalently, any one-dimensional projection*** $\langle x, \xi \rangle$***!*** *For instance, this was crucial in [18], and the necessity to have a net with such a property motivated this construction.*

2.8 High-dimensional phenomenon: some themes

Weirdness of high dimension

Consider the cube B_∞^n and place a copy of B_2^n centered at each vertex of B_∞^n (see the Figure below). Note that the largest ball you can place at the center of B_∞ without intersecting any of the copies of B_2^n will have radius $\sqrt{n} - 1$. This might be somewhat surprising, since in many directions the ball extends out much further from the origin than the cube does, when n is large.



Central Limit Theorem and its geometric meaning

The following result is a very classical fact in Probability, see e.g. Durrett [7].

Theorem 2.27 (Central Limit Theorem)**.** *Let $X_1, X_2, \dots$ be i.i.d. random variables with finite second moment. Let $\mu = \mathbb{E}\, X_1$ and $\sigma^2 = \mathrm{Var}(X_1)$. Let $S_n = \sum_{1 \le k \le n} X_k$. Then,*

$$\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \to \mathcal{N}(0,1),$$

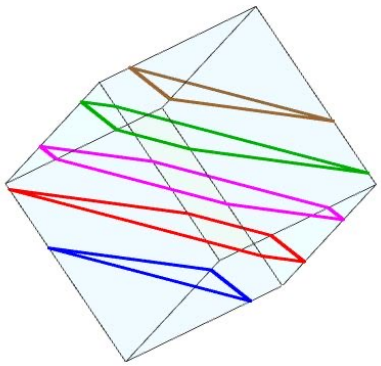
where the convergence is in distribution as $n \to \infty$.

To get a geometric interpretation of this, note that the vector $X := (X_1, \cdots, X_n)$ is distributed uniformly over the unit cube $\mathbf{B}_n^\infty := [-1,1]^n$. Given the vector $\theta = (1/\sqrt{n}, \cdots, 1/\sqrt{n})$ the above example says that the random variable $\langle X, \theta \rangle$ is distributed roughly as a normal random variable.

What is the geometric meaning of the density $f(t)$ of $\langle X, \theta \rangle$? After thinking a little, we see that

$$f(t) = |\mathbf{B}_n^\infty \cap \{\langle x, \theta \rangle = t\}|_{n-1}.$$

Thus $f(t)$ the $n-1$ dimensional area of the hyperplane section of the cube perpendicular to θ, distance t from the origin. Although sections of cubes are hard to compute exactly, as the dimension goes to infinity they resemble a normal random variable, which is a simple object.



This phenomenon appears to stem from *independence*. However, this fact is true not just about cubes! For any convex body there exists a direction θ (in fact, many of them) for which $\langle X, \theta \rangle$ behaves a bit like a Gaussian (in the appropriate sense). This is the content of the Central Limit Theorem for convex bodies from 2007 due to Klartag [17].

About the thin shell type concentration

Suppose that X is a random vector that is distributed uniformly over the sphere. Then with high probability, its Euclidean norm is very close to 1 when n is large!

Indeed, note that

$$\begin{aligned}\mathbb{P}\left(|X| \geq 1 - \frac{10}{n}\right) = \frac{|\mathbf{B}_2^n \setminus (1-10/n)\mathbf{B}_2^n|}{|\mathbf{B}_2^n|} &= 1 - \left(1 - \frac{10}{n}\right)^n \\ &\geq 1 - e^{-10} \\ &\geq 0.99.\end{aligned}$$

3 Concentration Inequalities for sums of independent random variables

3.1 What is a concentration inequality?

In this section, we closely follow Vershynin [53]; see also the references therein. Recall the Law of Large Numbers (see Theorem 2.4.1 in Durrett [7]).

Theorem 3.1 (LLN)**.** *Let $(X_n)_{n\geq 1}$ be iid random variables with $\mathbb{E}\,|X_1| < \infty$. Then,*

$$\frac{X_1 + \ldots + X_N}{N} \xrightarrow{a.s.} \mathbb{E}\, X_1$$

as $N \to \infty$.

Definition 3.2. Let X be a random variable. We say that X satisfies a **concentration inequality** if

$$\mathbb{P}\left(|X - \mathbb{E}\, X| > t\right) \leq \text{☺}(t).$$

We usually think of ☺(t) as a function that decays to 0 fast.

In some sense, a concentration inequality is a quantitative version of the Law of Large Numbers. A classical example is the Chebychev inequality:

Theorem 3.3 (Chebychev)**.** *Let X be a random variable with $\mathbb{E}\,|X|^2 < \infty$ and $t \geq 0$. Then,*

$$\mathbb{P}\left(|X - \mathbb{E}\, X| \geq t\right) \leq \frac{\mathrm{Var}(X)}{t^2}.$$

Proof. First, note that

$$\mathbb{E}\,|X - \mathbb{E}\, X|^2 = \mathbb{E}\, X^2 - (\mathbb{E}\, X)^2 \leq \mathbb{E}\, X^2 < \infty.$$

Applying Markov's inequality to the random variable $|X - \mathbb{E}\, X|$, we see that

$$\begin{aligned}\mathbb{P}\left(|X - \mathbb{E}\, X| \geq t\right) &= \mathbb{P}\left(|X - \mathbb{E}\, X|^2 \geq t^2\right) \\ &= \frac{1}{t^2}\,\mathrm{Var}(X).\end{aligned}$$

□

3.2 Hoeffding's inequality

However, Chebyshev's inequality provides a weak bound when it comes to random variables which are sums of independent random variables with certain properties. Below, we present the first example of a much stronger result.

Theorem 3.4 (Hoeffding's Inequality)**.** *Let $X_1, \ldots, X_n$ be independent symmetric Bernoulli random variables (taking values 1 and -1 with probability 0.5 each). Let $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$. Then, for all $t \geq 0$, we have*

$$\mathbb{P}\left(\sum_{1\leq i\leq n} a_i X_i \geq t\right) \leq e^{-\frac{t^2}{2|a|^2}}$$

Proof of Theorem 3.4. Without loss of generality, we may assume that $a \in \mathbb{S}^{n-1}$ (otherwise, replace a by $a/|a|$ in the proof below).

Then, using the so-called Chernoff's trick, we have

$$\begin{aligned}\mathbb{P}\left(\sum_i a_i X_i \geq t\right) &= \mathbb{P}\left(e^{\lambda \sum a_i X_i} \geq e^{\lambda t}\right) \\ &\leq e^{-\lambda t} \prod_{1\leq i\leq n} \mathbb{E}\left(e^{\lambda a_i X_i}\right)\end{aligned}$$

Recall that $X_1, \ldots, X_n$ are symmetric Bernoulli random variables, i.e.,

$$X_i = \begin{cases} 1, & \text{w.p. } 1/2 \\ -1, & \text{w.p. } 1/2 \end{cases}$$

Hence, we have

$$\begin{aligned}\mathbb{E}\left(e^{\lambda a_i X_i}\right) &= \frac{1}{2}e^{\lambda a_i} + \frac{1}{2}e^{-\lambda a_i} \\ &= \cosh(\lambda a_i)\end{aligned}$$

Hence, for any $\lambda > 0$,

$$\begin{aligned}\mathbb{P}\left(\sum_{1\leq i\leq n} a_i X_i \geq t\right) &\leq e^{-\lambda t} \prod_{1\leq i\leq n} \cosh(\lambda a_i) && (9) \\ &\leq e^{-\lambda t} \prod_{1\leq i\leq n} e^{\lambda^2 a_i^2/2} && (\text{prove } \cosh(x) \leq e^{x^2/2} \text{ in HW}) \\ &= e^{-\lambda t} e^{\lambda^2/2 \sum_{1\leq i\leq n} a_i^2} && (10) \\ &= e^{-\lambda t + \lambda^2/2} && (\ |a| = 1, \forall \lambda \geq 0)\end{aligned}$$

By picking the optimal $\lambda = t$, the upper bound becomes $e^{-t^2/2}$. □

Remark 3.5. *There is a nice geometric interpretation for this inequality, Let $X \in \{-1,1\}^n$, $t > 0$ and $a \in \mathbb{S}^{n-1}$.*

$$\mathbb{P}\left(\langle X, a\rangle \geq t\right) \leq e^{-t^2/2},$$

and therefore, the relative number of the points of the hypercube located in a half-space orthogonal to a and distance t from the origin behaves as (or, rather, better than) the Gaussian function.

Theorem 3.6 (Hoeffding's inequality for Bounded R.V.)**.** *Let $X_1, \ldots, X_n$ be bounded independent random variables, i.e., $X_i \in [m_i, M_i]$ for some $m_i, M_i \in \mathbb{R}$. Then, for all $\beta > 0$, we have*

$$\mathbb{P}\left(\left|\sum_i X_i - \mathbb{E}\, X_i\right| \geq \beta\right) \leq 2e^{-\frac{c\beta^2}{\sum_i (M_i - m_i)^2}}$$

Proof of Theorem 3.6. Homework.

Hint:

$$\mathbb{P}\left(\left|\sum_i X_i - \mathbb{E}\, X_i\right| \geq \beta\right) = \mathbb{P}\left(\sum_{1\leq i\leq n} X_i - \mathbb{E}\, X_i \geq \beta\right) + \mathbb{P}\left(\sum_{1\leq i\leq n} X_i - \mathbb{E}\, X_i \leq -\beta\right)$$

and estimate both parts. □

Remark 3.7. *(Hoeffding vs. Chebyshev) Theorem 3.6 tells us that*

$$\mathbb{P}\left(\left|\sum_{1\leq i\leq n} X_i - \mathbb{E}\, X_i\right| \geq t\right) \leq 2e^{-ct^2}$$

On the other hand, Chebyshev's inequality yields that

$$\mathbb{P}\left(\left|\sum_i X_i - \mathbb{E}\, X_i\right| \geq t\right) \leq \frac{\mathrm{Var}(\sum_i X_i)}{t^2} = \frac{\sum_i \mathrm{Var}(X_i)}{t^2}$$

Since $\frac{c_1}{t} \geq e^{-ct^2 + \log 2}$ for large enough t and some positive constants c_1 and c_2, we see that Hoeffding's inequality gives a tighter bound than Chebyshev's inequality.

Remark 3.8. *(Hoeffding vs. Central Limit Theorem) According to CLT, we know $\frac{X_1+\cdots+X_n}{\sqrt{n}} \xrightarrow{d} Z \sim \mathcal{N}(0,1)$. Let X_i be independent Bernoulli random variables, i.e.,*

$$X_i = \begin{cases} 1, & w.p.\ p \\ -1, & w.p.\ 1-p \end{cases}$$

For $t > 0$, observe that

$$\frac{1}{\sqrt{2\pi}}\left(\frac{1}{t} - \frac{1}{t^3}\right) e^{-\frac{t^2}{2}} \leq \mathbb{P}(Z > t) \leq \frac{1}{\sqrt{2\pi}} \frac{1}{t} e^{-\frac{t^2}{2}}. \tag{11}$$

The lower bound is for $t \geq 1$ is left as homework. For the upper bound, note that

$$\begin{aligned} P(Z > t) &= \int_t^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} \, \mathrm{d}s \\ &= \frac{1}{\sqrt{2\pi}} \int_t^\infty s \cdot \frac{1}{s} \cdot e^{-\frac{s^2}{2}} \, \mathrm{d}s \\ &\leq \frac{1}{t} \frac{1}{\sqrt{2\pi}} \int_t^\infty s \cdot e^{-\frac{s^2}{2}} \, \mathrm{d}s \\ &= \frac{1}{t} \frac{1}{\sqrt{2\pi}} \int_{t^2/2}^\infty e^{-s} \, \mathrm{d}s \\ &= \frac{1}{t} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}. \end{aligned}$$

But we cannot directly deduce Hoeffding from CLT: although

$$\left| \mathbb{P}\left(\frac{X_1 + \cdots + X_n}{\sqrt{n}} \geq t \right) - \mathbb{P}(Z \geq t) \right| \to_{n \to \infty} 0,$$

For fixed n, it can be polynomial in t. Hoeffding is not a stronger theorem than CLT, but it implies the behaviour predicted by the CLT when it comes to tail probabilities.

Remark 3.9. *(HW) In fact, Hoeffding is not tight for non-symmetric Bernoulli random variables*

$$X_i = \begin{cases} 1 & w.p.\ p_i \\ -1 & w.p.\ 1 - p_i, \end{cases}$$

where $p_i \approx 0.99$. Chernoff's Inequality below is better.

3.3 Chernoff's inequality

Theorem 3.10. *(Chernoff's Inequality) Let*

$$X_i = \begin{cases} 1, & w.p.\ p_i \\ 0, & w.p.\ 1 - p_i \end{cases}$$

be independent Bernoulli random variables with parameters p_i. Denote $S_N = \sum_{i=1}^N X_i$ and

$$\mathbb{E}\, S_N = \sum_{i=1}^N \mathbb{E}\, X_i = \sum_{i=1}^N p_i =: \mu.$$

Then for all $t > \mu$,

$$\mathbb{P}(S_N > t) \leq e^{-\mu} \left(\frac{e\mu}{t} \right)^t.$$

Proof. For any $\lambda \in \mathbb{R}$,

$$\begin{aligned}
\mathbb{P}\left(\sum X_i > t\right) = \mathbb{P}\left(e^{\lambda \sum X_i} > e^{\lambda t}\right) &\le e^{-\lambda t}\, \mathbb{E}\, e^{\lambda \sum X_i} \\
&= e^{-\lambda t} \prod_{i=1}^{N} \mathbb{E}\, e^{\lambda X_i} && \text{(by independence)} \\
&= e^{-\lambda t} \prod_{i=1}^{N} (p_i e^{\lambda} + 1 - p_i) \\
&= e^{-\lambda t} \prod_{i=1}^{N} (1 + p_i(e^{\lambda} - 1)) \\
&\le e^{-\lambda t} \prod_{i=1}^{N} e^{p_i(e^{\lambda}-1)} && (1 + x \le e^x, \forall x \in \mathbb{R}) \\
&= e^{-\lambda t} e^{(e^{\lambda}-1)\sum p_i} \\
&= e^{-\lambda t} e^{\mu(e^{\lambda}-1)}
\end{aligned}$$

By choosing the optimal $\lambda = \log \frac{t}{\mu}$ (HW: show that this choice is indeed optimal), the upper bound becomes $e^{-\mu}(\frac{e\mu}{t})^t$. □

Corollary 3.11 (HW, "small deviation Chernoff")**.** *Suppose*

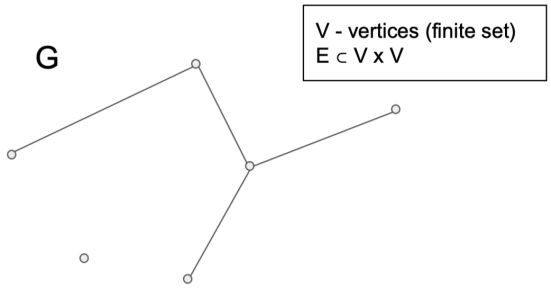
$$X_i = \begin{cases} 1, & w.p.\ p_i \\ 0, & w.p.\ 1 - p_i \end{cases},$$

and $X_1, \cdots, X_N$ are independent. Let $S_N = \sum_{i=1}^{N} X_i$ and $\mathbb{E}\, S_N = \mu$. Then, for all $\delta \in [0, 1]$,

$$\mathbb{P}\left(|S_N - \mu| \ge \delta\mu\right) \le 2e^{-c\mu\delta^2}.$$

3.4 Application to Random Graphs

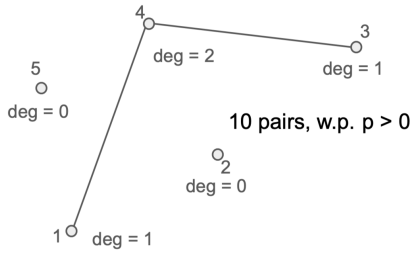
A graph is a pair (V, E) where V (vertices) is some finite set and E (edges) is a subset of $V \times V$. Schematically, if a pair (v_i, v_j) (where $v_i, v_j \in V$) belongs to E then there is an edge between v_i and v_j.



A random graph is a graph selected in some random way.

Definition 3.12 (Erdös-Rényi Random Graph)**.** Given a positive integer n and $p \in [0,1]$, a **random graph** $\mathbb{G}(n,p)$ satisfies the properties:

1. the set of vertices is $\{1,\ldots,n\}$.
2. For a given pair of vertices, there is an edge with probability $p \in (0,1)$.
3. The edge events are independent.



Recall that the degree of a vertex v, denoted by $deg(v)$, is the number of edges that terminate at v.

Theorem 3.13 (The degree of each vertex of $G(n,p)$ is well concentrated around its mean)**.** *Assume that there exists $c > 0$ such that for any vertex v of $G \sim \mathbb{G}(n,p)$ one has*

$$\mathbb{E}[\deg(v)] = d \geq c \log n.$$

Then, for all $v \in G$,

$$\mathbb{P}\left(\deg(v) \in [0.9, 1.1d]\right) \geq 0.9$$

Proof of Theorem 3.13. For $i \neq j$, let

$$X_{ij} = \begin{cases} 1, & \text{w.p. } p \\ 0, & \text{w.p. } 1-p \end{cases}$$

denote the existence of an edge between i-th and j-th vertex. Then $\deg(i) = \sum_{j \neq i} X_{ij}$. By Chernoff's inequality Corollary 3.11,

$$\mathbb{P}\left(|\deg(i) - d| \geq 0.1d\right) = \mathbb{P}\left(\left|\sum_{j \neq i} X_{ij} - d\right| \geq 0.1d\right) \leq 2e^{-cd} = 0.9.$$

for right c, $c > 0$. □

Remark 3.14. *Since the X_i are i.i.d, we see that $\mathbb{E}[\deg(v)] = p(n-1)$, and thus $d \geq c \log n$ is equivalent to $p \geq \frac{c' \log n}{n}$.*

3.5 Sub-Gaussian Random Variables

Next, we busy ourselves with the question:

Could we prove Hoeffding's inequality in a more general situation than bounded random variables? What sort of generality could we hope for?

Note that if the inequality

$$\mathbb{P}(|\sum_{i}^{n} a_i X_i| \geq t) \leq 2e^{-ct^2}$$

holds for all $a \in \mathbb{S}^{n-1}$, then by taking $a = e_i$, we see that necessarily

$$\mathbb{P}(|X_i| \geq t) \leq 2e^{-ct^2}. \tag{12}$$

We will show that (12) is in fact also a sufficient condition for Hoeffding's inequality to hold when $X_1, \cdots, X_n$ are independent. As a first step, we show the following (see Proposition 2.5.2 of [53]):

Proposition 3.15. *Let X be a random variable. Then, the following are equivalent (the parameters $K_1, K_2, K_3, K_4, K_5 \geq 0$ differ from one another at most by an absolute non-negative constant multiple):*

1. $\mathbb{P}(|X| \geq t) \leq 2e^{-t^2/K_1^2}$, *for all $t \geq 0$ and a fixed $K_1 > 0$;*
2. $||X||_p = (\mathbb{E}\,|X|^p)^{1/p} \leq K_2\sqrt{p}$ *for all $p \geq 1$;*
3. $\mathbb{E}\, e^{\lambda^2 X^2} \leq e^{K_3^2 \lambda^2}$ *for all $\lambda \in [-\frac{1}{K_3}, \frac{1}{K_3}]$;*
4. $\mathbb{E}\, e^{\frac{X^2}{K_4^2}} \leq 2$.

   Moreover, if $\mathbb{E}\, X = 0$, then conditions (1) - (4) are also equivalent to

5. $\mathbb{E}\, e^{\lambda X} \leq e^{K_5^2 \lambda^2}$ *for all $\lambda \in \mathbb{R}$.*

Proof. Firstly show (1) $\Rightarrow$ (2). Suppose that $\mathbb{P}(|X| \geq t) \leq 2e^{-t^2}$. We would like to show that in this case, $\mathbb{E}\,|X|^p \leq (C\sqrt{p})^p$. Indeed, by (3), and using a change of variables, we get

$$\begin{aligned}\mathbb{E}\,|X|^p &= \int_0^\infty \mathbb{P}(|X|^p > t)\ \mathrm{d}t \\ &= p\int_0^\infty \mathbb{P}(|X| > s)s^{p-1}\ \mathrm{d}s \\ &\leq 2p\int_0^\infty s^{p-1}e^{-s^2}\ \mathrm{d}s = p\Gamma\left(\frac{p}{2}\right)\cdot C\end{aligned}$$

In the last passage we used the assumption. It remains to recall that

$$\Gamma(m) = \int_0^\infty t^{m-1}e^{-t}dt \quad \Gamma\left(\frac{p}{2}\right) = (C\sqrt{p})^p.$$

Next we show (2) $\Rightarrow$ (3). Suppose $(\mathbb{E}|X|^p)^{\frac{1}{p}} \leq \sqrt{p}$. By Taylor's formula, $e^a = \sum_{k=0}^\infty \frac{a^k}{k!}$ for any $a \in [0,1]$, and therefore

$$\begin{aligned}\mathbb{E}e^{\lambda^2X^2} &= \mathbb{E}\sum_{k=0}^\infty \frac{(\lambda^2X^2)^k}{k!} = \sum_{k=0}^\infty \frac{\lambda^{2k}\mathbb{E}X^{2k}}{k!} \\ &\leq \sum_{k=0}^\infty \frac{(\lambda)^{2k}(2k)^k}{k!} \\ &\leq \sum_{k=0}^\infty (C\lambda)^{2k} \qquad \text{(by Stirling's formula } k! = \sqrt{2\pi k}\left(\tfrac{k}{e}\right)^k\left(1+O\left(\tfrac{1}{k}\right)\right)\text{)} \\ &= \frac{1}{1-(C\lambda)^2} \\ &\leq e^{C\lambda^2} \quad \text{for} \quad C|\lambda| \in [0,1).\end{aligned}$$

In the last passage we used the inequality $1 - x \leq e^{-x}$ (see home work).

It is straightforward to show (3) $\Rightarrow$ (4) : take $s = \frac{c}{K_3}$.

To verify (4) $\Rightarrow$ (1), note that

$$\mathbb{P}(X \geq t) = \mathbb{P}(e^{x^2} \geq e^{t^2}) \leq \mathbb{E}e^{x^2}e^{-t^2} \leq 2e^{-t^2},$$

where the last inequality uses the assumption $\mathbb{E}e^{x^2} \leq 2$.

Now check (3) $\Rightarrow$ (5). Suppose $\mathbb{E}X = 0$ and $\mathbb{E}e^{\lambda^2x^2} \leq e^{\lambda^2},\ \forall\lambda \in [-1,1]$. We use the inequality $e^x \leq x + e^{x^2}, \forall x \in \mathbb{R}$ (which we leave as a home work).

$$\begin{aligned}\mathbb{E}e^{\lambda x} &\leq E(\lambda x + e^{\lambda^2x^2}) \\ &= \mathbb{E}e^{\lambda^2x^2} && \text{(Since we assumed } \mathbb{E}x = 0\text{)} \\ &\leq e^{\lambda^2}. && \text{(if } |\lambda| < 1\text{, apply (3))}\end{aligned}$$

Now assume $|\lambda| \geq 1$. Then we use the inequality $2\lambda x \leq x^2 + \lambda^2, \forall \lambda, x \in \mathbb{R}$, which is also left as a home work.

$$\begin{aligned} \mathbb{E}e^{\lambda x} \leq & e^{\lambda^2/2}\mathbb{E}e^{x^2/2} && \text{(apply (3))} \\ \leq & e^{\lambda^2/2}\sqrt{e} \leq e^{\lambda^2}. \end{aligned}$$

Lastly, we show (5) $\Rightarrow$ (1). Suppose $\mathbb{E}X = 0$, and $\mathbb{E}e^{\lambda x} \leq e^{\lambda^2}$, for all $\lambda \in \mathbb{R}$. Then

$$\mathbb{P}(X \geq t) = \mathbb{P}(e^{\lambda X} \geq e^{\lambda t}) \leq e^{-\lambda t}\mathbb{E}e^{\lambda X} \leq e^{-\lambda t}e^{\lambda^2}.$$

We plug the optimal value of $\lambda = \frac{t}{2}$, which gives $\mathbb{P}(X \geq t) \leq e^{-t^2/4}$. Applying the same argument to the lower tail, we get

$$\mathbb{P}(|X| \geq t) = \mathbb{P}(X \geq t) + \mathbb{P}(X \leq -t) \leq 2 \cdot e^{-t^2/4}.$$

□

This brings us to a

Definition 3.16. (Sub-Gaussian Random Variables) If a random variable X satisfies (either of the) properties (1) - (4), it is called a **sub-Gaussian** random variable. The sub-Gaussian norm of X, denoted $\|X\|_{\psi_2}$, is defined to be the smallest K_4 in property 4. In other words, we define

$$\|X\|_{\psi_2} = \inf\{t > 0 : \mathbb{E}\exp(X^2/t^2) \leq 2\}.$$

The sub-Gaussian norm is indeed a norm, which is left as a home work.

Example 3.17 (Examples of sub-Gaussian variables)**.** *The following random variables are sub-Gaussian:*

- ***Gaussian****: $X \sim N(0,1)$ is a sub-gaussian random variable with $\|X\|_{\psi_2} \leq C$, where C is an absolute constant.*
- ***Bernoulli****: Let X be symmetric Bernoulli random variable. Since $|X| = 1$, it follows that X is a sub-Gaussian random variable with*

  $$\|X\|_{\psi_2} = \frac{1}{\sqrt{\ln 2}}.$$

- ***Bounded****: More generally, any bounded random variable X with $|X| \leq M$ almost surely, is sub-Gaussian with*

  $$\|X\|_{\psi_2} \leq C\|X\|_\infty$$

  where $C = \frac{1}{\sqrt{\ln 2}}$.

Next, we present

Example 3.18 (Examples of NON sub-Gaussian random variables)**.** *The following random variables are NON sub-Gaussian:*

- ***Exponential****: An exponential random variable is not sub-Gaussian, since*

$$\mathbb{P}(|X| \geq t) \not\leq e^{-Ct^2}$$

The probability density function (pdf) of an exponential random variable is:

$$f_X(x) = \begin{cases} 0 & \text{if } x < 0 \\ e^{-x} & \text{if } x \geq 0 \end{cases}.$$

And the tail probability is given by:

$$\mathbb{P}(X \geq t) = \int_t^\infty e^{-s} \, \mathrm{d}s = e^{-t}.$$

Thus we could see that it violates the definition of sub-Gaussian random variables.

- ***Poisson:*** *Poisson random variables are not sub-Gaussian.*
- ***Cauchy:*** *Cauchy random variables are not sub-Gaussian.*

3.6 General Hoeffding's inequality and Khinchine's inequality

Recall the fact that a sum of independent normal random variables X_i is normal. More precisely, if $X_i \sim N(0, \sigma_i^2)$ are independent then

$$\sum_{i=1}^N X_i \sim N\left(0, \sum_{i=1}^N \sigma_i^2\right). \tag{13}$$

This property of the normal distribution extends to general sub-Gaussian distributions (see also Proposition 2.6.1 of [53]):

Proposition 3.19. *(Sums of independent sub-Gaussians) Let $X_1, \ldots, X_N$ be independent, mean zero, sub-Gaussian random variables. Then $\sum_{i=1}^N X_i$ is also a sub-Gaussian random variable, and*

$$\left\| \sum_{i=1}^N X_i \right\|_{\psi_2}^2 \leq C \sum_{i=1}^N ||X_i||_{\psi_2}^2.$$

Proof.

$$\mathbb{E}\exp\left(\lambda\sum_{i=1}^{N}X_i\right)=\prod_{i=1}^{N}\mathbb{E}\exp(\lambda X_i) \qquad \text{(by independence)}$$
$$\leq\prod_{i=1}^{N}\exp\left(C^2\|\lambda X_i\|_{\psi_2}^2\right)=\exp\left(\lambda^2K^2\right),$$

where $K^2:=C^2\sum_{i=1}^{N}\|X_i\|_{\psi_2}^2$. By equivalence of properties (4) and (5) in Proposition 3.15 and Definition 3.16, we see that the sum $\sum_{i=1}^{N}X_i$ is sub-Gaussian, and

$$\left\|\sum_{i=1}^{N}X_i\right\|_{\psi_2}\leq C_1K.$$

□

Remark 3.20. *If $X_i\sim N(0,\sigma_i^2)$ then $||X_i||_{\psi_2}^2=\sigma_i^2$, and an equality holds in place of the inequality above.*

Proposition 3.19 is, in fact, nothing but the generalized form of the Hoeffding inequality!

Corollary 3.21. *(General Hoeffding's inequality) Let $X_1,\ldots,X_N$ be independent, mean zero, sub-Gaussian random variables, and $a=(a_1,\ldots,a_N)\in\mathbb{R}^N$. Then, for every $t\geq 0$, we have*

$$\mathbb{P}\left(\left|\sum_{i=1}^{N}a_iX_i\right|\geq t\right)\leq 2\exp\left(-\frac{ct^2}{K^2|a|^2}\right),$$

where $K=\max_i\|X_i\|_{\psi_2}$.

Proof.

$$\mathbb{P}\left(\left|\sum_{i=1}^{N}a_iX_i\right|\geq t\right)\leq 2\exp\left(-\frac{ct^2}{K^2|a|^2}\right)$$

is equivalent to (1) in Proposition 3.15. By Proposition 3.19,

$$\left\|\sum_{1}^{N}a_iX_i\right\|_{\psi_2}^2\leq C\sum_{1}^{N}\|a_iX_i\|_{\psi_2}^2=C\sum_{1}^{N}a_i^2\|X_i\|_{\psi_2}^2\leq CK^2|a|^2.$$

□

Remark 3.22. *Note that in some cases, a tighter version of Hoeffding's inequality, which is evident from the argument above, may be useful: if $X_1,\ldots,X_N$ be independent, mean zero, sub-Gaussian random variables, and $a=(a_1,\ldots,a_N)\in\mathbb{R}^N$, then, for every $t\geq 0$, we have*

$$\mathbb{P}\left(\left|\sum_{i=1}^{N}a_iX_i\right|\geq t\right)\leq 2\exp\left(-\frac{ct^2}{\sum_1^N a_i^2\|X_i\|_{\psi_2}^2}\right).$$

As an immediate corollary of Hoeffding's inequality Proposition 3.19 and the equivalence of the sub-Gaussian properties (2) and (4) in Proposition 3.15, we get:

Theorem 3.23. *(Khintchine's inequality). Let $X_1, \dots, X_N$ be independent sub-gaussian random variables with zero means and unit variances, and let $a = (a_1, \dots, a_N) \in \mathbb{R}^N$. Prove that for every $p \in [2, \infty)$ we have*

$$\left(\sum_{i=1}^N a_i^2 \right)^{1/2} \leq \left\| \sum_{i=1}^N a_i X_i \right\|_{L_p} \leq CK\sqrt{p} \left(\sum_{i=1}^N a_i^2 \right)^{1/2}$$

where $K = \max_i \|X_i\|_{\psi_2}$ and C is an absolute constant.

Remark 3.24 (centering)**.** *In many results above we impose the assumption $\mathbb{E}\, X = 0$ for convenience. But note that this assumption is often non-essential. Indeed, in the case where $X_1, X_2, \dots, X_n$ are sub-Gaussian and $\mathbb{E}X_i \neq 0$, we may center X_i by taking $Y_i = X_i - \mathbb{E}X_i$. It can be verified (see Lemma 2.6.8 of [53]) that Y_i's are also sub-Gaussian, and*

$$||Y_i||_{\psi_2} = ||X_i - \mathbb{E}(X_i)||_{\psi_2} \leq \|X_i\|_{\phi_2} + \|\mathbb{E}(X_i)\|_{\psi_2} \leq C\|X_i\|_{\psi_2}.$$

3.7 Sub-Exponential Random Variables

When independent random variables are not sub-Gaussian, we can no longer apply Hoeffding's inequality to bound the tail probability for their sum. However, there are important situations when this needs to be done, and is also (as we will see) possible. For example, for a standard Gaussian random vector $X \in \mathbb{R}^n$, that is, a random vector with i.i.d. coordinates $X_i \sim \mathcal{N}(0, 1)$, consider the random variable $|X|$ – the length of the standard Gaussian random vector. We could study $|X|^2 = \sum_{i=1}^n X_i^2$, which is the sum of i.i.d. random variables, but unfortunately we cannot apply Hoeffding's inequality: X_i^2 are not sub-Gaussian. Indeed, the tails of X_i^2 decrease only exponentially fast:

$$\mathbb{P}(X_i^2 \geq t) = \mathbb{P}(|X_i| \leq \sqrt{t}) \approx 2e^{-\frac{(\sqrt{t})^2}{2}} = 2e^{-\frac{t}{2}}.$$

In this section, we are going to discuss such "sub-exponential" random variables, and study their properties, which, as it turns out, are also very nice.

Definition 3.25 (Sub-exponential random variable)**.** a random variable X is called **sub-exponential** if there exists $K > 0$ such that

$$\mathbb{P}(|X| > t) \leq 2e^{-\frac{t}{K}}$$

for all $t > 0$.

Example 3.26 (Examples of sub-exponential random variables)**.**

- *Let X be an exponential random variable, i.e., X has probability density function*

$$f_X(t) = \begin{cases} 0, & t < 0 \\ e^{-t}, & t \geq 0 \end{cases}$$

Since $\mathbb{P}(X > s) = e^{-s}$, X is sub-exponential.

- *If $X \sim \mathcal{N}(0,1)$, then X^2 is sub-exponential.*

Proposition 3.27 (Equivalent definitions of sub-exponential random variables)**.** *For a random variable X, the following are equivalent (here the non-negative constants K_i differ from each other at most by an absolute constant factors):*

(a) $\mathbb{P}(|X| \geq t) \leq e^{-\frac{t}{K_1}}$, *for all* $t > 0$;

(b) $(\mathbb{E}|X|^p)^{\frac{1}{p}} \leq K_2 p$ *for all* $p \geq 1$;

(c) $\mathbb{E}(e^{\lambda |X|}) \leq e^{K_3 \lambda}$, *for all* $\lambda \in [0, \frac{1}{K_3}]$;

(d) $\mathbb{E}(e^{\frac{|X|}{K_4}}) \leq 2$;

If we further assume $\mathbb{E}X = 0$, then properties (a)-(d) are equivalent to

(e) $\mathbb{E}(e^{\lambda X}) \leq e^{K_5^2 \lambda^2}$ *if* $|\lambda| \leq \frac{1}{K_5}$.

The proof of the proposition is left as an exercise.

Definition 3.28. (ψ_1-norm) We define the sub-exponential norm as $\|X\|_{\psi_1} := \inf\{K > 0 : \mathbb{P}(|X| \geq t) \leq 2e^{-\frac{t}{K}}, \forall t \geq 0\}$. Checking that this is a norm is left as an exercise.

Note that a sub-Gaussian random variable is always sub-exponential. But more is true:

Lemma 3.29. *A random variable X is sub-Gaussian if and only X^2 is sub-exponential, and in fact $\|X^2\|_{\psi_1} = \|X\|_{\psi_2}^2$.*

Proof. Suppose that X is sub-Gaussian. Then, for all $t > 0$, $\mathbb{P}(|X| \geq t) \leq 2e^{-t^2/K}$ for some constant $K > 0$. Therefore,

$$\mathbb{P}(X^2 \geq t) \leq \mathbb{P}(|X| \geq \sqrt{t}) \leq 2e^{-\frac{t}{K^2}}.$$

Therefore, X^2 is sub-exponential. Reversing the calculation above establishes the reverse implication. The relation between the norms also follows. □

More generally,

Lemma 3.30. *Suppose that X, Y are sub-Gaussian random variables. Then XY is sub-exponential, and*

$$\|XY\|_{\psi_1} \leq \|X\|_{\psi_2} \cdot \|Y\|_{\psi_2} \tag{14}$$

Proof. WLOG, we can assume $\|X\|_{\psi_2} = \|Y\|_{\psi_2} = 1$, otherwise replace X by $\tilde{X} = \frac{X}{\|X\|_{\psi_2}}$ and Y by $\tilde{Y} = \frac{Y}{\|Y\|_{\psi_2}}$. Note that

$$\|\tilde{X}\tilde{Y}\|_{\psi_1} \leq \|\tilde{X}\|_{\psi_2}\|\tilde{Y}\|_{\psi_2}$$

if and only if

$$\|XY\|_{\psi_1} \leq \|X\|_{\psi_2}\|Y\|_{\psi_2}.$$

Now, using property (d) in Proposition 3.27, we have $\mathbb{E}(e^{X^2}) \leq 2, \mathbb{E}(e^{Y^2}) \leq 2$. We claim that

$$\mathbb{E}(e^{|XY|}) \leq 2.$$

Indeed, recall Young's inequality

$$\frac{a^2 + b^2}{2} \geq |ab|$$

for all real numbers a, b, as follows from the fact that $(a-b)^2 \geq 0$ and $(a+b)^2 \geq 0$. Using this inequality twice, we get:

$$\begin{aligned}
\mathbb{E}(e^{|XY|}) &\leq \mathbb{E}\, e^{\frac{X^2+Y^2}{2}} && (\text{ Young's inequality}) \\
&= \mathbb{E}\left(e^{\frac{X^2}{2}} \cdot e^{\frac{Y^2}{2}}\right) \\
&\leq \frac{1}{2} \cdot \mathbb{E}(e^{X^2} + e^{Y^2}) && (\text{ Young's inequality}) \\
&= \frac{1}{2}\left(\mathbb{E}\, e^{X^2} + \mathbb{E}\, e^{Y^2}\right) \\
&\leq 2.
\end{aligned}$$

□

Remark 3.31. *Note that the constant* 1 *in Lemma 3.30 can only be there when the definitions of* $\|\cdot\|_{\psi_1}$ *and* $\|\cdot\|_{\psi_2}$ *correspond to each other. Above, we used definitions (1) from Proposition 3.15 and (a) from Proposition 3.27.*

Lemma 3.30 can also be derived using other equivalent definition of sub-Gaussian and sub-exponential random variables correspondingly, so long as they match. This is left as a home work.

Remark 3.32. *Analogously to the centering trick of sub-Gaussians, if* X *is sub-exponential, then* $X - \mathbb{E}X$ *is also sub-exponential, and*

$$\|X - \mathbb{E}(X)\|_{\psi_1} \leq C \cdot \|X\|_{\psi_1}$$

for some absolute constant C*. (home work)*

3.8 Bernstein's inequality

The Bernstein inequality for sub-exponential distributions provides a bound on the probability that the sum of independent sub-exponential random variables deviates from its expected value. One may draw a parallel between Bernstein's inequality and Hoeffding's inequality, both in terms of the statement and the proof method. Historically, Berstein's work was published between 1920 and 1930, Chernoff's paper appeared in 1952, and Hoeffding's inequality was published by him in 1963.

Theorem 3.33 (Bernstein's inequality)**.** *Let $X = (X_1, X_2, \cdots, X_n) \in \mathbb{R}^n$ be a random vector, where the X_i are independent sub-exponential random variables with $\mathbb{E}\, X = 0$. Then, for all $t \geq 0$,*

$$\mathbb{P}\left(\left|\sum_{i=1}^n X_i\right| \geq t\right) \leq 2\exp\left(-c \cdot \min\left(\frac{t^2}{\sum\limits_{i=1}^n \|X_i\|_{\psi_1}}, \frac{t}{\max\limits_i \|X_i\|_{\psi_1}}\right)\right).$$

Here $c > 0$ is an absolute constant.

Proof. Let $K = \max\limits_{1\leq i\leq n} \|X_i\|_{\psi_1}$. Then, for all $t \in \mathbb{R}$, using Markov's inequality, we see that

$$\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^n X_i \geq t\right) &= \mathbb{P}\left(\exp\left[\lambda \sum_{i=1}^n X_i\right] \geq e^{\lambda t}\right) \\
&\leq e^{-\lambda t} \cdot \mathbb{E}\exp\left(\lambda \sum_{i=1}^n X_i\right) \\
&= e^{-\lambda t} \cdot \prod_{i=1}^n \mathbb{E}\, e^{\lambda X_i}.
\end{aligned}$$

Since the X_i are centered, property (e) in Proposition 3.27 implies that

$$\mathbb{E}\, e^{\lambda X_i} \leq e^{c\lambda^2 \|X_i\|_{\psi_1}^2}$$

for all λ satisfying $|\lambda| \leq c/K$. So for $\lambda \leq \frac{1}{K}$,

$$\begin{aligned}
\mathbb{E}\left(\sum_{i=1}^n X_i \geq t\right) &\leq e^{-\lambda t} \cdot \prod_{i=1}^n \mathbb{E}\, e^{\lambda X_i} \\
&\leq e^{-\lambda t} \cdot \prod_{i=1}^n e^{c\lambda^2 \|X_i\|_{\psi_1}^2} \\
&= e^{-\lambda t + c\lambda^2 \sigma^2},
\end{aligned}$$

where $\sigma^2 = \sum_{i=1}^{n} \|X_i\|_{\psi_1}^2$. To minimize the expression in λ under the constraint $\lambda \leq \frac{1}{K}$, we take $\lambda = \min\{\frac{t}{2c\sigma^2}, \frac{c}{K}\}$. Therefore,

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i \geq t\right) \leq \exp\left(\min\left\{\frac{t^2}{4c\sigma^2}, \frac{ct}{2K}\right\}\right).$$

By a similar reasoning, we see that

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i < -t\right) \leq \exp\left(\min\left\{\frac{t^2}{4c\sigma^2}, \frac{ct}{2K}\right\}\right).$$

The theorem then follows from combining the two bounds above. □

Applying the Bernstein's inequality to $a_i X_i$, we obtain the following inequality:

Corollary 3.34. *Let $X = (X_1, X_2, \cdots, X_n) \in \mathbb{R}^n$ be a random vector, where the X_i are independent and sub-exponential, $\mathbb{E}X_i = 0$, and $a \in \mathbb{R}^n$. Let $K = \max_i \|X_i\|_{\psi_1}$. Then for all $t \geq 0$,*

$$\mathbb{P}(|\langle X, a\rangle| \geq t) \leq 2e^{-c\cdot\min\left\{\frac{t^2}{K^2|a|^2}, \frac{t}{K\|a\|_\infty}\right\}}$$

If $a = (\frac{1}{n}, \ldots, \frac{1}{n})$, then we obtain the following inequality:

Corollary 3.35.

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i\right| \geq t\right) \leq 2\exp\left(-c\cdot\min\left\{\frac{t^2}{K^2}, \frac{nt}{K}\right\}\right).$$

If the X_i are bounded by a universal constant, then the bound in Bernstein's inequality can be further strengthened. For example, see Theorem 2.8.4 in [53].

3.9 Concentration of the norm of a random vector with independent sub-Gaussian coordinates

A model application of Bernstein's inequality is the following inequality:

Theorem 3.36 (Concentration of the norm of a random vector with independent sub–Gaussian coordinates)**.** *Let $X = (X_1, X_2, \ldots, X_n)$ be a random vector, where the X_i are independent. Suppose that each X_i is sub-Gaussian with the constant at most $K > 0$, and $\mathbb{E}\, X_i^2 = 1$. Then,*

$$\|X - \sqrt{n}\|_{\psi_2} \leq CK^2,$$

where C is an absolute constant.

