Algorithmic Convex Geometry

August 2011

Contents

Chapter 1

Overview

Algorithmic problems in geometry often become tractable with the assumption of convexity. Optimization, volume computation, geometric learning and finding the centroid are all examples of problems that are significantly easier for convex sets.

We will study this phenomenon in depth, pursuing three tracks that are closely connected to each other. The first is the theory of geometric inequalities. We begin with classical topics such as the Brunn-Minkowski inequality, and later deal with more recent developments such as isoperimetric theorems for convex bodies and their extensions to logconcave functions. The second track is motivated by the problem of sampling a geometric distribution by a random walk. Here we will develop some general tools and use them to analyze geometric random walks. The inequalities of the first track play a key role in bounding the rate of convergence of these walks. The last track is the connection between sampling and various algorithmic problems, most notably, that of computing the volume of a convex body (or more generally, integrating a logconcave function). Somewhat surprisingly, random sampling will be a common and essential feature of polynomial-time algorithms for these problems. In some cases, including the volume problem, sampling by a random walk is the *only* known way to get a polynomial-time algorithm.

1.1 Learning by random sampling

We will now see our first example of reducing an algorithmic problem to a random sampling problem. In a typical *learning* problem, we are presented with samples $X^1, X^2, \ldots$ from the domain of the function, and have to guess the values $f(X^i)$. After after each guess we are told whether it was right or wrong. The objective is to minimize the number of wrong guesses. One assumes there is an unknown function f belonging to some known restricted class of functions.

As a concrete example, suppose there is a fixed unknown vector $\vec{a} \in \mathbb{R}^n$, and our function f is defined by

$$f(\vec{x}) = \begin{cases} \text{True} & \text{if } \vec{a} \cdot \vec{x} \geq 0 \\ \text{False} & \text{if } \vec{a} \cdot \vec{x} < 0 \end{cases}$$

Assume the right answer has components $a_i \in \{-2^b, ..., 2^b\} \subset \mathbb{Z}$. Consider the following algorithm. At each iteration, choose a random $\vec{a}$ from those that have made no mistakes so far, and use that to make the next guess.

If, on every step, we pick the answer according the majority vote of those $\vec{a}$ which have made no mistake so far, then every mistake would cut down the field of survivors by at least a factor of 2. As there are $2^{b(n+1)}$ candidates at the outset, you would make at most $(b+1)n$ mistakes.

Exercise 1. *Show that for the randomized algorithm above,*

$$\mathsf{E}(\textit{number of mistakes}) \leq 2(b+1)n.$$

Chapter 2

The Brunn-Minkowski Inequality

In this lecture, we will prove a fundamental geometric inequality – the Brunn-Minkowski inequality. This inequality relates the volumes of sets in high-dimensional spaces under convex combinations. Let us first recall the definition of convexity.

Definition 1. *Let $K \subseteq \mathbb{R}^n$. K is a convex set if for any two points $x, y \in K$, and any $0 \le \lambda \le 1$, $\lambda x + (1-\lambda)y \in K$.*

To motivate the inequality, consider the following version of cutting a (convex) cake: you pick a point x on the cake, your brother makes a single knife cut and you get the piece that contains x. A natural choice for x is the centroid. For a convex set K, it is

$$x = \frac{1}{\mathsf{Vol}(K)} \int_{y \in K} y \, dy.$$

What is the minimum fraction of the cake that you are guaranteed to get?

For convenience, let K be a convex body whose centroid is the origin. Let $u \in \mathbb{R}^n$ be the normal vector defining the following halfspaces:

$$\begin{aligned} H_1 &= \{v \in \mathbb{R}^n : u \cdot v \ge 0\} \\ H_2 &= \{v \in \mathbb{R}^n : u \cdot v < 0\} \end{aligned}$$

Now we can consider the two portions that u cuts out of K:

$$\begin{aligned} K_1 &= H_1 \cap K \\ K_2 &= H_2 \cap K \end{aligned}$$

We would like to compare $\mathsf{Vol}(K_1)$ and $\mathsf{Vol}(K_2)$ with $\mathsf{Vol}(K)$.

Consider first one dimension – a convex body in one dimension is just a segment on the real number line. It's clear that any cut through the centroid of the segment (i.e. the center) divides the area of the segment into two sides of exactly half the area of the segment.

For two dimensions, this is already a non-trivial problem. Let's consider an isosceles triangle, whose side of unique length is perpendicular to the x axis. If we make a cut through the centroid perpendicular to the x axis, it is readily checked that the volume of the smaller side is $\frac{4}{9}$'ths of the total volume. Is this the least possible in $\mathbb{R}^2$? What about in $\mathbb{R}^n$?

The Brunn-Minkowski inequality will be very useful in answering these questions.

2.1 The inequality

We first define the Minkowski sum of two sets:

Definition 2. *Let $A, B \subseteq \mathbb{R}^n$. The Minkowski sum of A and B is the set*

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

How is the volume of $A + B$ related to the volume of A or B? The Brunn-Minkowski inequality relates these quantities.

