18.409 An Algorithmist's Toolkit	October 22, 2009

Lecture 12

Lecturer: Jonathan Kelner — *Scribes: Alex Levin (2009)*

1 Outline

Today we'll go over some of the details from last class and make precise many details that were skipped. We'll then go on to prove Fritz John's theorem. Finally, we will start discussing the Brunn-Minkowski inequality.

2 Separating Hyperplanes

Given a convex body $K \subseteq \mathbb{R}^n$ and a point p, a *separating hyperplane* for K and p is a hyperplane that has K on one side of it and p on the other. More formally, for a vector ν, the hyperplane $H = \{x | \nu \cdot x = 1\}$ is a separating hyperplane for K and p if

1. $\nu \cdot x \leq 1$ for all $x \in K$, and
2. $\nu \cdot p \geq 1$.

Note that if we replace the right hand side of both the above conditions by 0 or any other constant, we get an equivalent formulation.

We call a separating hyperplane H a *strongly separating hyperplane* if the second inequality is strict.

Last time, we sketched a proof of the following theorem:

Theorem 1 (Separating Hyperplane Theorem) *If K is a convex body and p is a point not contained in K, then there exists a hyperplane that strongly separates them.*

We'll use the above result to show why the polar of the polar of a convex body is the body itself. Recall that for a convex body K, we had defined its polar K^* to be $\{p | k \cdot p \leq 1 \forall k \in K\}$.

Theorem 2 *Let K be a convex body. Then $K^{**} = K$.*

Proof We know that $K^* = \{p | k \cdot p \leq 1 \forall k \in K\}$. Similarly $K^{**} = \{y | p \cdot y \leq 1 \forall p \in k^*\}$. Let y be any point in K. Then, by the definition of the polar, for all $p \in K^*$ we have that $p \cdot y \leq 1$. The definition of the polar of K^* implies that $y \in K^{**}$. Since this is true for every $y \in K$, we conclude that $K \subseteq K^{**}$.

The other direction of the proof is the nontrivial one and we'll have to use the convexity of the body and the separating hyperplane theorem. Suppose that we can find a $y \in K^{**}$ such that $y \notin K$. Since $y \in K^{**}$, we have that $p \cdot y \leq 1$ for all $p \in K^*$. Since $y \notin K$, there exists a strongly separating hyperplane for y and K. Let it be $H = \{x | v \cdot x = 1\}$. By the definition of separating hyperplane, we have $v \cdot k \leq 1$ for all $k \in K$. Hence, $v \in K^*$. Also, $v \cdot y > 1$ (since H is a separating hyperplane), and we just showed that $v \in K^*$. This contradicts our assumption that $y \in K^{**}$. Hence $K^{**} \subseteq K$. ■

3 Banach–Mazur Distance

Recall from last time the definition of the Banach–Mazur distance between two convex bodies:

Definition 3 *Let K and L be two convex bodies. The Banach–Mazur distance $d(K, L)$ is the least positive $d \in \mathbb{R}$ for which there is a linear image L' of L such that $L' \subseteq K \subseteq dL'$, where dL' is the convex body obtained by multiplying every vector in L' by the scalar d.*

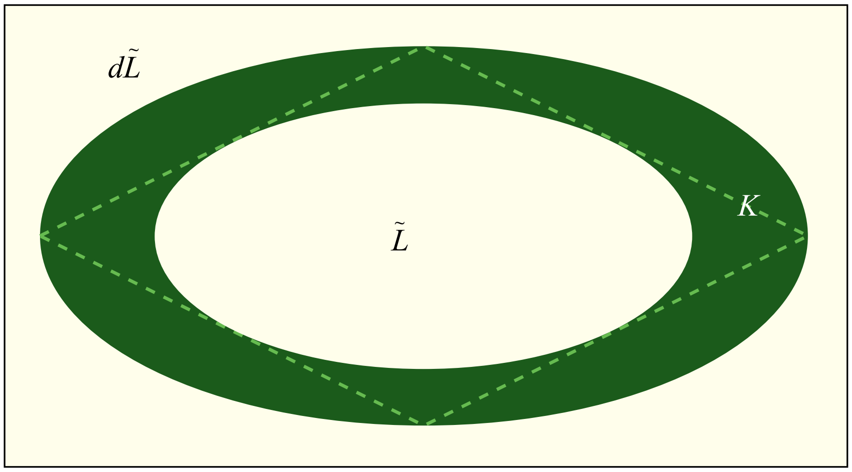

Figure 1: Defining the distance between K and L.

