Lecture 11

Lecturer: Jonathan Kelner *Scribe: Chaithanya Bandi*

1 Outline

Today we'll introduce and discuss

- Polar of a convex body.
- Correspondence between norm functions and origin-symmetric bodies (and see how convex geometry can be a powerful tool for functional analysis).
- Fritz-John's Theorem

2 The Polar of a Polytope

Given a bounded polytope $C \subset \mathbb{R}^n$ that contains the origin in its interior, we can represent C as

$$C = \{x | a_i \cdot x \leq b_i, i = 1, \dots, k\},$$

where $b_i > 0$.

Without loss of generality, by appropriately scaling each constraint, we can assume $b_i = 1, \forall i = 1, \dots, k$. Now the polar of C is given by

$$C^* = conv(a_1, \dots, a_k).$$

2.1 Examples

Let C be the square with corners at $(1, 1), (1, -1), (-1, 1), (-1, -1)$. Then $\{a_i\} = \{(1, 0), (0, 1), (-1, 0), (0, -1)\}$. The polar has corners at $(1, 0), (0, 1), (-1, 0), (0, -1)$. Note that the polar is a square rotated and shrunk into a diamond. This polytope is also referred to as the "cross polytope". Note that the facets of C become the vertices of C^* and vice versa. For example, the three dimensional cube's polar is the octahedron. Six facets and eight vertices correspond to eight facets and six vertices.

The size and shape of a polar tends to be the reverse of that of the original set. For example, a short bulging rectangle with corners at $(100, 3), (100, -3), (-100, 3), (-100, -3)$ would have a tall compressed polar with corners at $(\pm 1/100, 0), (0, \pm 1/3)$. Also note that polars of simplices are simplices.

2.2 Properties of a polar

Some of the useful properties of a polar is summarised here. The properties will be illustrated using pictures.

- $(C^*)^* = C$ (proof later).
- If C is origin-symmetric, then so is C^*.
- If $A \subseteq B$ then $B^* \subseteq A^*$.
- If A is scaled up, then A^* is scaled down.
- If the polar is low-dimensional, that would mean the original polytope had to be unbounded in some directions.

- Translation has a very drastic effect on the polar. It can become unbounded just by translating the polytope.

All these properties can be illustrated using the pictures below.

3 Polars of General Convex Bodies

Any convex body can be thought of as the intersection of a (possibly infinite) set of half spaces. These are called "'suporting hyperplanes"'. Therefore, the polar of a convex body can be seen as the convex hull of a (possibly infinite) set of points, coming from all of the supporting hyperplanes. With this intuition one can guess about the following :

- Polar of a sphere is a sphere.
- Polar of a sphere of radius r is a sphere of radius $1/r$.
- Polar of an ellipse is an ellipse with axes reversed.

Definition 1 *The polar of a convex body C is given by*

$$C^* = \{x \in \mathbb{R}^n | x \cdot c \leq 1 \forall c \in C\}$$

We observe that this definition is equivalent to the previous definition.

Proposition 2 *For a polytope C given by $C = \{x | a_i \cdot x \leq b_i, i = 1, \ldots, k\}$, the sets $C_1 = C_2$ where $C_1 = \{x \in \mathbb{R}^n | x \cdot c \leq 1 \forall c \in C\}$ and $C_2 == conv(a_1, \ldots, a_k)$.*

We skip the proof as it is easy to verify that if $x \in C_1$ then $x \in C_2$ and vice versa.

We will now prove that $(C^*)^* = C$. We would be needing the concept of a separating hyperplane for the proof which we introduce now.

3.1 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.

Theorem 3 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.*

Proof We'll sketch an outline of the proof. It can be made rigorous. Consider a point $x \in K$ that is the closest to p in ℓ_2 distance. Consider the plane H that is perpendicular to the line joining x to p and is passing through the midpoint of x and p. H must separate K from p because if there is some point of K, say y, that is on the same side of H as p, then we can use the convexity of K to conclude that the point x' which is the intersection of the hyperplane with the line joining x and y is also in K. x' is closer to p that x since px' forms the side of a right angled triangle of which xp is the hypotenuse. This contradicts the assumption that x is the point closest to p. ■