Theorem 3 (Brunn-Minkowski)**.** *Let $0 \leq \lambda \leq 1$, and suppose that A, B, and $\lambda A + (1-\lambda)B$ are measurable subsets of $\mathbb{R}^n$. Then,*

$$\mathsf{Vol}(\lambda A + (1-\lambda)B)^{1/n} \geq \lambda \mathsf{Vol}(A)^{1/n} + (1-\lambda)\mathsf{Vol}(B)^{1/n}.$$

Recall that for a measurable set A and a scaling factor λ, we have that:

$$\mathsf{Vol}(\lambda(A)) = \lambda^n \mathsf{Vol}(A).$$

It follows that an equivalent statement of the inequality is the following: for measurable sets A, B and $A + B$ over $\mathbb{R}^n$:

$$\mathsf{Vol}(A + B)^{1/n} \geq \mathsf{Vol}(A)^{1/n} + \mathsf{Vol}(B)^{1/n}.$$

2.1.1 Proving the Brunn-Minkowski inequality

For some intuition, let's first consider the Brunn-Minkowski inequality when A and B are axis-aligned cuboids in $\mathbb{R}^n$. A cuboid in $\mathbb{R}^n$ is a generalization of the familiar rectangle in two dimensions. An axis-aligned cuboid with side lengths $(a_1, a_2, \dots a_n)$ is the set

$$A = \{x \in \mathbb{R}^n : l_i \leq x_i \leq l_i + a_i\}$$

for $l = (l_1, \dots, l_n) \in \mathbb{R}^n$.

Cuboids

Let A be a cuboid with side lengths $(a_1, a_2, \dots, a_n)$ and B be a cuboid with side lengths $(b_1, b_2, \dots, b_n)$. Let us prove the Brunn-Minkowski inequality for A and B.

This proof will follow easily because if A and B are cuboids, then $A+B$ is a cuboid with side lengths $(a_1 + b_1, a_2 + b_2, \dots a_n + b_n)$. Since these are cuboids, it is easy to compute their volumes:

$$\mathsf{Vol}(A) = \prod_{i=1}^{n} a_i, \mathsf{Vol}(B) = \prod_{i=1}^{n} b_i, \mathsf{Vol}(A+B) = \prod_{i=1}^{n} a_i + b_i$$

We want to show that

$$\mathsf{Vol}(A+B)^{1/n} \geq \mathsf{Vol}(A)^{1/n} + \mathsf{Vol}(B)^{1/n}.$$

To this end, consider the ratio between the volumes

$$\begin{aligned} \frac{\mathsf{Vol}(A)^{1/n} + \mathsf{Vol}(B)^{1/n}}{\mathsf{Vol}(A+B)^{1/n}} &= \frac{\left(\prod_{i=1}^{n} a_i\right)^{1/n} + \left(\prod_{i=1}^{n} b_i\right)^{1/n}}{\left(\prod_{i=1}^{n} a_i + b_i\right)^{1/n}} \\ &= \left(\prod_{i=1}^{n} \frac{a_i}{a_i + b_i}\right)^{1/n} + \left(\prod_{i=1}^{n} \frac{b_i}{a_i + b_i}\right)^{1/n} \\ &\leq \sum_{i=1}^{n} \frac{a_i}{a_i + b_i} + \sum_{i=1}^{n} \frac{b_i}{a_i + b_i} \\ &= 1 \end{aligned}$$

Which proves our claim; the inequality used is the standard inequality between the geometric and the arithmetic mean.

Now that we have the result for cuboids, how can we generalize this to arbitrary measurable sets? The key is that any measurable set can be approximated arbitrarily well with unions of cuboids. We will prove that if $A \cup B$ is the union of a finite number of cuboids, the Brunn-Minkowski inequality holds. We can prove the general result by approximating arbitrary measurable sets with unions of cuboids.

Finite unions of cuboids

We prove the Brunn-Minkowski inequality by induction on the number of disjoint cuboids that $A \cup B$ contain. Note that our result for A and B being cuboids is the base case.

We first translate A so that there exists a cuboid in A fully contained in the halfspace $\{x : x_1 \geq 0\}$ and there exists a cuboid in A fully contained in $\{x : x_1 < 0\}$. Now, translate B so that the following inequality holds:

$$\frac{\mathsf{Vol}(A_+)}{\mathsf{Vol}(A)} = \frac{\mathsf{Vol}(B_+)}{\mathsf{Vol}(B)}$$

where:

$$A_+ = A \cap \{x : x_1 \geq 0\}, A_- = A \setminus A_+$$
$$B_+ = B \cap \{x : x_1 \geq 0\}, B_- = B \setminus B_+$$

