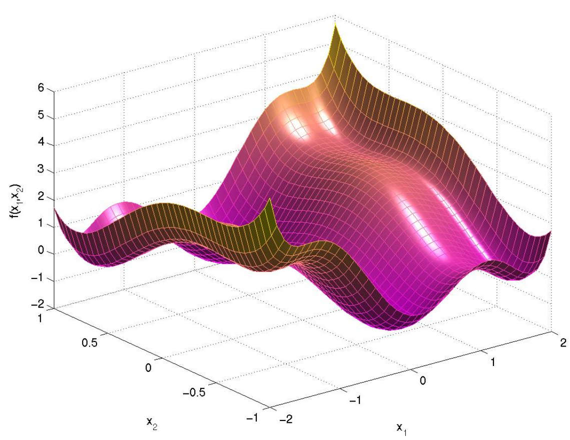
COURSE ON LMI OPTIMIZATION
WITH APPLICATIONS IN CONTROL
PART I.4

LMI RELAXATIONS

Denis Arzelier
www.laas.fr/~arzelier
arzelier@laas.fr

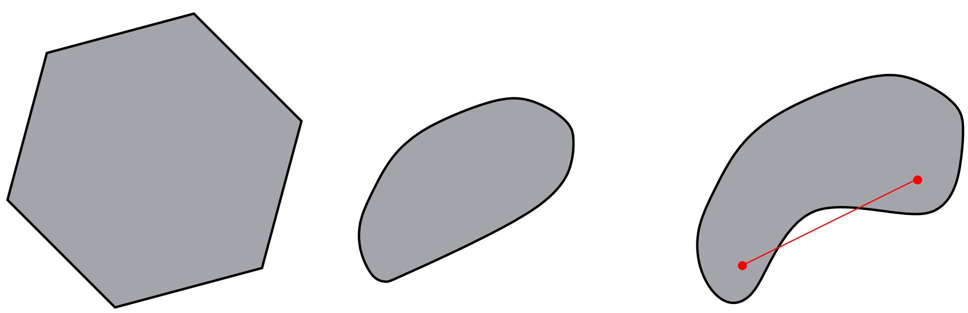


January 2005

Handling nonconvexity

So far we have studied convex LMI sets

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

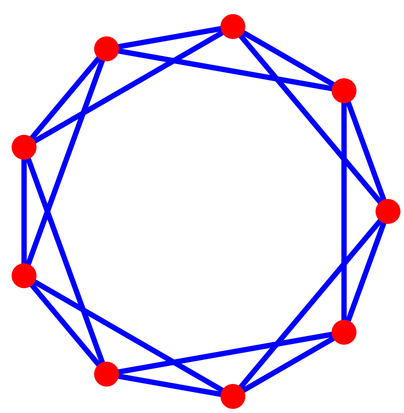
- convex LMI relaxations, and
- liftings = additional variables

Example of combinatorial optimization (1)

Typical combinatorial optimization problem

$$\begin{array}{ll} \min & x'Qx \\ \text{s.t.} & x_i \in \{-1, 1\} \end{array}$$

Examples: MAXCUT, knapsack..



Antiweb AW_9^2 graph

Basic non-convex constraints

$$x_i^2 = 1$$

Exponential # of points = NP-hard problem !

Example of combinatorial optimization (2)
LMI relaxation (1)

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

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

Not bad idea, but we can do better..

Example of combinatorial optimization (3)
LMI relaxation (2)

Replace all non-convex constraints

$$x_i^2 = 1, \quad i = 1, 2, \ldots, n$$

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

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

for each $i = 1, 2, \ldots, n$

Example of combinatorial optimization (4)
Rank constrained LMI (1)

In the original problem

$$\begin{array}{rcll} g^\star & = & \min & x'Qx \\ & & \text{s.t.} & x_i^2 = 1 \end{array}$$

Let $X = xx'$ then

$$x'Qx = \text{trace}\, Qxx' = \text{trace}\, QX$$

and

$$x_i^2 = X_{ii} = 1$$

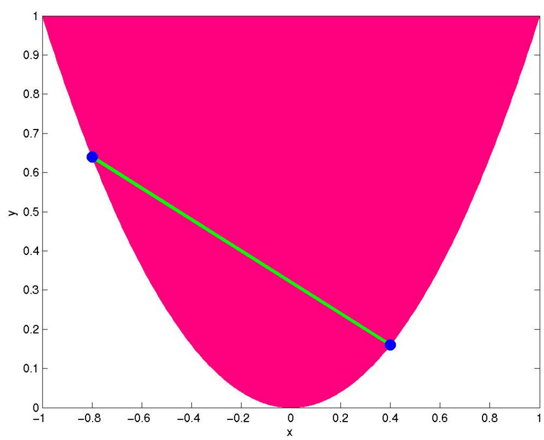
so that the problem can be written as a rank constrained LMI

