COURSE ON LMI OPTIMIZATION
WITH APPLICATIONS IN CONTROL
PART I.2

GEOMETRY OF LMI SETS

Didier HENRION
www.laas.fr/~henrion
henrion@laas.fr

Macchina Comica (1928)
František Kupka (1871-1957)

November 2003

Geometry of LMI sets

Given symmetric matrices F_i we want to characterize the shape in $\mathbb{R}^n$ of the LMI set

$$\mathcal{F} = \{x \in \mathbb{R}^n \; : \; F(x) = F_0 + \sum_{i=1}^n x_i F_i \succeq 0\}$$

Matrix $F(x)$ is PSD iff its diagonal minors $f_i(x)$ are nonnegative

Diagonal minors are multivariate polynomials of indeterminates x_i

So the LMI set can be described as

$$\mathcal{F} = \{x \in \mathbb{R}^n \; : \; f_i(x) \geq 0, \; i = 1, \ldots, n\}$$

which is a semialgebraic set

Moreover, it is a convex set

Example of 2D LMI feasible set

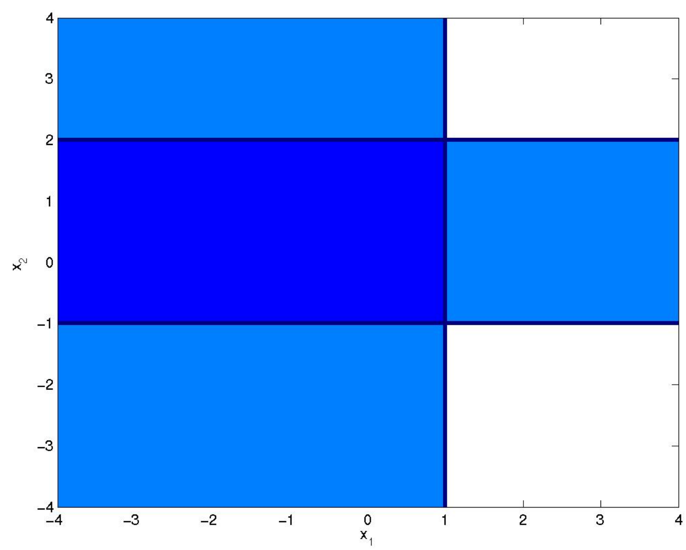
$$F(x) = \begin{bmatrix} 1-x_1 & x_1+x_2 & x_1 \\ x_1+x_2 & 2-x_2 & 0 \\ x_1 & 0 & 1+x_2 \end{bmatrix} \succeq 0$$