Note that we can now proceed by induction on the sets $A_+ + B_+$, $A_- + B_-$, since the number of cuboids in $A_+ \cup B_+$ and $A_- \cup B_-$ is fewer than the number of cuboids in $A \cup B$. This allows us to complete the proof:

$$\begin{aligned}
\mathsf{Vol}(A+B) &\geq \mathsf{Vol}(A_+ + B_+) + \mathsf{Vol}(A_- + B_-) \\
&\geq \left(\mathsf{Vol}(A_+)^{1/n} + \mathsf{Vol}(B_+)^{1/n}\right)^n + \left(\mathsf{Vol}(A_-)^{1/n} + \mathsf{Vol}(B_-)^{1/n}\right)^n \\
&= \mathsf{Vol}(A_+)\left(1 + \left(\frac{\mathsf{Vol}(B)}{\mathsf{Vol}(A)}\right)^{1/n}\right)^n + \mathsf{Vol}(A_-)\left(1 + \frac{\mathsf{Vol}(B)}{\mathsf{Vol}(A)}^{1/n}\right)^n \\
&= \mathsf{Vol}(A)\left(1 + \left(\frac{\mathsf{Vol}(B)}{\mathsf{Vol}(A)}\right)^{1/n}\right)^n
\end{aligned}$$

The first inequality above follows from the observation that since A_+, B_+ are separated from A_-, B_- by a halfspace, $A_+ + B_+$ is dijoint from $A_- + B_-$.

It follows that

$$\mathsf{Vol}(A+B)^{1/n} \geq \mathsf{Vol}(A)^{1/n} + \mathsf{Vol}(B)^{1/n}$$

which was our desired result.

2.2 Grunbaum's inequality

We are now ready to go back to our original question: what is the smallest volume ratio attainable for the smaller half of a cut through the centroid of any convex body? Somewhat surprisingly, the generalization of the two dimensional case we covered in the introduction is the worst case in high dimensions.

Theorem 4 (Grunbaum's inequality)**.** *Let K be a convex body. Then any half-space defined by a normal vector v that contains the centroid of K contains at least a $\frac{1}{e}$ fraction of the volume of K.*

The overview of the proof is that we will consider the body K and a halfspace defined by v going through the centroid. We will simplify K by performing a symmetrization step – replacing each $(n-1)$-dimensional slice S of K perpendicular to v with an $(n-1)$-dimensional ball of the same volume as S to obtain a new convex body K'. We will verify the convexity of K' by using the Brunn-Minkowski inequality. This does not modify the fraction of volume contained in the halfspace. Next, we will construct a cone from K'. The cone will have the property that the fraction of volume in the halfspace can only *decrease*, so it will still be a lower bound on K. Then we will easily calculate the ratio of the volume of a halfspace defined by a normal vector v cutting through the centroid of a cone, which will give us the desired result.

The two steps of symmetrization and conification can be seen in figure 2.1.

Figure 2.1: The first diagram shows the symmetrization step – each slice of K is replaced with an $(n-1)$ dimensional ball. The second diagram shows how we construct a cone from K'.

Proof. Assume without loss of generality that v is the x_1 axis, and that the centroid of K is the origin. From the convex body K, we construct a new set K'. For all $\alpha \in \mathbb{R}$, we place an $(n-1)$ dimensional ball centered around $c = (\alpha, 0, \ldots 0)$ on the last $(n-1)$ coordinates with radius r_α such that

$$\mathsf{Vol}(B(c, r_\alpha)) = \mathsf{Vol}(K \cap \{x : x_1 = \alpha\}).$$

See figure 2.1 for a visual of the construction. Recall that the volume of a $n-1$ dimensional ball of radius r is $f(n-1)r^{n-1}$ (where f is a function not depending on r and only on n), so that we set

$$r_\alpha = \left(\frac{\mathsf{Vol}(K \cap \{x : x_1 = \alpha\})}{f(n-1)}\right)^{1/(n-1)}$$

We first argue that K' is convex. Consider the function $r : \mathbb{R} \to \mathbb{R}$, defined as $r(\alpha) = r_\alpha$. If r is a concave function (i.e. $\forall x, y \in \mathbb{R}, 0 \le \lambda \le 1, r(\lambda x + (1-\lambda)y) \ge \lambda r(x) + (1-\lambda)r(y)$), then K' is convex.

Note that r_α is proportional to $\mathsf{Vol}(K \cap \{x : x_1 = \alpha\})^{1/(n-1)}$. Let $A = K \cap \{x : x_1 = a\}$, and let $B = K \cap \{x : x_1 = b\}$. For some $0 \le \lambda \le 1$, let $C = K \cap \{x : x_1 = \lambda a + (1-\lambda)b\}$. r is concave if

$$\mathsf{Vol}(C)^{1/n-1} \ge \lambda \mathsf{Vol}(A)^{1/n-1} + (1-\lambda)\mathsf{Vol}(B)^{1/n-1} \tag{2.1}$$

It is easily checked that

$$C = \lambda A + (1-\lambda)B$$

so that (2.1) holds by the Brunn-Minkowski inequality, and K' is convex. Let

$$\begin{aligned} K'_+ &= K' \cap \{x : x_1 \ge 0\} \\ K'_- &= K' \cap \{x : x_1 < 0\}. \end{aligned}$$