3.2 Polar of a Polar

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 4 *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. If possible, let y be such that $y \in K^{**}$ and $y \notin K$. Since $y \in K^{**}$, we have that $P \cdot y \leq 1 \forall p \in K^*$. 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 $v \cdot k \leq 1 \forall k \in K$. 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 $K^{**} \subseteq K$. ■

4 Norms and Symmetric Convex Bodies

We will show how norms and symmetric convex bodies co-exist. This provides us a way to use the results of Convex Geometry in Functional Analysis and vice versa. Recall that a norm on $\mathbb{R}^n$ is a map $q : \mathbb{R}^n \to \mathbb{R}$ such that:

1. $q(ax) = aq(x)$ for $a \in \mathbb{R}$ (homogeneity)
2. $q(x + y) \leq q(x) + q(y)$ (triangle inequality)
3. $q(x) \geq 0$ for all x (nonnegativity) (actually implied by 1 and 2)
4. $q(x) = 0$ if and only if $x = 0$ (positivity) (without this conditions, q is a "seminorm")

Note that given a norm, one can construct a convex body. The simplest being the unit ball $B_q = \{x \in \mathbb{R}^n | q(x) \leq 1\}$. It is an easy exercise to verify the convexity of B_q.

Also as we will show now, given a convex body C, we can come up with a norm under which C is the unit ball. Note that C has to be origin symmetric.

Definition 5 *The Minkowski functional of an origin symmetric convex body C is the map $p_C : \mathbb{R}^n \to \mathbb{R}$ defined by*

$$p_C(x) = \inf_{\lambda > 0} \{x \in \lambda C\}$$

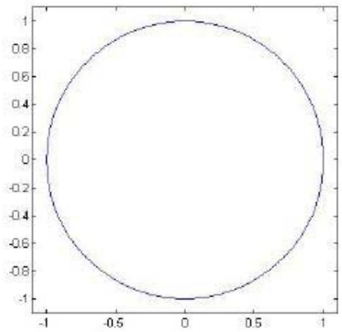
(We will sometimes denote this by $||x||_C$, because it is a norm.)

To prove that this is a norm, one needs to verify the properties of homogeneity, triangle inequality etc. These follow from the convexity of the body.

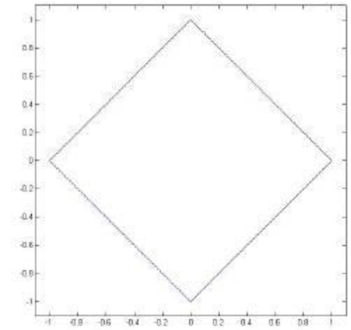
4.1 Norms, Duals, and the Polar

For any norm q, we can define its dual by $q^*(x) = \sup_{v \neq 0} |\frac{v \cdot x}{q(v)}|$. It is an exercise to see that the unit ball with respect to the dual norm of q is the polar of the unit ball with respect to q. This provides us a direct relation between convex geometry and functional analysis.

The following pictures allow us to have a geometric intuition of the norms and their duals.