Remark 3.37. *Equivalently, the theorem states that*

$$\mathbb{P}(|X - \sqrt{n}| \geq t) \leq 2e^{-\frac{ct^2}{K^2}}$$

i.e. X belongs, with high probability, to a thin shell around the sphere of radius $\sqrt{n}$. This result holds, and is interesting already when X is Gaussian.

Proof. Since each X_i is sub-Gaussian, we know that X_i^2 is sub-exponential and so is $X_i^2 - 1$. Furthermore,

$$\begin{aligned}\left\|X_i^2 - 1\right\|_{\psi_1} &\leq C \left\|X_i^2\right\|_{\psi_1} \\ &= C \left\|X_i\right\|_{\psi_2}^2 \\ &\leq CK^2.\end{aligned}$$

Without loss of generality, we may assume $K \geq 1$. Applying Bernstein's inequality (Theorem 3.33) to $\frac{1}{n}|X|^2 - 1$, we see that for any $s > 0$

$$\mathbb{P}\left(\left|\frac{1}{n}|X|^2 - 1\right| \geq s\right) \leq e^{-\frac{cn}{K^4}\min\{s, s^2\}},$$

where we used that $K \geq 1$. Note that for all $z \geq 0$ and $\delta > 0$, $|z - 1| \geq \delta$ implies $|z^2 - 1| \geq \max(\delta^2, \delta)$. Therefore,

$$\begin{aligned}\mathbb{P}\left(\left|\frac{1}{\sqrt{n}}|X| - 1\right| \geq \delta\right) &\leq \mathbb{P}\left(\left|\frac{1}{n}|X|^2 - 1\right| \geq \max(\delta^2, \delta)\right) \\ &\leq 2e^{-\frac{cn}{K^4}\delta^2}\end{aligned}$$

Finally, taking $t = \delta\sqrt{n}$ yields

$$\mathbb{P}\left(\left||X| - \sqrt{n}\right| \geq t\right) \leq 2e^{-\frac{ct^2}{K^4}}.$$

□

3.10 Sub-Gaussian Random vectors

A random vector is called **sub-Gaussian** if each of its one-dimensional projections is sub-Gaussian:

Definition 3.38. Let $X \in \mathbb{R}^n$ be a random vector. Then, we say that X is **sub-Gaussian** if for all $\theta \in \mathbb{S}^{n-1}$, the random variable $\langle X, \theta\rangle$ is sub-Gaussian. The associated sub-Gaussian norm $\|\cdot\|_{\psi_2}$ is defined by

$$\|X\|_{\psi_2} := \sup_{\theta \in \mathbb{S}^{n-1}} \|\langle X, \theta\rangle\|_{\psi_2}.$$

We will see that sub-Gaussian random vectors form a rich class which includes many examples. Firstly, consider the standard Gaussian random vector $X =\sim \mathcal{N}(0, \mathrm{Id})$ on $\mathbb{R}^n$: the random vector whose coordinates X_i are independent standard normal random variables. Then X is sub-Gaussian since all the $\langle X, \theta\rangle \sim \mathcal{N}(0,1)$ by (13), and therefore, they are also sub-Gaussian.

Proposition 3.39 (An example of a sub-Gaussian random vector)**.** *Let $X \in \mathbb{R}^n$ be a random vector such that $X = (X_1, \cdots, X_n)$ is a random vector, the coordinates X_i are independent, $\mathbb{E}\, X_i = 0$ and $\|X_i\|_{\psi_2} \leq K$ for $K > 0$. Then X is a sub-Gaussian random vector and $\|X\|_{\psi_2} \leq c \cdot K$.*

Proof. By Proposition 3.19, for any $\theta \in \mathbb{S}^{n-1}$,

$$\|\langle X, \theta\rangle\|_{\psi_2}^2 = \left\|\sum_{i=1}^n \theta_i X_i\right\|_{\psi_2}^2 \leq C \sum_{i=1}^n |\theta_i|^2 \|X_i\|_{\psi_2}^2 \leq C \max_{1\leq i\leq n} \|X_i\|_{\psi_2}^2 = CK^2.$$

This implies the Proposition. □

Therefore, the random vector uniformly distributed in $\{-1, 1\}^n$ is sub-Gaussian. Also the uniform random vector in a cube is sub-Gaussian. More generally, any random vector whose coordinates are independent and bounded is sub-Gaussian.

Remark 3.40. *What if we do not assume independence in general? Let $X = (X_1, ..., X_n)$ and all of its coordinates are $K-$sub-Gaussian, but possibly dependent. Since $\|\cdot\|_{\psi_2}$ is a norm, we have*

$$\|\langle X, \theta\rangle\|_{\psi_2} \leq \sum_{i=1}^n |\theta_i| \cdot \|X_i\|_{\psi_2} \leq K \cdot \sum_{i=1}^n |\theta_i| \leq \sqrt{n}K.$$

In the last passage we used the fact that for any $\theta \in \mathbb{S}^{n-1}$,

$$\sum_{i=1}^n |\theta_i| \leq \sqrt{n}\sqrt{\sum_{i=1}^n \theta_i^2} = \sqrt{n},$$

and in fact the equality in the above is attained for $\theta = (1/\sqrt{n}, ..., 1/\sqrt{n})$. Therefore,

$$\mathbb{P}(|\langle X, \theta\rangle| \geq t) \leq 2e^{\frac{ct^2}{nK^2}},$$

or in other words, X is a $\sqrt{n}K$-sub-Gaussian random vector. The loss of $\sqrt{n}$ can be necessary (as an example, one may consider $X = (a, ..., a)$ where a is a fixed random variable), but it could be problematic for applications.

In light of Remark 3.40 one may wonder if there is a sub-Gaussian random vector with the sub-Gaussian constant which does not depend on the dimension but with dependencies among the coordinates. Below we show one such example:

Proposition 3.41. *The uniform distribution on the sphere $X \sim Unif(\sqrt{n}S^{n-1})$ is sub-Gaussian and $\|X\|_{\psi_2} \leq C$ for some absolute constant $C > 0$ that does not depend on the dimension.*

Remark 3.42. *The normalization $\sqrt{n}\mathbb{S}^{n-1}$ is natural because this means that $\mathbb{E}\, X_i^2 = 1$ for all i, so the Propositions 3.39 and 3.41 compare in a natural way.*

Proof. For the standard Gaussian random vector $g \sim N(0, \mathrm{Id})$, the normalized random vector $\frac{g}{\|g\|} \sim Unif(\mathbb{S}^{n-1})$ (we live this as a home work). Therefore we can represent $X = \frac{\sqrt{n}g}{|g|}$. Therefore, using the notation Z for the standard normal random variable, we get

$$\begin{aligned}
\mathbb{P}\left(\langle X, \theta\rangle \geq t\right) &= \mathbb{P}\left(\left\langle \frac{\sqrt{n}g}{\|g\|}, \theta \right\rangle \geq t\right) \\
&= \mathbb{P}\left(\frac{\sqrt{n}}{\|g\|} \cdot Z \geq t\right) \\
&= \mathbb{P}\left(\left\{\frac{\sqrt{n}}{\|g\|} \cdot Z \geq t\right\} \cap \left\{\left|\|g\| - \sqrt{n}\right| \geq ct\sqrt{n}\right\}\right) \\
&\qquad + \mathbb{P}\left(\left\{\frac{\sqrt{n}}{\|g\|} \cdot Z \geq t\right\} \cap \left\{\left|\|g\| - \sqrt{n}\right| < ct\sqrt{n}\right\}\right) \\
&\leq \mathbb{P}\left(Z \geq (1-c)t\right) + e^{-c't^2} \leq 2e^{-c''t^2},
\end{aligned}$$

where in the last passage we used Propositions 11 and 3.36. □

3.11 Grothendieck's inequality

We now apply our results about sub-Gaussian random vectors to deduce the following fact from linear algebra, which is useful in semi-definite programming (see Vershynin [53]).

Theorem 3.43 (Grothendieck's inequality)**.** *Suppose A is an $m \times n$ matrix with real entries such that for all $x, y \in \{-1, 1\}^n$ (set of vectors with coordinates with $+1$ and -1) we have*

$$|\langle Ax, y\rangle| \leq 1.$$

Then, for all vectors $u_i, v_j \in \mathbb{S}^{n-1}$ we have

$$\left|\sum_{i,j} a_{ij}\langle u_i, v_j\rangle\right| \leq K,$$

where $K > 0$ is an absolute constant.

Proof. First of all, note that the assumption can be equivalently stated as

$$\left|\sum_{i,j} a_{ij} x_i y_j\right| \leq \max_i |x_i| \max_j |y_j|$$

for any collection $x_1, \dots, x_m, y_1, \dots, y_n$. The conclusion appropriately changes to

$$\left| \sum_{i,j} a_{ij} \langle u_i, v_j \rangle \right| \leq K \max_i |u_i| \max_j |v_j|.$$

If we did not wish for an absolute constant, we could take $K = \sum_{i,j} |a_{ij}|$. However, our goal is to obtain an absolute constant K (bounded above by 288, following [53]).

Suppose $K > 0$ is the optimal number for which Grothendick's theorem holds *for any Hilbert space,* and fix a collection of vectors $u_i, v_j \in \mathbb{S}^{n-1}$ where $\sum_{i,j=1}^n a_{ij} \langle u_i, v_j \rangle = K$.

Consider random variables U_i and V_j, given by $U_i = \langle g, u_i \rangle$ and $V_j = \langle g, v_j \rangle$ where $g \sim N(0, \mathrm{Id})$. One can show that $\mathbb{E}(U_i V_j) = \langle u_i, v_j \rangle$ (which is left as a home work).

By construction, we have

$$K = \sum_{i,j} a_{ij} \langle u_i, v_j \rangle = \sum_{i,j} a_{ij} \, \mathbb{E}(U_i V_j) = \mathbb{E} \sum_{i,j} a_{ij} U_i V_j.$$

If $|U_i| \leq R$ and $|V_j| \leq R$ almost surely for some $R > 0$, then

$$K \leq R^2 \sum_{i,j} a_{ij} \leq R^2,$$

where the last passage follows from the assumption of the Theorem: indeed, if we take all the coordinates of x and y to be 1, we get $\sum a_{ij} \leq 1$.

Write $U_i = U_i^+ + U_i^-$, where $U_i^+ = U_i \cdot \mathbf{1}_{\{|U_i| \leq R\}}$ and $U_i^- = U_i \cdot \mathbf{1}_{\{|U_i| > R\}}$. Similarly, decompose $V_i = V_i^+ + V_i^-$.

Note that $|U_i^+| \leq R$ almost surely. We have

$$\begin{aligned} K = \mathbb{E} \sum_{i,j} a_{ij} U_i V_j &= \mathbb{E} \left(\sum_{i,j} a_{ij} U_i^+ V_j^+ + \sum_{i,j} a_{ij} U_i^+ V_j^- + \sum_{i,j} a_{ij} U_i^- V_j^+ + \sum_{i,j} a_{ij} U_i^- V_j^- \right) \\ &\leq R^2 \cdot \sum_{i,j} a_{ij} + \mathbb{E} \left(\sum_{i,j} a_{ij} U_i^+ V_j^- + \sum_{i,j} a_{ij} U_i^- V_j^+ + \sum_{i,j} a_{ij} U_i^- V_j^- \right) \\ &\leq R^2 + \mathbb{E} \left(\sum_{i,j} a_{ij} U_i^+ V_j^- + \sum_{i,j} a_{ij} U_i^- V_j^+ + \sum_{i,j} a_{ij} U_i^- V_j^- \right). \end{aligned}$$

The first can be bounded using the fact that $U_i^+, V_i^+ \leq R$ by construction. For the other three terms, we use the fact that $\mathbb{E}\,|U_i^-|^2 \leq \frac{4}{R^2}$ and $\mathbb{E}\,|V_j^-|^2 \leq \frac{4}{R^2}$ (which we leave as a home work). Using Grothendieck's inequality with the constant K (indeed, by our assumption it holds with the constant K on any Hilbert space, so we are using it on the space of random variables with the scalar product $\mathbb{E}(XY)$), and using also the bound $\mathbb{E}\,|V_j^+|^2 \leq \mathbb{E}\,|V_j|^2 \leq 1$, we get

$$\mathbb{E} \sum_{i,j} a_{ij} U_i^- V_j^+ = \sum_{i,j} a_{ij} \, \mathbb{E}(U_i^- V_j^+) \leq K \left(\mathbb{E}\,|U_i^-|^2 \right)^{1/2} \left(\mathbb{E}\,|V_j^+|^2 \right)^{1/2} \leq \frac{2K}{R}.$$

The other two sums can be bounded in the same way. Putting everything together, we get

$$K \leq R^2 + \frac{6K}{R}.$$

Choosing $R = 12$ and solving the resulting inequality leads to $K \leq 288$. □

4 Random Matrices

Definition 4.1. A random $N \times n$ matrix $A \in \mathbb{R}^{N \times n}$ is a matrix drawn in some random way. We shall use notation

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{Nn} \end{bmatrix} \tag{15}$$

where the entries a_{ij} are random variables.

Example 4.2. *Here are some examples of random matrices:*

- *When $a_{ij} \sim N(0,1)$ and are independent, A is sometimes called a Gaussian random matrix;*
- *More generally, we may consider a_{ij} to be independent random variables selected according to some distributions;*
- *We may also consider a symmetric random matrix, by selecting the upper corner entries independently, and reinforcing the rule $a_{ij} = a_{ji}$;*
- *We may also select one random variable a and fill each entry with it;*
- *We may select some specific entries randomly independently, while other entries would have fixed values (for example zeroes);*
- *another example of a random matrix is a random rotation (selected uniformly from the compact set of rotations).*

We think about the random $N \times n$ matrix as about the operator on $\mathbb{R}^n$ into $\mathbb{R}^N$. We will study various properties of random matrices assuming that N and n are very large (but not reaching the infinity limit). This is what is informally called a *non-asymptotic random matrix theory*. We will be using the methods of High-dimensional Probability which is only one out of the myriad of methods and theories that come useful to study random matrices. We will leave a lot of the relevant methods and questions and theories beyond the scope of this course.

For convenience we will assume throughout that $N \geq n$ since most of the properties of a matrix A which we study can be easily transferred to properties of A^T.

What properties of random matrices can be studied? One example is singular values. Recall that the singular values of a matrix $A \in \mathbb{R}^{N\times n}$ are given by $\sigma_i(A) = \sqrt{\lambda_i(AA^T)}$, where λ_i are the eigenvalues, and we suppose that $\sigma_1(A) \geq \cdots \geq \sigma_n(A)$. Recall that

$$\begin{aligned} \sigma_1(A) &= \|A\| = \sup_{x\in\mathbb{S}^{n-1}} |Ax|; \\ \sigma_n(A) &= \inf_{x\in\mathbb{S}^{n-1}} |Ax|; \\ \sigma_i(A) &= \sup_{\dim(E)=i} \inf_{x\in\mathbb{S}(E)} |Ax|, \end{aligned} \tag{16}$$

where $\mathbb{S}(E)$ denotes the unit sphere in the subspace E. See Vershynin [53] for the details.

Definition 4.3. The condition number of a matrix A is defined as $\kappa(A) = \frac{\sigma_1(A)}{\sigma_n(A)}$.

The condition number measures how far a matrix A is from an isometry: indeed, when A is an isometry, we have $\kappa(A) = 1$, and if A stretches the space in some direction then $\kappa(A)$ is large. The parameter $\kappa(A)$ is directly involved in the speed of various algorithms for solving systems of linear equations, and it is important in many applications to know that a certain random matrix model has small enough condition number with high probability.

4.1 Norm of a sub-Gaussian random matrix

We start by studying the norm of a class of random matrices.

Theorem 4.4 (Norm of matrices with independent mean zero sub-Gaussian entries)**.** *Let A be an $N\times$ n random matrix whose entries a_{ij} are independent, mean zero, sub-Gaussian random variables. Suppose $N \geq n$. Then, for any $t > 0$ we have*

$$\|A\| \leq CK(\sqrt{N} + t) \tag{17}$$

with probability at least $1 - 2e^{-t^2}$. Here $K = \max_{i,j} \|a_{ij}\|_{\psi_2}$ and $C > 0$ is some constant.

Proof. According to Theorem 5, let $\epsilon = \frac{1}{4}$, we can find an ϵ-net $\mathcal{N} \subset \mathbb{S}^{n-1}$ and an ϵ-net $\mathcal{M} \subset \mathbb{S}^{N-1}$ where

$$\#N \leq \left(\frac{2+\epsilon}{\epsilon}\right)^n, \ \#M \leq \left(\frac{2+\epsilon}{\epsilon}\right)^N. \tag{18}$$

Note that

$$\sup_{x\in\mathbb{S}^{n-1}, y\in\mathbb{S}^{N-1}} \langle Ax, y\rangle \geq \sup_{x\in\mathcal{N}, y\in\mathcal{M}} \langle Ax, y\rangle,$$

since the supremum over a larger set is larger than the supremum over a smaller set. However,

$$\sup_{x\in\mathbb{S}^{n-1}, y\in\mathbb{S}^{N-1}} \langle Ax, y\rangle \cdot (1-2\epsilon) \leq \sup_{x\in\mathcal{N}, y\in\mathcal{M}} \langle Ax, y\rangle. \tag{19}$$

Indeed, suppose $x \in \mathbb{S}^{n-1}, \tilde{x} \in \mathcal{N}$ so that $|x-\tilde{x}| \le \epsilon$, and $y \in \mathbb{S}^{N-1}, \tilde{y} \in \mathcal{M}$ so that $|y-\tilde{y}| \le \epsilon$. We have

$$\begin{aligned} |\langle Ax, y\rangle - \langle A\tilde{x}, \tilde{y}\rangle| &= |\langle Ax, y\rangle - \langle Ax, \tilde{y}\rangle + \langle Ax, \tilde{y}\rangle - \langle A\tilde{x}, \tilde{y}\rangle| \\ &\le |\langle Ax, y-\tilde{y}\rangle| + |\langle A(x-\tilde{x}), \tilde{y}\rangle| \\ &\le 2\|A\| \cdot \epsilon \\ &= 2\varepsilon \cdot \sup_{x\in\mathbb{S}^{n-1}, y\in\mathbb{S}^{N-1}} \langle Ax, y\rangle, \end{aligned} \tag{20}$$

which implies (19).

Going back to our aim, we have

$$\begin{aligned} \mathbb{P}\left(\|A\| \ge C\cdot(\sqrt{N}+t)\right) &= \mathbb{P}\left(\sup_{x\in\mathbb{S}^{n-1}, y\in\mathbb{S}^{N-1}} \langle Ax, y\rangle \ge C\cdot(\sqrt{N}+t)\right) \\ &\le \mathbb{P}\left(\sup_{x\in\mathcal{N}, y\in\mathcal{M}} \langle Ax, y\rangle \ge \frac{C}{2}\cdot(\sqrt{N}+t)\right) \\ &= \mathbb{P}\left(\bigcup_{x\in\mathcal{N}, y\in\mathcal{M}} \{\langle Ax, y\rangle \ge \tilde{C}\cdot(\sqrt{N}+t)\}\right) \\ &\le \sum_{x\in\mathcal{N}, y\in\mathcal{M}} \mathbb{P}\left(\langle Ax, y\rangle \ge \tilde{C}\cdot(\sqrt{N}+t)\right) \\ &\le \sup_{x\in\mathcal{N}, y\in\mathcal{M}} \mathbb{P}\left(\langle Ax, y\rangle \ge \tilde{C}\cdot(\sqrt{N}+t)\right)\cdot 9^{n+N}. \end{aligned}$$

The first inequality holds because of the inequality (19) and plugging $1-2\epsilon = \frac{1}{2}$.

It remains to bound the term $\mathbb{P}\left(\langle Ax, y\rangle \ge \tilde{C}\cdot(\sqrt{N}+t)\right)$ for fixed $x \in S^{n-1}$, $y \in S^{N-1}$. Note that

$$\mathbb{P}\left(\langle Ax, y\rangle \ge \tilde{C}\cdot(\sqrt{N}+t)\right) = \mathbb{P}\left(\sum_{ij} a_{ij}\cdot x_i y_j \ge \tilde{C}\cdot(\sqrt{N}+t)\right). \tag{21}$$

We claim that $\sum_{ij} a_{ij}\cdot x_i y_j$ is sub-Gaussian. Indeed,

$$\begin{aligned} \left|\sum_{ij} a_{ij}\cdot \|x_i y_j\|_{\psi_2}^2\right| &\le C'\sum_{ij} |x_i y_j|^2 \cdot \|a_{ij}\|_{\psi_2}^2 \\ &\le C'\sum_{ij} x_i^2 y_j^2 \\ &= C'\left(\sum_{i=1}^n x_i^2\right)\left(\sum_{j=1}^N y_j^2\right) = C'. \end{aligned} \tag{22}$$

The first inequality is because of Hoeffding inequality and the last equality is because x, y are unit vectors. As a result, we have

$$\mathbb{P}\left(\langle Ax, y\rangle \ge \tilde{C}\cdot(\sqrt{N}+t)\right) \le 2e^{-C''\cdot(\sqrt{N}+t)^2}. \tag{23}$$

Combining the previous results we have

$$\mathbb{P}\left(\|A\| \geq C(\sqrt{N}+t)\right) \leq 9^{n+N} \cdot 2e^{-C''\cdot(\sqrt{N}+t)^2} \leq 2e^{-t^2}, \tag{24}$$

provided that $C > 0$ is selected appropriately. □

Corollary 4.5. *Under the assumptions of Theorem 3.36 we have $\|A\| \leq C\sqrt{N}$ with probability at least $1 - e^{-N}$ (select $t = \sqrt{N}$). This means that $\mathbb{E}\|A\| \leq C\sqrt{N}$.*

Corollary 4.6. *A is a $n \times n$ symmetric random matrix with the upper corner entries a_{ij} being independent mean zero and K-sub-Gaussian. Then for for all $t \geq 0$, we have*

$$\|A\| \leq CK(\sqrt{N}+t)$$

with probability at least $1 - 4e^{-t^2}$.

Proof. Home work! □

4.2 Two-sided bounds for intermediate singular values of tall enough random matrices

Recall the following notion:

Definition 4.7. $X \in \mathbb{R}^n$ is called an isotropic random vector in $\mathbb{R}^n$ if $\mathbb{E}X = 0$, and for all $\theta \in S^{n-1}$,

$$\mathbb{E}\langle X, \theta\rangle^2 = 1.$$

Note some similarity of this notion and the notion of sub-Gaussian random vectors – the vectors for which $\langle X, \theta\rangle$ is sub-Gaussian for every $\theta \in \mathbb{S}^{n-1}$.

Theorem 4.8 (Two-sided bound on sub-Gaussian matrices)**.** *Let A be an $N \times n$ matrix whose rows A_i are independent, mean zero, sub-gaussian isotropic random vectors in $\mathbb{R}^n$. Then for any $t \geq 0$ we have*

$$\sqrt{N} - CK^2(\sqrt{n}+t) \leq \sigma_n(A) \leq \sigma_1(A) \leq \sqrt{N} + CK^2(\sqrt{n}+t) \tag{25}$$

with probability at least $1 - 2e^{-t^2}$. Here $K = \max_i \|A_i\|_{\psi_2}$.

Theorem 4.8 is stronger than Theorem 3.36 in the following ways:

- it is more general: for instance, it includes projections composed with matrices from Theorem 3.36, random matrices whose rows are independent and sampled from the unit sphere, and much more.
- it is a two-sided bound for all singular values (rather than just an upper bound for only $\sigma_1(A)$.)

- If $N \gg n$, the bound is more precise, recovering the constant 1 in front of $\sqrt{N}$. Note also that the taller the matrix, the lesser is the impact of the sub-Gaussian constant.

However, note that Theorem 4.8 only becomes applicable when $N \geq Cn$ for an appropriate $C > 0$ that only depends on $K > 0$. This is very important to note.

We start the proof by pointing out

Claim 4.9. *Let A be an $N \times n$ matrix and $\delta > 0$. Suppose that*

$$\|A^\top A - I_n\| \leq \max(\delta, \delta^2). \tag{26}$$

Then

$$(1-\delta)|x| \leq |Ax| \leq (1+\delta)|x| \quad \text{for all } x \in \mathbb{R}^n. \tag{27}$$

The proof of this elementary fact is left as a homework.

Therefore, to establish Theorem 4.8, it is enough to prove that

$$\|\frac{1}{N} A^T A - I_n\| \leq K^2 \max(\delta, \delta^2)$$

with high probability, where $\delta = C\left(\sqrt{\frac{n}{N}} + \frac{t}{\sqrt{N}}\right)$. The proof will be done via the epsilon-net argument. As a first step, we shall show the point-wise bound, which relies on the Bernstein inequality which we proved before.

Lemma 4.10. *Let A be as in Theorem 4.8. Fix $s > 0$ and $x \in \mathbb{S}^{n-1}$. Then*

$$\mathbb{P}\left\{\left|\frac{1}{N}|Ax|^2 - 1\right| \geq s\right\} \leq 2e^{-CN\min\{\frac{s^2}{K^4}, \frac{s}{K^2}\}}. \tag{28}$$

Proof. Consider the random vector

$$Ax = \begin{pmatrix} \langle A^T e_1, x\rangle \\ \langle A^T e_2, x\rangle \\ \vdots \\ \langle A^T e_N, x\rangle \end{pmatrix}. \tag{29}$$

Note that the coordinates of Ax are independent by assumption. Also, by assumption, since $x \in \mathbb{S}^{n-1}$, we have $\|\langle A^T e_i, x\rangle\|_{\psi_2} \leq K$. Now, isotropicity implies that $\mathbb{E}\langle A^T e_i, x\rangle^2 = 1$. Thus, $\frac{1}{N}|Ax|^2 - 1$ is a sum of mean-zero sub-exponential independent random variables with sub-Gaussian norms bounded by $\frac{K^2}{N}$. An application of Bernstein's inequality (Theorem 3.33) thus finishes the proof. □

Proof of Theorem 4.8. Recall that there exists a 1/4-net $\mathcal{N} \subset \mathbb{S}^{n-1}$ with $\#\mathcal{N} \leq 9^n$ such that for any $x \in \mathbb{S}^{n-1}$ there exists a $y \in \mathcal{N}$ such that $|x - y| \leq \frac{1}{4}$, and therefore,

$$\sup_{x \in \mathbb{S}^{n-1}} \| \left(\frac{1}{N} A^\top A - I_n \right) x \| \leq 2 \max_{x \in \mathcal{N}} \left| \frac{1}{N} |Ax|^2 - 1 \right|. \tag{30}$$

The above conclusion is obtained in a similar fashion to the argument in the proof of the Theorem 4.4. Therefore, using the union bound as before, we see:

$$\begin{aligned} \mathbb{P}\left\{ \left\| \frac{1}{N} A^T A - I_n \right\| \geq K^2 \max(\delta, \delta^2) \right\} &\leq \mathbb{P}\left\{ \max_{x \in \mathcal{N}} \left| \frac{1}{N} \|Ax\|^2 - 1 \right| \geq 2K^2 \max(\delta, \delta^2) \right\} \\ &\leq 9^n \cdot \sup_{x \in \mathbb{S}^{n-1}} \mathbb{P}\left\{ \left| \frac{1}{N} \|Ax\|^2 - 1 \right| \geq CK^2 \max(\delta, \delta^2) \right\} \\ &\leq 2e^{-Ct^2}, \end{aligned} \tag{31}$$

where $\delta = C\left(\sqrt{\frac{n}{N}} + \frac{t}{\sqrt{N}} \right)$. The last passage follows from plugging (31) into Lemma 4.10. The result now follows in view of the Claim 4.9. □

4.3 Matrix Bernstein Inequality

Theorem 4.11 (Matrix Bernstein Inequality)**.** *Let $X_1, X_2, \cdots, X_N$ be independent mean-zero $n \times n$ positive definite symmetric random matrices. Suppose $\|X_i\| \leq K$ for all $i = 1, 2, \ldots, n$ almost surely (i.e. with probability 1). Then, for all $t \geq 0$,*

$$\mathbb{P}\left\{ \left\| \sum_{i=1}^N X_i \right\| \geq t \right\} \leq 2n \cdot e^{-\frac{t^2/2}{\sigma^2 + Kt/3}}, \tag{32}$$

where $\sigma^2 = \| \sum_{i=1}^N \mathbb{E} X_i^2 \|$. Equivalently,

$$\mathbb{P}\left\{ \left\| \sum_{i=1}^N X_i \right\| \geq t \right\} \leq 2n \cdot e^{-C \min\{\frac{t^2}{\sigma^2}, \frac{t}{K}\}}. \tag{33}$$

Note that $\| \sum_{i=1}^N X_i \|$ is the norm of a random matrix and $\| \sum_{i=1}^N \mathbb{E} X_i^2 \|$ is a norm of a deterministic matrix.

Is it possible to deduce Corollary 4.6 from Theorem 4.11?

Suppose that

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix}$$

and we decompose A into $N - Cn^2$ matrices where

$$X_1 = \begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, \; X_2 = \begin{pmatrix} 0 & a_{12} & 0 & \cdots & 0 \\ a_{21} & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

$$X_3 = \begin{pmatrix} 0 & 0 & a_{13} & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ a_{31} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, \; \ldots$$

Then, the statement of Corollary 4.6 follows:

$$\mathbb{P}\{\|A\| \leq (\sqrt{n} + t)K\} \geq 1 - 2e^{-Ct^2}. \tag{34}$$

We leave as a homework exercise to check the above consideration.

Theorem 4.12 (Lieb's Inequality)**.** *Let H be an $n \times n$ symmetric matrix. Consider $f(X) = \mathrm{tr}(e^{H+\log X})$ to be a function of the symmetric positive definite matrix X. Then f is a concave on this space:*

$$\mathrm{tr}(e^{H+\log \frac{X+Y}{2}}) \leq \frac{1}{2}\left(\mathrm{tr}(e^{H+\log X}) + \mathrm{tr}(e^{H+\log Y})\right) \tag{35}$$

Recall Jensen's Inequality (which works also for matrices): if f in concave, then $\mathbb{E}f(X) \leq f(\mathbb{E}X)$. As a result, we get

Corollary 4.13 (Lieb's Inequality + Jensen's Inequality)**.** *Let H be a fixed $n \times n$ symmetric matrix and Z be a random $n \times n$ symmetric matrix. Then we have*

$$\mathbb{E}\left(tr\left(e^{H+Z}\right)\right) \leq tr\left(e^{H+\log \mathbb{E}e^{Z}}\right) \tag{36}$$

We are now ready to prove Theorem 4.11.

Proof of Theorem 4.11. Denote $S = \sum_{i=1}^{N} X_i$, then

$$\|S\| = \max_{i=1,\cdots,N} |\lambda_i S| = \max\left\{\lambda_{\max}(S), \lambda_{\max}(-S)\right\} \tag{37}$$

Applying Chernoff trick, for all $\lambda \in \mathbb{R}$,

$$\begin{aligned} \mathbb{P}\{\lambda_{\max}(S) \geq t\} \; &= \mathbb{P}\{e^{\lambda\lambda_{\max}(S)} \geq e^{\lambda t}\} \\ &\leq e^{-\lambda t}\mathbb{E}e^{\lambda\lambda_{\max}(S)} \end{aligned} \tag{38}$$

Using (37), we have $P\{\|S\| \geq t\} \leq 2nP\{\lambda_{\max}(S) \geq t\}$.

Our goal is

$$\mathbb{E}\, e^{\lambda \lambda_{\max}(S)} = \mathbb{E}\, \lambda_{\max}(e^{\lambda S}) \le \mathbb{E}\left(tr(e^{\lambda S})\right). \tag{39}$$

The first equation is by definition 1.5 of functions on matrices, and the following inequality is because for all non-negative definite matrices A with eigenvalues $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n \ge 0$, we have $\lambda_1 \le \lambda_1 + \cdots + \lambda_n = tr(A)$.

Applying Theorem 4.12, we get

$$\begin{aligned}
\mathbb{E}\left(tr(e^{\lambda S})\right) &= \mathbb{E}\left(tr(e^{\sum_{i=1}^{N-1} \lambda X_i + \lambda X_N})\right) \\
&\le \mathbb{E}\left(tr(e^{\sum_{i=1}^{N-1} \lambda X_i + \log \mathbb{E} e^{\lambda X_N}})\right) \qquad \text{(applying } N \text{ times on } X_{N-1}, X_{N-2}, \dots) \\
&\le tr\left(e^{\sum_{i=1}^{N} \log \mathbb{E} e^{\lambda X_i}}\right) \\
&= tr\left(e^{\log \prod_{i=1}^{N} \mathbb{E} e^{\lambda X_i}}\right) \\
&= tr\left(\prod_{i=1}^{N} \mathbb{E} e^{\lambda X_i}\right).
\end{aligned}$$

The second inequality here follows from *conditioning*: considering the random matrix $X = \lambda X_N$ and the fixed realization of the matrix $H = \sum_{i=1}^{N-1} \lambda X_i$, and then integrating the expectation.

Now, all that remains is to bound $\mathbb{E}\, e^{\lambda X_i}$.

Lemma 4.14 (Homework)**.** *Let X be an $n \times n$ symmetric random matrix, $\mathbb{E}\, X = 0$. Suppose that $\mathbb{E}\, \|X\| \le K > 0$ almost surely, then*

$$\mathbb{E}\, e^{\lambda X} \le e^{g(\lambda)\, \mathbb{E}\, X^2} \tag{40}$$

where $g(\lambda) = \frac{\lambda^2/2}{1 - |\lambda| K/3}$ for $|\lambda| \le 3/K$.

Remark 4.15. *Note that the inequality is in the matrix sense (that is, $A \ge 0$ if for all $x \in \mathbb{R}^n$, one has $\langle Ax, x \rangle \ge 0$).*

We will leave the remaining proof of Theorem 4.11 to homework. □

4.4 Non-asymptotic bounds for the smallest singular value of random matrices

4.4.1 General discussion about the smallest singular value of a random matrix

Theorem 4.8 implies in particular that for $N \times n$ matrices with sub-Gaussian independent entries a_{ij} such that $\mathbb{E}\, a_{ij} = 0$, and $\mathbb{E}\, a_{ij}^2 = 1$, with probability $\ge 1 - e^{ct^2}$, we have

$$\sqrt{N} - cK(\sqrt{n} - t) \le \sigma_n(A) \le \cdots \le \sigma_1(A) \le \sqrt{N} + cK(\sqrt{n} + t).$$

Keep in mind that this is only meaningful when the matrix is tall enough, that is, $N \geq CKn$. We discussed that in order to have the upper-bound on $\sigma_1(A)$ with high probability for a random matrix A whose entries are independent, one really needs strong assumptions such as sub-Gaussian entries, and even to guarantee the bound of order $\sqrt{N}$ on average for $\sigma_1(A)$, the boundedness of $\mathbb{E}\, a_{ij}^4$ is required (see e.g. Litvak, Spektor [24]). In sharp contrast, it turns out that the sub-Gaussian assumption is not necessary for bounding the smallest singular value

$$\sigma_n(A) = \inf_{x \in \mathbb{S}^{n-1}} |Ax|$$

from below, and in fact, much weaker assumptions on the matrix suffice! This phenomenon was discovered in a series of works of Tikhomirov [49], [50], followed by Rebrova, Tikhomirov [32], Livshyts [25], Guedon, Litvak, Tatarko [12], Livshyts, Tikhomirov, Vershynin [26], and others.

Let us investigate how the bound of the type

$$\sigma_n(A) \geq \clubsuit, \tag{41}$$

can be proven (on average or with high probability), and what might be required of the random matrix A for this. For a square $n \times n$ matrix A, the condition $\sigma_n(A) = 0$ is equivalent to A being non-invertible. Therefore, in the case $N = n$, the condition (41) means that the matrix A is "well invertible", and having such an information about a random matrix could be valuable for applications in various situations.

Let us start by making a naive attempt to use the net argument to bound $\sigma_n(A)$ from below. Let $\mathcal{N} \subset S^{n-1}$ be an ε-net of size $\left(\frac{3}{\varepsilon}\right)^n$. Then for all $x \in S^{n-1}$, there exits $y \in \mathcal{N}$ such that $|x - y| \leq \varepsilon$. That implies that $|A(x-y)| \leq \|A\| \cdot \varepsilon$. By triangle inequality, $|Ax| \geq |Ay| - \|A\| \cdot \varepsilon$. Taking infimum on both sides, we get

$$\inf_{x \in S^{n-1}} |Ax| \geq \inf_{y \in \mathcal{N}} |Ay| - \|A\| \cdot \varepsilon.$$

Therefore,

$$\begin{aligned}
\mathbb{P}\left(\sigma_n(A) \leq \clubsuit\right) &= \mathbb{P}\left(\inf_{x \in \mathbb{S}^{n-1}} |Ax| \leq \clubsuit\right) \leq \mathbb{P}\left(\inf_{y \in \mathcal{N}} |Ay| \leq \clubsuit + \|A\| \cdot \varepsilon\right) \\
&= \mathbb{P}\left(\bigcup_{y \in \mathcal{N}} \left\{|Ay| \leq \clubsuit + \|A\| \cdot \varepsilon\right\}\right) \\
&\leq \#\mathcal{N} \cdot \sup_{y \in \mathcal{N}} \mathbb{P}\left(|Ay| \leq \clubsuit + \|A\| \cdot \varepsilon\right).
\end{aligned}$$

In case the assumptions on the random matrix allow us to have a good control of $\|A\|$ then this is a promising start! However, as we mentioned earlier, one should not need to make such strong assumptions in principle. Thus we are going to apply Theorem 2.23 (about a more involved net argument) instead, in order to aim for a more general result. Theorem 2.23 was valid for all deterministic matrices, so we point out the following Corollary for random matrices:

Corollary 4.16. *Suppose A is any random matrix, with $\mathbb{E}\,\|A\|_{HS}^2 < \infty$. Then for any $\varepsilon > 0$, there is a net $\mathcal{N} \subset \frac{3}{2}\mathbf{B}_2^n \setminus \frac{1}{2}\mathbf{B}_2^n$ with $\#\mathcal{N} \le (C/\varepsilon)^n$ such that with probability at least 0.9 we have for all $x \in \mathbb{S}^{n-1}$ there is some $y \in \mathcal{N}$ such that*

$$|A(x-y)| \le \sqrt{10}\varepsilon \cdot \frac{\sqrt{\mathbb{E}\,\|A\|_{HS}^2}}{\sqrt{n}}.$$

Proof. Take the net $\mathcal{N}$ from Theorem 2.23. Then for all $x \in \mathbb{S}^{n-1}$ there is some $y \in \mathcal{N}$ such that

$$|A(x-y)| \le \varepsilon \cdot \frac{\|A\|_{HS}}{\sqrt{n}}.$$

By Markov's inequality,

$$\mathbb{P}\Big(\|A\|_{HS}^2 \le 10\,\mathbb{E}\,\|A\|_{HS}^2\Big) = 1 - \mathbb{P}\Big(\|A\|_{HS}^2 \ge 10\,\mathbb{E}\,\|A\|_{HS}^2\Big) \ge 1 - 0.1 = 0.9.$$

This implies that with probability at least 0.9, for all $x \in \mathbb{S}^{n-1}$ there is some $y \in \mathcal{N}$ such that

$$|A(x-y)| \le \varepsilon \cdot \frac{\sqrt{10\,\mathbb{E}\,\|A\|_{HS}^2}}{\sqrt{n}}.$$

□

4.4.2 Small ball (or anti-concentration) assumption and the tensorization lemma

In order to apply any kind of net argument, we need to be able to upper bound $\mathbb{P}(|Ay| \le \clubsuit)$ for a fixed $y \in \frac{3}{2}\mathbf{B}_2^n \setminus \frac{1}{2}\mathbf{B}_2^n$. To this end, we shall now have a discussion about a small ball (or anti-concentration) assumption for a random variable and for a random vector.

Definition 4.17. Suppose ξ is a random variable. We say that it satisfies a small ball (or an anti-concentration) assumption if for any $z \in \mathbb{R}$, $\mathbb{P}(|\xi - z| \le a) \le b$ for some $a > 0$, $b \in (0,1)$.

In other words, ξ does not concentrate around some point z ("small ball") with too high probability. Here are some examples:

- if ξ has a bounded density f (say, bounded by some constant K), then for any interval I with length $2a$,
$$\mathbb{P}(\xi \in I) = \int_I f \le K|I| = 2aK.$$
One may let $a = \frac{1}{4K}$, $b = \frac{1}{2}$, and then for any $z \in \mathbb{R}$, $\mathbb{P}(|\xi - z| \le a) \le b$.

- Consider the symmetric Bernoulli

$$\xi = \begin{cases} 1, \text{ with probability } \frac{1}{2}, \\ -1, \text{ with probability } \frac{1}{2}. \end{cases}$$

  Then for any $z \in \mathbb{R}$, $\mathbb{P}(|\xi - z| \leq 0.9) \leq \frac{1}{2}$.

- If ξ is a sub-Gaussian (in particular, if it is bounded), then it satisfies small ball estimate for some a, b depending on K (this is left as a home work).

We shall rely heavily on the following Lemma which appears in Rudelson, Vershynin [35].

Lemma 4.18 (Tensorization Lemma)**.** *Let $\xi_1, \xi_2, \ldots, \xi_n$ be independent non-negative random variables. Fix $\varepsilon_0 > 0$, $K > 0$, and suppose that for any $\varepsilon \geq \varepsilon_0$,*

$$\mathbb{P}(\xi_k \leq \varepsilon) \leq K\varepsilon, k = 1, 2, \ldots, n.$$

Then for some constant $C > 0$

$$\mathbb{P}\left(\sum_{j=1}^{n} \xi_j^2 \leq \varepsilon^2 n\right) \leq (CK\varepsilon)^n.$$

In other words, the length of a random vector with independent coordinates satisfying small ball satisfies small ball estimate.

