High-Dimensional Convex Geometry

Amit Rajaraman

Last updated December 14, 2020

Contents

§1. Introduction

Definition 1.1. A subset S of a Euclidean space is said to be *convex* if for any $u_1, \ldots, u_r \in S$ and non-negative $\lambda_1, \ldots, \lambda_r$ such that $\lambda_1 + \cdots + \lambda_r = 1$, the *affine combination* $\sum_{i=1}^r \lambda_i u_i$ is in S as well.

We primarily consider convex bodies, that is, compact and convex subsets of Euclidean spaces here. To put it more succinctly, a convex body is something that "behaves a bit like a Euclidean ball".
A few simple examples of convex bodies on $\mathbb{R}^n$ are:

- the cube $[-1, 1]^n$. Here, the ratio of the radii of the circumscribed ball to the inscribed ball is $\sqrt{n}$, so it is not much like a Euclidean ball.
- the *n-dimensional regular solid simplex* which is the convex hull of $n+1$ equally spaced points. Here, the ratio of the radii of the circumscribed ball to the inscribed ball is n. This ratio is "maximal" in some sense.
- the n-dimensional "octahedron" or *cross-polytope* which is the convex hull of the $2n$ points $(\pm 1, 0, \ldots, 0)$, $(0, \pm 1, 0, \ldots, 0)$, $\ldots, (0, 0, \ldots, 0, \pm 1)$. Note that this is the unit ball on the ℓ_1 norm on $\mathbb{R}^n$ so we denote it as B_1^n. Here, the ratio of the radii of the circumscribed ball to the inscribed ball is $\sqrt{n}$.

Definition 1.2. A *cone* in $\mathbb{R}^n$ is the convex hull of a single point and a convex body of dimension $n-1$. In $\mathbb{R}^n$, the volume of a cone of "height" h over a base of $(n-1)$-dimensional volume B is Bh/n.

Since B_1^n is made up of 2^n pieces similar to the piece with non-negative coordinates, which is a cone of height 1 with base analogous to the similar piece in $\mathbb{R}^{n-1}$, the volume of the non-negative section is $1/n!$. Therefore, the $\operatorname{vol}(B_1^n) = 2^n/n!$.

1.1. The Euclidean Ball

The fourth and final example is the Euclidean ball itself, namely

$$B_2^n = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^n x_i^2 \leq 1 \right\}.$$

Let us now attempt to calculate $v_n = \operatorname{vol}(B_2^n)$. Note that we can easily get the "surface area" of the ball from the volume by splitting it into "thin" cones from 0 and observing that the volume of each cone is equal to $1/n$ times its base area. Therefore, the surface area of the ball is nv_n.
We perform integration in spherical polar coordinates using two variables - r, which denotes the distance from 0 and θ, which is a point on the unit ball that represents the direction of the point. We obviously have $x = r\theta$. The point θ carries the information of $n-1$ coordinates.
We can then write the integral of a general function on $\mathbb{R}^n$ by

$$\int_{\mathbb{R}^n} f = \int_{r=0}^{\infty} \int_{S^{n-1}} f(r\theta) r^{n-1} \, \mathrm{d}\theta \, \mathrm{d}r \tag{1.1}$$

Here, $\mathrm{d}\theta$ represents the area measure on the sphere. From our earlier observation, its total mass is nv_n. The r^{n-1} factor appears because the sphere of radius r has r^{n-1} times that of the sphere of radius 1.
An important thing to note about the measure corresponding to $\mathrm{d}\theta$ is that it is *rotation-invariant.* If A is a subset of the sphere and U is orthogonal to A, then UA has the same measure as A. Therefore, we often simplify integrals such as 1.1 by pulling out the nv_n factor to get

$$\int_{\mathbb{R}^n} f = nv_n \int_{r=0}^{\infty} \int_{S^{n-1}} f(r\theta) r^{n-1} \, \mathrm{d}\sigma_{n-1}(\theta) \, \mathrm{d}r \tag{1.2}$$

where σ_{n-1} is the rotation-invariant measure on $\mathbb{R}^{n-1}$ of total mass 1. Now, to evaluate v_n, we choose a suitable f such that the integrals on either side can easily be calculated, namely

$$f : x \mapsto \exp\left(-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right).$$