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

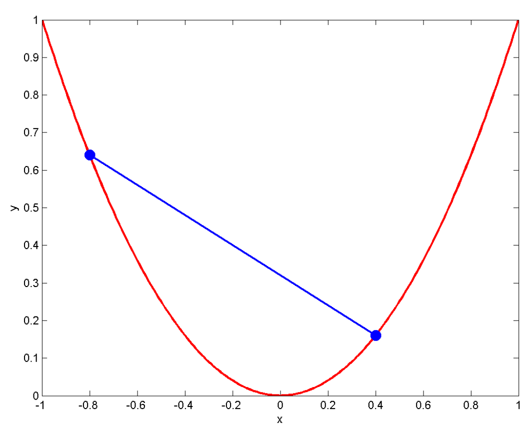
Remember introduction on combinatorial optimization!

Example of combinatorial optimization (5)
Rank constrained LMI (2)

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



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



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

Example of combinatorial optimization (6)
Rank constrained LMI (3)

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

The obtained LMI relaxation is called Shor's relaxation

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

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:

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

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

LMI relaxation and Lagrangian duality (1)

Consider again the original problem

$$\begin{array}{ll} \text{min} & x'Qx \\ \text{s.t.} & x_i^2 = 1 \end{array}$$

and build Lagrangian

$$\begin{array}{rcl} L(x,y) & = & x'Qx - \sum_i y_i(x_i^2 - 1) \\ & = & x'(Q-Y)x + \text{trace}\, Y \end{array}$$

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 !

LMI relaxation and Lagrangian duality (2)

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
=
Lagrangian relaxation

Lagragian duality is a very general tool to built LMI relaxations

Beyond Shor's relaxation

Recent work (2000) to narrow relaxation gap

- gradually adding lifting variables
- hierarchy of nested LMI relaxations
- theoretical proof of convergence

Dual point of views:

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

Tradeoff between conservatism and computational effort

Optimizing with polynomials

Let the polynomial optimization problem

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

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) = NP-hard

Notation

$$\mathbb{P} = \{x \in \mathbb{R}^n \mid g_i \geq 0,\ i = 1,\ldots,m\}$$

Consider the problem without constraints

Since $g^\star$ is the global optimum, polynomial

$$\boxed{g_0(x) - g^\star \geq 0}$$

must be globally positive (non-negative)

Polynomial non negativity

$p \in \mathbb{R}[x_1, \cdots, x_n]$ is globally non negative iff

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

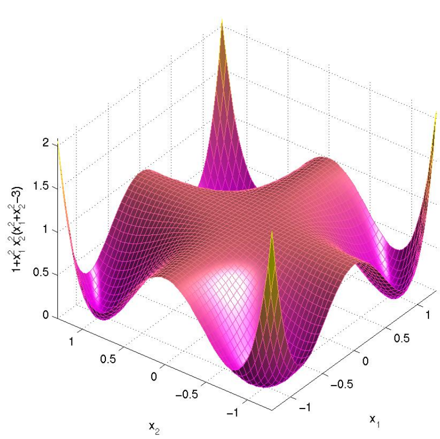
p is called positive semidefinite or PSD

- The set of PSD polynomials of degree $\leq d$

$$\mathcal{P}_n^d = \{p \in \mathbb{R}[x_1, \cdots, x_n] \mid p \text{ is PSD}\}$$

is a convex cone in $\mathbb{R}^N$ where $N = \begin{pmatrix} n+d \\ d \end{pmatrix}$

- Testing if a particular $p \in \mathcal{P}_n^d$ is NP-hard



Motzkin's polynomial

SOS polynomials

A polynomial $p \in \mathbb{R}[x_1, \cdots, x_n]$ is called a sum-of-squares (SOS) if

$$p(x) = \sum_{i=1}^{r} q_i(x)^2$$

for some polynomials $q_1, \cdots, q_r$ and some $r \geq 0$

- The set of SOS polynomials of degree $\leq d$

$$\mathcal{S}_n^d = \{p \in \mathbb{R}[x_1, \cdots, x_n] \mid p \text{ is SOS}\}$$

is a convex cone in $\mathbb{R}^N$ where $N = \binom{n+d}{d}$

- $\mathcal{S}_n^d \subset \mathcal{P}_n^d$ and testing if a particular $p \in \mathcal{S}_n^d$ is an SDP

Condition for $p(x) \in \mathcal{P}_n^d$ is there exist polynomials $q_i(x)$ s.t.

