Lecture 6

John ellipsoids

- John theorem
- Dikkin ellipsoid

Banach-Mazur distance between two non-empty compact convex bodies

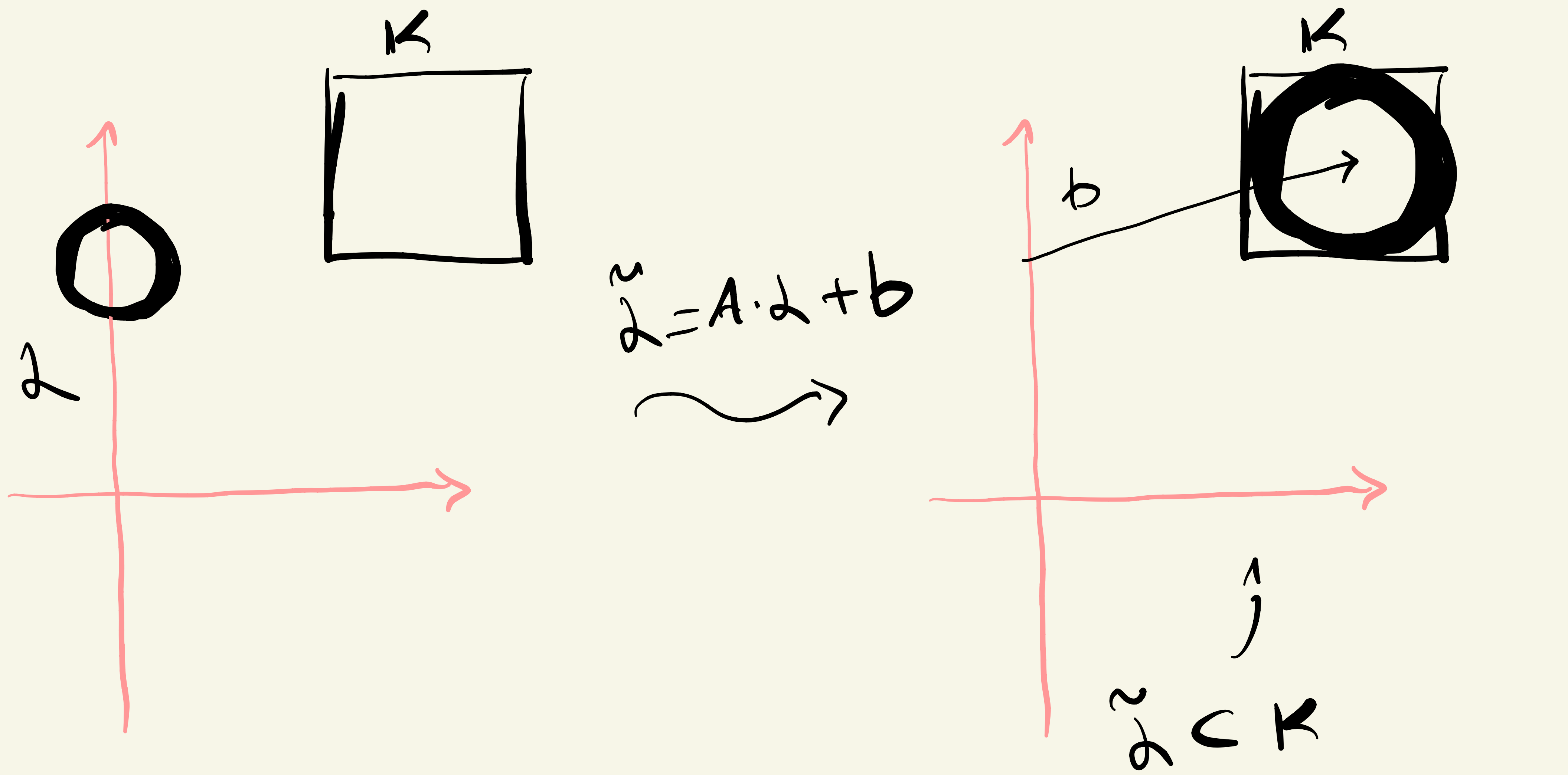
the term distance is used to denote how "similar" are the shapes of two convex bodies

