Example: Min Volume Ellipsoid (Löwner-John ellipsoid) ①

Goal: Given convex set C,
find min volume ellipsoid containing C.
(We had C = convex hull of k points in class.)

Parameterize ellipsoid $\mathcal{E} = \{v \mid \|Av+b\|_2 \le 1\}$
i.e. inverse image of Euclidean unit ball under an affine mapping.
Note we may assume $A \in S^n_{++}$,
so $\log \mathrm{Vol}(\mathcal{E}) \propto -\log|A|$ (as in class)

So we want to solve:

$$\left\{ \begin{array}{l} \min\limits_{A,b} \; -\log|A| \quad \text{s.t.} \quad \sup\limits_{v\in C} \|Av+b\|_2 \le 1 \\ \\ \text{(Note implicit constraint } A \succ 0.) \end{array} \right\}$$

This leads to a problem without the semidefinite constraints used in class when $C = \mathrm{conv}\{x_1, \ldots, x_k\}$:

$$\left\{ \begin{array}{l} \min\limits_{A,b} \; -\log|A| \quad \text{s.t.} \quad \|Ax_i+b\|_2^2 \le 1, \; i=1,\ldots,k \\ \\ \text{since we can replace } \|\cdot\|_2 \le 1 \text{ with } \|\cdot\|_2^2 \le 1 \end{array} \right\}$$

Now, we'll look at the efficiency of the Löwner-John ellipsoidal approx. of sets C.
Let $\mathcal{E}$ be L-J ellipsoid of convex set $C \subseteq \mathbb{R}^n$ where C is bounded with a nonempty interior.
Let x_0 be center of C. If we shrink $\mathcal{E}$ by a factor of n around x_0, the new ellipsoid lies within C: $x_0 + \frac{1}{n}(\mathcal{E} - x_0) \subseteq C \subseteq \mathcal{E}$

So $\mathcal{E}$ approximates any C by a factor which depends only on dimension n.

This factor of n is tight for general C; we will prove it for $C = \operatorname{conv}\{x_1, \ldots, x_m\}$.

Rewrite opt. from before:

$$\min_{A,b} \ \log\det A^{-1} \quad \text{s.t.} \quad \|Ax_i + b\|_2^2 \le 1, \quad i = 1, \ldots, m$$

Introduce vars $\tilde{A} = A^2$ and $\tilde{b} = Ab$ to get:

$$\min_{\tilde{A}, \tilde{b}} \ \log\det \tilde{A}^{-1} \quad \text{s.t.} \quad x_i^T \tilde{A} x_i + 2\tilde{b}^T x_i + \tilde{b}^T \tilde{A}^{-1} \tilde{b} \le 1, \quad i = 1, \ldots, m$$

$$\left(\text{since } \log|\tilde{A}^{-1}| = \log|A^{-2}| = \log|A^{-1}|^2 = 2\log|A^{-1}|\right)$$

Derive KKT conditions:

$$\mathcal{L} = \log|\tilde{A}^{-1}| - \sum_{i=1}^m \lambda_i \left(1 - x_i^T \tilde{A} x_i - 2\tilde{b}^T x_i - \tilde{b}^T \tilde{A}^{-1} \tilde{b}\right), \quad \lambda \ge 0$$

Abuse of notation?

$$\frac{\partial}{\partial A} \mathcal{L} = 0 \implies -\tilde{A}^{-1} + \sum_{i=1}^m \lambda_i \left(x_i x_i^T + \tilde{b}\tilde{b}^T(-\tilde{A}^{-2})\right) = 0$$

$$\frac{\partial}{\partial b} \mathcal{L} = 0 \implies -\sum_{i=1}^m \lambda_i \left(-2x_i - 2\tilde{A}^{-1}\tilde{b}\right) = 0$$

using $\nabla \log|A| = A^{-1}$ and $\nabla A^{-1} = -A^{-2}$

These give:

$$\text{KKT} \begin{cases} \sum_{i=1}^m \lambda_i \left(x_i x_i^T - \tilde{A}^{-1}\tilde{b}\tilde{b}\tilde{A}^{-1}\right) = \tilde{A}^{-1} \\ \sum_{i=1}^m \lambda_i \left(x_i + \tilde{A}^{-1}\tilde{b}\right) = 0 \\ \left.\begin{array}{l} \lambda_i \ge 0 \\ x_i^T \tilde{A} x_i + 2\tilde{b}^T x_i + \tilde{b}^T \tilde{A}^{-1} \tilde{b} \le 1 \\ \lambda_i \left(1 - x_i^T \tilde{A} x_i - 2\tilde{b}^T x_i - \tilde{b}^T \tilde{A}^{-1}\tilde{b}\right) = 0 \end{array}\right\} i = 1, \ldots, m \end{cases}$$

We may do an affine change of coordinates so we may assume $\tilde{A} = I$ and $\tilde{b} = 0$, i.e. min vol ellipsoid is unit ball at origin.

Note: This does not affect the quantities we are interested in (i.e. how much we need to scale $\mathcal{E}$ to have it lie within C) since the coordinate change is affine.