Then the ratio of volumes on either side of $x_1 = 0$ has not changed, i.e.

$$\begin{aligned} \mathsf{Vol}(K'_+) &= \mathsf{Vol}(K \cap \{x : x_1 \ge 0\}) \\ \mathsf{Vol}(K'_-) &= \mathsf{Vol}(K \cap \{x : x_1 < 0\}). \end{aligned}$$

We argue that

$$\frac{\mathsf{Vol}(K'_+)}{\mathsf{Vol}(K')} \ge \frac{1}{e}$$

and the other side will follow by symmetry. To this end, we transform K' into a cone (see figure 2.1). To be precise, we replace K'_+ with a cone C of equal volume with the same base as K'_+. We replace K'_- with the extension E of the cone C, such that $\mathsf{Vol}(E) = \mathsf{Vol}(K'_-)$. The centroid of K' was at the origin; the centroid of $C \cup E$ can only be pushed to the right along x_1, because r is concave. Let the position of the centroid of $C \cup E$ along x_1 be $\alpha \ge 0$. Then we have that

$$\frac{\mathsf{Vol}(K'_+)}{\mathsf{Vol}(K')} \ge \frac{\mathsf{Vol}((C \cup E) \cap \{x : x_1 \ge \alpha\})}{\mathsf{Vol}(C \cup E)}$$

so a lower bound on the latter ratio implies a lower bound on $\frac{\mathsf{Vol}(K'_+)}{\mathsf{Vol}(K')}$. To this end, we compute the ratio of volume on the right side of a halfspace cutting through the centroid of a cone.

Let the height of the cone $C \cup E$ be h, and the length of the base be R. The centroid of the cone is at position α along x_1 given by:

$$\begin{aligned}\frac{1}{\mathsf{Vol}(C \cup E)} \int_{t=0}^{h} t f(n-1) \left(\frac{tR}{h}\right)^{n-1} dt &= \frac{f(n-1)}{\mathsf{Vol}(C \cup E)} \left(\frac{R}{h}\right)^{n-1} \int_{t=0}^{h} t^n dt \\ &= \frac{f(n-1)}{\mathsf{Vol}(C \cup E)} \cdot \frac{R^{n-1} h^2}{n+1}\end{aligned}$$

where $f(\cdot)$ is the function independent of r such that $f(n)r^n$ gives the volume of an n dimensional ball with radius r. By noting that, for a cone,

$$\mathsf{Vol}(C \cup E) = \frac{f(n-1)R^{n-1}h}{n}$$

we have that $\alpha = \frac{n}{n+1}h$. Now, plugging in to compute the volume of C (the right half of the cone), we have our desired result.

$$\begin{aligned}\mathsf{Vol}(C) &= \int_{t=0}^{\frac{n}{n+1}h} f(n-1) \left(\frac{tR}{h}\right)^{n-1} dt \\ &= \frac{R^{n-1} f(n-1)}{h^{n-1}} \int_{t=0}^{\frac{n}{n+1}h} t^{n-1} dt \\ &= \left(\frac{n}{n+1}\right)^n \mathsf{Vol}(C \cup E) \\ &\geq \frac{1}{e} \mathsf{Vol}(C \cup E)\end{aligned}$$

□

Chapter 3

Convex Optimization

3.1 Optimization in Euclidean space

Let $S \subset \mathbb{R}^n$, and let $f : S \to \mathbb{R}$ be a real-valued function. The problem of interest may be stated as

$$\begin{array}{ll} \min & f(x) \\ & x \in S, \end{array} \tag{3.1}$$

that is, find the point $x \in S$ which minimizes by f. For convenience, we denote by x^* a solution for the problem (3.1) and by $f^* = f(x^*)$ the associated cost.

Restricting the set S and the function f to be convex[1], we obtain a class of problems which are polynomial-time solvable in

$$n \quad \text{and} \quad \log\left(\frac{1}{\epsilon}\right),$$

where ϵ defines an optimality criterion such as

$$\|x - x^*\| \leq \epsilon. \tag{3.2}$$

Linear programming (LP) is one important subclass of this family of problems. Let $A : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$. We state a LP problem as

$$\begin{array}{ll} \min\limits_{x} & c^T x \\ & Ax \geq b. \end{array} \tag{3.3}$$

[1] In fact, one could easily extend to the case of quasi-convex f.

Here, the feasible set is a polyhedron, for many CO problems it is also bounded, i.e., it is a polytope.

Here we illustrate other classical examples.

Example 5. *Minimum weight perfect matching problem; given graph $G = (V, E)$, with costs $\omega : E \to \mathbb{R}$, find*

$$\begin{array}{ll} \min & \sum_{e\in E} c_e x_e \\ & \sum_{e\in\delta(v)} x_e = 1, \quad \text{for all} \ \ v \in V \\ & x_e \in \{0,1\}, \quad \text{for all} \ \ e \in E, \end{array} \tag{3.4}$$

where $\delta(v) = \{e = (i,j) \in E : v = i \ \text{ or } \ v = j\}$. Due to the last set of constraints, the so called integrality constraints, (3.4) is not a LP.

