Sums of squares relaxations for Convex Hulls of Semialgebraic Sets

João Gouveia

University of Washington

September 30th 2010

IPAM Workshop: Convex Optimization and Algebraic Geometry

Acknowledgments

Work done with:

Rekha Thomas - University of Washington

Pablo Parrilo - MIT

Tim Netzer - University of Leipzig

Optimization over algebraic sets

Problem

$$\begin{aligned} \max_{x \in \mathbb{R}^n} \quad & \langle c, x \rangle \\ \text{s.t.} \quad & h_i(x) = 0, \quad i = 1, \cdots, k, \end{aligned}$$

where the h_i are polynomials over $\mathbb{R}$.

Optimization over algebraic sets

Problem

$$\max_{x\in\mathbb{R}^n} \quad \langle c, x\rangle$$
$$\text{s.t.} \quad h_i(x) = 0, \quad i = 1, \cdots, k,$$

where the h_i are polynomials over $\mathbb{R}$.

Let $I = \langle h_1, ..., h_k\rangle$ and $\mathcal{V}_{\mathbb{R}}(I)$ the set of real solutions of the system.

Optimization over algebraic sets

Problem

$$\begin{aligned} \max_{x \in \mathbb{R}^n} \quad & \langle c, x \rangle \\ \text{s.t.} \quad & h_i(x) = 0, \quad i = 1, \cdots, k, \end{aligned}$$

where the h_i are polynomials over $\mathbb{R}$.

Let $I = \langle h_1, ..., h_k \rangle$ and $\mathcal{V}_{\mathbb{R}}(I)$ the set of real solutions of the system. We can rewrite the problem as

Problem*

$$\begin{aligned} \min_{\lambda \in \mathbb{R}} \quad & \lambda \\ \text{s.t.} \quad & \lambda - \langle c, x \rangle \geq 0, \quad \forall x \in \mathcal{V}_{\mathbb{R}}(I). \end{aligned}$$

Sums of squares modulo an ideal

As usual we use sums of squares as an algebraic stand-in for nonnegativity.

Sums of squares modulo an ideal

As usual we use sums of squares as an algebraic stand-in for nonnegativity.

Definition

Let $f \in \mathbb{R}[x]$ we say that f is **sos modulo** I if there are polynomials $p_i \in \mathbb{R}[x]$ such that

$$f = \sum p_i^2 + g,$$

for $g \in I$.

Sums of squares modulo an ideal

As usual we use sums of squares as an algebraic stand-in for nonnegativity.

Definition

Let $f \in \mathbb{R}[x]$ we say that f is **sos modulo** I if there are polynomials $p_i \in \mathbb{R}[x]$ such that

$$f = \sum p_i^2 + g,$$

for $g \in I$. If all the degrees of the p_i's are smaller than k, we say that f is k**-sos modulo** I.

Sums of squares modulo an ideal

As usual we use sums of squares as an algebraic stand-in for nonnegativity.

Definition

Let $f \in \mathbb{R}[x]$ we say that f is **sos modulo I** if there are polynomials $p_i \in \mathbb{R}[x]$ such that

$$f = \sum p_i^2 + g,$$

for $g \in I$. If all the degrees of the p_i's are smaller than k, we say that f is k**-sos modulo I**.

- Verifying the k-sos property can be done efficiently given a 'nice' basis for the quotient space $\mathbb{R}[x]/I$.

Sums of squares modulo an ideal

As usual we use sums of squares as an algebraic stand-in for nonnegativity.

Definition

Let $f \in \mathbb{R}[x]$ we say that f is **sos modulo** I if there are polynomials $p_i \in \mathbb{R}[x]$ such that

$$f = \sum p_i^2 + g,$$

for $g \in I$. If all the degrees of the p_i's are smaller than k, we say that f is k**-sos modulo** I.

- Verifying the k-sos property can be done efficiently given a 'nice' basis for the quotient space $\mathbb{R}[x]/I$.
- We can now approximate the original program by a hierarchy of relaxations.

Sums of squares relaxations

Recall we had the problem

Problem*

$$\begin{aligned} \min_{\lambda \in \mathbb{R}} \quad & \lambda \\ \text{s.t.} \quad & \lambda - \langle c, x \rangle \geq 0, \quad \forall x \in \mathcal{V}_{\mathbb{R}}(I). \end{aligned}$$

Sums of squares relaxations

Recall we had the problem

Problem*

$$\begin{aligned} \min_{\lambda \in \mathbb{R}} \quad & \lambda \\ \text{s.t.} \quad & \lambda - \langle c, x \rangle \geq 0, \quad \forall x \in \mathcal{V}_{\mathbb{R}}(I). \end{aligned}$$

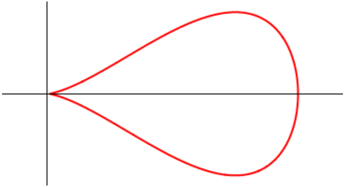
We can now replace it by the hierarchy

Problem-SOS$_k$

$$\begin{aligned} \min_{\lambda \in \mathbb{R}} \quad & \lambda \\ \text{s.t.} \quad & \lambda - \langle c, x \rangle \text{ is } k\text{-sos modulo } I. \end{aligned}$$

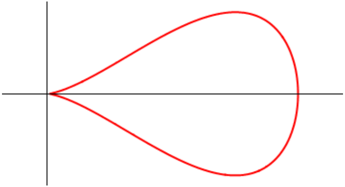
Example

Consider the ideal $I = \langle x^4 - x^3 + y^2 \rangle$ and its variety.



Example

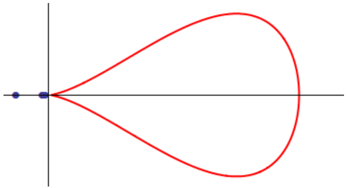
Consider the ideal $I = \langle x^4 - x^3 + y^2 \rangle$ and its variety.



We want to minimize x in this variety.

Example

Consider the ideal $I = \langle x^4 - x^3 + y^2 \rangle$ and its variety.



We want to minimize x in this variety. The relaxation gives us

	$k = 1$	$k = 2$	$k = 3$	$k = 4$
λ_{sos_k}	$-\infty$	$-.1250$	$-.0208$	$-.0091$

.

Theta Bodies

The geometric set underlying these relaxations is called the k-**th theta body** of the ideal I.

Definition

$$\mathrm{TH}_k(I) := \bigcap_{\ell \text{ linear }, \ell \ k\text{-sos modulo } I} \{x \in \mathbb{R}^n : \ell(x) \geq 0\}$$

Theta Bodies

The geometric set underlying these relaxations is called the k**-th theta body** of the ideal I.

Definition

$$\mathrm{TH}_k(I) := \bigcap_{\ell \text{ linear }, \ell \ k\text{-sos modulo } I} \{x \in \mathbb{R}^n : \ell(x) \geq 0\}$$

One can see it alternatively has a relaxation of the closure of the convex hull of the variety.

