CS 798 - Convexity and Optimization , Winter 2017 , Waterloo

Lecture 19 : Brunn-Minkowski inequality

We will see a proof of Brunn-Minkowski inequality by proving the more general Prékopa-Leindler inequality.

We will then see an application of Brunn-Minkowski inequality in proving the Grünbaum's theorem.

Multiplicative Brunn-Minkowski inequality

First, recall the Brunn-Minkowski inequality.

Theorem (Brunn-Minkowski inequality) If A and B are nonempty compact subsets of $\mathbb{R}^n$ then

$$\mathrm{vol}((1-\lambda)A+\lambda B)^{\frac{1}{n}} \geq (1-\lambda)\,\mathrm{vol}(A)^{\frac{1}{n}} + \lambda\,\mathrm{vol}(B)^{\frac{1}{n}}.$$

By AM-GM inequality that $(1-\lambda)a+\lambda b \geq a^{1-\lambda} b^{\lambda}$ for any $a,b \geq 0$ and $0 \leq \lambda \leq 1$ (see L02),

we have $(1-\lambda)\,\mathrm{vol}(A)^{\frac{1}{n}} + \lambda\,\mathrm{vol}(B)^{\frac{1}{n}} \geq \mathrm{vol}(A)^{(1-\lambda)/n}\,\mathrm{vol}(B)^{\lambda/n}$ and the following corollary.

Corollary (multiplicative Brunn-Minkowski inequality) If A and B are compact subsets of $\mathbb{R}^n$ then

$$\mathrm{vol}\Big((1-\lambda)A+\lambda B\Big) \geq \mathrm{vol}(A)^{1-\lambda}\,\mathrm{vol}(B)^{\lambda}.$$

For given A, B, λ, the multiplicative version is weaker than the original version, but we will soon see that the multiplicative version (for all A, B, λ) actually implies the original version.

Remarks:

① Recall the setting of Brunn's theorem: let K be a convex body in $\mathbb{R}^n$, let $u \in \mathbb{R}^n$, and for each $r \in \mathbb{R}$ let H_r be the hyperplane $\{x \in \mathbb{R}^n \mid \langle x,u\rangle = r\}$.

Brunn's theorem says that the function $r \Rightarrow \mathrm{vol}(K \cap H_r)^{\frac{1}{n-1}}$ is concave on its support.

The multiplicative Brunn-Minkowski inequality says that $r \Rightarrow \log \mathrm{vol}(K \cap H_r)$ is concave, or in other words, the function $r \Rightarrow \mathrm{vol}(K \cap H_r)$ is log-concave, a concept that is important later.

② One advantage of the multiplicative version is that we don't require that A and B are non-empty.

In the original version, the non-empty assumption is needed as $A + \phi = \phi$.

In Brunn's setting, the function $r \Rightarrow \mathrm{vol}(K \cap H_r)^{\frac{1}{n-1}}$ is only concave on its support, and the assumption that A and B are non-empty ensure that they correspond to slices in the support.

Removing the non-emptyness assumption makes the multiplicative version easier to use.

③ There is no dimension in the inequality. which also makes it cleaner.

④ As we shall see soon, the Prékopa-Leindler inequality is a functional generalization of the multiplicative version, which is more useful and also simpler to prove.

Before we see Prékopa-Leindler inequality. we first see that the multiplicative version is equivalent to the original version.

The proof uses a trick similar to the one in L02 in proving Hölder's inequality from AM-GM inequality.

Proposition The Brunn-Minkowski inequality is equivalent to the multiplicative Brunn-Minkowski inequality.

Proof We have already seen one direction, so we focus on the other direction.

Assume the multiplicative Brunn-Minkowski inequality holds for all A, B, and $0 \le \lambda \le 1$, we would like to prove that the original Brunn-Minkowski inequality holds for all A, B and $0 \le \lambda \le 1$.

Suppose $\mathrm{vol}(A)=0$ or $\mathrm{vol}(B)=0$ (e.g. A consists of one point), then the inequality holds trivially.

