I.5. NONCONVEX LMI MODELLING

Didier HENRION
henrion@laas.fr

Belgian Graduate School on
Systems, Control, Optimization and Networks

Leuven - April and May 2010

BMI - Bilinear Matrix Inequality

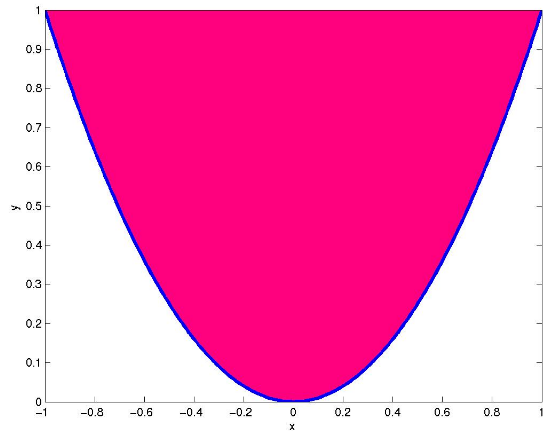
$$F(x) = F_0 + \sum_i x_i F_i + \sum_i \sum_j x_i x_j F_{ij} \succeq 0$$

Symmetric matrices F_i, F_{ij} given

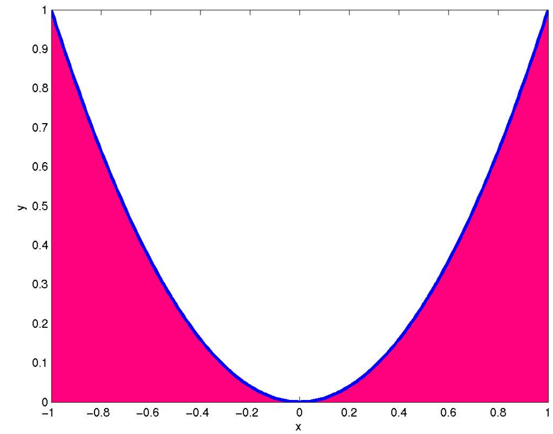
Decision variables x_i

Actually QMI with quadratic terms x_i^2

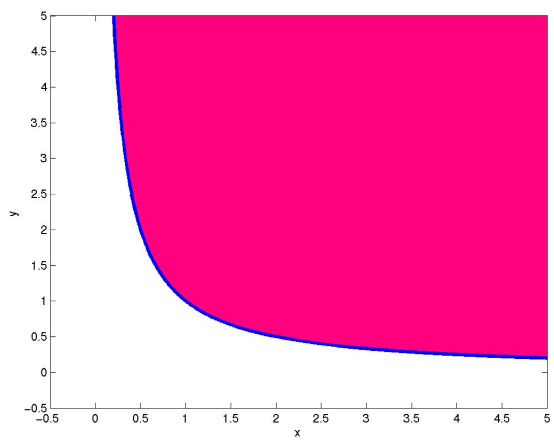
Contrary to LMIs, BMIs may have non-convex feasible sets



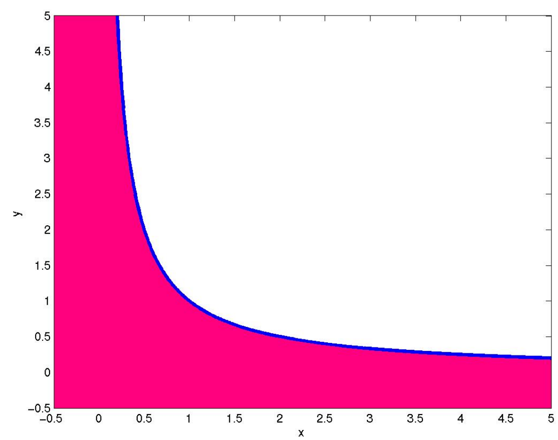
Convex LMI
$x^2 \leq y$



Nonconvex BMI
$x^2 \geq y$



Convex LMI
$xy \geq 1$



Nonconvex BMI
$xy \leq 1$

PMI - Polynomial Matrix Inequality

$$F(x) = \sum_\alpha x^\alpha F_\alpha \succeq 0$$

More general than BMI ?

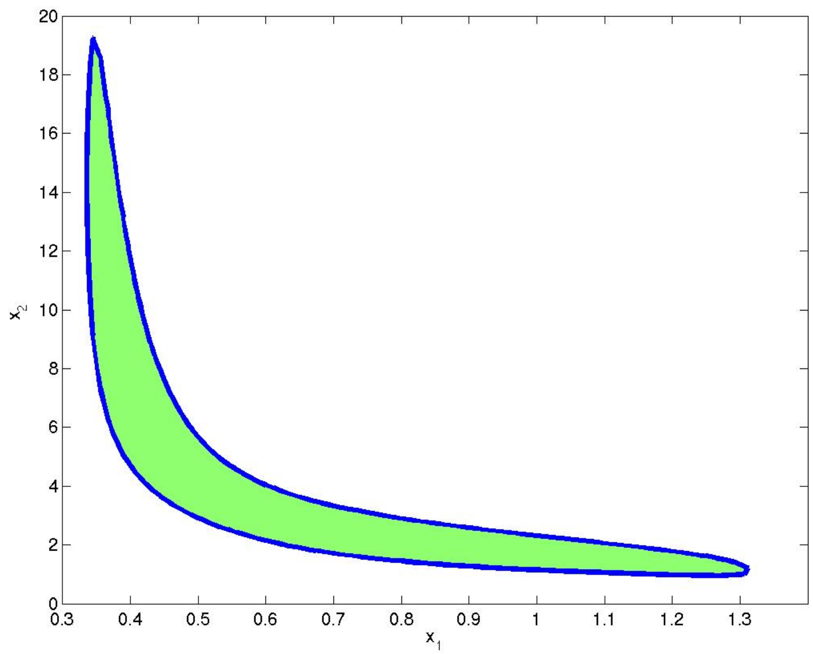
By appropriate changes of variables
any PMI can be written as a BMI

Example

$F(x) = F_0 + F_1 x_1 + F_{12} x_1 x_2^2 + F_{03} x_2^3$
can be written as the BMI
$F(x) = F_0 + F_1 x_1 + F_{12} x_1 x_3 + F_{03} x_2 x_3$
with lifting variable x_3 constrained by $x_3 = x_2^2$

Example of a nonconvex 2D BMI

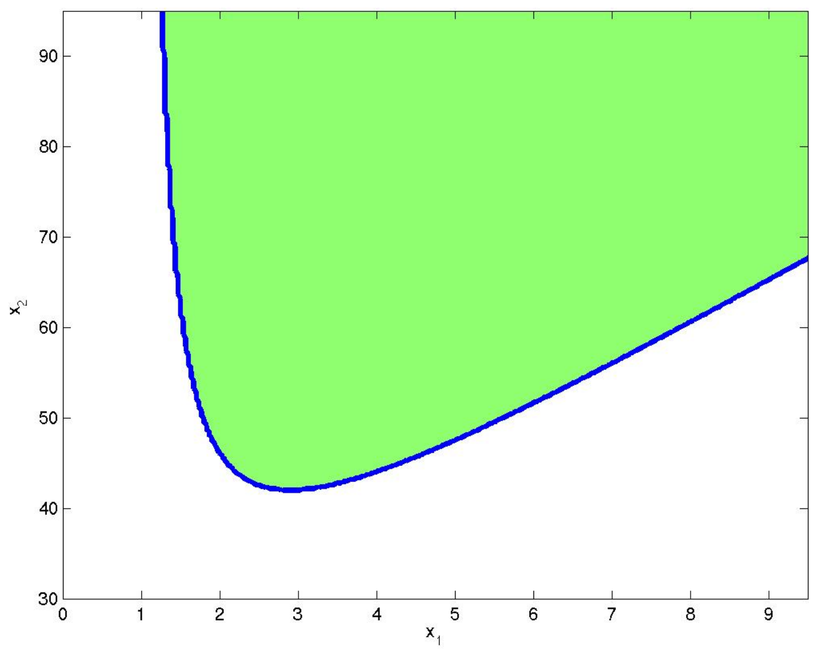
$$F(x) = \begin{bmatrix} 10 & 0.5 & 2 \\ 0.5 & -4.5 & 0 \\ 2 & 0 & 0 \end{bmatrix} + \begin{bmatrix} -9 & -0.5 & 0 \\ -0.5 & 0 & 3 \\ 0 & 3 & 1 \end{bmatrix} x_1 +$$
$$\begin{bmatrix} 1.8 & 0.1 & 0.4 \\ 0.1 & -1.2 & 1 \\ 0.4 & 1 & 0 \end{bmatrix} x_2 + \begin{bmatrix} 0 & 0 & -2 \\ 0 & 5.5 & -3 \\ -2 & -3 & 0 \end{bmatrix} x_1 x_2 \succeq 0$$