Theta Bodies

The geometric set underlying these relaxations is called the *k*-**th theta body** of the ideal I.

Definition

$$\mathrm{TH}_k(I) := \bigcap_{\ell \text{ linear }, \ell \ k\text{-sos modulo } I} \{x \in \mathbb{R}^n : \ell(x) \geq 0\}$$

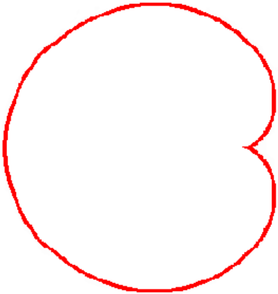
One can see it alternatively has a relaxation of the closure of the convex hull of the variety.

Convex Hull

$$\mathrm{cl}(\mathrm{conv}(\mathcal{V}_{\mathbb{R}}(I))) = \bigcap_{\ell \text{ linear }, \ell|_{\mathcal{V}_{\mathbb{R}}(I)} \geq 0} \{x \in \mathbb{R}^n : \ell(x) \geq 0\}$$

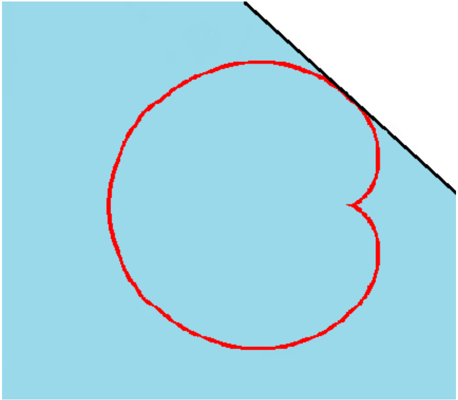
Example 2

Consider the cardioid given by
$I = \langle (x^2 + 2x + y^2)^2 - 4(x^2 + y^2) \rangle$.



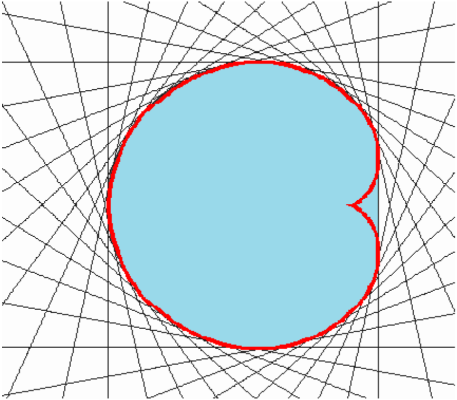
Example 2

Consider the cardioid given by $I = \langle (x^2 + 2x + y^2)^2 - 4(x^2 + y^2) \rangle$. By sweeping through different directions and solving the optimization problem we can visualize the second theta body:



Example 2

Consider the cardioid given by $I = \langle (x^2 + 2x + y^2)^2 - 4(x^2 + y^2) \rangle$. By sweeping through different directions and solving the optimization problem we can visualize the second theta body:



Combinatorial Moment Matrices

Another possible line of reasoning is the moment approach.

Combinatorial Moment Matrices

Another possible line of reasoning is the moment approach. Let I be a polynomial ideal and

$$\mathcal{B} = \{1 = f_0, x_1 = f_1, ..., x_n = f_n, f_{n+1}, ...\}$$

be a basis of $\mathbb{R}[x]/I$ and $\mathcal{B}_k = \{f_i : \deg(f_i) \leq k\}$ for all k.

Combinatorial Moment Matrices

Another possible line of reasoning is the moment approach. Let I be a polynomial ideal and

$$\mathcal{B} = \{1 = f_0, x_1 = f_1, ..., x_n = f_n, f_{n+1}, ...\}$$

be a basis of $\mathbb{R}[x]/I$ and $\mathcal{B}_k = \{f_i : \deg(f_i) \leq k\}$ for all k. Consider the polynomial vector $f^k(x) = (f_i(x))_{f_i \in \mathcal{B}_k}$ then

$$(f^k(x))(f^k(x))^t = \sum_{f_i \in \mathcal{B}} A_i f_i(x)$$

for some symmetric matrices A_i.

Combinatorial Moment Matrices

Another possible line of reasoning is the moment approach. Let I be a polynomial ideal and

$$\mathcal{B} = \{1 = f_0, x_1 = f_1, ..., x_n = f_n, f_{n+1}, ...\}$$

be a basis of $\mathbb{R}[x]/I$ and $\mathcal{B}_k = \{f_i : \deg(f_i) \leq k\}$ for all k. Consider the polynomial vector $f^k(x) = (f_i(x))_{f_i \in \mathcal{B}_k}$ then

$$(f^k(x))(f^k(x))^t = \sum_{f_i \in \mathcal{B}} A_i f_i(x)$$

for some symmetric matrices A_i. Given a vector y indexed by the elements in $\mathcal{B}$ we define the ***k*-th truncated combinatorial moment matrix** of y as

$$M_{\mathcal{B},k}(y) = \sum_{f_i \in \mathcal{B}} A_i y_{f_i}.$$

Combinatorial Moment Matrices - Example

Let $I = \langle x_1^2 - x_1, x_2^2 - x_2, x_3^2 - x_3 \rangle \subset \mathbb{R}[x_1, x_2, x_3]$,

Combinatorial Moment Matrices - Example

Let $I = \langle x_1^2 - x_1, x_2^2 - x_2, x_3^2 - x_3 \rangle \subset \mathbb{R}[x_1, x_2, x_3]$, pick

$$\mathcal{B} = \{ \quad 1, \quad x_1, \quad x_2, \quad x_3, \quad x_1x_2, \quad x_1x_3, \quad x_2x_3, \quad x_1x_2x_3 \quad \}$$

Combinatorial Moment Matrices - Example

Let $I = \langle x_1^2 - x_1, x_2^2 - x_2, x_3^2 - x_3 \rangle \subset \mathbb{R}[x_1, x_2, x_3]$, pick

$$\begin{array}{rcccccccccc} \mathcal{B} = \{ & 1, & x_1, & x_2, & x_3, & x_1x_2, & x_1x_3, & x_2x_3, & x_1x_2x_3 & \} \\ y = (& y_0, & y_1, & y_2, & y_3, & y_{12}, & y_{13}, & y_{23}, & y_{123} &). \end{array}$$

Combinatorial Moment Matrices - Example

Let $I = \langle x_1^2 - x_1, x_2^2 - x_2, x_3^2 - x_3 \rangle \subset \mathbb{R}[x_1, x_2, x_3]$, pick

$$\begin{array}{rcccccccccl} \mathcal{B} = \{ & 1, & x_1, & x_2, & x_3, & x_1x_2, & x_1x_3, & x_2x_3, & x_1x_2x_3 & \} \\ y = (& y_0, & y_1, & y_2, & y_3, & y_{12}, & y_{13}, & y_{23}, & y_{123} &). \end{array}$$