Observe that the above definition takes into consideration only the intrinsic shape of the body, and it is independent of any particular choice of coordinate system. Also observe that the Banach–Mazur distance is symmetric in its input arguments. If $L' \subseteq K \subseteq dL'$, then by scaling everything by d, we get that $dL' \subseteq dK$. Hence $K \subseteq dL' \subseteq dK$, which implies the symmetry property.

4 Fritz John's Theorem

Let B_2^n denote the n-dimensional unit ball. For any two convex bodies K and K', let $d(K, K')$ denote the Banach–Mazur distance between them. In the rest of this lecture, we will state and prove Fritz John's theorem.

Theorem 4 *For any n-dimensional, origin-symmetric convex body K, we have $d(K, B_2^n) \leq \sqrt{n}$.*

In other words, the theorem states that for every origin-symmetric convex body K, there exists some ellipsoid E such that $E \subseteq K \subseteq \sqrt{n}E$. We will prove that the ellipsoid of maximal volume that is contained in K will satisfy the above containment.

Informally, the theorem says that up to a factor of $\sqrt{n}$, every convex body looks like a ball. The above bound of $\sqrt{n}$ is tight for the cube. If we didn't require the condition that K is origin symmetric, then the bound would be n, which would be tight for a simplex.

The theorem can also be rephrased as the following: there exists a change of the coordinate basis for which $B_2^n \subseteq K \subseteq \sqrt{n}B_2^n$.

4.1 A slightly stronger version of the Fritz John Theorem

We will actually state and prove a more technical and slightly stronger version of the Fritz John theorem that implies our previous formulation. From now on, we assume that all the convex bodies we consider are origin-symmetric.

Theorem 5 *Let K be an origin-symmetric convex body. Then K contains a unique ellipsoid of maximal volume. Moreover, this largest ellipsoid is B_2^n if and only if the following conditions hold:*

- $B_2^n \subseteq K$
- *There are unit vectors $u_1, u_2, \ldots, u_m$ on the boundary of K and positive real numbers $c_1, c_2, \ldots, c_m$ such that*

1. $\sum_{i=1}^m c_i u_i = 0$, *and*

2. For all vectors x, *we have* $\sum_{i=1}^m c_i \langle x, u_i\rangle^2 = |x|^2$. *It is not hard to show that this condition is equivalent to the requirement that* $\sum_{i=1}^m c_i u_i u_i^T = \mathrm{Id}_n$, *where* Id_n *is the* $n \times n$ *identity matrix.*

Since the u_i are unit vectors, they are points on the convex body K that also belong to the sphere B_2^n. Also, the first identity, i.e. $\sum_{i=1}^m c_i u_i = 0$, is actually redundant, since for origin-symmetric bodies it can be derived from the second identity. This is because for every u_i, its reflection in the origin (namely $-u_i$) is also contained in $K \cap B_2^n$; further we can take the constants in the second identity corresponding to u_i and $-u_i$ to be the same, and this establishes the first equation.

The second identity says that the contact points of the sphere with K act somewhat like an orthonormal basis. They can be weighted so that they are completely isotropic. In other words, the points are not concentrated near some proper subspace, but are pretty evenly spread out in all directions. Together they mean that the u_i can be weighted so that their center of mass is the origin and their inertia tensor is the identity. Also, a simple rank argument shows that there need to be at least n such contact points, since the second identity can only hold for x in the span of the u_i.

Note that Theorem 4 easily follows from Theorem 5. Indeed, assume without loss of generality that B_2^n is the ellipsoid of maximal volume contained in K. We can make this assumption since the particular choice of basis is not important for the proof. We need to show that $B_2^n \subseteq K \subseteq \sqrt{n} B_2^n$. Now, for all $x \in K$, we have $x \cdot u_i \le 1$ for all i. Hence, $|x|^2 = \sum c_i (x \cdot u_i)^2 \le \sum c_i$. In the course of the proof below, we will see that $\sum c_i = n$. This shows that $|x| \le \sqrt{n}$, and hence $K \subseteq \sqrt{n} B_2^n$.

Thus, once we prove Theorem 5, we will have shown the existence of an ellipsoid E such that

$$E \subseteq K \subseteq \sqrt{n} E.$$

4.2 Proof of John's Theorem

As part of the proof Theorem 5, we will prove the following things:

1. If there exist contact points $\{u_i\}$ as required in the statement of Theorem 5, then B_2^n is the unique ellipsoid of maximal volume that is contained in K.
2. If B_2^n is the unique ellipsoid of maximal volume that is contained in K, then there exist points $\{u_i\}$ such that they satisfy the two identities in Theorem 5.