Example of a convex 2D BMI

Coming from a static output feedback problem

$$\begin{bmatrix} x_2(-13-5x_1+x_2) & x_2 & 0 \\ x_2 & x_1 & 0 \\ 0 & 0 & x_1(-13-5x_1+x_2)-x_2 \end{bmatrix} \succ 0$$



Converting BMI into LMI

Can we detect or exploit convexity ?

Sometimes a BMI problem can be reformulated as an LMI problem

For example the static output feedback BMI can be reformulated as the LMI

$$\begin{bmatrix} -1 + x_1 & 1 \\ 1 & -1 - \frac{5}{18}x_1 + \frac{1}{18}x_2 \end{bmatrix} \succ 0$$

Can we systematically detect whether such a reformulation is possible ?

Can we design a systematic reformulation algorithm ?

History of BMIs

Interest in BMIs originated in systems control
Mid 1990s, Safonov's team

Typical BMI: static output feedback

$$(A + BKC)^T P + P(A + BKC) \prec 0, \quad P \succ 0$$

More intricate BMI arise for reduced order controller design, H_2, H_∞ performance..

Main criticisms:
- too general
- no good algorithm

..in sharp contrast with LMIs..

BMI as a rank-one LMI

Defining liftings $x_{ij} = x_i x_j$ the BMI

$$F_0 + \sum_i x_i F_i + \sum_i \sum_j x_i x_j F_{ij} \succeq 0$$

can be written as an LMI

$$F_0 + \sum_i x_i F_i + \sum_i \sum_j x_{ij} F_{ij} \succeq 0$$

$$X = \begin{bmatrix} 1 & x_1 & x_2 & \\ x_1 & x_{11} & x_{12} & \\ x_2 & x_{12} & x_{22} & \\ & & & \ddots \end{bmatrix} \succeq 0$$

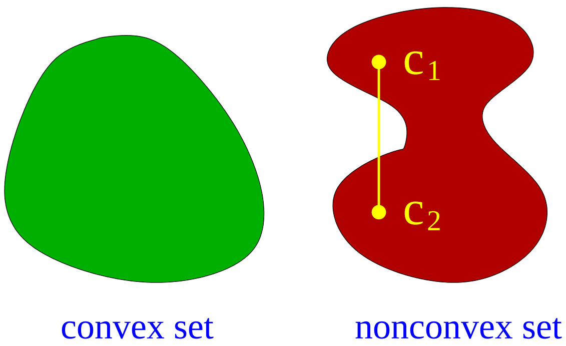
with an additional rank constraint

$$\operatorname{rank} X = 1$$

All the non-convexity is concentrated in this rank constraint

Handling nonconvexity

We have seen that additional variables, or liftings can prove useful in describing convex sets with LMIs



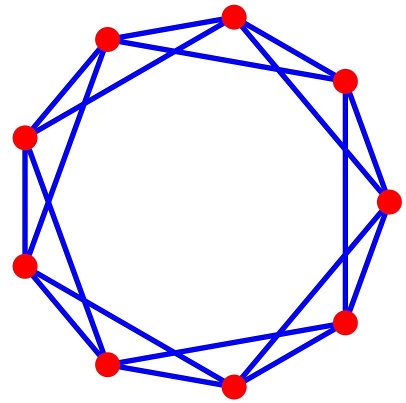
But LMI are also frequently used to cope with non-convex sets

This chapter is dedicated to the joint use of
- convex LMI relaxations, and
- additional variables = liftings

Combinatorial optimization

MAXCUT: typical combinatorial optimization problem

$$\begin{array}{ll} \text{min} & x^T Q x \\ \text{s.t.} & x_i \in \{-1,\ 1\} \end{array}$$



Antiweb AW_9^2 graph

Basic non-convex constraints

$$x_i^2 = 1$$

Exponential number of points = difficult problem

LMI relaxation

Basic idea..

For each i replace non-convex constraint

$$x_i^2 = 1$$

with relaxed convex constraint

$$x_i^2 \leq 1$$

which is an LMI constraint

$$\begin{bmatrix} 1 & x_i \\ x_i & 1 \end{bmatrix} \succeq 0$$

What about cross terms $x_i x_j$?

Dealing with cross terms

Replace all non-convex constraints $x_i^2 = 1$ for $i = 1, 2, \ldots, n$ with relaxed LMI constraint

$$X = \begin{bmatrix} 1 & x_1 & x_2 & \cdots & x_n \\ x_1 & 1 & x_{12} & & x_{1n} \\ x_2 & x_{12} & 1 & & x_{2n} \\ \vdots & & & \ddots & \vdots \\ x_n & x_{1n} & x_{2n} & \cdots & 1 \end{bmatrix} \succeq 0$$

where x_{ij} are additional variables = liftings

Always less conservative than previous relaxation because $X \succeq 0$ implies for all i

$$\begin{bmatrix} 1 & x_i \\ x_i & 1 \end{bmatrix} \succeq 0$$

Rank constrained LMI

In the original problem

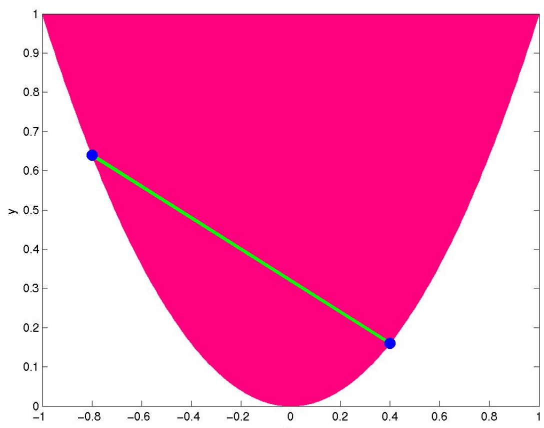
$$\begin{aligned} g^\star \;=\; & \min \;\; x^T Q x \\ & \text{s.t.} \;\; x_i^2 = 1 \end{aligned}$$