Then the integral on the left of 1.2 is

$$\int_{\mathbb{R}^n} f = \int_{\mathbb{R}^n} \prod_{i=1}^{n} \exp\left(-\frac{x_i^2}{2}\right) = \prod_{i=1}^{n} \int_{-\infty}^{\infty} \exp\left(-\frac{x_i^2}{2}\right) = \left(\sqrt{2\pi}\right)^n$$

and the integral on the right is

$$nv_n \int_0^\infty \int_{S^{n-1}} e^{-r^2/2} r^{n-1} \, \mathrm{d}\sigma_{n-1} \, \mathrm{d}r = nv_n \int_0^\infty e^{-r^2/2} r^{n-1} \, \mathrm{d}r = v_n 2^{n/2} \Gamma\left(\frac{n}{2}+1\right).$$

Equating the two,

$$v_n = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}$$

Using Stirling's Formula, we can approximate this slightly better as

$$v_n \approx \frac{\pi^{n/2}}{\sqrt{2\pi} e^{-n/2} \left(\frac{n}{2}\right)^{(n+1)/2}} \approx \left(\frac{2\pi e}{n}\right)^{n/2}.$$

This is quite small for large n. The radius of a ball of volume 1 would be approximately $\sqrt{n/2\pi e}$, which is very large!

This is possibly the first hint one should take that following your intuition is probably not a good idea when dealing with high dimensional spaces.

Let us now restrict ourselves to considering the ball of volume 1.
What is the $(n-1)$-dimensional volume of a slice through the center of the ball? Since the slice is an $(n-1)$ dimensional ball, it is equal to

$$v_{n-1} r^{n-1} = v_{n-1} \left(\frac{1}{v_n}\right)^{(n-1)/n}.$$

This is approximately equal to $\sqrt{e}$ (using Stirling's formula once again). More generally, the volume of the slice that is at distance x from the center of the ball is equal to

$$\sqrt{e}\left(\frac{\sqrt{r^2-x^2}}{r}\right)^{n-1} = \sqrt{e}\left(1-\frac{x^2}{r^2}\right)^{(n-1)/2} \approx \sqrt{e}\left(1-\frac{2\pi e x^2}{n}\right)^{(n-1)/2} \approx \sqrt{e}\exp(-\pi e x^2)$$

Note that this is normally distributed but the variance $1/2\pi e$ *does not depend on* n*!*[1] So despite the fact that the radius grows as $\sqrt{n}$, the distribution of the volume stays the same. For example, nearly all the volume (around 96%) is concentrated in the slab with $|x_1| \leq 1/2$.

This might lead us to believe that since the volume is concentrated around any such equator[2] around a subspace, the volume should be concentrated around the intersection of all such equators, which seems to suggest that it should be concentrated around the center. However, for large n, we obviously know that most of the volume should be concentrated on the surface of the sphere[3]. These two points seem to be directly contradictory! However, as might

[1] exclamation mark, not factorial
[2] since we could have equally well taken something other than x_1
[3] the volume of a ball of radius dr $(d < 1)$ is $d^n \ll 1$ times that of a ball of radius r

be expected, this is once again because our intuition fails when dealing with high-dimensional spaces.
The measure of unit ball is "concentrated" both near the surface *and* around the equator, for any equator. To make more sense of this[4], while each x_i is small, the overall distance from 0 is quite large since the small individual coordinates are compensated by the large dimension n. The former leads to the point being close to the equator and the latter leads to the point being close to the surface of the ball.

Another fun[5] thing to think about is the following. Consider the cube $[-1,1]^n$. Construct a ball of radius $1/2$ at each of the 2^n vertices $(\pm 1, \ldots, \pm 1)$. Now, construct the ball with center $(\frac{1}{2}, \ldots, \frac{1}{2})$ that touches each of these 2^n balls. Then note that for $n = 4$, this ball touches (the center of each face of) the cube, and for $n \geq 5$, it actually goes *outside* the cube!

To conclude, let us write the volume of a general convex body K in spherical polar coordinates. Assume that K has 0 in its interior and for each direction $\theta \in S^{n-1}$, let $r(\theta)$ be the radius of K (in that direction). Then,

$$\operatorname{vol}(K) = n v_n \int_{S^{n-1}} \int_0^{r(\theta)} s^{n-1} \, \mathrm{d}s \, \mathrm{d}\sigma = v_n \int_{S^{n-1}} r(\theta)^n \, \mathrm{d}\sigma.$$

Definition 1.3. A convex body K is said to be *(centrally) symmetric* if $-x \in K$ whenever $x \in K$.