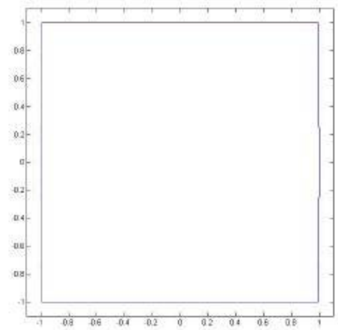


p = 2

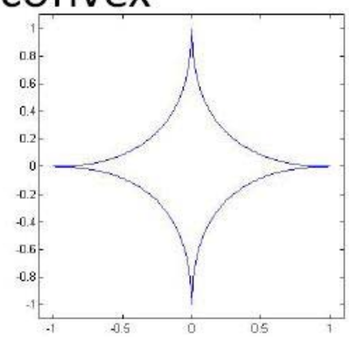


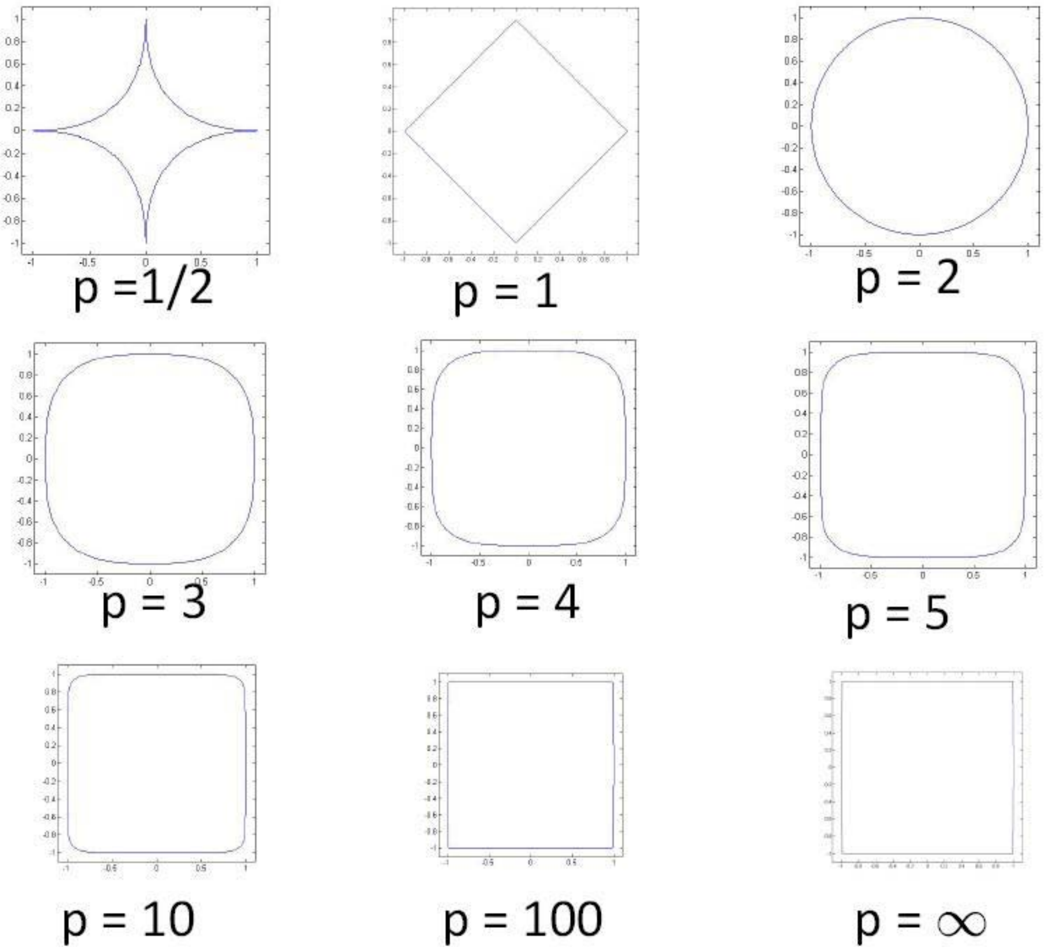
p = 1

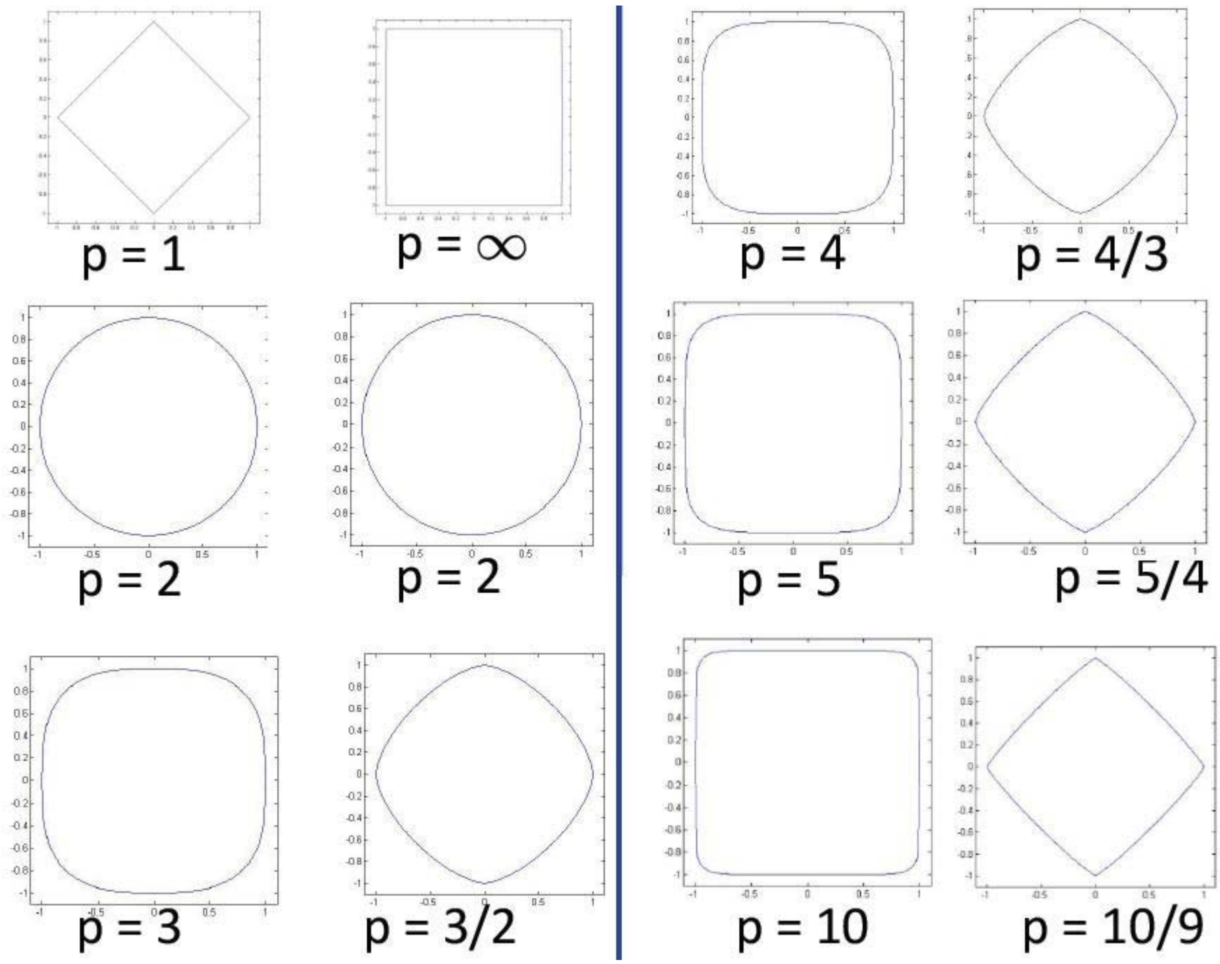
p = ∞



p = ½: not a norm, and not convex




p =1/2
p = 1
p = 2
p = 3
p = 4
p = 5
p = 10
p = 100
p = ∞



5 Banach–Mazur Distance

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

Definition 6 *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 theres 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.*

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 it's 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.

6 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'll state and prove the Fritz John's theorem.

Theorem 7 *For any n-dimensional, origin-symmetric convex body K, $d(K, B_2^n) \leq \sqrt{n}$.*

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'll 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 $B_2^n \subseteq K \subseteq \sqrt{n}B_2^n$.

Theorem 8 *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, $\sum_{i=1}^m c_i \langle x, u_i \rangle^2 = |x|^2$.*

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, it's reflection in the origin is also contained in $K \cap B_2^n$.