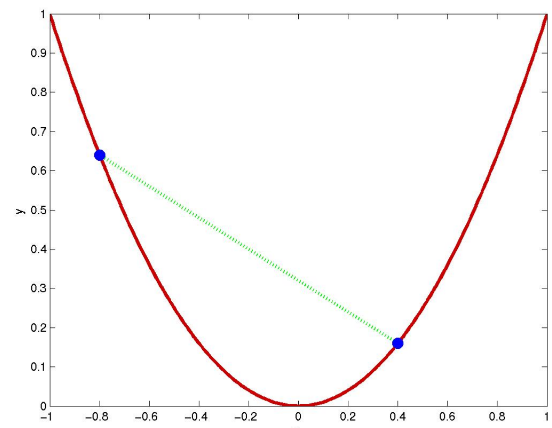
let $X = xx^T$ and then $x^T Q x = \operatorname{trace} Q x x^T = \operatorname{trace} QX$ and $x_i^2 = X_{ii} = 1$ so that the problem can be written as a rank constrained LMI

$$\begin{aligned} g^\star \;=\; & \min \;\; \operatorname{trace} QX \\ & \text{s.t.} \;\; X_{ii} = 1 \\ & \qquad X \succeq 0 \\ & \qquad \operatorname{rank} X = 1 \end{aligned}$$

Example of rank constrained LMI

$$X = \begin{bmatrix} y & x \\ x & 1 \end{bmatrix}$$





Convex set $X \succeq 0$ $(x^2 \leq y)$

Non-convex set $X \succeq 0$, rank $X = 1$ $(x^2 = y)$

Relaxing the rank constraint

All the nonconvexity is concentrated into the rank constraint, so we just drop it !

The obtained LMI relaxation is called Shor's relaxation

$$
\begin{array}{rcll}
p^\star & = & \min & \text{trace}\, QX \\
 & & \text{s.t.} & X_{ii} = 1 \\
 & & & X \succeq 0
\end{array}
$$

Naum Zuselevich Shor (Inst Cybernetics, Kiev) in the 1980s was among the first to recognize the relevance of this approach

Since the feasible set is relaxed = enlarged, we get a lower bound for the original non-convex optimization problem: $p^\star \leq g^\star$

Shor's relaxation

Systematic approach: can be applied to general polynomial optimization problems

Example:

$$x_1^2 x_2 = x_1 \left\{ \begin{array}{c} x_1^2 = x_3 \\ x_3 x_2 = x_1 \end{array} \right. \left\{ \begin{array}{c} X_{11} = X_{30} \\ X_{32} = X_{10} \\ X \succeq 0 \\ \mathrm{rank}\, X = 1 \end{array} \right. \left\{ \begin{array}{c} X_{11} = X_{30} \\ X_{32} = X_{10} \\ X \succeq 0 \end{array} \right.$$

Algorithm:

- introduce lifting variables to reduce polynomials to quadratic and linear terms
- build the rank-one LMI problem
- solve the LMI problem by relaxing the non-convex rank constraint

Relaxed LMI via duality

Consider again the original problem

$$\begin{array}{ll} \text{min} & x^T Q x \\ \text{s.t.} & x_i^2 = 1 \end{array}$$

and build Lagrangian $L(x,y) = x^T Q x - \sum_i y_i(x_i^2 - 1) = x^T(Q - Y)x + \text{trace}\, Y$ where Y is a diagonal matrix and $Q - Y \succeq 0$ must be enforced to ensure that Lagrangian is bounded below

Associated dual problem reads

$$\begin{array}{ll} \text{max} & \text{trace}\, Y \\ \text{s.t.} & Q - Y \succeq 0 \\ & Y \text{ diagonal} \end{array}$$

This is an LMI problem !

Relaxed LMI via duality

The dual LMI problem

$$\begin{array}{ll} \max & \text{trace}\, Y \\ \text{s.t.} & Q \succeq Y \\ & Y \text{ diagonal} \end{array}$$

has for dual the primal LMI problem

$$\begin{array}{ll} \min & \text{trace}\, QX \\ \text{s.t.} & X_{ii} = 1 \\ & X \succeq 0 \end{array}$$

which is Shor's original LMI relaxation !

More generally it can be shown that LMI rank dropping and Lagrangian relaxation are equivalent

Example of LMI relaxation

Original nonconvex 0-1 quadratic problem

$$g^\star = \min \; 2x_1x_2 + 4x_1x_3 + 6x_2x_3 \quad \text{s.t.} \; x_i^2 = 1$$

$$Q = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 2 & 3 & 0 \end{bmatrix}$$

Primal and dual LMI solutions

$$X = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix} \qquad Y = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -5 \end{bmatrix}$$

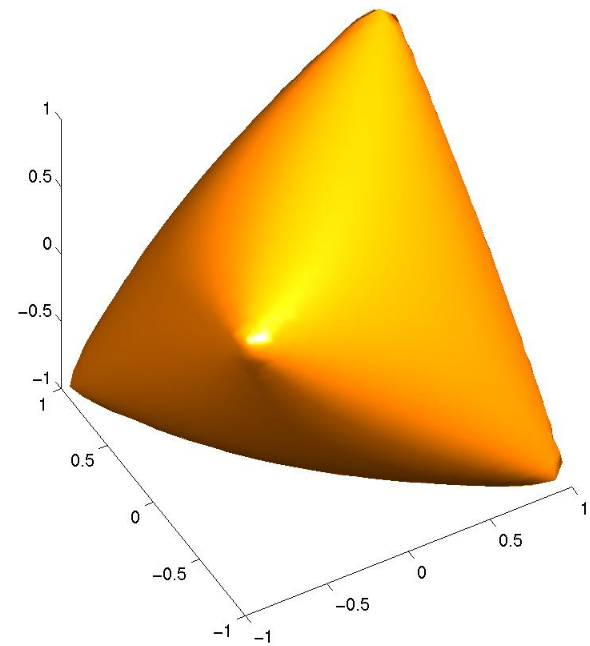
yield lower bound $p^\star = \mathrm{trace}\, QX = d^\star = \mathrm{trace}\, Y = -8$
(strong duality holds here)

Since rank $X = 1$ we recover here the optimum $x = [1 \;\; 1 \;\; -1]^T$
such that $X = xx^T$ and hence $g^\star = p^\star = d^\star$
the relaxation is exact !

Example of LMI relaxation

LMI relaxation of ± 1 constraints

$$X = \begin{bmatrix} 1 & X_{12} & X_{13} \\ X_{12} & 1 & X_{23} \\ X_{13} & X_{23} & 1 \end{bmatrix} \succeq 0\}$$



So we optimize the linear objective function
trace $QX = 2X_{12} + 4X_{13} + 6X_{23}$
and the optimum is a vertex $[1 \;\; -1 \;\; -1]$

How good are LMI relaxations ?

We have seen that we can obtain lower bounds for non-convex polynomial minimization with the help of liftings and relaxations

But can we measure the gap between the global optimum and the relaxed optimum ?

In other words, how much conservative are LMI relaxations ?

Answers only in a (too) few specific cases..