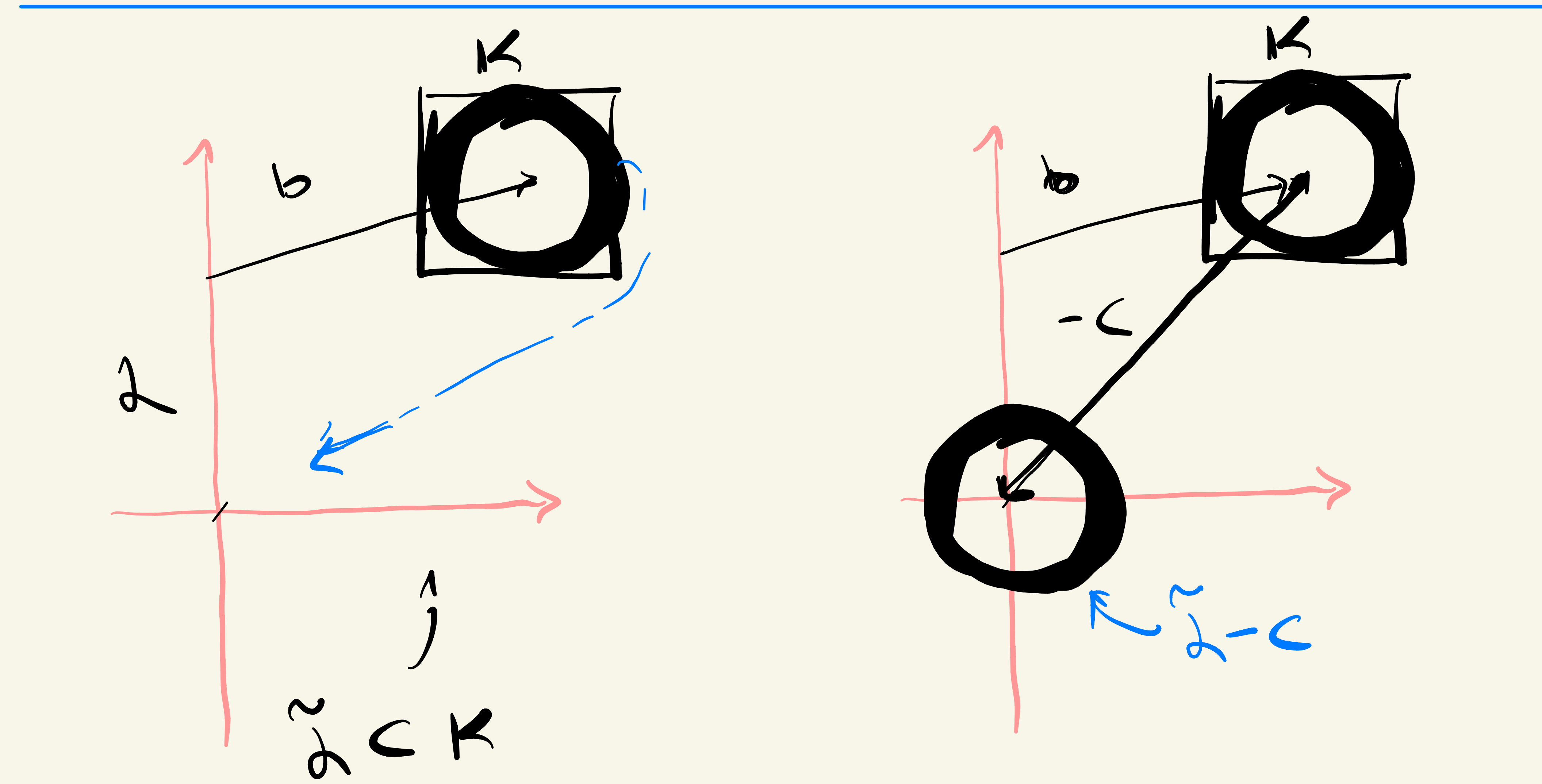
we will assume that both convex bodies "live" in the same # of dimensions
+
the affine transformations considered are invertible