Then $M_{\mathcal{B}}(y)$ is given by

Combinatorial Moment Matrices - Example

Let $I = \langle x_1^2 - x_1, x_2^2 - x_2, x_3^2 - x_3 \rangle \subset \mathbb{R}[x_1, x_2, x_3]$, pick

$$\begin{array}{llllllllll} \mathcal{B} = \{ & 1, & x_1, & x_2, & x_3, & x_1x_2, & x_1x_3, & x_2x_3, & x_1x_2x_3 & \} \\ y = (& y_0, & y_1, & y_2, & y_3, & y_{12}, & y_{13}, & y_{23}, & y_{123} &). \end{array}$$

Then $M_{\mathcal{B}}(y)$ is given by

$$\begin{array}{c} \\ 1 \\ x_1 \\ x_2 \\ x_3 \\ x_1x_2 \\ x_1x_3 \\ x_2x_3 \\ x_1x_2x_3 \end{array} \begin{array}{c} \begin{array}{cccccccc} 1 & x_1 & x_2 & x_3 & x_1x_2 & x_1x_3 & x_2x_3 & x_1x_2x_3 \end{array} \\ \left[\begin{array}{cccccccc} & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \end{array} \right] \end{array}$$

Combinatorial Moment Matrices - Example

Let $I = \langle x_1^2 - x_1, x_2^2 - x_2, x_3^2 - x_3 \rangle \subset \mathbb{R}[x_1, x_2, x_3]$, pick

$$\begin{array}{rcccccccccc} \mathcal{B} = \{ & 1, & x_1, & x_2, & x_3, & x_1x_2, & x_1x_3, & x_2x_3, & x_1x_2x_3 & \} \\ y = (& y_0, & y_1, & y_2, & y_3, & y_{12}, & y_{13}, & y_{23}, & y_{123} &). \end{array}$$

Then $M_{\mathcal{B}}(y)$ is given by

$$\begin{array}{c} \\ 1 \\ x_1 \\ x_2 \\ x_3 \\ x_1x_2 \\ x_1x_3 \\ x_2x_3 \\ x_1x_2x_3 \end{array} \begin{array}{c} \begin{array}{cccccccc} 1 & x_1 & x_2 & x_3 & x_1x_2 & x_1x_3 & x_2x_3 & x_1x_2x_3 \end{array} \\ \left[\begin{array}{cccccccc} y_0 & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \end{array} \right] \end{array}$$

Combinatorial Moment Matrices - Example

Let $I = \langle x_1^2 - x_1, x_2^2 - x_2, x_3^2 - x_3 \rangle \subset \mathbb{R}[x_1, x_2, x_3]$, pick

$$\begin{array}{lcccccccccc} \mathcal{B} = \{ & 1, & x_1, & x_2, & x_3, & x_1x_2, & x_1x_3, & x_2x_3, & x_1x_2x_3 & \} \\ y = (& y_0, & y_1, & y_2, & y_3, & y_{12}, & y_{13}, & y_{23}, & y_{123} &). \end{array}$$

Then $M_{\mathcal{B}}(y)$ is given by

	1	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$
1	y_0	y_1						
x_1								
x_2								
x_3								
x_1x_2								
x_1x_3								
x_2x_3								
$x_1x_2x_3$								

Combinatorial Moment Matrices - Example

Let $I = \langle x_1^2 - x_1, x_2^2 - x_2, x_3^2 - x_3 \rangle \subset \mathbb{R}[x_1, x_2, x_3]$, pick

$$\begin{array}{rccccccccc} \mathcal{B} = \{ & 1, & x_1, & x_2, & x_3, & x_1x_2, & x_1x_3, & x_2x_3, & x_1x_2x_3 & \} \\ y = (& y_0, & y_1, & y_2, & y_3, & y_{12}, & y_{13}, & y_{23}, & y_{123} &). \end{array}$$

Then $M_{\mathcal{B}}(y)$ is given by

$$\begin{array}{r} \\ 1 \\ x_1 \\ x_2 \\ x_3 \\ x_1x_2 \\ x_1x_3 \\ x_2x_3 \\ x_1x_2x_3 \end{array} \begin{array}{c} \begin{array}{cccccccc} 1 & x_1 & x_2 & x_3 & x_1x_2 & x_1x_3 & x_2x_3 & x_1x_2x_3 \end{array} \\ \left[\begin{array}{cccccccc} y_0 & y_1 & y_2 & y_3 & y_{12} & y_{13} & y_{23} & y_{123} \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \end{array} \right] \end{array}$$

Combinatorial Moment Matrices - Example

Let $I = \langle x_1^2 - x_1, x_2^2 - x_2, x_3^2 - x_3 \rangle \subset \mathbb{R}[x_1, x_2, x_3]$, pick

$$\begin{array}{rlcccccccc l} \mathcal{B} = & \{ & 1, & x_1, & x_2, & x_3, & x_1x_2, & x_1x_3, & x_2x_3, & x_1x_2x_3 & \} \\ y = & (& y_0, & y_1, & y_2, & y_3, & y_{12}, & y_{13}, & y_{23}, & y_{123} &). \end{array}$$

Then $M_{\mathcal{B}}(y)$ is given by

$$\begin{array}{c} \\ 1 \\ x_1 \\ x_2 \\ x_3 \\ x_1x_2 \\ x_1x_3 \\ x_2x_3 \\ x_1x_2x_3 \end{array} \begin{array}{c} \begin{array}{cccccccc} 1 & x_1 & x_2 & x_3 & x_1x_2 & x_1x_3 & x_2x_3 & x_1x_2x_3 \end{array} \\ \left[\begin{array}{cccccccc} y_0 & y_1 & y_2 & y_3 & y_{12} & y_{13} & y_{23} & y_{123} \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & ? & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \end{array} \right] \end{array}$$

Combinatorial Moment Matrices - Example

Let $I = \langle x_1^2 - x_1, x_2^2 - x_2, x_3^2 - x_3 \rangle \subset \mathbb{R}[x_1, x_2, x_3]$, pick

$$\begin{array}{rcccccccccc} \mathcal{B} = \{ & 1, & x_1, & x_2, & x_3, & x_1x_2, & x_1x_3, & x_2x_3, & x_1x_2x_3 & \} \\ y = (& y_0, & y_1, & y_2, & y_3, & y_{12}, & y_{13}, & y_{23}, & y_{123} &). \end{array}$$

Then $M_{\mathcal{B}}(y)$ is given by

$$\begin{array}{r} \\ 1 \\ x_1 \\ x_2 \\ x_3 \\ x_1x_2 \\ x_1x_3 \\ x_2x_3 \\ x_1x_2x_3 \end{array} \begin{array}{c} \begin{array}{cccccccc} 1 & x_1 & x_2 & x_3 & x_1x_2 & x_1x_3 & x_2x_3 & x_1x_2x_3 \end{array} \\ \left[\begin{array}{cccccccc} y_0 & y_1 & y_2 & y_3 & y_{12} & y_{13} & y_{23} & y_{123} \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & y_{123} & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \end{array} \right] \end{array}$$