So, assume $\mathrm{vol}(A)>0$ and $\mathrm{vol}(B)>0$ and they are subsets in $\mathbb{R}^n$

We define $A' = A/\mathrm{vol}(A)^{\frac{1}{n}}$, $B' = B/\mathrm{vol}(B)^{\frac{1}{n}}$, and $\lambda' = \dfrac{\lambda\,\mathrm{vol}(B)^{\frac{1}{n}}}{(1-\lambda)\mathrm{vol}(A)^{\frac{1}{n}} + \lambda\,\mathrm{vol}(B)^{\frac{1}{n}}}$.

Applying the multiplicative Brunn-Minkowski inequality on A', B' and λ' gives

$$\mathrm{vol}\left(\frac{(1-\lambda)A + \lambda B}{(1-\lambda)\mathrm{vol}(A)^{\frac{1}{n}} + \lambda\mathrm{vol}(B)^{\frac{1}{n}}}\right) = \mathrm{vol}\left((1-\lambda')A' + \lambda' B'\right) \ge \mathrm{vol}(A')^{1-\lambda'}\mathrm{vol}(B')^{\lambda'} = 1.$$

This implies that $\mathrm{vol}\left((1-\lambda)A + \lambda B\right) \ge \left((1-\lambda)\mathrm{vol}(A)^{\frac{1}{n}} + \lambda\mathrm{vol}(B)^{\frac{1}{n}}\right)^n$, the original Brunn-Minkowski. □

Henceforth. we focus on proving the multiplicative Brunn-Minkowski inequality.

Prékopa-Leindler inequality

Prékopa-Leindler inequality generalize the Brunn-Minkowski inequality to the function setting.

To motivate it, let f, g, m be the characteristic functions of $A, B, \lambda A + (1-\lambda)B$ respectively. i.e.

$f(x)=1$ if $x \in A$ and $f(x)=0$ if $x \notin A$, and similarly for g and m.

Then, the volumes of A, B, and $(1-\lambda)A+\lambda B$ are simply $\int_{\mathbb{R}^n} f$, $\int_{\mathbb{R}^n} g$, and $\int_{\mathbb{R}^n} m$ respectively.

The conclusion of the Brunn-Minkowski inequality is that $\int m \ge \left(\int f\right)^{1-\lambda}\left(\int g\right)^{\lambda}$.

Obviously it is not true for any three functions f, g, m.

In the Brunn-Minkowski setting, if $x \in A$ and $y \in B$, then $(1-\lambda)x + \lambda y \in (1-\lambda)A + \lambda B$.

Translating in the characteristic function setting, if $f(x)=1$ and $g(y)=1$, then $m((1-\lambda)x+\lambda y)=1$

One way to write it in the functional setting is $m((1-\lambda)x+\lambda y) \geq f(x)^{1-\lambda} g(y)^{\lambda}$, so that the inequality still holds if we scale f, g, m by the same factor.

From this discussion, it is clear that Brunn-Minkowski inequality is a special case of the following theorem.

<u>Theorem.</u> (Prékopa-Leindler inequality) Let f, g, m be nonnegative measurable functions on $\mathbb{R}^n$.

For any $\lambda \in (0,1)$, if for all $x, y \in \mathbb{R}^n$, it holds that $m((1-\lambda)x+\lambda y) \geq f(x)^{1-\lambda} g(y)^{\lambda}$,

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

<u>Proof</u> When you see a general statement is easier to prove than a special case, one typical reason is that the more general structure is easier to prove by induction. Which is the case here.

First, we start with the one-dimensional Brunn-Minkowski inequality (not Prékopa-Leindler yet), as it will be used in the one-dimensional Prékopa-Leindler inequality.

Given two sets $A, B \subseteq \mathbb{R}$, note that $\mathrm{vol}(A+B) = \mathrm{vol}(A+(B+c))$ where $B+c$ is a "shift" of B by $c \in \mathbb{R}$, because there is a bijection between the points in $A+B$ and $A+(B+c)$.

So, without loss of generality, we can assume (by shifting) that the maximum point of A is zero, and also the minimum point of B is zero, so that A and B are (almost) disjoint sets.

Then $A+B$ contains A (as $A+\{0\}=A$) and B, and this implies that $\mathrm{vol}(A+B) \geq \mathrm{vol}(A)+\mathrm{vol}(B)$.

Next, we prove the one-dimensional case of Prépoka-Leindler inequality.

We may assume that f and g are bounded, as otherwise the statement holds.