MAXCUT

Given a graph with arcs (i,j) with weights $a_{ij} \geq 0$,
find a partition maximizing total weight of linking arcs

Non-convex quadratic problem

$$g^\star = \max \ \tfrac{1}{4}\sum_{i,j} a_{ij}(1 - x_i x_j) \quad \text{s.t.} \ x_i^2 = 1$$

with convex LMI relaxation

$$d^\star = \max \ \tfrac{1}{4}\sum_{i,j} a_{ij}(1 - X_{ij}) \quad \text{s.t.} \ X_{ii} = 1, \ X = X^T \succeq 0$$

With a geometric proof, Goemans and Williamson showed (1994)
that independently of the data (graph) $1 \geq \dfrac{g^\star}{d^\star} \geq 0.8786$

LMI relaxations for quadratic problems

Non-convex quadratic problem

$$\begin{array}{rcl} g^\star & = & \max \quad x^T A x \\ & & \text{s.t.} \quad x_i^2 = 1 \end{array}$$

with convex LMI relaxation

$$\begin{array}{rcl} d^\star & = & \max \quad \operatorname{trace} AX \\ & & \text{s.t.} \quad X_{ii} = 1 \\ & & \qquad\; X = X^T \succeq 0 \end{array}$$

For $A \succeq 0$ Nesterov showed that

$$1 \geq \frac{g^\star}{d^\star} \geq \frac{2}{\pi} = 0.6366$$

Beyond Shor's relaxation

Recent work (2000) to narrow relaxation gap

- gradually adding lifting variables
- hierarchy of nested LMI relaxations
- theoretical proof of convergence
- tradeoff between conservatism and computational effort

Dual point of views:

- theory of moments (Lasserre)
- sum-of-squares decompositions (Parrilo)

Higher order LMI relaxations
Illustration

Non-convex quadratic problem

$$\begin{array}{ll} \min & g_0(x) = -2x_1^2 - 2x_2^2 + 2x_1x_2 + 2x_1 + 6x_2 - 10 \\ \text{s.t.} & g_1(x) = -x_1^2 + 2x_1 \geq 0 \\ & g_2(x) = -x_1^2 - x_2^2 + 2x_1x_2 + 1 \geq 0 \\ & g_3(x) = -x_2^2 + 6x_2 - 8 \geq 0. \end{array}$$

LMI relaxation built by replacing each monomial $x_1^i x_2^j$ with lifting variable y_{ij}

For example, quadratic expression $g_2(x) = -x_1^2 - x_2^2 + 2x_1x_2 + 1$ is replaced with linear expression $-y_{20} - y_{02} + 2y_{11} + 1$

Lifting variables y_{ij} satisfy non-convex relations such as $y_{10}y_{01} = y_{11}$ or $y_{20} = y_{10}^2$

LMI relaxations: illustration (2)

Relax these non-convex relations by enforcing LMI constraint

$$M_1(y) = \left[\begin{array}{c|cc} 1 & y_{10} & y_{01} \\ \hline y_{10} & y_{20} & y_{11} \\ y_{01} & y_{11} & y_{02} \end{array}\right] \succeq 0$$

Moment matrix of first order relaxing quadratic monomials
You have recognized Shor's relaxation !

First LMI (=Shor's) relaxation of original global optimization problem is given by

$$\begin{array}{ll} \min & -2y_{20} - 2y_{02} + 2y_{11} + 2y_{10} + 6y_{01} - 10 \\ \text{s.t.} & -y_{20} + 2y_{10} \geq 0 \\ & -y_{20} - y_{02} + 2y_{11} + 1 \geq 0 \\ & -y_{02} + 6y_{01} - 8 \geq 0 \\ & M_1(y) \succeq 0 \end{array}$$

LMI relaxations: illustration (3)

To build second LMI relaxation, we must increase size of moment matrix so that it captures expressions of degrees up to 4

Second order moment matrix reads

$$M_2(y) = \left[\begin{array}{c|cc|ccc} 1 & y_{10} & y_{01} & y_{20} & y_{11} & y_{02} \\ \hline y_{10} & y_{20} & y_{11} & y_{30} & y_{21} & y_{12} \\ y_{01} & y_{11} & y_{02} & y_{21} & y_{12} & y_{03} \\ \hline y_{20} & y_{30} & y_{21} & y_{40} & y_{31} & y_{22} \\ y_{11} & y_{21} & y_{12} & y_{31} & y_{22} & y_{13} \\ y_{02} & y_{12} & y_{03} & y_{22} & y_{13} & y_{04} \end{array}\right] \succeq 0$$

LMI relaxations: illustration (4)

Constraints are localized on moment matrices, meaning that original constraint $g_1(x) = -x_1^2 + 2x_1 \geq 0$ becomes localizing matrix constraint

$$M_1(g_1 y) = \left[\begin{array}{c|cc} -y_{20}+2y_{10} & -y_{30}+2y_{20} & -y_{21}+2y_{11} \\ \hline -y_{30}+2y_{20} & -y_{40}+2y_{30} & -y_{31}+2y_{21} \\ -y_{21}+2y_{11} & -y_{31}+2y_{21} & -y_{22}+2y_{12} \end{array}\right] \succeq 0$$

Second LMI feasible set included in first LMI feasible set, thus providing a tighter relaxation

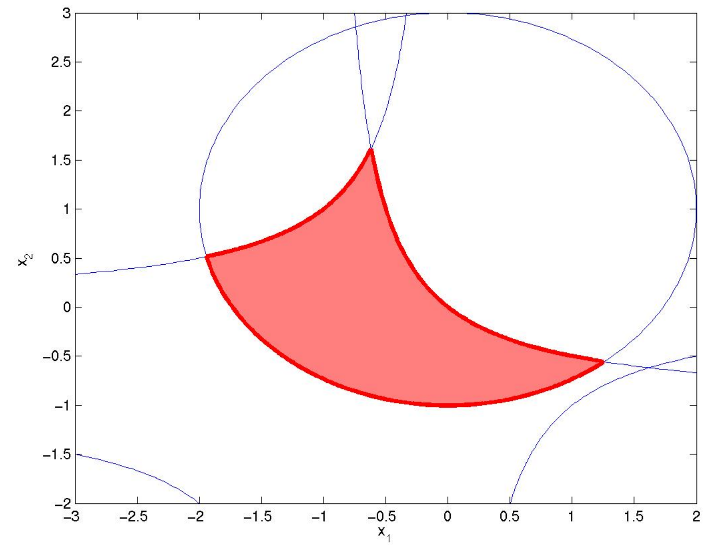
$$\begin{array}{ll} \text{min} & -2y_{20} - 2y_{02} + 2y_{11} + 2y_{10} + 6y_{01} - 10 \\ \text{s.t.} & M_1(g_1 y) \succeq 0, \;\; M_1(g_2 y) \succeq 0, \;\; M_1(g_3 y) \succeq 0 \\ & M_2(y) \succeq 0 \end{array}$$

Similary, we can build up 3rd, 4th, 5th LMI relaxations..

Geometric illustration

Non-convex quadratic problem with linear objective function

$$
\begin{array}{ll}
\max & x_2 \\
\text{s.t.} & 3 - 2x_2 - x_2^1 - x_2^2 \geq 0 \\
 & -x_1 - x_2 - x_1 x_2 \geq 0 \\
 & 1 + x_1 x_2 \geq 0
\end{array}
$$

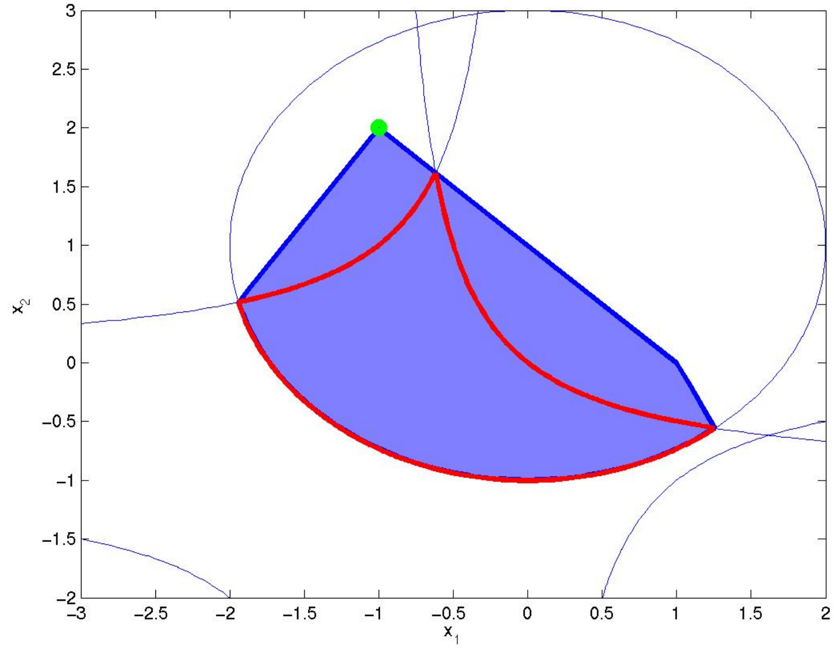


Non-convex feasible set delimited by circular and hyperbolic arcs

Geometric illustration (2)