Combinatorial Moment Matrices - Example

Let $I = \langle x_1^2 - x_1, x_2^2 - x_2, x_3^2 - x_3 \rangle \subset \mathbb{R}[x_1, x_2, x_3]$, pick

$$\begin{array}{lllllllll} \mathcal{B} = \{ & 1, & x_1, & x_2, & x_3, & x_1x_2, & x_1x_3, & x_2x_3, & x_1x_2x_3 \;\} \\ y = (& y_0, & y_1, & y_2, & y_3, & y_{12}, & y_{13}, & y_{23}, & y_{123} \;). \end{array}$$

Then $M_{\mathcal{B}}(y)$ is given by

$$\begin{array}{c} \\ 1 \\ x_1 \\ x_2 \\ x_3 \\ x_1x_2 \\ x_1x_3 \\ x_2x_3 \\ x_1x_2x_3 \end{array} \begin{array}{c} \begin{array}{cccccccc} 1 & x_1 & x_2 & x_3 & x_1x_2 & x_1x_3 & x_2x_3 & x_1x_2x_3 \end{array} \\ \begin{bmatrix} y_0 & y_1 & y_2 & y_3 & y_{12} & y_{13} & y_{23} & y_{123} \\ y_1 & y_1 & y_{12} & y_{13} & y_{12} & y_{13} & y_{123} & y_{123} \\ y_2 & y_{12} & y_2 & y_{23} & y_{12} & y_{123} & y_{23} & y_{123} \\ y_3 & y_{13} & y_{23} & y_3 & y_{123} & y_{13} & y_{23} & y_{123} \\ y_{12} & y_{12} & y_{12} & y_{123} & y_{12} & y_{123} & y_{123} & y_{123} \\ y_{13} & y_{13} & y_{123} & y_{13} & y_{123} & y_{13} & y_{123} & y_{123} \\ y_{23} & y_{123} & y_{23} & y_{23} & y_{123} & y_{123} & y_{23} & y_{123} \\ y_{123} & y_{123} & y_{123} & y_{123} & y_{123} & y_{123} & y_{123} & y_{123} \end{bmatrix} \end{array}$$

Combinatorial Moment Matrices - Example

Let $I = \langle x_1^2 - x_1, x_2^2 - x_2, x_3^2 - x_3 \rangle \subset \mathbb{R}[x_1, x_2, x_3]$, pick

$$\begin{array}{rcccccccccc} \mathcal{B} = \{ & 1, & x_1, & x_2, & x_3, & x_1x_2, & x_1x_3, & x_2x_3, & x_1x_2x_3 & \} \\ y = (& y_0, & y_1, & y_2, & y_3, & y_{12}, & y_{13}, & y_{23}, & y_{123} &). \end{array}$$

$M_{\mathcal{B},1}(y)$ is given by:

	1	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$
1	$\mathbf{y_0}$	$\mathbf{y_1}$	$\mathbf{y_2}$	$\mathbf{y_3}$	y_{12}	y_{13}	y_{23}	y_{123}
x_1	$\mathbf{y_1}$	$\mathbf{y_1}$	$\mathbf{y_{12}}$	$\mathbf{y_{13}}$	y_{12}	y_{13}	y_{123}	y_{123}
x_2	$\mathbf{y_2}$	$\mathbf{y_{12}}$	$\mathbf{y_2}$	$\mathbf{y_{23}}$	y_{12}	y_{123}	y_{23}	y_{123}
x_3	$\mathbf{y_3}$	$\mathbf{y_{13}}$	$\mathbf{y_{23}}$	$\mathbf{y_3}$	y_{123}	y_{13}	y_{23}	y_{123}
x_1x_2	y_{12}	y_{12}	y_{12}	y_{123}	y_{12}	y_{123}	y_{123}	y_{123}
x_1x_3	y_{13}	y_{13}	y_{123}	y_{13}	y_{123}	y_{13}	y_{123}	y_{123}
x_2x_3	y_{23}	y_{123}	y_{23}	y_{23}	y_{123}	y_{123}	y_{23}	y_{123}
$x_1x_2x_3$	y_{123}	y_{123}	y_{123}	y_{123}	y_{123}	y_{123}	y_{123}	y_{123}

Combinatorial Moment Matrices - Example

Let $I = \langle x_1^2 - x_1, x_2^2 - x_2, x_3^2 - x_3 \rangle \subset \mathbb{R}[x_1, x_2, x_3]$, pick

$$\begin{array}{lllllllll} \mathcal{B} = \{ & 1, & x_1, & x_2, & x_3, & x_1x_2, & x_1x_3, & x_2x_3, & x_1x_2x_3 \quad \} \\ y = (& y_0, & y_1, & y_2, & y_3, & y_{12}, & y_{13}, & y_{23}, & y_{123} \quad). \end{array}$$

$M_{\mathcal{B},2}(y)$ is given by:

	1	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$
1	y_0	y_1	y_2	y_3	y_{12}	y_{13}	y_{23}	y_{123}
x_1	y_1	y_1	y_{12}	y_{13}	y_{12}	y_{13}	y_{123}	y_{123}
x_2	y_2	y_{12}	y_2	y_{23}	y_{12}	y_{123}	y_{23}	y_{123}
x_3	y_3	y_{13}	y_{23}	y_3	y_{123}	y_{13}	y_{23}	y_{123}
x_1x_2	y_{12}	y_{12}	y_{12}	y_{123}	y_{12}	y_{123}	y_{123}	y_{123}
x_1x_3	y_{13}	y_{13}	y_{123}	y_{13}	y_{123}	y_{13}	y_{123}	y_{123}
x_2x_3	y_{23}	y_{123}	y_{23}	y_{23}	y_{123}	y_{123}	y_{23}	y_{123}
$x_1x_2x_3$	y_{123}	y_{123}	y_{123}	y_{123}	y_{123}	y_{123}	y_{123}	y_{123}

Moment relaxation

Define the convex body

$$Q_k(I) = \{y \in \mathbb{R}^{\mathcal{B}} : y_0 = 1, M_{\mathcal{B},k}(y) \succeq 0\}.$$

Moment relaxation

Define the convex body

$$Q_k(I) = \{y \in \mathbb{R}^{\mathcal{B}} : y_0 = 1, M_{\mathcal{B},k}(y) \succeq 0\}.$$

Definition

The k-th moment relaxation of $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$ is the set

$$L_k(I) = \{(y_1, ..., y_n) : y \in Q_k(I)\}.$$

Moment relaxation

Define the convex body

$$Q_k(I) = \{y \in \mathbb{R}^{\mathcal{B}} : y_0 = 1, M_{\mathcal{B},k}(y) \succeq 0\}.$$

Definition

The k-th moment relaxation of $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$ is the set

$$L_k(I) = \{(y_1, ..., y_n) : y \in Q_k(I)\}.$$

Note that for all $p = (p_1, ..., p_n) \in \mathcal{V}_{\mathbb{R}}(I)$ we have

$$M_{\mathcal{B},k}(f^k(p)) = \sum_{f_i \in \mathcal{B}_k} A_i f_i(p) = (f^k(p))(f^k(p))^t \succeq 0.$$

Relation between approaches

We say that an ideal I is real radical if

$$I = \sqrt[\mathbb{R}]{I} := \{p : -p^{2m} \text{ sos modulo } I \text{ for some } m\}$$

or equivalently if $I = \mathcal{I}(\mathcal{V}_{\mathbb{R}}(I))$ (Real Nullstellensatz).