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.

6.1 Proof of John's Theorem

Proof As part of the proof of John's Theorem, we'll prove the following things:

1. If there exist contact points $\{u_i\}$ as required in the statement of Theorem 8, 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 8.

Proof of 1: 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, $\sum_{i=1}^m c_i \langle x, u_i \rangle^2 = |x|^2$. 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 we have that for all $k \in K$, $u_i \cdot k \leq 1$. Hence $u_i \in K^*$. Let E be any axis-aligned ellipsoid such that $E \in K$. Then $K^* \subseteq E^*$. Hence $\{u_1, u_2, \ldots, u_m\} \subseteq E^*$. Since E is axis-aligned, it is of the form $\{x | \sum_{i=1}^n \frac{x_i^2}{\alpha_i^2} \leq 1\}$.

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

Observe that $E^* = \{Y | \sum_{i=1}^n \alpha_i^2 y_i^2 \leq 1\}$. Also, condition 2 of Theorem 8 is equivalent to the following: $\sum_{i=1}^m c_i u_i u_i^T = Id_n$, where Id_n is the identity matrix of size n. Now, since $u_i \cdot u_i = 1$, we have $\text{Trace}(\sum_{i=1}^m c_i u_i u_i^T) = \sum_{i=1}^n c_i$. Since $\text{Trace}(Id_n) = n$, this implies that $\sum_{i=1}^n c_i = n$.

