Lecture 5: Concentration of Measure and Isoperimetry

Tianpei Xie

Jan. 19th., 2023

Contents

1 The Classic Isoperimetry Inequalities

1.1 Brunn-Minkowski Inequality

- **Definition** (***Minkowski Sum of Sets***)
  Consider sets $A, B \subseteq \mathbb{R}^n$ and define ***the Minkowski sum*** of A and B as the set of all vectors in $\mathbb{R}^n$ formed by sums of elements of A and B:

$$A + B := \{x + y : x \in A, y \in B\}$$

  Similarly, for $c \in \mathbb{R}$, let $cA = \{cx : x \in A\}$. Denote by $\mathrm{Vol}(A)$ the ***Lebesgue measure*** of a *(measurable) set* $A \subset \mathbb{R}^n$.

- **Theorem 1.1** *(**Brunn-Minkowski Inequality**) [Boucheron et al., 2013, Vershynin, 2018, Wainwright, 2019]*
  *Let $A, B \subset \mathbb{R}^n$ be **non-empty compact sets**. Then for all $\lambda \in [0, 1]$,*

$$Vol\,(\lambda A + (1-\lambda)B)^{\frac{1}{n}} \ge \lambda\, Vol(A)^{\frac{1}{n}} + (1-\lambda)\, Vol(B)^{\frac{1}{n}}. \tag{1}$$

  Note: a convex body in $\mathbb{R}^n$ is closed and compact set.

  Proof: (***Part 1,*** $n = 1$)
  Note that if $A \subset \mathbb{R}$, and $c \ge 0$ then $\mathrm{Vol}(cA) = c\mathrm{Vol}(A)$. Thus it suffice to prove

$$\mathrm{Vol}\,(A + B) \ge \mathrm{Vol}(A) + \mathrm{Vol}(B).$$

  To see this, observe that none of the three volumes involved changes if the sets A and B are ***translated*** arbitrarily. Since A, B are compact subsets in $\mathbb{R}$, it is closed and bounded. Let $a = \max\{a' : a' \in A\}$ and $b = \min\{b' : b' \in B\}$. Let $A' = A + \{-a\}$ and $B' = B + \{-b\}$ so that $A' \subset (-\infty, 0]$ and $B' \subset [0, +\infty)$. Also $\mathrm{Vol}(A') = \mathrm{Vol}(A)$ and $\mathrm{Vol}(B') = \mathrm{Vol}(B)$. However,

$$\begin{aligned} A' \cup B' &\subset A' + B' \\ \Rightarrow \mathrm{Vol}(A') + \mathrm{Vol}(B') = \mathrm{Vol}(A' \cup B') &\le \mathrm{Vol}(A' + B') \end{aligned}$$

  This prove the 1-dimensional case for *the Brunn-Minkowski inequality.* ∎

  To prove $n > 1$ case, we need the following inequalities:

- **Theorem 1.2** *(**The Prékopa-Leindler Inequality**). [Boucheron et al., 2013, Wainwright, 2019]*
  *Let $\lambda \in (0, 1)$, and let $f, g, h : \mathbb{R}^n \to [0, \infty)$ be **non-negative measurable functions** such that for all $x, y \in \mathbb{R}^n$,*

$$h\,(\lambda x + (1-\lambda)y) \ge f(x)^{\lambda} g(y)^{1-\lambda}.$$

  Then

$$\int_{\mathbb{R}^n} h(x)dx \ge \left(\int_{\mathbb{R}^n} f(x)dx\right)^{\lambda} \left(\int_{\mathbb{R}^n} g(x)dx\right)^{1-\lambda}. \tag{2}$$

  Proof: The proof goes by induction with respect to the dimension n.