Proof. Assume $\varepsilon \geq \varepsilon_0$. Using Markov's inequality and the independence of ξ_j's we have:

$$\begin{aligned} \mathbb{P}\left(\sum_{j=1}^{n} \xi_j^2 \leq \varepsilon^2 n\right) &= \mathbb{P}\left(\exp\left\{-\frac{1}{\varepsilon^2}\sum_{j=1}^{n} \xi_j^2\right\} \geq e^{-n}\right) \\ &\leq e^n \cdot \mathbb{E}\exp\left\{-\frac{1}{\varepsilon^2}\sum_{j=1}^{n} \xi_j^2\right\} \\ &= e^n \cdot \prod_{j=1}^{n} \mathbb{E}\exp\left\{-\frac{1}{\varepsilon^2}\xi_j^2\right\}. \end{aligned}$$

Writing out the expectation and changing variables $t = e^{-s^2}$, we get

$$\begin{aligned} \mathbb{E}\exp\left\{-\frac{1}{\varepsilon^2}\xi_j^2\right\} &= \int_0^\infty \mathbb{P}\left(\exp\left\{-\frac{1}{\varepsilon^2}\xi_j^2\right\} \geq t\right) dt \\ &= \int_0^\infty 2se^{-s^2}\mathbb{P}\left(\xi_j \leq \varepsilon s\right) ds. \end{aligned}$$

Since $\varepsilon \geq \varepsilon_0$, if $s \geq 1$, then $\varepsilon s \geq \varepsilon_0$ and $\mathbb{P}\left(\xi_j \leq \varepsilon s\right) \leq Ks\varepsilon$. If $s \leq 1$, then

$$\mathbb{P}\left(\xi_j \leq \varepsilon s\right) \leq \mathbb{P}\left(\xi_j \leq \varepsilon\right) \leq K\varepsilon.$$

So we can break the integral into two parts and get, for some constant $C_1 > 0$:

$$\int_0^\infty 2se^{-s^2}\mathbb{P}\left(\xi_j \le \varepsilon s\right) ds \le \int_0^1 2K\varepsilon s e^{-s^2} ds + \int_1^\infty 2K\varepsilon s^2 s e^{-s^2} ds \le C_1 K\varepsilon.$$

Combining all of the above and letting $C = e \cdot C_1$, we get

$$\mathbb{P}\left(\sum_{j=1}^n \xi_j^2 \le \varepsilon^2 n\right) \le e^n \cdot (C_1 K\varepsilon)^n = (CK\varepsilon)^n.$$

□

Definition 4.19 (High-dimensional version of the small ball assumption)**.** A random vector X is called anti-concentrated if there exist $a > 0$, $b \in (0,1)$ such that for any $\theta \in \mathbb{S}^{n-1}$, $\mathbb{P}\Big(|\langle x, \theta\rangle| < a\Big) < b$.

Claim 4.20 (Home work)**.** *Let X be a random vector with independent entries. If for each $i = 1, 2, \ldots, n$, $\mathbb{P}(|X_i| < a) < b$, then there is some constant C such that for any $\theta \in \mathbb{S}^{n-1}$,*

$$\mathbb{P}\Big(|\langle x, \theta\rangle| < a\Big) < Cb.$$

The following Lemma is a slightly stronger version of Lemma 4.18, whose proof we leave as a home work.

Lemma 4.21. *Suppose A is an $N \times n$ random matrix with independent rows $A^\top e_i$, $i = 1, 2, \cdots, N$, and suppose for any $\theta \in S^{n-1}$,*

$$\mathbb{P}\Big(\big|\langle A^\top e_i, \theta\rangle\big| < a\Big) < b.$$

Then for any $x \in \mathbb{S}^{n-1}$,

$$\mathbb{P}\Big(|Ax| \le c\sqrt{N}\Big) \le e^{-c_1 N},$$

where $c, c_1 > 0$ only depend on a and b.

We point out also the following

Corollary 4.22 (Claim 4.20 combined with Lemma 4.21)**.** *Suppose A is an $N \times n$ random matrix with independent entries a_{ij}. If for all i, j and some $a > 0$, $b \in (0,1)$,*

$$\mathbb{P}(|a_{ij}| < a) < b,$$

then for any $x \in \mathbb{S}^{n-1}$,

$$\mathbb{P}\Big(|Ax| \le c\sqrt{N}\Big) \le e^{-c_1 N},$$

where $c, c_1 > 0$ only depend on a and b.

4.4.3 The smallest singular value of tall random matrices

We are finally ready to establish our first bound on the smallest singular value of a random matrix. So far, we will require the strong assumption that $N \geq Cn$. The statement below appears as Proposition 1 in Livshyts [25].

Proposition 4.23 (Tall matrices with possibly dependent columns)**.** *Let A be an $N \times n$ random matrix whose rows $A^T e_i$ are independent. Given $a > 0$ and $b \in (0,1)$, suppose that for every row $A^\top e_i$ and any $\theta \in S^{n-1}$,*

$$\mathbb{P}\Big(|\langle A^\top e_i, \theta\rangle| < a\Big) < b.$$

Also assume that

$$\mathbb{E}\,\|A\|_{HS}^2 = \sum_{i,j} \mathbb{E}\, a_{ij}^2 \leq KNn$$

for some $K > 0$. Then there exist $C_1 \geq 1$, and $c_2, c_3 > 0$ which only depend on a, b and K such that whenever $N \geq C_1 n$ then

$$\mathbb{P}\Big(\sigma_n(A) \leq c_2\sqrt{N}\Big) \leq 0.2.$$

Consequently, $\mathbb{E}\,\sigma_n(A) \geq c_3\sqrt{N}$.

Proof. Since $\mathbb{E}\,\|A\|_{HS}^2 \leq KNn$, applying Corollary 4.16 with $\varepsilon = 1/4$, we get that there is a net $\mathcal{N}$ such that with probability at least 0.9 for any $x \in \mathbb{S}^{n-1}$, there is some $y \in \mathcal{N}$ such that

$$|A(x-y)| \leq \sqrt{10}\varepsilon \frac{\sqrt{KNn}}{\sqrt{n}} = \tilde{C}\sqrt{N}.$$

We use the net argument to upper bound the probability

$$\begin{aligned}
\mathbb{P}\left(\sigma_n(A) \leq c_2\sqrt{N}\right) &= \mathbb{P}\left(\inf_{x\in\mathbb{S}^{n-1}} |Ax| \leq c_2\sqrt{N}\right) \\
&\leq \mathbb{P}\left(\inf_{y\in\mathcal{N}} |Ay| \leq \tilde{c_2}\sqrt{N}\right) + 0.1 \\
&\leq \#\mathcal{N} \cdot \sup_{x\in\mathbb{S}^{n-1}} \mathbb{P}\left(|Ay| \leq \tilde{c_2}\sqrt{N}\right) + 0.1.
\end{aligned}$$

By Corollary 4.16 we have $\#\mathcal{N} \leq e^{Cn}$ for some constant $C > 0$. By Lemma 4.21, we have

$$\sup_{x\in\mathbb{S}^{n-1}} \mathbb{P}\left(|Ay| \leq \tilde{c_2}\sqrt{N}\right) \leq e^{-c_1 N}.$$

Hence if $N \geq C_1 n$ for $C_1 \geq 1$ large enough, we are able to conclude that

$$\mathbb{P}\left(\sigma_n(A) \leq c_2\sqrt{N}\right) \leq e^{cn} \cdot e^{-c_1 N} + 0.1 \leq 0.2.$$

□

Proposition 4.23 should be compared with Theorem 4.8: while it only discusses the smallest singular value, the assumptions on the matrix are a lot less demanding. Keep in mind also that Theorem 4.8 implies sharpness of the estimate in the Proposition 4.23 at least in many important situations. In both cases the assumption $N \geq Cn$ is crucial. Soon we will see that the situation is quite different for square matrices, and the related results are far more complicated.

Let us formulate a corollary of Proposition 4.23 combined with Claim 4.20:

Corollary 4.24. *Let A be an $N \times n$ random matrix with the following properties:*

1. *All entries are independent;*
2. *The entries are Uniformly Anti-Concentrated (UAC), that is, there exist $a > 0$ and $b \in (0,1)$ such that $\mathbb{P}(|a_{ij}| \leq a) \leq b$;*
3. *$\mathbb{E}\,\|A\|_{HS}^2 \leq K \cdot Nn$.*

If $N \geq Cn$ for an appropriate $C \geq 1$ that depends on a and b, then $\mathbb{E}\,\sigma_n(A) \geq C_1\sqrt{N}$.

What if we want this estimate to hold with high probability rather than on average, that is, would it be possible to prove

$$\mathbb{P}(\sigma_n(A) \leq C\sqrt{N}) \leq e^{-GN}$$

in place of

$$\mathbb{P}(\sigma_n(A) \leq C\sqrt{N}) \leq 0.1?$$

Turns out, this is possible! Analyzing the proof, one may note that the 0.1 error in the probability estimate came from the fact that our net only works with constant probability (as was obtained in Corollary 4.16 using Markov's inequality). But it is possible to construct a net with similar properties which would in fact work well with high probability! We discuss this construction below.

4.4.4 A net construction which works with high probability for matrices with independent columns

Theorem 4.25 (Livshyts [25])**.** *Fix $n \in N$, $\epsilon \in (0, 0.1)$. Then there exists a net $\mathcal{N} \subset \frac{3}{2}\mathbf{B}_2^n \setminus \frac{1}{2}\mathbf{B}_2^n$ with $\#\mathcal{N} \leq (\frac{50}{\epsilon})^n$ such that whenever we consider an $N \times n$ random matrix A with independent columns then with probability at least $1 - e^{-5n}$, the following holds: for all $x \in \mathbb{S}^{n-1}$, there exists $y \in \mathcal{N}$ such that $|A(x-y)| \leq \frac{10}{\sqrt{n}} \cdot \sqrt{\mathbb{E}\|A\|_{HS}^2}$*

One should compare Theorem 4.25 to Corollary 4.16. On one hand, the probability in the new result is much larger: $1 - e^{-5n} >> 0.9$ when $n \to \infty$. On the other hand, we pay by requiring the matrix to have independent columns.

Before we discuss the proof of Theorem 4.25, we state the following improvement of Corollary 4.24:

Corollary 4.26. *Let A be an $N \times n$ random matrix with the following properties*

1. *All entries are independent*
2. *The entries are Uniformly Anti-Concentrated (UAC).*
   There exists $a > 0, b \in (0,1)$ such that $P(\|a_{ij}\| \leq a) \leq b$
3. $\mathbb{E}\,\|A\|_{HS}^2 \leq K \cdot Nn$

This implies that if $N \geq Cn$, then $\mathbb{P}(b_n)(A) \leq C\sqrt{N}) \leq e^{-Cn}$

Corollary 4.26 follows from Theorem 4.25 verbatim as Corollary 4.24 follows from Corollary 4.16, so we leave the details to the reader.

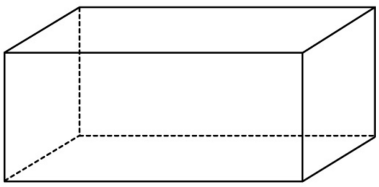
4.5 Proof of Theorem 4.25.

Recall that earlier, we constructed a net on the sphere which was a subset of the cubic lattice net. Here, inspired by Rebrova and Tikhimorov [32], we will instead cover the sphere by coordinate boxes, that is, sets of the form $[a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$.

Step 1: Cover the sphere by coordinate boxes and consider the associated random rounding. Up to translation, a coordinate box P is determined by the lengths of its sides $\alpha_1, \alpha_2, \ldots, \alpha_n \in [0,1]$, or in other words, a coordinate box

$$P = [0, \alpha_1] \times ... \times [0, \alpha_n]$$

is determined by the vector $\alpha = (\alpha_1, \ldots \alpha_n)$.



Fix also a parameter $\kappa \geq 1$.
Consider

$$\Omega_\kappa = \{\alpha \in \mathbb{R}^n : \alpha_i \in [0,1] \ \forall i \text{ and } \prod_{i=1}^{n} \alpha_i \geq \kappa^{-n}\}$$

This is the "set of admissible parallelepipeds", a set of coordinate boxes that fit inside the cube with volume no smaller than κ.

Lemma 4.27. *Fix $\kappa \geq 1, \epsilon \in (0, 0.1), \alpha \in \Omega_\kappa$, and consider $P_\alpha = [0, \alpha_1] \times \cdots \times [0, \alpha_n]$. Then, there exist $y_1, \ldots, y_m \in 1.4B_2^n$ so that $\mathbb{S}^{n-1} \subset \bigcup_{i=1}^m (y_i + \frac{\epsilon}{\sqrt{n}} P)$ and $m \leq (\frac{10}{\epsilon}\kappa)^n$.*

This fact is proved by considering

$$\mathcal{F}_\alpha = \left(\frac{\epsilon\alpha_1}{\sqrt{n}}\mathbb{Z} \times ... \times \frac{\epsilon\alpha_n}{\sqrt{n}}\mathbb{Z} \right) \cap \left(\frac{3}{2}B_2^n \setminus \frac{1}{2}B_2^n \right), \tag{42}$$

and letting y_i to be the centers of the lattice boxes forming $\mathcal{F}_\alpha$. The details are left as a home work.

Remark 4.28. *Compare with the corresponding statement about covering the sphere with the cubes: the size of the net is κ^n larger.*

Using the net from the Lemma 4.27 and the random rounding construction, we deduce the following.

Lemma 4.29. *Pick an $\alpha \in \Omega_\kappa$. Let A be any $N \times n$ matrix. There exists a finite set $\mathcal{F}_\alpha \subset \frac{3}{2}\mathbf{B}_2^n \setminus \frac{1}{2}\mathbf{B}_2^n$ satisfying $\#\mathcal{F}_\alpha \leq (\frac{10\kappa}{\epsilon})^n$, such that for all $x \in \mathbb{S}^{n-1}$, there is $y \in \mathcal{F}_\alpha$ with*

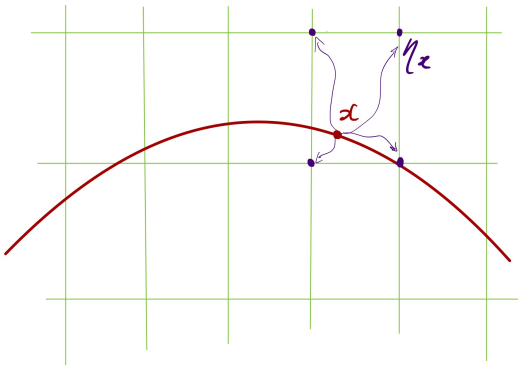
$$|A(x-y)| \leq \frac{\varepsilon}{\sqrt{n}} \sqrt{\sum_{i=1}^n \alpha_i^2 \, |Ae_i|^2}.$$

Keep in mind here that Ae_i are columns of A and $|Ae_i|$ are lengths of columns. Recall that we earlier proved this when all $\alpha_1 = \alpha_2 = \ldots = \alpha_n$, in which case we get $\sum_{i=1}^n \alpha_i^2 \, |Ae_i|^2 = \|A\|_{HS}^2$.

Proof. Consider $\mathcal{F}_\alpha$ as in (42). By Lemma 4.27, the size of this net is appropriate, and we have the covering

$$\mathbb{S}^{n-1} \subset \bigcup_{i=1}^m (y_i + \frac{\epsilon}{\sqrt{n}} P).$$

Fix $x \in \mathbb{S}^{n-1}$ and consider a box P from this covering which contains x. Construct an "α-associated random rounding" η_x, that is, a random vector η_x which takes values in the vertices of P, such that the coordinates of η_x are independent and $\mathbb{E}\,\eta_x = x$.



Then, considering the expected value with respect to the randomness of the random rounding, we get, as before,

$$\begin{aligned}\mathbb{E}\,|A(\eta_x - x)|^2 &= \mathbb{E}\sum_{i=1}^{N}\langle A^T e_i, \eta_x - x\rangle^2 \\ &\leq \frac{\epsilon^2}{n}\sum_{i=1}^{n}\alpha_i^2 \cdot |Ae_i|^2,\end{aligned}$$

by the same reasoning as in the proof of Theorem 2.23, in the equation (8). Therefore, by Markov's inequality, for all $x \in \mathbb{S}^{n-1}$, there exists $y \in \mathcal{F}_\alpha$ such that

$$|A(x-y)| = \frac{\epsilon}{\sqrt{n}}\sqrt{\sum_{i=1}^{n}\alpha_i^2 \cdot |Ae_i|^2}.$$

Note that we can indeed guarantee that $y \in \mathcal{F}_\alpha$ because we know that y takes values in the vertices of one of the boxes forming the net $\mathcal{F}_\alpha$. □

This inspires us to formulate the following definition:

Definition 4.30 (A proxy for the Hilbert Schmidt Norm)**.** Let A be an $N \times n$ matrix. We define a functional that will serve as a proxy for the squared Hilbert-Schmidt norm by

$$\mathcal{B}_\kappa(A) := \min_{\alpha \in \Omega_\kappa}\sum_{i=1}^{n}\alpha_i^2 \cdot |Ae_i|^2.$$

Note that $\mathcal{B}_\kappa(A) \leq \|A\|_{HS}^2 = \sum_{i=1}^{n}|Ae_i|^2$. Therefore, using Lemma 4.29 we can extend Theorem 2.23 as follows:

Corollary 4.31. *Fix $\kappa \geq 1$ and $\epsilon \in (0, 0.1)$. Let A be an $N \times n$ matrix. Then there exists a net $\mathcal{F} \subset \frac{3}{2}B_2^n \setminus \frac{1}{2}B_2^n$ such that for all $x \in S^{n-1}$ there exists $y \in \mathcal{F}$ with*

$$|A(x-y)| \leq \frac{\epsilon}{\sqrt{n}}\sqrt{\mathcal{B}_\kappa(A)},$$

such that $\#\mathcal{F} \leq \left(\frac{10\kappa}{\epsilon}\right)^n$.

Unfortunately, however, Corollary 4.31 is useless! Indeed, the value of the vector α which gives the minimum in $\mathcal{B}_\kappa(A)$ depends on the matrix A, and therefore the net in Corollary 4.31 depends on the matrix. In order to apply this to study random matrices, we would like to consider the union of all the nets associated with the admissible boxes in Ω_κ, but there is infinitely many of them! However, it turns out that we can discretize the set Ω_κ, and only consider finitely many nets which would "represent" all the possible nets. We will then be able to construct our desired net as a union of those representative nets.

Step 2: "Nets on nets " Next, we would like to switch the quantifiers in the previous statement: in place of the net that depends on the matrix, we need to have a fixed net, which serves all matrices. For that purpose we shall consider a net on the set of admissible nets.

Lemma 4.32 (nets on nets)**.** *There exist absolute constants $C, C', C'' > 0$ such that for any $\kappa > 1$ and $\mu \in (0, \sqrt{n})$ there exists a collection $\mathcal{F} \subset \Omega_{\kappa^{1+\mu}}$ of cardinality*

$$\max\left(\left(\frac{C}{\mu}\right)^{n-1}, (C'\mu)^{\frac{C''n}{\mu^2}}\right), \tag{43}$$

such that for any $\alpha \in \Omega_\kappa$ there exists a $\beta \in \mathcal{F}$ such that for all $i = 1, ..., n$ we have $\alpha_i^2 \geq \beta_i^2$.

In particular, for any $N \times n$ matrix A, we have

$$\mathcal{B}_\kappa(A) \geq \min_{\beta \in \mathcal{F}} \sum_{i=1}^n \beta_i^2 |Ae_i|^2.$$

Proof. Consider a transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ given by

$$T\alpha = \left(..., \sqrt{\frac{\log|\frac{1}{\alpha_i}|}{n \log \kappa}}, ...\right).$$

Denote $B = B_2^n \cap \{x_i \geq 0 \, \forall i = 1, ..., n\}$. Then, by definition of Ω_κ we have

$$T\Omega_\kappa = B,$$

and

$$T^{-1}\left((1+\mu)B\right) = \Omega_{\kappa^{1+\mu}}.$$

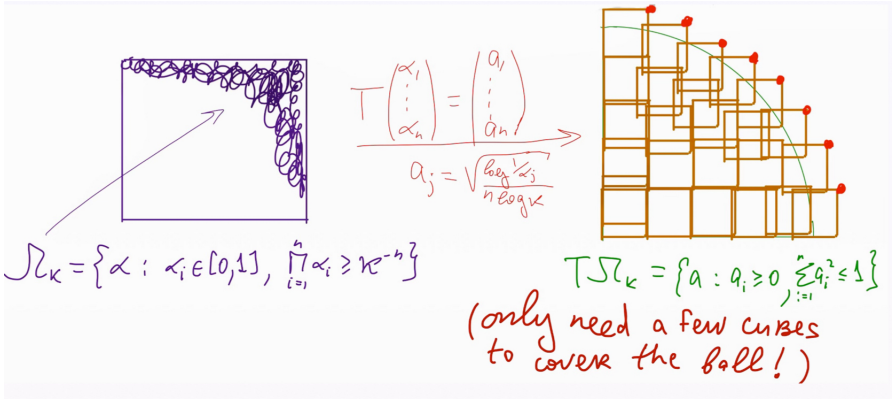
Note that this mapping is a bijection on Ω_κ as well as on $\Omega_{\kappa^{1+\mu}}$.

Consider a lattice covering $\mathcal{N}$ of the boundary of B with translates of $\frac{\mu}{\sqrt{n}} B_\infty^n$. In each cube $x + \frac{\mu}{\sqrt{n}} B_\infty^n$ from this covering, pick such a vertex $v(x)$ that for all $y \in x + \frac{\mu}{\sqrt{n}} B_\infty^n$, and for all $i = 1, ..., n$, one has $y_i \leq v(x)_i$. Define $\mathcal{S} = \{v(x) : x \in \mathcal{N}\}$. Note that $\mathcal{S} \subset (1+\mu)B$, and that

$$\#\mathcal{S} = \#\mathcal{N} \leq \min\left(\left(\frac{C}{\mu}\right)^{n-1}, (C'\mu)^{\frac{C''n}{\mu^2}}\right).$$

Note that the power $n-1$ comes from the fact that we are covering the sphere rather than the ball.

Let $\mathcal{F} = T^{-1}\mathcal{S} \subset \Omega_{\kappa^{1+\mu}}$. For every $\alpha \in \Omega_\kappa$ let $a = T\alpha \in B$. Then take the $b \in \mathcal{S} \subset (1+\mu)B$ such that $a_i^2 \leq b_i^2$; consider $\beta \in \mathcal{F}$ defined as $\beta = T^{-1}b$. Since T is coordinate-vise decreasing, we have, for all $i \in \{1, ..., n\}$, the inequality $\alpha_i^2 \geq \beta_i^2$, as desired. □



Finally, we deduce the result about a net which serves all deterministic matrices.

Theorem 4.33 (A net for deterministic matrices)**.** *Fix $n \in \mathbb{N}$, $\epsilon \in (0, \frac{1}{10})$, $\kappa \geq 1$ Consider any $\mathcal{S} \subset \mathbb{S}^{n-1}$. There exists a net $\mathcal{N} \subset \frac{3}{2}\mathbf{B}_2^n \backslash \frac{1}{2}\mathbf{B}_2^n$, such that for any $N \times n$ matrix A, the following holds: for every $x \in \mathcal{S}$ there exists $y \in \mathcal{N}$ such that*

$$|A(x-y)| \leq \frac{100}{\sqrt{n}} \sqrt{\mathcal{B}_\kappa(A)},$$

and

$$\#\mathcal{N} \leq \left(\frac{50\kappa \log \kappa}{\epsilon}\right)^n.$$

Proof. Let $\mu = 2$ and consider a net $\beta_1, \beta_2, \ldots, \beta_m$ on Ω_κ with $m \leq 5^n$, as was described in Lemma 4.32. For each β_i, we consider a box P_{β_i}, and the lattice net $\mathcal{F}_i$ generated by $\frac{\epsilon}{\sqrt{n}} P_{B_i}$.

By Lemma 4.27 applied with $\alpha = \beta_i$, for all $x \in \mathbb{S}^{n-1}$, there exists $y \in \mathcal{F}_i$ such that

$$|A(x-y)| \leq \frac{\epsilon}{\sqrt{n}} \sqrt{\sum_{j=1}^n (\beta_i^j)^2 |A_j e_j|^2}.$$

By Lemma 4.32,

$$\min_i \sum_{j=1}^n (\beta_i^j)^2 |A_j e_j|^2 \leq \mathcal{B}_\kappa(A).$$

Therefore, if we consider our net to be $\mathcal{N} = \cup_{i=1}^m \mathcal{F}_i$, the conclusion follows. □

Note that Theorem 4.33 improves upon Theorem 2.23 since $\mathcal{B}_\kappa(A) \leq \|A\|_{HS}^2$. It turns out that when the matrix A is random and it has independent columns, this improvement is really crucial, since $\mathcal{B}_\kappa(A)$ has strong large deviation properties, while $\|A\|_{HS}^2$ is only larger than a multiple of its average with constant probability (as follows from Markov's inequality). We explore these strong properties in our next step.

Step 3: Large deviation of $\mathcal{B}_\kappa$. Note that if $y_1, \dots, y_n \geq 0$ are fixed and we constrain $\prod_{i=1}^n a_i = C_1$ for some constant $C_1 > 0$, then the quantity $\sum_i a_i y_i$ is minimized precisely when $a_i = \frac{C_2}{y_i}$ for all i and some constant $C_2 > 0$. This inspires our next proof.

Lemma 4.34. *Let A be an $N \times n$ matrix with independent columns. Fix $\kappa \geq 1$. It holds that*

$$\mathbb{P}\left\{\mathcal{B}_\kappa(A) \geq 10\,\mathbb{E}\,\|A\|_{HS}^2\right\} \leq (C\kappa)^{-2n},$$

for some absolute constant $C > 0$.

Proof. For all $i \in \{1, \dots, n\}$, define the random variables $Y_i := |Ae_i|$ and

$$a_i := \sqrt{\min\left\{1, \frac{\mathbb{E}\,Y_i^2)}{Y_i^2}\right\}}.$$

Let a be a vector with coordinates a_i. We note that

$$\begin{aligned}
\mathbb{P}\left\{\mathcal{B}_\kappa(A) \geq 10\,\mathbb{E}\,\|A|\|_{HS}^2\right\} &= \mathbb{P}\left\{\min_{\alpha\in\Omega_\kappa} \sum_{i=1}^n \alpha_i^2 Y_i^2 \geq 10\,\mathbb{E}\,\|A\|_{HS}^2\right\} \\
&\leq \mathbb{P}\left\{\sum_{i=1}^n a_i^2 Y_i^2 \geq 10\,\mathbb{E}\,\|A\|_{HS}^2\right\} + \mathbb{P}\{a \notin \Omega_\kappa\} \\
&= \mathbb{P}\left\{\sum_{i=1}^n \min\left\{1, \frac{\mathbb{E}\,Y_i^2)}{Y_i^2}\right\} Y_i^2 \geq 10\,\mathbb{E}\,\|A\|_{HS}^2\right\} + \mathbb{P}\{a \notin \Omega_\kappa\} \\
&\leq \mathbb{P}\left\{\sum_{i=1}^n \mathbb{E}\,Y_i^2 \geq 10\,\mathbb{E}\,\|A\|_{HS}^2\right\} + \mathbb{P}\left\{\prod_{i=1}^n \frac{\mathbb{E}\,Y_i^2}{Y_i^2} \leq \kappa^{-2n}\right\} \\
&\leq (C\kappa)^{-2n}\,\mathbb{E}\prod_{i=1}^n \frac{Y_i^2}{\mathbb{E}\,Y_i^2} = (C\kappa)^{-2n},
\end{aligned}$$

where the last inequality follows by noting $\sum_i \mathbb{E}\,Y_i^2 = \mathbb{E}\,\|A\|_{HS}^2$ (and therefore, the first summand is zero), and by applying Markov's inequality to the second term. □

$$|A(x-y)| \leq \frac{c\sqrt{\mathbb{E}\,\|A\|_{HS}^2}}{\sqrt{n}},$$

where $c > 0$ is an absolute constant.

Proof of Theorem 4.25. The assertion follows by applying Lemma 4.34 and Theorem 4.33 with $\kappa = 5$. □

Remark 4.35 (Important remark)**.** *Suppose $\epsilon = 0.1$ and $\kappa = 3$. Then the net described in Theorem 4.25 is roughly of order 200^n whereas the usual $\epsilon-$net would be of order roughly 30^n. Note furthermore that the net from Theorem 4.25 cannot be smaller than of size 2^n,*

because each box in the net has as many as 2^n vertices, and the net is formed by the vertices of many such boxes. Often this aspect does not matter.

However, suppose $S \subset \mathbb{R}^n$ is a "small set" which can be covered by m $\epsilon-$balls, where $m \leq 1.01^n$. So S is in this sense a lot smaller than the whole sphere (which, like we said, would require 30^n balls). Then S can also be covered by parallelepipeds and one can constrict a net on S with properties like the ones in Theorem 4.25. And furthermore, for any $\gamma > 0$ one can have a net $\mathcal{F}$ near S of size $\#\mathcal{F} \leq (1+\gamma)^n m$ (recall that m is the number of $\epsilon-$balls covering S), and such that for any $x \in S$ one has $y \in \mathcal{F}$ so that with probability at least $1 - e^{-cn}$ one has

$$|A(x-y)| \leq \frac{C(\gamma)\epsilon\sqrt{\mathbb{E}\,\|A\|_{HS}^2}}{\sqrt{n}}.$$

Here the assumptions on A are the same as in Theorem 4.25. For the full statement and the details, see Livshyts [25].

At the first glance, this may seem surprising: how could one guarantee a net made out of vertices of some boxes of size less than, say 1.2^n, if there is 2^n vertices already for one box? The trick is that a lot of the vertices could be "dismissed" from this construction. This requires some further ideas and technicalities which we leave beyond the scope of this course. See the details in [25], particularly in Lemma 3.10.

4.6 The smallest singular value of square random matrices

The above concentration results for the smallest singular value of A rely upon taking the number of rows of A to be sufficiently large relative to the number of columns. Having $N \geq Cn$ allowed the term

$$\sup_{x \in \mathbb{S}^{n-1}} \mathbb{P}(|Ax| \leq \clubsuit) \leq e^{-cN}$$

compensate for the $e^{C_1 n}$ factor coming from the union bound applied over the ϵ-net. The standard ϵ-net argument no longer works if A is a square matrix, even if one imposes additional strong assumptions on the matrix – it doesn't even work for Gaussian random matrices. However, when an idea does not work, it might be a good strategy to find a way to still apply it to make at least some progress.

In what follows we describe a foundational idea of Rudelson and Vershynin about a decomposition of the sphere. Instead of using an ϵ-net to cover the whole sphere $\mathbb{S}^{n-1}$, we can construct an ϵ-net of prescribed size e^{cn} for a "small but well spread" subset $S \subseteq \mathbb{S}^{n-1}$. "Small" in this context means that the set has a small covering number, and "well spread" means that its complement is in some sense predictable and controllable using different ideas. The smallest singular value can then be analyzed by noticing that

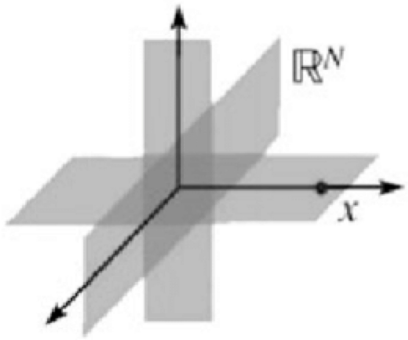
$$\mathbb{P}\left(\inf_{x \in \mathbb{S}^{n-1}} |Ax| \leq \clubsuit\right) \leq \mathbb{P}\left(\inf_{x \in S} |Ax| \leq \clubsuit\right) + \mathbb{P}\left(\inf_{x \in S^c} |Ax| \leq \clubsuit\right),$$

and considering the infima $\inf_{x \in S} |Ax|$ and $\inf_{x \in \mathbb{S}^{n-1}} |Ax|$ separately.

Definition 4.36. For all $\delta > 0$ define the set

$$\mathrm{Sparse}(\delta) := \left\{ x \in \mathbb{S}^{n-1} : \#\{i : x_i = 0\} \geq \delta n \right\}.$$

In other words, the set $\mathrm{Sparse}(\delta)$ is the intersection of $\mathbb{S}^{n-1}$ with the union of all coordinate sub-spaces of dimension δn.



The following Lemma is a crucial fact about sparse vectors: the set of sparse vectors can be covered by a relatively small number of ϵ-balls.

Lemma 4.37. *For all $\epsilon > 0$ and $\delta \in (0, 1/2)$ one has*

$$\mathrm{Sparse}(\delta) \subset \bigcup_{i=1}^{m} (y_i + \epsilon \mathbf{B}_2^n),$$

where $m \leq \epsilon^{-c_1 \delta \log \frac{1}{\delta} n + c_2}$ and $c_1, c_2 > 0$ are absolute constants.

Proof. Suppose for simplicity that δn is an integer (the proof works along similar lines either way). There are $\binom{n}{\delta n}$ sub-spaces in $\mathbb{R}^n$ of dimension δn. Note, using Stirling's formula:

$$\binom{n}{\delta n} = \frac{n!}{(\delta n)!(n-\delta n)!} = C \frac{n^n e^{-n}}{(\delta n)^{\delta n} e^{-\delta n} ((n-\delta n)^{n-\delta n} e^{-n+\delta n}} \leq e^{n \delta \log \frac{1}{\delta} + c},$$

where in the last passage one may use elementary calculus, see e.g. [58].

We deduce that

$$m \leq \left(\frac{3}{\epsilon}\right)^{\delta n} \binom{n}{\delta n} \leq \epsilon^{-\delta n} e^{H(\delta) n} = \epsilon^{-\delta n} e^{-(\delta \log \delta + (1-\delta) \log(1-\delta)) n} \leq \epsilon^{-\delta n} e^{2\delta \log(\frac{1}{\delta}) n + c},$$

which yields the conclusion. □

4.6.1 Rudelson–Vershynin decomposition of the sphere

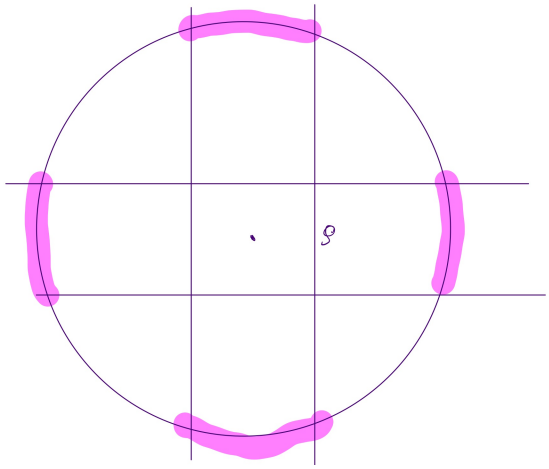
As we discussed earlier, we want to decompose the unit sphere into two disjoint subsets, one of which would have a small covering. Sparse vectors are a good candidate for this first set. However, that is not enough: the set of sparse vectors has measure zero, and removing it does not achieve much... So instead we will do a net argument on the set of vectors which are *close to sparse vectors*. These are called *compressible vectors*. Here is the formal

Definition 4.38 (the Rudelson-Vershynin decomposition of the sphere)**.** For all $\delta > 0$ and $\rho > 0$, define the *compressible vectors* to be the set

$$\begin{aligned}\mathrm{Comp}(\delta, \rho) &= \{x \in \mathbb{S}^{n-1} : d(x, \mathrm{Sparse}(\delta)) \le \rho\} \\ &= \{x \in \mathbb{S}^{n-1} : \exists\, y \in \mathrm{Sparse}(\delta) \text{ s.t. } |x - y| \le \rho\},\end{aligned}$$

and the *incompressible vectors* to be their complement

$$\mathrm{Incomp}(\delta, \rho) = \mathbb{S}^{n-1} \setminus \mathrm{Comp}(\delta, \rho).$$



Lemma 4.37 implies:

Lemma 4.39. *For all $c \in [0,1]$ and $\epsilon \in [0,1]$ there exist $\delta > 0$ and $\rho > 0$ such that*

$$\mathrm{Comp}(\delta, \rho) \subset \bigcup_{i=1}^{m} (y + \epsilon \mathbf{B}_2^n)$$

where $m \le \frac{\rho}{\epsilon} e^{cn} \le e^{c_1 n}$, with c depending on ϵ and δ, and c_1 depending on ϵ, ρ, δ.

The following claim follows from Lemma 4.39 and Remark 4.35:

Claim 4.40. *For any $c \in [0,1]$ there exist $\epsilon \in (0, \frac{1}{4})$, $\delta, \rho \in (0,1)$, and an ϵ-net $\mathcal{N} \subset \frac{3}{2}\mathbf{B}_2^n \setminus \frac{1}{2}\mathbf{B}_2^n$ with $\#\mathcal{N} \le e^{cn}$, such that the following holds. Let A be an $N \times n$ random*

matrix with independent columns. With probability at least $1 - e^{cn}$*, one has that for all* $x \in \mathrm{Comp}(\delta, \rho)$*, there exists* $y \in \mathcal{N}$ *such that*

$$\|A(x-y)\| \leq \frac{C\sqrt{\mathbb{E}\,\|A\|_{HS}^2}}{\sqrt{n}}.$$

Here $C > 0$ *depends only on* C, δ, ρ*.*

The Claim above will be useful for us when we estimate the infimum of $|Ax|$ over the set of compressible vectors. But for the time being, let us discuss how the smallest singular value of square random matrices actually behaves, and state some relevant results.

4.6.2 Survey of results regarding the smallest singular value of square random matrices

Theorem 4.41 (Rudelson-Vershynin)**.** *Let* A *be an* $n \times n$ *random matrix with entries that are iid,* K*-sub-Gaussian, mean-zero, and variance-one. For all* $\epsilon > 0$ *one has*

$$\mathbb{P}\left\{\sigma_n(A) \leq \frac{\epsilon}{\sqrt{n}}\right\} \leq C\epsilon + e^{-c_1 n},$$

where C *and* c_1 *are positive constants depending only on* K*.*

In the analogous setting for tall matrices $N \geq Cn$, we established that $\mathbb{E}\,\sigma_n(A) \geq \widetilde{c}\sqrt{N}$, which tends to infinity. Theorem 4.41 states something quite different for square matrices: the smallest singular value turns out to be of order $\frac{1}{\sqrt{n}}$ on average, and in fact this estimate is sharp (up to an absolute constant), see Rudelson, Vershynin [36], Tatarko [44]. Let us consider two key examples which demonstrate the sharpness also of the probability estimate in Theorem 4.41.

Example 4.42. *Suppose the entries of* A *are i.i.d. distributed as standard Gaussian random variables. Szarek [43] and Edelman [8] proved that for all* $\epsilon > 0$ *one has*

$$\mathbb{P}\left\{\sigma_n(A) \leq \frac{\epsilon}{\sqrt{n}}\right\} \leq \epsilon.$$

Note that this implies $\mathbb{P}\{\sigma_n(A) = 0\} = 0$*, and this phenomenon would more generally be true for any random matrix whose entries have continuous distribution, even without the mean zero assumption, see Tikhomirov [51].*

Example 4.43. *Suppose the entries* a_{ij} *are i.i.d. symmetric Bernoulli random variables with parameter* $1/2$*, i.e.* $a_{ij} \sim \mathrm{Unif}\{-1, 1\}$*. What can be said about the invertibility of* A*? One has for instance*

$$\begin{aligned}\mathbb{P}\{\sigma_n(A) = 0\} &\geq \mathbb{P}\left\{\textit{two rows are the same or two columns are the same}\right\} \\ &\geq (n^2 + o(1))2^{-n}.\end{aligned}$$

A conjecture of Erdős stated that this bound is sharp up to a polynomial error. This conjecture was essentially resolved by Tikhomirov, who proved that

$$\mathbb{P}\{\sigma_n(A) = 0\} \leq (2 + o(1))^{-n}.$$

This shows that for random matrices with independent discrete entries, one generally expects to have

$$\mathbb{P}\{\sigma_n(A) = 0\} = e^{-n}.$$

Combining the terms in Examples 1 and 2 shows that the probability estimate in Theorem 4.41 is natural.

Following the result from Theorem 4.41, Rebrova and Tikhomirov were able to relax the sub-Gaussian assumption, instead considering uniformly anti-concentrated random variables.

Theorem 4.44 (Rebrova-Tikhomirov [32])**.** *Let A be a random matrix, whose entries a_{ij} are zero-mean, unit variance, and i.i.d. If a_{ij} are uniformly anti-concentrated (UAC),* i.e., *$\mathbb{P}(a_{ij} < a) < b$ for fixed $a > 0$, $b \in (0, 1)$, then for every $\varepsilon > 0$,*

$$\mathbb{P}\left(\sigma_n(A) \leq \frac{\varepsilon}{\sqrt{n}}\right) \leq c\varepsilon + e^{-c_1 n},$$

where c, c_1 are constants that depend on a and b.

The proof technique for Theorem 4.44 uses a more clever ε-net argument than Theorem 4.41, which depends on the norm of the matrix. In particular, every assumption including $\mathbb{E}(a_{ij}) = 0$, $\mathbb{E}(a_{ij}^2) = 1$, and UAC are used in the net construction. A further generalization of the result investigates removing the assumption that the entries are all identically distributed, zero-mean, and unit variance.

Theorem 4.45 (Livshyts [25])**.** *Let A be a random matrix, whose entries a_{ij} are independent. If $\mathbb{E}\,\|A\|_{HS}^2 \leq Kn^2$, and a_{ij} are uniformly anti-concentrated (UAC),* i.e., *$\mathbb{P}(a_{ij} < a) < b$ for fixed $a > 0$, $b \in (0, 1)$, then for every $\varepsilon > 0$,*

$$\mathbb{P}\left(\sigma_n(A) \leq \frac{\varepsilon}{n}\right) \leq c\varepsilon + \frac{c_1}{\sqrt{n}}, \tag{44}$$

where c, c_1 are constants that depend on a and b. Moreover, if the rows of A are i.i.d., then

$$\mathbb{P}\left(\sigma_n(A) \leq \frac{\varepsilon}{\sqrt{n}}\right) \leq c\varepsilon + e^{-c_1 n}, \tag{45}$$

where c, c_1 are constants that depend on a and b.

Remark 4.46. *The polynomial bound from* (44)*, while easier to prove, is a strictly weaker bound than the exponential bound of* (45)*, which recovers the bound from Theorem 4.44 for the i.i.d. case. Recovering the exponential bound* (45) *requires an additional tool called the "Least Common Denominator" (LCD) of vectors. If one only cares about the limit as $n \to \infty$, this is unimportant, however, in the case of square matrices, we may care about the conservatism of this bound. The proof of the polynomial bound* (44)*, is worked out in the next subsection.*

In a future result, Livshyts, Tikhomirov, and Vershynin were able to prove the same exponential bound without the i.i.d. row assumption.

Theorem 4.47 (Livshyts, Tikhomirov, Vershynin [26]). *Let A be a random matrix, whose entries a_{ij} are independent. If $\mathbb{E}\|A\|_{HS}^2 \leq Kn^2$, and a_{ij} are uniformly anti-concentrated (UAC),* i.e., *$\mathbb{P}(a_{ij} < a) < b$ for fixed $a > 0$, $b \in (0,1)$, then for every $\varepsilon > 0$,*

$$\mathbb{P}\left(\sigma_n(A) \leq \frac{\varepsilon}{n}\right) \leq c\varepsilon + e^{-c_1 n},$$

where c, c_1 are constants that depend on a and b.

The proof of Theorem 4.47 requires the use of the "randomized LCD", similar to the LCD used in the proof of Theorem 4.45.

Arbitrary aspect ratios We briefly state some results for matrices of arbitrary aspect ratios. Consider the random matrix A, of dimension $N \times n$, with i.i.d., zero mean, and unit variance entries a_{ij}. Consider the case $N \geq n$, but possibly not $N >> n$.