The previous problem had integrality constraints which usually increase the difficulty of the problem. One possible approach is to solve a linear relaxation to get (at least) bounds for the original problem. In our particular case,

$$\begin{array}{ll} \min & \sum_{e\in E} c_e x_e \\ & \sum_{e\in\delta(v)} x_e = 1, \quad \text{for all} \ \ v \in V \\ & 0 \le x_e \le 1, \quad \text{for all} \ \ e \in E. \end{array} \tag{3.5}$$

This problem is clearly an LP, and thus, it can be efficiently solved. Unfortunately, the solution for the LP might not be interesting for the original problem (3.4). For example, it is easy to see that the problem (3.4) is infeasible when $|V| = 3$, while the linear relaxation is not[2].

Surprisingly, it is known that the graph G is bipartite if and only if each vertex of the polytope defined in (3.5) is a feasible solution for (3.4). Thus, for this particular case, the optimal cost coincide for both problems.

For general graphs, we need to add another family of constraints in order to obtain the previous equivalence[3].

$$\sum_{e\in(S,S^c)} x_e \ge 1, \quad \text{for all} \ \ S \subset V, |S| \ \ \text{odd}. \tag{3.6}$$

Note that there are an exponential number of constraints in this family. Thus, it is not obvious that it is polynomial-time solvable.

Example 6. *Another convex optimization problem when a polyhedra is intersected with a ball of radius R, i.e., the feasible set is nonlinear.*

[2] Just let $x = (0.5, 0.5, 0.5)$.

[3] The proof is due to Jack Edmonds.

$$\begin{array}{ll} \min & c^T x \\ & Ax \geq b \\ & \sum_{i=1}^n x_i^2 \leq R^2. \end{array} \tag{3.7}$$

Example 7. *Semi-definite programming (SDP). Here the variable, instead of being a vector in $\mathbb{R}^n$, is constrained to being a semi-definite positive matrix.*

$$\begin{array}{ll} \min & C \cdot X \\ & A \cdot X \leq b_i, \quad \textit{for} \;\; i = 1, \ldots, m \\ & X \succcurlyeq 0, \end{array} \tag{3.8}$$

where $A \cdot B = \sum_{i=1}^n \sum_{j=1}^n A_{ij} B_{ij}$ denotes a generalized inner product, and the last constraints represent

$$\forall y \in \mathbb{R}^n, y^T X y \geq 0$$

which defines an infinite number of linear constraints.

One might notice that it is impossible to explicitly store all constraints in (3.6) or (7). It turns out that there is an extremely convenient way to describe the input of a convex programming problem in an implicit way.

Definition 8. *A separation oracle for a convex set $K \subset \mathbb{R}^n$ is an algorithm that for each $x \in \mathbb{R}^n$, either states that $x \in K$, or provides a vector $a \in \mathbb{R}^n$ such that*

$$a^T x < b \quad \textit{and} \quad a^T y \geq b \quad \textit{for all} \;\; y \in K$$

Exercise Prove that for any convex set, there exists a separation oracle. Returning to our examples,

5. All that is needed is to check each constraint and return a violated constraint if it is the case. In the complete formulation, finding a violating inequality can be done in polynomial time but is nontrivial.

6. A hyperplane tangent to the ball of radius R can be easily constructed.

7. Given $\bar{X}$, the separation oracle reduces to finding a vector v such that $\bar{X} v = -\lambda v$ for some positive scalar λ. This would imply that

$$v^T \bar{X} v = \lambda \Rightarrow \bar{X} \cdot (vv^T) = -\lambda,$$

while $X \cdot (vv^T) \geq 0$ for all feasible points.

3.1.1 Reducing Optimization to Feasibility

Before proceeding, we will reduce the problem (3.1), under the convexity assumptions previous mentioned, to a feasibility problem. That is, given a convex set K, find a point on K or prove that K is empty ("give a certificate" that K is empty).

For a scalar parameter t, we define the convex set

$$K \cap \{x \in \mathbb{R}^n : f(x) \leq t\}$$

If we can find a point in such a convex set, then we can solve (3.1) via binary search for the optimal t.

3.1.2 Feasibility Problem

Now, we are going to investigate a convex feasibility algorithm with the following rules:

Input: separate oracle for K, and $r, R \in \mathbb{R}_{++}$
(where a ball of radius r contained in K, K is contained in a ball of radius R)
Output: A point in K if K is not empty, or give a certificate that K is empty.