To prove the first statement, suppose that we are given unit vectors $u_1, u_2, \ldots, u_m$ on the boundary of K and positive real numbers $c_1, c_2, \ldots, c_m$ such that $\sum_{i=1}^m c_i u_i = 0$, and for all vectors x, it is the case that $\sum_{i=1}^m c_i \langle x, u_i\rangle^2 = |x|^2$. We wish to show that B_2^n is the unique ellipsoid of maximal volume that is contained in K. Observe that it suffices to show that among all axis-aligned ellipsoids contained in K, B_2^n is the unique ellipsoid of maximal volume. This is because what we are trying to prove doesn't mention any basis and is only in terms of dot products. Hence, since the statement will remain true under rotations, proving it for axis-aligned ellipsoids is enough.

For each u_i it is the case that $u_i \cdot k \le 1$ for all $k \in K$, Hence $u_i \in K^*$. Let E be any axis-aligned ellipsoid such that $E \subseteq K$. Then $K^* \subseteq E^*$. Hence $\{u_1, u_2, \ldots, u_m\} \subseteq E^*$. Since E is axis-aligned, it is of the form $\left\{x \middle| \sum_{i=1}^n \frac{x_i^2}{\alpha_i^2} \le 1\right\}$.

We also have that $\mathrm{Vol}(E)/\mathrm{Vol}(B_2^n) = \prod_{i=1}^n \alpha_i$. Therefore, to show that $\mathrm{Vol}(E) < \mathrm{Vol}(B_2^n)$, we must show that $\prod_{i=1}^n \alpha_i < 1$ for any such E that is not B_2^n.

Observe that $E^* = \{y | \sum_{i=1}^n \alpha_i^2 y_i^2 \le 1\}$. Now, since $u_i \cdot u_i = 1$, we have $\mathrm{Tr}\left(\sum_{i=1}^m c_i u_i u_i^T\right) = \sum_{i=1}^n c_i$. Since $\mathrm{Tr}(\mathrm{Id}_n) = n$, this implies that $\sum_{i=1}^n c_i = n$.

Let e_j denote the vector which has a 1 in the j^{th} coordinate and 0 in the other coordinates. Clearly $\langle u_i, e_j\rangle$ is the j^{th} coordinate of u_i. For $1 \le i \le m$, since $u_i \in E^*$, we get that $\sum_{j=1}^n \alpha_j^2 \langle u_i, e_j\rangle^2 \le 1$. Summing it over all i, we get

$$\sum_{i=1}^m c_i \sum_{j=1}^n \alpha_j^2 \langle u_i, e_j\rangle^2 \le \sum_{i=1}^n c_i = n. \tag{1}$$

Now, switching the order of summation on the left-hand side of (1) gives us

$$\sum_{j=1}^{n} \alpha_j^2 \sum_{i=1}^{m} c_i \langle u_i, e_j \rangle^2,$$

and by the above we know that this is at most n. Further, by condition 2 of Theorem 5, we know that $\sum_{i=1}^{m} c_i \langle u_i, e_j \rangle^2 = |e_j|^2 = 1$. Therefore, we get $\sum_{j=1}^{n} \alpha_j^2 \leq n$. By the AM-GM inequality, we get that $\left(\prod_{i=1}^{n} \alpha_i^2\right)^{1/n} \leq \frac{\sum_{i=1}^{n} \alpha_i^2}{n} \leq 1$, which implies that $\prod_{i=1}^{n} \alpha_i \leq 1$. Equality only holds if all the α_i are equal. This shows that $\prod_{i=1}^{n} \alpha_i < 1$ for any such E that is not B_2^n, completing the first part of the proof.

For the second part, assume that we are given that B_2^n is the unique ellipsoid of maximal volume that is contained in K. We want to show that for some m, there exist c_i and u_i for $1 \leq i \leq m$ (as in the statement of Theorem 5), such that for all vectors x, $\sum_{i=1}^{m} c_i \langle x, u_i \rangle^2 = |x|^2$. Again, this is equivalent to showing that

$$\sum_{i=1}^{m} c_i u_i u_i^T = \mathrm{Id}_n .$$

We already observed that for origin-symmetric bodies, the condition that $\sum_{i=1}^{m} c_i u_i = 0$ is implied by the previous requirement.