Theorem 4.48 (Rudelson, Vershynin [37]). *If a_{ij} are sub-Gaussian, then for every $\varepsilon > 0$,*

$$\mathbb{P}(\sigma_n(A) \leq \varepsilon(\sqrt{N+1} - \sqrt{n})) \leq (c\varepsilon)^{N-n+1} + e^{-c_1 n},$$

for constants c, c_1.

Theorem 4.49 (Livshyts). *If a_{ij} are UAC, then for every $\varepsilon > 0$,*

$$\mathbb{P}(\sigma_n(A) \leq \varepsilon(\sqrt{N+1} - \sqrt{n})) \leq \left(c\varepsilon \log \frac{1}{\varepsilon}\right)^{N-n+1} + e^{-c_1 n},$$

for constants c, c_1 dependent on a and b.

4.7 Proof of Theorem 4.45 Part 1

Recall the statement of the theorem. Let A be a random matrix, whose entries a_{ij} are independent. If $\mathbb{E}\|A\|_{HS}^2 \leq Kn^2$, and a_{ij} are uniformly anti-concentrated (UAC), *i.e.*, $\mathbb{P}(a_{ij} < a) < b$ for fixed $a > 0$, $b \in (0,1)$, then for every $\varepsilon > 0$,

$$\mathbb{P}\left(\sigma_n(A) \leq \frac{\varepsilon}{n}\right) \leq c\varepsilon + \frac{c_1}{\sqrt{n}},$$

where c, c_1 are constants that depend on a and b.

To prove this result, we will need to use additional tools to separately handle the compressible and incompressible vectors from the Rudelson-Vershynin decomposition of the sphere from Definition 4.38. Recall the following definitions of sparse vectors, compressible vectors (close to sparse), and incompressible vectors (far from sparse). For $\delta, \rho > 0$,

$$\begin{aligned}
\text{Sparse}(\delta) &= \left\{x \in \mathbb{S}^{n-1} : \#\{i : x_i = 0\} \geq \delta n\right\}, \\
\text{Comp}(\delta, \rho) &= \{x \in \mathbb{S}^{n-1} : \exists\ y \in \text{Sparse}(\delta) \text{ s.t. } |x - y| \leq \rho\}, \\
\text{Incomp}(\delta, \rho) &= \mathbb{S}^{n-1} \setminus \text{Comp}(\delta, \rho).
\end{aligned}$$

4.7.1 Compressible Vectors

After taking the Rudelson-Vershynin decomposition of the sphere, we first consider the compressible vectors in $\mathrm{Comp}(\delta, \rho)$. The following Lemma bounds the behavior of the compressible vectors.

Lemma 4.50 (Compressible vectors)**.** *Let A be a random matrix, whose entries a_{ij} are independent. If $\mathbb{E}\,\|A\|_{HS}^2 \leq Kn^2$, and a_{ij} are uniformly anti-concentrated (UAC),* i.e., *$\mathbb{P}(a_{ij} < a) < b$ for fixed $a > 0$, $b \in (0,1)$, then for every $\varepsilon > 0$,*

$$\mathbb{P}\left(\inf_{x \in \mathrm{Comp}(\delta,\rho)} |Ax| \leq \frac{\varepsilon}{\sqrt{n}}\right) \leq \mathbb{P}\left(\inf_{x \in \mathrm{Comp}(\delta,\rho)} \leq c\sqrt{n}\right) \leq e^{-\widetilde{c}n}$$

Proof. Let $\mathcal{N}$ be a net from Claim 4.40, such that $\mathcal{N} \subset \frac{3}{2}B_2^n \setminus \frac{1}{2}B_2^n$, $\#\mathcal{N} \leq e^{c_1 n}$, and with probability $e^{-c_1 n}$, there exists a $y \in \mathcal{N}$ such that

$$|A(x-y)| \leq \frac{C\,\mathbb{E}\,\|A\|_{HS}^2}{\sqrt{n}} \leq \widetilde{C}\sqrt{n},$$

since $\mathbb{E}\,\|A\|_{HS}^2 \leq Kn^2$, with $\widetilde{C} = C\sqrt{K}$. This implies that

$$\mathbb{P}\left(\inf_{x \in \mathrm{Comp}(\delta,\rho)} |Ax| \leq \frac{\varepsilon}{\sqrt{n}}\right) \leq \mathbb{P}\left(\inf_{x \in \mathcal{N}} |Ax| \leq \widetilde{C}\sqrt{n}\right).$$

However, since $\#\mathcal{N} \leq e^{c_1 n}$ and $\mathcal{N} \subset \frac{3}{2}B_2^n \setminus \frac{1}{2}B_2^n$,

$$\begin{aligned}\mathbb{P}\left(\inf_{x \in \mathcal{N}} |Ax| \leq \widetilde{C}\sqrt{n}\right) &\leq \#\mathcal{N} \sup_{x \in \frac{3}{2}B_2^n \setminus \frac{1}{2}B_2^n} \mathbb{P}\left(|Ax| \leq \widetilde{C}\sqrt{n}\right) \\ &\leq e^{c_1 n} \sup_{x \in \frac{3}{2}B_2^n \setminus \frac{1}{2}B_2^n} \mathbb{P}\left(|Ax| \leq \widetilde{C}\sqrt{n}\right).\end{aligned}$$

Finally, by the UAC assumption and the Tensorization lemma 4.18 (which we can use by independence), $\mathbb{P}\left(|Ax| \leq \widetilde{C}\sqrt{n}\right) \leq e^{-c'n}$, which implies that

$$\mathbb{P}\left(\inf_{x \in \mathcal{N}} |Ax| \leq \widetilde{C}\sqrt{n}\right) \leq e^{c_1 n} e^{-c'n} \leq e^{-\widetilde{c}n},$$

if $c_1 > 0$ is chosen small enough. □

Remark 4.51. *We used all the assumptions from the Theorem statement, including:*

- *Independent columns, which were used to constrct the net;*
- *$\mathbb{E}\,\|A\|_{HS}^2 \leq Kn^2$, which crucially ensured that $|A(x-y)| \leq \widetilde{C}\sqrt{n}$;*
- *Independent rows and uniform anti-concentration, which were used to ensure that $\sup_{x \in} \mathbb{P}\left(|Ax| \leq \widetilde{C}\sqrt{n}\right) \leq e^{-c'n}$.*

4.7.2 Incompressible Vectors

While the result for the compressible vectors followed immediately from previous results, we will need to use different tools to handle the case of incompressible vectors in $\text{Incomp}(\delta, \rho)$. The following Example demonstrates some crucial behavior of incompressible vectors which will be useful for bounding their behavior.

Example 4.52 (An incompressible vector)**.** *Consider the vector* $x = \left(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}}\right) \in \mathbb{S}^{n-1}$*. It is clear to see that for any* $\delta > 0$*, any* $y \in \text{Sparse}(\delta)$ *is such that*

$$|x - y| \geq \|x - y\|_\infty \geq \frac{1}{\sqrt{n}},$$

since y *has at least one zero element. Thus,* x *is in* $\text{Incomp}(\delta, \frac{1}{\sqrt{n}})$*. The vector is "spread", in the sense that there are many nonzero elements where* $|x_k| = \frac{1}{\sqrt{n}}$*.*

Example 4.52 provides some intuition for how an incompressible vector behaves. In what follows, we will try to find some characterization of this behavior, where incompressible vectors will similarly have a substantial subset of coordinates where $|x_k| \sim \frac{1}{\sqrt{n}}$.

Lemma 4.53 (Incompressible vectors are spread)**.** *For every* $x \in \text{Incomp}(\delta, \rho)$*, for* $\delta, \rho > 0$*, there exists a subset of indices* $\sigma \subset \{1, \ldots, n\}$ *with* $\#\sigma \geq \frac{1}{2}\rho^2\delta n$ *such that for every* $k \in \sigma$*,*

$$\frac{\rho}{\sqrt{2n}} \leq |x_k| \leq \frac{1}{\sqrt{\delta n}}.$$

Proof. Let $x \in \text{Incomp}(\delta, \rho)$, and consider $\sigma_1, \sigma_2 \subset \{1, \ldots, n\}$ such that

$$\sigma_1 := \left\{k : |x_k| \leq \frac{1}{\sqrt{\delta n}}\right\}, \quad \sigma_2 := \left\{k : |x_k| \geq \frac{\rho}{\sqrt{2n}}\right\}.$$

We would like to show that the cardinality of $\sigma = \sigma_1 \cap \sigma_2$ is controlled, *i.e.*, that $\#\sigma \geq cn$ for some c. First, since $x \in \mathbb{S}^{n-1}$, $|x| = 1$, and thus,

$$1 = |x|^2 = \sum_{k=1}^{n} |x_k|^2 \geq \sum_{k \in \sigma_1^\complement} |x_k|^2 \geq \sum_{k \in \sigma_1^\complement} \left|\frac{1}{\sqrt{\delta n}}\right|^2 \geq \#\sigma_1^\complement \frac{1}{\delta n}.$$

Thus, $\#\sigma_1^\complement \leq \sqrt{\delta n}$, so $\#\sigma_1 \geq n - \delta n$. Next, consider the following projection operator P_σ,

$$P_\sigma(x) = (y_1, y_2, \ldots, y_n), \quad y_k = \begin{cases} x_k & k \in \sigma \\ 0 & k \notin \sigma \end{cases}.$$

Let $y = P_{\sigma_1^\complement}(x)$. Since $\#\sigma_1 \geq n - \delta n$, this implies that $y \in \text{Sparse}(\delta)$. But since $x \in \text{Incomp}(\delta, \rho)$, $|x - y| \geq \rho$. Note that

$$|P_{\sigma_2^\complement}(x)|^2 = \sum_{k : |x_k| \leq \frac{\rho}{\sqrt{2n}}} |x_k|^2 \leq n \left|\frac{\rho}{\sqrt{2n}}\right|^2 = \frac{\rho^2}{2}, \text{ and } |P_{\sigma_1}(x)|^2 = |x - y|^2 \geq \rho^2,$$

which implies that

$$|P_\sigma(x)|^2 \geq |P_{\sigma_1}(x)|^2 - |P_{\sigma_2^\complement}(x)|^2 \geq \frac{\rho^2}{2}.$$

On the other hand,

$$|P_\sigma(x)|^2 \leq \#\sigma \max_{k\in\sigma} |x_k|^2 \leq \frac{1}{\delta n}\#\sigma,$$

and therefore, $\#\sigma \geq \delta n \frac{\rho^2}{2}$. □

Lemma 4.54 (Invertibility via distance)**.** *Let A be a random matrix, let $X_j = Ae_j$ be the columns of A, and $H_j = \operatorname{span}\{X_i : i \neq j\}$. Then for every $\delta, \rho \in (0, \frac{1}{2})$, $\varepsilon > 0$,*

$$\mathbb{P}\left(\inf_{x\in \mathrm{Incomp}(\delta,\rho)} |Ax| \leq \frac{\rho\varepsilon}{\sqrt{n}}\right) \leq \frac{1}{\delta n}\sum_{j=1}^n \mathbb{P}(\operatorname{dist}(X_j, H_j) < \varepsilon).$$

Proof. Let $x \in \mathrm{Incomp}(\delta, \rho)$, and let $X_k = Ae_k$ be the k-th column of A. Note that for any vector a, and any subspace (passing through the origin) H, $\operatorname{dist}(a, H) = \inf\{|a - h| : h \in H\} \leq |a - 0| = |a|$, since $0 \in H$. Thus, since Ax is a vector, and H_k is a subspace,

$$|Ax| \geq \max_{k=1,\dots,n} \operatorname{dist}(Ax, H_k) = \max_{k=1,\dots,n} \operatorname{dist}\left(\sum_{j=1}^n x_j X_j, H_k\right).$$

Note that by definition of H_k, $X_j \in H_k$ unless $j \neq k$. Thus, $\operatorname{dist}(X_j, H_k) = 0$ for every $j \neq k$, so those components vanish, and

$$|Ax| \geq \max_{k=1,\dots,n} \operatorname{dist}(x_k X_k, H_k) = \max_{k=1,\dots,n} |x_k| \operatorname{dist}(X_k, H_k). \tag{46}$$

Let $p_k := \mathbb{P}(\operatorname{dist}(X_k, H_k) \leq \varepsilon)$. Consider the event $\mathcal{U}$ where $\sigma_1 = \{k : \operatorname{dist}(X_k, H_k) > \varepsilon\}$ contains more than $(1 - \delta)n$ elements. Then, using Markov's inequality,

$$\begin{aligned}\mathbb{P}(\mathcal{U}^\complement) = \mathbb{P}(\#\sigma_1^\complement \geq \delta n) \leq \frac{1}{\delta n}\mathbb{E}(\#\sigma_1^\complement) &= \frac{1}{\delta n}\mathbb{E}\,\#\{k : \operatorname{dist}(X_k, H_k) \leq \varepsilon\} \\ &= \frac{1}{\delta n}\mathbb{E}\sum_{k=1}^n \mathbf{1}_{\{\operatorname{dist}(X_k,H_k)\leq\varepsilon\}} = \frac{1}{\delta n}\sum_{k=1}^n p_k. \end{aligned} \tag{47}$$

Let $\sigma_2(x) = \{k : |x_k| \geq \frac{\rho}{\sqrt{n}}\}$. Note that $\|P_{\sigma_2(x)^\complement}(x)\|_2^2 = \sum_{k:|x_k|\leq\frac{\rho}{\sqrt{n}}} |x_k|^2 \leq n\left|\frac{\rho}{\sqrt{n}}\right|^2 = \rho^2$. This implies that $\sigma_2(x)$ has at least δn elements, since otherwise we would have $|x - y| = |P_{\sigma_2(x)^\complement}(x)| \leq \rho$ for sparse vector $y = P_{\sigma_2(x)}(x)$, contradicting the incompressibility of x.

Next, suppose $\mathcal{U}$ occurs, and consider any $x \in \mathrm{Incomp}(\delta, \rho)$. We have that

$$\#\sigma_1 + \#\sigma_2(x) > (1 - \delta)n + \delta n = n,$$

which implies that $\sigma_1 \cap \sigma_2(x) \neq \emptyset$. Let $k \in \sigma_1 \cap \sigma_2(x)$, then, using (46),

$$|Ax| \geq |x_k| \operatorname{dist}(X_k, H_k) > \frac{\rho}{\sqrt{n}}\varepsilon,$$

using $|x_k| \geq \frac{\rho}{\sqrt{n}}$ from $\sigma_2(x)$ and $\operatorname{dist}(X_k, H_k) > \varepsilon$ from σ_1. We have shown that $\mathcal{U} \implies \{\inf_{x \in \operatorname{Incomp}(\delta,\rho)} |Ax| \geq \frac{\rho\varepsilon}{\sqrt{n}}\}$, or equivalently, that $\{\inf_{x \in \operatorname{Incomp}(\delta,\rho)} |Ax| \leq \frac{\rho\varepsilon}{\sqrt{n}}\} \implies \mathcal{U}^{\complement}$, so

$$\mathbb{P}\left(\inf_{x \in \operatorname{Incomp}(\delta,\rho)} |Ax| \leq \frac{\rho\varepsilon}{\sqrt{n}}\right) \leq \mathbb{P}(\mathcal{U}^{\complement}) = \frac{1}{\delta n}\sum_{k=1}^{n} p_k,$$

which follows from (47). □

Lemma 4.54 reduces the problem from dealing with the incompressible vectors to finding an estimate of the following form

$$\mathbb{P}(\operatorname{dist}(X, H) \leq \varepsilon) \leq c\varepsilon + \frac{c_1}{\sqrt{n}},$$

where X is a random vector with independent UAC entries, and H is the span of $n-1$ independent random vectors with independent UAC entries.

4.7.3 Distance Theorem

Recall that our goal is to prove the following 'distance' theorem

Theorem 4.55 (Distance theorem)**.** *Let X be a random vector with uniformly anti-concentrated (UAC) entries . Let H be the span of $n-1$ independent random vectors with independent UAC entries. Then*

$$\mathbb{P}\left(\operatorname{dist}(X, H) \leq \varepsilon\right) \leq c\varepsilon + \frac{c_1}{\sqrt{n}}. \tag{48}$$

Remark 4.56. *We remark that 4.55 bounds the probability of the distance to the subspace by $c\varepsilon + \frac{c_1}{\sqrt{n}}$, but one can in fact improve the bound to $c\varepsilon + e^{-c_1 n}$ using more sophisticated techniques [26].*

As motivation for our next theorem, we note that the distance between X and H is exactly the length of the component of X that is orthogonal to H. In particular $\operatorname{dist}(X, H) = \langle x, n\rangle$, where n is the normal unit vector of H.

Theorem 4.57 (Rogozin's Theorem)**.** *Let $v = (v_1, \cdots, v_n)$ be a random vector with independent UAC entries, say*

$$\sup_{z \in \mathbb{R}} \mathbb{P}\left(|v_i - z| < a\right) < b. \tag{49}$$

Then for any $u \in \mathbb{R}^n$ and any $\varepsilon > ca\|u\|_\infty$

$$\sup_{z\in\mathbb{R}} \mathbb{P}\left(|\langle u, v\rangle - z| < \varepsilon\right) \leq \frac{C\varepsilon}{|u|}, \tag{50}$$

where C, c depend only on a and b.

Corollary 4.58. *Let $c_1, c_2 > 0$ and let u be a random vector satisfying*

$$\#\{i \ : \ |u_i| \geq \frac{c_1}{\sqrt{n}}\} \geq c_2 n.$$

If v is a random vector with independent UAC entries then for all $\varepsilon > \frac{c}{\sqrt{n}}$ one has

$$\sup_{z\in\mathbb{R}} \mathbb{P}[|\langle u, v\rangle - z| < \varepsilon] \leq c_2\varepsilon.$$

Proof. Let $\sigma := \{i : |u_i| \geq \frac{c_1}{\sqrt{n}}\}$. By assumption $|\sigma| \geq c_2 n$. Note that we may write $\langle u, v\rangle = R + \sum_{i\in\sigma} u_i v_i$, where $R = \sum_{i\notin\sigma} u_i v_i$. Therefore by Theorem 4.57

$$\begin{aligned}
\sup_{z\in\mathbb{R}} \mathbb{P}\left(|\langle u, v\rangle - z| < \varepsilon\right) &= \sup_{z\in\mathbb{R}} \mathbb{P}\left(\left|\sum_{i\in\sigma} u_i v_i - (z - R)\right| < \varepsilon\right) \\
&= \sup_{z\in\mathbb{R}} \mathbb{E}_R\, \mathbb{P}\left(\left|\sum_{i\in\sigma} u_i v_i - (z - R)\right| < \varepsilon\right) \\
&\leq \mathbb{E}_R \sup_{z\in\mathbb{R}} \mathbb{P}\left(\left|\sum_{i\in\sigma} u_i v_i - (z - R)\right| < \varepsilon\right) \\
&= \sup_{y\in\mathbb{R}} \mathbb{P}\left(\left|\sum_{i\in\sigma} u_i v_i - y\right| < \varepsilon\right) \\
&\leq \frac{c\varepsilon}{\sqrt{\sum_{i\in\sigma} u_i^2}} = \tilde{c}\varepsilon,
\end{aligned}$$

whenever $\varepsilon \geq \sup_{i\in\sigma} |u_i| \geq \frac{c_1}{\sqrt{n}}$. □

Using Corollary 4.58 we will show that

$$\sup_{z\in\mathbb{R}} \mathbb{P}\left(|\langle X, n\rangle - z| < \varepsilon\right) \leq C\varepsilon + \frac{c_1}{\sqrt{n}}$$

Where $X = Ae_j$ and n is the the unit normal to $\operatorname{span}(Ae_j \ : \ i \neq j)$. This, however, will require that n is incompressible.

Lemma 4.59 (Random normal is incompressible)**.** *Let $H = \operatorname{span}(Ae_i \ : \ i \neq j)$, where A is a matrix satisfying the assumptions of Theorem 4.45. Let $n \perp H$ be a unit vector. Then n is incompressible with probability $1 - e^{-cn}$.*

Proof. Note that the condition $n \perp H$ is equivalent to the condition $B^\top n = 0$, where $B = [Ae_1, \cdots, Ae_{j-1}, Ae_{j+1}, \cdots, Ae_n]$. Then

$$\mathbb{P}(n \in \mathrm{Comp}(\delta, \rho)) \leq \mathbb{P}\left(\inf_{x \in \mathrm{Comp}(\delta,\rho)} |B^\top x| = 0\right) \leq e^{-cn},$$

where the exponential failure probability follows from the net-argument for compressible vectors, as done in Lemma 4.50. □

This lemma, in conjunction with Corollary 4.58, will give us the distance theorem with $X = Ae_i$ and $H = \mathrm{span}(Ae_j \ : \ j \neq i)$.

4.7.4 Proof of the first part of Theorem 4.45

Recall that A has independent UAC entries and satisfies $\mathbb{E}\,\|A\|_{HS}^2 \leq Kn^2$. Therefore

$$\begin{aligned}
\mathbb{P}\left(\sigma_n(A) \leq \frac{\varepsilon}{\sqrt{n}}\right) &\leq \mathbb{P}\left(\inf_{x \in \mathrm{Comp}(\delta,\rho)} |Ax| < \frac{\varepsilon}{\sqrt{n}}\right) + \mathbb{P}\left(\inf_{x \in \mathrm{Incomp}(\delta,\rho)} |Ax| < \frac{\varepsilon}{\sqrt{n}}\right), \\
&\leq e^{-c_1 n} + \mathbb{P}\left(\inf_{x \in \mathrm{Incomp}(\delta,\rho)} |Ax| < \frac{\varepsilon}{\sqrt{n}}\right), \\
&\leq e^{-c_1 n} + \frac{1}{\delta n} \sum_{i=1}^{n} \mathbb{P}(\mathrm{dist}(Ae_i, H) < \varepsilon).
\end{aligned}$$

Note that in the first line we used the Rudelson-Vershynin decomposition of the sphere, in the second line we used Lemma 4.50, and in the third line we used Lemma 4.54. Next we have that

$$\begin{aligned}
\mathbb{P}\left(\mathrm{dist}(Ae_i, H) < \varepsilon\right) &\leq \mathbb{P}\left(|\langle Ae, n\rangle| \leq \varepsilon\right), \\
&\leq e^{-c_1 n} + \mathbb{P}\left(|\langle Ae, n\rangle| \leq \varepsilon, \ n \in \mathrm{Incomp}(\delta, \rho)\right), \\
&\leq e^{-c_1 n} + \sup_{\substack{u \text{ s.t.} \\ \#\{i \ : \ |u_i| \geq c_1/\sqrt{n}\} \geq c_2 n}} \mathbb{P}(|\langle Ae_i, u\rangle| \leq \varepsilon), \\
&\leq e^{-c_1 n} + c\varepsilon + \frac{c_2}{\sqrt{n}}.
\end{aligned}$$

Note that in the second line we used Lemma 4.59, in the third line we used the fact that incompressible vectors are spread, and in the fourth line used Rogozin's theorem. Plugging this estimate in our bound for the smallest singular value yields

$$\mathbb{P}\left(\sigma_n(A) \leq \frac{\varepsilon}{\sqrt{n}}\right) \leq e^{-c_1 n} + \left(\frac{e^{-c_1 n}}{\delta} + \frac{c\varepsilon}{\delta} + \frac{c_2}{\delta\sqrt{n}}\right), \tag{51}$$

$$\leq c'\varepsilon + \frac{c''}{\sqrt{n}}. \tag{52}$$

5 Gaussian Random Processes

Recall that a random vector is a collection of n random variables, forming a vector (i.e. $x = (x_1, \cdots, x_n)$) . A random walk is a sequence of random variables (i.e. $\{x_1, x_2, \dots\}$). In 1-D a random process is a collection of random variables indexed by "time". Examples include

$$\{X_t \ : \ t \in \mathbb{R}\}, \ \ \{X_t \ : \ t \geq 0\}, \ \ \{X_t \ : \ t \in [a, b]\}.$$

5.1 Basic Concepts and Examples

Example 5.1 (Brownian Motion)**.** *A brownian motion $\{X_t \ : \ t \geq 0\}$ is a random process having the following properties:*

1. *For all $s \geq t \geq 0$ the random variable $X_s - X_t$ is distributed as the normal variable $N(0, t-s)$. This is known as having "gaussian increments".*
2. *The function $f(t) = X_t$ is continuous in t almost surely.*
3. *For $\gamma \leq \tau \leq t \leq s$ the increments $X_s - X_t$ and $X_\tau - X_\gamma$ are independent.*

We will be interested in high dimensional random processes (i.e. random processes where "time" is a subset of $\mathbb{R}^n$).

Example 5.2 (Ocean temperature)**.** *To represent the temperature of the ocean as a random process we can take $T \subset \mathbb{R}^n$ and for $t \in T$ let X_t denote the temperature at t.*

Example 5.3 (Random Projection)**.** *Let $g \sim N(0, I_d)$ be a standard gaussian vector. In other words $g = (g_1, \dots, g_n)$ where the cooordinates g_i are independent standard normal gaussians. Then for $t \in T \subset \mathbb{R}^n$ we define $X_t := \langle g, t \rangle$.*

We now recall the definition of a gaussian random vector. Given a non-negative definite $n \times n$ matrix A, and a vector b $\in \mathbb{R}^n$, we define the random variable $X \sim N(b, A)$ whose law has density $e^{\langle A(x-b), x-b \rangle} \cdot c_n$, where c_n is chosen so that the density integrates to 1.

Remark 5.4. ***(Homework)*** *Let $X \sim N(0, \Sigma)$ be a random gaussian vector where $\Sigma = AA^\top$. Then there exists $g \sim N(0, \mathrm{Id}_n)$ such that $X_i = \langle g, u_i \rangle$ for some $u_1, \dots, u_n \in \mathbb{R}^n$. Note that the coordinates of X may be dependent.*

Definition 5.5. The covariance matrix Σ of a random vector $X \in \mathbb{R}^n$ is the $n \times n$ matrix with entries

$$\Sigma_{ij} = \mathbb{E}[(X - \mathbb{E}\, X_i)(X - \mathbb{E}\, X_j)].$$

Remark 5.6. *If X has independent entries then Σ is a diagonal matrix.*

As a cool fact (HW), we remark that the distribution of a Gaussian vector $X \sim N(0, \Sigma)$ is uniquely determined by its covariance matrix. In general, knowing that a random vector belongs to a certain class of vectors (say Poisson or Exponential) and knowing its covariance matrix are not enough to recover its distribution.

Definition 5.7 (Covariance Function)**.** Let $\{X_t : t \in T\}$ be a "mean zero" random process (i.e. $\mathbb{E}[X_t] = 0$ for all $t \in T$). We define $\Sigma \colon T \times T \to \mathbb{R}$ according to

$$\Sigma(t, s) = \mathbb{E}[X_t X_s].$$

Σ is known as the covariance function and is the random process analogue to the covariance matrix.

Definition 5.8 (Gaussian Random Process)**.** Let $T \subset \mathbb{R}^n$. A random process $\{X_t : t \in T\}$ is called a Gaussian Random Process (GRP) if for every finite subset $T_0 \subset T$, the vector $(X_t)_{t \in T_0}$ is a gaussian vector. An equivalent characterization is that for every finite subset $T_0 \subset T$ and vector $(a_t)_{t \in T_0}$ the linear combination $\sum_{t \in T_0} a_t X_t$ is a gaussian random variable. This equivalence is because the projection of a gaussian vector in any direction is a normal random variable, and a random vector whose projection in every direction is a normal random variable must be a gaussian vector. (HW) A gaussian random process is determined by its covariance function.

Definition 5.9 ("Canonical GRP")**.** Let $T \subset R^n$ and let $g \sim N(0, \mathrm{Id_n})$. For every $t \in T$ define $X_t = \langle g, t \rangle$. Then $\{X_t : t \in T\}$ is known as a Canonical Gaussian Process. Note that this is indeed a Gaussian Random Process since, by the definition of X_t, each X_t is the projection of a gaussian vector and therefore a normal random variable and therefore any linear combination of a finite number of X_t is a normal random variable.

Lemma 5.10 (All GRP are canonical)**.** *Let Y_t be a mean zero Gaussian Random Process. Then there exists $T \subset \mathbb{R}^n$ such that $Y_t = \langle g, t \rangle$ for all $t \in T$, where $g \sim N(0, 1)$.*

5.2 Slepian's Inequality

In applications, it is useful to have a uniform control on a random process $\{X_t : t \in T\}$, i.e. to have a bound on $\mathbb{E} \sup_{t \in T} X_t$.

For some processes, this quantity can be computed exactly. For example, if $\{X_t : t \in T\}$ is a standard Brownian motion, then by reflection principle, we have $\mathbb{E} \sup_{t \le t_0} X_t = \sqrt{2t_0/\pi}$ for every $t_0 \ge 0$. For general random processes, even if they are Gaussian, the problem is emphatically nontrivial.

The first general bound we will prove is Slepian's comparison inequality for Gaussian processes. Intuitively, it states that the faster the process grows (in terms of the magnitude of the increments), the farther it gets.

Theorem 5.11 (Slepian's Inequality)**.** *Let $\{X_t : t \in T\}, \{Y_t : t \in T\}$ be mean zero gaussian processes indexed by T. Suppose that for all $s, t \in T$ it holds that*

$$\mathbb{E}\, X_t^2 = \mathbb{E}\, Y_t^2, \quad \mathbb{E}(X_t - X_s)^2 \le \mathbb{E}(Y_t - Y_s)^2.$$

Then for all $\tau \in \mathbb{R}$ it follows that

$$\mathbb{P}\left[\sup_{t\in T} X_t \ge \tau\right] \le \mathbb{P}\left[\sup_{t\in T} Y_t \ge \tau\right], \tag{53}$$

and thus

$$\mathbb{E}\sup_{t\in T} X_t \le \mathbb{E}\sup_{t\in T} Y_t. \tag{54}$$

Remark 5.12 (Homework)**.** *Instead of the Gaussian processes $\{X_t : t \in T\}$ and $\{Y_t : t \in T\}$, it suffices to prove the above inequality for the Gaussian random vectors X and Y in $\mathbb{R}^n$. Furthermore, it suffices to prove for the case when X and Y are independent. Hence, Theorem 5.11 is equivalent to Theorem 5.19 which will be proved later.*

Then the inequalities (53) and (54) are equivalent to

$$\mathbb{P}\left[\sup_{i\le n} X_i \ge \tau\right] \le \mathbb{P}\left[\sup_{i\le n} Y_i \ge \tau\right], \qquad \mathbb{E}\sup_{i\le n} X_i \le \mathbb{E}\sup_{i\le n} Y_i,$$

where the first inequality always guarantees the second since

$$\mathbb{E}\max_{i\le n} X_i^2 = \int_0^\infty \mathbb{P}\left(\max_{i\le n} X_i^2 \ge \tau\right) d\tau \le \int_0^\infty \mathbb{P}\left(\max_{i\le n} Y_i^2 \ge \tau\right) d\tau = \mathbb{E}\max_{i\le n} Y_i^2$$

by the tail formula for non-negative random variables.

5.2.1 Gaussian Interpolation

The proof of Slepian's inequality will be based on the technique of Gaussian Interpolation which is described as follows.

Definition 5.13 (Gaussian Interpolation)**.** For any pair of independent Gaussian random vectors $X, Y \in \mathbb{R}^n$, not necessarily standard, define a Gaussian random vector $Z(u)$ in $\mathbb{R}^n$ that continuous interpolates between $Z(0) = Y$ and $Z(1) = X$:

$$Z(u) := \sqrt{u}\, X + \sqrt{1-u}\, Y, \quad u \in [0,1].$$

Remark 5.14 (Homework)**.** *The covriance matrix of $Z(u)$ interpolates linearly between the covariance matrices of Y and X. Namely, if $\Sigma(X)$ is the covariance matrix for X and $\Sigma(Y)$ is the covariance matrix for Y, then*

$$\Sigma(Z(u)) = u\Sigma(X) + (1-u)\Sigma(Y)$$

Consider the indicator function for vector $x = (x_1, \cdots, x_n)$

$$f(x) := \mathbf{1}_{\{\max_i x_i < \tau\}}$$

which satisfies $\mathbb{E}\, f(Z(1)) = \mathbb{P}\left(\max_{i\le n} X_i < \tau\right)$ and $\mathbb{E}\, f(Z(0)) = \mathbb{P}\left(\max_{i\le n} Y_i < \tau\right)$. Now, if we can show

$$\mathbb{E}\, f(Z(1)) \ge \mathbb{E}\, f(Z(0))$$

then inequality (53) can be concluded. Our goal now shifts to study how the quantity $\mathbb{E}\, f(Z(u))$ changes as u increases from 0 to 1. We approach this goal by starting with the following identity which is a version of integration by parts in Gaussian expectations.

Lemma 5.15 (Gaussian integration by parts)**.** *Let $X \sim N(0,1)$. Then for any differentiable function $f : \mathbb{R} \to \mathbb{R}$ we have*

$$\mathbb{E}\, f'(X) = \mathbb{E}\, X f(X) \tag{55}$$

Proof. It suffices to argue for the case when f has bounded support, and this identity then can be extended to general functions by a standard approximation argument. By density of standard normal and Integration by parts, we have

$$\mathbb{E}\, f'(X) = \int_{\mathbb{R}} f'(t)\frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt = 0 + \int_{\mathbb{R}} t f(t)\frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt = \mathbb{E}\, X f(X)$$

as claimed, where the zero term comes from the fact that $f(t)e^{-t^2/2}$ has limits equal to zero when t goes to both positive and negative infinity if f has bounded support. □

Corollary 5.16 (Homework: Multivariate Gaussian Integration by parts)**.** *Let $X \sim N(0, \Sigma)$. Then for any differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, we have*

$$\mathbb{E}\, X f(X) = \Sigma\ \ \mathbb{E}\, \nabla f(X),$$

where $\nabla f(X)$ is the n-dimensional gradient vector with entries $\mathbb{E}\, \partial_i f(X)$.

Lemma 5.17 (Gaussian Interpolation)**.** *Consider two independent Gaussian random vectors $X \sim N(0, \Sigma^X)$ and $Y \sim N(0, \Sigma^Y)$. Define the interpolation Gaussian vector*

$$Z(u) := \sqrt{u}X + \sqrt{1-u}Y, \qquad u \in [0,1]. \tag{56}$$

Then for any twice-differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, we have

$$\frac{d}{du}\mathbb{E}\, f(Z(u)) = \frac{1}{2}\sum_{i,j=1}^{n}\left(\Sigma^X_{ij} - \Sigma^Y_{ij}\right)\mathbb{E}\,\frac{\partial^2 f}{\partial x_i \partial x_j}(Z(u)), \tag{57}$$

where Σ^X_{ij} stands for the (i,j) entry of Σ^X.

Proof. By the (multivariate) chain rule, we have

$$\frac{d}{du}\,\mathbb{E}\,f(Z(u)) = \sum_{i=1}^{n} \mathbb{E}\,\frac{\partial f}{\partial x_i}(Z(u))\frac{dZ_i}{du} = \frac{1}{2}\sum_{i=1}^{n} \mathbb{E}\,\frac{\partial f}{\partial x_i}(Z(u))\left(\frac{X_i}{\sqrt{u}} - \frac{Y_i}{\sqrt{1-u}}\right),$$

where the second equality is from the definition (56) of $Z(u)$. Now break this sum into two, and first compute the contribution of terms containing X_i. To this end, we condition on Y and express

$$\sum_{i=1}^{n} \frac{1}{\sqrt{u}}\,\mathbb{E}\,X_i\frac{\partial f}{\partial x_i}(Z(u)) = \sum_{i=1}^{n} \frac{1}{\sqrt{u}}\,\mathbb{E}\,X_i g_i(X),$$

where

$$g_i(X) = \frac{\partial f}{\partial x_i}\left(\sqrt{u}X + \sqrt{1-u}Y\right).$$

Apply the multivariate Gaussian integration by parts (Corollary 5.16), we have

$$\mathbb{E}\,X_i g_i(X) = \sum_{j=1}^{n} \Sigma_{ij}^X\,\mathbb{E}\,\frac{\partial g_i}{\partial x_j}(X) = \sum_{j=1}^{n} \Sigma_{ij}^X\,\mathbb{E}\,\frac{\partial^2 f}{\partial x_i \partial x_j}(X)\left(\sqrt{u}X + \sqrt{1-u}Y\right)\sqrt{u},$$

where the second equality is by definition of g_i. Substituting this back into the previous equation gives

$$\sum_{i=1}^{n} \frac{1}{\sqrt{u}}\,\mathbb{E}\,X_i\frac{\partial f}{\partial x_i}(Z(u)) = \sum_{i,j=1}^{n} \Sigma_{ij}^X\,\mathbb{E}\,\frac{\partial^2 f}{\partial x_i \partial x_j}(Z(u)).$$

Taking expectation of both sizes with respect to Y, we remove the conditioning on Y. Similar discussion works for the contribution of terms containing Y_i, and that should yield

$$\sum_{i=1}^{n} \frac{1}{\sqrt{1-u}}\,\mathbb{E}\,X_i\frac{\partial f}{\partial x_i}(Z(u)) = \sum_{i,j=1}^{n} \Sigma_{ij}^Y\,\mathbb{E}\,\frac{\partial^2 f}{\partial x_i \partial x_j}(Z(u)).$$

Combining these two equalities, together with the first line of the proof, gives the desired relation (57). □

5.2.2 Proof of Slepian's Inequality

We are ready to establish the key lemma in proving Slepian's inequality which is also known as a preliminary functional form of Slepian's inequality.

Lemma 5.18 (Slepian 's inequality: functional form)**.** *Consider two mean-zero independent Gaussian random vectors X and Y in $\mathbb{R}^n$. Assume that for all $i, j = 1, \cdots, n$, we have*

$$\mathbb{E}\,X_i^2 = \mathbb{E}\,Y_i^2 \qquad \text{and} \qquad \mathbb{E}(X_i - X_j)^2 \le \mathbb{E}(Y_i - Y_j)^2. \tag{58}$$

Consider a twice-differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ such that

$$\frac{\partial^2 f}{\partial x_i \partial x_j} \geq 0 \qquad for\ all\ i \neq j.$$

Then

$$\mathbb{E}\, f(X) \geq \mathbb{E}\, f(Y).$$

Proof. The assumptions (58) imply that the entries of the covariance matrices Σ^X and Σ^Y of X and Y satisfy

$$\Sigma^X_{ii} = \Sigma^Y_{ii} \qquad and \qquad \Sigma^X_{ij} \geq \Sigma^Y_{ij}, \qquad for\ all\ i, j = 1, \cdots, n,$$

where for the second relation we used $ab = (a^2 + b^2 - (a-b)^2)/2$. Applying Lemma 5.17 with our assumptions gives

$$\frac{d}{du}\, \mathbb{E}\, f(Z(u)) = \frac{1}{2} \sum_{i,j=1}^{n} \left(\Sigma^X_{ij} - \Sigma^Y_{ij}\right) \mathbb{E}\, \frac{\partial^2 f}{\partial x_i \partial x_j}(Z(u)) \geq 0$$

by assumptions, where $Z(u)$ is the Gaussian interpolation we defined in (56). This means $\mathbb{E}\, f(Z(u))$ increases in u. But $Z(0) = Y$ and $Z(1) = X$ by the way we constructed the interpolation, hence $\mathbb{E}\, f(X) = \mathbb{E}\, f(Z(1)) \geq \mathbb{E}\, f(Z(0)) = \mathbb{E}\, f(Y)$ as desired. □

Eventually, now we are ready to prove Slepian's inequality (53) in it's equivalent form which is in terms of random vectors rather than random processes.

Theorem 5.19 (Slepian's inequality: random vector)**.** *Let X and Y be two mean zero independent Gaussian random vectors in $\mathbb{R}^n$. Suppose that for all $s, t \in T$ it holds that*

$$\mathbb{E}\, X_i^2 = \mathbb{E}\, Y_i^2, \quad \mathbb{E}(X_i - X_j)^2 \leq \mathbb{E}(Y_i - Y_j)^2.$$

Then for every $\tau \in \mathbb{R}$ we have

$$\mathbb{P}\left[\max_{i \leq n} X_i \geq \tau\right] \leq \mathbb{P}\left[\max_{i \leq n} Y_i \geq \tau\right]. \tag{59}$$

Consequently,

$$\mathbb{E} \max_{i \leq n} X_i \leq \mathbb{E} \max_{i \leq n} Y_i. \tag{60}$$

Proof. Let $h : \mathbb{R} \to [0, 1]$ be a twice-differentiable non-increasing approximation to the indicator function of the interval $(-\infty, \tau)$ satisfying

$$h(x) \approx \mathbf{1}_{(-\infty, \tau)}.$$

Define the function $f : \mathbb{R}^n \to \mathbb{R}$ by $f(x) = h(x_1) \cdots h(x_n)$ for any $x = [x_1, \cdots, x_n] \in \mathbb{R}^n$, then

$$f(x) \approx \mathbf{1}_{\{\max_i x_i < \tau\}}$$

To apply the functional form of Slepian's inequality, we need to check the assumption for f. Note that, for every $i \neq j$, we have

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = h'(x_i)h'(x_j) \prod_{k \notin \{i,j\}} h(h_k).$$

The first two factors are non-positive and the other are non-negative by assumption. So the second mixture derivative of f is always non-negative. Hence we can apply the key Lemma above to conclude that

$$\mathbb{E}\, f(X) \geq \mathbb{E}\, f(Y).$$

Now, by approximation, we have

$$\mathbb{P}\left(\max_{i \leq n} X_i < \tau\right) = \mathbb{E}\, \mathbf{1}_{\{\max_{i \leq n} X_i < \tau\}} \approx \mathbb{E}\, f(X) \geq \mathbb{E}\, f(Y) \approx \mathbb{E}\, \mathbf{1}_{\{\max_{i \leq n} Y_i < \tau\}} = \mathbb{P}\left(\max_{i \leq n} Y_i < \tau\right)$$

which implies (59). And relation (60) follows from (59) as we discussed before. □

5.2.3 The Sudakov-Fernique Inequality

In theorem 5.11, Slepian's inequality has two assumptions on the random processes $\{X_t : t \in T\}$ and $\{Y_t : t \in T\}$: the equality of variances and the dominance of increments. It turns out that, even if we drop the assumption on the equality of variances, we will still be able to obtain the inequality on expectations. This more pratically useful result is due to Sudakov and Fernique.

Theorem 5.20 (Sudakov-Fernique Inequality)**.** *Let $\{X_t : t \in T\}$ and $\{Y_t : t \in T\}$ be two mean zero Gaussian processes. Assume that, for all $t, s \in T$, we have*

$$\mathbb{E}(X_t - X_s)^2 \leq \mathbb{E}(Y_t - Y_s)^2.$$

Then

$$\mathbb{E} \sup_{t \in T} X_t \leq \mathbb{E} \sup_{t \in T} Y_t.$$

Proof. It is enough to prove this theorem for Gaussian random vectors X and Y in $\mathbb{R}^n$, just as we did for Slepian's inequality. We again deduce the result form the Gaussian Interpolation lemma 5.17. But this time, instead of choosing a function $f(x)$ that approximates the indicator function of $\{\max_i x_i < \tau\}$, we want $f(x)$ to approximate $\max_i x_i$.