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

Sufficient non-negativity condition only..

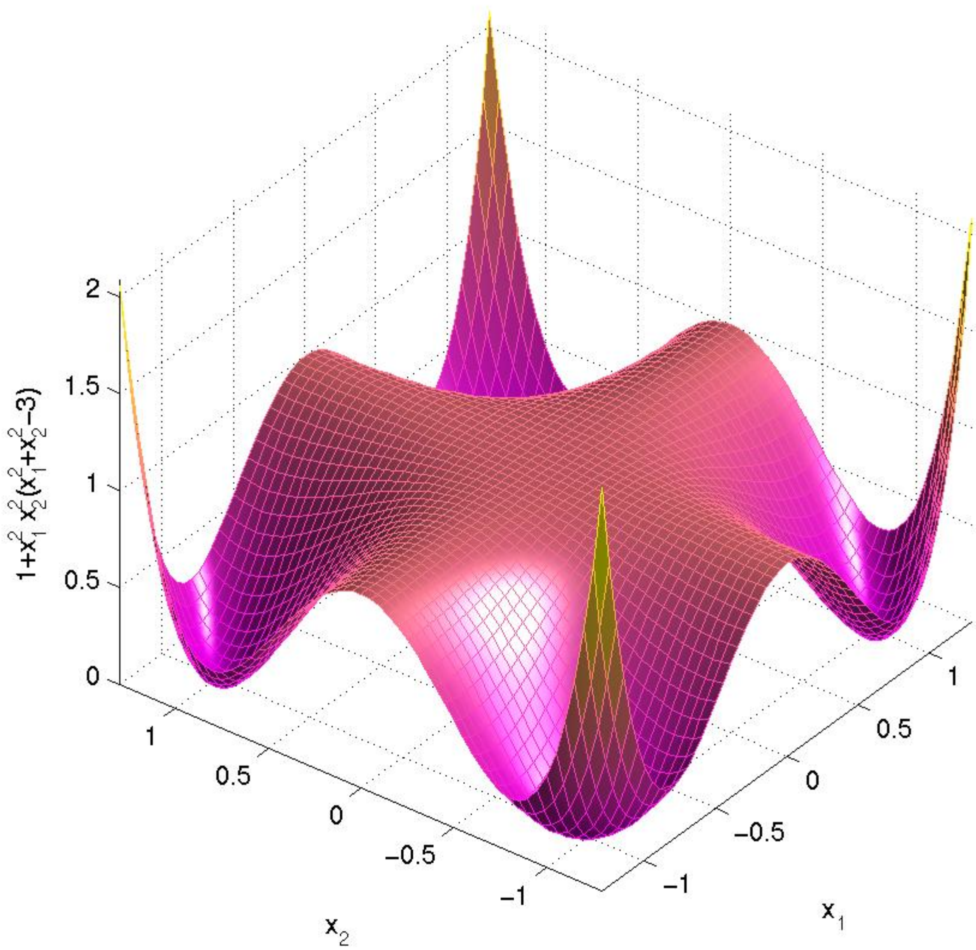
$$p(x) \text{ SOS} \Longrightarrow p(x) \text{ PSD}$$

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 $|x_1| = |x_2| = 1$)



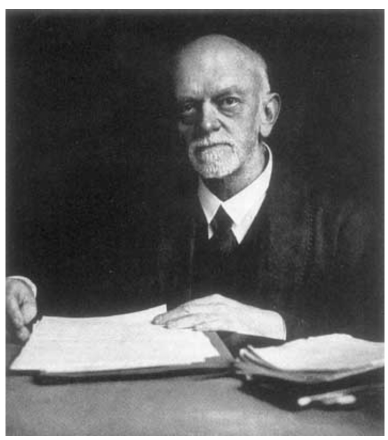
PSD and SOS polynomials

Let n denote the number of variables and d the degree

In 1888, David Hibert showed that $\mathcal{P}_n^d = \mathcal{S}_n^d$ iff

$n = 1$	univariate polynomials	$d = 2, 4 \cdots$
$n = 2$	bivariate polynomials	$d = 2, 4$
n	quadratic forms	$d = 2$

Hilbert's 17th pb about algebraic sum-of-squares decompositions of rational functions (solved by Artin)



David Hilbert
(1862 Königsberg - 1943 Göttingen)

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

LMI formulation of SOS polynomials (1)

Polynomial

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

of degree $|\alpha| \leq 2d$ (α = vector of powers of indeterminates x) is SOS iff $\exists$ X s.t.