Let $U_i = u_i u_i^T$. Also, observe that we can view the space of $n \times n$ matrices as a vector space of dimension n^2. Hence we can parametrize the space of $n \times n$ matrices by $\mathbb{R}^{n^2}$. Thus, $\sum_{i=1}^{m} c_i u_i u_i^T = \mathrm{Id}_n$ for $c_i > 0$ means that Id_n / n is in the convex hull of the U_i (if the identity holds, we know that the c_i are positive and sum to n).

If we cannot find c_i, u_i such that $\sum_{i=1}^{m} c_i u_i u_i^T = \mathrm{Id}_n$, it means that Id_n / n is not in the convex hull of the U_i. Hence, there must be a hyperplane in the space of matrices that separates Id_n / n from the convex hull of the U_i.

For two $n \times n$ matrices A and B, let $A \bullet B$ denote their dot product in $\mathbb{R}^{n^2}$, i.e. $A \bullet B = \sum_{i,j} A_{ij} \cdot B_{ij}$. Thus, the separating hyperplane gives a matrix H such that $A \bullet H \geq 1$ for all $A \in \mathrm{conv}(U_i)$ and $(\mathrm{Id}_n / n) \bullet H < 1$.

Let $t = \mathrm{Tr}(H) = H \bullet \mathrm{Id}_n$. Let $H' = H - (t/n)(\mathrm{Id}_n)$. Then $(\mathrm{Id}_n / n) \bullet H' = (\mathrm{Id}_n / n) \bullet (H - (t/n)\,\mathrm{Id}_n) = t/n - ((\mathrm{Id}_n / n) \bullet (t/n)\,\mathrm{Id}_n) = 0$. Similarly, since $\mathrm{Tr}(A) = 1$ for all A in $\mathrm{conv}(U_i)$, we get that $A \bullet H' > 0$. Hence, H' is such that:

1. $\mathrm{Tr}(H') = 0$, and

2. $H' \bullet (u_i u_i^T) > 0$ for all i.

Now, let $E_\delta = \left\{x \in \mathbb{R}^n | x^T (\mathrm{Id}_n + \delta H') x \leq 1\right\}$. For all i, we have $u_i^T (\mathrm{Id}_n + \delta H') u_i = 1 + \delta u_i^T H' u_i$, which is greater than 1 since $u_i^T H' u_i > 0 = H' \bullet (u_i u_i^T) > 0$. Hence $u_i \notin E_\delta$. Also, since $H' \bullet (u_i u_i^T) > 0$ for all i, by continuity, there exists $\epsilon > 0$ such that for all vectors w in the ϵ-neighborhood of the set of all u_i satisfy $H' \bullet (w w^T) > 0$. Hence, by the previous argument, any such w is not contained in E_δ.

Note that when $\delta = 0$, we get the unit ball B_2^n. For every $\delta > 0$ we have that all w in the ϵ-neighborhood of the contact points of B_2^n are not contained in E_δ. Hence, as we increase δ continuously starting from 0, the continuity of the transformation of E_δ implies that for sufficiently small δ, $\mathrm{boundary}(K) \cap E_\delta = \emptyset$.

Hence $\exists \epsilon' > 0$ such that $(1 + \epsilon') E_\delta \subseteq K$. Therefore, to conclude the proof, it suffices to show that $\mathrm{Vol}(E_\delta) \geq \mathrm{Vol}(B_2^n)$, which gives give us a contradiction (as $(1 + \epsilon') E_\delta$ is an ellipse of volume larger than B_2^n contained in K).

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of $\mathrm{Id}_n + \delta H'$. Since $\mathrm{Vol}(E_\delta) = (\prod_{i=1}^{n} \lambda_i)^{-1}$, to show that $\mathrm{Vol}(E_\delta) \geq \mathrm{Vol}(B_2^n)$, we need to show that $\prod_{i=1}^{n} \lambda_i \leq 1$. However we know that $\sum_{i=1}^{n} \lambda_i = \mathrm{Tr}(\mathrm{Id}_n + \delta H') = \mathrm{Tr}(\mathrm{Id}_n) = n$. By the AM-GM inequality, $(\prod_{i=1}^{n} \lambda_i)^{1/n} \leq (\sum_{i=1}^{n} \lambda_i)/n = 1$. Hence $\prod_{i=1}^{n} \lambda_i \leq 1$. This concludes the proof of part 2.