First LMI relaxation given by

$$\begin{array}{ll} \max & y_{01} \\ \text{s.t.} & \begin{bmatrix} 1 & y_{10} & y_{01} \\ y_{10} & y_{20} & y_{11} \\ y_{01} & y_{11} & y_{02} \end{bmatrix} \succeq 0 \\ & 3 - 2y_{01} - y_{20} - y_{02} \geq 0 \\ & -y_{10} - y_{01} - y_{11} \geq 0 \\ & 1 + y_{11} \geq 0 \end{array}$$



Projection of the LMI feasible set onto the plane y_{10}, y_{01} of first-order moments

LMI optimum = 2 = upper-bound on global optimum

Geometric illustration (3)

To build second LMI relaxation, the moment matrix must capture expressions of degrees up to 4

$$M_2^2(y) = \left[\begin{array}{c|cc|ccc} 1 & y_{10} & y_{01} & y_{20} & y_{11} & y_{02} \\ \hline y_{10} & y_{20} & y_{11} & y_{30} & y_{21} & y_{12} \\ y_{01} & y_{11} & y_{02} & y_{21} & y_{12} & y_{03} \\ \hline y_{20} & y_{30} & y_{21} & y_{40} & y_{31} & y_{22} \\ y_{11} & y_{21} & y_{12} & y_{31} & y_{22} & y_{13} \\ y_{02} & y_{12} & y_{03} & y_{22} & y_{13} & y_{04} \end{array}\right]$$

Constraints are also lifted and relaxed with the help of localization matrices

Geometric illustration (4)

Second LMI provides tighter relaxation

$$\max \quad y_{01}$$

$$\text{s.t.} \quad \begin{bmatrix} 1 & y_{10} & y_{01} & y_{20} & y_{11} & y_{02} \\ y_{10} & y_{20} & y_{11} & y_{30} & y_{21} & y_{12} \\ y_{01} & y_{11} & y_{02} & y_{21} & y_{12} & y_{03} \\ y_{20} & y_{30} & y_{21} & y_{40} & y_{31} & y_{22} \\ y_{11} & y_{21} & y_{12} & y_{31} & y_{22} & y_{13} \\ y_{02} & y_{12} & y_{03} & y_{22} & y_{13} & y_{04} \end{bmatrix} \succeq 0$$

$$\begin{bmatrix} 3-2y_{01}-y_{20}-y_{02} & 3y_{10}-2y_{11}-y_{30}-y_{12} & 3y_{01}-2y_{02}-y_{21}-y_{03} \\ 3y_{10}-2y_{11}-y_{30}-y_{12} & 3y_{20}-2y_{21}-y_{40}-y_{22} & 3y_{11}-2y_{12}-y_{31}-y_{13} \\ 3y_{01}-2y_{02}-y_{21}-y_{03} & 3y_{11}-2y_{12}-y_{31}-y_{13} & 3y_{02}-2y_{03}-y_{22}-y_{04} \end{bmatrix} \succeq 0$$

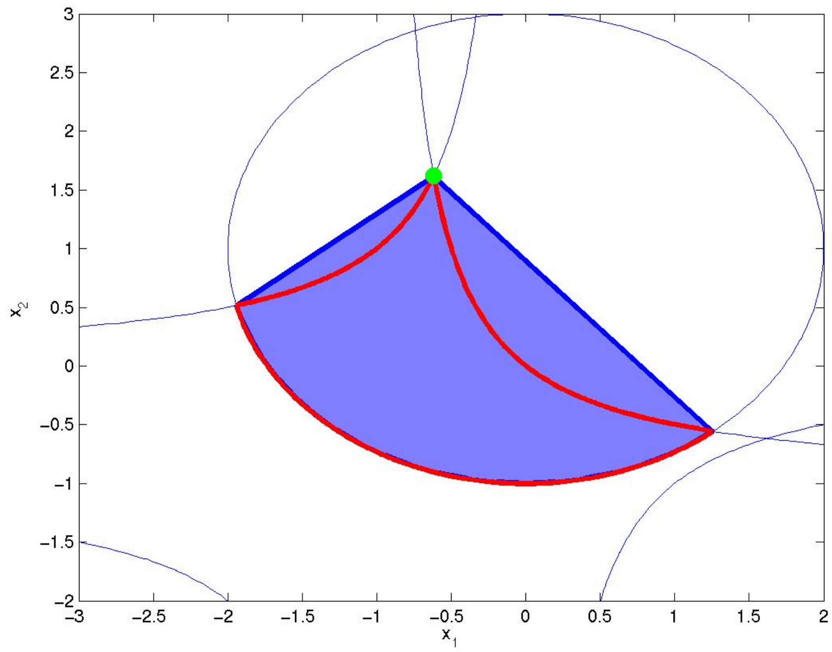
$$\begin{bmatrix} -y_{10}-y_{01}-y_{11} & -y_{20}-y_{11}-y_{21} & -y_{11}-y_{02}-y_{12} \\ -y_{20}-y_{11}-y_{21} & -y_{30}-y_{21}-y_{31} & -y_{21}-y_{12}-y_{22} \\ -y_{11}-y_{02}-y_{12} & -y_{21}-y_{12}-y_{22} & -y_{12}-y_{03}-y_{13} \end{bmatrix} \succeq 0$$

$$\begin{bmatrix} 1+y_{11} & y_{10}+y_{21} & y_{01}+y_{12} \\ y_{10}+y_{21} & y_{20}+y_{31} & y_{11}+y_{22} \\ y_{01}+y_{12} & y_{11}+y_{22} & y_{02}+y_{13} \end{bmatrix} \succeq 0$$

Geometric illustration (5)

Optimal value of 2nd LMI relaxation = 1.6180 = global optimum
Numerical certificate = moment matrix has rank one

First order moments
$(y^*_{10}, y^*_{01}) = (-0.6180, 1.6180)$
provide optimal solution
of original problem



Polynomial multipliers

Polynomial optimization problem

$$
\begin{array}{rcl}
g^\star & = & \min \quad g_0(x) \\
 & & \text{s.t.} \quad g_i(x) \geq 0, \; i = 1, \ldots, m
\end{array}
$$

where $g_i(x)$ are real-valued multivariate polynomials in vector indeterminate $x \in \mathbb{R}^n$

Non-convex problem in general (includes 0-1 or quadratic problems) = difficult problem

If $g^\star$ is the global optimum, polynomial $g_0(x) - g^\star$ is non-negative whenever $g_i(x) \geq 0$

In particular we want to maximize such a lower bound $g^\star$

Polynomial multipliers

The positivity condition is satisfied if we can find polynomials $q_i(x)$ such that

$$g_0(x) - g^\star = q_0(x) + \sum_{i=1}^{m} g_i(x) q_i(x)$$

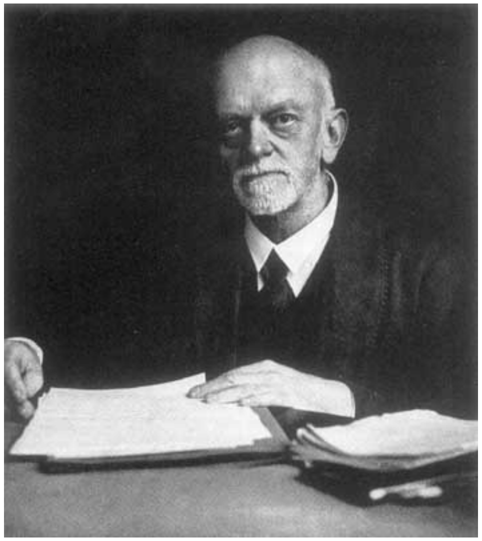
Recall Lagrangian when building dual..