Any symmetric body is the unit ball under some $\|\cdot\|_K$ on $\mathbb{R}^n$ (for example, the octahedron was the unit ball under the ℓ_1 norm). For a general symmetric body with associated norm $\|\cdot\|$, the volume is given by

$$\operatorname{vol}(K) = v_n \int_{S^{n-1}} \|\theta\|^{-n} \, \mathrm{d}\sigma_{n-1}(\theta)$$

1.2. The Cube

So for example. since the volume of the cube $[-1,1]^n$ is 2^n, we can use it to estimate the average radius of the cube[6] as

$$v_n \int_{S^{n-1}} r(\theta)^n = 2^n \implies \int_{S^{n-1}} r(\theta)^n \approx \left(\sqrt{\frac{2n}{\pi e}}\right)^n \text{ so the average radius is approximately } \sqrt{\frac{2n}{\pi e}}$$

That is, the volume of the cube is far more concentrated towards the corners (where the radius is closer to $\sqrt{n}$), rather than the middles of facets (where the radius is closer to 1).
It can actually be shown[7] that the fraction of volume of the intersection of the cube and the ball is less than $\exp(-4n/45)$, which further emphasizes the point that nearly all the volume lies in the corners.

Definition 1.4. A body which is bounded by a finite number of flat facets is called a *polytope*.

A polytope is essentially the intersection of a finite number of half-spaces (we shall later look at the Hahn-Banach separation theorem, which is a huge generalization of this to convex bodies in general).
Note that the cube is a polytope with $2n$ facets.

Earlier, we remarked that the cube is not much like a Euclidean ball. So a question that might come to mind is: If K is a polytope with m facets, how close can K be to the Euclidean ball?

Let us define this "closeness" more concretely.

[4]The answers to this mathoverflow question might further aid understanding
[5]subject to debate
[6]more concretely, the expectation of a random variable that corresponds to the distance of a point uniformly randomly chosen from the unit cube from the origin
[7]Consider the random variable $z_i = x_i^2$ where x_i is drawn uniformly randomly from $[-1,1]$. Show that $\mathbf{E}[z_i] = 1/3$ and $\mathbf{Var}[z_i] = 4/45$ and use the Chernoff bound to get a bound on $\Pr[\sum_i z_i \leq 1]$.

Definition 1.5 (Banach-Mazur Distance)**.** The *Banach-Mazur distance* $d(K, L)$ between symmetric convex bodies K and L is the least positive d for which there is a linear image $\tilde{L}$ of L such that $\tilde{L} \subseteq K \subseteq d\tilde{L}$.

Henceforth, we refer to the Banach-Mazur distance as just *distance*.
This corresponds to the how we thought of inscribing/circumscribing a ball earlier, since the ratio of the two radii we considered is just this distance.
If we wanted to make this distance a metric, then we should consider $\log d$ instead of d (the current distance is multiplicative and for any K, $d(K, K) = 1$).

From what we mentioned earlier, we know that the distance between the cube and the Euclidean ball in $\mathbb{R}^n$ is at most $\sqrt{n}$. We shall prove later that it is indeed equal to $\sqrt{n}$.

As might be expected, if we want a polytope that approximates the ball very well, we would need a very large number of facets.

Definition 1.6. For a fixed unit vector v and some $\varepsilon \in [0, 1)$, the set

$$C(\varepsilon, v) = \{\theta \in S^{n-1} : \langle \theta, v \rangle \geq \varepsilon\}$$

is called the *ε-cap about v* or more generally, a *spherical cap* (or just *cap*).
It is often better to write a cap in terms of its radius rather than in terms of ε. The *cap of radius r about v* is

$$\{\theta \in S^{n-1} : |\theta - v| \leq r\}$$

It is easy to see that a cap of radius r is a $(1 - \frac{r^2}{2})$-cap.
As we shall see in the proof of Theorem 1.3, it is useful to know some upper and lower bounds on the area of an ε-cap.

Lemma 1.1 (Upper bound on the area of spherical caps)**.** For $\varepsilon \in [0, 1)$, the cap $C(\varepsilon, u)$ on S^{n-1} has measure (under σ_{n-1}) at most $e^{-n\varepsilon^2/2}$.

Lemma 1.2 (Lower bound on the area of spherical caps)**.** For $r \in [0, 2]$, a cap of radius r on S^{n-1} has measure (under σ_{n-1}) at least $\frac{1}{2}(r/2)^{n-1}$.