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

where z is a vector with all monomials with degree not greater than d

For a feasible X, Cholesky factorization

$$X = Q'Q \quad Q' = \left[\; q_1, \quad \cdots \quad , q_r \; \right]$$

such that

$$\begin{aligned} p(x) &= z'Q'Qz = \|Qz\|_2 = \sum_{i=1}^{r} (q_i' z)^2 \\ &= \sum_{i=1}^{r} q_i^2(x) \end{aligned}$$

Number of squares r = rank X

LMI formulation of SOS polynomials (2)

Comparing monomial coefficients in expression

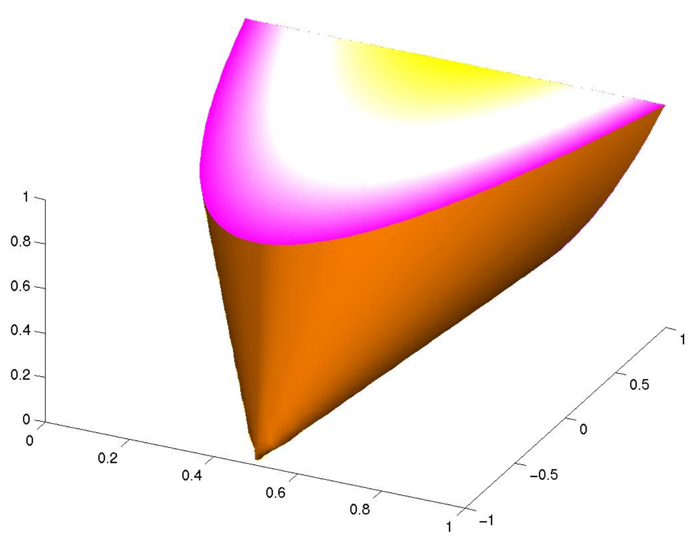
$$p(x) = z'Xz = \sum_{\alpha} p_{\alpha} x^{\alpha} \geq 0$$

we get an LMI

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

where H_{α} are Hankel-like matrices

SOS polynomials described by an intersection between a subspace and the PSD cone



LMI formulation of SOS polynomials (3)
Example (1)

Consider the homogeneous form

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

With monomial vector

$$z = \begin{bmatrix} x_1^2 & x_2^2 & x_1x_2 \end{bmatrix}'$$

A general bivariate form of degree 4 reads

$$\begin{aligned} z'Xz &= X_{11}x_1^4 + X_{22}x_2^4 + 2X_{31}x_1^3x_2 \\ &\quad +2X_{32}x_1x_2^3 + (X_{33} + 2X_{21})x_1^2x_2^2 \end{aligned}$$

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

$X_{11} = 2$	$X_{22} = 5$
$2X_{31} = 2$	$2X_{32} = 0$
$X_{33} + 2X_{21} = -1$	

LMI formulation of SOS polynomials (4) Example (2)