Relation between approaches

We say that an ideal I is real radical if

$$I = \sqrt[\mathbb{R}]{I} := \{p : -p^{2m} \text{ sos modulo } I \text{ for some } m\}$$

or equivalently if $I = \mathcal{I}(\mathcal{V}_{\mathbb{R}}(I))$ (Real Nullstellensatz).

Theorem (G.,Parrilo,Thomas)

Let I be a polynomial ideal:

- $\mathrm{cl}(L_k(I)) \subseteq \mathrm{TH}_k(I)$.

Relation between approaches

We say that an ideal I is real radical if

$$I = \sqrt[\mathbb{R}]{I} := \{p : -p^{2m} \text{ sos modulo } I \text{ for some } m\}$$

or equivalently if $I = \mathcal{I}(\mathcal{V}_{\mathbb{R}}(I))$ (Real Nullstellensatz).

Theorem (G.,Parrilo,Thomas)

Let I be a polynomial ideal:

- $\mathrm{cl}(L_k(I)) \subseteq \mathrm{TH}_k(I)$.
- *If I is real radical,* $\mathrm{cl}(L_k(I)) = \mathrm{TH}_k(I)$.

Relation between approaches

We say that an ideal I is real radical if

$$I = \sqrt[\mathbb{R}]{I} := \{p : -p^{2m} \text{ sos modulo } I \text{ for some } m\}$$

or equivalently if $I = \mathcal{I}(\mathcal{V}_{\mathbb{R}}(I))$ (Real Nullstellensatz).

Theorem (G.,Parrilo,Thomas)

Let I be a polynomial ideal:

- $\mathrm{cl}(L_k(I)) \subseteq \mathrm{TH}_k(I)$.
- *If I is real radical,* $\mathrm{cl}(L_k(I)) = \mathrm{TH}_k(I)$.
- *If I is real radical and $\mathcal{V}_{\mathbb{R}}(I) \subseteq \{0,1\}^n$, then $L_k(I) = \mathrm{TH}_k(I)$.*

Real Radicalness matters?

How important is real radicalness?

Real Radicalness matters?

How important is real radicalness?

Lemma (G.,Thomas)

Let p be a polynomial of degree d such that $-p^{2m}$ is k-sos modulo I (i.e. $p \in \sqrt[\mathbb{R}]{I}$).

Real Radicalness matters?

How important is real radicalness?

Lemma (G.,Thomas)

Let p be a polynomial of degree d such that $-p^{2m}$ is k-sos modulo I (i.e. $p \in \sqrt[\mathbb{R}]{I}$). Then $p + \varepsilon$ is $(k + 4dm)$-sos modulo I for all $\varepsilon > 0$.

Real Radicalness matters?

How important is real radicalness?

Lemma (G.,Thomas)

Let p be a polynomial of degree d such that $-p^{2m}$ is k-sos modulo I (i.e. $p \in \sqrt[\mathbb{R}]{I}$). Then $p + \varepsilon$ is $(k + 4dm)$-sos modulo I for all $\varepsilon > 0$.

Fixing the ideal there exist bounds for the Real Nullstellensatz that bound m and k as a function of d. Using this we get:

Real Radicalness matters?

How important is real radicalness?

Lemma (G.,Thomas)

Let p be a polynomial of degree d such that $-p^{2m}$ is k-sos modulo I (i.e. $p \in \sqrt[\mathbb{R}]{I}$). Then $p + \varepsilon$ is $(k + 4dm)$-sos modulo I for all $\varepsilon > 0$.

Fixing the ideal there exist bounds for the Real Nullstellensatz that bound m and k as a function of d. Using this we get:

Theorem (G.,Thomas)

Fix an ideal I and an integer k. Then there exists $m_k \in \mathbb{N}$ such that $\mathrm{TH}_{m_k}(I) \subseteq \mathrm{TH}_k(\sqrt[\mathbb{R}]{I})$.

Proof of Lemma

Suppose $-p^{2m}$ is k-sos.

Proof of Lemma

Suppose $-p^{2m}$ is k-sos.

- For $l \geq m$ and $\xi > 0$ we have

$$p^l + \xi = \frac{1}{\xi}\left((p^l/2 + \xi)^2 + \frac{1}{2}(-p^{2m})p^{2(l-m)}\right) \text{ is sos.}$$

Proof of Lemma

Suppose $-p^{2m}$ is k-sos.

- For $l \geq m$ and $\xi > 0$ we have

$$p^l + \xi = \frac{1}{\xi}\left((p^l/2 + \xi)^2 + \frac{1}{2}(-p^{2m})p^{2(l-m)}\right) \text{ is sos.}$$

- For $\sigma > 0$, consider the Taylor series of $\sqrt{\sigma + t}$ and let $f(t)$ be the truncation of this series at the $(2m-1)$-th term.

Proof of Lemma

Suppose $-p^{2m}$ is k-sos.

- For $l \geq m$ and $\xi > 0$ we have

$$p^l + \xi = \frac{1}{\xi}\left((p^l/2 + \xi)^2 + \frac{1}{2}(-p^{2m})p^{2(l-m)}\right) \text{ is sos.}$$

- For $\sigma > 0$, consider the Taylor series of $\sqrt{\sigma + t}$ and let $f(t)$ be the truncation of this series at the $(2m - 1)$-th term.One can check

$$(f(p(x)))^2 = \sigma + p(x) + \sum_{i=0}^{m-1} a_i p(x)^{2m+2i} - \sum_{i=0}^{m-2} b_i p(x)^{2m+2i+1},$$

for positive a_i's and b_i's,

Proof of Lemma

Suppose $-p^{2m}$ is k-sos.

- For $l \geq m$ and $\xi > 0$ we have

$$p^l + \xi = \frac{1}{\xi}\left((p^l/2 + \xi)^2 + \frac{1}{2}(-p^{2m})p^{2(l-m)}\right) \text{ is sos.}$$

- For $\sigma > 0$, consider the Taylor series of $\sqrt{\sigma + t}$ and let $f(t)$ be the truncation of this series at the $(2m-1)$-th term.One can check

$$(f(p(x)))^2 = \sigma + p(x) + \sum_{i=0}^{m-1} a_i p(x)^{2m+2i} - \sum_{i=0}^{m-2} b_i p(x)^{2m+2i+1},$$

for positive a_i's and b_i's,hence

$$\sigma + p(x) = (f(p(x)))^2 - \sum_{i=0}^{m-1} a_i p(x)^{2m+2i} + \sum_{i=0}^{m-2} b_i p(x)^{2m+2i+1}.$$

Convergence

By the previous theorem the asymptotic/finite convergence of the theta bodies depends only on the real variety.