Feasible iff all principal minors nonnegative

System of polynomial inequalities $f_i(x) \geq 0$

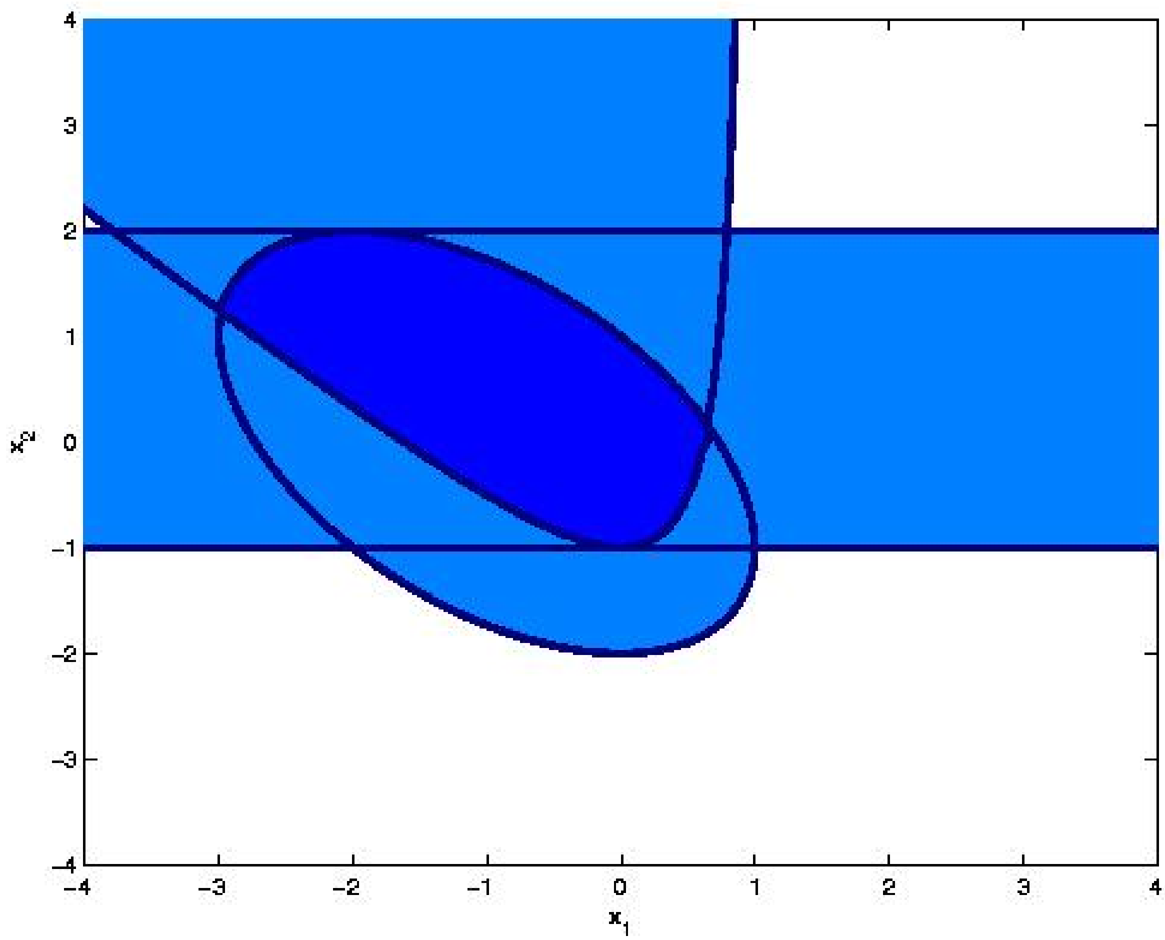
1st order minors

$f_1(x) = 1 - x_1 \geq 0$
$f_2(x) = 2 - x_2 \geq 0$
$f_3(x) = 1 + x_2 \geq 0$



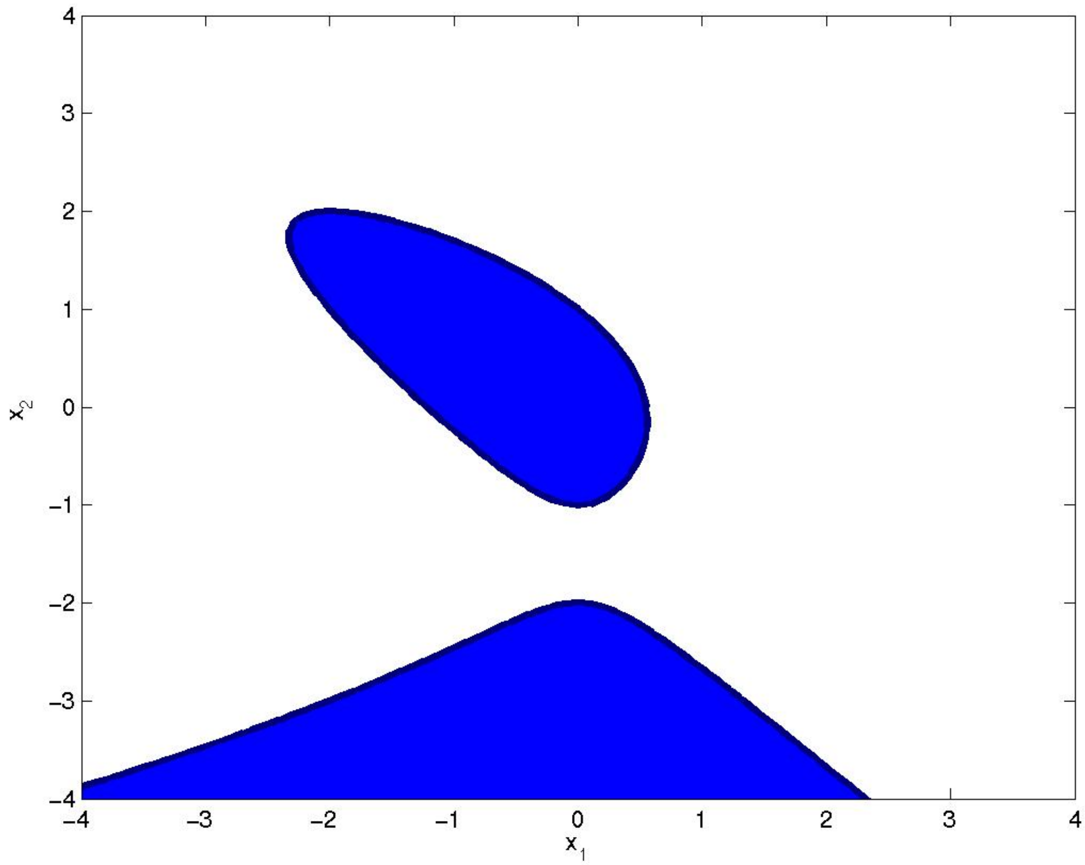
2nd order minors

$$f_4(x) = (1 - x_1)(2 - x_2) - (x_1 + x_2)^2 \geq 0$$
$$f_5(x) = (1 - x_1)(1 + x_2) - x_1^2 \geq 0$$
$$f_6(x) = (2 - x_2)(1 + x_2) \geq 0$$

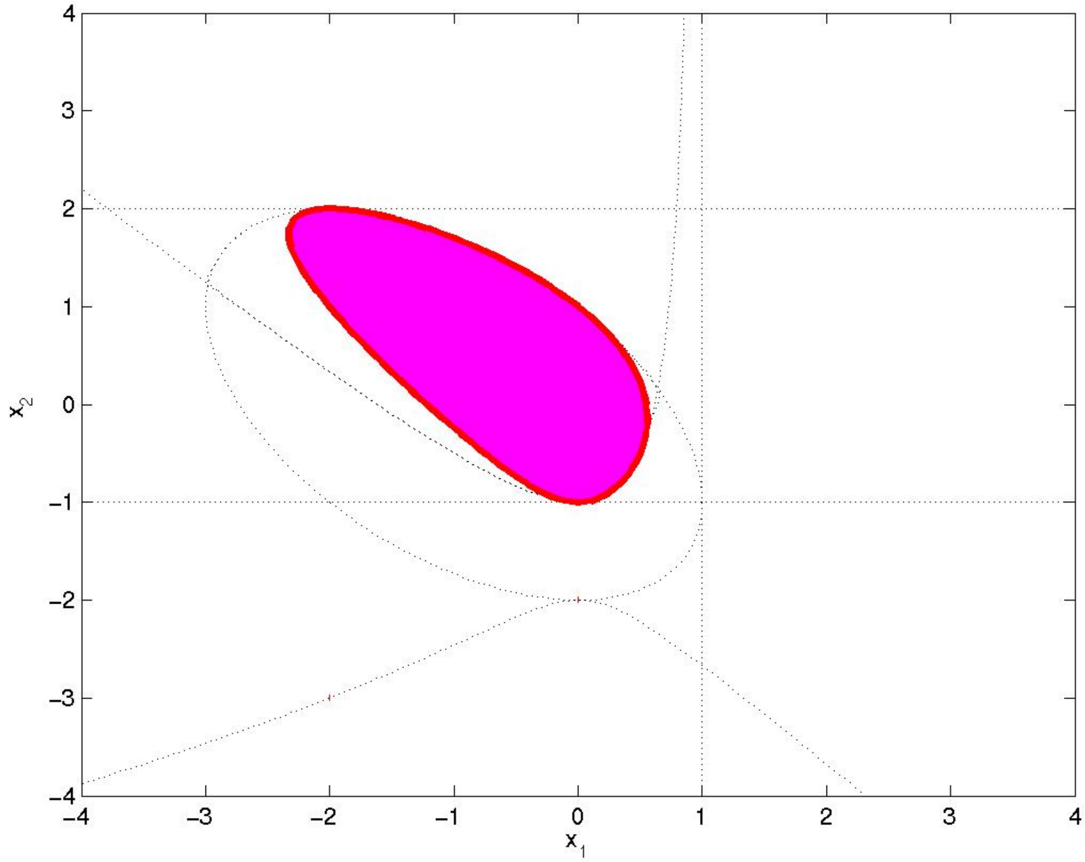


3rd order minor

$$f_7(x) = (1 + x_2)((1 - x_1)(2 - x_2) - (x_1 + x_2)^2) \\ -x_1^2(2 - x_2) \geq 0$$



LMI feasible set = intersection of semialgebraic sets $f_i(x) \geq 0$ for $i = 1, \ldots, 7$

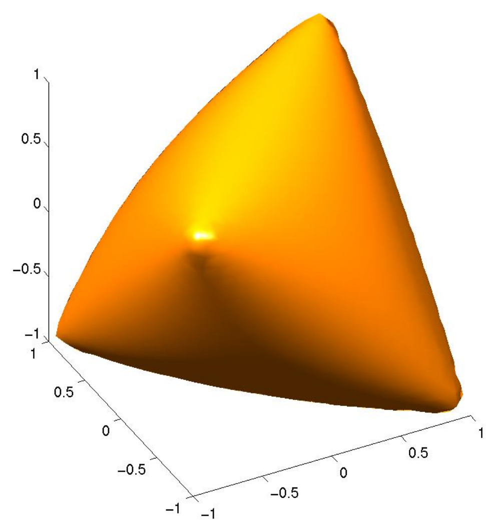


Example of 3D LMI feasible set

LMI set

$$\mathcal{F} = \{x \in \mathbb{R}^3 \ : \ \begin{bmatrix} 1 & x_1 & x_2 \\ x_1 & 1 & x_3 \\ x_2 & x_3 & 1 \end{bmatrix} \succeq 0\}$$