To this end, let $\beta > 0$ be a parameter and define the function

$$f_\beta(x) := \frac{1}{\beta} \log \sum_{i=1}^{n} e^{\beta x_i}.$$

One can check (Homework!) that this function is twice differentiable and

$$f_\beta(x) \to \max_{i \leq n} x_i \quad \text{as } \beta \to \infty.$$

Substitute $f(x)$ into the Gaussian interpolation formula and simplifying the expression shows that (Homework!)

$$\frac{d}{du}\,\mathbb{E}\,f(Z(u)) \le 0, \qquad for\ all\ u.$$

The proof can then be completed in just the same way as the proof of Slepian's inequality. □

5.2.4 Application of Sudakov-Fernique Inequality to Gaussian Random Matrices

Theorem 5.21. *Let A be an $m \times n$ matrix with entries (a_{ij}) such that a_{ij} are independent and $a_{ij} \sim \mathcal{N}(0,1)$. Then we have*

$$\mathbb{E}\|A\| \le \sqrt{m} + \sqrt{n}.$$

Remark 5.22. *Note $\|\cdot\|_{op} = \sigma_1(A)$. This is consistent with our past results as we have proved that sub-Gaussian ensures $\mathbb{P}\left(\|A\| \ge t(\sqrt{m}+\sqrt{n})\right) \le e^{-cm}$ with $m \ge n$. But here we see that the constant is 1, i.e., $\mathbb{E}\|A\| \le C\left(\sqrt{m}+\sqrt{n}\right)$ with $C = 1$.*

Proof of Theorem 5.21. Let A be an $m \times n$ matrix with entries (a_{ij}) such that a_{ij} are independent and $a_{ij} \sim \mathcal{N}(0,1)$. Then

$$\|A\| = \sup_{x\in\mathbb{S}^{n-1}} |Ax| = \sup_{\substack{y\in\mathbb{S}^{m-1}\\ x\in\mathbb{S}^{n-1}}} \langle Ax, y\rangle = \max_{\substack{t=(x,y)\\ t\in T\\ T=\mathbb{S}^{n-1}\times\mathbb{S}^{m-1}}} X_t,$$

since if we have $a \in \mathbb{R}^m$, then $|a| = \sup_{y\in\mathbb{S}^{m-1}} \langle a, y\rangle$. Here we have X_t as a random Gaussian process, indeed, if we fix x and y, then $\langle Ax, y\rangle$ is a Gaussian random variable. If A has all the Gaussian entries, then all rows of A are Gaussian processes, i.e.,

$$Ax = \begin{pmatrix} \langle A^\top e_1, x\rangle \\ \vdots \\ \langle A^\top e_m, x\rangle \end{pmatrix}$$

with $A^\top e_i \sim \mathcal{N}(0, Id)$ and $A^\top e_i$ are all independent. Thus, Ax is a vector with independent coordinates and each of them is Gaussian, so Ax is Gaussian, and $\langle Ax, y\rangle$ is a Gaussian random variable. The idea of the proof is that we can apply Sudakov-Fernique Inequality 5.20 to find Y_t indexed by $\mathbb{S}^{m-1}\times\mathbb{S}^{n-1}$ which is also Gaussian, then compare X_t to Y_t to get the upper bound of $\mathbb{E}\|A\| = \mathbb{E}\sup X_t \le \mathbb{E}\sup Y_t$ and find the increment by estimating from above of $\mathbb{E}\left(X_t - X_s\right)^2$.

Let $t = (u, v) \in \mathbb{S}^{n-1} \times \mathbb{S}^{m-1}$ and let $s = (w, z) \in \mathbb{S}^{n-1} \times \mathbb{S}^{m-1}$. Then

$$\begin{aligned}
\mathbb{E}\left(X_t - X_s\right)^2 &= \mathbb{E}\left(\langle Au, v\rangle - \langle Aw, z\rangle\right)^2 \\
&= \mathbb{E}\left(\sum_{i,j} a_{ij}\left(u_i v_j - w_i z_j\right)\right)^2 \\
&= \sum_{i,j} \mathbb{E} a_{ij}^2 \left(u_i v_j - w_i z_j\right)^2 \\
&= \sum_{i,j} \left(u_i v_j - w_i z_j\right)^2 \\
&= \|u \otimes v - w \otimes z\|_{HS}^2 \\
&\overset{\text{HW}}{\leq} |u - w|^2 + |v - z|^2.
\end{aligned}$$

Recall that for independent and mean zero $\xi_1, \ldots, \xi_k$, we have $\mathbb{E}\left(\sum \xi_l\right)^2 = \sum \mathbb{E}\xi_l^2$ so here $a_{ij}\left(u_i v_j - w_i z_j\right)$ are independent and mean zero. We can conclude that

$$\mathbb{E}|X_{uv} - X_{wz}|^2 \leq |u - w|^2 + |v - z|^2.$$

In fact, we can construct Y_t with $t \in \mathbb{S}^{n-1} \times \mathbb{S}^{m-1}$ such that

$$\mathbb{E}|Y_{uv} - Y_{wz}|^2 \leq |u - w|^2 + |v - z|^2.$$

Indeed, consider $Y_{uv} = \langle g, u\rangle + \langle h, v\rangle$ where $h \sim \mathcal{N}(0, Id_m)$, $g \sim \mathcal{N}(0, Id_n)$, h, g are independent, and $(u, v) \in \mathbb{S}^{n-1} \times \mathbb{S}^{m-1}$. Note that here Y_{uv} is a Gaussian random process, then by definition

$$\begin{aligned}
\mathbb{E}|Y_{uv} - Y_{wz}|^2 &= \mathbb{E}\left|\langle g, u\rangle + \langle h, v\rangle - \langle g, w\rangle - \langle h, z\rangle\right|^2 \\
&= \mathbb{E}\langle g, u - w\rangle^2 + \mathbb{E}\langle h, v - z\rangle^2 \qquad \text{(by independence and mean zero)} \\
&= |u - w|^2 + |v - z|^2
\end{aligned}$$

since if $X \sim \mathcal{N}(0, Id)$, we have $\langle x, \theta\rangle \sim \mathcal{N}(0, |\theta|^2)$.
By the results above, we can conclude that $\mathbb{E}|X_t - X_s|^2 \leq \mathbb{E}|Y_t - Y_s|^2$. Since X_t and Y_t are Gaussian, by Sudakov-Fernique Inequality 5.20, we have

$$\begin{aligned}
\mathbb{E}\|A\|_{op}^2 = \mathbb{E}\sup_{t \in \mathbb{S}^{n-1} \times \mathbb{S}^{m-1}} X_t &\leq \mathbb{E}\sup_{t \in \mathbb{S}^{n-1} \times \mathbb{S}^{m-1}} Y_t \\
&= \mathbb{E}\sup_{\substack{u \in \mathbb{S}^{n-1} \\ v \in \mathbb{S}^{m-1}}} \left(\langle g, u\rangle + \langle h, v\rangle\right) \\
&= \mathbb{E}\sup_{u \in \mathbb{S}^{n-1}} \langle g, u\rangle + \mathbb{E}\sup_{v \in \mathbb{S}^{m-1}} \langle h, v\rangle \\
&= \mathbb{E}|g| + \mathbb{E}|h| \\
&\overset{\text{Cauchy}}{\leq} \sqrt{\mathbb{E}|g|^2} + \sqrt{\mathbb{E}|h|^2} \\
&= \sqrt{n} + \sqrt{m}
\end{aligned}$$

as $\mathbb{E}|g|^2 = \mathbb{E}\left(\sum_{i=1}^{n} g_i^2\right) = n$ and $g \sim \mathcal{N}(0, Id)$. □

Definition 5.23 (Sub-Gaussian Random Process)**.** For X_t with $t \in T$ and metric d on T, we say that the random process X_t is *sub-Gaussian* if for some constant $K \geq 0$,

$$\|X_t - X_s\|_{\psi_2} \leq K \cdot d(t, s),$$

i.e., all increments are sub-Gaussian random vectors.

Note that Gaussian random process are sub-Gaussian.

Theorem 5.24 (Dudley's Inequality)**.** *Suppose X_t with $t \in T$ is mean zero random process on the metric (T, d) and it is sub-Gaussian with some constant K. Then*

$$\mathbb{E} \sup_{t \in T} X_t \leq CK \sum_{j \in \mathbb{Z}} 2^{-j} \sqrt{\log N\left(T, d, 2^{-j}\right)},$$

where $\log N\left(T, d, 2^{-j}\right)$ is the metric entropy, i.e., $N\left(T, d, 2^{-j}\right)$ is the smallest number of balls of radius 2^{-j} required to cover T in metric d.

Remark 5.25. *If T is compact, there is $K \in \mathbb{Z}$ such that for all $j \leq K$, we have $N\left(T, d, 2^{-j}\right) = 1$ such that all summands are zero.*

Example 5.26. *Let $T = \mathbb{S}^{n-1} \subset \mathbb{R}^n$ and let d be the Euclidean metric. For any $\epsilon > 0$, we have $N(\mathbb{S}^{n-1}, d, \epsilon) \leq \left(\frac{3}{\epsilon}\right)^{n-1}$. Then*

$$\begin{aligned} C \sum_{j \in \mathbb{Z}} 2^{-j} \sqrt{\log N\left(\mathbb{S}^{n-1}, d, 2^{-j}\right)} &= C \sum_{j \geq 0} 2^{-j} \sqrt{(n-1)\left(\log 2^j \cdot 3\right)} \\ &= c' \sqrt{n} \sum_{j \geq 0} \sqrt{j}^{2^{-j}} = c'' \sqrt{n}. \end{aligned}$$

In conclusion, we have $\mathbb{E} \sup X_t \leq c\sqrt{n}$ if X_t is 1-sub-Gaussian. Dudley's Inequality 5.24 is applicable to random walks on Hamming cube.

6 The Semigroup method

The semigroup method is a powerful method to prove interesting "isoperimetric-type" inequalities.

6.1 Basic definitions and set up

Definition 6.1 (Markov Process)**.** Consider X_t as a random process on time $T \subset \mathbb{R}$. A *Markov process* is a stochastic process with the property that

$$\mathbb{P}\left(X_{t_n} \leq x_n \,|\, X_{t_{n-1}}, X_{t_{n-2}}, \ldots, X_{t_1}\right) = \mathbb{P}\left(X_{t_n} \leq x_n \,|\, X_{t_{n-1}}\right),$$

where $x_n \in \mathbb{R}$ and $X_t \in \mathbb{R}$, i.e., this stochastic process "does not see the past".

Recall the conditional probability is $\mathbb{P}(A \,|\, B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$, and the conditional expectation $\mathbb{E}(X \,|\, Y)$ of random variables X and Y is also a random variable which is the best prediction of X given some behavior of Y. Suppose X has density function f_X and Y has density function f_Y. Then (X, Y) has joint density f_{XY}, we can define the conditional density as $f_{X\,|\,Y}(x \,|\, y) = \frac{f_{XY}(x,y)}{f_Y(y)}$, so $\mathbb{E}(X \,|\, Y = y) = \int_{-\infty}^{\infty} x \cdot f_{X\,|\,Y}(x \,|\, y)\, dx$ which is a number that depends on y. If let y be vary, then we get a random variable.

Definition 6.2 (Conditional Expectation)**.** Consider X on σ-algebra $\mathcal{F}_0$. Consider $\mathcal{F} \subset \mathcal{F}_0$ as another σ-algebra. The *conditional expectation of X with respect to $\mathcal{F}$*, $\mathbb{E}(X \,|\, \mathcal{F})$, is such an L_1 random variable/vector that $\mathbb{E}(X \,|\, \mathcal{F}) \in \mathcal{F}$, so that all events that relates to the random variable are sets of σ-algebra $\mathcal{F}$, i.e., $\{\mathbb{E}(X \,|\, \mathcal{F}) < t\} \in \mathcal{F}$. And for any event $A \in \mathcal{F}$,

$$\mathbb{E}(X \cdot \mathbb{1}_A) = \mathbb{E}(\mathbb{E}(X \,|\, \mathcal{F}) \cdot \mathbb{1}_A), \qquad \text{i.e., } \int_A X\, dP = \int_A \mathbb{E}(X \,|\, \mathcal{F})\, dP.$$

If Y is another random vector, then the conditional expectation of X with respect to Y is

$$\mathbb{E}(X \,|\, Y) := \mathbb{E}(X \,|\, \sigma(Y)),$$

where $\sigma(Y)$ is the σ-algebra generated by Y.

Definition 6.3 (Alternative Definition of Markov Process)**.** Consider $X_t \in \mathbb{R}^n$ with $t \geq 0$ as a random process. Assume for any bounded measurable function $f : \mathbb{R}^n \to \mathbb{R}$ and for any times $t, s > 0$, there is a bounded measurable function $P_s f : \mathbb{R}^n \to \mathbb{R}$ such that

$$\mathbb{E}(f(X_{t+s}) \,|\, \{X_z\}_{z \leq t}) = P_s f(X_t), \tag{61}$$

i.e., the behavior of X_{t+s} only depends on X_t and no earlier times.

Homework: find an example of a non-Markov process.

Definition 6.4 (Markov Semigroup)**.** Suppose P_s is an operator on bounded measurable functions such that

$$f : \mathbb{R}^n \to \mathbb{R} \quad \xrightarrow{P_s} \quad P_s f : \mathbb{R}^n \to \mathbb{R}$$

as defined in (61). P_s is what we called a *Markov semigroup*.

Definition 6.5 (Stationary measure of a Markov Process)**.** Consider X_t on $\mathbb{R}^n$ as a Markov process indexed by $T \subset \mathbb{R}^+$. A measure μ on $\mathbb{R}^n$ is called a *stationary measure of* X_t if for any bounded measurable function $f : \mathbb{R}^n \to \mathbb{R}$, we have

$$\int f \, d\mu = \int P_t \cdot f \, d\mu.$$

Here the semigroup $\{P_t f\}_{t \geq 0}$ is a collection of functions that describes some evolution of a function in time. We say the measure is stationary whatever the evolution it is, the average of the function does not change.

Remark 6.6. *Consider initial random vector $X_0 \sim \mu$ where μ is a stationary measure. Then*

$$\begin{aligned}
\mathbb{E}f(X_t) &= \mathbb{E}\left(\mathbb{E}\left(f(X_t) \mid X_0\right)\right) \\
&= \mathbb{E}P_t f(X_0) \\
&= \int P_t f \, d\mu \qquad (\textit{as } X_0 \sim \mu) \\
&= \int f \, d\mu \qquad (\textit{by stationary}) \\
&= \mathbb{E}f(X_0).
\end{aligned}$$

Hence, for any function f that has a stationary measure, we have $\mathbb{E}f(X_t) = \mathbb{E}f(X_0)$. If $f = \mathbb{1}_\Omega$, then $\mathbb{P}(X_t \in \Omega) = \mathbb{P}(X_0 \in \Omega)$. In other words, if $X_0 \sim \mu$ with stationary μ, then for any $t \geq 0$, $X_t \sim \mu$.

6.2 Properties of Markov semigroups, and some examples

Lemma 6.7. *Let μ be a stationary measure of a Markov process X_t indexed by $t \geq 0$. Then the following hold for all $p \geq 1$, $\alpha, \beta \in \mathbb{R}$, bounded measurable function f, g:*

1. *$\|P_t f\|_{L^p(\mu)} \leq \|f\|_{L^p(\mu)}$ (contraction).*
2. *P_t is a linear operator, i.e. $P_t(\alpha f + \beta g) = \alpha P_t f + \beta P_t g$ (linearity).*
3. *$P_{t+s} f = P_t P_s f \quad \mu$-a.s. (semigroup property).*
4. *$P_t 1 = 1 \quad \mu$-a.s. (conservativeness).*

Proof. 1. Suppose $X_0 \sim \mu$, we have

$$\int (P_t f)^p d\mu = \mathbb{E}_\mu(\mathbb{E}(f(X_t)|X_0)^p)) \leq \mathbb{E}_\mu(\mathbb{E}(f(X_t)^p|X_0))) = \int (f)^p d\mu,$$

where we have used Jensen's inequality.

2. Linearity follows similarly as

$$P_t(\alpha f + \beta g) = \mathbb{E}(\alpha f(X_t) + \beta g(X_t)|X_0) = \alpha\,\mathbb{E}(f(X_t)|X_0) + \beta\,\mathbb{E}(g(X_t)|X_0) = \alpha P_t f + \beta P_t g,$$

where we have used the linearity of conditional expectation.

3. For the semigroup property, we have

$$P_{t+s}f = \mathbb{E}(f(X_{t+s})|X_0) = \mathbb{E}(\mathbb{E}(f(X_{t+s})|\{X_r\}_{r\leq t})|X0) = \mathbb{E}(P_s f(X_t)|X_0) = P_t P_s f,$$

where the third equality is based on markovianity.

4. Conservativeness follows $P_s 1 = \mathbb{E}(1|X_0) = 1$.

□

Remark 6.8. *If X_t has a stationary measure μ, then the above lemma is true for all $f \in L^1(\mu)$, not only for bounded measurable functions f. From now on, we will assume the $P_t f$ is defined in this manner for every $f \in L^1(\mu)$.*

Remark 6.9. *Not every Markov process has a stationary measure; the questions of existence are complicated. We will consider concrete examples when things work well, the stationary measure exists and has nice properties, and we will explain the existence in these examples.*

Definition 6.10 (Variance)**.** Let μ be a stationary measure and $f \in L^2(\mu)$. The variance is defined as

$$\mathrm{Var}_\mu(f) := \int f^2 d\mu - \left(\int f d\mu\right)^2 = \mathbb{E}_\mu f^2 - (\mathbb{E}_\mu f)^2. \tag{62}$$

Note, by Cauchy's inequality, that $Var_\mu(f) \geq 0$.

Lemma 6.11. *Let μ be a stationary measure of a Markov process X_t and $f \in L^2(\mu)$. Then $\mathrm{Var}_\mu(f)$ decreases in $t \geq 0$*

Proof. Note that

$$\mathrm{Var}_\mu(P_t f) = \int (P_t f - \int P_t f d\mu)^2 d\mu = \int (P_t f - \int f d\mu)^2 d\mu = \int (P_t(f - \int f d\mu))^2 d\mu \tag{63}$$

$$\leq \int (f - \int f d\mu)^2 d\mu = \mathrm{Var}_\mu(f). \tag{64}$$

Here we use the definition of stationary measure for the second equality, linearity and conservativeness for the third equality and contraction for the fourth. Also we have

$$\int (P_t(f - \int f d\mu))^2 d\mu = \int (P_{t-s}P_s(f - \int f d\mu))^2 d\mu \leq \mathrm{Var}_\mu(P_s f) \tag{65}$$

Hence, we get $\mathrm{Var}_\mu(P_t f) \leq \mathrm{Var}_\mu(P_s f)$ for all $t \geq s \geq 0$. □

Remark 6.12. $\mathrm{Var}_\mu(f)$ *measures how far is* f *from a constant, as*

$$\mathrm{Var}_\mu(C) = \int C^2 d\mu - (\int C d\mu)^2 = C^2 - C^2 = 0.$$

Therefore, variance is a measure of the distance of the function to a constant function. The fact that it decreases along the semi-group means that the function $P_t f$ *becomes closer and closer to a constant function – namely, to the function* $\int f d\mu$ *(since this integral is preserved). Soon we will see that in some nice situations, not only does the variance decrease, but it decreases all the way to zero, and* $P_t f \underset{t\to\infty}{\to} \int f d\mu$*; however, this is not necessarily the case for an arbitrary Markov process.*

Definition 6.13 (Generator of Markov process)**.** Given a Markov process X_t with stationary measure μ on $\mathbb{R}^n$. For every $f \in L^2(\mu)$, the generator $\mathscr{L}$ is defined as

$$\mathscr{L} f := \lim_{t\to 0} \frac{P_t f - f}{t}. \tag{66}$$

Here, $\mathscr{L}$ is an operator on functions from $L^2(\mu)$ for which this limit makes sense. $\mathscr{L}$ is called the Generator associated with X_t.

Remark 6.14. $\mathscr{L}$ *is a linear operator, since* $P_t f$ *is linear.*

Remark 6.15 (Important)**.** *One can in fact define the Markov semigroup using a given linear operator* $\mathscr{L}$*. Indeed,* $\frac{d}{dt} P_t f = \lim_{\delta\to 0} \frac{P_{t+\delta} f - P_t f}{\delta} = \lim_{\delta\to 0} P_t(\frac{P_\delta f - f}{\delta}) = P_t \mathscr{L} f$*. Equivalently,* $\frac{d}{dt} P_t f = \lim_{\delta\to 0} \frac{P_\delta P_t f - P_t f}{\delta} = \mathscr{L} P_t f$*.*

Corollary 6.16.

$$\mathscr{L} P_t f = P_t \mathscr{L} f. \tag{67}$$

Consider the following PDE:

$$\begin{cases} \dfrac{d}{dt}(P_t f) = \mathscr{L}(P_t f), \\ P_0 f = f, \end{cases}$$

when $\mathscr{L}$ is fixed, $P_t f$ can then be defined as a solution of the above PDE. When the linear operator $\mathscr{L}$ is such that the PDE methods allow to conclude existence, probabilistic methods also provide a way of solving this PDE.

Example 6.17 (Finite state space)**.** *Let* $(X_t)_{t\in\mathbb{R}_+}$ *be a Markov process with values in a finite state space* $X_t \in \{1, \dots, d\}$*. Such processes are typically described in terms of their transition rates* $\lambda_{ij} \geq 0$ *for* $i \neq j$*:*

$$\mathbb{P}[X_{t+\delta} = j \mid X_t = i] = \lambda_{ij}\delta + o(\delta) \quad \text{for } i \neq j.$$

Evidently, the transition rates λ_{ij} *describe the infinitesimal rate of growth of the probability of jumping from state* i *to state* j *(informally if* $X_t = i$ *then the probability that* $X_{t+dt} = j$ *is* $\lambda_{ij} dt$*).*

Let us organize the transition probabilities $q_{tij} = \mathbb{P}[X_t = j \mid X_0 = i]$ and rates λ_{ij} into matrices $Q_t = (q_{tij})_{1\le i,j\le d}$ and $\Lambda = (\lambda_{ij})_{1\le i,j\le d}$ respectively where we define the diagonal entries of Λ as $\lambda_{ii} = -\sum_{j\neq i}\lambda_{ij} \le 0$. Then

$$\lim_{t\to 0}\frac{q_{tij} - q_{0ij}}{t} = \lambda_{ij}$$

for every $1 \le i, j \le d$ (the diagonal entries λ_{ii} were chosen precisely to enforce the law of total probability $\sum_j q_{tij} = 1$). In particular, we have

$$\mathscr{L}f(i) = \lim_{t\to 0}\frac{1}{t}\sum_{j=1}^{d}\lambda_{ij}f(j) = (\Lambda f)_i$$

where we identify the function f with the vector $(f(1), \dots, f(d)) \in \mathbb{R}^d$. We therefore conclude that the generator of a Markov process in a finite state space corresponds precisely to the matrix of transition rates. The Kolmogorov equation now reduces to the matrix differential equation (semigroup)

$$\frac{d}{dt}Q_t = Q_t\Lambda, \quad Q_0 = I.$$

This differential equation is the basic tool for computing probabilities of finite state space Markov processes. The solution is in fact easily obtained as $Q_t = e^{t\Lambda}$ from which we readily see why P_t and L must commute.

Example 6.18 (Heat semigroup)**.** *Suppose the generator $\mathscr{L} = \Delta$ be the Laplace operator and $u : \mathbb{R}^n \to R$.*

$$\Delta u = \sum_{i=1}^{n}\partial_{ii}u = \operatorname{tr}(\nabla^2 u). \tag{68}$$

Here, $\partial_{ii}u$ is the partial derivative and $\nabla^2 u$ is the Hessian matrix. We have Dom $= \overline{L^2(\mathbb{R}^n)\cap C^2(\mathbb{R}^n)}$*. Pick $f \in$* Dom*, the heat semigroup is defined as*

$$\begin{cases}\partial_t(P_t f) = \Delta(P_t f),\\ P_0 f = f.\end{cases}$$

There exists a solution and this defines a Markov process. Then, what is the invariant measure? We have the condition $\int_{\mathbb{R}^n} P_t f d\mu = \int_{\mathbb{R}^n} f d\mu$, i.e., $\frac{d}{dt}(\int_{\mathbb{R}^n} P_t f d\mu) = 0 = \int_{\mathbb{R}^n}\frac{d}{dt}(\int_{\mathbb{R}^n} P_t f d\mu = \int_{\mathbb{R}^n}\Delta(P_t f)d\mu$.

For Lebesgue measure on $\mathbb{R}^n$, for all $g \in C^2(\mathbb{R}^n)$ such that $\Delta g \in L^1(\mathbb{R}^n)$, we have $\int_{\mathbb{R}^n}\Delta g dx = 0$. Indeed, recall Green's formula: for measurable u, v, we have

$$\int u\Delta v dx = -\int\langle\nabla u, \nabla v\rangle dx.$$

By taking $u = 1$ and $\nabla u = 0$, we got $\int \nabla v dx = 0$. The conclusion is: Lebesgue measure is stationary for the heat semigroup.

Example 6.19 (Heat semigroup on the circle/torus)**.** *Consider a nice enough one dimensional function $f : [-\pi, \pi] \to \mathbb{R}$ (2π periodic). Then we have the PDE*

$$\begin{cases} \partial_t(P_t f) = \Delta(P_t f) \\ P_0 f = f \end{cases}$$

Then the stationary measure is uniform on circle.

Example 6.20 (Main example: Ornstein–Uhlenbeck semigroup)**.** *Firstly, Ornstein–Uhlenbeck operator is defined as a second order linear operator on nice enough function on $\mathbb{R}^n$ following:*

$$\mathscr{L}u = \Delta u - \langle \nabla u, x \rangle$$

Ornstein–Uhlenbeck semigroup is defined as: $\frac{d}{dt}(P_t f) = \mathscr{L}(P_t f)$, $P_0 f = f$. The question here is: What is the stationary measure for this process? Observe that

$$\int P_t f d\mu = \int f d\mu, \tag{69}$$

$$\frac{d}{dt} \int P_t f d\mu = 0, \tag{70}$$

$$\int \frac{d}{dt}(P_t f) d\mu = \int \mathscr{L}(P_t f) d\mu = 0. \tag{71}$$

The answer is the Gaussian measure!!!

$$d\gamma = \frac{1}{\sqrt{2\pi}^n} \exp -\frac{|x|^2}{2} dx, \tag{72}$$

$$\int u \mathscr{L} v d\gamma = -\int \langle \nabla u, \nabla v \rangle d\gamma, \tag{73}$$

for nice enough u, v. Plug in $u = 1$ and $\int \mathscr{L} v d\gamma = 0$, we have $\int \mathscr{L}(P_t f)) d\mu = 0$. Hence we got the conclusion: Gaussian measure is stationary for the Ornstein–Uhlenbeck process. The proof of 73 is left as homework (use first order gaussian integration by parts twice). It could also be proved by using Green's formula.

Definition 6.21 (Reversibility of semigroups)**.** A Markov semigroup P_t with stationary measure μ is called reversible if:

$$\int f \cdot P_t g \, d\mu = \int P_t f \cdot g \, d\mu$$

for all $f, g \in \mathrm{Dom}(P_t)$.

The name of reversibility indicates that if we assume our Markov process X_t is such that $X_0 \sim \mu$ as we usually do, then

$$\mathbb{E}(f(X_0)\, \mathbb{E}(g(X_t)|X_0)) = \mathbb{E}(g(X_t)\, \mathbb{E}(f(X_0)|X_t))$$

where $\mathbb{E}(f(X_0)\, \mathbb{E}(g(X_t)|X_0)) = \int f P_t g d\mu$, and $\mathbb{E}(g(X_t)\, \mathbb{E}(f(X_0)|X_t)) = \int P_t f g d\mu$. One can show that, *left as homework*, $P_t f(x) = \mathbb{E}(f(X_t)|X_0 = x) = \mathbb{E}(f(X_0)|X_t = x)$. That is to say, time goes in both directions in the same way.

Definition 6.22 (Ergodicity)**.** A Markov semigroup semigroup P_t is called ergodic if for all $f \in \mathrm{Dom}(P_t)$,

$$P_t f \xrightarrow[t\to\infty]{L_2} \mathbb{E}_\mu f = \int f d\mu$$

Here, $P_t f \xrightarrow[t\to\infty]{L_2} = C = \mathbb{E}_\mu f$, since $\int P_t f d\mu = \int f d\mu$ for all t. In particular, $\int c d\mu = \int P_\infty f d\mu = \int f d\mu$.

Remark 6.23. *Recall that for all Markov semigroup, the* $\mathrm{Var}(P_t f)$ *decreases as* $t \to \infty$*. If* P_t *is ergodic, that is equivalent to* $\mathrm{Var}(P_t f) \to 0$*.*

Definition 6.24 (Dirichlet form of semigroups)**.** Assume X_t is a Markov process with a stationary measure μ and generator $\mathscr{L}$, i.e., $\partial_t(P_t f) = \mathscr{L} P_t f$, and $P_0 f = f$. The Dirichlet form $\mathcal{E}$ is defined by:

$$\mathcal{E}(f, g) := -\int f \mathscr{L} g \, d\mu$$

Example 6.25 (Dirichlet form of Heat semigroup)**.** *Recall that Heat semigroup is* $\mathscr{L} = \nabla$ *on* $\mathbb{R}^n$*, then* $\partial_t(P_t f) = \nabla(P_t f)$ *and* $P_0 f = f$*, and the stationary measure* μ *is the Lebesgue measure because it satisfies the integral by parts:* $\int u \Delta v dx = -\int \langle \nabla u, \nabla v \rangle dx$*, in particular,* $\int \nabla v dx = 0$*, then we have:*

$$\mathcal{E}(f, g) = -\int f \nabla g dx = \int \langle \nabla u, \nabla v \rangle dx$$

Remark 6.26. *Notice that one can show the Heat semigroup is reversible: given* f *and* g*:* $\int P_t f g d\mu = \int g P_t f d\mu$*.*

Proof left as homework.

Theorem 6.27 (Abstract Theorem about "Poincare inequality")**.** P_t *is a reversible and ergodic markov semigroup with stationary measure* μ*, fix a constant* $c \geq 0$*, the followings are all equivalent:*

1. *For all* $f \in \mathrm{Dom}(P_t)$, $\mathrm{Var}_\mu(f) \leq c\mathcal{E}(f, f)$ *(Poincare inequality).*
2. P_t *is "Hypercontractive". That is,*

$$\int_{\mathbb{R}^n} \left(P_t f - \int_{\mathbb{R}^n} f d\mu \right)^2 d\mu \leq e^{-\frac{2t}{c}} \int_{\mathbb{R}^n} \left(f - \int_{\mathbb{R}^n} f d\mu \right)^2 d\mu$$

   for all $t \geq 0$ *and* $f \in \mathrm{Dom}(P_t)$*. This is the same as* $\mathrm{Var}_\mu(P_t f) \leq e^{-\frac{2t}{c}} \mathrm{Var}_\mu(f)$*.*

3. $\mathcal{E}(P_t f, P_t f) \leq e^{-\frac{2t}{c}} \mathcal{E}(f, f)$*.*
4. *For all* f*, there exists a constant* $\kappa(f)$ *such that* $\sqrt{\mathrm{Var}_\mu(P_t f)} \leq \kappa(f) e^{-\frac{t}{c}}$*.*
5. *for all* f*, there exists a constant* $k(f)$ *such that* $\mathcal{E}(P_t f, P_t f) \leq k(f) e^{-\frac{2t}{c}}$

Before we complete the proof, we discuss the most important example in detail, and then the ideas of the proof will become clearer.

6.3 The Ornstein–Uhlenbeck semigroup

Recall $\mathscr{L}u = \Delta u - \langle \nabla u, x \rangle$, where $u \colon \mathbb{R}^n \to \mathbb{R}$ is an appropriate function here. The semigroup P_t is defined as: $\partial_t(P_t f) = \mathscr{L}(P_t f)$, $P_0 f = f$, given $f \in \mathrm{Dom}(P_t)$.

Lemma 6.28. *A very nice and concrete representation for the Ornstein-Uhlenbeck semigroup:*

1. *$P_t f(x) = \mathbb{E}\, f(e^{-t}x + \sqrt{1-e^{-2t}}Z)$, where $Z \sim \mathcal{N}(0, \mathrm{Id})$ is the Ornstein-Uhlenbeck semigroup, such that it satisfies $\partial_t(P_t f) = \mathscr{L}(P_t f)$, $P_0 f = f$, which is true because we can check: when $t = 0$, $P_0 f = \mathbb{E}\, f(x)$, when $t \to \infty$, $P_\infty f = \mathbb{E}\, f(z) = \int f d\gamma$.*
2. *P_t is ergodic.*
3. *P_t is reversible.*
4. *γ is the stationary measure.*

Claim 6.29. *This claim of second order integration by parts is used in the following proof. For all f, g,*

$$\int g \mathscr{L} f \, d\gamma = -\int \langle \nabla f, \nabla g \rangle \, d\gamma.$$

In particular, the Dirichlet form is given by $\mathcal{E}(f,g) = -\int \langle \nabla f, \nabla g \rangle d\gamma$.

Proof of Claim 6.29. By Green's formula, we have

$$\begin{aligned}
\int_{\mathbb{R}^n} g \Delta f \, d\gamma &= c_n \int_{\mathbb{R}^n} \Delta f \cdot g e^{\frac{-|x|^2}{2}} \, dx \\
&= -c_n \int_{\mathbb{R}^n} \left\langle \nabla f, \nabla \left(g e^{\frac{-|x|^2}{2}} \right) \right\rangle dx \\
&= -\int_{\mathbb{R}^n} \langle \nabla f, \nabla g \rangle c_n e^{\frac{-|x|^2}{2}} dx + \int_{\mathbb{R}^n} g c_n \langle \nabla f, x \rangle e^{\frac{-|x|^2}{2}} \, dx \\
&= -\int_{\mathbb{R}^n} \langle \nabla f, \nabla g \rangle d\gamma + \int_{\mathbb{R}^n} g \langle \nabla f, x \rangle \, d\gamma.
\end{aligned}$$

Therefore,

$$\int_{\mathbb{R}^n} g \mathscr{L} f \, d\gamma = \int_{\mathbb{R}^n} g \Delta f \, d\gamma - \int_{\mathbb{R}^n} g \langle \nabla f, x \rangle \, d\gamma = -\int_{\mathbb{R}^n} \langle \nabla f, \nabla g \rangle \, d\gamma.$$

□

Proof of Lemma 6.28. **Proof of Property 4.** Let $f \in \mathrm{Dom}(P_t)$. Then,

$$\frac{d}{dt} \int P_t f d\gamma = \int \partial_t \left(P_t f \right) d\gamma = \int \mathscr{L}(P_t f) \, d\gamma.$$

Here, if we let $g = 1$, then

$$\int \mathscr{L} f d\gamma = -\int \langle \nabla f, 0 \rangle d\gamma = 0.$$

Thus, $\frac{d}{dt} \int P_t f d\gamma = 0$, implying that $\int P_t f d\gamma = \int f d\gamma$ and P_t is constant in time.

Proof of Property 1. We want to show that $P_t f(x) = \mathbb{E}\, f(e^{-t}x + \sqrt{1-e^{-2t}}z)$ satisfies the following

$$\begin{cases} \partial_t (P_t f) = \mathscr{L}(P_t f) \\ P_0 t = f. \end{cases}$$

By direct calculation, we see that $P_0 f(x) = \mathbb{E}\, f(x) = f(x)$. Next, by the chain rule, we have

$$\begin{aligned} \partial_t (P_t f) &= \partial_t \, \mathbb{E}\, f(e^{-t}x + \sqrt{1-e^{-2t}}Z) \\ &= \mathbb{E}\, \partial_t f(e^{-t}x + \sqrt{1-e^{-2t}}Z) \\ &= \mathbb{E} \left\langle \nabla f(e^{-t}x + \sqrt{1-e^{-2t}}Z), -e^{-t}x + (1-e^{-2t})^{-1/2} e^{-2t} Z \right\rangle \\ &= \int_{\mathbb{R}^n} \left\langle \nabla f(e^{-t}x + \sqrt{1-e^{-2t}}z), -e^{-t}x + (1-e^{-2t})^{-1/2} e^{-2t} z \right\rangle d\gamma(z). \end{aligned}$$

Using Claim 6.29, we see that

$$\begin{aligned} &\int_{\mathbb{R}^n} \left\langle \nabla f(e^{-t}x + \sqrt{1-e^{-2t}}z), -e^{-t}x + (1-e^{-2t})^{-1/2} e^{-2t} z \right\rangle d\gamma(z) \\ &= \int_{\mathbb{R}^n} \left\langle \nabla f(e^{-t}x + \sqrt{1-e^{-2t}}z), -e^{-t}x \right\rangle d\gamma(z) + \int_{\mathbb{R}^n} e^{-2t} \Delta f(e^{-t}x + \sqrt{1-e^{-2t}}z)\, d\gamma(z) \\ &= -\left\langle \nabla_x \left[\int_{\mathbb{R}^n} f(e^{-t}x + \sqrt{1-e^{-2t}}z) d\gamma(z) \right], x \right\rangle + \Delta_x \left[\int_{\mathbb{R}^n} f(e^{-t}x + \sqrt{1-e^{2t}}z)\, d\gamma(z) \right] \\ &= \mathscr{L}\, \mathbb{E}\, f(e^{-t}x + \sqrt{1-e^{-2t}}Z). \end{aligned}$$

Proof of Property 3. For $t = 0$, note that

$$\int_{\mathbb{R}^n} f(P_0 g)\, d\gamma = \int_{\mathbb{R}^n} fg\, d\gamma = \int_{\mathbb{R}^n} g(P_0 f)\, d\gamma$$

Furthermore, applying Gaussian integration by parts, we see that

$$\frac{d}{dt} \int_{\mathbb{R}^n} f P_t g\, d\gamma = \int_{\mathbb{R}^n} f \mathscr{L}(P_t g)\, d\gamma = -\int_{\mathbb{R}^n} \langle \nabla f, \nabla (P_t g) \rangle d\gamma.$$

Similarly,

$$\frac{d}{dt} \int g P_t f d\gamma = -\int \langle \nabla g, \nabla P_t f \rangle\, d\gamma$$

Now, via a change of variable, we see that

$$\begin{aligned}\int_{\mathbb{R}^n} \langle \nabla f, \nabla (P_t g)\rangle \, d\gamma &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-t}\langle \nabla f(x), \nabla g(e^{-t}x + \sqrt{1-e^{-2t}}z)\rangle \, d\gamma(z) d\gamma(x) \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} c_n \frac{e^{-t}}{\sqrt{1-e^{-2t}}} e^{-\frac{(\xi - e^{-t}x)^2}{2(1-e^{-2t})}} \langle \nabla f(x), \nabla g(\xi)\rangle \, d\xi \, d\gamma(x) \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-t}\langle \nabla g(x), \nabla f(e^{-t}x + \sqrt{1-e^{-2t}}z)\rangle \, d\gamma(z) d\gamma(x) \\ &= \int_{\mathbb{R}^n} \langle \nabla g, \nabla (P_t f)\rangle \, d\gamma\end{aligned}$$

Hence,

$$\langle f, P_t g\rangle_{L^2(\gamma)} = \int_0^t \frac{d}{d\tau}\left[\langle f, P_\tau g\rangle_{L^2(\gamma)}\right] d\tau = \int_0^t \frac{d}{d\tau}\left[\langle g, P_\tau f\rangle_{L^2(\gamma)}\right] d\tau = \langle P_t f, g\rangle_{L^2(\gamma)}.$$

Therefore, P_t follows accordingly.

Proof of Property 2. By the Bounded Convergence Theorem, note that

$$P_t f = \mathbb{E}\, f(e^{-t} + \sqrt{1-e^{-2t}}Z) \xrightarrow[t\to\infty]{L_2} \mathbb{E}\, f(Z).$$

□

Remark 6.30. *For the students who have background in stochastic partial differential equation, when X_t is a Markov process, $dX_t = -X_t dt + \sqrt{2}dB_t$ where B_t is Brownian motion. Then X_t is the Ornstein-Uhlenbeck semigroup.*

Lemma 6.31 (Hypercontactivity)**.** *Let $\{P_t\}$ be the Ornstein-Ulhenbeck semigroup. Then, the following hold:*

1. *$\nabla P_t f = e^{-t} P_t \nabla f$, where $P_t \nabla f = (P_t \partial_1 f, \ldots, P_t \partial_n f)$. Here the $\partial_i f$ denote the partial derivatives of f and $\nabla f = (\partial_1 f, \ldots, \partial_n f)$.*
2. *$\int |\nabla P_t f|^2 \, d\gamma \le e^{-2t} \int |\nabla f|^{-2} \, d\gamma$.*

Proof. 1. Recall that $P_t f = \mathbb{E}\, f(e^{-t}x + \sqrt{1-e^{-2t}}Z)$. Then, the chain rule yields that

$$\partial_i (P_t f) = \mathbb{E}\, \partial_{x^i} f\left(e^{-t}x + \sqrt{1-e^{-2t}}Z\right) = e^{-t}\,\mathbb{E}\, \partial_i f\left(e^{-t}x + \sqrt{1-e^{-2t}}Z\right) = e^{-t} P(\partial_i f).$$

2. Note that

$$\begin{aligned}\int_{\mathbb{R}^n} |\nabla P_t f|^2 \, d\gamma = e^{-2t}\int_{\mathbb{R}^n} \sum_{1\le i\le n} |P_t \partial_i f|^2 \, d\gamma &= e^{-2t} \sum_{1\le i\le n} \int |P_t \partial_i f|^2 \, d\gamma \\ &\le e^{-2t} \sum_{1\le i\le n} \int_{\mathbb{R}^n} |\partial_i f|^2 \, d\gamma \\ &= e^{-2t} \int_{\mathbb{R}^n} |\nabla f|^2 \, d\gamma,\end{aligned}$$

where we have used $\int |P_t g|^2 d\gamma \le \int g^2 d\gamma$. □

6.4 Gaussian Poincare inequality via the semigroup method

Theorem 6.32 (Gaussian Poincaré Inequality)**.** *Let $f \in \mathscr{L}^2(\mathbb{R}^n) \cap \mathcal{C}^2(\mathbb{R}^n)$. Then,*

$$\mathrm{Var}_\gamma(f) = \int_{\mathbb{R}^n} f^2 d\gamma - \left(\int_{\mathbb{R}^n} f \, d\gamma \right)^2 \leq \mathbb{E}_\gamma \left| \nabla f \right|^2 = \int_{\mathbb{R}^n} \left| \nabla f \right|^2 \, d\gamma,$$

where γ is Gaussian measure.