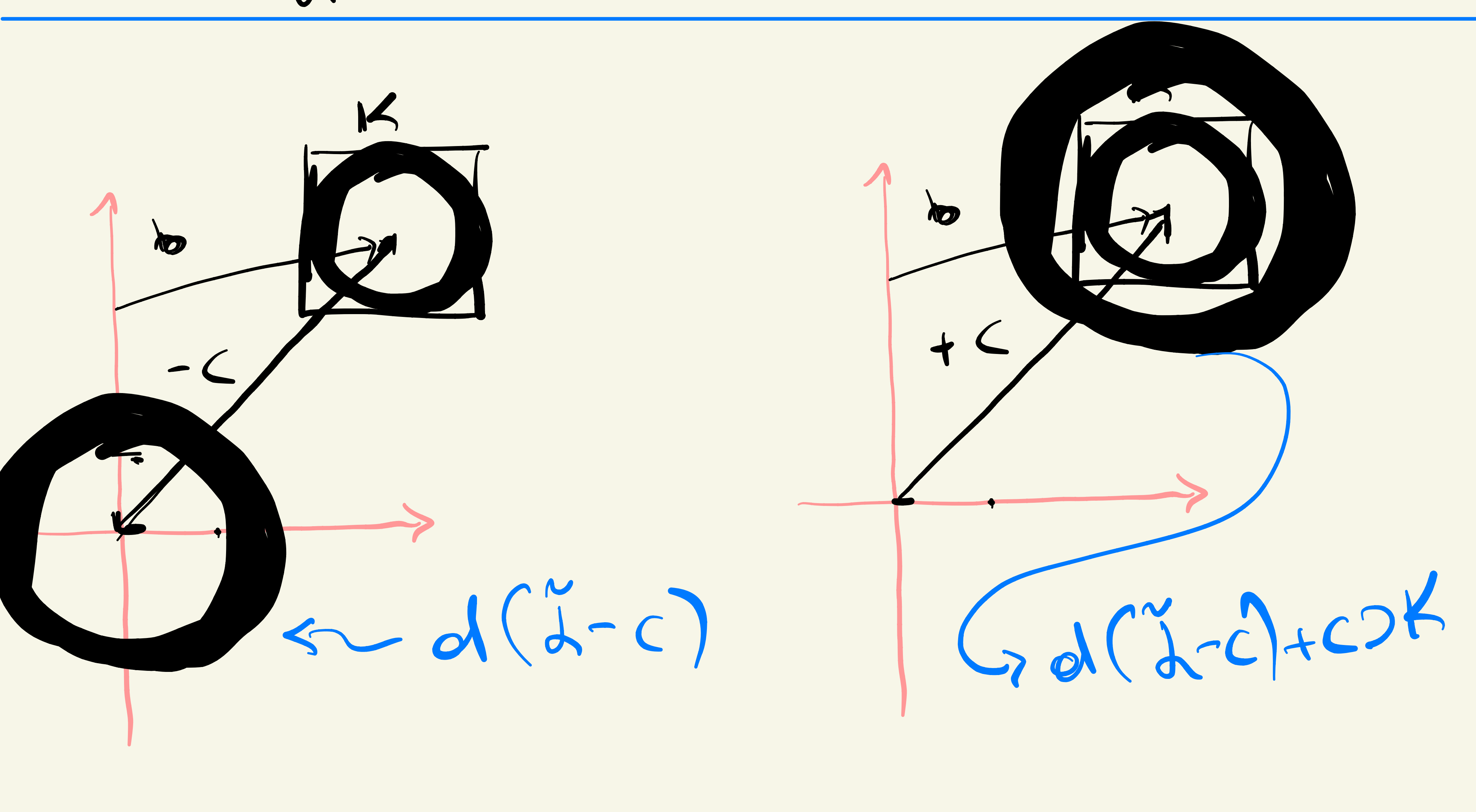
$$d(K,L) = \arg\min \{ d > 0 \text{ s.t. } \tilde{L} \subset K \subset d(\tilde{L} - c) + c$$

where $\tilde{L}$ is an affine transformation of L an $c \in \tilde{L}$ }

linear + offset


K
λ
$\tilde{\lambda} = A \cdot \lambda + b$
b
K
$\tilde{\lambda} \subset K$


K
b
λ
$\tilde{\lambda} \subset K$
K
b
$-c$
$\tilde{\lambda} - c$


K
b
$-c$
$d(\tilde{\lambda} - c)$
b
$+c$
$d(\tilde{\lambda} - c) + c \supset K$

John's theorem

$x \in C \Rightarrow -x \in C$

For any symmetric convex body C there is an ellipsoid $\mathcal{E}$ such that $\mathcal{E} \subseteq C \subseteq \sqrt{n}\,\mathcal{E}$

For any (non-symmetric) convex body C there is an ellipsoid $\mathcal{E}$ s.t.

$$\mathcal{E} \subseteq C \subseteq n(\mathcal{E} - \text{center}) + \text{center}$$

center of the ellipsoid $\mathcal{E}$

because the optimization problem is simpler

We will prove John's theorem for the case where the convex body C is a symmetric polytope.