Proof. Suppose $n \geq 2$ and let $\alpha = 2\sin^{-1}(r/2)$. We can assume that $\alpha \in [0, \frac{\pi}{2}]$ since we can prove the other case similarly. Then the measure of the cap is given by

$$\begin{aligned}
A(n, \alpha) &= \int_0^\alpha \frac{(n-1)v_{n-1}}{nv_n} (\sin\theta)^{n-2} \,\mathrm{d}\theta \\
&= \frac{(n-1)\Gamma\left(\frac{n}{2}+1\right)}{n\Gamma\left(\frac{n-1}{2}+1\right)\sqrt{\pi}} \int_0^\alpha \sin^{n-2}(\theta)\,\mathrm{d}\theta \\
&= \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)\sqrt{\pi}} \int_0^\alpha \sin^{n-2}(\theta)\,\mathrm{d}\theta \\
&\geq \frac{1}{\sqrt{\pi}} \int_0^\alpha \left(\frac{2\theta}{\pi}\right)^{n-2} \mathrm{d}\theta \\
&= \frac{1}{\sqrt{\pi}} \cdot \left(\frac{2}{\pi}\right)^{n-2} \frac{\alpha^{n-1}}{n-1} \\
&= \frac{4}{\sqrt{\pi}(n-1)} \left(\frac{4}{\pi}\right)^{n-2} \cdot \frac{1}{2}(\alpha/2)^{n-1} \\
&\geq \frac{4}{\sqrt{\pi}(n-1)} \left(\frac{4}{\pi}\right)^{n-2} \cdot \frac{1}{2}(r/2)^{n-1} \\
&= \frac{\sqrt{\pi}}{n-1} \left(\frac{4}{\pi}\right)^{n-1} \cdot \frac{1}{2}(r/2)^{n-1}
\end{aligned}$$

It is easily shown that

$$\frac{\sqrt{\pi}}{n-1}\left(\frac{4}{\pi}\right)^{n-1} \geq 1$$

for all $n \geq 2$, thus proving the inequality. ■

Theorem 1.3. Let K be a symmetric polytope in $\mathbb{R}^n$ with $d(K, B_2^n) = d$. Then K has at least $\exp(n/2d^2)$ facets. On the other hand, for each n, there is a polytope with 4^n facets whose distance from the ball is at most 2.

Before proving the above theorem, let us reformulate what a symmetric polytope is in another way. Suppose you have a symmetric polytope K with m pairs of facets. Then it is basically the intersection of m slabs in $\mathbb{R}^n$ each of the form $\{x : |\langle x, v_i\rangle| \leq 1\}$ for some $v_i \in \mathbb{R}^n$. That is,

$$K = \{x : |\langle x, v_i\rangle| \leq 1 \text{ for } 1 \leq i \leq m\} \tag{1.3}$$

We can then consider a linear map from $K \to \mathbb{R}^m$ given by

$$T : x \mapsto (\langle x, v_1\rangle, \ldots, \langle x, v_m\rangle)$$

This maps $\mathbb{R}^n$ to a subspace of $\mathbb{R}^m$. By the formulation of K given in 1.3 , the intersection of this subspace with the unit cube is just the image of K under T! This is just an n-dimensional slice of $[-1,1]^m$. Even conversely, any n-dimensional slice of $[-1,1]^m$ is a convex body with at most m pairs of facets.

Proof. For the proof, let us write what it means for each v_i (following the above notation) if $B_2^n \subseteq K \subseteq dB_2^n$.

- The first inclusion just says that each v_i is of length at most 1 (otherwise, one could consider $v_i/|v_i|$, which would be in B_2^n but not in K).
- The latter says that if $|x| > d$, then there is some i for which $\langle x, v_i\rangle > 1$. That is, for any unit vector θ, there is some i such that
$$\langle \theta, v_i\rangle > \frac{1}{d}.$$

Since we want to minimize m while satisfying the above two conditions, we can clearly do no better than have $|v_i| = 1$ for each i. We want that every $\theta \in S^{n-1}$ is in one of the m $(1/d)$-caps about the (v_i).

- Obviously, to do this, we should attempt to estimate the area of a general ε-cap ($\varepsilon = 1/d$ here). Given Lemma 1.1, we get that
$$m \geq \frac{1}{\exp(-n\varepsilon^2/2)} = \exp\left(\frac{n}{2d^2}\right).$$
- To show that there exists a polytope with the given number of facets, it is enough to find $2 \cdot 4^{n-1}$ points $v_1, \ldots, v_m$ such that the caps of radius 1 centered at these points covers the sphere. Such a set is called a 1-*net*. Now, suppose we choose a set of points on the sphere such that any two of them are at least distance 1 apart. Such a set is called a 1-*separated set*.

  Note that the caps of radius 1/2 centered at each of the points in a 1-separated set are disjoint. Since the measure of a cap of radius 1/2 is at least $1/2 \cdot 4^{n-1}$ (by Lemma 1.2), the number of points in a 1-separated set is at most $2 \cdot 4^{n-1}$.
  It is then enough to choose a "maximal" 1-separated set (a 1-separated set S such that $S \cup x$ is not 1-separated for any $x \in S^{n-1}$) since it is then automatically a 1-net!