By Fubini's theorem, we can write $\int f(x)dx = \int_0^\infty \mathrm{vol}((f \geq t))dt$ as f is non-negative, and similarly for g, where we use $(f \geq t)$ to denote the set $\{x \in \mathbb{R} \mid f(x) \geq t\}$.

By our assumption, if $f(x) \geq t$ and $g(x) \geq t$, then $m((1-\lambda)x+\lambda y) \geq t$.

This implies that $(m \geq t) \supseteq (1-\lambda)(f \geq t) + \lambda(g \geq t)$.

Therefore,
$$\begin{aligned}
\int_{\mathbb{R}} m &= \int_0^\infty \mathrm{vol}((m \geq t))dt && \text{(by Fubini's theorem)} \\
&\geq \int_0^\infty \mathrm{vol}((1-\lambda)(f \geq t) + \lambda(g \geq t))\,dt && \text{(follows from our assumption)} \\
&\geq \int_0^\infty \mathrm{vol}((1-\lambda)(f \geq t))\,dt + \int_0^\infty \mathrm{vol}(\lambda(g \geq t))\,dt && \text{(by one-dimensional Brunn-Minkowski)} \\
&= (1-\lambda)\int_0^\infty \mathrm{vol}((f \geq t))\,dt + \lambda \int_0^\infty \mathrm{vol}((g \geq t))\,dt \\
&= (1-\lambda)\int_{\mathbb{R}} f + \lambda \int_{\mathbb{R}} g && \text{(by Fubini's theorem)}
\end{aligned}$$

$= (1-\lambda)\int_0^\infty \mathrm{vol}((f\geq t))\,dt + \lambda\int_0^\infty \mathrm{vol}((g\geq t))\,dt$

$= (1-\lambda)\int_{\mathbb{R}} f + \lambda \int_{\mathbb{R}} g$ (by Fubini's theorem)

$\geq \left(\int_{\mathbb{R}} f\right)^{1-\lambda} \left(\int_{\mathbb{R}} g\right)^{\lambda}$ (by AM-GM inequality)

We are ready to prove Prépoka-Leindler inequality by induction.

The idea is to reduce to the $(n-1)$-dimensional case by recording the volume of each slice as in the original statement of Brunn's theorem.

For f, g, m and $z \in \mathbb{R}$, we write $f_z : \mathbb{R}^{n-1} \to \mathbb{R}$ as $f_z(x) = f(z, x)$ for $x \in \mathbb{R}^{n-1}$ and similarly for g and m.

Also, we write $F_z = \int_{\mathbb{R}^{n-1}} f_z(x)\,dx$ as the "total sum" of f when the first coordinate is fixed to z.

Then $\int_{\mathbb{R}^n} m \geq \left(\int_{\mathbb{R}^n} f\right)^{1-\lambda} \left(\int_{\mathbb{R}^n} g\right)^{\lambda}$ can be rewritten as $\int_{-\infty}^{\infty} M_z\,dz \geq \left(\int_{-\infty}^{\infty} F_z\,dz\right)^{1-\lambda} \left(\int_{-\infty}^{\infty} G_z\,dz\right)^{\lambda}$.

This will follow from the one-dimensional Prékopa-Leindler inequality if $M_{(1-\lambda)a+\lambda b} \geq (F_a)^{1-\lambda}(G_b)^{\lambda}$ $\forall a, b \in \mathbb{R}$, which we will argue below that it follows from the induction hypothesis and our assumption.

For all $x, y \in \mathbb{R}^{n-1}$, our assumption implies that $m((1-\lambda)a+\lambda b, (1-\lambda)x+\lambda y) \geq (f(a,x))^{1-\lambda}(g(b,y))^{\lambda}$, and this means that $m_{(1-\lambda)a+\lambda b}((1-\lambda)x+\lambda y) \geq (f_a(x))^{1-\lambda}(g_b(y))^{\lambda}$.