arising in SDP relaxation of MAXCUT



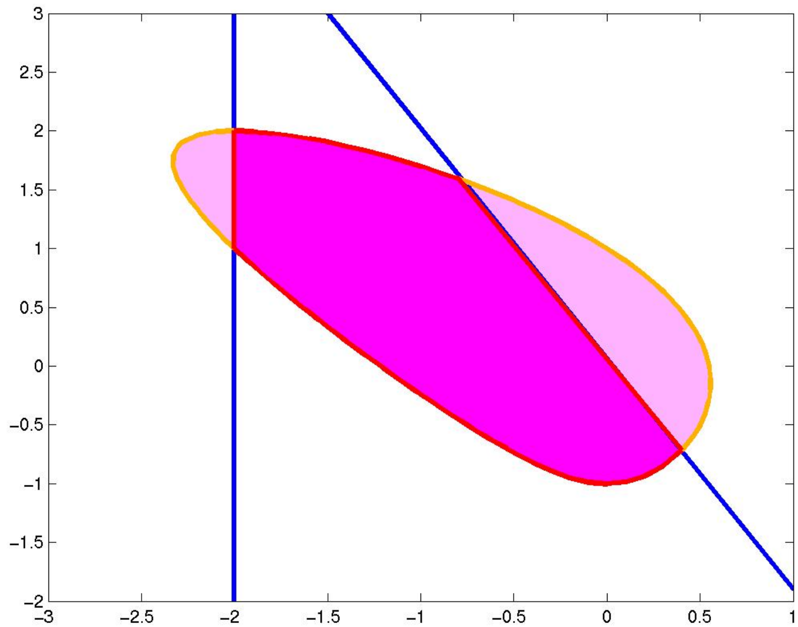
Semialgebraic set

$$\mathcal{F} = \{x \in \mathbb{R}^3 \ : \ 1 + 2x_1x_2x_3 - (x_1^2 + x_2^2 + x_3^2) \geq 0, \\ x_1^2 \leq 1,\ x_2^2 \leq 1,\ x_3^2 \leq 1$$

Intersection of LMI sets

Intersection of LMI feasible sets

$$F(x) \succeq 0 \quad x_1 \geq -2 \quad 2x_1 + x_2 \leq 0$$



is also an LMI

$$\begin{bmatrix} F(x) & 0 & 0 \\ 0 & x_1 + 2 & 0 \\ 0 & 0 & -2x_1 - x_2 \end{bmatrix} \succeq 0$$

SDP representability

LMI sets are convex semialgebraic sets.. but are all convex semialgebraic sets representable by LMIs ?

We say that a convex set $X \subset \mathbb{R}^n$ is SDP representable if there exists an affine mapping $F(x,u)$ such that

$$x \in X \iff \exists u \,:\, F(x,u) \succeq 0$$

In words, if X is the projection of the solution set of the LMI $F(x,u) \succeq 0$ onto the x-space and u are additional, or lifting variables

We say that a convex set $X \subset \mathbb{R}^n$ is LMI representable if there exists an affine mapping $F(x)$ such that

$$x \in X \iff F(x) \succeq 0$$

In other words, additional variables u are not allowed

Similarly, a convex function $f : \mathbb{R}^n \to \mathbb{R}$ is SDP or LMI representable if its epigraph

$$\mathcal{F} = \{x,\, t \,:\, f(x) \leq t\}$$

is an SDP or LMI representable set

Conic quadratic forms

The Lorentz, or ice-cream cone

$$\mathcal{L} = \{ \begin{bmatrix} x \\ t \end{bmatrix} \in \mathbb{R}^{n+1} \;:\; \|x\|_2 \leq t \}$$

is SDP (LMI) representable as

$$\mathcal{L} = \left\{ \begin{bmatrix} x \\ t \end{bmatrix} \;:\; \begin{bmatrix} tI_n & x \\ x^T & t \end{bmatrix} \succeq 0 \right\}$$

As a result, all (convex quadratic) conic representable sets are also SDP representable

The SOCP cone is included in the SDP cone

In the sequel we first give a list of conic representable sets (following Ben-Tal and Nemirovski)

Quadratic forms

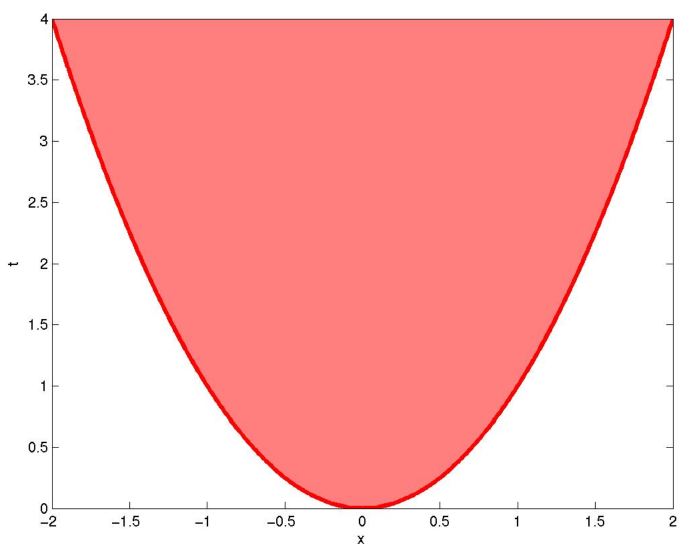
The Euclidean norm $\{x, t \in \mathbb{R}^2 : \|x\|_2 \le t\}$ is conic representable by definition

The squared Euclidean norm

$$\{x, t \in \mathbb{R}^2 : x^T x \le t\}$$

is conic representable as

$$\left\| \begin{bmatrix} x \\ \frac{t-1}{2} \end{bmatrix} \right\|_2 \le \frac{t+1}{2}$$



Quadratic forms (2)

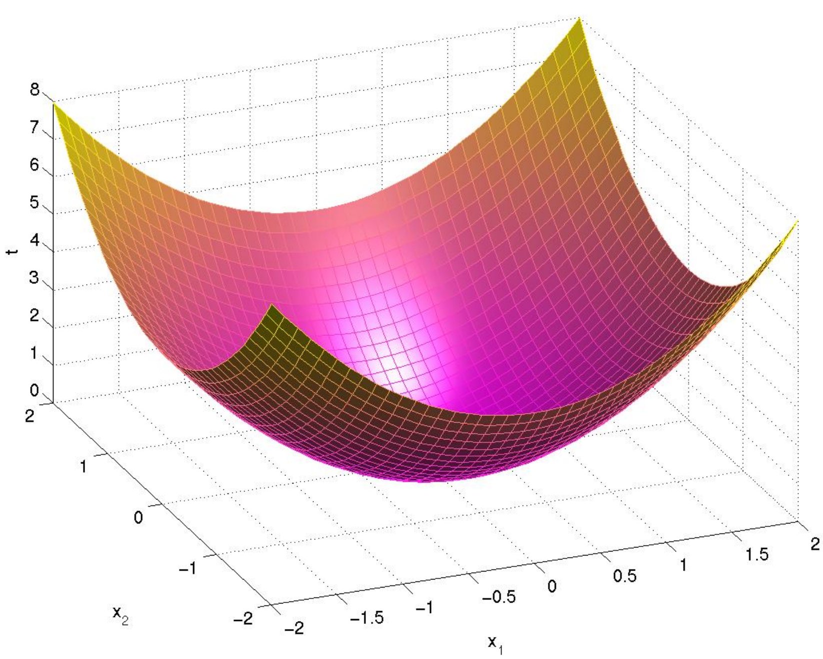
More generally, the convex quadratic set