(this should be intuitively enough as any convex body can be written as the intersection of halfspaces) proof

sketch → (1) try to find the Ellipsoid that is contained in P with maximum volume

(2) prove that $\sqrt{n}\cdot\mathcal{E}$ contains C (using the optimality of $\mathcal{E}$

$$C = \{ x \mid a_i^T x \le b_i \quad i \in [m] \}$$

(since $x \in C \rightsquigarrow -x \in C \Rightarrow a_i^T x \le b_i$ are $(-a_i^T) x \le b_i$

both hyperplanes of the polytope

---

from lecture 1 we know that a symmetric ellipsoid can be represented as

$$\mathcal{E} = \{ x \mid x^T Q^{-1} x \le 1 \} \text{ where } Q \succ 0$$

$$\Rightarrow \mathcal{E} = \{ x \mid \| \underbrace{Q^{-1/2} \cdot x}_{u} \|_2^2 \le 1 \} \overset{u = Q^{-1/2} x}{\equiv} \{ Q^{1/2} u \mid \|u\|_2^2 \le 1 \}$$

$$\mathrm{vol}(\mathcal{E}) = \det(Q^{1/2}) \, \mathrm{vol}(B_2^n)$$

$\Rightarrow$ maximizing the volume of $\mathcal{E}$ is the same as maximizing $\det(Q^{1/2})$

We know what we should aim to maximize!

how do we describe an ellipsoid which is inscribed into a symmetric polytope?

$$\sup\{a_i^T x \mid x \in \mathcal{E}\} =$$

easier to think about the supremum

$$= \sup\{a_i^T Q^{1/2} u \mid \|u\| \leq 1\},$$

$$\overset{u = \frac{Q^{1/2}a_i}{\|Q^{1/2}a_i\|}}{=\!=} \frac{a_i^T Q^{1/2} Q^{1/2} a_i}{\|Q^{1/2} a_i\|} = \|Q^{1/2} a_i\|$$

therefore the constraints

$$a_i^T x \leq b_i \quad \forall x \in \mathcal{E} \iff \|Q^{1/2} a_i\|_2 \leq b_i$$

Now that we know both what we should optimize and a nice form for the constraints. We are ready to formulate the optimization problem.

$\Longrightarrow$

$$\max \ \det(Q^{1/2})$$

$$\text{s.t.} \quad a_i^T x \leq b_i \ \forall x \in \mathcal{E}, \forall i$$

$$\mathcal{E} = \{x \mid x^T Q^{-1} x \leq 1\}$$

$$Q \succ 0$$

$\iff$

$$\max \ \det(Q^{1/2})$$

$$\|Q^{1/2} a_i\|_2 \leq b_i \ \forall i$$

$$Q \succ 0$$

write it as a convex program

$$\min \ -\log\det(Q)$$

convex function

$$\text{s.t.} \quad \frac{a_i^T Q a_i}{b_i^2} \leq 1 \quad \forall i$$

linear in Q constraint

$$Q \succ 0$$

convex set

we will try to find a closed form solution using the KKT conditions

$$L(Q,\lambda) = -\log\det(Q) + \sum_i \lambda_i \left(\frac{a_i^T Q a_i}{b_i^2} - 1 \right)$$

primal feasibility: $\frac{a_i^T Q a_i}{b_i^2} \leq 1 \ \forall i$

dual feasibility: $\lambda_i \geq 0 \ \forall i$

complementary slackness:

$$\lambda_i \left(\frac{a_i^T Q a_i}{b_i^2} - 1 \right) = 0 \quad \forall i$$

Lagrangian optimality:

$$\nabla_Q \mathcal{L}(Q,\lambda) = 0 \iff -(Q^{-1})^T + \sum_i \frac{\lambda_i a_i a_i^T}{b_i^2} = 0$$