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

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

However, since by condition 2 of Theorem 8, $\sum_{i=1}^m \langle u_i, e_j\rangle^2 = |e_j|^2$, we get $\sum_{i=1}^n \alpha_i^2 \leq n$. By the AM-GM inequality, we get that $(\prod_{i=1}^n \alpha_i^2)^{1/n} \leq \frac{\sum_{i=1}^n \alpha_i^2}{n} \leq 1$, 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 which is not B_2^n, completing the first part of the proof.

Proof of 2**:** Assume that we are give 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 8), such that for all vectors x, $\sum_{i=1}^m c_i \langle x, u_i\rangle^2 = |x|^2$. This is equivalent to showing that

$$\sum_{i=1}^m c_i u_i u_i^T = Id_n.$$

Also, taking trace of both sides, we get that $\sum_{i=1}^m c_i = n$. 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 condition.

Let $U_i = u_i u_i^T$. Also, observe that we can view the space of $n \times n$ matrices as a vector of n^2 real numbers. Hence we can parametrize the space of $n \times n$ matrices by $\mathbb{R}^{n^2}$. Hence $\sum_{i=1}^m c_i u_i u_i^T = Id_n$ means that Id_n/n is in the convex hull of the U_i (recall that the c_i are positive and sum to 1).

If possible, let there be no c_i, u_i such that $\sum_{i=1}^m c_i u_i u_i^T = Id_n$. This means that Id_n/n is not in the convex hull of the U_i. Hence, there must be a separating hyperplane H 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 is a matrix H such that $\forall A \in \text{conv}(U_i)$, $A \bullet H \geq 1$, and $Id_n/n \bullet H < 1$.

Let $t = \text{Trace}(H) = H \bullet Id_n$. Let $H' = H - t/n(Id_n)$. Then $Id_n/n \bullet H' = Id_n/n \bullet (H - t/nId_n) = t/n - (Id_n/n \bullet t/nId_n) = 0$. Similarly, since $\forall A \in \text{conv}(U_i)$, $\text{Trace}(A) = 1$, we get that $A \bullet H' > 0$. Hence, H' is such that:

1. $\text{Trace}(H') = 0$, and
2. $H' \bullet (u_i u_i^T) > 0$ for all i.

Now, let $E_\delta = \{x \in \mathbb{R}^n | x^T(Id_n + \delta H')x \leq 1$. For all i, we have $u_i^T(Id_n + \delta H')u_i = 1 + \delta u_i^T H' u_i > 1$, since $H' \bullet (u_i u_i^T) > 0 \Rightarrow u_i^T H' u_i > 0$. Hence $u_i \notin E_\delta$. Also, since $H' \bullet (u_i u_i^T) > 0$ for all i, by compactness, there exists $\epsilon > 0$ such that for all matrices w in the ϵ-neighborhood of the set of all u_i satisfy $H' \bullet (ww^T) > 0$. 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 for all w in the ϵ-neighborhood of the contact points of B_2^n, $w \notin E_\delta$. Hence, as we increase δ continuously starting from 0, the continuity of the transformation of E_δ implies that for sufficiently small δ, $\text{boundary}(K) \cap E_\delta = \phi$.

Hence $\exists \epsilon' > 0$ such that $(1 + \epsilon')E_\delta \subseteq K$. Therefore, to conclude the proof, it suffices to show that $Vol(E_\delta \geq Vol(B_2^n)$.

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of $Id_n + \delta H'$. Since $Vol(E_\delta = (\prod_{i=1}^n \lambda_i)^{-1}$, to show that $Vol(E_\delta \geq Vol(B_2^n)$, we need to show that $\prod_{i=1}^n \lambda_i \leq 1$. However we know that $\sum_{i=1}^n \lambda_i = \text{Trace}(Id_n + \delta H') = \text{Trace}(Id_n) = n$. 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.

To wrap up the proof of John's Theorem, 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 $x \cdot u_i \leq 1$ for all i. Hence, $|x|^2 = \sum c_i (x \cdot u_i)^2 \leq \sum c_i = n$. This shows that $|x| \leq \sqrt{n}$, and hence $K \subseteq \sqrt{B_2^n}$.

Thus, we have proven the existence of an ellipse E such that

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

■

18.409 Topics in Theoretical Computer Science: An Algorithmist's Toolkit
Fall 2009