Semidefinite Programming Characterization and Computation of Real Radical Ideals

MEGA 2011

Monique Laurent, CWI, Amsterdam & Tilburg University

Joint work with
Jean Lasserre, LAAS-CNRS Toulouse
Philipp Rostalski, UC Berkeley, now Dräger, Lübeck

May 30, 2011

The problem

Given polynomials $h_1, \ldots, h_m \in \mathbb{R}[\mathbf{x}] = \mathbb{R}[x_1, \ldots, x_n]$ *generating the ideal*

$$I = (h_1, \ldots, h_m)$$

1. *Compute a base of the real radical ideal* $\sqrt[\mathbb{R}]{I}$ *of the ideal* I
2. *Compute the real variety* $V_{\mathbb{R}}(I)$

Assuming that $V_{\mathbb{R}}(I)$ *is finite.*

The complex case: $V_{\mathbb{C}}(I)$ is well understood

- **Homotopy continuation methods**

  [Sommese, Verschelde, Wampler,..]

- **Elimination methods:** Find polynomials in I with special structure that can be used to represent the roots:

  $\rightsquigarrow$ Compute polynomials in I in **triangular shape**: $f_1 \in \mathbb{R}[x_1]$, $f_2 \in \mathbb{R}[x_1, x_2]$, $\ldots$, $f_n \in \mathbb{R}[x_1, \ldots, x_n]$ (via Gröbner bases)

  $\rightsquigarrow$ Compute a rational univariate representation **(RUR)** of the roots: $x_i = h_i(t)/h(t)$, $f(t) = 0$ [Rouillier,...]

  $\rightsquigarrow$ Compute a border base and reduce to some **eigenvalue computations**

  [Kehrein-Kreuzer-Robbiano, Mourrain, Möller, Stetter,...]

The real case: $V_{\mathbb{R}}(I)$ is less well understood

- **Subdivision methods** combined with search methods and real root counting [Mourrain-Pavone, ...]

- **Khovanskii-Rolle continuation:** exploiting sharp bounds for real roots of fewnomials [Bates-Sottile,...]

Our contribution:

- A characterization of the real radical ideal $\sqrt[\mathbb{R}]{I}$, *as kernel of a positive semidefinite moment matrix*
- When $|V_{\mathbb{R}}(I)| < \infty$, an algorithm for computing a base of $\sqrt[\mathbb{R}]{I}$ and the real variety $V_{\mathbb{R}}(I)$

Remarks about our method:

- *Real algebraic* in nature: no complex roots are computed
- Works if $V_{\mathbb{R}}(I)$ is *finite* (while $V_{\mathbb{C}}(I)$ could be infinite)
- *Numerical:* uses semidefinite programming (SDP)

Outline

- Recap: (Real) Nullstellensatz, sums of squares of polynomials and semidefinite programming (SDP), eigenvalue method

- Moment matrices and real radical ideals

- Moment matrix approach for $\sqrt[\mathbb{R}]{I}$

- Extension to the complex case and links to the elimination method of Zhi-Reid

Some notation

- Polynomial ring: $R = \mathbb{K}[\mathbf{x}]$ $[\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, mostly $\mathbb{R}]$
- **(Complex) variety** of $I \subseteq R$:

$$V_{\mathbb{C}}(I) = \{v \in \mathbb{C}^n \mid f(v) = 0 \ \forall f \in I\}$$

- **Real variety** of I:

$$V_{\mathbb{R}}(I) = V_{\mathbb{C}}(I) \cap \mathbb{R}^n$$

- **Radical ideal** of I:

$$\sqrt{I} = \{f \in R \mid \exists m \in \mathbb{N} \ \ f^m \in I\}$$

- **Real radical** of I:

$$\sqrt[\mathbb{R}]{I} = \{f \in R \mid \exists m \in \mathbb{N} \ \exists s_i \in R \ \ f^{2m} + \sum_i s_i^2 \in I\}$$

- The ideal I is **radical** if $I = \sqrt{I}$ and **real radical** if $I = \sqrt[\mathbb{R}]{I}$.
- **Vanishing ideal** of $V \subseteq \mathbb{K}^n$:

$$I(V) = \{f \in R \mid f(v) = 0 \ \forall v \in V\}$$

A small example

Consider the ideal $\mathbf{I} = ((\mathbf{x_1^2} + \mathbf{x_2^2})^\mathbf{2})$ in $\mathbb{R}[x_1, x_2]$.

- $V_{\mathbb{C}}(I) = \{(x_1, \pm \mathbf{i} x_1) \mid x_1 \in \mathbb{C}\}$

  Radical ideal: $\sqrt{I} = (x_1^2 + x_2^2)$.