Proof. Let P_t be the Ornstein-Uhlenbeck semigroup. Then, note that $\int_{\mathbb{R}^n} f^2 d\gamma = \int (P_0 f)^2 d\gamma$ and $\left(\int_{\mathbb{R}^n} f d\gamma \right)^2 = (P_\infty f)^2$ since ergodicity of $\{P_t\}$ implies that $P_\infty f = \int f \, d\gamma$. Since γ is a probability measure, the Fundamental Theorem of Calculus implies that

$$\begin{aligned}
\mathrm{Var}_\gamma(f) = \int_{\mathbb{R}^n} (P_0 f)^2 d\gamma - \int_{\mathbb{R}^n} (P_\infty f)^2 \, d\gamma &= \int_{\mathbb{R}^n} (P_0 f)^2 - (P_\infty f)^2 \, d\gamma \\
&= - \int_{\mathbb{R}^n} \int_0^\infty \frac{\partial}{\partial t} (P_t f)^2 \, dt \, d\gamma(x) \\
&= - \int_0^\infty \int_{\mathbb{R}^n} \frac{\partial}{\partial t} (P_t f)^2 \, d\gamma(x) \, dt \\
&= - \int_0^\infty \int_{\mathbb{R}^n} 2 P_t f \cdot \mathscr{L}(P_t f) \, d\gamma(x) \, dt \\
&= - \int_0^\infty \int_{\mathbb{R}^n} 2 P_t f \cdot \mathscr{L}(P_t f) \, d\gamma(x) \, dt
\end{aligned}$$

Applying Gaussian integration by parts, we obtain

$$\begin{aligned}
-2 \int_0^\infty \int_{\mathbb{R}^n} P_t f \cdot \mathscr{L}(P_t f) \, d\gamma(x) \, dt &= 2 \int_0^\infty \int_{\mathbb{R}^n} \langle \nabla P_t f, \nabla P_t f \rangle \, d\gamma \, dt \\
&= 2 \int_0^\infty \int_{\mathbb{R}^n} e^{-2t} \left| P_t \nabla f \right|^2 \, d\gamma \, dt \\
&\leq 2 \int_0^\infty \int_{\mathbb{R}^n} e^{-2t} \left| \nabla f \right|^2 \, d\gamma \, dt \\
&= \int_{\mathbb{R}^n} \left| \nabla f \right|^2 \, d\gamma.
\end{aligned}$$

□

Remark 6.33. *When is the Gaussian Poincaré inequality sharp? - It is when we use the hypercontactivity. The step we used $\int |P_t g|^2 \leq \int g^2 d\gamma$ was the only step we used the inequality. In hypercontactivity, the equality is achieved when g is a constant, that is to say, $P_t g = g$ is constant. That is, when all of $\partial_i f$ in Gaussian Poincaré inequality are constant or f is a linear function. $f(x) = \langle x, \theta \rangle$ where $\theta \in \mathbb{R}^n$.*

Indeed,

$$\mathrm{Var}_\gamma \langle x, \theta \rangle = \sum \theta_i^2 \, \mathrm{Var}_\gamma(x_i) = |\theta|^2 = \int_{\mathbb{R}^n} |\nabla \langle x, \theta \rangle|^2 \, d\gamma.$$

6.5 A discussion on Poincare inequalities, the example of the circle and periodic functions

More generally, let $d\mu = e^{-V(x)}\,dx$ for $V \in \mathcal{C}^2(\mathbb{R}^n)$ such that $\int d\mu = 1$. Then we define the operator $\mathscr{L}_\mu u = \Delta u - \langle \nabla V, \nabla u \rangle$. In the case of the gaussian measure $d\gamma = e^{-|x|^2/2}\,dx$ so $V(x) = |x|^2/2$, $\nabla V = x$ and $\mathscr{L}_\gamma = \Delta u - \langle x, \nabla u \rangle$, which is exactly the Ornstein-Uhlenbeck operator. For a given measure μ we have the integration by parts formula

$$\int f \cdot L_\mu g \, d\mu = -\int \langle \nabla f, \nabla g \rangle \, d\mu.$$

Definition 6.34. A function f is the *first eigenfunction* of $\mathscr{L}_\mu$ if

$$\mathscr{L}_\mu f = -\lambda f \tag{$\star$}$$

and $\lambda > 0$ is the smallest number such that there is f satisfying $(\star)$.

The number λ is called the *first eigenvalue* of $\mathscr{L}_\mu$.

Definition 6.35. A number $c_p(\mu) > 0$ is called the Poincaré constant of μ if it is the smallest number such that

$$\int f^2 \, d\mu - \left(\int f \, d\mu \right)^2 \leq c_p \cdot \int |\nabla f|^2 \, d\mu.$$

for all reasonable functions f.

Remark 6.36. *For the Gaussian measure, the Poincaré inequality (Theorem 6.32) implies that $c_p(\gamma) = 1$.*

Claim 6.37. *For a measure μ $c_p(\mu) = \frac{1}{\lambda}$ where λ is the first eigenvalue of $\mathscr{L}_\mu$.*

Remark 6.38. *The first eigenfunction may not exist, for example*

- *when $V(t) = |t|$, or*
- *for the Lebesgue measure.*

The above claim also shows that linear functions are the first eigenfunctions of the Ornstein-Uhlenbeck operator.

It is also possible to use harmonic analysis to prove the Gaussian Poincaré inequality. We will illustrate this with a simpler example. Consider 2π-periodic functions with two continuous derivatives which are themselves also 2π-periodic.

Theorem 6.39. *Let f be a 2π-periodic function with two continuous and 2π-periodic derivatives. Then*

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} f^2(\theta) \, d\theta - \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) \, d\theta \right)^2 \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} \dot{f}(\theta)^2 \, d\theta,$$

where $\dot{f}$ is the derivative of f.

Proof. From basic Fourier analysis we can write

$$f(\theta) = \sum_{i=-\infty}^{\infty} \hat{f}(i) e^{-2\pi i \theta}.$$

The Plancherel identity states that

$$\frac{1}{2\pi} \int f^2 \, d\theta = \sum_{i=-\infty}^{\infty} \hat{f}(i)^2,$$

and we also know that

$$\frac{1}{2\pi} \int f \, d\theta = \hat{f}(0).$$

With these facts we can evaluate the variance of the function f

$$\frac{1}{2\pi} \int f^2 \, d\theta - \left(\frac{1}{2\pi} \int f \, d\theta \right)^2 = \sum_{i=-\infty}^{\infty} \hat{f}(i)^2 - \hat{f}(0)^2 = \sum_{i \neq 0} \hat{f}(i)^2.$$

On the other hand, by differentiating, we know that

$$\dot{f}(\theta) = \sum_{i=-\infty}^{\infty} i \hat{f}(i) e^{-2\pi i},$$

so applying the Plancherel identity we see that

$$\frac{1}{2\pi} \int \dot{f}(\theta) \, d\theta = \sum_{i=-\infty}^{\infty} i^2 \hat{f}(i) \geq \sum_{i \neq 0} \hat{f}(i),$$

where the final inequality is true term-by-term. □

Remark 6.40. *If $f : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz then $\mathrm{Var}_\gamma(f) \leq 1$, since f is Lipschitz if and only if $|\nabla f| \leq 1$.*

6.6 Proof of the abstract Poincare-type inequality for stationary measures of semigroups

Now we go back to our general Markov process and see how much of the previous theory carries over.

Theorem 6.41. *Let P_t be a reversible ergodic Markov semigroup with stationary measure μ. The following are equivalent for a given constant $c > 0$.*

1. *$\mathrm{Var}_\mu(f) \leq c \cdot \mathcal{E}(f, f)$ for all reasonable f,*

2. $\left\|P_t f - \int (P_t f)\, d\mu\right\|_{L^2(\mu)} \le e^{-t/c} \left\|f - \int f\, d\mu\right\|_{L^2(\mu)}$ *for all* $t \ge 0$ *and* $f \in L^2$,

3. $\mathcal{E}(P_t f, P_t f) \le e^{-2t/c}\mathcal{E}(f,f)$ *for all* $t \ge 0$ *and* $f \in L^2$,

4. *for all* $f \in L^2$ *there is a constant* $\kappa(f)$ *such that* $\left\|P_t f - \int (P_t f)\, d\mu\right\|_{L^2(\mu)} \le \kappa(f) \cdot e^{-t/c}$ *for all* $t \ge 0$,

5. *for all* $f \in L^2$ *there is a constant* $\kappa(f)$ *such that* $\mathcal{E}(P_t f, P_t f) \le \kappa(f) \cdot e^{-2t/c}$ *for all* $t \ge 0$.

So far, for the Gaussian measure, we have used the implication 3. $\Rightarrow$ 1 to prove the Gaussian Poincaré inequality (Theorem 6.32). It is important to note that the above result is not strictly speaking a generalization of the Poincaré inequality. It is only the equivalence of several conditions, and we had to do work by hand to show hyperconstractivity (condition 3.) in the case of the Gaussian measure.

Lemma 6.42. *Let* P_t *be a semigroup satisfying the conditions of Theorem 6.41. Then*

$$\frac{d}{dt}\operatorname{Var}_\mu(P_t f) = -2\mathcal{E}(P_t f, P_t f)$$

for all reasonable functions f.

Proof. Recall

$$\operatorname{Var}_\mu(f) = \int (P_t f)^2\, d\mu - \left(\int P_t f\, d\mu\right)^2.$$

Since the chain is ergodic, the second term is constant, so

$$\begin{aligned}
\frac{d}{dt}\operatorname{Var}_\mu(f) &= \frac{d}{dt}\int (P_t f)^2\, d\mu \\
&= \int \frac{d}{dt}(P_t f)^2\, d\mu \\
&= 2\int (P_t f)\frac{d}{dt}(P_t f)\, d\mu \\
&= 2\int P_t f \mathscr{L}(P_t f)\, d\mu \\
&= -2\mathcal{E}(P_t f, P_t f).
\end{aligned}$$

□

Remark 6.43. *Recall that* $\operatorname{Var}_\mu(P_t f)$ *decreases to zero as* $t \to \infty$ *and so the above implies that* $\mathcal{E}(g,g) \ge 0$ *for all reasonable* g.

Corollary 6.44 (Integral representation of variance)**.** *Let $\{P_t\}$ be a semigroup satisfying the conditions of Theorem 6.41. Then*

$$\mathrm{Var}_\mu(f) = 2\int_0^\infty \mathcal{E}(P_t f, P_t f)\,dt.$$

Proof. We can make a direct calculation, using ergodicity and the above lemma

$$\begin{aligned}\int (P_t f)^2\,d\mu - \left(\int P_t f\,d\mu\right)^2 &= \int (P_0 f)^2\,d\mu - (P_\infty f)^2 \\ &= \int (P_0 f)^2 - (P_\infty f)^2\,d\mu \\ &= -\int\int_0^\infty \frac{d}{dt}(P_t f)^2\,dt\,d\mu \\ &= 2\int_0^\infty \mathcal{E}(P_t f, P_t f)\,dt.\end{aligned}$$

□

With these lemmas, we can prove some of the implications required for Theorem 6.41. For now we will not use the reversibility assumption.

Proof of Theorem 6.41 Part 1. In this part of the proof we will show

$$5 \Leftarrow 3 \Rightarrow 1 \Leftrightarrow 2 \Rightarrow 4,$$

without using reversibility.
$(2 \Rightarrow 4)$: This implication is immediate by setting $\kappa(f) = \left\|f - \int f\,d\mu\right\|_{L^2(\mu)}$.
$(3 \Rightarrow 5)$: This time setting $\kappa(f) = \mathcal{E}(f,f)$ is sufficient.
$(3 \Rightarrow 1)$: We can make a direct calculation using Corollary 6.44

$$\begin{aligned}\mathrm{Var}_\mu(f) &= 2\int_0^\infty \mathcal{E}(P_t f, P_t f)\,dt \\ &\le 2\int_0^\infty e^{-2t/c}\mathcal{E}(f,f)\,dt \\ &\le c\cdot\mathcal{E}(f,f).\end{aligned}$$

$(1 \Rightarrow 2)$: Suppose that $\mathrm{Var}_\mu(f) \le c\mathcal{E}(f,f)$ for all reasonable functions f. Then, by Lemma 6.42

$$\frac{d}{dt}\mathrm{Var}_\mu(P_t f) = -2\mathcal{E}(P_t f, P_t f) \le -\frac{2}{c}\mathrm{Var}_\mu(P_t f).$$

We can now apply Grönwall's inequality viewing $\mathrm{Var}_\mu(P_t f)$ as a function of t to deduce that

$$\mathrm{Var}_\mu(P_t f) \le e^{-2t/c}\,\mathrm{Var}_\mu(f).$$

By taking the square root, we see that

$$\left\| P_t f - \int P_t f \, d\mu \right\|_{L^2(\mu)} \leq e^{-t/c} \left\| f - \int f \, d\mu \right\|_{L^2(\mu)}.$$

$(2 \Rightarrow 1)$: We will use Lemma 6.42 in reverse and expand the definition of derivative.

$$\begin{aligned} 2\mathcal{E}(f,f) &= \lim_{t \to 0} \frac{\mathrm{Var}_\mu(f) - \mathrm{Var}_\mu(P_t f)}{t} \\ &\leq \mathrm{Var}_\mu(f) \cdot \lim_{t \to \infty} \frac{1 - e^{-2t/c}}{t} = \frac{2}{c} \mathrm{Var}_\mu(f), \end{aligned}$$

where the first inequality follows from the assumption 2.. Rearranging the constants, 1 follows. □

To prove the implication $1 \Rightarrow 2$ we used Grönwall's inequality.

Lemma 6.45 (Grönwall's inequality)**.** *Let $g : \mathbb{R} \to \mathbb{R}$ be differentiable. If*

$$\frac{d}{dt} g(t) \leq c \cdot g(t),$$

then

$$g(t) \leq e^{ct} g(0).$$

Proof. Rearranging the assumption we see that

$$\frac{d}{dt} \log(g(t)) = \frac{\frac{d}{dt} g(t)}{g(t)} \leq c.$$

Now we can integrate to see that

$$\log(g(t)) - \log(g(0)) \leq c \cdot t,$$

and taking the exponential implies that

$$g(t) \leq e^{ct} g(0).$$

□

Next, we will suppose that the chain P_t is reversible.

Lemma 6.46. *If $\{P_t\}$ is a reversible Markov semigroup, then for all reasonable functions f*

- $\log \mathrm{Var}_\mu(P_t f)$ *is convex in $t \geq 0$, and*
- $\log \mathcal{E}(P_t f, P_t f)$ *is convex in $t \geq 0$.*

Recall that g is called convex if for all t, s and all $\lambda \in [0,1]$

$$g(\lambda t + (1-\lambda)s) \leq \lambda g(t) + (1-\lambda)g(s),$$

and that if $g \in \mathcal{C}^2$ then this is equivalent to $g'' \geq 0$.

Proof. We establish convexity by computing the second derivative, and proving it is positive. Fix $t \geq 0$. We thus compute:

$$\begin{aligned}\frac{d}{dt}\log \mathrm{Var}_\mu(P_t f) &= \frac{\frac{d}{dt}\mathrm{Var}_\mu(P_t f)}{\mathrm{Var}_\mu(P_t f)} \\ &= -\frac{2\mathcal{E}(P_t f, P_t f)}{\mathrm{Var}_\mu(P_t f)},\end{aligned}$$

and therefore,

$$\begin{aligned}\frac{d^2}{dt^2}\log \mathrm{Var}_\mu(P_t f) &= \frac{d}{dt}\left(-\frac{2\mathcal{E}(P_t f, P_t f)}{\mathrm{Var}_\mu(P_t f)}\right) \\ &= -\frac{2\frac{d}{dt}\mathcal{E}(P_t f, P_t f)}{\mathrm{Var}_\mu(P_t f)} - \frac{4\mathcal{E}(P_t f, P_t f)^2}{\mathrm{Var}_\mu(P_t f)^2}.\end{aligned} \tag{74}$$

Now, by definition of our Dirichlet form,

$$\begin{aligned}\frac{d}{dt}\mathcal{E}(P_t f, P_t f) &= \mathcal{E}\left(\frac{d}{dt}P_t f, P_t f\right) + \mathcal{E}\left(P_t f, \frac{d}{dt}P_t f\right) \text{ by bilinearity of } \mathcal{E}, \\ &= 2\mathcal{E}(\frac{d}{dt}P_t f, P_t f) \text{ since } \mathcal{E} \text{ is symmetric}, \\ &= 2\mathcal{E}\left(\mathscr{L}P_t f, P_t f\right) \text{ since } \frac{d}{t}P_t f = \mathscr{L}f, \\ &= -2\int (\mathscr{L}P_t f)^2 \, d\mu \text{ by definition of } \mathcal{E}.\end{aligned} \tag{75}$$

Hence, inserting Equation (75) in Equation (74)we obtain:

$$\begin{aligned}\frac{d^2}{dt^2}\log \mathrm{Var}_\mu(P_t f) &= \frac{4}{\mathrm{Var}_\mu(P_t f)^2}\left(\int (\mathscr{L}P_t f)^2 \,\mathrm{Var}_\mu(P_t f) - \mathcal{E}(P_t f, P_t f)^2\right) \\ &= \frac{4}{\mathrm{Var}_\mu(P_t f)^2}\left(\int (\mathscr{L}P_t f)^2 \,\mathrm{Var}_\mu(P_t f) - \int P_t f \mathscr{L}P_t f \, d\mu)^2\right) \\ &\geq 0\end{aligned}$$

since, by Cauchy-Schwarz,

$$\int (\mathscr{L}P_t f)\, d\mu \leq \int \left(\mathscr{L}(P_t f)^2 \, d\mu\right) \cdot \underbrace{\int (P_t f)^2 \, d\mu}_{=\mathrm{Var}_\mu(\mathbb{P}_t f)}.$$

This completes the proof of convexity of $t \in [0,\infty) \mapsto \log \mathrm{Var}_\mu(P_t f)$. □

6.7 The Gaussian Log-Sobolev Inequality of Gross via the semigroup method

In this section, we will establish the log Sobolev inequality via properties of the Orstein-Uhlenbeck semigroup. To start, we recall that some basic properties of this semigroup:

- for all $t \geq 0$, $P_t f = \mathbb{E}(f(e^{-t}x + \sqrt{1-e^{2t}}Z)$ where $Z \sim \mathcal{N}(0, I_n)$; in particular, $P_0 f = f$,
- $d\gamma = (2\pi)^{-\frac{n}{2}} \exp\left(-\frac{|x|^2}{2}\right) dx$ — the Gaussian probability measure is stationary,
- $\frac{d}{dt} P_t f = \mathscr{L} P_t f$, where
$$\mathscr{L} u = \Delta u - \langle x, \nabla u \rangle.$$
In particular,
$$\int_{\mathbb{R}^n} f \mathscr{L} g \, d\gamma = -\int_{\mathbb{R}^n} \langle \nabla f, \nabla g \rangle \, d\gamma.$$

Furthermore, note that this process is reversible and ergodic:

$$\lim_{t\to\infty} P_t f = \int_{\mathbb{R}^n} f \, d\gamma = \mathbb{E}(f(Z)).$$

Theorem 6.47 (Gross's Log-Sobolev Inequality)**.** *Suppose that $f \geq 0$ and*

$$\int_{\mathbb{R}^n} f \log f \, d\gamma - \int_{\mathbb{R}^n} f \, d\gamma \cdot \log\left(\int_{\mathbb{R}^n} f d\,\gamma\right) \leq \frac{1}{2} \int_{\mathbb{R}^n} \frac{|\nabla f|^2}{f} \, d\gamma$$

To establish this inequality, we will reproduce the proof of the Gaussian-Poincaré inequality in a more general setting. Let $\phi\colon \mathbb{R} \to \mathbb{R}$ be a non-negative function with $\phi \in C^2(\mathbb{R})$. As in the proof of Theorem 6.32, we compute

$$\begin{aligned}
&\int_{\mathbb{R}^n} \phi(f) \, d\gamma - \phi\left(\int_{\mathbb{R}^n} f \, d\gamma\right) \\
&= \int_{\mathbb{R}^n} \phi\left(P_0 f\right) - \phi\left(P_\infty f\right) \, d\gamma \\
&= -\int_{\mathbb{R}^n} \int_0^\infty \frac{d}{dt}\Big[\phi(P_t f)\Big] \, dt \, d\gamma \\
&= -\int_{\mathbb{R}^n} \int_0^\infty \phi'(P_t f) \mathcal{L}(P_t f) \, dt \, d\gamma \\
&= \int_0^\infty \int_{\mathbb{R}^n} \langle \nabla_x \left(\phi' \circ P_f(\,x), \nabla_x (P_t f(x))\right\rangle \, d\gamma(x) \, dt && \text{(Gaussian integration by parts)} \\
&= \int_0^\infty \int_{\mathbb{R}^n} \phi''(P_t f) \left|\nabla(P_t f)\right|^2 \, d\gamma \, dt && \text{(chain rule)} \\
&= \int_0^\infty e^{-2t} \, dt \cdot \int_{\mathbb{R}^n} \phi''(P_t f) \left|P_t\left(\nabla f\right)\right|^2 \, d\gamma && \text{(property of Orstein-Uhlenbeck semigroup)}
\end{aligned}$$

Proof of Theorem 6.47. Taking $\phi(t) = t \log t$ in the computation above, we see that

$$\int_{\mathbb{R}^n} f \log f \, d\gamma - \int_{\mathbb{R}^n} f \, d\gamma \cdot \log \left(\int_{\mathbb{R}^n} f \, \mathrm{d}\gamma \right) = \int_0^\infty e^{-2t} \, \mathrm{d}t \cdot \int_{\mathbb{R}^n} \frac{1}{P_t f} \cdot |P_t \nabla f|^2 \, \mathrm{d}\gamma.$$

Using the explicit representation for P_t and applying Cauchy-Schwarz, we see that for any functions h, g

$$[P_t(hg)]^2 \leq P_t(h^2) \cdot P_t(g^2).$$

Therefore, we see that

$$\begin{aligned}
\int_{\mathbb{R}^n} \frac{1}{P_t f} \cdot |P_t \nabla f|^2 \, d\gamma &= \int_{\mathbb{R}^n} \frac{1}{P_t f} \cdot \sum_{1 \leq i \leq n} (P_t \partial_i f)^2 \, d\gamma \\
&= \int_{\mathbb{R}^n} \frac{1}{P_t f} \cdot \sum_{1 \leq i \leq n} \left[P_t \left(\frac{\partial_i f}{\sqrt{f}} \cdot \sqrt{f} \right) \right]^2 d\gamma \\
&\leq \int_{\mathbb{R}^n} \frac{1}{P_t f} \cdot \sum_{1 \leq i \leq n} P_t \left(\frac{(\partial_i f)^2}{f} \right) \cdot P_t(f) \, d\gamma \\
&= \int_{\mathbb{R}^n} P_t \left(\frac{|\nabla f|^2}{f} \right) d\gamma \\
&= \int_{\mathbb{R}^n} \frac{|\nabla f|^2}{f} \, d\gamma,
\end{aligned}$$

where the last equality follows from the stationarity of γ. □

Remark 6.48. *For a given probability measure μ, we may define the **entropy** of f with respect μ as*

$$\mathrm{Ent}_\mu(f) = \int f \log f \, \mathrm{d}\mu - \left(\int f \, d\mu \right) \cdot \log \left(\int f \, d\mu \right)$$

By applying Jensen's inequality to $t \mapsto t \log t$, we see that $\mathrm{Ent}_\mu(\cdot) \geq 0$.

Remark 6.49. *The quantity $\int |\nabla f|^2 / f \, d\mu$ is commonly referred to as the **Fisher information**.*

Remark 6.50. *If we let $f = g^2$, then*

$$2 \int_{\mathbb{R}^n} g^2 \log g \, d\gamma - \int_{\mathbb{R}^n} g^2 \, d\gamma \cdot \log \left(\int_{\mathbb{R}^n} g^2 d\, \gamma \right) \leq 2 \int_{\mathbb{R}^n} |\nabla g|^2 \, d\gamma$$

Remark 6.51. *See the HW file for a discussion of other related inequalities.*

6.8 Gaussian isoperimetry via the semigroup method

Definition 6.52. Let $A \subseteq \mathbb{R}^n$ be a Borel measurable set. We define the **Minkowski content** of $\partial A \subseteq \mathbb{R}^n$ by

$$|\partial A| = \liminf_{\varepsilon \searrow 0} \frac{|(A + \varepsilon \mathbf{B}_2^n) \setminus A|}{\varepsilon}.$$

The classical isoperimetric inequality states the following:

Theorem 6.53 (Classical Isoperimetric Inequality)**.** *Let $A \subseteq \mathbb{R}^n$ be a Borel measurable set such that $|A| = |\mathbf{B}_2^n|$. Then, $|\partial A| \geq |\partial \mathbf{B}_2^n|$.*

Remark 6.54. *By scaling, we see that if $|A| = |R\mathbf{B}_2^n|$, then*

$$|\partial A| \geq |\partial (R\mathbf{B}_2^n)|.$$

Definition 6.55. For a Borel measurable set $A \subseteq \mathbb{R}^n$, we define its **Gaussian perimeter** by

$$\gamma^+(\partial A) = \liminf_{\varepsilon \searrow 0} \frac{\gamma((A + \varepsilon \mathbf{B}_2^n) \setminus A)}{\varepsilon},$$

where $\gamma(\cdot)$ denotes the standard Gaussian measure on $\mathbb{R}^n$.

Remark 6.56. *If A is a "nice enough" set, then*

$$|\partial A| = \int_{\partial A} \mathbb{1} \, d\mathcal{H}^{n-1},$$

where $\mathcal{H}^{n-1}$ denotes the $(n-1)$-dimensional Hausdorff measure. Similarly,

$$\gamma^+(\partial A) = \int_{\partial A} \frac{1}{(2\pi)^{n/2}} e^{-|x|^2/2} \, d\mathcal{H}^{n-1}(x).$$

Heuristically, we also have

$$\gamma^+(\partial A) \approx \frac{1}{\varepsilon} \mathbb{P}\{Z \notin A : \operatorname{dist}(Z, A) \leq \varepsilon\},$$

where Z is a random variable on $\mathbb{R}^n$ with law γ.

Theorem 6.57 (Gaussian Isoperimetric Inequality, Tsirelson and Sudakov (1974), Borell (1975))**.** *Let $A \subseteq \mathbb{R}^n$ be a Borel measurable set and $a \in (0, 1]$. Then, $\gamma^+(\partial A) \geq \gamma^+(\partial H)$, where $H = \{x \in \mathbb{R}^n : \langle x, \theta \rangle \leq t\}$ for some $\theta \in \mathbb{S}^{n-1}$ $t \in \mathbb{R}^n$ such that $\gamma^+(\partial H) = a$.*

Claim 6.58. *Let $H = \{x \in \mathbb{R}^n : \langle x, \theta \rangle \leq \Phi^{-1}(a)\}$, where $\Phi(s) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{s} e^{-t^2/2} \, dt$ denotes the CDF of the standard Gaussian distribution. Then, $\gamma(H) = a$.*

Proof. Without loss of generality, we may assume that $\theta = e_1$ (otherwise, replace x by Ax, where $A \in \mathbb{R}^{n\times n}$ is such that $\langle Ax, \theta\rangle = \langle x, e_1\rangle$). Using the formula for $\gamma(\cdot)$, we have

$$\begin{aligned}\gamma(H) &= \frac{1}{(2\pi)^{n/2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{0}^{\Phi^{-1}(a)} e^{-\frac{x_1^2+\cdots+x_n^2}{2}}\, dx_1\, dx_2\, \ldots\, dx_n. \\ &= \int_0^{\Phi^{-1}(a)} e^{-x_1^2/2}\, dx_1 \\ &= a.\end{aligned}$$

□

Remark 6.59. *This claim implies that* $\gamma^+(\partial H) = \frac{1}{\sqrt{2\pi}} e^{-\Phi^{-1}(a)^2/2}$. *Therefore, the Gaussian isoperimetric inequality implies that*

$$\gamma^+(\partial A) \geq \frac{1}{\sqrt{2\pi}} e^{-\Phi^{-1}(a)^2/2} = \frac{1}{\sqrt{2\pi}} e^{-\Phi^{-1}(\gamma(A))^2/2}.$$

Definition 6.60. Let $a \in [0,1]$. We define the **Gaussian isoperimetric profile** as

$$I(a) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\Phi^{-1}(a)^2}{2}} = \inf_{\substack{\gamma(A)=a,\\ A\subseteq\mathbb{R}^n}} \gamma^+(\partial A).$$

Remark 6.61. *Here are some basic properties of* $I(\cdot)$*:*

- *The function* $a \mapsto I(a)$ *is convex.*
- *The function* $a \mapsto I(a - 1/2)$ *is even.*
- I *is increasing on* $[0, 1/2]$ *and decreasing on* $[1/2, 1]$.
- $I'' \cdot I = -1$.

6.9 Bobkov's Inequality via the semigroup method

This section is closely related to the Gaussian isoperimetry, specifically the Gaussian Isoperimetric Inequality and the Gaussian isoperimetric profile in the last section. You may review them before proceeding to this section.

Theorem 6.62 (Bobkov's Inequality - "functional version of the Gaussian Isoperimetric Inequality")**.** *If* f *is a measurable function on* $\mathbb{R}^n$*, and for any* x*,* $f(x) \in (0, 1]$*. Suppose* $f \in C^1(\mathbb{R}^n)$ *and the integrals below exist, then*

$$I\left(\int_{\mathbb{R}^n} f\, d\gamma\right) \leq \int_{\mathbb{R}^n} \sqrt{|\nabla f|^2 + I(f)^2}\, d\gamma$$

Note here since $f(x) \in (0, 1]$, so we may apply the $I(\cdot)$ on it.

Claim 6.63. *Bobkov's inequality implies Gaussian isoperimetric inequality for connected sets with smooth boundary*

proof sketch of the claim. For a Borel measurable set $A \in \mathbb{R}^n$, let

$$f = \mathbb{1}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

And we could approximate f with $\tilde{f}$ which is smooth and very close to f.

$$\nabla \mathbb{1}_A(x) = \begin{cases} n_x & \text{if } x \in \partial A \\ 0 & \text{if } x \notin \partial A \end{cases}$$

where n_x is the standard unit normal vector pointing outwards on the boundary

$$\text{RHS: } \int_{\mathbb{R}^n} \sqrt{|\nabla f|^2 + I(f)^2}\, d\gamma \overset{(a)}{=} \int_{\mathbb{R}^n} |\nabla \mathbb{1}_A|\, d\gamma = \int_{\partial A} |n_x|\, d\gamma \overset{(b)}{=} \gamma^+(\partial A)$$

(a) is because $I(f) = \begin{cases} I(1) = 0 \\ I(0) = 0 \end{cases}$ and (b) is because $|n_x| = 1$.

$$\text{LHS: } I\left(\int_{\mathbb{R}^n} f\, d\gamma\right) = I\left(\int_{\mathbb{R}^n} \mathbb{1}_A\, d\gamma\right) = I(\gamma(A))$$

Therefore, we get $I(\gamma(A)) \leq \gamma^+(\partial A)$ □

6.9.1 Semigroup proof of Bobkov's inequality

Let $\{P_t\}_{t \geq 0}$ be the Ornstein-Uhlenbeck semigroup, then the contraction property implies that

$$f \in [0,1] \Rightarrow P_t f \in [0,1].$$

Recall that

$$P_0 f = f,\ P_\infty f = \int f\, d\gamma$$

and with $\mathscr{L}$ being Ornstein-Uhlenbeck operator

$$\frac{d}{dt} P_t f = \mathscr{L}(P_t f) \text{ s.t. } \int f \mathscr{L} g\, d\gamma = -\int \langle \nabla f, \nabla g \rangle\, d\gamma$$

Proof idea: We will show $\frac{d}{dt} \int \sqrt{|\nabla(P_t f)|^2 + I(P_t f)^2}\, d\gamma \leq 0$.

This implies that $\int \sqrt{|\nabla(P_t f)|^2 + I(P_t f)^2}\, d\gamma$ is non-increasing in $t \leq 0$. Thus,

$$\int \sqrt{|\nabla(P_0 f)|^2 + I(P_0 f)^2}\, d\gamma \geq \int \sqrt{|\nabla(P_\infty f)|^2 + I(P_\infty f)^2}\, d\gamma$$

Since $P_\infty f = \int f\, d\gamma$, $\nabla P_\infty f = \nabla(\text{constant}) = 0$. Then,

$$\int \sqrt{|\nabla f|^2 + I(f)^2}\, d\gamma \geq \int \sqrt{I\left(\int f\, d\gamma\right)^2}\, d\gamma = I\left(\int f\, d\gamma\right)$$

which implies Bobkov's inequality.

It remains to prove $\frac{d}{dt}\int \sqrt{|\nabla(P_t f)|^2 + I(P_t f)^2}\, d\gamma \leq 0$. And it is enough to show at $t = 0$.

Setup and Preliminary Identities: The following identities will be used in the proof of Bobkov's inequality:

1. Define $h(x) = I(f(x))^2 + |\nabla f(x)|^2$. Then, chain rule implies that $\nabla h = 2I(f)I'(f)\nabla f + 2\nabla^2 f \cdot \nabla f$.
2. $\nabla(\mathscr{L} f) = \mathscr{L}\nabla f - \nabla f$. Indeed,

$$\nabla(\mathscr{L} f) = \nabla(\Delta f - \langle \nabla f, x\rangle) = \nabla\Delta f - \nabla^2 f x - \nabla f = \mathscr{L}\nabla f - \nabla f.$$

   Write $\mathscr{L}\nabla f = (\mathscr{L}\partial_1 f, \ldots, \mathscr{L}\partial_n f)$.
3. $I \cdot I'' = -1$. To verify this, note that

$$\begin{aligned} &I(a) = \frac{1}{\sqrt{2\pi}} e^{-\Phi^{-1}(a)^2/2};\ I'(a) = \frac{-1}{\sqrt{2\pi}}\Phi^{-1}(a)' \cdot \Phi^{-1}(a) e^{-\Phi^{-1}(a)^2/2} \\ &\left(\Phi^{-1}\right)' = \frac{1}{\Phi'\left(\Phi^{-1}(a)\right)} = \sqrt{2\pi} e^{\Phi^{-1}(a)^2/2} = \frac{1}{I(a)} \\ &\Rightarrow I'(a) = -\Phi^{-1}(a);\ I''(a) = -\frac{1}{I(a)} \end{aligned}$$

4. Item 3 also implies the identity: $(I(a)I'(a))' = (I')^2 - 1$.

Proof of Theorem 6.62. **Start the computation:**

$$\begin{aligned} \frac{d}{dt}\left(I\left(P_t f\right)^2 + |\nabla P_t f|^2\right)\Big|_{t=0} &= 2 \cdot I\left(P_t f\right) \cdot I'\left(P_t f\right) \frac{d}{dt} P_t f\Big|_{t=0} + 2\left\langle \nabla f, \nabla \frac{d}{dt} P_t f\right\rangle\Big|_{t=0} \\ &\overset{(a)}{=} 2I(f) \cdot I'(f)\mathscr{L} f + 2\langle \nabla f, \nabla(\mathscr{L} f)\rangle \\ &\overset{(b)}{=} 2I(f) \cdot I'(f) \cdot \mathscr{L} f + 2\langle \nabla f, \mathscr{L}\nabla f\rangle - 2|\nabla f|^2 \end{aligned}$$

(a) is due to $\frac{d}{dt} P_t f = \mathscr{L}\left(P_t f\right)$ and plug in $t = 0$. (b) is due to 2) in the set up. Therefore,

$$D := \frac{d}{dt}\int \sqrt{I\left(P_t f\right)^2 + |\nabla P_t f|^2}\, d\gamma\Big|_{t=0} \overset{(c)}{=} \int \frac{I(f)I'(f)\mathscr{L} f + \langle \nabla f, \mathscr{L}\nabla f\rangle - |\nabla f|^2}{\sqrt{h}}\, d\gamma$$

(c) is due to $h = I(f)^2 + |\nabla f|^2$ and $(\sqrt{g})' = \frac{g'}{2\sqrt{g}}$. Then, we calculate them in parts.

$$\begin{aligned}\int \frac{I(f)I'(f)\cdot \mathscr{L}f}{\sqrt{h}}\,d\gamma &\overset{(d)}{=} -\int \left\langle \nabla f, \nabla\left(\frac{I(f)I'(f)}{\sqrt{h}}\right)\right\rangle d\gamma \\ &= \int -\frac{|\nabla f|^2}{\sqrt{h}}\cdot\left[I'(f)^2-1\right] + \left\langle \frac{\nabla h}{2}, \frac{I(f)\cdot I'(f)}{(\sqrt{h})^3}\nabla f\right\rangle - \frac{|\nabla f|^2}{\sqrt{h}}\,d\gamma\end{aligned}$$

(d) is by $\int f\mathscr{L}g\,d\gamma = -\int\langle\nabla f,\nabla g\rangle\,d\gamma$. Then, the last term in D can be canceled out, and we are left with

$$D = \int \frac{-I'(f)^2|\nabla f|^2}{\sqrt{h}} + \left\langle \frac{\nabla h}{2}, \frac{I(f)I'(f)}{(\sqrt{h})^3}\nabla f\right\rangle d\gamma + \int \frac{\langle\nabla f, \mathscr{L}\nabla f\rangle}{\sqrt{h}}d\gamma$$

Next, we deal with the second term in D.

$$\begin{aligned}\int \frac{\langle\nabla f,\mathscr{L}\nabla f\rangle}{\sqrt{h}}\,d\gamma &= \sum_{i=1}^n \int \frac{\partial_i f}{\sqrt{h}}\cdot\mathscr{L}\partial_i f\,d\gamma \\ &\overset{(e)}{=} -\sum_{i=1}^n\int\left(\left\langle\nabla\partial_i f, \frac{\nabla\partial_i f}{\sqrt{h}}\right\rangle - \frac{\partial_i f}{(\sqrt{h})^3}\nabla h\right)d\gamma \\ &= \int \frac{-\|\nabla^2 f\|_{HS}^2}{\sqrt{h}} + \frac{\sum\partial_i f\,\langle\nabla\partial_i f,\nabla h\rangle}{2(\sqrt{h})^3}\,d\gamma \\ &= \int \frac{-\|\nabla^2 f\|_{Hs}^2}{\sqrt{h}} + \frac{1}{(\sqrt{h})^3}\left\langle\frac{\nabla h}{2}, \nabla^2 f\nabla f\right\rangle d\gamma\end{aligned}$$

(e) follows from applying the integration by part again. Plug the expression for $\nabla h = 2I(f)I'(f)\nabla f + 2\nabla^2 f\cdot\nabla f$ into the last expression. Simplifying, we get

$$\begin{aligned}D &= \int \frac{1}{(\sqrt{h})^3}(-I'(f)^2I(f)^2|\nabla f|^2 - I(f)^2\left\|\nabla^2 f\right\|_{HS}^2 - I'(f)^2|\nabla f|^4 + I'(f)^2I(f)^2|\nabla f|^2 \\ &\quad + 2I(f)I'(f)\langle\nabla^2 f\nabla f,\nabla f\rangle)\,d\gamma \\ &= -\int\frac{M}{h^{3/2}}\,d\gamma\end{aligned}$$

After some cancellations, we obtain

$$M = I(f)^2\left\|\nabla^2 f\right\|_{HS}^2 + \left[I'(f)\right]^2|\nabla f|^4 - 2I(f)I'(f)\langle\nabla^2 f\nabla f,\nabla f\rangle$$

Then, if $M\geq 0$, then $D = -\int\frac{M}{h^{3/2}}\,d\gamma \leq 0$, which will imply Bobkov's inequality. Therefore, It remains to prove $M\leq 0$. Set

$$A = I(f)\nabla^2 f;\ v = \sqrt{I'(f)}\nabla f$$

where A is a matrix and v is a vector. Then,

$$2\langle Av, v\rangle \leq 2\|A\|_{\text{op}} \cdot |v|^2 \leq \|A\|_{op}^2 + |v|^4 \leq \|A\|_{HS}^2 + |v|^4$$

Therefore,

$$2I(f)I'(f)\left\langle \nabla^2 f \nabla f, \nabla f\right\rangle \leq I(f)^2 \left\|\nabla^2 f\right\|_{HS}^2 + I'(f)^2 |\nabla f|^4 \Rightarrow M \geq 0$$

□

6.9.2 Applications to concentration of measure

Definition 6.64. Let $A \in \mathbb{R}^n$ be a Borel measurable set and $t > 0$. Define

$$A_t := \{x \in \mathbb{R}^n : \text{dist}(x, A) \leq t\} = A + tB_2^n.$$

Theorem 6.65 (Borell's noise stability)**.** *Suppose the Borel measurable set A has a smooth boundary and if $\gamma(A) \geq 1/2$, then $\gamma(A_t) \geqslant 1 - \frac{1}{2}e^{-t^2/2}$. Moreover, if H is a half-space with $\gamma(A) = \gamma(H)$, then $\gamma(A_t) \geq \gamma(H_t) \overset{(a)}{=} \Phi\left(\Phi^{-1}(\gamma(A)) + t\right)$.*

Remark 6.66. *Indeed, $\gamma(A) = \gamma(H)$, $H_t = \{x \in \mathbb{R}^n : \ x \leq \Phi^{-1}(\gamma(A)) + t\}$*

$$\gamma(H_t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\Phi^{-1}(\gamma(A))+t} e^{-s^2/2}\, ds = \Phi(\Phi^{-1}(\gamma(A)) + t)$$

Proof. Idea: Write A_t in layers

$$h(t) = \Phi^{-1}\left(\gamma\left(A_t\right)\right)$$

$$h'(t) = \sqrt{2\pi} e^{\Phi^{-1}(\gamma(A_t))^2/2} \cdot \frac{d}{dt}\gamma\left(A_t\right) = \frac{\gamma^+\left(\partial A_t\right)}{I\left(\gamma\left(A_t\right)\right)} \geqslant 1$$

$$h(t) = h(0) + \int_0^t h'(s)ds \geqslant h(0) + \int_0^t ds = h(0) + t$$

$$\Phi^{-1}\left(\gamma\left(A_t\right)\right) \geqslant \Phi^{-1}(\gamma(A)) + t$$

$$\Rightarrow \gamma\left(A_t\right) \geqslant \Phi\left(\Phi^{-1}(\gamma(A)) + t\right) = \gamma\left(H_t\right)$$

Remaining is to prove $\Phi\left(\Phi^{-1}(\gamma(A)) + t\right) \geqslant 1 - e^{-t^2/2}$. Note

$$\frac{1}{\sqrt{2\pi}} \int_{\Phi^{-1}(\gamma(A))+t}^{\infty} e^{-s^2/2}\, ds \leq \int_t^{\infty} \frac{1}{\sqrt{2\pi}} e^{-s^2/2}\, ds$$

since if $\gamma(A) \geq \frac{1}{2}$, then $\Phi^{-1}(\gamma(A)) \geq 0$. It is enough to show that $\int_t^{\infty} \frac{1}{\sqrt{2\pi}} e^{-s^2/2}\, ds \leq e^{t^2/2}$ and it is left as a homework problem.