Multipliers $q_i(x)$ are now polynomials !
How can we enforce their positivity ?

SOS polynomials

How can we ensure that a polynomial is globally non-negative ?

$$p(x) \geq 0, \forall x \in \mathbb{R}^n$$



David Hilbert
(1862 Königsberg - 1943 Göttingen)

Hilbert's 17th pb about algebraic sum-of-squares decompositions of rational functions (ICM, Paris, 1900)

SOS polynomials

A form is a homogeneous polynomial, i.e. all monomials have same degree

An obvious condition for a polynomial (form) $p(x)$ to be non-negative is that is a sum-of-squares (SOS) of other polynomials (forms)

$$p(x) = \sum_i q_i^2(x)$$

However, not every non-negative polynomial or form is SOS

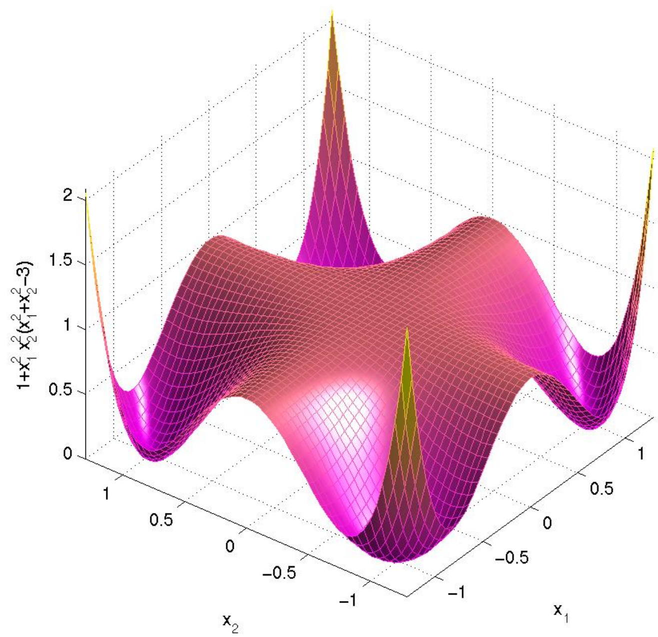
$$p(x) \text{ SOS} \implies p(x) \geq 0$$

Sufficient non-negativity condition only..

Motzkin's polynomial

Counterexample:

$$p(x) = 1 + x_1^2 x_2^2 (x_1^2 + x_2^2 - 3)$$



cannot be written as an SOS but it is globally non-negative (vanishes at $|x_1| = |x_2| = 1$)

SOS polynomials

Let n denote the number of variables and d the degree

Non-negativity and SOS are sometimes equivalent:

$n = 2$	bivariate forms
	univariate polynomials (dehomogen)
$d = 2$	quadratic forms
$n = 3$, $d = 4$	quartic forms of 3 variables

In all other cases, the set of SOS polynomials (a cone) is a subset of the set of non-negative polynomials

We do not know polynomial-time algorithms to check whether a polynomial is non-negative when $d \geq 4$

Note however that the set of SOS polynomials is dense in the set of polynomials nonnegative over the n-dimensional box $[-1, 1]^n$

Most importantly

> The cone of SOS polynomials
> is lifted-LMI representable

as we will see in the sequel..

LMI formulation of SOS polynomials

Polynomial

$$p(x) = \sum_{\alpha} p_\alpha x^\alpha$$

of degree $|\alpha| \leq 2d$ (α = vector of powers) is SOS iff

$$p(x) = z^T X z \quad X \succeq 0$$

where z is a vector with all monomials with degree $\leq d$

Cholesky factorization $X = Q^T Q$ such that

$$\begin{aligned} p(x) &= z^T Q^T Q z = \|Qz\|_2^2 = \textstyle\sum_i (Qz)_i^2 \\ &= \textstyle\sum_i q_i^2(x) \end{aligned}$$

Number of squares $q_i^2(x)$ = rank X

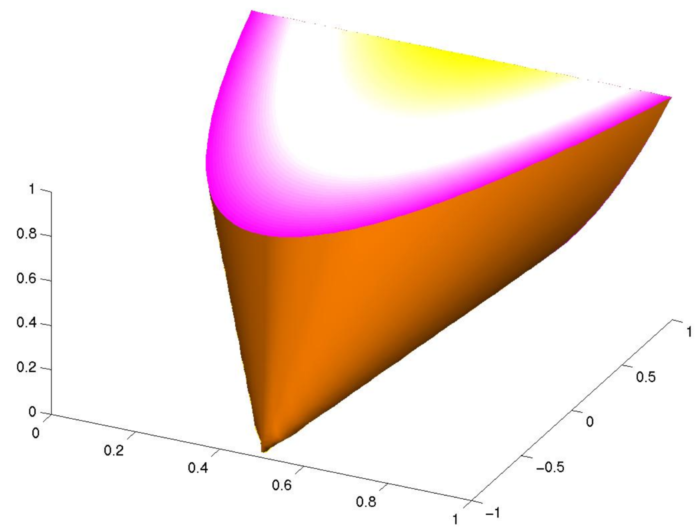
LMI formulation of SOS polynomials

Comparing monomial coefficients in expression

$$p(x) = z^T X z = \sum_{\alpha} p_\alpha x^\alpha \geq 0$$

we get an LMI

$$\text{trace}\, H_\alpha X = p_\alpha \quad \forall \alpha$$
$$X \succeq 0$$



where H_α are Hankel-like matrices

SOS example

Consider the homogeneous form

$$\begin{aligned} p(x) &= 2x_1^4 + 5x_2^4 + 2x_1^3x_2 - x_1^2x_2^2 \\ &= z^T X z \end{aligned}$$

With monomial vector $z = [x_1^2\ x_2^2\ x_1x_2]^T$ a general bivariate form of degree 4 reads

$$z^T X z = X_{11}x_1^4 + X_{22}x_2^4 + 2X_{31}x_1^3x_2 + 2X_{32}x_1x_2^3 + (X_{33} + 2X_{21})x_1^2x_2^2$$

$p(x)$ SOS iff there exists $X \succeq 0$ such that

$$\begin{array}{ll} X_{11} = 2 & X_{22} = 5 \\ 2X_{31} = 2 & 2X_{32} = 0 \\ X_{33} + 2X_{21} = -1 & \end{array}$$