- $V_{\mathbb{R}}(I) = \{(0,0)\}$

  Real radical ideal: $\sqrt[\mathbb{R}]{I} = (x_1, x_2)$.

  Indeed: $x_1^4 + \underbrace{2x_1^2 x_2^2 + x_2^4}_{\text{sum of squares}} \in I \Longrightarrow x_1 \in \sqrt[\mathbb{R}]{I}$

Feature of this example:

$V_{\mathbb{R}}(I)$ is finite while $V_{\mathbb{C}}(I)$ is infinite.

A small example

Consider the ideal $\mathbf{I} = ((\mathbf{x_1^2} + \mathbf{x_2^2})^\mathbf{2})$ in $\mathbb{R}[x_1, x_2]$.

- $V_{\mathbb{C}}(I) = \{(x_1, \pm \mathbf{i} x_1) \mid x_1 \in \mathbb{C}\}$
  Radical ideal: $\sqrt{I} = (x_1^2 + x_2^2) = I(V_{\mathbb{C}}(I))$.

- $V_{\mathbb{R}}(I) = \{(0, 0)\}$
  Real radical ideal: $\sqrt[\mathbb{R}]{I} = (x_1, x_2) = I(V_{\mathbb{R}}(I))$.

Nullstellensatz and Real Nullstellensatz

Theorem

1. **[Hilbert's Nullstellensatz]** *For an ideal* $I \subseteq \mathbb{C}[\mathbf{x}]$,

$$\sqrt{I} = I(V_{\mathbb{C}}(I)).$$

2. **[Real Nullstellensatz, Krivine (1964)]**
For an ideal $I \subseteq \mathbb{R}[\mathbf{x}]$,

$$\sqrt[\mathbb{R}]{I} = I(V_{\mathbb{R}}(I)).$$

Hence, for an ideal $I = (h_1, \ldots, h_m)$

1. $V_{\mathbb{C}}(I) = \emptyset \iff 1 = \sum_{j=1}^m u_j h_j \in I$ [with LP]
2. $V_{\mathbb{R}}(I) = \emptyset \iff 1 + \sum_i s_i^2 = \sum_{j=1}^m u_j h_j \in I$ [with SDP]

for some polynomials s_i, u_j.

Semidefinite programming

Semidefinite programming (SDP) is linear optimization (LP) over the cone of positive semidefinite matrices.

- **LP**: vector variable $x \in \mathbb{R}^n$, $x \geq 0$
- **SDP:** matrix variable $X \in \mathbb{R}^{n \times n}$, $X \succeq 0$ (positive semidefinite)

(Semidefinite program)

Given symmetric matrices $C, A_j \in \mathbb{R}^{n \times n}$ *and* $b \in \mathbb{R}^m$*, compute:*

$$\max \;\; Tr(CX) \;\; \textit{such that } Tr(A_j X) = b_j \; (j = 1, \ldots, m), \;\; X \succeq 0$$

Dual SDP:

$$\min \;\; b^T y \;\; \textit{such that } \sum_{j=1}^{m} y_j A_j - C \succeq 0$$

There are efficient (interior-point) algorithms to solve semidefinite programs (to arbitrary precision).

Recognizing sums of squares of polynomials with SDP

Gram-matrix method of Powers-Wörmann [1998]:

$$f(x) = \sum_{|\alpha| \leq 2d} f_\alpha x^\alpha$$ is a **sum of squares of polynomials**

$$\Updownarrow$$

$f(x) = \sum_i s_i(x)^2$ [write $s_i(x) = \overline{s_i}^T [x]_d$]

$$\Updownarrow$$

$$f(x) = \sum_i [x]_d^T \overline{s_i}\ \overline{s_i}^T [x]_d = [x]_d^T \underbrace{\left(\sum_i \overline{s_i}\ \overline{s_i}^T \right)}_{X \succeq 0} [x]_d$$

$$\Updownarrow$$

The SDP: $\begin{cases} \displaystyle\sum_{\beta,\gamma | \beta+\gamma=\alpha} X_{\beta,\gamma} = f_\alpha \quad (|\alpha| \leq 2d) \\ X \succeq 0 \end{cases}$ **is feasible**

Example: $f(x,y) = x^4 + 2x^3y + 3x^2y^2 + 2xy^3 + y^4$ SOS?

$$f(x,y) = (x^2 \ xy \ y^2) \underbrace{\begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}}_{X \succeq 0} \begin{pmatrix} x^2 \\ xy \\ y^2 \end{pmatrix}$$