  Therefore, there is a 1-net with at most $2 \cdot 4^{n-1}$ points.

  The corresponding polytope then has twice as many, that is, 4^n facets (since each v_i corresponds to a pair of facets).

■

§2. More Convex Geometry!

2.1. Fritz John's Theorem

At the very beginning, we had mentioned that the distance of the cube $[-1,1]^n$ and the regular solid simplex are at distance at most $\sqrt{n}$ and n from the ball respectively. However, how would one go about proving that the distances are *exactly* $\sqrt{n}$ and n?

Fritz John's Theorem aids us in this pursuit.
He considered ellipsoids inside convex bodies. If (e_i) is an orthonormal basis of $\mathbb{R}^n$ and (α_i) are positive numbers, then the ellipsoid defined by

$$\left\{x : \sum_{i=1}^{n} \frac{\langle x, e_j \rangle^2}{\alpha_j^2} \leq 1\right\}$$

has volume equal to $v_n \prod_i \alpha_i$. The theorem states that there is a *unique* maximal ellipsoid contained in any convex body, and further, he characterized this ellipsoid! Also, if K is a symmetric convex body and $\mathcal{E}$ is its maximal ellipsoid, then $K \subseteq \sqrt{n}\mathcal{E}$!
We can then use this characterization combined with an affine transformation to prove that the distance between the cube and the ball is $\sqrt{n}$.
We state John's Theorem after performing the affine transformation, since it is easier to understand what's going on then. Roughly, it says that there should be several points of contact between the ball and the boundary of K.

Theorem 2.1 (Fritz John's Theorem)**.** Each convex body K contains a unique ellipsoid of maximal volume. This ellipsoid is B_2^n iff $B_2^n \subseteq K$ and for some m, there are unit vectors $(u_i)_1^m$ on the boundary of K and positive numbers $(c_i)_1^m$ such that

$$\sum_i c_i u_i = 0 \tag{2.1}$$

and for each $x \in \mathbb{R}^n$,

$$\sum_i c_i \langle x, u_i \rangle^2 = |x|^2. \tag{2.2}$$

Before proving the theorem, let us discuss some of its implications.

The first condition essentially says that the (u_i) are not all on one side of the body[8]. Intuitively, this makes sense because if the points were concentrated towards one side of the body, then we could move the ball a little bit in the opposite direction and then expand it a little.
The second says that the (u_i) are something like an orthonormal basis, in that we can resolve the norm as a weighted sum of squares of inner products.
On the other hand, the second point Equation (2.2) is equivalent to saying that for all $x \in \mathbb{R}^n$,

$$x = \sum_i c_i \langle x, u_i \rangle u_i.$$

This ensures that the points do not lie close to a (proper) subspace of $\mathbb{R}^n$. This makes sense intuitively as well since if they did, we could contract the ellipsoid a bit in this direction and expand it orthogonally.

Equation (2.2) is written more compactly as

$$\sum_i c_i u_i \otimes u_i = I_n. \tag{2.3}$$

[8] note that it says more than simple linear independence since the c_i are positive.

Here, $u_i \otimes u_i$ represents the (rank-1) orthogonal projection onto the span of u_i, the map given by $x \mapsto \langle x, u_i\rangle u_i$. Note that this map is just equal to $u_i u_i^\top$. This implies that the trace of this projection is equal to $|u_i|^2 = 1$[9]. Equating the traces of either side of Equation (2.3), we get

$$\sum_i c_i = n. \tag{2.4}$$

Finally, note that if K is a *symmetric* convex body, then the first condition is obsolete since we can just find any (u_i) satisfying the second condition and replacing it with $+u_i$ and $-u_i$ with each having half the original weight.

Let us now consider a couple of examples to better understand the implications of the theorem.