KKT conditions simplify to:

$$\begin{cases} ① \sum_{i=1}^m \lambda_i x_i x_i^T = I & ② \sum_{i=1}^m \lambda_i x_i = 0 \\ ③ \lambda_i \geq 0 & ④ x_i^T x_i = \|x\|_2^2 \leq 1 \\ ⑤ \lambda_i (1 - x_i^T x_i) = 0 & \end{cases} \quad \Bigg\} \; i = 1, \ldots, m$$

Also, take trace of 1^{st} condition:

$$\sum_{i=1}^m \lambda_i \|x_i\|_2^2 = n$$

By complementary slackness, (last condition)

~~If $\lambda_i \neq 0$, then $\|x_i\|_2^2 = 1$ (using 4th condition)~~

this becomes $\sum_{i=1}^m \lambda_i = n$

Now, show that the shrunk ellipsoid (which is a ball with radius $1/n$ at origin) contains C.

i.e. $\|x\|_2 \leq 1/n \implies x \in C = \mathrm{conv}\{x_1, \ldots, x_n\}$

Suppose $\|x\|_2 \leq 1/n$.

From KKT,

$$x = \underbrace{\sum_{i=1}^m \lambda_i (x^T x_i) x_i}_{\text{multiply } 1^{st} \text{ condition by } x} = \underbrace{\sum_{i=1}^m \lambda_i \left(x^T x_i + \tfrac{1}{n}\right) x_i}_{\text{add } 2^{nd} \text{ condition}} = \sum_{i=1}^m \mu_i x_i$$

where $\mu_i = \lambda_i \left(x^T x_i + \frac{1}{n}\right)$

From Cauchy-Schwartz inequality,

$$\mu_i = \lambda_i (x^T x_i + \tfrac{1}{n}) \underset{\uparrow \text{ by Cauchy-Schwartz}}{\geq} \lambda_i (-\|x\|_2 \|x_i\|_2 + \tfrac{1}{n}) \underset{\uparrow \text{ by condition 4}}{\geq} \lambda_i (-1/n + 1/n) = 0$$

Also, $$\sum_{i=1}^{m} \mu_i \underset{\uparrow \text{ by def.}}{=} \sum_{i=1}^{m} \lambda_i (x^T x_i + \tfrac{1}{n}) \underset{\uparrow \text{ by condition 2}}{=} \sum_{i=1}^{m} \lambda_i / n = 1$$ ← proved earlier

So we have shown:

If $\|x\|_2 \leq \frac{1}{n}$

then $\exists \mu_i$ s.t. $x = \sum_{i=1}^{m} \mu_i x_i$, $\mu_i \geq 0$, $\sum_{i=1}^{m} \mu_i = 1$

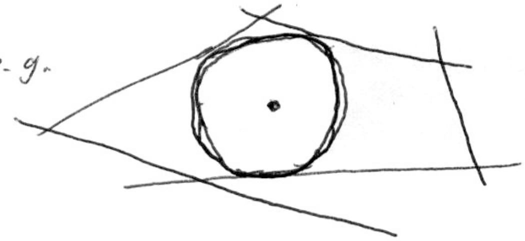
i.e. $x \in C = \text{conv}\{x_1, \dots, x_m\}$

Example: Chebyshev center

Idea: (1) Given convex set C, find point farthest from exterior.

(2) Find center of largest Euclidean ball in C.

e.g.



(1) + (2) define the Chebyshev center of C.

Definition: $\text{depth}(x, C) = \text{dist}(x, \mathbb{R}^n \setminus C)$

$$x_{cheb}(C) = \underset{x}{\text{argmax}} \ \underbrace{\text{depth}(x, C)}_{\text{concave}}$$

If C is convex set $C = \{x \mid f_1(x) \le 0, \dots, f_m(x) \le 0\}$ where f_i are convex, then solve:

$$\max_{R,x} R \quad \text{s.t.} \quad g_i(x,R) = \sup_{\|u\| \le 1} f_i(x + Ru) \le 0, \quad i = 1, \dots, m$$

This is convex since each g_i is pointwise max of family of convex functions of R, x.

If C is a polyhedron where $a_i^T x \le b_i$, $i = 1, \dots, m$, then ~~solve~~

$$g_i(x,R) = \sup_{\|u\| \le 1} a_i^T(x + Ru) - b_i$$

$$= a_i^T x + R\|a_i\|_* - b_i \quad (\text{if } R \ge 0)$$

where $\|\cdot\|_*$ is dual norm of $\|\cdot\|$

So solve:

$$\max_{R,x} R \quad \text{s.t.} \quad a_i^T x + R\|a_i\|_* \le b_i, \quad i = 1, \dots, m$$
$$R \ge 0$$

(This is an LP.)