1. ($n = 1$ ***case***). Consider measurable non-negative functions f, g, h satisfying the condition of the theorem. By *the monotone convergence theorem*, it suffices to prove the statement for ***bounded functions*** f and g. Without loss of generality, assume that $\sup_{x\in\mathbb{R}^n} f(x) = \sup_{x\in\mathbb{R}^n} g(x) = 1$. Then

$$\int_{\mathbb{R}} f(x)dx = \int_0^1 \mathrm{Vol}\,\{x : f(x) \geq t\}\,dt$$
$$\int_{\mathbb{R}} g(x)dx = \int_0^1 \mathrm{Vol}\,\{x : g(x) \geq t\}\,dt.$$

For any fixed $t \in [0,1]$, if $f(x) \geq t$ and $g(y) \geq t$, then by the hypothesis of the theorem, $h\,(\lambda x + (1-\lambda)y) \geq t$. This implication may be re-written as

$$\lambda\,\{x : f(x) \geq t\} + (1-\lambda)\,\{x : g(x) \geq t\} \subset \{x : h(x) \geq t\}\,.$$

Thus

$$\begin{aligned}
\int_{\mathbb{R}} h(x)dx &= \int_0^\infty \mathrm{Vol}\,\{x : h(x) \geq t\}\,dt \\
&\geq \int_0^1 \mathrm{Vol}\,\{x : h(x) \geq t\}\,dt \\
&\geq \int_0^1 \mathrm{Vol}\,(\lambda\,\{x : f(x) \geq t\}) + \mathrm{Vol}\,((1-\lambda)\,\{x : g(x) \geq t\})\,dt \\
&(\text{ by 1-dimensional } \textit{Brunn-Minkowski inequality}) \\
&\geq \lambda \int_0^1 \mathrm{Vol}\,(\{x : f(x) \geq t\})\,dt + (1-\lambda) \int_0^1 \mathrm{Vol}\,(\{x : g(x) \geq t\})\,dt \\
&= \lambda \int_{\mathbb{R}} f(x)dx + (1-\lambda) \int_{\mathbb{R}} g(x)dx \\
&\geq \left(\int_{\mathbb{R}} f(x)dx\right)^{\lambda} \left(\int_{\mathbb{R}} g(x)dx\right)^{1-\lambda} \quad (\text{by the } \textit{arithmetic-geometric mean inequality})
\end{aligned}$$

2. For the induction step, assume that the theorem holds for all dimensions $1, \ldots, n-1$ and let $f, g, h : \mathbb{R}^n \to [0,\infty)$, $\lambda \in (0,1)$ be such that they satisfy the assumption of the theorem. Now let $x, y \in \mathbb{R}^{n-1}$ and $a, b \in \mathbb{R}$. Then

$$h\,(\lambda\,(x,a) + (1-\lambda)\,(y,b)) \geq f\,((x,a))^{\lambda}\, g((y,b))^{1-\lambda},$$

so by the inductive hypothesis

$$\int_{\mathbb{R}^{n-1}} h\,((x, \lambda a + (1-\lambda)b))\,dx \geq \left(\int_{\mathbb{R}^{n-1}} f\,((x,a))\,dx\right)^{\lambda} \left(\int_{\mathbb{R}^{n-1}} g((x,b))dx\right)^{1-\lambda}$$

In other words, introducing

$$F(a) := \int_{\mathbb{R}^{n-1}} f\,((x,a))\,dx, \quad G(b) := \int_{\mathbb{R}^{n-1}} g((x,b))dx$$
$$H((\lambda a + (1-\lambda)b)) := \int_{\mathbb{R}^{n-1}} h\,((x, \lambda a + (1-\lambda)b))\,dx.$$

We have

$$H((\lambda a + (1-\lambda)b)) \geq (F(a))^{\lambda} (G(b))^{1-\lambda},$$

so by *Fubini's theorem* and the one-dimensional inequality, we have

$$\begin{aligned} \int_{\mathbb{R}^n} h(x)dx = \int_{\mathbb{R}} H(a)da &\geq \left(\int_{\mathbb{R}} F(a)da\right)^{\lambda} \left(\int_{\mathbb{R}} G(a)da\right)^{1-\lambda} \\ &= \left(\int_{\mathbb{R}^n} f(x)dx\right)^{\lambda} \left(\int_{\mathbb{R}^n} g(x)dx\right)^{1-\lambda}. \quad \blacksquare \end{aligned}$$