Since f, g, m are $(n-1)$-dimensional, by the induction hypothesis, we have $\int_{\mathbb{R}^{n-1}} m_{(1-\lambda)a+\lambda b} \geq \left(\int_{\mathbb{R}^{n-1}} f_a\right)^{1-\lambda}\left(\int_{\mathbb{R}^{n-1}} g_b\right)^{\lambda}$ and thus $M_{(1-\lambda)a+\lambda b} \geq (F_a)^{1-\lambda}(G_b)^{\lambda}$ for any $a, b \in \mathbb{R}$, and this completes the proof. □

Even though we only wanted to prove Brunn-Minkowski inequality, to use induction, we need to keep track of the volumes of the slices and this requires us to use the one-dimensional Prépoka-Leindler inequality (see [Ball, lecture 5]).

There is an alternative proof of Brunn-Minkowski inequality using cuboids, which is easier to understand (see [Grüber, chapter 8] and [Vempala, chapter 2]).

The original proof of Brunn uses a symmetrization argument (see [Grüber, chapter 8]).

We will use the Prépoka-Leindler inequality next time and so we prefer this approach to Brunn-Minkowski.

Grünbaum's theorem

Recall Grünbaum's theorem that we used in the center of gravity method in L15.

Theorem (Grünbaum) Let K be a centered convex set, i.e. $\int_{x\in K} x\,dx = 0$.

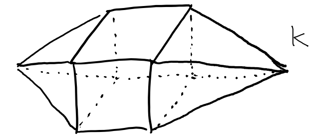
For any $v \in \mathbb{R}^n$ with $v \neq 0$, we have $\mathrm{vol}(K \cap \{x \in \mathbb{R}^n \mid \langle v, x\rangle \geq 0\}) \geq \frac{1}{e}\mathrm{vol}(K)$.

We can now prove it using Brunn-Minkowski inequality.

Without loss let's assume that $v = e_1$.

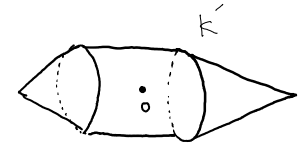
Let K be a centered convex set.

We will do some transformations of K until the theorem becomes clear.



(a cube with two cones)

Symmetrization



Let $K \subseteq \mathbb{R}^n$ be the original centered convex set.

We replace each slice of K by an $(n-1)$-dimensional ball with the same volume.

More precisely, for each $z \in \mathbb{R}$, let K_z be $\{x \in \mathbb{R}^n \mid x \in K \text{ and } x_1 = z\}$ be the slice of K at $x_1 = z$

Then K' is constructed in such a way that $\mathrm{vol}(K_z) = \mathrm{vol}(K'_z)$ for all $z \in \mathbb{R}$ and K'_z is an $(n-1)$-dimensional ball.

Obviously, K' is still centered. Less obviously, we argue that K' is convex by the Brunn-Minkowski inequality.

Claim K' is convex.

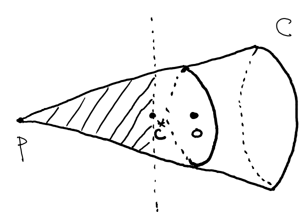
proof Since each slice is an $(n-1)$-dimensional ball, to show that K' is convex, it is enough to establish that $r(z) = \mathrm{radius}(K_z)$ is concave (recall the Brunn's theorem in the 2-dimensional case).

Note that $r(z)^{n-1} \cdot v_{n-1} = \mathrm{vol}(K_z)$ where v_{n-1} is the volume of the $(n-1)$-dimensional unit ball.

So, $r(z) = \left(\mathrm{vol}(K_z)/v_{n-1}\right)^{\frac{1}{n-1}}$.

By Brunn's theorem, $\mathrm{vol}(K_z)^{\frac{1}{n-1}}$ is a concave function, and so is $r(z)$, proving the claim. □

Cone



We now transform K' into a cone C.

The center slice of K' and C will be the same $(n-1)$-dimensional ball.

We choose a point p so that the convex hull of the central slice plus p is a cone with the same volume as $\{x \in \mathbb{R}^n \mid x \in K' \text{ and } x_1 \le 0\}$, i.e. $\mathrm{vol}(K'_{\le 0}) = \mathrm{vol}(C_{\le 0})$.

We extend this cone to the right so that we also have $\mathrm{vol}(K'_{\ge 0}) = \mathrm{vol}(C_{\ge 0})$, where we defined $K'_{\le a} = \{x \in \mathbb{R}^n \mid x \in K' \text{ and } x_1 \le a\}$ and similarly define $K'_{\ge a}$

By doing so, we claim that the mass of K' has only move to the left.