Convergence

By the previous theorem the asymptotic/finite convergence of the theta bodies depends only on the real variety.

Theorem

Let I be a polynomial ideal:

- *If $\mathcal{V}_{\mathbb{R}}(I)$ is compact,* $\mathrm{TH}_k(I) \to \mathrm{cl}(\mathrm{conv}(\mathcal{V}_{\mathbb{R}}(I)))$. *(Schmügden)*

Convergence

By the previous theorem the asymptotic/finite convergence of the theta bodies depends only on the real variety.

Theorem

Let I be a polynomial ideal:

- *If* $\mathcal{V}_{\mathbb{R}}(I)$ *is compact,* $\mathrm{TH}_k(I) \to \mathrm{cl}(\mathrm{conv}(\mathcal{V}_{\mathbb{R}}(I)))$. *(Schmügden)*
- *If* $\mathcal{V}_{\mathbb{R}}(I)$ *is finite, then the convergence is finite. (Lasserre, Laurent, Rostalski)*

Convergence

By the previous theorem the asymptotic/finite convergence of the theta bodies depends only on the real variety.

Theorem

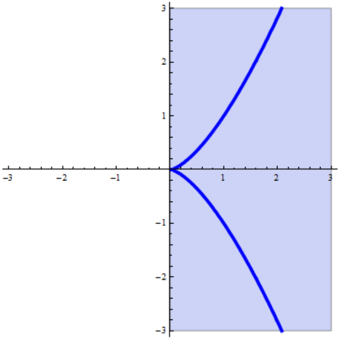
Let I be a polynomial ideal:

- *If* $\mathcal{V}_{\mathbb{R}}(I)$ *is compact,* $\mathrm{TH}_k(I) \to \mathrm{cl}(\mathrm{conv}(\mathcal{V}_{\mathbb{R}}(I)))$. *(Schmügden)*
- *If* $\mathcal{V}_{\mathbb{R}}(I)$ *is finite, then the convergence is finite. (Lasserre, Laurent, Rostalski)*

However we do not always have convergence.

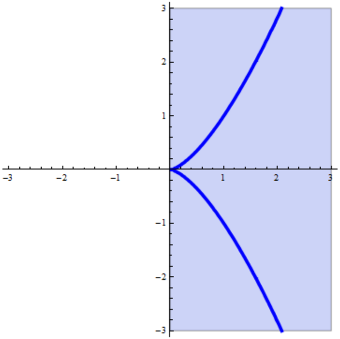
Bad Example

Consider the ideal $I = \langle y^2 - x^3 \rangle$.



Bad Example

Consider the ideal $I = \langle y^2 - x^3 \rangle$.



It is not hard to see that no linear polynomial is sos modulo this ideal, hence $\mathrm{TH}_k(I) = \mathbb{R}^2$ for all k.

Singularities and Convergence

Definition

For $x \in \mathcal{V}_{\mathbb{R}}(I)$, the tangent space of x, $T_x(I)$ is the affine space passing through x and perpendicular to ∇g for all $g \in \sqrt[\mathbb{R}]{I}$.

Singularities and Convergence

Definition

For $x \in \mathcal{V}_{\mathbb{R}}(I)$, the tangent space of x, $T_x(I)$ is the affine space passing through x and perpendicular to ∇g for all $g \in \sqrt[\mathbb{R}]{I}$.
x is **convex-singular** if $x \in \partial(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$ and $T_x(I)$ intersects the relative interior of $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$.

Singularities and Convergence

Definition

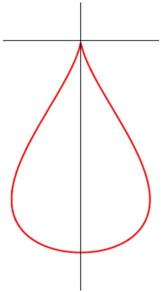
For $x \in \mathcal{V}_{\mathbb{R}}(I)$, the tangent space of x, $T_x(I)$ is the affine space passing through x and perpendicular to ∇g for all $g \in \sqrt[\mathbb{R}]{I}$.
x is **convex-singular** if $x \in \partial(\mathrm{conv}(\mathcal{V}_{\mathbb{R}}(I)))$ and $T_x(I)$ intersects the relative interior of $\mathrm{conv}(\mathcal{V}_{\mathbb{R}}(I))$.

Proposition (G.,Netzer)

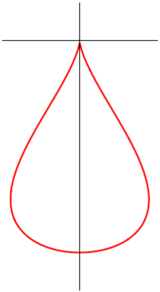
If I has a convex-singularity then, for all k,

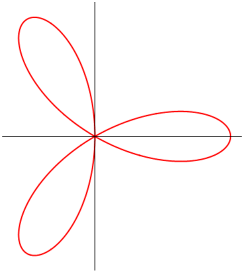
$$\mathrm{TH}_k(I) \neq \mathrm{cl}(\mathrm{conv}(\mathcal{V}_{\mathbb{R}}(I))).$$

Examples

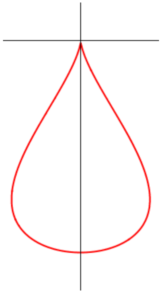


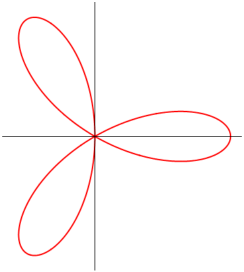
Examples

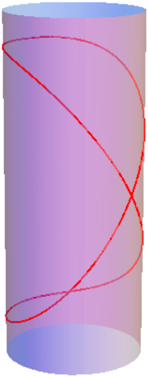




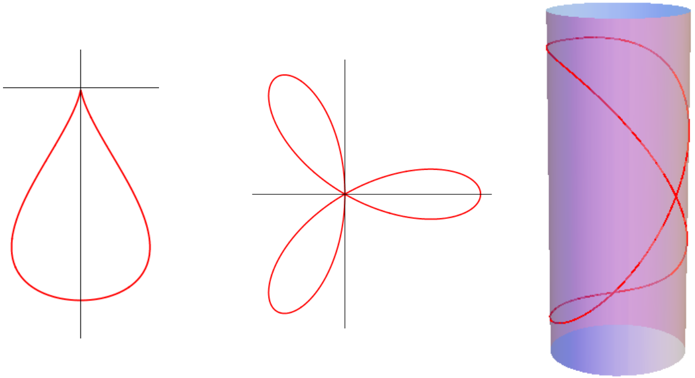
Examples







Examples



The first variety has a convex-singularity, but none of the other varieties have it.

Zero-dimensional varieties

In combinatorial optimization, zero-dimensional varieties ($0/1$-optimization) play an important role.

Zero-dimensional varieties

In combinatorial optimization, zero-dimensional varieties ($0/1$-optimization) play an important role.