Note: dual norms

The dual of the ℓ_p-norm is the ℓ_q-norm where $1/p + 1/q = 1$

If C is an intersection of ellipsoids,

i.e. $C = \{x \mid x^T A_i x + 2 b_i^T x + c_i \le 0,\ i=1,\dots,m\}$

where $A_i \in S_{++}^n$,

then

$$g_i(x,R) = \sup_{\|u\|_2 \le 1} \left((x+Ru)^T A_i (x+Ru) + 2 b_i^T (x+Ru) + c_i \right)$$

$$= x^T A_i x + 2 b_i^T x + c_i + \sup_{\|u\|_2 \le 1} \left(R^2 u^T A_i u + 2R(A_i x + b_i)^T u \right)$$

From §B.1 of Boyd + V., $g_i(x,R) \le 0$

iff $\exists \lambda_i$ such that

$$\begin{bmatrix} -x^T A_i x_i - 2 b_i^T x - c_i - \lambda_i & R(A_i x + b_i)^T \\ R(A_i x + b_i) & \lambda_i I - R^2 A_i \end{bmatrix} \succeq 0$$

Finally, we can rewrite this constraint using the Schur complement (backwards) to get this SDP:

$$\begin{cases} \max\limits_{R,\lambda,x} R \quad \text{s.t.} \\ \begin{bmatrix} -\lambda_i - c_i + b_i^T A_i^{-1} b_i & 0 & (x + A_i^{-1} b_i)^T \\ 0 & \lambda_i I & RI \\ x + A_i^{-1} b_i & RI & A_i^{-1} \end{bmatrix} \succeq 0, \quad i=1,\dots,m \end{cases}$$

To see this constraint is equivalent to the previous constraint, take the Schur complement w.r.t. A_i^{-1}:

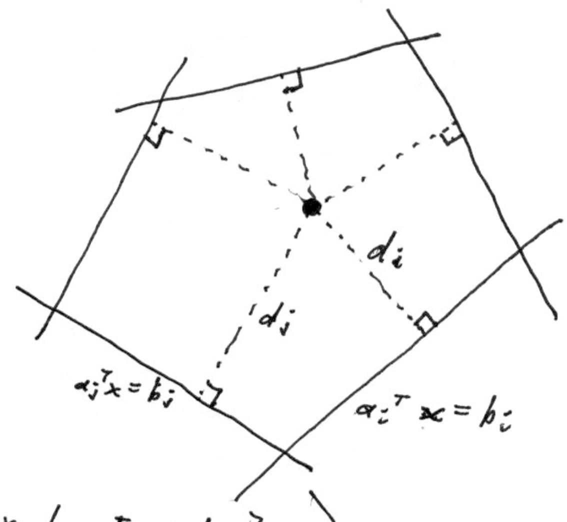
$$S = \begin{bmatrix} -\lambda_i - c_i + b_i^T A_i^{-1} b_i & 0 \\ 0 & \lambda_i I \end{bmatrix} - \begin{bmatrix} x + A_i^{-1} b_i & RI \end{bmatrix}^T A_i \begin{bmatrix} x + A_i^{-1} b_i & RI \end{bmatrix}$$

$$= \begin{bmatrix} -\lambda_i - c_i + b_i^T A_i^{-1} b_i & 0 \\ 0 & \lambda_i I \end{bmatrix} - \begin{bmatrix} x^T A_i x + 2 b_i^T x + b_i^T A_i^{-1} b_i & R(A_i x + b_i)^T \\ R(A_i x + b_i) & R^2 A_i \end{bmatrix}$$

= (above constraint)

Example: Analytic center (of a set of linear inequalities)

Goal: Given set of linear inequalities defining a (closed) ~~polygon~~ polyhedron, find point within polyhedron which maximizes ~~distance~~ ~~to~~ the product of distances to hyperplanes defined by inequalities.



i.e. $\left\{ \begin{array}{l} \text{Given hyperplanes } \mathcal{H}_i = \{x \mid a_i^T x = b_i\}, \\ \text{find } \underset{x}{\arg\max} \prod_i \operatorname{dist}(x, \mathcal{H}_i) \end{array} \right\}$

We ~~may~~ will assume $\|a_i\|_2 = 1$ ~~since positive scaling of~~ and that a_i point towards interior of ~~polygon~~ polyhedron.
So distance $d_i = \operatorname{dist}(x, \mathcal{H}_i) = b_i - a_i^T x$

~~So~~ We want: ~~the~~

$$\left\{ \max_x \prod_i (b_i - a_i^T x) \quad \text{s.t. } b_i - a_i^T x > 0, \ i = 1, \dots, m \right.$$

Transform this and make constraint implicit:

$$\left\{ \min_x - \sum_i \log(b_i - a_i^T x) \right.$$

This is convex if polyhedron is bounded.

This may be generalized:

The <u>analytic center</u> of the system $\left\{ \begin{array}{l} f_i(x) \le 0, \ i = 1, \dots, m, \\ Fx = g \end{array} \right.$ where f_i convex is optimum of:

$$\min_x - \sum_{i=1}^m \log(-f_i(x))$$

$$\text{s.t. } Fx = g$$

(with implicit constraints $f_i(x) < 0$, $i = 1, \dots, m$)

We assume $\{x \mid f_i(x) < 0, \ i = 1, \dots, m, \ Fx = g\}$ is non-empty and bounded.