Theorem 9. *Any algorithm needs $n \log_2 \frac{R}{r}$ oracle calls in the worst case.*

Proof. Since we only have access to the oracle, the set K may be in the largest remaining set at each iteration. In this case, we can reduce the total volume at most by 1/2. It may happen for iteration k as long,

$$\left(\frac{1}{2}\right)^k < \left(\frac{r}{R}\right)^n$$

which implies that

$$k < n \log_2 \frac{R}{r}$$

□

In the seminal work of Khachiyan and its extension by GLS, the Ellipsoid Algorithm was proved to require $\mathcal{O}(n^2 \log \frac{R}{r})$ iterations for the convex feasibility problem. Subsequently, Karmarkar introduced interior points methods for the special case of linear programming; the latter turned out to be more practical (and polynomial-time).

We are going to consider the following algorithm for convex feasibility.

1. $P = \text{cube}(R)$, $z = 0$.
2. Call Separation Oracle for z. If $z \in K$ stop.
3. $P := P \cap \{x \in \mathbb{R}^n : a^T x \leq a^T z\}$
4. Define the new point $z \in P$.
5. Repeat steps 2, 3, and 4 N times
6. Declare that K is empty.

Two of the steps are not completely defined. It is intuitive that the choice of the points z will define the number of iterations needed, N. Many choices are possible here, e.g., the centroid, a random point, the average of random points, analytic center, point at maximum distance from the boundary, etc.

For instance, consider the centroid and an arbitrary hyperplane crossing it. As we proved in the previous lecture, P will be divided into two convex bodies, each having at least $\frac{1}{e}$ of the total volume of the starting set (Grunbaum's theorem). Since the initial volume is R^n and the final volume is at least r^n, we need at most

$$\log_{1-\frac{1}{e}} \left(\frac{R}{r}\right)^n = \mathcal{O}\left(n \log \frac{R}{r}\right)$$

iterations.

So, we have reduced convex optimization to convex feasibility, and the latter we reduced to finding centroid. Unfortunately, the problem of finding the centroid is #P-hard.

Next, let us consider z to be defined as a random point from P. For simplicity, consider the case where P is a ball of radius R centered at the origin. We would like z to be close to the center, so it is of interest to estimate its expected norm.

$$\begin{aligned} \mathsf{E}[\|x\|] &= \int_0^R \frac{\mathsf{Vol}(S_n) t^{n-1}}{\mathsf{Vol}(B_n) R^n} t dt && (3.9) \\ &= \frac{n}{R^n} \int_0^R t^n dt = \frac{n}{n+1} R. && (3.10) \end{aligned}$$

That is, a random point in the ball is close to the boundary. Now, we proceed to estimate how much the volume is decreasing at each iteration.

Defining $b = \mathsf{E}[\|x\|] = \frac{n}{n+1} R$, we obtain $a \leq R\sqrt{\frac{2}{n+1}}$. This implies that the volume is falling by

$$\left(\sqrt{\frac{2}{n+1}}\right)^n,$$

Figure 3.1: Sphere cap.

which is decreasing exponentially in n. Thus, it will not give us a polynomial-time algorithm.

Now, consider the average of random points,

$$z = \frac{1}{m} \sum_{i=1}^{m} y^i,$$

where y^i's are i.i.d. points in P.

We will prove shortly the following theorem.

Theorem 10. *Let z and $\{y^i\}_{i=1}^m$ be defined as above, we have*

$$\mathsf{E}[\mathsf{Vol}(P')] \leq \left(1 - \frac{1}{e} + \sqrt{\frac{n}{m}}\right) \mathsf{Vol}(P).$$

Thus, the following corollary will hold.

Corollary 11. *With $m = 10n$, the algorithm finishes in $\mathcal{O}(n \log \frac{R}{r})$ iterations with high probability.*

We will begin to prove Theorem (10) for balls centered at the origin. This implies that $\mathsf{E}[z] = 0$ and, using the radial symmetry,

$$\begin{aligned} \mathsf{var}(\|z\|) \leq \mathsf{E}[\|z\|^2] &= \frac{1}{m}\mathsf{E}[|y_i|^2] \\ &= \frac{1}{m}\int_0^R \frac{\mathsf{Vol}(S_n)t^{n-1}}{\mathsf{Vol}(B_n)R^n} t^2 dt \\ &= \frac{n}{mR^n}\int_0^R t^{n+1} dt = \frac{n}{m(n+2)}R^2. \end{aligned}$$

To conclude the proof for this special case, using the inequality

$$P\left(\|z\| > \frac{R}{\sqrt{n}}\right) \leq \frac{Var(z)n}{R^2} = \frac{nR^2}{m(n+2)}\frac{n}{R^2} \leq \frac{n}{3m}, \tag{3.11}$$

using $m = \mathcal{O}(n)$, with a fixed probability, we obtain $\|z\| \leq \frac{R}{\sqrt{n}}$. Now $b = \frac{R}{\sqrt{n}}$, thus,

$$a^2 = \left(R + \tfrac{R}{\sqrt{n}}\right)\left(R - \tfrac{R}{\sqrt{n}}\right) = R^2\left(1 - \tfrac{1}{n}\right). \tag{3.12}$$

This implies that the volume is falling by

$$\left(1 - \frac{1}{n}\right)^{n/2} \geq \frac{1}{\sqrt{e}}.$$

To extend the previous result to an ellipsoid E, it is enough to observe that if $A : \mathbb{R}^n \to \mathbb{R}^n$ is an affine linear transformation, $E = AB$, which scales the volume of every subset by the same factor, namely, $|\det(A)|$.