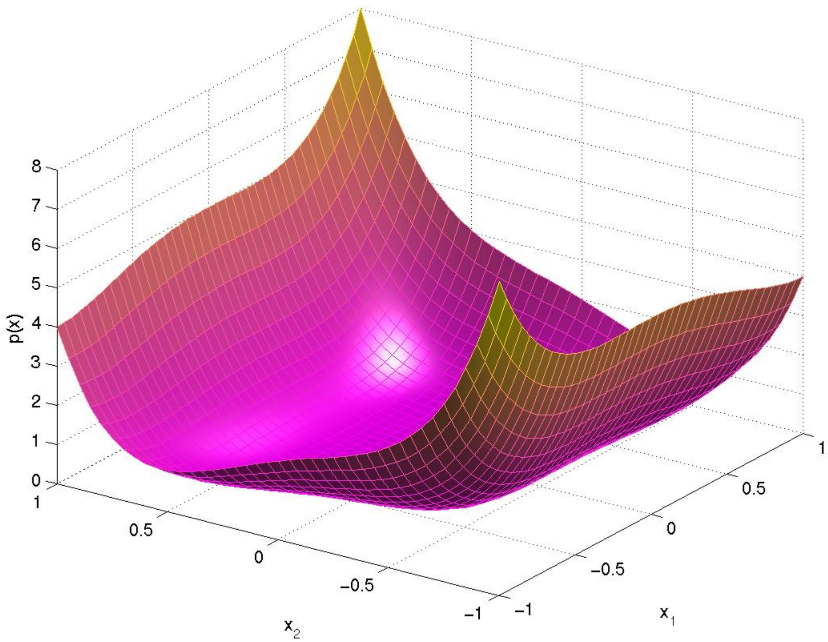
SOS example

One particular solution is

$$X = \begin{bmatrix} 2 & -3 & 1 \\ -3 & 5 & 0 \\ 1 & 0 & 5 \end{bmatrix} = Q^T Q, \;\; Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & -3 & 1 \\ 0 & 1 & 3 \end{bmatrix}$$

So $p(x)$ is the sum of rank $X = 2$ squares

$$\begin{aligned} p(x) &= \tfrac{1}{2}(2x_1^2 - 3x_2^2 + x_1 x_2)^2 \\ &\quad + \tfrac{1}{2}(x_2^2 + 3x_1 x_2)^2 \end{aligned}$$



Finding polynomial multipliers

Returning to our global optimization problem

$$
\begin{array}{rcl}
g^\star & = & \min \quad g_0(x) \\
 & & \text{s.t.} \quad g_i(x) \geq 0,\ i = 1, \ldots, m
\end{array}
$$

the problem of finding SOS polynomials $q_i(x)$ such that

$$p(x) = g_0(x) - g^\star = q_0(x) + \sum_{i=1}^{m} g_i(x) q_i(x)$$

can be formulated as an LMI as soon as the degrees of the $q_i(x)$ are fixed

Depending on parity let $\deg p(x) = 2k - 1$ or $2k$; then the LMI problem of finding an SOS $p(x)$ is referred to as the LMI relaxation of order k

Hierarchy of LMI relaxations

The LMI relaxation of order k reads

$$
\begin{array}{rcll}
d_k^\star & = & \min & \sum_\alpha (g_0)_\alpha y_\alpha \\
 & & \text{s.t.} & M_k(y) = \sum_\alpha A_\alpha y_\alpha \succeq 0 \\
 & & & M_{k-d_i}(g_i y) = \sum_\alpha A_\alpha^{g_i} y_\alpha \succeq 0 \quad \forall i
\end{array}
$$

with $y_0 = 1$ (normalization), d_i is half the degree of $g_i(x)$, $M_k(y)$ is the moment matrix, $M_{k-d_i}(g_i y)$ are the localizing matrices

The dual LMI

$$
\begin{array}{rcll}
p_k^\star & = & \max & \text{trace}\, A_0 X + \sum_i \text{trace}\, A_0^{g_i} X_i \\
 & & \text{s.t.} & \text{trace}\, A_\alpha X + \sum_i \text{trace}\, A_\alpha^{g_i} X_i = (g_0)_\alpha \quad \forall \alpha \neq 0
\end{array}
$$

corresponds to the condition $p(x)$ SOS

Hierarchy of LMI relaxations

If feasible set $g_i(x) \geq 0$ is compact, and under mild additional assumptions, Lasserre could use results by Putinar (on SOS representations of positive polynomials) and Curto/Fialkow (on flat extension of moment matrices) to prove in 2000 that

$$p_k^\star = d_k^\star \leq g^\star$$

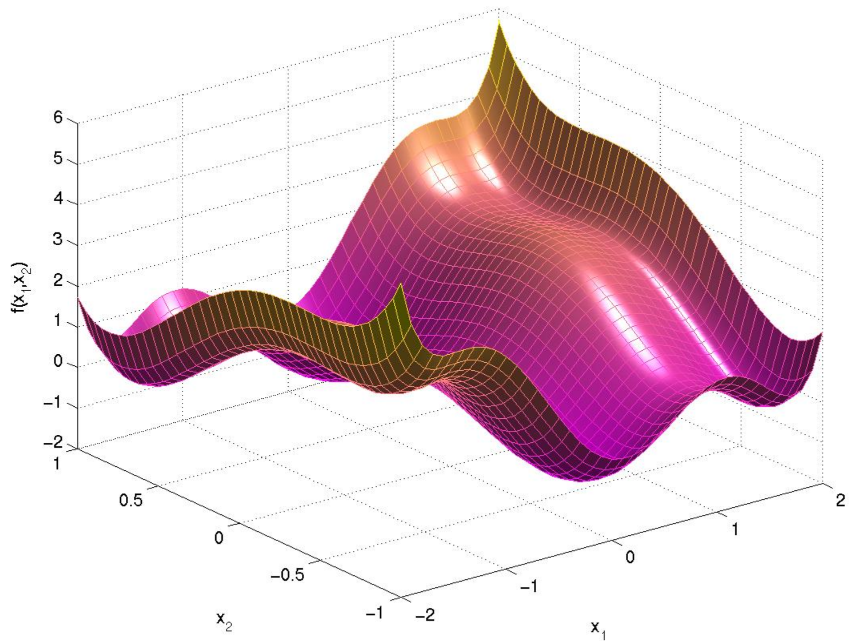
with asymptotic convergence guarantee

$$\lim_{k\to\infty} p_k^\star = g^\star$$

Moreover, in practice, convergence is fast:
$p_k^\star$ is very close to $g^\star$ for small k

Camelback function

For the six-hump camelback function



with two global optima and six local optima, the global optimum is reached at the first LMI relaxation ($k = 1$) without any problem splitting

LMI hierarchy: example

Quadratic problem

$$\begin{array}{ll} \min & -2x_1 + x_2 - x_3 \\ \text{s.t.} & x_1(4x_1 - 4x_2 + 4x_3 - 20) + x_2(2x_2 - 2x_3 + 9) \\ & \qquad +x_3(2x_3 - 13) + 24 \geq 0 \\ & x_1 + x_2 + x_3 \leq 4, \quad 3x_2 + x_3 \leq 6 \\ & 0 \leq x_1 \leq 2, \quad 0 \leq x_2, \quad 0 \leq x_3 \leq 3. \end{array}$$

Computational burden increases quickly with relaxation order

order	1	2	3	4	5	6
bound	-6.0000	-5.6923	-4.0685	-4.0000	-4.0000	-4.0000
`size(y)`	9	34	83	164	285	454

..yet fourth LMI relaxation solves globally the problem

Complexity

d: overall polynomial degree ($2\delta = d$ or $d+1$)
m: number of polynomial constraints
n: number of polynomial variables
M: number of primal variables (moments)
N: number of dual variables (LMI size)

$$
\begin{aligned}
M &= \binom{n+2\delta}{2\delta} - 1 \\
N &= \binom{n+\delta}{\delta} + m\binom{n+\delta-1}{\delta-1}
\end{aligned}
$$

When n is fixed:

- M grows polynomially in $O(\delta^n)$
- N grows polynomially in $O(m\delta^n)$

Solving BMIs with LMI relaxations

Two approaches: scalarization or PMI relaxations

Scalarization:
- scalarize using characteristic polynomial
- polynomials with generally large degree