Equating coefficients on both sides:

x^4: $a = 1$ x^3y: $2b = 2$ x^2y^2: $2c + d = 3$
xy^3: $2e = 2$ y^4: $f = 2$

$$X = \begin{pmatrix} 1 & 1 & c \\ 1 & 3-2c & 1 \\ c & 1 & 2 \end{pmatrix} \succeq 0 \iff -1 \leq c \leq 1$$

$c = -1 \rightsquigarrow f = (x^2 + xy - y^2)^2 + (y^2 + 2xy)^2$

$c = 0 \rightsquigarrow f = (x^2 + xy)^2 + \frac{3}{2}(xy + y^2)^2 + \frac{1}{2}(xy - y^2)^2$

The eigenvalue method: Linear algebra in $\mathcal{A} = R/I$

$I \subseteq R$ ideal $\rightsquigarrow$ $\mathcal{A} = R/I$ its quotient algebra.

Lemma

- $|V_{\mathbb{C}}(I)| < \infty \iff \dim \mathcal{A} < \infty$.
- $|V_{\mathbb{C}}(I)| \leq \dim \mathcal{A}$, *with equality if and only if I is radical.*

Theorem (Stickelberger)

For $h \in R$, consider the **multiplication operator** *(by h):*

$$\begin{array}{rccc} M_h : & \mathcal{A} & \to & \mathcal{A} \\ & [f] & \mapsto & [hf] \end{array}$$

- *The* **eigenvalues** *of M_h are $\{h(v) \mid v \in V_{\mathbb{C}}(I)\}$.*
- *The* **eigenvectors** *of M_h^T give the points $v \in V_{\mathbb{C}}(I)$.*

Multiplication matrices M_{x_i} $\rightsquigarrow$ base of I

- $\mathcal{B} = \{b_1 = 1, b_2, \ldots, b_N\}$: a monomial base of $\mathcal{A} = R/I$

- Write any **(border) monomial**: $x_i b_j = \underbrace{\sum_{k=1}^{N} c_k^{(ij)} b_k}_{\in \mathrm{Span}(\mathcal{B})} + \underbrace{g^{(ij)}}_{\in I}$

- $M_{x_i} = \begin{array}{c c} & \begin{array}{ccc} x_i b_1 & x_i b_j & x_i b_N \end{array} \\ \begin{array}{c} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_N \end{array} & \begin{pmatrix} & \vdots & \\ & & \\ \ldots & c_k^{(ij)} & \ldots \\ & & \\ & \vdots & \end{pmatrix} \end{array}$: **multiplication matrix**

- $G = \{g^{(ij)} \mid i = 1, \ldots, n,\ j = 1, \ldots, N\}$: **(border) base** of I.

Therefore: *To find a base of $\sqrt[\mathbb{R}]{I}$, it suffices to compute a base $\mathcal{B}$ of $\mathcal{A} = R/\sqrt[\mathbb{R}]{I}$ and the multiplication matrices M_{x_i}*

Characterizing border bases

Theorem (Mourrain 1999)

Given:

- $\mathcal{B}$: *set of monomials* **connected to** 1*, i.e.,* $1 \in \mathcal{B}$ *and*
  $\forall m \in \mathcal{B} \;\; \exists m' \in \mathcal{B} \;\; \exists x_i \;\; m = x_i m'$

  $\mathcal{B}^+ = \mathcal{B} \cup x_1\mathcal{B} \cup \ldots \cup x_n\mathcal{B}$: **prolongation** *of* $\mathcal{B}$ *(by one degree).*

  $\partial\mathcal{B} = \mathcal{B}^+ \setminus \mathcal{B}$: **border** *of* $\mathcal{B}$.

- $G \subseteq R$: **rewriting family** *for* $\mathcal{B}$, *permitting to express any monomial of* $\partial\mathcal{B}$ *in* $\mathrm{Span}(\mathcal{B})$.

Then:

$\mathcal{B}$ *is a* **base** *of* $\mathcal{A} = R/(G)$ $\rightsquigarrow$ *[* G *is a* **border base***]*
i.e., any polynomial can be **uniquely** *written in* $\mathrm{Span}(\mathcal{B})$ *modulo the ideal* (G)

$\iff$ **The (formal) multiplication matrices** $M_{x_1}, \ldots, M_{x_n}$ **commute pairwise.**

Count (real) roots and compute $\sqrt{I}$ with the Hermite form

Consider the **Hermite quadratic form:**

$$\begin{array}{rccc} H: & \mathcal{A} & \rightarrow & \mathcal{A} \\ & (f,g) & \mapsto & Tr(M_{fg}) \end{array}$$

Theorem

Let I be a zero-dimensional ideal, i.e., $|V_{\mathbb{C}}(I)| < \infty$.