5 Sketch of a Simpler Proof

If we just wish to prove the existence of an ellipse E that satisfies the conditions of Fritz John's Theorem without actually characterizing it, then the picture below suggests an alternative and possibly simpler proof of the result.

If any point of K is more than $\sqrt{n}$ distance away from the origin, then we can find an ellipse of larger volume than B_2^n that is contained in K.

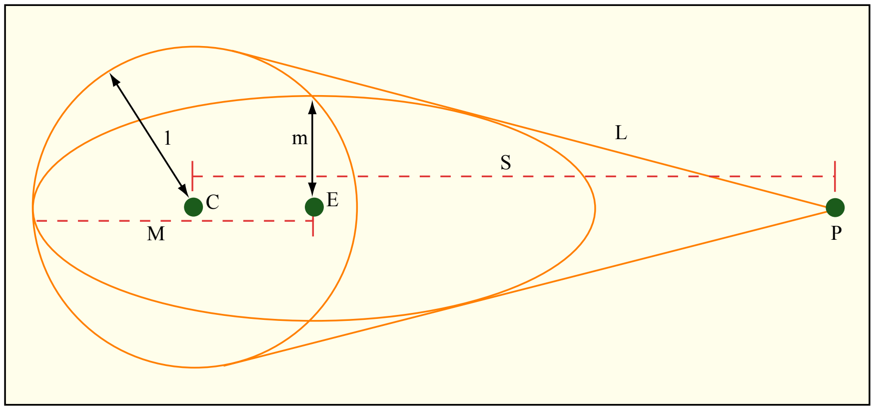

Image by MIT OpenCourseWare.

Figure 2: A simpler proof of the "Rounding" result.

6 The Brunn-Minkowski Inequality

Definition 6 *For $A, B \in \mathbb{R}^n$, the Minkowski sum $A \oplus B$ is given by*

$$A \oplus B = \{a + b | a \in A, b \in B\}.$$

The Minkowski sum can be defined for any subsets of $\mathbb{R}^n$, but it is nicely behaved if A and B are convex. Intuitively, the Minkowski sum is obtained by moving one of the sets around the boundary of the other one.

The Brunn-Minkowski inequality, which relates the volume of $A \oplus B$ to the volumes of A and B, implies many important theorems in convex geometry. The goal is to bound $\mathrm{Vol}(A \oplus B)$ in terms of $\mathrm{Vol}(A)$ and $\mathrm{Vol}(B)$. The following are some loose bounds that can be simply verified.

Fact 7 $\mathrm{Vol}(A \oplus B) \geq \max\{\mathrm{Vol}(A), \mathrm{Vol}(B)\}$

Proof Let $a \in A$. We have $\{a\} \oplus B \subseteq A \oplus B$, by definition. Hence,

$$\mathrm{Vol}(A \oplus B) \geq \mathrm{Vol}(\{a\} \oplus B) = \mathrm{Vol}(B).$$

Similarly, $\mathrm{Vol}(A \oplus B) \geq \mathrm{Vol}(B)$. ■

Fact 8 $\mathrm{Vol}(A \oplus B) \geq \mathrm{Vol}(A) + \mathrm{Vol}(B)$

Proof By moving one of the sets around the other one (summing the extreme points), we can get disjoint copies of A and B in $A \oplus B$. ■

The bound given by Fact 8 is loose. To see that, consider the case that $A = B$. In this case, $A \oplus A = 2A$ and hence, $\mathrm{Vol}(A \oplus A) = 2^n \mathrm{Vol}(A)$. So the volume of $A \oplus A$ grows exponentially with n, while the lower bound given in the above fact do not. This suggests taking the n-th roots and still get a valid bound. Let us first prove it for boxes.

Lemma 9 *Let A and B be boxes in $\mathbb{R}^n$. Then*

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

Proof Let A have sides of length $a_1, \ldots, a_n$ and B have sides of length $b_1, \ldots, b_n$. It directly follows from the definition of Minkowski sums that $A \oplus B$ has sides of length $a_1 + b_1, \ldots, a_n + b_n$.

We just need to show the following:

$$\frac{\mathrm{Vol}(A)^{1/n} + \mathrm{Vol}(B)^{1/n}}{\mathrm{Vol}(A \oplus B)^{1/n}} \leq 1. \tag{2}$$

We can rewrite the left-hand side of (2) as

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

where the inequality is just an application of AM-GM. ■

Next time, we will prove the Brunn-Minkowski inequality for more general bodies, and study some of its applications.