□

Remark 6.67. *We saw this theorem for $X \sim N(0, I)$, where $\mathbb{P}(|x| - \mathbb{E}[x]) \geq t) \leq 2e^{-ct^2}$*

Remark 6.68. *This theorem is in fact stronger than the Gaussian isoperimetric inequality.*

Remark 6.69. *This means that if $Z \sim N(0, I)$, then*

$$\mathbb{P}(Z \in A) \geq 1/2 \Rightarrow \mathbb{P}\left(Z \notin A_t\right) \leq 1/2e^{-t^2/2}$$

Corollary 6.70. *Consider a Borel measurable set $A \subseteq \mathbb{S}^{n-1}$ and $A_t = \{x \in S^{n-1} : \operatorname{dist}_{\mathbb{S}^{n-1}}(x, A) \leq t\}$. Suppose $\mathbb{P}(\theta \in A) \geq 0.51$, where $\theta \sim \operatorname{Unif}(\mathbb{S}^{n-1})$. Then, $\mathbb{P}\left(\theta \notin A_t\right) \leq 2e^{-ct^2 n}$, where $c > 0$ is an absolute constant.*

Remark 6.71. *We define the spherical distance between two points $\theta, \eta \in \mathbb{S}^{n-1}$ as*

$$\operatorname{dist}_{\mathbb{S}^{n-1}}(\theta, \eta) := \sphericalangle(\theta, \eta).$$

Note that dist(*between any two points on $\mathbb{S}^{n-1}$ is at most π.*

Proof Sketch of Corollary 6.70. Recall that $X \sim N(0, \mathrm{Id})$, thus by Bernstein's inequality (cite):

$$\mathbb{P}\left(\left|\,|X| - \mathbb{E}\,|X|\,\right| \geq t\right) \leq e^{-ct^2}.$$

Since $\mathbb{E}\,|X|^2 = n$, there exists a constant c such that $\mathbb{E}\,|X| = \sqrt{n-c}$. Thus, we can re-normalize the sphere by considering $\sqrt{n-c} \cdot \mathbb{S}^{n-1}$ instead of $\mathbb{S}^{n-1}$.

So, the statement we want to prove now is equivalent to: let $A \subset \sqrt{n-c} \cdot \mathbb{S}^{n-1}$, $X \sim \operatorname{Unif}\left(\sqrt{n-c} \cdot \mathbb{S}^{n-1}\right)$, $\mathbb{P}(X \in A) \geq 0.51$, and $A_s = \{x \in S^{n-1} : \operatorname{dist}_{\mathbb{S}^{n-1}}(x, A) \leq s\}$, where $s = \sqrt{n-c} \cdot t$. Then, $\mathbb{P}\left(X \notin A_t\right) \leq 2e^{-cs^2}$.

Now consider the set $\tilde{A}$ defined as the cone in $\mathbb{R}^n$ spanned by A, that is

$$\tilde{A} := \left\{x \in \mathbb{R}^n : \frac{x}{|x|}\sqrt{n-c} \in A\right\},$$

and define $\tilde{A}_s$ analogously as $\{x \in \mathbb{R}^n : \frac{x}{|x|}\sqrt{n-c} \in A_s\}$. If $\sigma(\cdot)$ is the uniform measure on $\sqrt{n-c}\mathbb{S}^{n-1}$, then it is clear that $\sigma(A) = \gamma(\tilde{A})$, and $\sigma(A) = \gamma(\tilde{A}_s)$, where we recall that γ is the Gaussian measure on $\mathbb{R}^n$. It is worth noting that $\tilde{A}_s \neq \{x \in \mathbb{R}^n : \operatorname{dist}_{\mathbb{R}^n}(x, \tilde{A}) \leq s\}$.

Now, we will focus on the points of $\tilde{A}$ outside the sphere, that is $\tilde{A} \setminus \{x \in \mathbb{R}^n : |x| \leq \sqrt{n-c}\}$. Then, as $0.51 \leq \sigma(A) = \gamma(\tilde{A})$, this implies that, by Bernstein's inequality, $\gamma(\tilde{A} \setminus \{x \in \mathbb{R}^n : |x| \leq \sqrt{n-c}\}) \geq 0.5$, for a suitable c. Now, we can apply Theorem 6.65 to the set $B = \tilde{A} \setminus \{x \in \mathbb{R}^n : |x| \leq \sqrt{n-c}\}$, thus, $\gamma(B_s) = \gamma(\{x \in \mathbb{R}^n : \operatorname{dist}_{\mathbb{R}^n}(x, B) \leq s\}) \geq 1 - 2e^{-s^2/2}$. Notice that $B_{cs} \subset \tilde{A}_s$ for c small enough, implying that $\sigma(\tilde{A}_s) \geq \gamma(B_{cs}) \geq 1 - 2^{cs^2/2}$. □

Homework: add the remaining details to the proof and remove the $\sigma(A) \geq 0.51$ assumption, particularly, replace it with $\sigma(A) \geq 0.5$.

Definition 6.72. Let μ be a measure on some metric space (X, d). Then we define $\alpha_\mu(t)$ as:

$$\alpha_\mu(t) := \sup_{A \subset X, \text{ measurable}} (1 - \mu(A_t) : \mu(A) \geq 1/2),$$

where, as usual, $A_t = \{x \in X : d(x, A) \leq t\}$.

Given this definition, it is immediate to see, by the previous results, that $\alpha_\gamma(t) = 2e^{-t^2/2}$, and that $\alpha_\gamma(t) \leq 2e^{-ct^2 n/2}$.

Theorem 6.73. *Let (X, d, μ) be a probability metric space, and $f \colon X \to \mathbb{R}$ a random variable which is 1-Lipschitz (i.e. $\forall x, y \in X$ we have $|f(x) - f(y)| \leq d(x, y)$). Then,*

$$\mu(x \in X : |f(x) - \text{med}(f)| > t) \leq 2\alpha_\mu(t).$$

Recall that $\text{med}(f)$ is **the median** of f, that is, a number in $\mathbb{R}$ such that $\mu(f \geq \text{med}(f)) \geq 1/2$ and $\mu(f \leq \text{med}(f)) \geq 1/2$. What this theorem states basically is: "Lipschitz functions are almost constant on metric-measure spaces whose concentration function decays fast".

Remark 6.74. *Vice versa the following is also true:*

$$\alpha_\mu(t) \leq \sup_{\text{Lip}(f)=1} (x \in X : |f(x) - \text{med}(f)| \geq t).$$

Proof of Theorem 6.73. Consider the sets:

$$A = \{x \in X : f(x) \geq med(f)\},$$
$$B = \{x \in X : f(x) \leq med(f)\}.$$

Now, $\forall y \in A_t \; \exists x \in A$ such that $d(x, y) \leq t$, so

$$f(y) = f(y) - f(x) + f(x) \geq -d(x, y) + med(f) \implies$$
$$f(y) - med(f) \geq -t,$$

where $f(y) - f(x) \geq -d(x, y)$ as f is 1-Lipschitz and $f(x) \geq med(f)$ as $x \in A$. Analogously, for B_t we have that if $y \in B_t$ then $f(y) - med(f) \leq t$, so if $y \in A_t \cap B_t$ then

$$|f(y) - med(f)| \leq t,$$

i.e., $\{y \in X : |f(y) - med(f)| \geq t\} \subset A_t^c \cup B_t^c$. However, by the definition of $\alpha_\mu(t)$ we know that $\mu(A_t) \geq 1 - \alpha_\mu(t)$ and $\mu(B_t) \geq 1 - \alpha_\mu(t)$, thus $\mu(A_t^c \cup B_t^c) \leq 1 - 2\alpha_\mu(t)$ implying that

$$\mu(y \in X : |f(y) - med(f)| \geq t) \leq 1 - 2\alpha_\mu(t),$$

and by taking complements we have what we needed. □

Homework: Prove that we can replace the median by the mean when μ is the Gaussian measure or the uniform measure on $\mathbb{S}^{n-1}$:

- $\gamma(x \in \mathbb{R}^n : |f(x) - \mathbb{E}_\gamma(f)| \geq t) \leq 2e^{-ct^2}$,
- $\gamma(x \in \mathbb{R}^n : |f(x) - \mathbb{E}_\sigma(f)| \geq t) \leq 2e^{-ct^2 n}$,

for any 1-Lipschitz function f.

Application 1. The Johnson-Lindenstrauss Lemma. Consider a set of points in $\mathbb{R}^n$, with n large, we can think of these points as data. We would like to reduce the complexity of the data by "reducing the dimension" of the data projecting it to a subspace, but not loosing too much information about the geometry of the point set.

Theorem 6.75 (Johnson-Lindenstrauss Lemma)**.** *Let $\mathcal{X}$ be a set of N points in $\mathbb{R}^n$. Take $\epsilon > 0$, and suppose m is an integer such that $m \geq \frac{c}{\epsilon^2} \log N$. Now, let E be a random subspace of $\mathbb{R}^n$ of dimension m sampled uniformly from $G_{n,m}$, the collection of all m-dimensional subspaces of $\mathbb{R}^n$. Let also P_E be the orthogonal projection on E, and $Q = \sqrt{\frac{n}{m}} P_E$ the normalized projection. Then, with probability at least $1 - 2e^{-c\epsilon^2 m}$, we have that for all $x, y \in \mathcal{X}$.*

$$(1 - \epsilon)\,|x - y| \leq |Qx - Qy| \leq (1 + \epsilon)\,|x - y|\,. \tag{76}$$

Rephrased in a general way the theorem states that "with high probability the normalized random projection is almost an isometry".

Corollary 6.76. *There exists a subspace of dimension m whose normalized orthogonal projection satisfies (76).*

Remark 6.77. *We call the collection of all m-dimensional subspaces of $\mathbb{R}^n$ as $G_{m,n}$. The random subspace is selected uniformly in the previous theorem.*

Lemma 6.78. *Let P be a projection of $\mathbb{R}^n$ onto a uniform random subsapce of dimension m. Fix a point $z \in \mathbb{R}^n$ and $\epsilon > 0$. Then*

1. $\mathbb{E}\,|Pz|^2 = \frac{m}{n}|z|^2$,
2. *with probability $\geq 1 - 2e^{-c\epsilon^2 n}$, we have that*

$$(1 - \epsilon)\sqrt{\frac{m}{n}} \leq |Pz| \leq (1 + \epsilon)\sqrt{\frac{m}{n}}.$$

Proof of the Lemma. Without loss of generality, let $|z| = 1$. Equivalently, instead of randomly selecting the subspace we can fix it as $F = \operatorname{span}(e_1, \ldots, e_m)$, and consider $Z \sim \operatorname{Unif}(\mathbb{S}^{n-1})$. Given this, we have

$$\mathbb{E}[|Pz|^2] = \mathbb{E}[\operatorname{proj}(Z, F)^2] = \mathbb{E}\left[\sum_{i=1}^{m} Z_i^2\right] = m\mathbb{E}[Z_1^2], \tag{77}$$

where we used the projection invariance and the fact that since Z is uniform on $\mathbb{S}^{n-1}$ then all of its coordinates have the same distribution, thus $\mathbb{E}[Z_i^2] = \mathbb{E}[Z_1^2]$ for all i. On the other hand,

$$1 = \mathbb{E}[|Z|^2] = \sum_{i=1}^{n} \mathbb{E}[Z_i^2] = n\mathbb{E}[Z_1^2],$$

implying that $\mathbb{E}[Z_1^2] = 1/n$. Substituting this in 77 we get that $\mathbb{E}[|Pz|^2] = m/n$ as wanted.

Now for part two, we will use Theorem 6.73 on $\mathbb{S}^{n-1}$ and with the uniform measure, that is $Z \sim \text{Unif}(\mathbb{S}^{n-1})$. Now note that the equation in part two is equivalent to

$$1 - \epsilon \leq \sqrt{\frac{m}{n}}\,|\text{Proj}_F(Z)| \leq 1 + \epsilon.$$

But this follows immediately by applying Theorem 77 to the function $\sqrt{\frac{m}{n}}|\text{proj}_F(Z)|$ which is clearly $\sqrt{\frac{m}{n}}$-Lipschitz if $\dim(F) = m$. That is

$$\mathbb{P}\left(\left|\sqrt{\frac{m}{n}}\,|\text{Proj}_F(Z)| - \sqrt{\frac{m}{n}}\right| \geq t\right) \leq e^{-cnt^2},$$

which implies what we want by considering $t = \epsilon\sqrt{\frac{n}{m}}$. □

Proof of the Johnson-Linderstrauss Lemma. Again let $\mathcal{X}$ be a set of N points, and consider $z \in \mathcal{X} - \mathcal{X} := \{z = x - y : x, y \in \mathcal{X}\}$, the "difference set". From the previous lemma, we have that for each z:

$$\mathbb{P}(Qz \in [(1-\epsilon)|z|, (1+\epsilon)|z|]) \geq 1 - 2e^{c\epsilon^2 m}.$$

Let E_z be the event $\{Qz \in [(1-\epsilon)|z|, (1+\epsilon)|z|]\}$, then

$$\mathbb{P}\left(\bigcap_{z\in\mathcal{X}-\mathcal{X}} E_z\right) \geq 1 - \#(\mathcal{X}-\mathcal{X}) \cdot 2e^{-c\epsilon^2 m} = 1 - N^2 \cdot 2e^{-c\epsilon^2 m} \geq 1 - 2e^{-c'm^2\epsilon},$$

where we used $m \geq e/\epsilon^2 \log N$. □

Application 2. Milman-Dvoretzky Theorem (1971).

Theorem 6.79 (Schetman's version of Milman-Dvoretzky Theorem)**.** *Let $\|\cdot\|$ be a norm in $\mathbb{R}^n$, and consider $\epsilon \in (0,1)$. Then $\exists k \geq c\epsilon^2 \log n$ such that with probability $\geq 1 - e^{c_1\epsilon^2 k}$ the random k-dimensional subspace F satisfies: $\exists \mu > 0$ and GL_n operator $T : \mathbb{R}^n \to \mathbb{R}^n$ such that $\forall x \in F$*

$$(1-\epsilon)\mu|x| \leq \|Tx\| \leq (1+\epsilon)\mu|x|.$$

Rephrasing the last theorem is "any norm is almost euclidean on a random subspace".

The proof of the Milman-Dvoretzky Theorem is based on the concentration of measure for Lipschitz functions and net argument. We state and prove intermediate lemmas and propositions before proving the main theorem. Throughout the following discussion we denote $b > 0$ as smallest positive number such that $\|x\| \leq b|x|$ for all $x \in \mathbb{R}^n$. Let $M := \int_{\mathbb{S}^{n-1}} \|x\| \, d\sigma(x)$ where σ is the uniform measure on $\mathbb{S}^{n-1}$. Let $O(n)$ be the n-dimensional orthogonal group and $\mathcal{V}_n$ be the uniform measure on $O(n)$. A useful fact is that the probability of hitting the area A on the sphere is equal to the probability that any point on the sphere is randomly rotated and lands into the area A. That is, for any fixed $x_0 \in \mathbb{S}^{n-1}$ and $A \subset \mathbb{S}^{n-1}$, we have $\sigma(A) = \mathcal{V}_n(\{u \in O(n) : ux_0 \in A\})$.

To begin with, we are going to prove the following lemma.

Lemma 6.80. *Fix $m \in \mathbb{N}$ such that $m \leq \frac{1}{4}e^{c_1\epsilon^2 n/2}$. Let $\{y_1, ..., y_m\}$ be a collection of m points in $\mathbb{S}^{n-1}$. Then there exists $B \subset O(n)$ with $\mathcal{V}_n(B) \geq 1 - \exp\{-c_1\epsilon^2 n/2\}$, such that for any $u \in B$ and any $i \in \{1, ..., m\}$, we have*

$$M - b\epsilon \leq \|uy_i\| \leq M + b\epsilon. \tag{78}$$

Proof of Lemma 6.80. Let $A = \{x \in \mathbb{S}^{n-1} : M - b\epsilon \leq \|x\| \leq M + b\epsilon\}$ and let $f(x) = \|x\|$. Note that $f(x)$ is b-Lipschitz and $\mathbb{E}\, f(X) = M$. That means we can apply concentration of Lipschitz functions, Theorem 6.73, and get

$$\sigma(A) \geq 1 - 4\exp\{-c_1\epsilon^2 n\}. \tag{79}$$

Now consider $B_i = \{u \in O(n) : M - b\epsilon \leq \|uy_i\| \leq M + b\epsilon\}$ and define our B to be $B = \cap_{i=1}^m B_i$. Recall that we have $\mathcal{V}_n(B_i) = \sigma(A)$ from previous argument. By union bound we conclude

$$\mathcal{V}_n(B) = 1 - \mathcal{V}_n(B^c) \geq 1 - \sum_{i=1}^m \mathcal{V}_n(B_i^c) = 1 - 4m\exp\{-c_1\epsilon^2 n\} \geq 1 - \exp\{-c_1\epsilon^2 n/2\}, \tag{80}$$

where in the last step we recall $m \leq 1/4\exp\{c_1 c^2 n/2\}$. □

The lemma above basically states the concentration of the norm of randomly rotated vectors on a sphere. It can be used to obtain similar results for vectors in subspaces, shown in the following Propositon.

Proposition 6.81. *Fix $\delta, \epsilon \in (0,1), k \in \mathbb{N}$ and suppose $(1 + \frac{2}{\delta})^k \leq \frac{1}{4}\exp\{c_1\epsilon^2 n/2\}$. Then there exists a set of k-dimensional subspaces, call it Γ, with $\mathcal{V}_{n,k}(\Gamma) \geq 1 - \exp\{-c_1\epsilon^2 n/2\}$[1], such that for all subspace $F \in \Gamma$, there exists a δ-Net $\mathcal{N}_F$ on $\mathbb{S}^{n-1} \cap F$, so that*

$$\forall x \in \mathcal{N}_F, \quad M - b\epsilon \leq \|x\| \leq M + b\epsilon. \tag{81}$$

[1] $\mathcal{V}_{n,k}$ is the uniform measure on all k-dimensional linear subspaces. In other words, here it is equivalent to say "we take a random subspace F, with probability at least $1 - \exp\{-c_1\epsilon^2 n/2\}$, ..."

Proof of Proposition 6.81. Fixed a k-dimensional subspace F_0. Consider the δ-Net $\{y_1, ..., y_m\}$ on $\mathbb{S}^{n-1} \cap F_0$. Since $\mathbb{S}^{n-1} \cap F_0$ has dimension $k-1$, we can take $m = (1+\frac{2}{\delta})^k$ by net argument. Now applying Lemma 6.80 to $\{y_1, ..., y_m\}$, we can find $B \subset O(n)$ such that

$$\forall u \in B, \forall i \in \{1, ..., m\}, \quad M - b\epsilon \leq \|uy_i\| \leq M + b\epsilon. \tag{82}$$

For a rotation $u \in B$, define $x_i = uy_i$ and $F_u = uF_0$. Note that $\{x_i\}$ forms a δ-Net for $\mathbb{S}^{n-1} \cap F_u$, since orthogonal transformation preserves Euclidean norm.

Now construct Γ by $\Gamma = \{F_u : u \in B\}$. We have $\mathcal{V}_{n,k}(\Gamma) = \mathcal{V}_n(B) \geq 1 - \exp\{-c_1\epsilon^2 n/2\}$. For every $F_u \in F$, the δ-Net is exactly $\{x_1, ..., x_m\} = \{uy_1, ..., uy_m\}$ and it satisfies

$$M - b\epsilon \leq \|x_i\| = \|uy_i\| \leq M + b\epsilon \tag{83}$$

by construction. □

The Proposition above is close to what we want for concentration of norm, except that it only works for δ-net. Now let us extend it to any points in the subspace.

Proposition 6.82. *Fix $\delta, \epsilon \in (0,1), k \in \mathbb{N}$ and suppose $(1+\frac{2}{\delta})^k \leq \frac{1}{4}\exp\{c_1\epsilon^2 n/2\}$. If we take a random k-dimensional subspace F, then with probability at least $1 - \exp\{-c_1\epsilon^2 n/2\}$, for all $y \in F \cap \mathbb{S}^{n-1}$,*

$$\frac{1-2\delta}{1-\delta}M - \frac{b\epsilon}{1-\delta} \leq \|y\| \leq \frac{M + b\epsilon}{1-\delta}. \tag{84}$$

Proof of Proposition 6.82. Let Γ be the set of k-dimensional subspace in Proposition 6.81. By definition of δ-net, For all $F \in \Gamma$, for all $y \in F \cap \mathbb{S}^{n-1}$, there exists $x_0 \in \mathcal{N}_F$ so that $|y - x_0| := \delta_1 \leq \delta$. Meanwhile, from Proposition 6.81 we know that

$$M - b\epsilon \leq \|x\| \leq M + b\epsilon. \tag{85}$$

Now note that $\frac{y-x_0}{\delta_1} \in \mathbb{S}^{n-1} \cap F$ because F is closed to linear combination and $|\frac{y-x_0}{\delta_1}| = 1$. So again we can find $x_1 \in \mathcal{N}_F$ so that

$$\left|\frac{y - x_0}{\delta_1} - x_1\right| := \delta_2 \leq \delta, \tag{86}$$

or equivalently $|y - x_0 - \delta_1 x_1| = \delta_1\delta_2 \leq \delta^2$. We can keep following this procedure and get a sequence $x_0, x_1, ..., x_N$ so that

$$\left|y - \sum_{i=0}^{N}\prod_{j=0}^{i}\delta_j x_i\right| = \prod_{i=0}^{N}\delta_i \leq \delta^{N+1}. \tag{87}$$

As $\delta \in (0,1)$, this implies

$$y = \sum_{i=0}^{\infty}\prod_{j=0}^{i}\delta_j x_i. \tag{88}$$

Therefore, using triangle inequality and the fact that $\delta_j \leq \delta$, we obtain

$$\begin{aligned} \|y\| &\leq \sum_{i=0}^{\infty} \delta^i \|x_i\| \\ &\leq \sum_{i=0}^{\infty} \delta^i (M + b\epsilon) = \frac{M + b\epsilon}{1 - \delta}, \end{aligned} \tag{89}$$

which finishes the upper bound. The proof of lower bound can be done in a similar manner (HW). □

What remains to prove Milman-Dvoretzky Theorem is to show there exists an invertible linear operator T such that if we replace $\|x\|$ by $\|Tx\|$ we can get

$$\frac{M}{b} \geq c \frac{\log n}{\sqrt{n}}. \tag{90}$$

Then the bound depends only on parameter δ. We can compute and plug the optimal δ to obtain the desired bound (HW).

Remark 6.83. *Equation (90) is sharp. For example $\|\cdot\| = \|\cdot\|_\infty$, in such case*

- $M = \int_{\mathbb{S}^{n-1}} \max_{i=1,\dots,n} |x_i| d\sigma(x) = \frac{c \log n}{\sqrt{n}}$ *(HW).*
- $b = 1$ *as* $\|x\|_\infty \leq |x|$.

Remark 6.84. *Milman-Dvoretzky Theorem (Theorem 6.79) is sharp for $\|\cdot\|_\infty$.*

7 The Mass Transport method

Imagine that there are several bakeries producing croissants and shops waiting for the products. The productivity of bakeries and the demand of shops are different. What is the best way to deliver croissants, in terms of delivery distance? This can be formalized as the Monge-Kantorovich problem, or transportation problem.

7.1 Basic definitions and set up

Definition 7.1 (Transport map)**.** let μ, ν be Borel measures on $\mathbb{R}^n$. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a function defined almost everywhere on the support of μ, i.e., a Borel measurable map. We say that T transports μ into ν, if for any subset $A \subset \mathbb{R}^n$,

$$\nu(A) = \mu(T^{-1}A), \tag{91}$$

where $T^{-1}A$ is the inverse image of T by A.

Remark 7.2 (HW)**.** *The definition above of transport map is equivalent to: for any function $\phi \in L'(\nu)$ (integrable function using measure ν),*

$$\int \phi d\nu = \int \phi \circ T d\mu. \tag{92}$$

Example 7.3 (Discrete measure)**.** *let $\mu = \sum_{i=1}^m \lambda_i \delta_{x_i}$ be a discrete measure supported on $\{x_1, ..., x_m\}$. Then for any function $T : \mathbb{R}^n \to \mathbb{R}^n$, let $\nu = \sum_{i=1}^m \lambda_i \delta_{T(x_i)}$, T transports μ into ν.*

Example 7.4 (Gaussian measure)**.** *Let μ be Gaussian measure on $\mathbb{R}$ and ν be uniform measure on $[0, 1]$. Let $T(x) = \Phi(x)$ be the cdf of Gaussian distribution. Then T transport μ into ν. Indeed, for any $0 \le a \le b \le 1$, we have*

$$\begin{aligned} \nu([a, b]) &= b - a \\ \mu(T^{-1}([a, b])) &= \Phi(T^{-1}(b)) - \Phi(T^{-1}(a)) = b - a. \end{aligned} \tag{93}$$

7.2 Useful Properties of Lipschitz Transport Maps

Moving forward, we can extend the above example (Example 7.4) to $\mathbb{R}^n$.

Example 7.5 (Gaussian measure on $\mathbb{R}^n$)**.** *Let γ be a Gaussian measure in $\mathbb{R}^n$. Consider the transport map $T : \mathbb{R}^n \to \mathbb{R}^n$ given by*

$$T(x) = \begin{bmatrix} \Phi(x_1) \\ \vdots \\ \Phi(x_n) \end{bmatrix}, \tag{94}$$

where Φ is the cdf of a Gaussian distribution. We have $T_\star \gamma = Unif([0, 1]^n)$, and T is $\frac{1}{\sqrt{2\pi}}$-Lipschitz continuous. Recall that a map $T : \mathbb{R}^n \to \mathbb{R}^n$ is L-Lipschitz continuous if there exists a real constant $L \ge 0$ such that

$$|T(x) - T(y)| \le L|x - y|, \qquad \forall x, y \in \mathbb{R}^n.$$

Since $\Phi(s) = \int_{-\infty}^s \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$, we know

$$\Phi'(s) = \left(\int_{-\infty}^s \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \right)' = \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} \le \frac{1}{\sqrt{2\pi}},$$

which implies T is $\frac{1}{\sqrt{2\pi}}$-Lipschitz continuous.

The following proposition demonstrates that a L-Lipschitz continuous transport map allows us to extend the properties of one measure to another.

Proposition 7.6. *Suppose T is L-Lipschitz continuous for any $L \ge 0$ and $T_\star \mu = \nu$ (i.e., T transports μ into ν) on $\mathbb{R}^n$. Then the following holds:*

1. $\alpha_\mu(t) \geq \alpha_\nu(tL)$ *; or equivalently,* $1-\alpha_\mu(t) \leq 1-\alpha_\nu(tL)$ *(recall from Definition 6.72)*

2. *Poincaré Inequality propagates:* $C_{poin}(\mu) \leq \frac{1}{L^2} C_{poin}(\nu)$*, where* $C_{poin}(\cdot)$ *denotes the Poincaré constant of a measure. In other words, if we know* $\forall f$ *vice*

$$\int f^2 d\mu - \left(\int f d\mu\right)^2 \leq C_{poin}(\mu) \int |\nabla f|^2 d\mu,$$

then $\forall g$ *vice*

$$\int g^2 d\nu - \left(\int g d\nu\right)^2 \leq L^2 C_{poin}(\mu) \int |\nabla g|^2 d\nu.$$

3. *Log-Sobolev Inequality propagates: (Recall the definition of entropy of a function given by Remark 6.48) If* $\forall f$ *reasonable with*

$$\mathrm{Ent}_\mu(f^2) \leq \frac{C}{2} \int |\nabla f|^2 d\mu,$$

then $\forall g$ *reasonable with*

$$\mathrm{Ent}_\nu(g^2) \leq \frac{L^2 C}{2} \int |\nabla g|^2 d\nu.$$

4. *Isoperimetry propagates: Consider* I_μ *the isoperimetric profile of* μ*, that is,* $\forall a \in [0,1]$*,*

$$I_\mu(a) = \inf_{\substack{A \subset \mathbb{R}^n \text{ Borel measurable,} \\ \mu(A)=a}} \mu^+(\partial A). \tag{95}$$

Here recall $\mu^+(\partial A) = \liminf_{t \to 0} \frac{\mu(A_t \setminus A)}{t}$ *where* $A_t = \{x \in \mathbb{R}^n : \mathrm{dist}(x,a) \leq t\}$*. We have*

$$I_\nu \geq \frac{1}{L} I_\mu \quad \text{on } [0,1]. \tag{96}$$

Proof Sketch. We first make an important observation that if the map T is L-Lipschitz continuous for some $L \geq 0$, then

$$T(A_t) \subset (TA)_{tL}, \qquad \forall t > 0. \tag{97}$$

This is because

$$\begin{aligned} T(A_t) &= \{Tx : \mathrm{dist}(x, A) \leq t\} \\ &= \{y : \mathrm{dist}(T^{-1}y, A) \leq t\} \subset \{y : \mathrm{dist}(y, TA) \leq Lt\} = (TA)_{tL}. \end{aligned}$$

Now using (97), we can prove the statements.

1. We know
$$1-\alpha_\mu(t)=\inf_{A:\mu(A)\geq 1/2}\mu(A_t)\leq \inf_{A:\mu(A)\geq 1/2}\nu((TA)_{tL}).$$
We let $B=TA$. It follows that
$$B=\inf_{B:\nu(B)\geq 1/2}\nu(B_{tL})=1-\alpha_\nu(tL).$$

2. Since T is a mass transport map, we have
$$\int f^2 d\mu=\int (f\circ T)^2 d\nu$$
and
$$\int f d\mu=\int (f\circ T) d\nu.$$
Consider $g=f\circ T$, we have $\mathrm{Var}_\mu(f)=\mathrm{Var}_\nu(g)$. By the change of variable $y=Tx$,
$$\int |\nabla g|^2 d\nu=\int |\nabla (f\circ T)|^2 d\nu\geq L^2\int |\nabla f|^2 d\mu. \tag{98}$$

3. The proof follows from (98) and the fact that $\mathrm{Ent}_\mu(f^2)=\mathrm{Ent}_\nu((f\circ T)^2)$.

4. The proof is left as HW.

□

The following corollary demonstrates the good properties of uniform measures on the cube.

Corollary 7.7. *Let $d\nu=\mathbb{1}_{[0,1]^n}(x)dx$. Then we have*

1. $\alpha_\nu(t)\geq 1-Ce^{-\frac{t^2}{2}},\ \forall t>0$.
2. *For any f in $L^2([0,1]^n)\cap C^2([0,1]^n)\cap\{\nabla f\in L^2([0,1]^n)\}$, we have*
$$\int_{[0,1]^n} f^2 dx-\left(\int_{[0,1]^n} f dx\right)^2\leq \frac{1}{2\pi}\int_{[0,1]^n}|\nabla f|^2 dx. \tag{99}$$
Note that this inequality is sharp for univariate functions like $f(x)=\sin(x_1)$ (with correct scaling).
3. *Log-Sobolev Inequality propagates*
4. *For any $A\subset\mathbb{R}^n$ with $|A\cap[0,1]^n|=a\in[0,1]$, we have*
$$|\partial A\cap[0,1]^n|_{n-1}\geq\sqrt{2\pi}I_\gamma(a)=e^{-\frac{\Phi^{-1}(a^2)}{2}}, \tag{100}$$
where $I_\gamma(a)$ is the Gaussian isoperimetric profile of $a\in[0,1]$.

Remark 7.8. *Statement (4) in the above corollary is sharp for A with $|A \cap [0,1]^n| = \frac{1}{2}$.*

Next, we present a special example that was previously studied by Vaaler, Hadwiger, and others.

Corollary 7.9 (Vaaler's theorem on the central sections of the cube)**.** *For any $\theta \in \mathbb{S}^{n-1}$, we have $|[0,1]^n \cap \theta^\perp| \geq 1$.*

At this point, we can conclude that it is nice to have a Lipschitz map transporting nice measures into measures. Now the question is, are there other measures besides the uniform measure of the cube where we can construct a Lipschitz map transporting between measures?

7.3 Optimal Transport Map with respect to quadratic cost

Definition 7.10 (Optimal Transport)**.** A map T_0 transporting μ into ν is said to be **optimal w.r.t. quadratic cost** if for all T with $T_\star \mu = \nu$,

$$\int_{\mathbb{R}^n} |T_0 x - x|^2 d\mu \leq \int_{\mathbb{R}^n} |Tx - x|^2 d\mu. \tag{101}$$

In other words, on average, T_0 minimizes the distance each point needs to travel.

Now that we have defined the optimal transport map (w.r.t. quadratic cost), the question is, how can we check if a map attains optimality? Also, how can we explicitly construct an optimal transport map between measures? To answer these questions, we first state an important theorem in optimal transport.

Theorem 7.11 (Brenier's theorem)**.** *Let μ, ν be absolutely continuous Borel finite measures. Suppose $T_\star \mu = \nu$. The map T is optimal w.r.t. quadratic cost if and only if there exists a convex function $F : \mathbb{R}^n \to \bar{\mathbb{R}}$ (finite a.e. on supp(μ)) such that*

$$T = \nabla F. \tag{102}$$

In this situation, such a map T always exists. This map is known as the "Brenier map".

In the following, we give some examples of the construction of optimal transport maps.

Example 7.12. *Let γ be a Gaussian measure. Suppose we have*

$$T_\star \gamma = \frac{1}{2}\delta_{e_1} + \frac{1}{2}\delta_{-e_1},$$

where $e_1 \in \mathbb{R}$, δ_x denotes the Dirac delta function at $x \in \mathbb{R}$. The optimal transport map w.r.t. quadratic cost, in this case, is given by

$$T_\star x = \operatorname{sign}(x_1) e_1. \tag{103}$$

We can see that $T_\star x = \nabla F(x)$ where $F(x) = |x_1|$ and $\nabla F(x) = \operatorname{sign}(x_1) e_1$.

Example 7.13. *Let us consider a more general example. Consider a map that transports μ to finitely many points, that is,*

$$T_\star \mu = \sum_{i=1}^{N} \lambda_i v_i, \qquad \text{for } v_i \in \mathbb{S}^{n-1}. \tag{104}$$

The optimal map is given by $Tx = \nabla F(x)$ where $F(x) = \max_{j\in\{1,\cdots,N\}}(y_j + \lambda_j \langle x, v_j\rangle)$.

Remark 7.14 (Caferelli's theorem)**.** *Let γ be a Gaussian measure. If $d\mu = e^{-V(x)}dx$ on $\mathbb{R}^n$ where $V \in C^2(\mathbb{R}^n)$ and $\nabla^2 V \geq kI_d$ for some constant $k \geq 0$ (here I_d denotes the identity matrix). Let $T_\star \gamma = \mu$ be the Brenier map. Then T is $\frac{1}{k}$-Lipschitz. Such measures are called* **strictly log-concave** *and have many nice properties (allowing us to transport Gaussian measures to log-concave measures). We will prove Caferelli's theorem later.*

Before we prove Brenier's theorem, let us first sketch the proof for the discrete version of Brenier.

Theorem 7.15 (Discrete version of Brenier's theorem)**.** *Let $x_1, \ldots, x_N, y_1, \ldots, y_N \in \mathbb{R}^n$.*

1. *Supposed there exists convex function $F : \mathbb{R}^n \to \mathbb{R}$ such that*

$$\nabla F(x_i) = y_i, \qquad \forall i = \{1, \ldots, N\}. \tag{105}$$

Then for any permutation σ on the set $\{1, \ldots, N\}$, we have

$$\sum_{i=1}^{N} |x_i - y_{\sigma(i)}|^2 \geq \sum_{i=1}^{N} |x_i - y_i|^2. \tag{106}$$

2. *Vice versa, if*

$$\sum_{i=1}^{N} |x_i - y_i|^2 = \min_{\sigma\in S_N} \sum_{i=1}^{N} |x_i - y_{\sigma(i)}|^2, \tag{107}$$

then there exists a convex function $F : \mathbb{R}^n \to \bar{\mathbb{R}}$ such that

$$y_i \in \partial F(x_i), \qquad \forall i = \{1, \ldots, N\}. \tag{108}$$

Here, we would like to introduce the concept of sub-gradient of a convex function:

Definition 7.16 (Sub-gradient of a convex function f)**.** The sub-gradient of a convex function f at x is given by

$$\partial f(x) = \{y \in \mathbb{R}^n : f(z) \geq f(x) + \langle y, z - x\rangle, \quad \forall z \in \mathbb{R}^n\}. \tag{109}$$

Note that if $f \in C^1(\mathbb{R}^n)$, then $\partial f(x) = \{\nabla f(x)\}$. Let us consider a quick example.

Example 7.17 (Sub-gradient of the absolute value function)**.** *Consider the function* $f : \mathbb{R} \to \mathbb{R}$ *given by* $f(t) = |t|$*. The sub-gradient of* f *at* 0 *is* $\partial f(0) = [-1, 1]$*. The sub-gradient of* f *on points not equal to* 0 *is*

$$\partial f(x) = \begin{cases} -1, & \forall\, x < 0, \\ 1, & \forall\, x > 0. \end{cases}$$

Now that we have introduced the concept of sub-gradient, we present the proof of Theorem 7.15.

Proof of Discrete version of Brenier's theorem (Theorem 7.15). Let us first prove statement (1) in the theorem.

Proof of Statement (1)

Given that F is convex, we know

$$F(y) - F(x) \geq \langle \nabla F(x), y - x \rangle, \qquad \forall x, y \in \mathbb{R}^n. \tag{110}$$

Also, since $\nabla F(x_i) = y_i$ for $i = \{1, \ldots, N\}$, it follows from (110) and the commutativity of addition that

$$\sum_{i=1}^{N} \langle \nabla F(x_i), x_{\sigma(i)} - x_i \rangle \leq \sum_{i=1}^{N} F(x_{\sigma(i)}) - F(x_i) = 0$$

This implies that

$$\sum_{i=1}^{N} \langle \nabla F(x_i), x_{\sigma(i)} \rangle \leq \sum_{i=1}^{N} \langle \nabla F(x_i), x_i \rangle, \tag{111}$$

or equivalently, $\sum_{i=1}^{N} \langle \nabla y_i, x_{\sigma(i)} \rangle \leq \sum_{i=1}^{N} \langle y_i, x_i \rangle$. This further implies that

$$\sum_{i=1}^{N} |y_i - x_{\sigma(i)}|^2 \geq \sum_{i=1}^{N} |y_i - x_i|^2 \tag{112}$$

because $\sum_{i=1}^{N} (x_{\sigma(i)})^2 = \sum_{i=1}^{N} x_i^2$. The proof of statement (1) is complete.

□

Definition 7.18. (Coupling of Measures - Generalization of the Mass Transport) Let μ and ν be probability measures on $\mathbb{R}^n$ Consider a measure γ on $\mathbb{R}^n \times \mathbb{R}^n$ to be a coupling (Borell, and bounded) of μ and ν. If $\forall \varphi : \mathbb{R}^n \to \mathbb{R}$. Then -

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} \varphi(x) d\gamma(x, y) = \int_{\mathbb{R}^n} \varphi(x) d\mu(x) \tag{113}$$

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} \varphi(y) d\gamma(x, y) = \int_{\mathbb{R}^n} \varphi(y) d\nu(y) \tag{114}$$

where $x, y \in \mathbb{R}^n$

Example 7.19. *Let* μ *and* ν *be probability measures on* $\mathbb{R}^n$*. Consider a transport map* $T : \mathbb{R}^n \to \mathbb{R}^n$*, which transports* μ *into* ν *(i.e.* $T_* \mu = \nu$*) on* $\mathbb{R}^n$*. Associate a coupling of* μ *and* ν *as* γ_T *on* $\mathbb{R}^n \times \mathbb{R}^n$ *with* T*. The measure* γ *is supported on* (x, Tx)

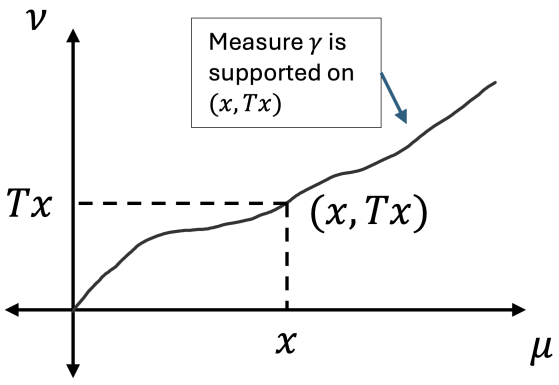


Figure 3: Measure

7.4 Monge Problem

Definition 7.20. (Monge Problem) Let a cost function $C : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty)$ be lower semicontinuous. Then -

$$\inf_{T_*\mu=\nu} \int_{\mathbb{R}^n} C(x, Tx) d\mu(x) = C_\mu(\mu, \nu) \tag{115}$$

If we use a quadratic cost, i.e.

$$C(x, y) = |x - y|^2 \tag{116}$$

then if the infimum in Eq. 115 is attained for some T, then T is optimal w.r.t quadratic cost

7.5 Kantorovich Problem

Let μ and ν be probability measures on $\mathbb{R}^n$ Consider the equation -

$$C_k(\mu, \nu) = \inf_{\gamma} \int_{\mathbb{R}^n \times \mathbb{R}^n} C(x, y) d\gamma(x, y) \tag{117}$$

where γ is a coupling of μ and ν, defined on $\mathbb{R}^n \times \mathbb{R}^n$. As we had explained, $C_k(\mu, \nu) \leq C_\mu(\mu, \nu)$, however in most cases - $C_k(\mu, \nu) = C_\mu(\mu, \nu)$, and in particular for a quadratic cost function (Eq. 116), and absolutely continuous measures.

Theorem 7.21. *(Without proof) Let a cost function $C : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty)$ be lower semicontinuous. This implies that $\exists$ a coupling γ^* between μ and ν (Can be any probability measures), which solves the Kantorovich problem.*

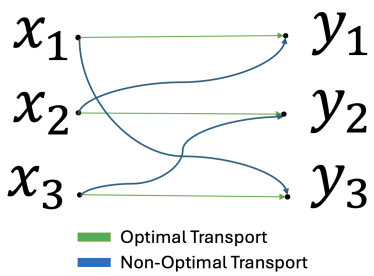


Figure 4: Cyclic Monotonicity Explanation

7.6 Cyclic Monotonicity

Definition 7.22. A set $\Lambda \subset \mathbb{R}^n \times \mathbb{R}^n$ is C-cyclically monotone if $\forall$ finite sequences $(x_i, y_i)_{i=1}^N$, s.t. $x_{N+1} = x_1$, then -

$$\sum_{i-1}^{N} C(x_i, y_i) \leq \sum_{i=1}^{N} C(x_{i+1}, y_i) \tag{118}$$

Theorem 7.23. *If γ^* is an optimal coupling of μ and ν w.r.t. continuous cost $C(x, y)$, then supp(γ^*) is* ***cyclically monotone****.*

Recall that supp(γ^*) is the interior of the set of points where $\gamma^* > 0$

Proof. (Homework!) □

Remark 7.24. *Consider a quadratic cost as shown in Eq. 116, and $x_{N+1} = x_1$. Then it follows $\forall (x_i, y_i)_{i=1}^N \subset A$, where* ***A is cyclically monotone*** *if and only if -*

$$\sum_{i=1}^{N} |x_i - y_i|^2 \leq \sum_{i=1}^{n} |x_{i+1} - y_i|^2 \Leftrightarrow \sum_{i=1}^{N} \langle x_i, y_i \rangle \geq \sum_{i=1}^{n} \langle x_{i+1}, y_i \rangle \Leftrightarrow \sum_{\mathbf{i=1}}^{\mathbf{N}} \langle \mathbf{x_i} - \mathbf{x_{i+1}}, \mathbf{y_i} \rangle \geq \mathbf{0}$$

Our aim is to relate cyclic monotonicity to convex functions. In dimension 1, cyclic monotonicity means that $A \subset \mathbb{R} \times \mathbb{R}$ supports a non-decreasing function.