why? (1) $\nabla_Q (a_i^T Q a_i) = \nabla_Q (\underbrace{\langle Q, a_i a_i^T \rangle}) =$
(calculations)

$$= a_i a_i^T$$

flatten the matrices and compute the inner product

(2) $\nabla_Q (\log\det(Q)) = \frac{1}{\det(Q)} \nabla_Q \det(Q)$

M_{ij} = determinant of Q removing i-th row and jth coloumn

$$\det(Q) = \sum_j (-1)^{i+j} M_{ij} \cdot Q_{ij} \quad \forall i$$

$$\frac{\partial \det(Q)}{\partial Q_{ij}} = (-1)^{i+j} M_{ij} = (\mathrm{adj}(Q))_{ji}$$

by definition

$\implies$

$$\frac{\partial \log\det(Q)}{\partial Q_{ij}} = \frac{(\mathrm{adj}(Q))_{ji}}{\det(Q)}$$

and

$$Q \cdot \mathrm{adj}(Q) = \det(Q) \cdot I \Rightarrow$$

$$\Rightarrow \frac{\mathrm{adj}(Q)}{\det(Q)} = Q^{-1} \Rightarrow \frac{(\mathrm{adj}(Q))^T}{\det(Q)} = (Q^{-1})^T$$

$$\text{So } -(Q^{-1})^T + \sum_i \frac{\lambda_i a_i a_i^T}{b_i^2} = 0 \Rightarrow$$

$$\Rightarrow Q^{-1} = \sum_i \lambda_i \frac{a_i a_i^T}{b_i^2}$$

now we know that a pair of optimal solutions Q, λ should satisfy

we want to prove that $\sqrt{n} \cdot \mathcal{E} \supseteq C$

$$\Rightarrow x \in C \Rightarrow x \in \sqrt{n}\,\mathcal{E} \Rightarrow x^T Q^{-1} x \leq n$$

$$\Longrightarrow$$

$x \in C$

C is a symmetric polytope

$$\begin{cases} a_i^T x \le b_i \\ -a_i^T x \le b_i \end{cases} \qquad \left(a_i^T x \in [-b_i, b_i] \right.$$

$$\Longrightarrow \quad (a_i^T x)^2 \le b_i^2 \Rightarrow$$

$$\Rightarrow \quad \boxed{\frac{a_i^T x \cdot x^T a_i}{b_i^2} \le 1}$$

$$\downarrow$$

$$\frac{x^T a_i a_i^T x}{b_i^2} \le 1$$

$$x^T Q^{-1} x = x^T \left(\sum \lambda_i \frac{a_i a_i^T}{b_i^2} \right) x =$$

this is what I want to bound

$$= \sum_i \lambda_i \frac{x^T a_i a_i^T x}{b_i^2} \overset{\lambda_i \ge 0}{\le} \sum \lambda_i =$$

complementary slackness:

$$\lambda_i \left(\frac{a_i^T Q a_i}{b_i^2} - 1 \right) = 0 \quad \forall i$$

$$= \sum \lambda_i \frac{a_i^T Q a_i}{b_i^2} =$$

$$= \left\langle Q, \sum \lambda_i \frac{a_i a_i^T}{b_i^2} \right\rangle = \langle Q, Q^{-1} \rangle$$

$$= Tr(Q \cdot Q^{-1}) = n \quad \blacksquare$$

Discussion

① The non-symmetric case is harder because also the center of the ellipsoid is a variable

$\rightsquigarrow \mathcal{E} = \{x \mid (x - \text{center})^T Q^{-1} (x - \text{center}) \leq 1\}$

② John's theorem is tight and the hard example for the symmetric case is the cube $[-1,1]^n$ and for the non symmetric case in the simplex.