To generalize for arbitrary convex sets, we introduce the following concept.

Definition 12. *A set S is in isotropic position if for a random point $x \in S$,*

1. $\mathsf{E}[x] = 0$;
2. $\forall v \in \mathbb{R}^n$, $\mathsf{E}[(v^T x)^2] = \|v\|^2$.

Next, we are going to prove a lemma which proves an equivalence between the second condition of isotropic position and conditions over the variance-covariance matrix of the random vector x.

Lemma 13. *Let x be a random point in a convex set S, we have $\forall v \in \mathbb{R}^n$, $\mathsf{E}[(v^T x)^2] = \|v\|^2$ if and only if $\mathsf{E}[xx^T] = I$.*

Proof. If $\mathsf{E}[(v^T x)^2] = \|v\|^2$, for $v = e_i$ we have

$$\mathsf{E}[(e_i^T x)^2] = \mathsf{E}[x_i^2] = 1.$$

For $v := (v_1, v_2, 0, \ldots, 0)$,

$$\begin{aligned} \mathsf{E}[(v^T x)^2] &= \mathsf{E}[v_1^2 x_1^2 + v_2^2 x_2^2 + 2 v_1 v_2 x_1 x_2] \\ &= v_1^2 + v_2^2 + 2 v_1 v_2 \mathsf{E}[x_1 x_2] \\ &= v_1^2 + v_2^2, \end{aligned}$$

where the last equality follows from the second condition of Definition 12. So, $\mathsf{E}[x_1 x_2] = 0$.

To prove the other direction,

$$\mathsf{E}[(v^T x)^2] = \mathsf{E}[(v^T x x^T v)] = v^T \mathsf{E}[xx^T] v = v^T v = \|v\|^2.$$

□

It is trivial to see that the first condition of isotropic position can be achieved by a translation. It turns out that is possible to obtain the second condition by means of a linear transformation.

Lemma 14. *Any convex set can be put in isotropic position by an affine transformation.*

Proof. $\mathsf{E}_K[xx^T] = A$ is positive definite. Then, we can write $A = B^2$ (or BB^T).

Defining, $y = B^{-1}x$, we get

$$\mathsf{E}[yy^T] = B^{-1} \mathsf{E}[xx^T] (B^{-1})^T = I.$$

The new set is defined as

$$K' := B^{-1}(K - z(K)),$$

where $z(K) := \mathsf{E}[x]$, the centroid of K. □

Note that the argument used before to relate the proof for balls to ellipsoids states that volume ratios are invariant under affine transformations. Thus, we can restrict ourselves to the case where the convex set is in isotropic position due to the previous lemma.

Lemma 15. *K is a convex body, z is the average of m random points from K. If H is a half space containing z,*

$$\mathsf{E}[\mathsf{Vol}(H \cap K)] \geq \left(\frac{1}{e} - \sqrt{\frac{n}{m}}\right) \mathsf{Vol}(K).$$

Proof. As state before, we can restrict ourselves to the case which K is in isotropic position. Since $z = \frac{1}{m}\sum_{i=1}^m y^i$,

$$\begin{aligned}
\mathsf{E}[\|z\|^2] = \frac{1}{m^2}\sum_{i=1}^m \mathsf{E}[\|y^i\|^2] &= \frac{1}{m}\mathsf{E}[\|y^i\|^2] \\
&= \frac{1}{m}\sum_{j=1}^n \mathsf{E}[(y^i_j)^2] = \frac{n}{m},
\end{aligned}$$

where the first equality follows from the independence between y^i's, and equalities of the second line follows from the isotropic position.

Let h be a unit vector normal to H, since the isotropic position is invariant under rotations, we can assume that $h = e_1 = (1, 0, \ldots, 0)$.

Define the $(n-1)$-volume of the slice $K \cap \{x \in \mathbb{R}^n : x_1 = y\}$ as $\mathsf{Vol}_{n-1}(K \cap \{x \in \mathbb{R}^n : x_1 = y\})$. We can also define the following density function,

$$f_K(y) = \frac{\mathsf{Vol}_{n-1}(K \cap \{x \in \mathbb{R}^n : x_1 = y\})}{\mathsf{Vol}(K)}. \tag{3.13}$$

Note that this function has the following properties:

1. $f_K(y) \geq 0$ for all $y \in \mathbb{R}$;
2. $\displaystyle\int_{\mathbb{R}} f_K(y)dy = 1$;
3. $\displaystyle\int_{\mathbb{R}} y f_K(y)dy = 0$;
4. $\displaystyle\int_{\mathbb{R}} y^2 f_K(y)dy = 1$;

The following technical lemma will give an useful bound to this density (we defer its proof).