$$\{x \in \mathbb{R}^n, t \in \mathbb{R} \,:\, x^T A x + b^T x + c \leq 0\}$$

with $A = A^T \succeq 0$ is conic representable as

$$\left\| \begin{bmatrix} Dx \\ \frac{t + b^T x + c}{2} \end{bmatrix} \right\|_2 \leq \frac{t - b^T x - c}{2}$$

where D is the Cholesky factor of $A = D^T D$



Who is Cholesky ?

André Louis Cholesky (1875-1918) was a French military officer (graduated from Ecole Polytechnique) involved in geodesy

He proposed a new procedure for solving least-squares triangulation problems

He fell for his country during World War I

Work posthumously published in
Commandant Benoît. Procédé du Commandant Cholesky. *Bulletin Géodésique*, No. 2, pp. 67-77, Toulouse, Privat, 1924.

Nice biography in
C. Brezinski. André Louis Cholesky. *Bulletin of the Belgian Mathematical Society*, Vol. 3, pp. 45-50, 1996.

I. — NOTICES SCIENTIFIQUES

Commandant BENOIT[1].

NOTE SUR UNE MÉTHODE DE RÉSOLUTION DES ÉQUATIONS NORMALES PROVENANT DE L'APPLICATION DE LA MÉTHODE DES MOINDRES CARRÉS A UN SYSTÈME D'ÉQUATIONS LINÉAIRES EN NOMBRE INFÉRIEUR A CELUI DES INCONNUES. — APPLICATION DE LA MÉTHODE A LA RÉSOLUTION D'UN SYSTÈME DEFINI D'ÉQUATIONS LINÉAIRES.

(Procédé du Commandant CHOLESKY[2].)

Le Commandant d'Artillerie Cholesky, du Service géographique de l'Armée, tué pendant la grande guerre, a imaginé, au cours de recherches sur la compensation des réseaux géodésiques, un procédé très ingénieux de résolution des équations dites *normales*, obtenues par application de la méthode des moindres carrés à des équations linéaires en nombre inférieur à celui des inconnues. Il en a conclu une méthode générale de résolution des équations linéaires.

Nous suivrons, pour la démonstration de cette méthode, la progression même qui a servi au Commandant Cholesky pour l'imaginer.

1. De l'Artillerie coloniale, ancien officier géodésien au Service géographique de l'Armée et au Service géographique de l'Indo-Chine, Membre du Comité national français de Géodésie et Géophysique.

2. Sur le Commandant Cholesky, tué à l'ennemi le 31 août 1918, voir la notice biographique insérée dans le volume du *Bulletin géodésique* de 1922 intitulé : *Union géodésique et géophysique internationale, Première Assemblée générale, Rome, mai 1922, Section de Géodésie*, Toulouse, Privat, 1922, in-8°, 241 p., pp. 159 à 161.

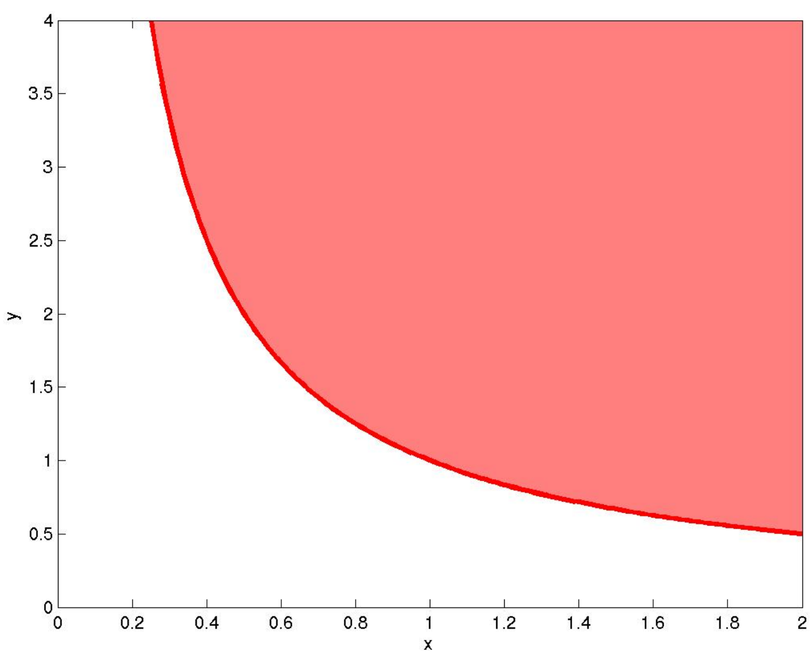
Hyperbola

The branch of hyperbola

$$\{x, y \in \mathbb{R}^2 \,:\, xy \geq 1,\, x > 0\}$$

is conic representable as

$$\left\| \begin{bmatrix} \frac{x-y}{2} \\ 1 \end{bmatrix} \right\|_2 \leq \frac{x+y}{2}$$



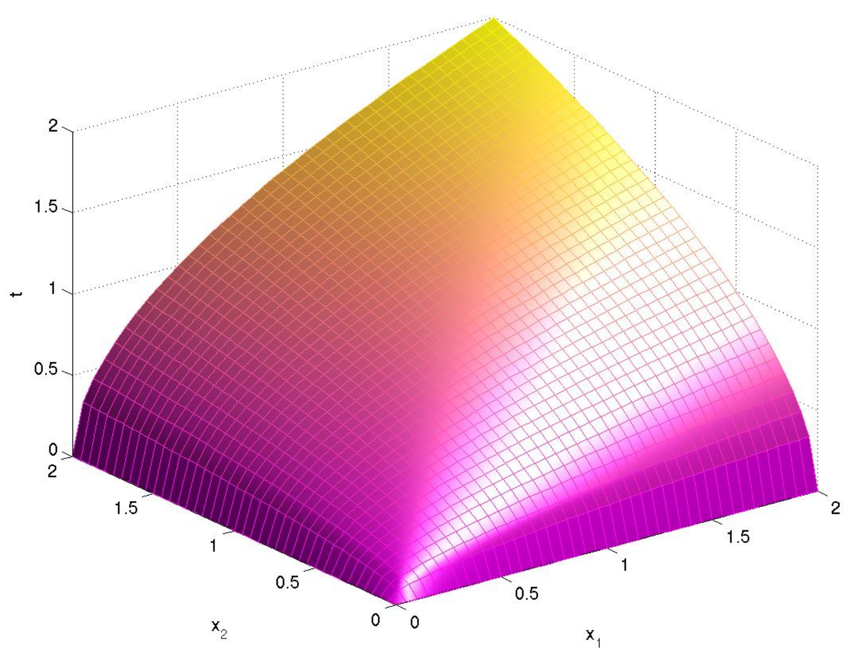
Geometric mean of two variables

The hypograph of the geometric mean of 2 variables

$$\{x_1, x_2, t \in \mathbb{R}^3 \;:\; x_1, x_2 \geq 0,\; \sqrt{x_1 x_2} \geq t\}$$

is conic representable as

$$\exists u \;:\; u \geq t,\; \left\| \begin{bmatrix} u \\ \frac{x_1 - x_2}{2} \end{bmatrix} \right\|_2 \leq \frac{x_1 + x_2}{2}$$