- For the cube $[0,1]^n$, the maximal ellipsoid is B_2^n as one would expect. The points of contact are the standard basis vectors $(e_i)_1^n$ of $\mathbb{R}^n$ with their negatives, and they do indeed satisfy

$$\sum_i e_i \otimes e_i = I_n.$$

- A slightly more nuanced example is that of the regular solid simplex. Unfortunately, there is no simple or standard way to represent the n-dimensional simplex in n dimensions. It is, however, far natural to represent it in $\mathbb{R}^{n+1}$ by considering the convex hull of the $n+1$ standard basis vectors $(e_i)_1^{n+1}$. We also scale it up by a factor of $\sqrt{n(n+1)}$ (so that the ball contained is B_2^n) such that the $n+1$ points (p_i) we take the convex hull of to get the simplex are given by
$$p_i = \sqrt{n(n+1)} e_i.$$
This simplex can be parametrized as
$$K = \left\{ x \in \mathbb{R}^{n+1} : \sum_{i=1}^{n+1} x_i = \sqrt{n(n+1)} \text{ and } x_i \geq 0 \text{ for each } i \right\}.$$
Similar to the cube, the contact points of the maximal ellipsoid are the centers of each of the facets. More precisely, these $n+1$ endpoints are given by
$$u_i = \frac{\sqrt{n(n+1)}}{n} \left(\sum_{j=1}^{n+1} e_j - e_i \right).$$
Affinely shifting the hyperplane such that it passes through the origin (making $x_0 = \frac{\sqrt{n(n+1)}}{n+1} \sum_i e_i$ the new origin) and setting the constants c_i as $c = \frac{n}{n+1}$ for each i, for any x in the (unshifted) body,
$$\begin{aligned} \sum_i n(n+1) c_i \langle x - x_0, u_i - x_0 \rangle^2 &= c \sum_i \left(\sum_{j=1}^{n+1} \frac{x_j}{n} - \frac{x_i}{n} - \frac{2}{n+1} + \frac{1}{n+1} \right)^2 \quad \left(\langle x, x_0 \rangle = \langle x_0, u_i \rangle = \langle x_0, x_0 \rangle = \frac{1}{n+1} \right) \\ &= n^2 \sum_i \left(\frac{x_i}{n} - \frac{1}{n(n+1)} \right)^2 \\ &= |x - x_0|^2. \end{aligned}$$
It is easily shown that $\sum_i c_i (u_i - x_0) = 0$ and that each $(u_i - x_0)$ is of magnitude unity, thus proving that the ball touching the centers of the facets (which is an affine shift of B_2^n) is the maximal ellipsoid inside the n-dimensional simplex.

[9] We could have also got this more directly by using the fact that the trace of (a matrix in some basis corresponding to) a linear transformation is the sum of its eigenvalues. For an orthogonal projection, this is just equal to the rank of the target space (which is 1 in this case)

Now, let us prove one of the claims that we made at the beginning of the section.

Theorem 2.2. Suppose that K is a symmetric convex body and B_2^n is the maximal ellipsoid contained in K. Then $K \subseteq \sqrt{n} B_2^n$.
Suppose that K is a convex body and B_2^n is the maximal ellipsoid contained in K. Then $K \subseteq n B_2^n$.

Note that while we have stated the above assuming that B_2^n is the maximal ellipsoid, any convex body in general can be brought to this form by performing an affine shift.

Proof.

- Let x be an arbitrary point in the symmetric body K. Our aim is to show that $|x| \leq \sqrt{n}$. Let $(u_i)_1^m$ be the points as described in Fritz John's Theorem. We may assume that if u is in this set, then so is $-u$.
Now, note that for any i, the tangent plane to K at u_i must coincide with the tangent plane to B_2^n at u_i (otherwise, we would get a contradiction to $B_2^n \subseteq K$). Then, since K is convex, any point in the body must be in the half-space defined by this tangent that contains 0 – this means that $\langle x, u_i \rangle \leq 1$ for each i.
Then, for each i, we have $\langle x, u_i \rangle \leq 1$ *and* $\langle x, -u_i \rangle \leq 1$ (since we've assumed that if u is in the (u_i), then so is $-u$). That is, $|\langle x, u_i \rangle| \leq 1$ for each i.
Using the above along with Equation (2.2) and Equation (2.4), we now have
$$|x|^2 = \sum_i c_i \langle x, u_i \rangle^2 \leq \sum_i c_i = n,$$
which is exactly what we set out to prove!

- Let x be an arbitary point in the convex body. From the first part, we already have that $\langle x, u_i \rangle \leq 1$ for each i. We also have $\langle x, u_i \rangle \geq -|x|$ (since $|u_i| = 1$). Then,
$$\begin{aligned}
0 &\leq \sum_i c_i \left(1 - \langle x, u_i \rangle\right)\left(|x| + \langle x, u_i \rangle\right) \\
\implies \sum_i c_i \langle x, u_i \rangle^2 &\leq \sum_i c_i |x| + (1 - |x|) \left\langle x, \sum_i c_i u_i \right\rangle \\
\implies \sum_i c_i \langle x, u_i \rangle^2 &\leq \sum_i c_i |x| && \left(\text{since } \sum_i c_i u_i = 0\right) \\
\implies |x| &\leq n. && \text{(by Equation (2.2) and Equation (2.4))}
\end{aligned}$$

∎

Let us now prove Fritz John's Theorem.