Lemma 16. *If f_K satisfies the previous properties for some convex body K,*

$$\max_y f(y) \leq \frac{n}{n+1}\sqrt{\frac{n}{n+2}} < 1$$

Combining this lemma with the bound on the norm of z,

$$\begin{aligned}
\frac{\mathsf{Vol}(K \cap \{x\mathbb{R}^n : x_1 \geq z_1\}}{\mathsf{Vol}(K)} &= \int_{z_1}^{\infty} f_K(y)dy \\
&= \int_0^{\infty} f_K(y)dy - \int_0^{z_1} f_K(y)dy \\
&\geq \frac{1}{e} - \int_0^{z_1} \max_y f_K(y)dy \\
&\geq \frac{1}{e} - \|z\| \\
&\geq \frac{1}{e} - \sqrt{\frac{n}{m}}.
\end{aligned} \tag{3.14}$$

□

Proof. (of Lemma 16) Let K be a convex body in isotropic position for which $\max_y f_K(y)$ is as large as possible. In Figure 3.1.2, we denote by $y^* \in \arg\max_y f_K(y)$, the centroid is zero, and the set K is the black line.

In Figure 3.1.2, the transformation in part I consists of replacing it with a cone whose base is the cross-section at zero and its apex is on the x_1 axis. The distance of the apex is defined to preserve the volume of part I. Such a transformation can only move mass to the left of the centroid.

The transformation in part II of figure 3.1.2 is to construct the convex hull between the cross sections at zero and y^*. This procedure may lose some mass (never increases). Part III is also replaced by a cone, with base being the cross-section at y^* and the distance of the apex is such that we maintain the total volume. Thus, the transformations in parts II and III move mass only to the right, i.e., away from the centroid.

Defining the moment of inertia[4], as $I(K) = \int_K (y - \mathsf{E}_K[y])^2 f_K(y)dy$, we make the following claim.

Claim 17. *Moving mass away from the centroid of a set can only increase the moment of inertia I.*

[4]Note that $\mathsf{E}_K[y] = 0$ since K is in isotropic position.

Figure 3.2: Construction to increase the moment of inertia I.

Let us define a transformation $C \to C'$, $x_1 \to x_1 + g(x_1)$, where $x_1 g(x_1) \geq 0$.

$$\begin{aligned} I(C) &= Var_C(y) = \mathsf{E}_C[y^2] \\ I(C') &= Var_{C'}(y) = \mathsf{E}_{C'}[y^2] - (\mathsf{E}_{C'}[y])^2 \\ &= \mathsf{E}_C[(y + g(y))^2] - (\mathsf{E}_C[y + g(y)])^2 \\ &= Var_C(y) + Var_C(g(y)) + 2\mathsf{E}_C[yg(y)] \\ &\geq I(C). \end{aligned}$$

In fact, we have strict inequality if g(y) is nonzero on a set of positive measure. In this case, we can shrink the support of f_K by a factor strictly smaller that one, and scale up f_K by a factor greater than one to get a set with moment of inertia equal to 1. In this process $\max f_K$ has increased. This is a contradiction since we started with maximum possible. Thus, K must have the shape of K'.

Now, we proceed for the transformation in Figure 3.1.2.

Figure 3.3: Construction to reduce to a double cone.

In Figure 3.1.2, y is chosen to preserve the volume in Part I, while Part II is not altered. Note that all the mass removed from the segment between $[-y, y]$ is replaced at a distance to zero greater than y. Thus, the moment of inertia increases. By the same argument used before, we obtain a contradiction. Hence, K must be a double cone.

Once we obtain a double cone, we can move y^* to the last point to the right, and build a cone with the same volume which has a greater moment of inertia. So K must be this cone with length h and with base area A at y^*. Then, it is an easy exercise in integration to get $h = (n+1)\sqrt{\frac{n+2}{n}}$. Since $\mathsf{Vol}(K) = \frac{Ah}{n}$,

$$f(y^*) = \frac{A}{\mathsf{Vol}(K)} = \frac{n}{h} = \frac{n}{n+1}\sqrt{\frac{n}{n+2}} < 1.$$

$\square$

3.1.3 Membership Oracle

We have reduced quasi-convex optimization over convex sets to a feasibility problem and solved the latter using a polynomial number of calls to a separation oracle.

We note that with only a membership oracle, given an arbitrary point x in the space, it will return a certificate that $x \in K$ or a certificate that $x \notin K$, and a point $x \in K$ such that $x+rB \subseteq K \subseteq x+RB$, the optimization problem can still be solved using a polynomial number of oracle calls. The idea is to use the oracle to generate random samples (this part needs only a membership oracle), then restrict the set using the objective function values at the sample points. The first polynomial algorithm in this setting was given by GLS using a variant of the Ellipsoid algorithm.