Example: Stable Set Problem

Given a graph $G = ([n], E)$ find the maximum set $S \subseteq [n]$ such that no two points in S are connected with an edge.

Zero-dimensional varieties

In combinatorial optimization, zero-dimensional varieties (0/1-optimization) play an important role.

Example: Stable Set Problem

Given a graph $G = ([n], E)$ find the maximum set $S \subseteq [n]$ such that no two points in S are connected with an edge.

This can be modeled by the ideal

$$I = \left\langle x_i^2 - x_i, x_j x_k \ : \ \forall i \in [n], \{j, k\} \in E \right\rangle$$

since $\mathcal{V}_{\mathbb{R}}(I)$ is the set of characteristic vectors of stable sets.

TH_1-Exactness

It is interesting to characterize convergence in one step of the theta body hierarchy, i.e, TH_1-exactness.

TH_1-Exactness

It is interesting to characterize convergence in one step of the theta body hierarchy, i.e, TH_1-exactness. In the zero-dimensional case a full characterization is possible.

Theorem (G.,Parrilo,Thomas)

Let $S \subset \mathbb{R}^n$ be finite, and $I = \mathcal{I}(S)$.

TH_1-Exactness

It is interesting to characterize convergence in one step of the theta body hierarchy, i.e, TH_1-exactness. In the zero-dimensional case a full characterization is possible.

Theorem (G.,Parrilo,Thomas)

Let $S \subset \mathbb{R}^n$ be finite, and $I = \mathcal{I}(S)$.
I is TH_1-exact if and only if for every facet of the polytope $\mathrm{conv}(S)$ all points of S are either on that face or on a unique plane parallel to it.

TH_1-Exactness

It is interesting to characterize convergence in one step of the theta body hierarchy, i.e, TH_1-exactness. In the zero-dimensional case a full characterization is possible.

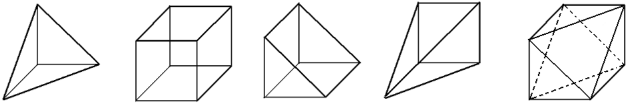
Theorem (G.,Parrilo,Thomas)

Let $S \subset \mathbb{R}^n$ be finite, and $I = \mathcal{I}(S)$.
I is TH_1*-exact if and only if for every facet of the polytope* $\mathrm{conv}(S)$ *all points of S are either on that face or on a unique plane parallel to it.*

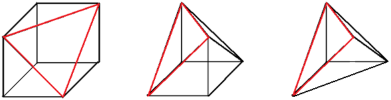
- For the stable set problem the ideal is TH_1-exact if and only if the graph is perfect. (Lovász)

Examples in $\mathbb{R}^3$

TH_1-exact



Not TH_1-exact



Optimization over semialgebraic sets

We now change our focus to basic closed semialgebraic sets.

Optimization over semialgebraic sets

We now change our focus to basic closed semialgebraic sets.

Problem

$$\begin{aligned} \max_{x \in \mathbb{R}^n} \quad & \langle c, x \rangle \\ \text{s.t.} \quad & g_i(x) \geq 0, \quad i = 1, \cdots, k, \end{aligned}$$

where the g_i are polynomials over $\mathbb{R}$.

Optimization over semialgebraic sets

We now change our focus to basic closed semialgebraic sets.

Problem

$$\begin{aligned} \max_{x \in \mathbb{R}^n} \quad & \langle c, x \rangle \\ \text{s.t.} \quad & g_i(x) \geq 0, \quad i = 1, \cdots, k, \end{aligned}$$

where the g_i are polynomials over $\mathbb{R}$.

Let $S = \{x \in \mathbb{R}^n : g_i(x) \geq 0, \ i = 1, ..., k\}$.

Optimization over semialgebraic sets

We now change our focus to basic closed semialgebraic sets.

Problem

$$\begin{aligned} \max_{x \in \mathbb{R}^n} \quad & \langle c, x \rangle \\ \text{s.t.} \quad & g_i(x) \geq 0, \quad i = 1, \cdots, k, \end{aligned}$$

where the g_i are polynomials over $\mathbb{R}$.

Let $S = \{x \in \mathbb{R}^n : g_i(x) \geq 0,\ i = 1, ..., k\}$. We can rewrite the problem as

Problem*

$$\begin{aligned} \min_{\lambda \in \mathbb{R}} \quad & \lambda \\ \text{s.t.} \quad & \lambda - \langle c, x \rangle \geq 0, \quad \forall x \in S. \end{aligned}$$

Optimization over semialgebraic sets

We now change our focus to basic closed semialgebraic sets.

Problem

$$\begin{aligned} \max_{x \in \mathbb{R}^n} \quad & \langle c, x \rangle \\ \text{s.t.} \quad & g_i(x) \geq 0, \quad i = 1, \cdots, k, \end{aligned}$$

where the g_i are polynomials over $\mathbb{R}$.

Let $S = \{x \in \mathbb{R}^n : g_i(x) \geq 0, \ i = 1, ..., k\}$. We can rewrite the problem as

Problem*

$$\begin{aligned} \min_{\lambda \in \mathbb{R}} \quad & \lambda \\ \text{s.t.} \quad & \lambda - \langle c, x \rangle \geq 0, \quad \forall x \in S. \end{aligned}$$

To do as in the algebraic set case we have to find an algebraic certificate of nonnegativity over S.

Nonnegativity over semialgebraic sets

A classic way of certifying nonnegativity of a polynomial p over S is to provide a representation

$$p(x) = \sigma_0(x) + \sum_{i=1}^{k} \sigma_i(x) g_i(x)$$

where the σ_i are sums of squares of polynomials.

Nonnegativity over semialgebraic sets

A classic way of certifying nonnegativity of a polynomial p over S is to provide a representation

$$p(x) = \sigma_0(x) + \sum_{i=1}^{k} \sigma_i(x) g_i(x)$$

where the σ_i are sums of squares of polynomials.

We will denote by Σ_d^S the set of all polynomials that have such a representation with $\deg(\sigma_i g_i) \leq 2d$ for all i.

Sums of squares relaxations

We then get the sums of squares problem

Problem-Σ_d

$$\begin{aligned} \min_{\lambda\in\mathbb{R}} \quad & \lambda \\ \text{s.t.} \quad & \lambda - \langle c, x \rangle \in \Sigma_d^S. \end{aligned}$$

Sums of squares relaxations

We then get the sums of squares problem

Problem-Σ_d

$$\begin{aligned} \min_{\lambda \in \mathbb{R}} \quad & \lambda \\ \text{s.t.} \quad & \lambda - \langle c, x \rangle \in \Sigma_d^S. \end{aligned}$$

Again we are interested in the underlying geometry of this problem.

Sums of squares relaxations

We then get the sums of squares problem

Problem-Σ_d

$$\begin{aligned} \min_{\lambda \in \mathbb{R}} \quad & \lambda \\ \text{s.t.} \quad & \lambda - \langle c, x \rangle \in \Sigma_d^S. \end{aligned}$$

Again we are interested in the underlying geometry of this problem.