Geometric mean of several variables

The hypograph of the geometric mean of 2^k variables

$$\{x_1, \ldots, x_{2^k}, t \in \mathbb{R}^{2^k+1} \,:\, x_i \geq 0,\ \sqrt{x_1 \cdots x_{2^k}} \geq t\}$$

is also conic representable

Proof: iterate the previous construction

Example with $k = 3$:

$$(x_1 x_2 \cdots x_8)^{1/8} \geq t$$

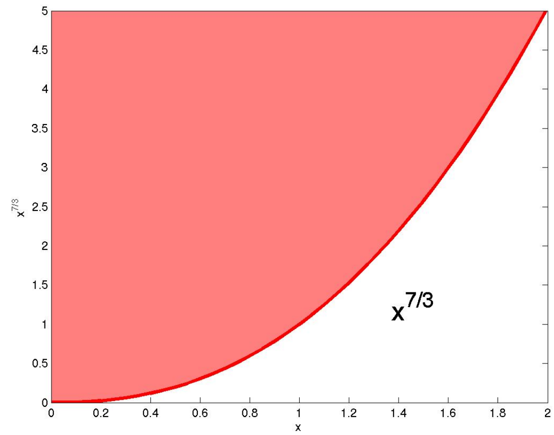
$$\left.\begin{array}{rcl} \sqrt{x_1 x_2} & \geq & x_{11} \\ \sqrt{x_3 x_4} & \geq & x_{12} \\ \sqrt{x_5 x_6} & \geq & x_{13} \\ \sqrt{x_7 x_8} & \geq & x_{14} \end{array}\right\} \left.\begin{array}{rcl} \sqrt{x_{11} x_{12}} & \geq & x_{21} \\ \sqrt{x_{13} x_{14}} & \geq & x_{22} \end{array}\right\} \sqrt{x_{21} x_{22}} \geq x_{31} \geq t$$

Useful idea in other SDP representability problems

Rational functions

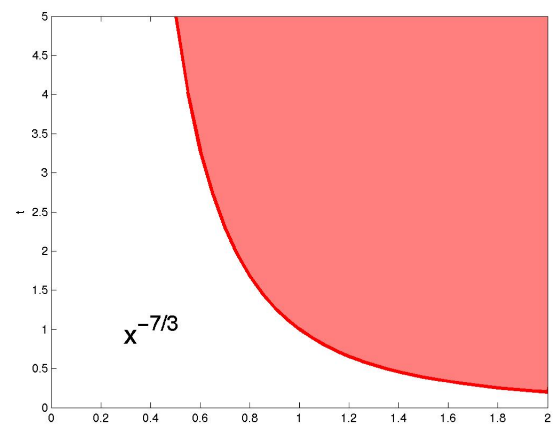
Usually similar ideas, we can show that the increasing rational power functions

$$f(x) = x^{p/q}, \quad x \geq 0$$



with rational $p/q \geq 1$, as well as the decreasing

$$g(x) = x^{-p/q}, \quad x \geq 0$$



with rational $p/q \geq 0$, are both conic representable

Rational functions

Example:

$$\{x, t \,:\, x \geq 0,\; x^{7/3} \leq t\}$$

Start from conic representable

$$\hat{t} \leq (\hat{x}_1 \cdots \hat{x}_8)^{1/8}$$

and replace

$$\begin{array}{l} \hat{t} = \hat{x}_1 = x \geq 0 \\ \hat{x}_2 = \hat{x}_3 = \hat{x}_4 = t \geq 0 \\ \hat{x}_5 = \hat{x}_6 = \hat{x}_7 = \hat{x}_8 = 1 \end{array}$$

to get

$$\begin{array}{rcl} x & \leq & x^{1/8} t^{3/8} \\ x^{7/8} & \leq & t^{3/8} \\ x^{7/3} & \leq & t \end{array}$$

Same idea works for any rational $p/q \geq 1$

- lift = use additional variables, and
- project in the space of original variables