- $\mathrm{rank}(H) = |V_{\mathbb{C}}(I)|$.
- $\mathrm{Sign}(H) = |V_{\mathbb{R}}(I)|$.
- $\sqrt{I} = \mathrm{Ker}(H) + I$.

Our strategy to compute $\sqrt[\mathbb{R}]{I}$: Work on the dual side

- **Goal:** *Compute the real radical* $\sqrt[\mathbb{R}]{I}$ *of* $I = (h_1, \ldots, h_m)$.
- **Strategy:** *Use the* **dual space** $R^* :=$ *linear forms* Λ *on* R.

$v \in V_{\mathbb{R}}(I) \rightsquigarrow \Lambda_v \in R^*$: **Evaluation** *at* v*, defined by*

$$\Lambda_v(f) = f(v) \quad \text{for} \quad f \in R$$

Properties of Λ_v:

1. Λ_v *vanishes on* I [True for **all complex** v]

   Indeed: $\Lambda_v(h_j x^\alpha) = h_j(v) v^\alpha = 0 \ \forall j \ \forall \alpha$
2. Λ_v *is positive on squares:* $\Lambda_v \succeq 0$ [True **only for real** v]

   Indeed: $\Lambda_v(f^2) = f(v)^2 \geq 0 \ \ \forall f \in R$

Moment matrices and Hankel operators

Definition

The **moment matrix** $M(\Lambda)$ of $\Lambda \in R^*$ is indexed by all monomials:

$$M(\Lambda) = (\Lambda(x^\alpha x^\beta))_{\alpha,\beta\in\mathbb{N}^n}.$$

$\rightsquigarrow$ $M(\Lambda)$ *is a* **generalized Hankel matrix.**

Lemma

Λ positive on squares $\Longleftrightarrow$ $M(\Lambda) \succeq 0$: positive semidefinite matrix.

Proof: $\Lambda(f^2) = \overline{f}^T M(\Lambda)\overline{f}$ where $f = \sum_\alpha f_\alpha x^\alpha$, $\overline{f} = (f_\alpha)$.

Remark

$M(\Lambda)$ is the matrix (in the monomial/dual bases of R, R^*) of the **Hankel operator** *H_Λ: $f \in R \mapsto f \cdot \Lambda \in R^*$, defined by*

$$(f \cdot \Lambda)(g) = \Lambda(fg) \quad \forall g \in R.$$

This lecture: Use the terminology of moment matrices.

Basic properties of moment matrices

$\Lambda \in R^* \rightsquigarrow$

$\color{blue}{\mathrm{Ker}\ M(\Lambda)} = \{f \in R \mid \overline{g}^T M(\Lambda)\overline{f} = 0, \text{ i.e., } \Lambda(fg) = 0\ \forall g \in R\}$

Proposition

1. $\mathrm{Ker}\ M(\Lambda)$ *is an ideal, with* $\dim R/\mathrm{Ker} M(\Lambda) = \mathrm{rank}\ M(\Lambda)$.
   $\mathcal{B}$ *column base of* $M(Ł.a) \Longleftrightarrow \mathcal{B}$ *is a base of* $R/\mathrm{Ker}\ M(\Lambda)$
2. $\Lambda \succeq 0 \Longrightarrow \mathrm{Ker}\ M(\Lambda)$ *is a real radical ideal.*
3. *Assume* $\Lambda \succeq 0$ *and* $\mathrm{rank}\ M(\Lambda) < \infty$. *Then,*
   $\Lambda \in cone\{\Lambda_v \mid v \in \mathbb{R}^n\}$

Proof of (2): Assume $\sum_i p_i^2 \in \mathrm{Ker}\ M(\Lambda)$. Then:

$$0 = \Lambda(\textstyle\sum_i p_i^2) = \sum_i \underbrace{\Lambda(p_i^2)}_{\geq 0} \Longrightarrow \Lambda(p_i^2) = 0 \Longrightarrow p_i \in \mathrm{Ker}\ M(\Lambda)$$

Curto-Fialkow (1996) show **(3)** with functional analysis

Next: Short proof using the Real Nullstellensatz

Short proof

Theorem (Finite rank moment matrix thm, Curto-Fialkow 1996)

Let $\Lambda \in R^*$. *Assume* $M(\Lambda) \succeq 0$ *and* $\mathrm{rank}\ M(\Lambda) = r < \infty$. *Then*

$$\Lambda = \sum_{i=1}^{r} \lambda_i \Lambda_{v_i}$$

for some $\lambda_i > 0$ *and* $\{v_1, \ldots, v_r\} = V(\mathrm{Ker}\ M(\Lambda)) \subseteq \mathbb{R}^n$.