- **Corollary 1.3** *(**Weaker Brunn-Minkowski Inequality**) [Boucheron et al., 2013, Wainwright, 2019]*
*Let $A, B \subset \mathbb{R}^n$ be **non-empty compact sets**. Then for all $\lambda \in [0,1]$,*

$$Vol(\lambda A + (1-\lambda)B) \geq Vol(A)^{\lambda} Vol(B)^{1-\lambda}. \tag{3}$$

Proof: We apply *the Prékopa-Leindler inequality* with $f(x) = \mathbb{1}\{x \in A\}$, $g(x) = \mathbb{1}\{x \in B\}$ and $h(x) = \mathbb{1}\{x \in \lambda A + (1-\lambda)B\}$. We see that

$$h(\lambda x + (1-\lambda)y) = \mathbb{1}\{\lambda x + (1-\lambda)y \in \lambda A + (1-\lambda)B\} \geq \mathbb{1}\{x \in A, y \in B\} = f(x)^{\lambda} g(y)^{1-\lambda}.$$

Thus the hypothesis of *the Prékopa-Leindler inequality* holds. $\blacksquare$

- **Proof:** ($n > 1$ ***case for Brunn-Minkowski Inequality***). First observe that it suffices to prove that for all *nonempty compact sets* A and B,

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

since $\mathrm{Vol}(cA)^{1/n} = c\mathrm{Vol}(A)^{1/n}$ for any $c \in \mathbb{R}$ and $A \subset \mathbb{R}^n$. Also notice that we may assume that $\mathrm{Vol}(A), \mathrm{Vol}(B) > 0$ because otherwise the inequality holds trivially. Defining $A' = \mathrm{Vol}(A)^{-\frac{1}{n}} A$ and $B' = \mathrm{Vol}(B)^{-\frac{1}{n}} B$, we have $\mathrm{Vol}(A') = \mathrm{Vol}(B') = 1$. By *weaker Brunn-Minkowski inequality*, for $\lambda \in (0,1)$,

$$\mathrm{Vol}\left(\lambda A' + (1-\lambda)B'\right) \geq 1.$$

Finally, we apply this *inequality* with the choice

$$\lambda = \frac{\mathrm{Vol}(A)^{\frac{1}{n}}}{\mathrm{Vol}(A)^{\frac{1}{n}} + \mathrm{Vol}(B)^{\frac{1}{n}}}$$

obtaining

$$\begin{aligned} \mathrm{Vol}\left(\frac{\mathrm{Vol}(A)^{\frac{1}{n}} A'}{\mathrm{Vol}(A)^{\frac{1}{n}} + \mathrm{Vol}(B)^{\frac{1}{n}}} + \frac{\mathrm{Vol}(B)^{\frac{1}{n}} B'}{\mathrm{Vol}(A)^{\frac{1}{n}} + \mathrm{Vol}(B)^{\frac{1}{n}}}\right) &\geq 1 \\ \Rightarrow \mathrm{Vol}\left(\frac{A}{\mathrm{Vol}(A)^{\frac{1}{n}} + \mathrm{Vol}(B)^{\frac{1}{n}}} + \frac{B}{\mathrm{Vol}(A)^{\frac{1}{n}} + \mathrm{Vol}(B)^{\frac{1}{n}}}\right) &\geq 1 \\ \Rightarrow \mathrm{Vol}\left(\frac{A+B}{\mathrm{Vol}(A)^{\frac{1}{n}} + \mathrm{Vol}(B)^{\frac{1}{n}}}\right) &\geq 1 \\ \Rightarrow \frac{\mathrm{Vol}(A+B)}{\left(\mathrm{Vol}(A)^{\frac{1}{n}} + \mathrm{Vol}(B)^{\frac{1}{n}}\right)^n} &\geq 1 \end{aligned}$$

which proves the theorem. $\blacksquare$

1.2 The Classical Isoperimetry Theorem

- **Definition** (***Blowup of Sets***)
  For any $t > 0$, and any (measurable) sets $A \subset \mathbb{R}^n$, ***the t-blowup*** *of* A is defined by
  $$A_t := \{x \in \mathbb{R}^n : d(x, A) < t\} = A + t\, B$$
  where $B = \{x \in \mathbb{R}^n : d(0, x) < 1\}$ is an *open unit ball* and $d(x, A) = \inf_{y \in A} d(x, y)$.