Even power monomial

The epigraph of even power monomial

$$\mathcal{F} = \{x, t \;:\; x^{2p} \le t\}$$

where p is a positive integer
is conic representable

Note that

$$\{x, t \;:\; x^{2p} \le t\}$$

$$\Longleftrightarrow$$

$$\{x, y, t \;:\; x^2 \le y\}$$
$$\{x, y, t \;:\; y \ge 0,\; y^p \le t\}$$

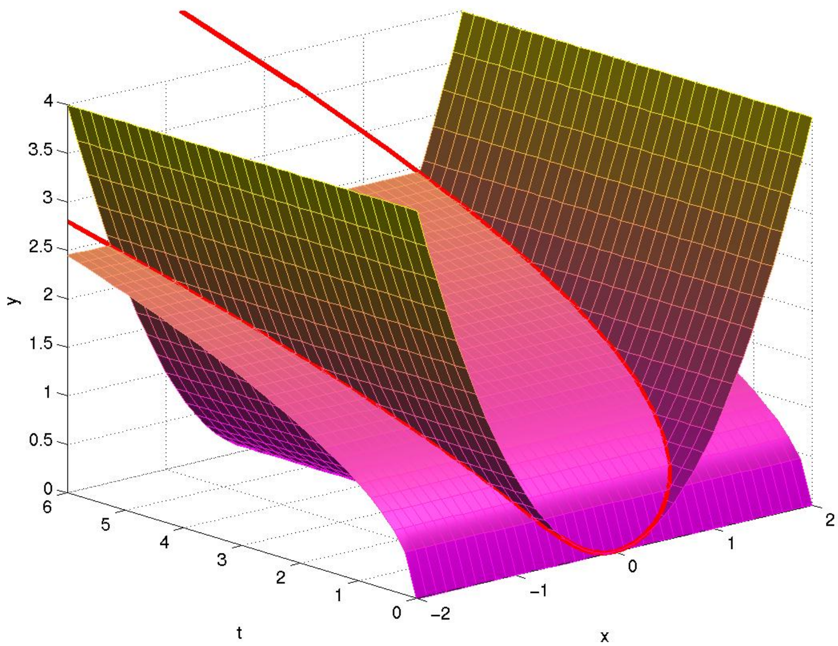
both conic representable

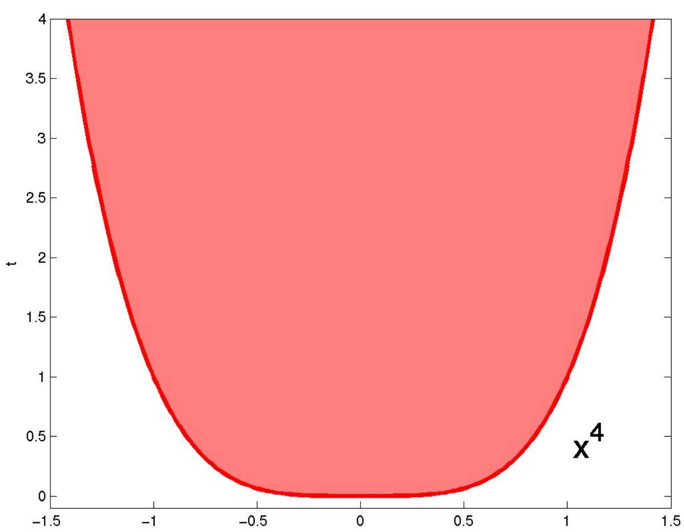
Use lifting y and project back onto x, t

Similarly, even power polynomials are conic representable (combinations of monomials)

Even power monomial

$$\mathcal{F} = \{x, t \; : \; x^4 \leq t\}$$





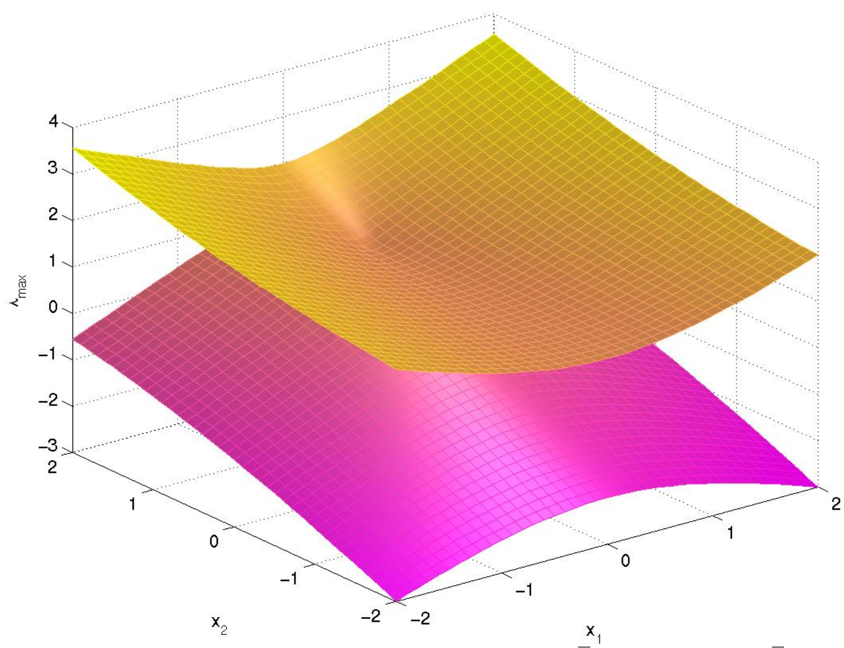
Largest eigenvalue

The epigraph of the function largest eigenvalue of a symmetric matrix

$$\{X = X^T \in \mathbb{R}^{n\times n}, t \in \mathbb{R} \,:\, \lambda_{\max}(X) \le t\}$$

is SDP (LMI) representable as

$$X \preceq tI_n$$



Eigenvalues of matrix $\begin{bmatrix} 1 & x_1 \\ x_1 & x_2 \end{bmatrix}$

Sums of largest eigenvalues

Let

$$S_k(X) = \sum_{i=1}^{k} \lambda_i(X), \quad k = 1, \ldots, n$$

denote the sum of the k largest eigenvalues of an n-by-n symmetric matrix X

The epigraph

$$\{X = X^T \in \mathbb{R}^{n\times n}, t \in \mathbb{R} \,:\, S_k(X) \leq t\}$$

is SDP representable as

$$\begin{array}{c} t - ks - \text{trace}\, Z \succeq 0 \\ Z \succeq 0 \\ Z - X + sI_n \succeq 0 \end{array}$$

where Z and s are additional variables

Determinant of a PSD matrix

The determinant

$$\det(X) = \prod_{i=1}^{n} \lambda_i(X)$$

is not a convex function of X, but the function

$$f_q(X) = -\det{}^q(X), \quad X = X^T \succeq 0$$

is convex when $q \in [0,\ 1/n]$ is rational

The epigraph

$$\{f_q(X) \le t\}$$

is SDP representable as

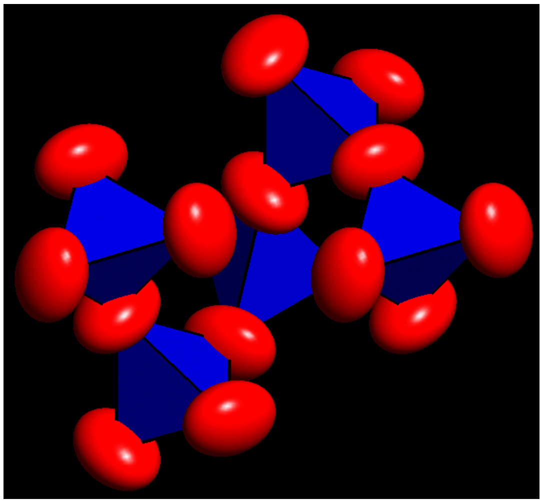
$$\begin{bmatrix} X & \Delta \\ \Delta^T & \operatorname{diag} \Delta \end{bmatrix} \succeq 0$$
$$t \le (\delta_1 \cdots \delta_n)^q$$

since we know that the latter constraint (hypograph of a concave monomial) is conic representable

Here Δ is a lower triangular matrix of additional variables with diagonal entries δ_i

Application: extremal ellipsoids

A little excursion in the world of ellipsoids and polytopes..



Crystal structure

Various representations of an ellipsoid in $\mathbb{R}^n$

$$\begin{aligned} E &= \{x \in \mathbb{R}^n \,:\, x^T P x + 2x^T q + r \leq 0\} \\ &= \{x \in \mathbb{R}^n \,:\, (x - x_c)^T P (x - x_c) \leq 1\} \\ &= \{x = Qy + x_c \in \mathbb{R}^n \,:\, y^T y \leq 1\} \\ &= \{x \in \mathbb{R}^n \,:\, \|Rx - x_c\| \leq 1\} \end{aligned}$$

where

$$Q = R^{-1} = P^{-1/2} \succ 0$$

Ellipsoid volume

Volume of ellipsoid $E = \{Qy + x_c \,:\, y^T y \leq 1\}$

$$\mathrm{vol}\, E = k_n \det Q$$

where k_n is volume of n-dimensional unit ball

$$k_n = \begin{cases} \dfrac{2^{(n+1)/2}\pi^{(n-1)/2}}{n(n-2)!!} & \text{for } n \text{ odd} \\ \dfrac{2\pi^{n/2}}{n(n/2-1)!} & \text{for } n \text{ even} \end{cases}$$

n	1	2	3	4	5	6	7	8
k_n	2.00	3.14	4.19	4.93	5.26	5.17	4.72	4.06

Unit ball has maximum volume for $n = 5$!