One particular solution is

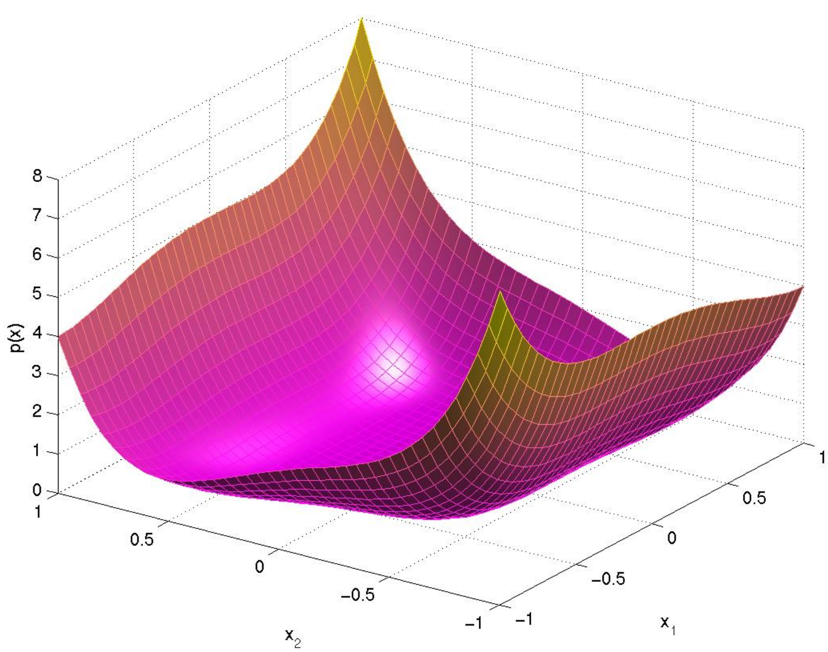
$$X = \begin{bmatrix} 2 & -3 & 1 \\ -3 & 5 & 0 \\ 1 & 0 & 5 \end{bmatrix} = Q'Q$$

with Cholesky factor

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

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

$$p(x) = \tfrac{1}{2}(2x_1^2 - 3x_2^2 + x_1x_2)^2 + \tfrac{1}{2}(x_2^2 + 3x_1x_2)^2$$



Parameterized SOS (1)

Consider the 4th-degree univariate polynomial

$$p(x) = 1 + 2ax + x^2 + 2bx^3 + x^4$$

parameterized in $(a,\ b) \in \mathbb{R}^2$

Which values of a and b make $p(x)$ non-negative or equivalently SOS ?

Primal LMI

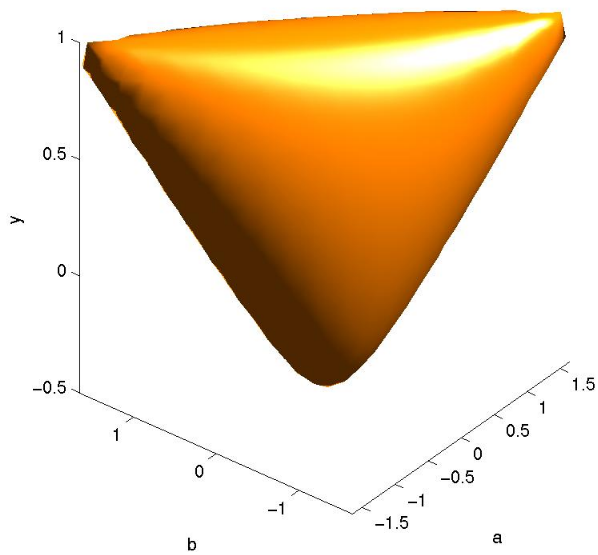
$$\text{trace}\, H_i X = p_i(a,b)$$
$$X \succeq 0$$

with H_i Hankel matrices and corresponding reduced LMI (null-space parameterization)

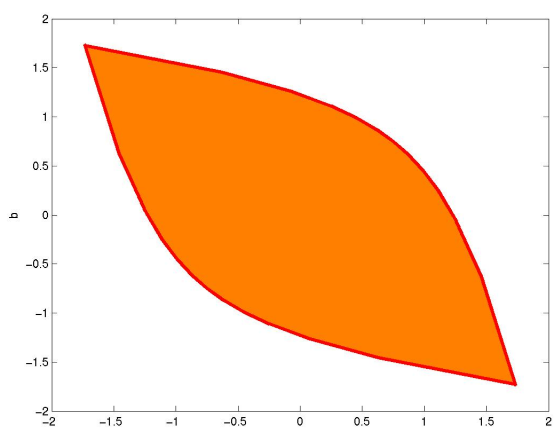
$$\begin{bmatrix} 1 & a & -y \\ a & 1+2y & b \\ -y & b & 1 \end{bmatrix} \succeq 0$$

Parametrized SOS (2)

For $y = 0$, $p(x)$ is SOS iff $a^2 + b^2 \leq 1$



For other values, LMI set in 3D space $(a,\, b,\, y)$



Projection in the plane $(a,\, b)$

Global optimization over polynomials (1)

Returning to our global optimization problem

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

Since g^* is a global minimizer of g_0 on $\mathbb{P}$, if there exist SOS polynomials $q_i(x)$, $i = 0, \cdots, m$ s.t.