Proof. (Homework!) □

We can consider a non-decreasing function as a "derivative" of a differentiable convex function

Recall that using sub-gradients/sub-differential definition, let us consider a function $\varphi : \mathbb{R}^n \to \overline{\mathbb{R}}$ is convex.

$$d\varphi(x) = \{y \in \mathbb{R}^n : \forall z \in \mathbb{R}^n \ \varphi(z) \geq \varphi(x) + \langle y, z - x \rangle\}$$

$$d\varphi = \{(x, y) : y \in d\varphi(x)\} \subset \mathbb{R}^n \times \mathbb{R}^n$$

Recall that $\varphi \in C^1(\mathbb{R}^n) \implies d\varphi(x) = \{\nabla \varphi(x)\}$

Theorem 7.25. *(Rockafellar Theorem) If $S \subset \mathbb{R}^n \times \mathbb{R}^n$, then S is cyclically monotone $\Leftrightarrow \exists$ a convex function $\varphi : \mathbb{R}^n \to \mathbb{R}$ s.t. $S \subset d\varphi$*

Proof. (Subgradients of convex functions are cyclically monotone.) Consider $(x_i, y_i)_{i=1}^N \subset S \subset d\varphi$ i.e. $\forall i\ y_i \in d\varphi(x_i)$, i.e. by definition $\varphi(z) \geq \varphi(x_i) + \langle y_i, z - x_i \rangle\ \forall z \in \mathbb{R}^n$
Plug $z = x_{i+1}$ $\varphi(x_{i+1}) \geq \varphi(x_i) + \langle y_i, x_{i+1} - x_i \rangle$
Sum in i - $\sum_i \varphi(x_{i+1}) \geq \sum_i \varphi(x_i) + \sum_i \langle y_i, x_{i+1} - x_i \rangle$
Now $\sum_i \varphi(x_{i+1}) = \sum_i \varphi(x_i)$ (Since $x_{N+1} = x_1$)
$\therefore \sum_i \langle y_i, x_{i+1} - x_i \rangle \leq 0$

□

Our goal is to construct a convex function φ such that set S (given that it isa cyclically monotone set) $\subset d\varphi$. We take -

$$\varphi(x) = sup(\langle y_N, x - x_N \rangle + \langle y_{N-1}, x_N - x_{N-1} \rangle + .. + \langle y_0, x_1 - x_0 \rangle | (x_i, y_i)_{i=1}^N \subset S) \quad (119)$$

Important notes -

- φ is convex since it is a *supremum* of linear functions.
- $\varphi(x_0) \geq 0$ as $\varphi \geq \langle y_1, x_1 - x_0 \rangle$, which follows from cyclic monotonicity
- $\varphi < \infty$ (Does not take the value ∞)

7.7 Kantorovich Duality

Definition 7.26. (Legendre Transform) Consider a convex function $\varphi : \mathbb{R}^n \to \overline{\mathbb{R}}$, then the dual function is -

$$\varphi^*(x) = \sup_{y \in \mathbb{R}^n} (\langle x, y \rangle - \varphi(y)) \quad (120)$$

7.7.1 Examples

Example 7.27.

$$f(x) = \frac{x^2}{2}$$

$\forall x \in \mathbb{R}$, *then*

$$f^*(x) = \left(\frac{x^2}{2}\right)^* = \frac{x^2}{2}$$

Example 7.28.

$$\varphi(x) = |x|$$

$\forall x \in \mathbb{R}^n$, *then*

$$\varphi^*(x) = \begin{cases} 0 & ; |x| \leq 1 \\ \infty & ; |x| > 1 \end{cases}$$

Example 7.29.

$$\left(\frac{t^p}{p}\right)^* = \frac{t^q}{q}$$

where $p, q \geq 0$, $\frac{1}{p} + \frac{1}{q} = 1$

Notes -

- If φ is convex $\implies \varphi^{**} = \varphi$
- $\forall x, y \in \mathbb{R}^n$

$$\varphi(x) + \varphi^*(y) \geq \langle x, y \rangle \tag{121}$$

- In the definition of φ^* (Eq. 120), the supremum (sup) is attained for $y = \nabla\varphi(x)$ if $\varphi \in C^1$. Then $\forall x$

$$\varphi(x) + \varphi^*(\nabla\varphi(x)) = \langle x, \nabla\varphi(x) \rangle$$

when $\varphi \in C^1$ is convex

Proof. (Homework!) □

Theorem 7.30. *(Kantorovich Duality) Consider probability measures* μ, ν *on* $\mathbb{R}^n$, *and* $\mathbb{E}_\mu X^2 < \infty$, $\mathbb{E}_\nu X^2 < \infty$.

Consider γ *to be a coupling of measures* μ *and* ν, *and two convex functions* φ *and* ψ *on* $\mathbb{R}^n$, *s.t.* $\varphi(x) + \psi(y) \geq \langle x, y \rangle$

$$\min_\gamma \int_{\mathbb{R}^n \times \mathbb{R}^n} -\langle x, y \rangle d\gamma(x, y) = \max_{\varphi, \psi} \left(\int_{\mathbb{R}^n} -\varphi(x) d\mu(x) + \int_{\mathbb{R}^n} -\psi(y) d\nu(y) \right)$$

$$= \max_\varphi \left(\int_{\mathbb{R}^n} -\varphi(x) d\mu(x) + \int_{\mathbb{R}^n} -\varphi^*(y) d\nu(y) \right)$$

In particular, these are attained and are finite.

Proof. Consider two functions $\varphi, \psi : \mathbb{R}^n \to \overline{\mathbb{R}}$, with γ begin the coupling between probability measures μ and ν. Now $\forall x, y \in \mathbb{R}^n$ -

$$\varphi(x) + \psi(y) \geq \langle x, y \rangle \tag{122}$$

Integrate this Eq. 122 w.r.t. any coupling γ of μ, ν.

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} -\langle x, y \rangle d\gamma(x, y) \geq \int_{\mathbb{R}^n \times \mathbb{R}^n} -(\varphi(x) + \psi(y)) d\gamma(x, y)$$

$$= \int_{\mathbb{R}^n} -\varphi(x) d\mu(x) + \int_{\mathbb{R}^n} -\psi(y) d\nu(y)$$

Take infimum (inf) in couplings (Since RHS does not see γ) Since LHS does not see φ and ψ, take supremum.

$$\inf_{\gamma} \int_{\mathbb{R}^n \times \mathbb{R}^n} -\langle x, y \rangle d\gamma(x,y) \geq \sup_{\varphi,\psi} \left(\int_{\mathbb{R}^n} -\varphi(x) d\mu(x) + \int_{\mathbb{R}^n} -\psi(y) d\nu(y) \right)$$

$$\geq \sup_{\varphi(\text{convex})} \left(\int_{\mathbb{R}^n} -\varphi(x) d\mu(x) + \int_{\mathbb{R}^n} -\varphi^*(y) d\nu(y) \right)$$

By Eq. 121, φ, φ^* satisfies the constraint. Consider $\overline{\gamma}$ to be the optimal coupling inf = min (By the theorem above)

By Rockafellar theorem, $\text{supp}(\overline{\gamma})$ is cyclically monotone. $\implies$ $\exists$ a convex function φ s.t.

$$\text{supp}(\overline{\gamma}) < d\varphi$$

$\implies$ a.e. w.r.t $\overline{\gamma}$ - Eq.121 is satisfied.

Hence the supremum is also attained and finite. □

Corollary 7.31. *Consider* $C(x,y) = \frac{|x-y|^2}{2}$. *Then,* $\overline{\gamma}$ *is the optimal coupling of probability measures* μ *and* ν.

$\Leftrightarrow \text{supp}(\overline{\gamma})$ *is cyclically monotone. (Proof: HW!)*

$\Leftrightarrow$ *(By Rockafellar)*$\text{supp}(\overline{\gamma})$ $\exists$ *a convex* $\varphi : \mathbb{R}^n \to \overline{\mathbb{R}}$ *s.t.* $\text{supp}(\overline{\gamma}) \subset d\varphi$

7.8 Brenier's Theorem

Theorem 7.32 (Brenier's Theorem)**.** *Suppose* $\int_{\mathbb{R}^n} |x|^2 \, d\mu + \int_{\mathbb{R}^n} |x|^2 \, d\nu < +\infty$, *where* μ, ν *are probability measures and* μ *is absolutely continuous. Then, there exists the unique optimal (with respect to quadratic loss) plan (coupling)* $\overline{\gamma}$ *of* μ, ν. *Moreover,* $\overline{\gamma} = (\text{Id} \times T)_{\#}\mu$, *i.e.,* $\overline{\gamma}$ *is realized by a unique transport map, and* $T = \nabla\varphi$ *for some convex function* φ.

Proof. (i) Existence (follows from general measure theory considerations)

Suppose $\overline{\gamma}$ is an optimal plan. By Corollary 7.31, we know that $\text{supp}(\overline{\gamma}) \subset \partial\varphi$ for some convex function φ. Also, we have

$$\varphi(x) + \varphi^*(y) = \langle x, y \rangle$$

on $\partial\varphi$ from Kantorovich duality. Thus, φ is finite μ-a.e., and by the Alexandrov Theorem (Convex functions are differentiable almost everywhere (w.r.t. Lebesgue) where they are not ∞), φ is C^1 μ-a.e. Let A be the set where φ is not differentiable. Then, $\mu(A) = 0$ and

$$\text{supp}(\overline{\gamma}) \cap (\mathbb{R}^n \setminus A \times \mathbb{R}^n) \subset \{(x, \nabla\varphi(x)) : x \in \mathbb{R}^n\}.$$

Consider any bounded $F \in C(\mathbb{R}^n \times \mathbb{R}^n)$. Then, we have

$$\begin{aligned}\int_{\mathbb{R}^n\times\mathbb{R}^n} F(x,y)\, d\bar{\gamma}(x,y) &= \int_{\mathbb{R}^n\times\mathbb{R}^n} F(x,\nabla\varphi(x))\, d\bar{\gamma}(x,y)\\ &= \int_{\mathbb{R}^n} F(x,\nabla\varphi(x))\, d\mu(x)\\ &= \int_{\mathbb{R}^n\times\mathbb{R}^n} F(x,y)\, d((\mathrm{Id}\times T)_{\#}\mu).\end{aligned}$$

Thus, we can conclude that $\overline{\gamma}$ is realized by optimal mass transport $T = \nabla\varphi$, with convex φ.

(ii) Uniqueness: HW □

Corollary 7.33. *Suppose μ, ν are probability measures, and μ is absolutely continuous. Then, there exists a map T with $T_\star\mu = \nu$ such that $Tx = \nabla\varphi(x)$ for some convex function φ. In particular,* $\mathrm{Jac}(T) = \nabla^2\varphi(x)$ *(non-negative definite) and*

$$\int f(x)\, d\nu(x) = \int f\circ T\, d\mu = \int f\cdot\det(\nabla^2\varphi)\, d\mu.$$

7.9 Mass transport proof of the Brunn-Minkowski inequality

Definition 7.34 (Minkowski sum)**.** Suppose K, L are Borel measurable sets in $\mathbb{R}^n$. The *Minkowski sum* of K and L is given as

$$K + L \coloneqq \{x + y : x \in K, y \in L\}.$$

Theorem 7.35 (Brunn-Minkowski inequality)**.** *Suppose K, L are Borel measurable sets. Then,*

$$|K+L|^{1/n} \geq |K|^{1/n} + |L|^{1/n}, \tag{123}$$

where $|\cdot|$ denotes the volume (Lebesgue measure).

Remark 7.36. *Note that for $\lambda > 0$, $|\lambda K| = \lambda^n |K|$. Thus, we can restate* (123) *as follows: For all $\lambda \in [0,1]$,*

$$|\lambda K + (1-\lambda)L|^{1/n} \geq |\lambda K|^{1/n} + |(1-\lambda)L|^{1/n} = \lambda\,|K|^{1/n} + (1-\lambda)\,|L|^{1/n}.$$

That is, $|\cdot|^{1/n}$ is concave with respect to the Minkowski addition.

Remark 7.37. *For $a, b > 0$ and $p \in (0,1)$. The Hölder inequality gives*

$$(\lambda a^p + (1-\lambda)b^p)^{1/p} \geq a^\lambda b^{1-\lambda},$$

(Check for HW). Substituting $p = \frac{1}{n}$, $a = |K|$ and $b = |L|$, we obtain

$$(\lambda\,|K|^{1/n} + (1-\lambda)\,|L|^{1/n})^n \geq |K|^\lambda\,|L|^{1-\lambda}.$$

Under the same assumptions as Theorem 7.35, the Brunn-Minkowski inequality implies

$$|\lambda K + (1-\lambda)L| \geq |K|^{\lambda} |L|^{1-\lambda} \tag{124}$$

In fact, (124) with all $\lambda \in (0,1)$ implies (123) (Check for HW. Idea: Select right λ and use homogeneity of $|\cdot|$).

Recall the isoperimetric inequality: Suppose A is a Borel measurable set, and pick $R > 0$ such that $|A| = |RB_2^n|$. Then, we have

$$|\partial A| \geq |\partial(RB_2^n)|,$$

that is, among all the sets of fixed volume, the Euclidean ball minimizes the perimeter. This is equivalent to saying

$$\frac{|\partial A|}{|A|^{(n-1)/n}} \geq \frac{|\partial(RB_2^n)|}{|RB_2^n|^{(n-1)/n}} = \frac{|\partial B_2^n|}{|B_2^n|^{(n-1)/n}}.$$

Indeed, for all $t > 0$,

$$\frac{|\partial A|}{|A|^{(n-1)/n}} = \frac{t^{n-1}|\partial A|}{t^{n\cdot(n-1)/n}|A|^{(n-1)/n}} = \frac{|\partial(tA)|}{|tA|^{(n-1)/n}}.$$

Also, note that

$$\left|S^{n-1}\right| = n \cdot |B_2^n|,$$

since $|B_2^n| = \int_{B_2^n} dx = \int_0^1 \int_{S^{n-1}} d\theta dt = |S^{n-1}| \cdot \int_0^1 t^{n-1} dt = \frac{1}{n}|S^{n-1}|$. Therefore, we have

$$\frac{|\partial B_2^n|}{|B_2^n|^{(n-1)/n}} = \frac{n|B_2^n|}{|B_2^n|^{(n-1)/n}} = n|B_2^n|^{1/n},$$

and thus, the isoperimetric inequality implies, for any Borel measurable set A,

$$\frac{|\partial A|}{|A|^{(n-1)/n}} \geq n|B_2^n|^{1/n}. \tag{125}$$

Proof of (125) *via the Brunn-Minkowski inequality.* Note that

$$\begin{aligned}
|\partial A| &= \liminf_{\varepsilon \to 0} \frac{|(A + \varepsilon B_2^n) \setminus A|}{\varepsilon} \\
&= \lim_{\varepsilon \to 0} \frac{|A + \varepsilon B_2^n| - |A|}{\varepsilon} && (A \text{ Borel}) \\
&\geq \lim_{\varepsilon \to 0} \frac{\left(|A|^{1/n} + |\varepsilon B_2^n|^{1/n}\right)^n - |A|}{\varepsilon} && (\text{B-M inequality, raised to the power } n) \\
&= n|A|^{(n-1)/n}|B_2^n|^{1/n}.
\end{aligned}$$

This gives

$$\frac{|\partial A|}{|A|^{(n-1)/n}} \geq n\,|B_2^n|^{1/n} = \frac{|\partial B_2^n|}{|B_2^n|^{(n-1)/n}}.$$

□

Proof of Theorem 7.35 via mass transport. Without loss of generality, assume that $|K| \neq 0, |L| \neq 0$, and let μ, ν be normalized uniform measures on K, L, respectively. Then, we have

$$d\mu(x) = \mathbb{1}_K(x)\frac{1}{|K|}dx,$$
$$d\nu(x) = \mathbb{1}_L(x)\frac{1}{|L|}dx.$$

Consider the Brenier map T such that $T_\star\mu = \nu$. Since T is a transport map and $T = \nabla\varphi$ for some convex φ, we know that for all $x \in K$, $f_\mu(x) = f_\nu(x) \cdot \det(\nabla^2\varphi)$. That is,

$$\frac{1}{|K|} = \frac{1}{|L|} \cdot \det(\nabla^2\varphi). \tag{126}$$

Then, consider $\lambda \in (0, 1]$, and let

$$T_\lambda = (1-\lambda)\mathrm{Id} + \lambda T.$$

That is, $x \overset{T_\lambda}{\longmapsto} (1-\lambda)x + \lambda Tx$. Then, we have

$$(T_\lambda)_\star\mu = \mu_\lambda,$$

which is supported on $(1-\lambda)K + \lambda L$ since

$$\mathrm{supp}(\mu_\lambda) = \{z : \exists x \in K, y \in L \text{ s.t. } \lambda x + (1-\lambda)y = z\}.$$

Let f_λ be the density of μ_λ. Then, the transport equation gives, for all $x \in K$,

$$\frac{1}{|K|} = f_\lambda(T_\lambda(x))\det(\mathrm{Jac}(T_\lambda)).$$

Recall that (123) is equivalent to

$$|(1-\lambda)K + \lambda L| \geq |K|^{1-\lambda}\,|L|^{\lambda}\,. \tag{127}$$

Therefore, since $|(1-\lambda)K + \lambda L| = |\mathrm{supp}(\mu_\lambda)|$, it suffices to show

$$f_\lambda(T_\lambda(x)) \leq \frac{1}{|K|^{1-\lambda}\,|L|^{\lambda}}. \tag{128}$$

Indeed, if (128) holds, then

$$\int_{(1-\lambda)K+\lambda L} f_\lambda(y)dy \leq |(1-\lambda)K + \lambda L| \sup_{x\in K} f_\lambda(T_\lambda(x)) \leq \frac{|(1-\lambda)K + \lambda L|}{|K|^{1-\lambda}\,|L|^{\lambda}},$$

while $\int_{(1-\lambda)K+\lambda L} f_\lambda(y)dy = 1$ since f_λ is a density of a probability measure. Thus, (127) and (123) would follow.

Note that (128) is equivalent to

$$\frac{1}{|K|\det(\mathrm{Jac}((1-\lambda)\mathrm{Id}+\lambda T))} \le \frac{1}{|K|^{1-\lambda}|L|^{\lambda}},$$

i.e.,

$$\det(\mathrm{Jac}((1-\lambda)\mathrm{Id}+\lambda T)) \ge \left(\frac{|L|}{|K|}\right)^{\lambda} = \det(\mathrm{Jac}(T))^{\lambda}. \tag{129}$$

We know that $\mathrm{Jac}(T) = \nabla^2\varphi$, where φ is a convex function. That is, $Jac(T)$ is a non-negative definite matrix. Without loss of generality, by choosing the right basis (depending on x), we can assume

$$T = \mathrm{diag}(t_1, \ldots, t_n),\ t_i \ge 0.$$

Then, (129) reduces to

$$\det(\mathrm{diag}(1-\lambda+\lambda t_1, \cdots, 1-\lambda+\lambda t_n)) \ge \det(\mathrm{diag}(t_1, \cdots, t_n))^{\lambda},$$

which is true by the AM-GM inequality ($1-\lambda+\lambda t_i \ge 1^{1-\lambda}t_i^{\lambda} = t_i^{\lambda}$). □

7.10 Log-concave measures

Definition 7.38. μ on $\mathbb{R}^n$ is called *log-concave* if for all Borel-measurable sets $K, L \in \mathbb{R}^n$,

$$\mu(\lambda K + (1-\lambda)L) \ge \mu(K)^{\lambda}\mu(L)^{1-\lambda},$$

for all $\lambda \in (0,1]$, i.e., if μ satisfies the Brunn-Minkowski inequality.

Theorem 7.39 (Borell Theorem)**.** *Suppose $d\mu = f(x)dx$ on $\mathbb{R}^n$ and $\mathrm{supp}(f) \neq \emptyset$. Also, suppose f is a log-concave function (i.e., $\log f$ is a concave function, i.e., for all $x, y \in \mathbb{R}^n$ and $\lambda \in [0.1]$, it holds that $f(\lambda x + (1-\lambda)y) \ge f(x)^{\lambda}f(y)^{1-\lambda}$). Then, μ is log-concave.*

Example 7.40. *The Gaussian measure is log-concave since*

$$\log e^{-|x|^2/2+c} = -\frac{|x|^2}{2} + c$$

Example 7.41. *All of the measures with the densities of the form $e^{-|x|}$, $e^{-\|x\|}$, $e^{-(\text{convex function})}$ are log-concave.*

Theorem 7.42 (Prékopa–Leindler inequality)**.** *Let $\lambda \in (0,1)$ and suppose $f, g, h : \mathbb{R}^n \to \mathbb{R}^{\ge 0}$ satisfy*

$$h(\lambda x + (1-\lambda)y) \ge f(x)^{\lambda}g(y)^{1-\lambda},$$

for all $x, y \in \mathbb{R}^n$. Then,

$$\int h \ge \left(\int f\right)^{\lambda}\left(\int g\right)^{1-\lambda}$$

Note that this implies the Borell Theorem for convex sets by substituting $f = \mathbb{1}_K d\mu$, $g = \mathbb{1}_L d\mu$, $h = \mathbb{1}_{\lambda K+(1-\lambda)L}d\mu$.

7.11 Applications of Brenier's therorem in HDP; Caffarelli's theorem

Remark 7.43. *Recall the following. If μ, ν are probability measures on $\mathbb{R}^n$ and μ is absolutely continuous, then there exists a transport map $T_{\#}\mu = \nu$ where $Tx = \nabla\varphi(x)$ for some convex function $\varphi : \mathbb{R}^n \to \mathbb{R}^n$.*

We can always do a change of variables. Suppose $g : \mathbb{R}^n \to \mathbb{R}$. Since T is a transport map, this means

$$\begin{aligned}\int g(x)d\mu(x) &= \int g(Tx)d\nu(x)\\ &= \int g(y)\mathrm{Jac}(T)d\nu(y)\\ &= \int g(y)\mathrm{det}\nabla^2\varphi(y)d\mu(\nabla\varphi)\end{aligned}$$

Since this is true for any function g (modulo some regularity issues), then

$$d\mu(x) = \mathrm{det}\nabla^2\varphi(x)d\nu(\nabla\varphi), \forall x \in \mathrm{supp}(\nu)$$

Since φ is convex, $\nabla^2\varphi$ is a non-negative definite matrix. This is referred to as the transport equation or Monge-Ampère type equation.

Remark 7.44. *Recall that we say a measure μ on $\mathbb{R}^n$ is strictly log-concave if $d\mu(x) = e^{-V(x)}dx$ for any $x \in \mathbb{R}^n$ where $\nabla^2 V \geq \mathbb{I}_d$. In the limit, any $d\mu(x) = \mathbb{1}_K(x)d\gamma(x)$, where $K \subset \mathbb{R}^n$ is a convex set, μ is a strictly log-concave measure.*

More generally, $V(x) = \frac{|x|^2}{2} + W(x)$, where W is a convex function. Then $\nabla^2 V = \mathbb{I}_d + \nabla^2 W \geq \mathbb{I}_d$. Also, $d\mu = \mathbb{1}_K(x)d\gamma(x) = \exp(-(\frac{|x|^2}{2} + W(x)))dx$ where

$$W(x) = \begin{cases}\inf & x \notin K\\ c & x \in K\end{cases}$$

Theorem 7.45 (Caffarelli's theorem)**.** *Let T be the Brenier map $T_{\#}\gamma = \mu$ for some strictly log-concave measure μ. Then T is 1-Lipshitz, i.e. for any $x, y \in \mathbb{R}^n$, $|Tx - Ty| \leq |x - y|$.*

Example 7.46. *(HW) Consider pair of convex sets $K \subset L$, $d\mu = \mathbb{1}_K dx$, $d\nu = \mathbb{1}_L dx$. If $T_{\#}\mu = \nu$ and T is Brenier, T is not necessarily Lipshitz.*

Remark 7.47. *Recall if there exists a 1-Lipshitz map T with $T_{\#}\mu = \nu$, then a lot of the nice properties of μ can be translated to ν.*

Remark 7.48. *Recall the Gaussian Poincaré inequality. For any $f : \mathbb{R}^n \to \mathbb{R}^n$,*

$$\int f^2 d\gamma - \left(\int f d\gamma\right)^2 \leq \int |\nabla f|^2 f\gamma.$$

Corollary 7.49. *Using Remark 7.48 and Caffarelli's theorem, $\forall \mu$ strictly log-concave (i.e. $d\mu = \exp(-(\frac{|x|^2}{2} + W(x)))dx$) where W is a convex function), we have $\forall f : \mathbb{R}^n \to \mathbb{R}$ reasonable,*

$$\int f^2 d\mu - \left(\int f d\mu\right)^2 \leq \int |\nabla f|^2 f\mu$$

Theorem 7.50 (Slightly more general than Caffarelli)**.** *Consider the probability measures on $\mathbb{R}^n$ $d\mu = e^{-V(x)}dx$ and $d\nu = e^{-W(x)}dx$. We assume that $V, W \in C^2(\mathbb{R}^n)$ and $\nabla^2 W \geq K \cdot \mathbb{I}_d$ where $K > 0$. Also, let $T_{\#}\mu = \nu$ s.t. $Tx = \nabla\varphi(x)$ for some φ convex. Then for any vector $e \in \mathbb{R}^n$,*

$$\sup_{x\in\mathbb{R}^n} \varphi_{ee}^2 \leq \frac{1}{K} \sup_{x\in\mathbb{R}^n} V_{ee},$$

where $f_e = \langle \nabla f, e\rangle$. In words, it denotes the derivative in the direction e. Consequently, $f_{ee} = \langle \nabla f e, e\rangle$ and more generally, $f_{ev} = \langle \nabla f e, v\rangle = \langle \nabla f v, e\rangle$.

Example 7.51. *If $\mu = \gamma$ Gaussian, then $V(x) = \frac{x^2}{2}$. So $V_{ee} = 1$ for all x. Then the theorem implies that*

$$sup_{x\in\mathbb{R}^n} \varphi_{ee}^2 \leq \frac{1}{K}$$

If we consider the situation where $K = 1$ which corresponds to $\nabla^2 W \geq 1$, or in other words ν is strictly log-concave, then

$$sup_{x\in\mathbb{R}^n} \varphi_{ee}^2 \leq 1,$$

which precisely means that $\nabla\varphi$ is a 1-Lipshitz map.

Let us now sketch out the proof by skipping some of the computation. We start with the transport equation:

$$e^{-V} = e^{-W(\nabla\varphi)} \det(\nabla^2\varphi)$$

We take a logarithm on both sides

$$V(x) = W(\nabla\varphi(x)) - log\det\nabla^2\varphi(x) \tag{130}$$

Let $e \in S^{n-1}$, and differentiate both sides in the direction e

$$\partial_e \log\det\nabla^2\varphi = \frac{\partial_e \det\nabla^2\varphi}{\det(\nabla^2\varphi)} = \operatorname{tr}\left((\nabla^2\varphi)^{-1} \cdot \nabla^2\varphi_e\right),$$

where the last step requires some involved matrix calculus (HW). Then, we differentiate again

$$\partial_{ev} \log \det \nabla^2 \varphi = \mathrm{tr}(\nabla^2\varphi)^{-1}\nabla^2\varphi_{ev} - \mathrm{tr}\left[(\nabla^2\varphi)^{-1}\nabla^2\varphi_e(\nabla^2\varphi)^{-1}\nabla^2\varphi_v\right] \tag{131}$$

Combining Equations 130 and 131,

$$V_{ee} = \langle(\nabla^2 W(\nabla\varphi)e, \nabla^2\varphi e\rangle + \langle \nabla W(\nabla\varphi), \nabla\varphi_{ee}\rangle - \mathrm{tr}\left((\nabla^2\varphi)^{-1}\nabla^2\varphi_{ee}\right) + \mathrm{tr}\left((\nabla^2\varphi)^{-1}\nabla^2\varphi_e\right)^2$$

Note that the above is true for any $x \in \mathbb{R}^n$. Suppose φ_{ee} attains maximum at $x_o \in \mathbb{R}^n$. Then $\nabla\varphi_{ee} = 0$ at x_o and $\nabla^2\varphi_{ee} \leq 0$ at x_o. Therefore,

$$V_{ee}(x_o) \geq \langle \nabla^2 W(\nabla\varphi)e, \nabla^2\varphi e\rangle|_{x_o}$$

But recall that $\nabla^2 W \geq K \cdot \mathbb{I}_d$. Then,

$$V_{ee}(x_o) \geq K \cdot ||\nabla^2\varphi(x_o)e||^2 \geq K \cdot \varphi_{ee}^2(x_o)$$

So,

$$\varphi_{ee}^2(x_o) = \sup_{x\in\mathbb{R}^n} \varphi_{ee}^2(x)$$

□

7.12 The Talagrand transport inequality

Theorem 7.52. *Suppose $T_{\#}\gamma = \mu$ where γ is a standard Gaussian and T is the Brenier map. Then,*

$$\frac{1}{2}\int |Tx - x|^2 d\gamma \leq \mathrm{Ent}(\mu||\gamma)$$

Let us prove this theorem. Let $g(x) = \sqrt{2\pi}^{-n} e^{-\frac{|x|^2}{2}}$ be the Gaussian density and $f(x) = \frac{d\mu}{dx}$ and $Tx = \nabla F(x)$ for some convex function F. Again, we ignore some regularity issues. Using the transport equation, for all $x \in \mathbb{R}^n$

$$f(\nabla F(x))\det\nabla^2 F(x) = g(x)$$

Changing the notation,

$$
\begin{aligned}
Ent(\mu||\gamma) &= \int f \log \frac{f}{g} dy \\
&= \int f(\nabla F(x)) \log \frac{f(\nabla F(x))}{g(\nabla F(x))} det \nabla^2 F(x) dx \qquad \text{change of variables } [y = \nabla f] \\
&= \int \log \frac{f(Tx)}{g(Tx)} \cdot g(x) dx \qquad \text{transport equation} \\
&= \int \log \frac{g(x)}{g(Tx) \det \nabla^2 F(x)} f\gamma \qquad \text{transport equation} \\
&= \left[\log g(x) = \frac{-|x|^2}{2} + c_n \right] \\
&= \int \left(\frac{|Tx|^2 - |x|^2}{2} - \log \det \nabla^2 F(x) \right) d\gamma
\end{aligned}
$$

We can conclude that

$$
\mathrm{Ent}(\mu||\gamma) = \int \left(\frac{|Tx|^2 - |x|^2}{2} - \log \det \nabla^2 F(x) \right) d\gamma
$$

But we need

$$
Ent(\mu||\gamma) \geq \frac{1}{2} \int |Tx - x|^2 d\gamma
$$

So the goal is to show that

$$
\int \log \det \nabla^2 F d\gamma \leq \int (\langle x, \nabla F \rangle - |x|^2) d\gamma
$$

Recall Gaussian integration by parts:

$$
\int \nu \mathrm{L} u d\gamma = - \int \langle \nabla \nu, \nabla u \rangle d\gamma,
$$

where $\mathrm{L}u = \nabla u - \langle \nabla u, x \rangle$. So in particular,

$$
0 = \int \mathrm{L} f d\gamma = \int (\Delta F - \langle \nabla F, x \rangle) d\gamma
$$

Thus,

$$
\int \langle \nabla F, x \rangle d\gamma = \int \Delta F d\gamma
$$

Also,

$$
\int x^2 d\gamma = n
$$

Our goal therefore is just to show

$$\int \log\det \nabla^2 F d\gamma \leq \int (\Delta F - n) d\gamma$$

Let's claim that the above identity holds pointwise. $\nabla^2 F = A$ where A is a non-negative definite matrix for all x. Pointwise, $\log\det \nabla^2 F \geq \Delta F - n$ because this becomes equivalent to $\log \prod_{i=1}^n t_i$ where t_i are eigenvalues of A, meaning $t_i \geq 0$. Then,

$$\sum_{i=1}^n \log t_i \geq \sum_{i=1}^n (t_i - 1)$$

The above is true because it is true for any $t \geq 0$, meaning $\log t \leq t - 1$. □

Remark 7.53. *Log-Sobolev implies that*

$$\mathrm{Ent}(f d\gamma \parallel \gamma) \leq \frac{1}{2} \int \frac{|\nabla f|^2}{f} d\gamma$$

This is also true for strictly log-concave measures by Caffarelli. Now combining Log-Sobolev and Talagrand,

$$W_2(\mu, \gamma)^2 \leq \frac{1}{2} \int \frac{|\nabla f|^2}{f} d\gamma$$

where $d\mu = f d\gamma$.

References

[1] S. Artstein-Avidan, A. Giannopoulos, V. D. Milman, *Asymptotic Geometric Analysis, part I,* 2015.

[2] N. Alon, B. Klartag, *Optimal compression of approximate inner products and dimension reduction*, Symposium on Foundations of Computer Science (FOCS 2017), 639-650.

[3] Z. D. Bai and Y. Q. Yin, *Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix,* Ann. Probab. 16 (1988), no. 4, 1729-1741. MR0958213

[4] Z. D. Bai,Y. Q. Yin, *Limit of the smallest eigenvalue of a large-dimensional sample covariance matrix,* Ann. Probab. 21 (1993), 1275-1294.

[5] J. Bourgain, V. H. Vu and P. M. Wood, *On the singularity probability of discrete random matrices,* J. Funct. Anal. 258 (2010), no. 2, 559-603. MR2557947

[6] N. Cook, *Lower bounds for the smallest singular value of structured random matrices,* Ann. Probab. **46** (2018), no. 6, 3442–3500.

[7] R. Durrett, *Probability: Theory and examples,* 2019.

[8] A. Edelman, *Eigenvalues and condition numbers of random matrices*, SIAM J. Matrix Anal. Appl. 9 (1988), 543-560.

[9] C. G. Esseen, *On the Kolmogorov-Rogozin inequality for the concentration function*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 5 (1966), 210-216

[10] O. N. Feldheim and S. Sodin, *A universality result for the smallest eigenvalues of certain sample covariance matrices,* Geom. Funct. Anal. 20 (2010), no. 1, 88-123. MR2647136

[11] Y. Gordon, *Some inequalities for Gaussian processes and applications, Israel J. Math. 50 (1985), 265-289.*

[12] O. Guedon, A. Litvak, K. Tatarko, *Random polytopes obtained by matrices with heavy tailed entries*, Commun. Contemp. Math. **22** (2020), no. 4, 1950027, 28 pp.

[13] R. van Handel, R. Latala, P. Youssef, *The dimension-free structure of nonhomogeneous random matrices*, Invent. Math. **214** (2018), no. 3, 1031–1080.

[14] V. Jain, S. Silwal. *A note on the universality of ESDs of inhomogeneous random matrices*, `arXiv preprint arXiv:2006.05418 (2020)`.

[15] J. Kahn, J. Komlos, E. Szemeredi, *On the probability that a random* ± 1 *matrix is singular,* J. Amer. Math. Soc. 8 (1995), 223-240.

[16] R. Kannan, S. Vempala, *Sampling Lattice Points*, Proc. 29th ACM Symposium on the Theory of Computing (STOC '97), El Paso, (1997), Invited for publication in Journal of Comp. and System Sciences.

[17] B. Klartag, *A central limit theorem for convex sets*, Invent. Math., Vol. 168, (2007), 91–131.

[18] B. Klartag, G. V. Livshyts, *The lower bound for Koldobsky's slicing inequality via random rounding*, Lect. notes GAFA seminar 2019.

[19] B. Klartag, V.D. Milman, *Geometry of log-concave functions and measures*, Geom. Dedicata 112 (2005) 169-182.

[20] J. E. Littlewood, A. C. Offord, *On the number of real roots of a random algebraic equation. III,* Rec. Math. [Mat. Sbornik] N.S. 12 (54), (1943), 277-286

[21] A. Litvak, A. Pajor, M. Rudelson, N. Tomczak-Jaegermann, Smallest singular value of random matrices and geometry of random polytopes, Adv. Math. **195** (2005), no. 2, 491–523.

[22] A. E. Litvak and O. Rivasplata, Smallest singular value of sparse random matrices, Studia Math. **212** (2012), no. 3, 195–218.

[23] G. Livshyts, G. Paouris, P. Pivovarov, *Small deviations for operator norms*, work in progress

[24] A. E. Litvak, S. Spektor, *Quantitative version of a Silverstein's result,* GAFA, Lecture Notes in Math., 2116 (2014), 335–340.

[25] G. V. Livshyts, *The smallest singular value of heavy-tailed not necessarily i.i.d. random matrices via random rounding*, Journal d'Analyze Mathematique 2020.

[26] G. V. Livshyts, K. Tikhomirov, R. Vershynin, *The smallest singular value of inhomogeneous square random matrices,* Ann. Probab. 49 (2021), no. 3, 1286-1309.

[27] A. Lytova, K. Tikhomirov, *On delocalization of eigenvectors of random non-Hermitian matrices*, Probab. Theory Related Fields **177** (2020), no. 1-2, 465–524.

[28] J. von Neumann, H. H. Goldstine, *Numerical inverting of matrices of high order*, Bull. Amer. Math. Soc. 53 (1947), 1021-1099.

[29] S. Mendelson, G. Paouris, *On the singular values of random matrices,* Journal of the European Mathematics Society, 16, 823-834, 2014.

[30] G. Pólya, Szegö, *Aufgaben und Lehrsätze aus der Analysis. Band I: Reihen. Integralrechnung. Funktionentheorie.* Dritte berichtigte Auflage. Die Grundlehren der Mathematischen Wissenschaften, Band 19 Springer-Verlag, Berlin-New York 1964.

[31] P. Raghavan, C. D. Thompson, *Randomized rounding: A technique for provably good algorithms and algorithmic proofs,* Combinatorica, 7(4):365-374.

[32] E. Rebrova, K. Tikhomirov, *Coverings of random ellipsoids, and invertibility of matrices with i.i.d. heavy-tailed entries*, Israel J. Math. **227** (2018), no. 2, 507–544.

[33] B. A. Rogozin, *An estimate for the maximum of the convolution of bounded densities*, Teor. Veroyatnost. i Primenen. 32 (1987), no. 1, 53-61, English translation:Theory Probab. Appl. 32 (1987), no. 1, 48-56.

[34] M. Rudelson, *Invertibility of random matrices: norm of the inverse*, Annals of Mathematics 168 (2008), 575-600.

[35] M. Rudelson, R. Vershynin, *The Littlewood-Offord problem and invertibility of random matrices*, Adv. Math. 218 (2008), no. 2, 600-633.

[36] M. Rudelson, R. Vershynin, *The least singular value of a random square matrix is* $O(\sqrt{n}^{-1})$, Comptes rendus de l'Academie des sciences - Mathematique 346 (2008), 893-896.

[37] M. Rudelson, R. Vershynin, *Smallest singular value of a random rectangular matrix,* Communications on Pure and Applied Mathematics 62 (2009), 1707-1739.

[38] M. Rudelson, R. Vershynin, *Non-asymptotic theory of random matrices: extreme singular values,* Proceedings of the International Congress of Mathematicians, 2010, pp. 83-120.

[39] M. Rudelson, R. Vershynin, *Small ball probabilities for linear images of high dimensional distributions,* Int. Math. Res. Not. 19 (2015), 9594-9617.

[40] M. Rudelson, R. Vershynin, *Delocalization of eigenvectors of random matrices with independent entries,* Duke Math. J. Volume 164, Number 13 (2015), 2507-2538.

[41] C. Schütt, *Entropy numbers of diagonal operators between symmetric Banach spaces,* J. Approx. Theory 40 (1984), 121–128.

[42] A. Srinivasan, *Approximation Algorithms via Randomized Rounding: a Survey*, Lectures on Approximation and Randomized Algorithms, Series in Advanced Topics in Mathematics, Polish Scientific Publishers PWN, Warsaw, 9-71, (1999).

[43] S. Szarek, *Condition numbers of random matrices, J. Complexity 7 (1991),* 131-149.

[44] K. Tatarko, *An upper bound on the smallest singular value of a square random matrix*, J. Complexity **48** (2018), 119–128.

[45] T. Tao and V. Vu, *On random* ± 1 *matrices: Singularity and Determinant,* Random Structures and Algorithms 28 (2006), no 1, 1-23.

[46] T. Tao, V. Vu, *On the singularity probability of random Bernoulli matrices*, J. Amer. Math. Soc. 20 (2007), 603-628.

[47] T. Tao and V. Vu, *Inverse Littlewood-Offord theorems and the condition number of random discrete matrices*, Annals of Mathematics 169 (2009), 595-632.

[48] T. Tao and V. Vu, *Random matrices: the distribution of the smallest singular values,* Geom. Funct. Anal. 20 (2010), no. 1, 260-297. MR2647142

[49] K. Tikhomirov, *The limit of the smallest singular value of random matrices with i.i.d. entries,* Adv. Math. 284 (2015), 1-20.

[50] K. Tikhomirov, *The smallest singular value of random rectangular matrices with no moment assumptions on entries*, Israel J. Math. 212 (2016), no. 1, 289-314.

[51] K. Tikhomirov, *Invertibility via distance for non-centered random matrices with continuous distributions*, Random Structures and Algorithms, to appear.

[52] K. Tikhomirov, *Singularity of random Bernoulli matrices*, Ann. of Math. (2) **191** (2020), no. 2, 593–634.

[53] R. Vershynin, *High-dimensional probability: an introduction with applications in data science*, Cambridge University Press, 2018.

[54] R. Vershynin, *Spectral norm of products of random and deterministic matrices,* Probability Theory and Related Fields 150 (2011), 471-509.

[55] M. S. Viazovska, *The sphere packing problem in dimension* 8, Annals of mathematics (2017), 991-1015.

[56] H. Cohn, A. Kumar, S. Miller, D. Radchenko, M. S. Viazovska, *The sphere packing problem in dimension* 24, Annals of mathematics (2017), 185 (3), 1017-1033.

[57] Shiryaev, A. N. *Probability, vol. 1*, Graduate Texts in Mathematics 95 (1989).

[58] S. Jukna *Extremal combinatorics: with applications in computer science*, Springer 512 (2011).