Claim For any K' with $\mathrm{vol}(K'_{\le 0}) = \mathrm{vol}(C_{\le 0})$ and $\mathrm{vol}(K'_{\ge 0}) = \mathrm{vol}(C_{\ge 0})$, we have $\mathrm{vol}(K'_{\le p}) = 0$ and $\mathrm{vol}(K'_{\le z}) \le \mathrm{vol}(C_{\le z})$ for all z in the support of C.

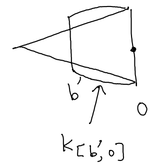
Suppose we scan from $-\infty$ to the largest z such that $\mathrm{vol}(C_z) > 0$.

Let b be a point such that $\mathrm{vol}(K'_{\le b}) > \mathrm{vol}(C_{\le b})$.

We will argue that such a point b does not exist and this clearly implies the claim.

Suppose by contradiction that $b \le 0$. Then there exists $b' \le b \le 0$ such that $\mathrm{vol}(K_{b'}) \geq \mathrm{vol}(C_{b'})$.

By Brunn's theorem, $r(z) = \mathrm{radius}(K'_z)$ is concave and so $\mathrm{vol}(K'_a) \geq \mathrm{vol}(C_a)$ for $b' \le a \le 0$.



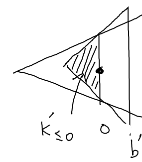
but this implies that $\mathrm{vol}(K'_{\le 0}) = \mathrm{vol}(K'_{\le b}) + \mathrm{vol}(K'_{[b,0]})$

$$> \mathrm{vol}(C'_{\le b}) + \mathrm{vol}(C_{[b,0]})$$

$$= \mathrm{vol}(C_{\le 0}),$$ contradicting our assumption that $\mathrm{vol}(K'_{\le 0}) = \mathrm{vol}(C_{\le 0})$.

Suppose by contradition that $b > 0$. Then there exists $0 \le b' \le b$ such that $\mathrm{vol}(K_{b'}) > \mathrm{vol}(C_{b'})$

Again, by Brunn's theorem, since K' is convex, $r(z) = \mathrm{radius}(K'_z)$ is concave, and this forces that $\mathrm{vol}(K'_{\le 0}) < \mathrm{vol}(C_{\le 0})$ (see picture). contradicting our assumption. □



Proof The proof of Grünbaum's theorem follows rather easily from here.

We have $\dfrac{\mathrm{vol}(K_{\le 0})}{\mathrm{vol}(K)} = \dfrac{\mathrm{vol}(K'_{\le 0})}{\mathrm{vol}(K')}$ since the first transformation preserves volumes on both sides

$= \dfrac{\mathrm{vol}(C_{\le 0})}{\mathrm{vol}(C)}$ since the second transformation preserves volumes on both sides

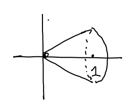
$\geq \dfrac{\mathrm{vol}(C_{\le c^*})}{\mathrm{vol}(C)}$ by the claim above where c^* is the center of gravity of C (as the mass has only moved left, so center of gravity only moved left)

$= \left(\dfrac{n}{n+1}\right)^n$ as we will argue below

$\geq \dfrac{1}{e}$ as $(1+\frac{1}{n})^n \le e$.

It remains to argue that the smaller half of the cone cut by the center of gravity is at least $\frac{1}{e}$ fraction.

Without loss, we assume that the cone is of height one with the tip at the origin and the base is B.



Then, $\bar{x}_1 = \dfrac{1}{\mathrm{vol(cone)}} \displaystyle\int_0^1 x \cdot (x)^{n-1} \mathrm{vol}(B)\, dx = \dfrac{n}{n+1}$ as $\mathrm{vol(cone)} = \mathrm{vol}(B)/n$.

Thus, the "left" side of the cone is of volume $\left(\frac{n}{n+1}\right)^n$ fraction of the whole cone. □

It is clear from the proof that Brünbaum's theorem is tight and the worst case is a cone, which is also the worst case of Brunn's theorem.

References : [Ball, lecture 5].

- Vempala, Algorithmic convex geometry, chapter 2.

- Gruber, Convex and Discrete Geometry, chapter 8 (Brunn-Minkowski), chapter 9 (symmetrization).