- **Definition** (***Surface Area of Sets***)
  let $A \subset \mathbb{R}^n$ be a measurable set and denote by $\mathrm{Vol}(A)$ its *Lebesgue measure. The* ***surface area*** *of* A is defined by
  $$\mathrm{Vol}(\partial A) = \lim_{t \to 0} \frac{\mathrm{Vol}(A_t) - \mathrm{Vol}(A)}{t}.$$
  provided that the limit exists. Here A_t denotes *the t-blowup* of A.

- **Theorem 1.4** ***(Isoperimetry Theorem)*** *[Boucheron et al., 2013, Vershynin, 2018, Wainwright, 2019]*
  Let $A \subset \mathbb{R}^n$ be such that Vol(A) = Vol(B) where $B := \{x \in \mathbb{R}^n : d(0, x) < 1\}$ is unit ball. Then for any $t > 0$,
  $$Vol(A_t) \geq Vol(B_t) \tag{4}$$
  Moreover, if Vol(∂A) exists, then
  $$Vol(\partial A) \geq Vol(\partial B). \tag{5}$$
  Proof: By *the Brunn-Minkowski inequality,*
  $$\begin{aligned} \mathrm{Vol}(A_t)^{1/n} = \mathrm{Vol}(A + tB)^{1/n} &\geq \mathrm{Vol}(A)^{1/n} + t\mathrm{Vol}(B)^{1/n} \\ &= (1 + t)\mathrm{Vol}(B)^{1/n} \\ &= \mathrm{Vol}(B_t)^{1/n}, \end{aligned}$$
  establishing the first statement. The second follows simply because
  $$\mathrm{Vol}(A_t) - \mathrm{Vol}(A) \geq \mathrm{Vol}(B)((1 + t)^n - 1) \geq nt\mathrm{Vol}(B)$$
  where $(1 + t)^n \geq 1 + nt$ for $t \geq 0$. Thus $\mathrm{Vol}(\partial A) \geq n\mathrm{Vol}(B)$. The isoperimetric theorem now follows from the fact that $\mathrm{Vol}(\partial B) = n\mathrm{Vol}(B)$. ∎

- **Remark** (***Isoperimetry Theorem***)
  The classical isoperimetric theorem in $\mathbb{R}^n$ states that, among all sets with ***a given volume***, ***the Euclidean unit ball minimizes the surface area***.

2 Concentration via Isoperimetry

2.1 Levy's Inequalities and Concentration Function

2.2 Isoperimetric Inequalities on the Unit Sphere

- **Remark** (***Volume Ratio of Unit Balls and its Interior***) [Vershynin, 2018]
  Let $B(0, 1) := \{x \in \mathbb{R}^n : \|x\| \leq 1\}$ be the unit ball in $\mathbb{R}^n$. The volume ratio between $B(0, 1)$

and its ϵ-interior $B(0, 1-\epsilon)$ is

$$\frac{\mathrm{Vol}(B(0, 1-\epsilon))}{\mathrm{Vol}(B(0, 1))} = (1-\epsilon)^n \leq \exp\left(-n\epsilon\right)$$

The inequality is due to $1 - x \leq e^{-x}$.

As $n \to \infty$, the above ratio goes to 0. In other words, most of volume in $B(0, 1)$ is ***concentrated*** in the ***boundary*** $\partial B = \mathbb{S}^{n-1} := \{x \in \mathbb{R}^n : \|x\| = 1\}$. This phenomenon is called "***the curse of dimensionality***".

- **Definition**

2.3 Gaussian Isoperimetric Inequalities and Concentration of Gaussian Measure

2.4 Edge Isoperimetric Inequality on the Binary Hypercube

2.5 Vertex Isoperimetric Inequality on the Binary Hypercube

2.6 Convex Distance Inequality

References

Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press, 2013.

Roman Vershynin. *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge university press, 2018.

Martin J Wainwright. *High-dimensional statistics: A non-asymptotic viewpoint*, volume 48. Cambridge University Press, 2019.