$x \to Q^{-1/2} x$

Geometric intuition

via the inverse affine transformation we can transform $\mathcal{E} \xrightarrow{Q^{-1/2}} B_2^2$ and

$C \xrightarrow{Q^{-1/2}} C'$ (to a different convex body)

let's assume again that $C = \{x \mid a_i^T x \leq b_i \ \forall i\}$

and $C' = \{x \mid a_i' x \leq b, \ a_i' = Q^{-1/2} a_i, \ \forall i\}$

now we know that B_2^2 solves the maximum volume inscribed ellipsoid problem for C'. Let's try to give a geometric interpretation of the KKT conditions in that case

$$\left(B_2^2 = \{ x \mid x^T x \leq 1 \} = \{ x \mid x^T I x \leq 1 \} \right)$$

complementary slackness $\quad \lambda_i \left(\frac{a_i'^T a_i'}{b_i^2} - 1 \right) = 0 \Rightarrow$

$$\Rightarrow \lambda_i > 0 \Rightarrow a_i'^T \left(\frac{a_i'}{b_i} \right) = b_i$$

thus $\quad a_i'^T \left(\frac{a_i'}{b_i} \right) = b_i \Rightarrow \frac{a_i'}{b_i}$ is on the boundary of K'

$$\frac{a_i'^T}{b_i} \cdot \frac{a_i'}{b_i} = 1 \Rightarrow \text{on the boundary of the unit ball}$$

when $\lambda_i > 0$ $\frac{a_i'^T}{b_i}$ lies both on the boundary of the polytope and the inscribed ellipsoid (in this case the ball)

Lagrangian optimality

$$I = \sum_{i=1}^{m} \lambda_i \frac{a_i a_i^T}{b_i^2} =$$

$\frac{a_i}{b_i} = u_i \quad i: \lambda_i > 0$

contact points

$$= \sum_{i:\lambda_i>0} \lambda_i \frac{a_i}{b_i} \cdot \frac{a_i^T}{b_i} =$$

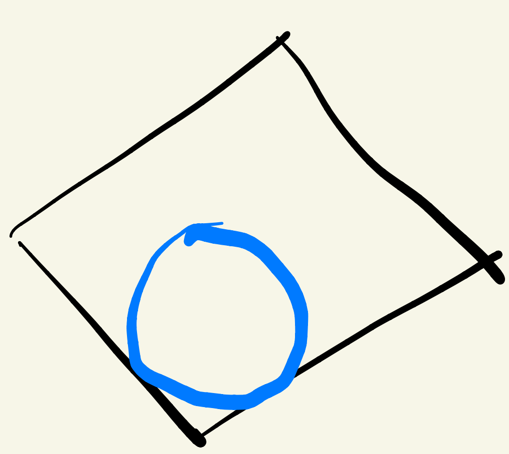
$$= \sum \lambda_i \cdot u_i \cdot u_i^T$$

The contact points ($\times \sqrt{\lambda_i}$) behave like an orthonormal basis

$$\Rightarrow \forall x \quad \left(\sum \lambda_i u_i u_i^T\right) x = I x = x \Rightarrow$$

$$\Rightarrow \sum \lambda_i \langle u_i, x \rangle u_i = x$$

guarantees that the contact points are spread in all directions



not maximum volume inscribed ellipsoid

Nice observation:

given a norm $\|\cdot\|$ function we can associate a ball $B = \{x \mid \|x\| \leq 1\}$ (which of course is a convex set)

Because of John's theorem we know that

$\exists\ \mathcal{E}$ (ellipsoid) $= \{x \mid x^T Q x \leq 1\}$ s.t

$$\mathcal{E}/\sqrt{n} \subseteq B \subseteq \mathcal{E} \qquad \left(\frac{\mathcal{E}}{\sqrt{n}} = \{x \mid x^T Q x \leq 1/n\}\right)$$

$\Rightarrow \|x\|_Q \leq \|x\| \leq \sqrt{n}\,\|x\|_Q$

fix an x

let $a > 0$ s.t $x = a x'$, $\|x'\| = 1$

$x' \in B \Rightarrow x' \in \mathcal{E} \Rightarrow \|x'\|_Q \leq 1$ $\Big\{$ $\|x'\| \geq \|x'\|_Q$, $x' = ax$

and $\|x\| \geq \|x\|_Q$

let $a' > 0$ s.t $x = a x'$, $\|x'\|_{\sqrt{n}Q} = 1$

$x' \in \frac{\mathcal{E}}{\sqrt{n}} \Rightarrow x' \in B \Rightarrow \|x'\| \leq 1$ $\Big\{$