Lemma 2.3 (Fritz John's Theorem Pt. 1)**.** Let K be a convex body and for some integer m, let there be unit vectors $(u_i)_1^m$ in ∂K and positive reals $(c_i)_1^m$ satisfying Equation (2.1) and Equation (2.2). Then B_2^n is the unique maximal ellipsoid contained in K.

Proof. Let
$$\mathcal{E} = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^n \frac{\langle x, e_j \rangle^2}{\alpha_j^2} \leq 1 \right\}$$
be an ellipsoid in K for some orthonormal basis (e_j) and positive (α_j). We must show that

- $\prod_j \alpha_j \leq 1$ (this implies that B_2^n is a maximal ellipsoid) and

- if $\prod_j \alpha_j = 1$, then for every j, $\alpha_j = 1$ (this implies that B_2^n is *the* maximal ellipsoid).

Now, consider the *dual* of $\mathcal{E}$ given by

$$\mathcal{E}^* = \left\{ y \in \mathbb{R}^n : \sum_{i=1}^n \alpha_j^2 \langle x, e_j \rangle^2 \leq 1 \right\}$$

Observe that we can more concisely describe $\mathcal{E}^*$ as $\{y \in \mathbb{R}^n : \langle y, x \rangle \leq 1 \text{ for all } x \in \mathcal{E}\}$ (Why? Try using the Cauchy-Schwartz inequality)[10].
Now, note that since $\mathcal{E} \subseteq K$, for any $x \in \mathcal{E}$ and any i, $\langle x, u_i \rangle \leq 1$ (as proved in the first part of Theorem 2.2). This implies that for every i, $u_i \in \mathcal{E}^*$! We then have

$$\begin{aligned}
\sum_j \alpha_j^2 &= \sum_j \alpha_j^2 |e_j|^2 \\
&= \sum_j \alpha_j^2 \sum_i c_i \langle u_i, e_j \rangle^2 \\
&= \sum_i \left(c_i \sum_j \alpha_j^2 \langle u_i, e_j \rangle^2 \right) \\
&\leq \sum_i c_i = n && \text{(since } u_i \in \mathcal{E}^*\text{)}
\end{aligned}$$

Then, using the AM-GM inequality,

$$\prod_j \alpha_j \leq \left(\frac{1}{n} \sum_j \alpha_j^2 \right)^{n/2} \leq 1.$$

This proves the first part. The second part follows directly as well, since if equality holds in the above equation, then every α_i^2 must be the same (the condition for equality to hold in the AM-GM inequality). ■

This is the easier of the two directions in Fritz John's Theorem. We now prove the harder.

Lemma 2.4 (Separation Theorem)**.** *Let X and Y be two disjoint closed convex bodies in $\mathbb{R}^n$ with at least one of them bounded. Then there exists some $v \in \mathbb{R}^n$ such that for all $x \in X$, $\langle x, v \rangle < b$ and for all $y \in Y$, $\langle y, v \rangle > b$.*

We leave the proof of the above to the reader.

Lemma 2.5 (Fritz John's Theorem Pt. 2)**.** *Let K be a convex body such that B_2^n is a maximal ellipsoid contained in K. Then, for some integer m, there exist unit vectors $(u_i)_1^m$ in ∂K and positive reals $(c_i)_1^m$ satisfying Equation (2.1) and Equation (2.3).*

Proof. We want to show that there exist unit vectors (u_i) in ∂K and positive constants (c_i) such that

$$\frac{1}{n} I_n = \sum_i \left(\frac{c_i}{n} \right) (u_i \otimes u_i)$$

Since $\sum_i c_i = n$, we essentially aim to show that $\frac{1}{n} I_n$ is in the convex hull of the $(u_i \otimes u_i)$ (in the space of matrices). To this end, define

$$T = \operatorname{Conv}\left(\{ u \otimes u : u \text{ is a unit vector in } \partial K \} \right).$$

We refer to such u as "contact points". We want to show that $\frac{1}{n} I_n \in T$. Suppose that it is not (we shall finally show that B_2^n is not a maximal ellipsoid). Then Lemma 2.4 implies that there exists a matrix $H = (h_{i,j})$ such that the linear map φ from the set of matrices to $\mathbb{R}$ defined by

$$(a_{i,j}) \mapsto \sum_{i,j} h_{i,j} a_{i,j}$$

[10] Interested readers can go through this source as well.

satisfies

$$\varphi\left(\frac{I_n}{n}\right) < \varphi(u \otimes u)$$

for all contact points u. Now, since the matrices on either side are symmetric, we may assume that H is symmetric as well (Why?). And since the matrices on either side have trace equal to 1, adding any constant to the diagonal elements of H leaves the inequality unchanged. Therefore, we may suppose that the trace of H is 0. But this just says that $\varphi(I_n) = 0$!