Outer and inner ellipsoidal approximations

Let $S \subset \mathbb{R}^n$ be a solid = a closed bounded convex set with nonempty interior

- the largest volume ellipsoid E_{in} contained in S is unique and satisfies

$$E_{\text{in}} \subset S \subset nE_{\text{in}}$$

- the smallest volume ellipsoid E_{out} containing S is unique and satisfies

$$E_{\text{out}}/n \subset S \subset E_{\text{out}}$$

These are Löwner-John ellipsoids

Factor n reduces to $\sqrt{n}$ if S is symmetric

How can these ellipsoids be computed ?

Ellipsoid in polytope

Let the intersection of hyperplanes

$$S = \{x \in \mathbb{R}^n \,:\, a_i^T x \leq b_i,\; i = 1, \ldots, m\}$$

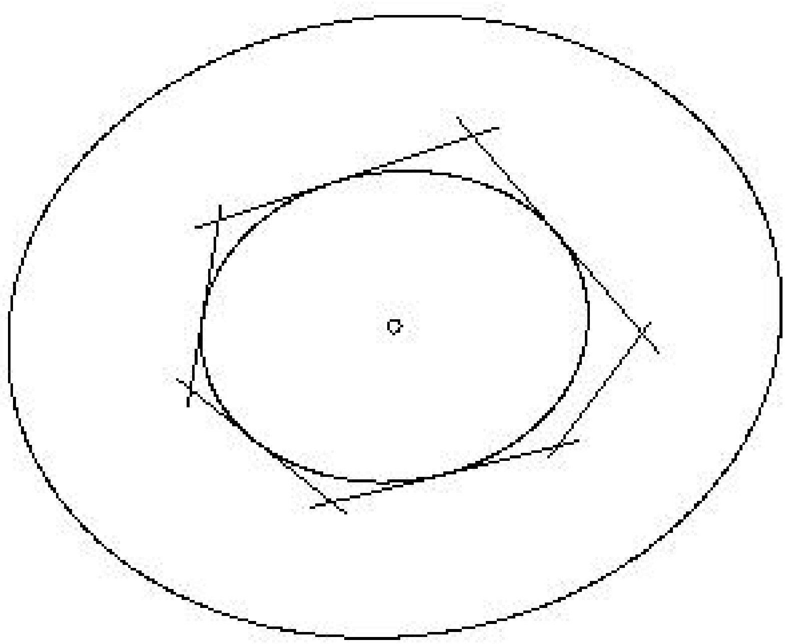
describe a polytope = bounded nonempty polyhedron

The largest volume ellipsoid contained in S is

$$E = \{Qy + x_c \,:\, y^T y \leq 1\}$$

where Q, x_c are optimal solutions of the LMI

$$\begin{array}{ll} \max & \det^{1/n} Q \\ & Q \succeq 0 \\ & \|Qa_i\|_2 \leq b_i - a_i^T x_c, \quad i = 1, \ldots, m \end{array}$$



Polytope in ellipsoid

Let the convex hull of vertices

$$S = \text{conv}\,\{x_1, \ldots, x_m\}$$

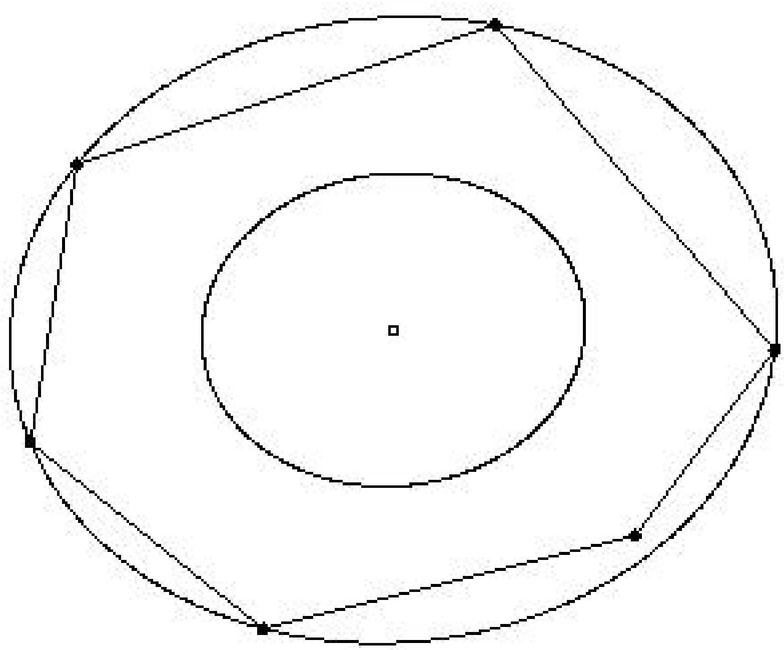
describe a polytope

The smallest volume ellipsoid containing S is

$$E = \{x\,:\,(x - x_c)^T P (x - x_c) \le 1\}$$

where P, $x_c = -P^{-1}q$ are optimal solutions of the LMI

$$\begin{array}{ll} \max & t \\ & t \le \det^{1/n} P \\ & \begin{bmatrix} P & q \\ q^T & r \end{bmatrix} \succeq 0 \\ & x_i^T P x_i + 2x_i^T q + r \le 1, \quad i = 1, \ldots, m \end{array}$$



Sums of largest singular values

Let

$$\Sigma_k(X) = \sum_{i=1}^{k} \sigma_i(X), \quad k = 1, \ldots, n$$

denote the sum of the k largest singular values of an n-by-m matrix X

Then the epigraph

$$\{X = X^T \in \mathbb{R}^{n \times n}, t \in \mathbb{R} \, : \, \Sigma_k(X) \leq t\}$$

is SDP representable since

$$\sigma_i(X) = \lambda_i \left(\begin{bmatrix} 0 & X^T \\ X & 0 \end{bmatrix} \right)$$

and the sum of largest eigenvalues of a symmetric matrix is SDP representable

Schur complement

We can use the Schur complement to convert a non-linear matrix inequality into an LMI

$$\begin{bmatrix} A(x) & B(x) \\ B^T(x) & C(x) \end{bmatrix} \succeq 0 \quad \Longleftrightarrow \quad A(x) - B(x)C^{-1}(x)B^T(x) \succeq 0$$
$$C(x) \succ 0 \qquad\qquad C(x) \succ 0$$

Issai Schur
(1875 Mogilyov - 1941 Tel Aviv)

Elimination lemma

To remove decision variables we can use the elimination lemma

$$A(x) + B(x)XC^T(x) + C(x)X^TB^T(x) > 0$$
$$\Longleftrightarrow$$
$$\tilde{B}^T(x)A(x)\tilde{B}(x) > 0 \quad \tilde{C}^T(x)A(x)\tilde{C}(x) > 0$$

where $\tilde{B}$ and $\tilde{C}$ are orthogonal complements of B and C respectively, and x is a decision variable independent of matrix X

Can be shown with SDP duality..