PMI relaxations:
- keep the matrix structure
- no degree growth
- theory for matrix polynomial SOS

Theory is ready, but experimentally at a very preliminary level

Numerical aspects (conditioning, solution extraction)
must be studied further

LMI modelling of convex hulls

Using the same technique, and the equivalence between nonnegative and SOS polynomials in specific cases, we can build lifted-LMI representations for convex hulls of rationally parametrized curves and surfaces

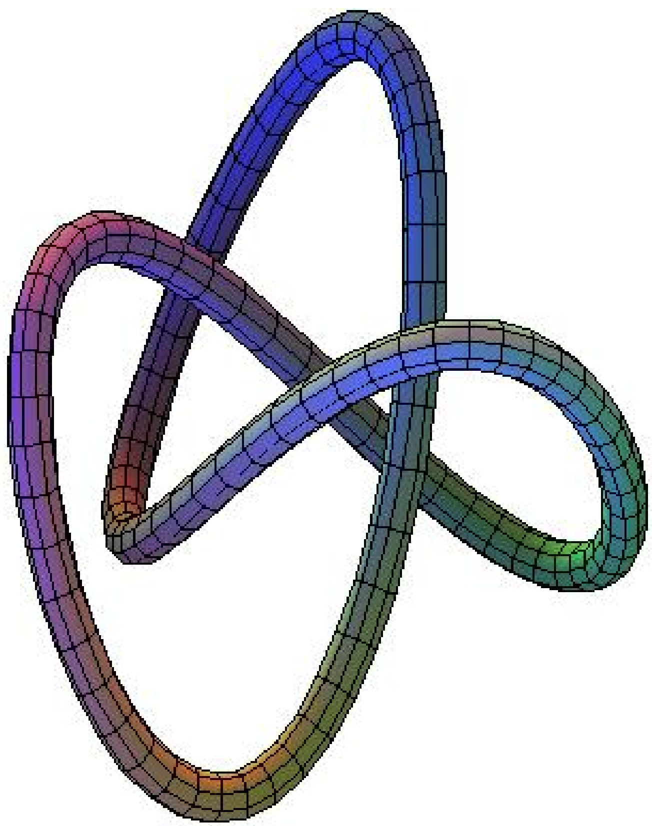
$$\{x \in \mathbb{R}^n \,:\, x_i = \frac{p_i(t)}{p_0(t)}\}$$

with given polynomials $p_0(t), p_1(t), \ldots, p_n(t)$

- $t \in \mathbb{R}$, any degree
- $t \in \mathbb{R}^2$, quartics $p_i(t)$
- quadratics $p_i(t)$

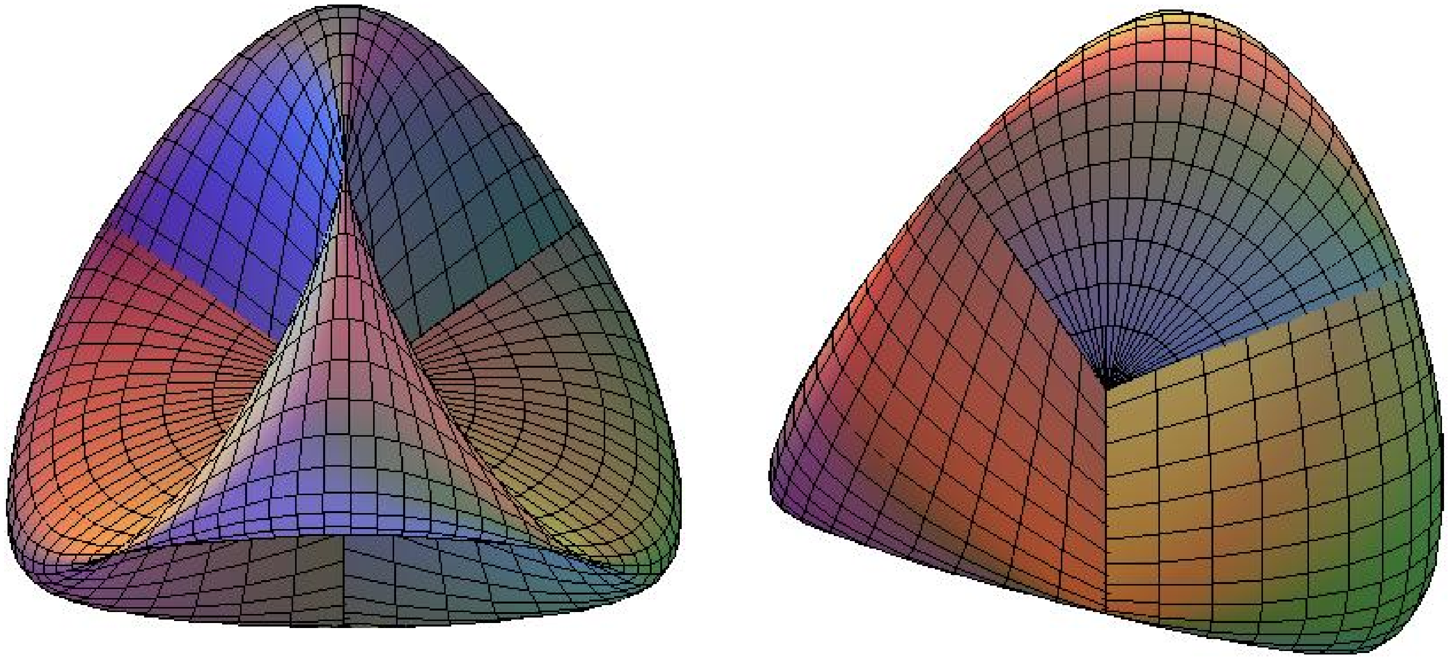
Ambient space dimension n is arbitrary

Trefoil knot curve



Convex hull lifted-LMI with 3 liftings

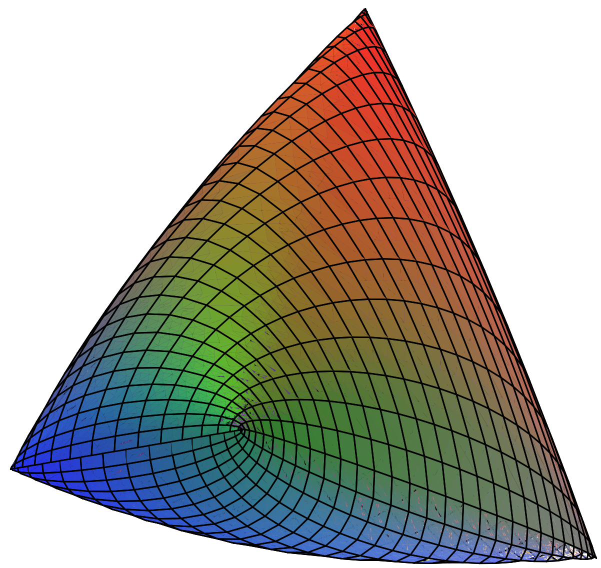
Steiner's Roman surface



Convex hull lifted-LMI with 2 liftings

Cayley's cubic surface

Projectively dual to Steiner's Roman surface



Lifted-LMI representable as an 6-by-6 LMI with 11 liftings..
yet we have another explicit 3-by-3 LMI with no lifting !

LMI relaxations: conclusion

LMI relaxations prove useful to solve general non-convex polynomial optimization problems

Shor's relaxation = rank dropping = Lagrangian relaxation = first order LMI relaxation

Sometimes one can measure the gap between the original problem and its relaxation

A hierarchy of successive LMI relaxations can be built with additional lifting variables and constraints

Theoretical guarantee of asymptotic convergence to global optimum without any problem splitting (no branch and bound scheme)