Therefore, we have essentially found a matrix H such that for any contact point u,

$$u^\top H u > 0. \quad \text{(check that } \varphi(u \otimes u) = u^\top H u)$$

Now, for $\delta > 0$, consider the ellipsoid defined by

$$\mathcal{E}_\delta = \left\{x \in \mathbb{R}^n : x^\top (I_n + \delta H) x \leq 1\right\}.$$

We claim that $\mathcal{E}_\delta$ is strictly inside K for sufficiently small δ. Note that for each contact point u,

$$u^\top (I_n + \delta H) u = 1 + \delta\left(u^\top H u\right) > 1$$

so no contact point (of B_2^n) is in $\mathcal{E}_\delta$. For each contact point u, consider a neighbourhood n_u such that for all $x \in n_u$, $x^\top (I_n + \delta H)x > 0$ – we know that such a neighbourhood exists due to the continuity of $x \mapsto x^\top (I_n + \delta H)x$. Let N be the union of all these neighbourhoods.

We now want to show that for any $x \in \partial K \setminus N$, $x^\top (I_n + \delta H)x > 1$. To this end, let $\lambda_{\min}$ be the minimum eigenvalue of H. For any $x \in \partial K \setminus N$, $x^\top H x \geq \lambda_{\min} |x|^2$. That is, for all such x,

$$x^\top (I_n + \delta H)x \geq (1 + \delta\lambda_{\min})|x|^2.$$

Now, observe that $\inf_{x \in \partial K \setminus N} |x|^2 > 1$[11]. We may also assume that $\lambda_{\min} < 0$, since the claim holds trivially otherwise (we have $|x|^2 > 1$). Then, we may set δ as a positive real which is less than $\frac{1}{|\lambda_{\min}|}\left(1 - \frac{1}{\inf_{x \in \partial K \setminus N} |x|^2}\right)$. Then for all $x \in \partial K \setminus N$,

$$(1 + \delta\lambda_{\min})|x|^2 > |x|^2 \left(1 - \left(1 - \frac{1}{\inf_{y \in \partial K \setminus N} |y|^2}\right)\right) \geq 1$$

Therefore, $\mathcal{E}_\delta$ does not intersect ∂K and is *strictly* inside K for sufficiently small δ!

Now, we claim that $\mathcal{E}_\delta$ has volume at least equal to that of B_2^n. Indeed, its volume is given by $v_n / \prod \lambda_i$, where (λ_i) are the eigenvalues of $(I_n + \delta H)$. Since the sum of the eigenvalues is equal to the trace of $I_n + \delta H$, which is n, we can use the AM-GM inequality to get

$$\prod_i \lambda_i \leq \left(\frac{1}{n}\sum_i \lambda_i\right)^n = 1,$$

which is exactly what we want, because equality holds iff the eigenvalues are all 1, that is, the ellipsoid is B_2^n (so this leads to a contradiction). ■

Note that we can concatenate the proofs of Lemma 2.5 and Lemma 2.3 to show that *a* maximal ellipsoid is *the* maximal ellipsoid (contained in a convex body).

There is an analogue of Fritz John's Theorem that characterizes the minimal ellipsoid that contains a given body – this is near-direct from the notion of duality that we used in the proof of Lemma 2.3. So for example, it follows form this analogue that the minimal ellipsoid that contains $[-1,1]^n$ is the ball of radius $\sqrt{n}$. This also enables us to say that $d([-1,1]^n, B_2^n)$ is *exactly* equal to $\sqrt{n}$.

[11] if it was equal to 1, then for any $\varepsilon > 0$, we would be able to find an x such that $|x|^2 < 1 + \varepsilon$ (we trivially have that $|x|^2 \geq 1$). However, this is not possible because ∂K is compact, we have removed a neighbourhood around each contact point u, and contact points are the only points in ∂K which have norm 1.

There are various extensions of this result. Recall how towards the beginning of these notes we had mentioned how a general convex body K is essentially a unit ball under some norm. Fritz John's Theorem essentially describes linear maps from the Euclidean space to a normed space (under which the unit ball is K) that have largest determinant under the constraing that the Euclidean ball is mapped into K. There is a more general theory that (attempts to) solve this problem under different constraints.