$\sqrt{n}\|x'\|_Q = \|x'\|_{\sqrt{n}Q} \geq \|x'\| \Rightarrow$

$\Rightarrow \sqrt{n}\,\|x'\|_Q \geq \|x'\|$

Dikin ellipsoid

an ellipsoid that approximates a polytope using the logarithmic barrier function

$$C = \{x \mid a_i^T x \leq b_i \ \forall i \in [m]\}$$

the analytic center of is defined

as $\min \underbrace{-\sum_{i=1}^{m} \log(b_i - a_i^T x)}$

the more you get close to a constraint => the more I penalize you

$\Longrightarrow$

unconstrained optimization

$$\nabla_x \left(- \sum \log(b_i - a_i^T x) \right) = 0$$

$$\Rightarrow \sum_i \frac{a_i}{b_i - a_i^T x} = 0$$

analytic center

x_{ac} is the point that satisfies

$$\sum_i \frac{a_i}{b_i - a_i^T x_{ac}} = 0$$

Dikin's ellipsoid = the Hessian of the logarithmic barrier function around x_{ac}

$$H = \nabla_x \left(\sum_i \frac{a_i}{b_i - a_i^T x} \right) = \sum_i \frac{a_i a_i^T}{(b_i - a_i^T x)^2}$$

$$\Longrightarrow$$

worse than Jordan's ellipsoid (since $m > n$)

Theorem

Let

$$\mathcal{E}_{inner} = \{ x \mid (x - x_{ac})^T H (x - x_{ac}) \le 1 \}$$

$$\mathcal{E}_{outer} = \{ x \mid (x - x_{ac})^T H (x - x_{ac}) \le m(m-1) \}$$

then

$$\mathcal{E}_{inner} \subseteq C \subseteq \mathcal{E}_{outer} \subseteq m \, \mathcal{E}_{inner}$$

Proof

easy $\mathcal{E}_{inner} \subseteq C$

$$x \in \mathcal{E}_{inner} \Rightarrow \sum_i \frac{((x - x_{ac})^T a_i)^2}{(b - a_i^T x_{ac})^2} \le 1 \Rightarrow$$

every element of the sum ≤ 1

$$\Rightarrow ((x - x_{ac})^T a_i)^2 \le (b - a_i^T x_{ac})^2 \Rightarrow$$

$$\Rightarrow (x - \cancel{x_{ac}})^T a_i \le b - \cancel{a_i^T x_{ac}} \Rightarrow$$

$$\Rightarrow a_i^T x \le b \;\; \forall i \Rightarrow x \in C$$

$$\Longrightarrow$$

$C \subseteq \mathcal{E}_{outer}$

$$\boxed{\sum_i \frac{a_i}{b_i - a_i^T x_{ac}} = 0}$$

$x \in C \Rightarrow a_i^T x \leq b_i \ \forall i$

let's see what we can say about $(x - x_{ac})^T H (x - x_{ac}) =$

$$= \sum_i \frac{\left((x - x_{ac})^T \cdot a_i\right)^2}{\underbrace{(b_i - a_i^T x_{ac})}_{d_i}^2} =$$

$$= \sum_i \frac{\left((x - x_{ac})^T a_i - d_i\right)^2 + d_i^2 - 2 d_i \left[(x - x_{ac})^T a_i - d_i\right]}{d_i^2} =$$

$$= \sum_i \frac{\left((x - x_{ac})^T a_i - d_i\right)^2}{d_i^2} - m - 2 \cancelto{0}{\sum_i \frac{(x - x_{ac})^T a_i}{d_i}}$$

$$= \sum_i \frac{(b_i - a_i^T x)^2}{(b_i - a_i^T x_{ac})^2} - m \leq$$

$\sum y_i^2 \leq (\sum y_i)^2$ when $y_i > 0$ $\forall i$

$$\leq \left(\sum_i \frac{b_i - a_i^T x}{b_i - a_i^T x_{ac}}\right)^2 - m =$$

$$\left(\sum_i \frac{b_i - a_i^T x}{b_i - a_i^T x_{ac}}\right)^2 - m =$$

$$= \left(\sum_i \frac{b_i - a_i^T x_{ac}}{b_i - a_i^T x_{ac}} - (x + x_{ac}) \cancelto{0}{\sum_i \frac{a_i^T}{b_i - a_i^T x_{ac}}}\right)^2 - m =$$

$$= m^2 - m$$

if C is symmetric
then we have
$m^2 - m \rightsquigarrow \sqrt{m}$