Lasserre Bodies

$$\mathcal{L}_d(S) = \bigcap_{\ell \text{ linear }, \ell \in \Sigma_d^S} \{x \in \mathbb{R}^n : \ell(x) \geq 0\}$$

which we call the d-th Lasserre Relaxation of $\text{conv}(S)$.

Remarks

- One can also use the moment approach as we have done in the algebraic sets case. This is the more traditional definition of Lasserre relaxations.

Remarks

- One can also use the moment approach as we have done in the algebraic sets case. This is the more traditional definition of Lasserre relaxations. If we assume that the set S has non-empty interior, these two definitions match.

Remarks

- One can also use the moment approach as we have done in the algebraic sets case. This is the more traditional definition of Lasserre relaxations. If we assume that the set S has non-empty interior, these two definitions match.
- The definition actually depends on a particular representation for S, and not only on S itself.

Remarks

- One can also use the moment approach as we have done in the algebraic sets case. This is the more traditional definition of Lasserre relaxations. If we assume that the set S has non-empty interior, these two definitions match.
- The definition actually depends on a particular representation for S, and not only on S itself. When we write S we are thinking of a fixed representation.

Remarks

- One can also use the moment approach as we have done in the algebraic sets case. This is the more traditional definition of Lasserre relaxations. If we assume that the set S has non-empty interior, these two definitions match.
- The definition actually depends on a particular representation for S, and not only on S itself. When we write S we are thinking of a fixed representation.
- If S is Archimedean (i.e., has an algebraic certificate of compactness) then we have asymptotic convergence of the hierarchy $\mathcal{L}_d(S)$. (Putinar)

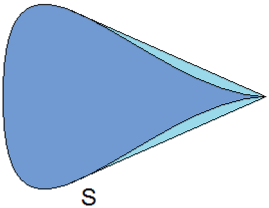
Remarks

- One can also use the moment approach as we have done in the algebraic sets case. This is the more traditional definition of Lasserre relaxations. If we assume that the set S has non-empty interior, these two definitions match.
- The definition actually depends on a particular representation for S, and not only on S itself. When we write S we are thinking of a fixed representation.
- If S is Archimedean (i.e., has an algebraic certificate of compactness) then we have asymptotic convergence of the hierarchy $\mathcal{L}_d(S)$. (Putinar)
- We are interested in finding when does finite convergence not hold.

Obstruction Lemma

Lemma (G., Netzer)

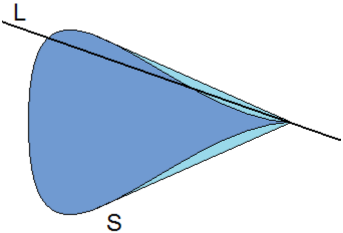
- *Let* $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_k(x) \geq 0\}$,



Obstruction Lemma

Lemma (G., Netzer)

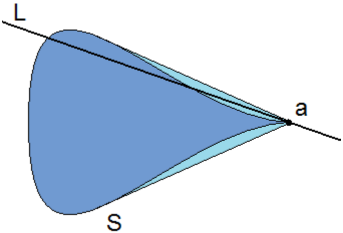
- *Let* $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_k(x) \geq 0\}$*,*
- L *a line in* $\mathbb{R}^n$ *s.t.* $\text{int}(S \cap L) \neq \emptyset$ *relative to L,*



Obstruction Lemma

Lemma (G., Netzer)

- *Let* $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_k(x) \geq 0\}$,
- L *a line in* $\mathbb{R}^n$ *s.t.* $\mathrm{int}(S \cap L) \neq \emptyset$ *relative to* L,
- $a \in S$ *be in the relative boundary of* $\overline{\mathrm{conv}(S)} \cap L$.

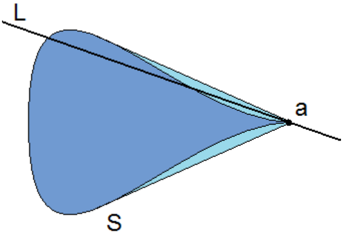


Obstruction Lemma

Lemma (G., Netzer)

- *Let* $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_k(x) \geq 0\}$,
- L *a line in* $\mathbb{R}^n$ *s.t.* $\text{int}(S \cap L) \neq \emptyset$ *relative to* L,
- $a \in S$ *be in the relative boundary of* $\overline{\text{conv}(S)} \cap L$.

If for all g_i *s.t.* $g_i(a) = 0$ *we have* $\bigtriangledown g_i(a) \perp L$ *then, for all* d*, we have* $\mathcal{L}_d(S) \neq \text{cl}(\text{conv}(S))$.



Singularities

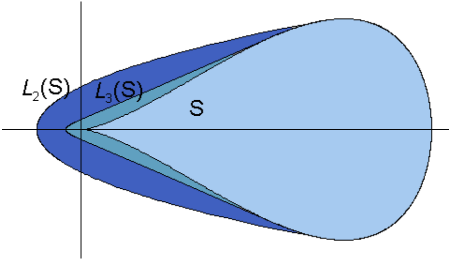
Corollary

If S has non-empty interior and there exists a point $a \in S$ that is on the boundary of $\operatorname{conv}(S)$ *s.t. all g_i verifying $g_i(a) = 0$ are singular at a, then we have $\mathcal{L}_d(S) \neq \operatorname{cl}(\operatorname{conv}(S))$ for all d.*

Singularities

Corollary

If S has non-empty interior and there exists a point $a \in S$ that is on the boundary of $\mathrm{conv}(S)$ *s.t. all g_i verifying $g_i(a) = 0$ are singular at a, then we have $\mathcal{L}_d(S) \neq \mathrm{cl}(\mathrm{conv}(S))$ for all d.*



$$g(x, y) = -x^4 + x^3 - y^2$$

Non-exposed faces

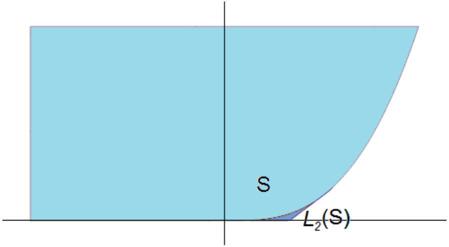
Corollary (Netzer-Plaumann-Schweighofer)

Suppose S is convex and has non-empty interior. If S has a non-exposed face then $\mathcal{L}_d(S) \neq \mathrm{cl}(\mathrm{conv}(S))$ for all d.

Non-exposed faces

Corollary (Netzer-Plaumann-Schweighofer)

Suppose S is convex and has non-empty interior. If S has a non-exposed face then $\mathcal{L}_d(S) \neq \mathrm{cl}(\mathrm{conv}(S))$ *for all d.*



$$g_1(x,y) = y - x^3,\ g_2(x,y) = y,\ g_3(x,y) = x+1,\ g_4(x,y) = 1-y$$

The End

Thank You