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

then

$$p(x) = g_0(x) - g^* \geq 0 \quad \forall\ x \in \mathbb{P}$$

Remember Lagrangian with SOS polynomials multipliers $q_i(x)$

Finding SOS polynomial multipliers $q_i(x)$ s.t.

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

is LMI problem when the degrees of $q_i(x)$ are fixed

Global optimization over polynomials (2)
Hierarchy of LMI relaxations (1)

For ($\deg p(x) = 2k$), the LMI problem of finding an SOS $p(x)$ is referred to as the LMI relaxation of order k

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

with $y_0 = 1$, $d_i = \deg(g_i(x))/2$,

$M_k(y)$ is the moment matrix,

$M_{k-d_i}(g_i y)$ are the localization matrices

The dual LMI

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

corresponds to $p(x)$ SOS

Global optimization over polynomials (3)
Hierarchy of LMI relaxations (2)

If $\mathbb{P}$ is compact (polytope) and there exists $u(x) \in \mathbb{R}[x_1, \cdots, x_n]$, s.t.

1 − $\{u(x) \geq 0\}$ is compact

2 − $u(x) = u_0(x) + \sum_{i=1}^{m} g_i(x) u_i(x) \quad \forall\, x \in \mathbb{R}^n$

where $u_i(x) \in \mathcal{S}_n^l,\ i = 0, \cdots, m$, Lasserre proved 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

Global optimization over polynomials (4)
Hierarchy of LMI relaxations: Example (1)

Non-convex quadratic problem

$$\begin{array}{ll} \min & h_0(x) = -2x_1^2 - 2x_2^2 + 2x_1x_2 + 2x_1 + 6x_2 - 10 \\ \text{s.t.} & g_1(x) = -x_1^2 + 2x_1 \geq 0 \\ & g_2(x) = -x_1^2 - x_2^2 + 2x_1x_2 + 1 \geq 0 \\ & g_3(x) = -x_2^2 + 6x_2 - 8 \geq 0. \end{array}$$

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

For example, quadratic expression

$$g_2(x) = -x_1^2 - x_2^2 + 2x_1x_2 + 1 \geq 0$$

is replaced with linear expression

$$-y_{20} - y_{02} + 2y_{11} + 1 \geq 0$$

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

Global optimization over polynomials (5)
Hierarchy of LMI relaxations: Example (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 monomials of degree up to 2

You have recognized Shor's relaxation !

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

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

Global optimization over polynomials (6)
Hierarchy of LMI relaxations: Example (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

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

Constraints are localized on moment matrices, meaning that original constraint

$$g_1(x) = -x_1^2 + 2x_1 \geq 0$$

becomes localizing matrix constraint

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

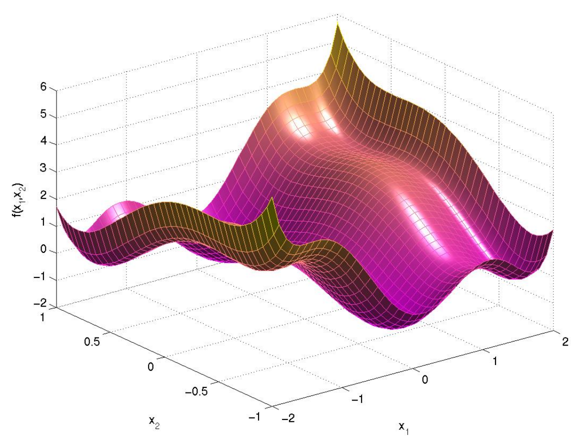
Global optimization over polynomials (7)
Hierarchy of LMI relaxations: Example (4)

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..

For the well-known six-hump camelback function



with two global optima and six local optima, the global optimum is reached at the first LMI relaxation ($k = 1$)

Global optimization over polynomials (8)
Hierarchy of LMI relaxations: Example (5)

Quadratic problem

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

Number of LMI variables (M) and size of relaxed LMI problem (N) increase quickly with relaxation order:

Relaxation	LMI opt	M	N
1	-6.0000	9	24
2	-5.6923	34	228
3	-4.0685	83	1200
4	-4.0000	164	4425
5	-4.0000	285	12936
6	-4.0000	454	32144

..yet fourth LMI relaxation solves globally the problem

Global optimization over polynomials (9)
Hierarchy of LMI relaxations: Complexity

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

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

When n is fixed:

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

Global optimization over polynomials (8)
Hierarchy of LMI relaxations: Conclusions

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)