Particular case: Finsler's theorem

Positive polynomials

The set of univariate polynomials that are positive on the real axis is a convex set that can be described by an LMI

Can be proved with cone duality (Nesterov) or with theory of moments (Lasserre)

The even polynomial

$$p(s) = p_0 + p_1 s + \cdots + p_{2n} s^{2n}$$

satisfies $p(s) \geq 0$ for all $s \in \mathbb{R}$ if and only if

$$\begin{aligned} p_k &= \textstyle\sum_{i+j=k} X_{ij}, \qquad k = 0, 1, \ldots, 2n \\ &= \operatorname{trace} H_k X \end{aligned}$$

for some matrix $X = X^T \succeq 0$

Sum-of-squares decomposition

The expression of p_k with Hankel matrices H_k comes from

$$p(s) = [1 \quad s \quad \cdots \quad s^n] X [1 \quad s \quad \cdots \quad s^n]^\star$$

hence $X \succeq 0$ naturally implies $p(s) \geq 0$

Conversely, existence of X for any polynomial $p(s) \geq 0$ follows from the existence of a sum-of-squares decomposition (with at most two elements) of

$$p(s) = \sum_k q_k^2(s) \geq 0$$

Matrix X has entries $X_{ij} = \sum_k q_{k_i} q_{k_j}$

Primal and dual formulations

Global minimization of polynomial

$$p(s) = \sum_{k=0}^{n} p_k s^k$$

Global optimum p^*: maximum value of $\widehat{p}$ such that $p(s) - \widehat{p}$ stays globally nonnegative

Primal LMI

$$\begin{array}{ll} \max & \widehat{p} = p_0 - \text{trace}\, H_0 X \\ \text{s.t.} & \text{trace}\, H_k X = p_k, \quad k = 1, \ldots, n \\ & X \succeq 0 \end{array}$$

Dual LMI

$$\begin{array}{ll} \min & p_0 + \sum_{k=1}^{n} p_k y_k \\ \text{s.t.} & H_0 + \sum_{k=1}^{n} H_k y_k \succeq 0 \end{array}$$

with Hankel structure (moment matrix)

Positive polynomials and LMIs

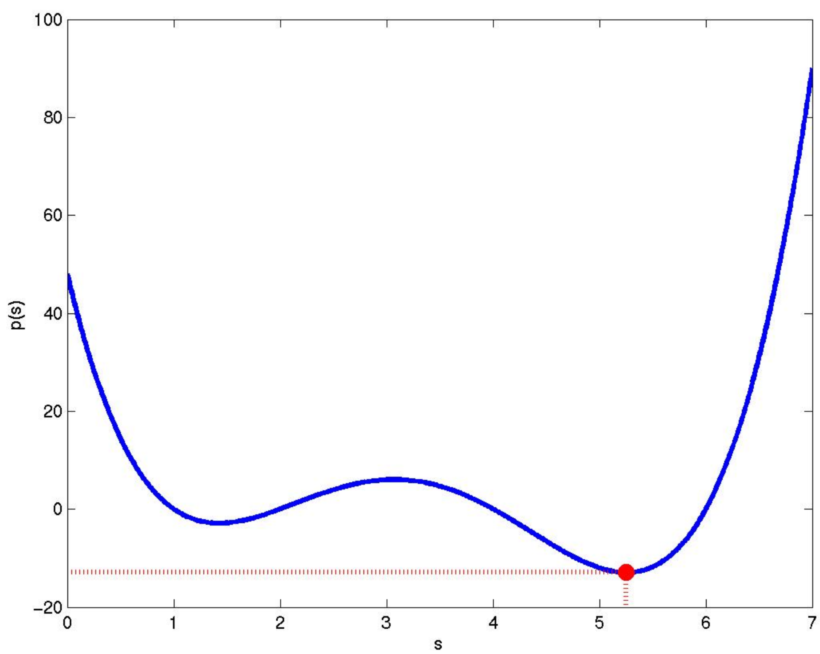
Example: Global minimization of the polynomial

$$p(s) = 48 - 92s + 56s^2 - 13s^3 + s^4$$

We just have to solve the dual LMI

$$\begin{array}{ll} \min & 48 - 92y_1 + 56y_2 - 13y_3 + y_4 \\ \text{s.t.} & \begin{bmatrix} 1 & y_1 & y_2 \\ y_1 & y_2 & y_3 \\ y_2 & y_3 & y_4 \end{bmatrix} \succeq 0 \end{array}$$

to obtain $p^* = p(5.25) = -12.89$



Complex LMIs

The complex valued LMI

$$F(x) = A(x) + jB(x) \succeq 0$$

is equivalent to the real valued LMI

$$\begin{bmatrix} A(x) & B(x) \\ -B(x) & A(x) \end{bmatrix} \succeq 0$$

If there is a complex solution to the LMI then there is a real solution to the same LMI

Note that matrix $A(x) = A^T(x)$ is symmetric whereas $B(x) = -B^T(x)$ is skew-symmetric

Rigid convexity

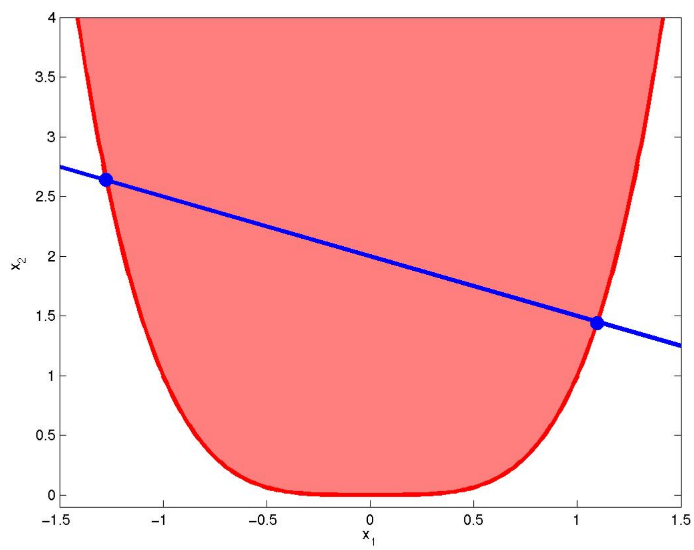
Helton & Vinnikov showed that a convex 2D set

$$\mathcal{F} = \{x \in \mathbb{R}^2 \,:\, p(x) \geq 0\}$$

defined by a polynomial $p(x)$ of minimum degree d
is LMI representable without lifting variables
iff $\mathcal{F}$ is rigidly convex, meaning that

for every point $x \in X$ and almost every line through x
then the line intersects $p(x) = 0$ in exactly d points

Example: $\mathcal{F} = \{x_1, x_2 \in \mathbb{R}^2 \,:\, p(x) = x_2 - x_1^4 \geq 0\}$
with 2 line intersections
is not rigidly convex because $2 < d = 4$



.. but it is LMI representable with lifting variables
see the previous construction for even power monomials