[Λ has a finite atomic representing measure]

Proof:

- $\mathrm{Ker}\ M(\Lambda)$ is real radical [as $M(\Lambda) \succeq 0$]
- $\mathrm{Ker}\ M(\Lambda)$ is zero-dimensional [as $\dim R/\mathrm{Ker}\ M(\Lambda) = r$]
- $V(\mathrm{Ker}\ M(\Lambda)) := \{v_1, \ldots, v_r\} \subseteq \mathbb{R}^n$
- $\mathrm{Ker}\ M(\Lambda) = I(\{v_1, \ldots, v_r\})$
- $\Lambda = \sum_{i=1}^r \Lambda(p_i^2)\Lambda_{v_i}$ [with p_i interpolation polynomials at v_i]

Our strategy: Work with truncated moment matrices

Given: $I = (\underbrace{h_1, \ldots, h_m}_{H})$, $D = \max_j \deg(h_j)$

For $t \geq D$, define:

- **The prolongation of H up to degree t:**

$$\langle H|t\rangle = \{h_j x^\alpha \mid \deg(h_j x^\alpha) \leq t\} \subseteq I \cap R_t.$$

- The **cone of truncated positive linear forms**:

$$\mathcal{L}_{t,\succeq} = \{\Lambda \in R_t^* \mid \Lambda(f) = 0 \ \forall f \in \langle H|t\rangle, \ \underbrace{\Lambda(f^2) \geq 0 \ \forall f \in R_{\lfloor t/2 \rfloor}}_{M_{\lfloor t/2 \rfloor}(\Lambda) \succeq 0}\}$$

Clearly: $\mathcal{L}_{t,\succeq} \supseteq \mathrm{cone}\{\Lambda_v \mid v \in V_{\mathbb{R}}(I)\}$

A crucial geometric property of the cone $\mathcal{L}_{t,\succeq}$

Lemma (Generic linear form)

The following properties are equivalent for $\Lambda \in \mathcal{L}_{t,\succeq}$:

1. *Λ lies in the relative interior of the cone $\mathcal{L}_{t,\succeq}$ (Λ is generic).*
2. $\mathrm{rank}\ M_{\lfloor t/2 \rfloor}(\Lambda)$ *is maximum.*
3. $\mathrm{Ker}\ M_{\lfloor t/2 \rfloor}(\Lambda)$ *is minimum, i.e.,*

$$\underbrace{\mathrm{Ker}\ M_{\lfloor t/2 \rfloor}(\Lambda)}_{\mathcal{N}_t\text{: generic kernel}} \subseteq \mathrm{Ker} M_{\lfloor t/2 \rfloor}(\Lambda') \quad \forall \Lambda' \in \mathcal{L}_{t,\succeq}.$$

Lemma

$\mathcal{N}_t \subseteq \mathcal{N}_{t+1} \subseteq \ldots \subseteq \sqrt[\mathbb{R}]{I}$

Proof: For all $v \in V_{\mathbb{R}}(I)$, $\mathcal{N}_t \subseteq \mathrm{Ker}\ M_{\lfloor t/2 \rfloor}(\Lambda_v) \subseteq I(v)$.

Hence: $\mathcal{N}_t \subseteq I(V_{\mathbb{R}}(I)) = \sqrt[\mathbb{R}]{I}$.

Semidefinite characterization of $\sqrt[\mathbb{R}]{I}$

$$I = (h_1, \ldots, h_m), \quad \mathcal{N}_t \subseteq \sqrt[\mathbb{R}]{I}$$

Theorem

$\sqrt[\mathbb{R}]{I} = (\mathcal{N}_t)$ *for t large enough.*

Sketch of proof: For t large enough, $\mathcal{N}_t$ contains a given basis $\{g_1, \ldots, g_L\}$ of $\sqrt[\mathbb{R}]{I}$:

- Real Nullstellensatz: $g_l^{2m} + \sum_i s_i^2 = \sum_{j=1}^m u_j h_j$
- $\mathcal{N}_t$ is "real ideal like": $g_l^{2m} + \sum_i s_i^2 \in \mathcal{N}_t \Longrightarrow g_l \in \mathcal{N}_t$

Question

How to recognize when $\mathcal{N}_t$ generates $\sqrt[\mathbb{R}]{I}$?

Next: An answer in the case $|V_{\mathbb{R}}(I)| < \infty$

Stopping criterion

$I = (h_1, \dots, h_m)$, $D = \max_j \deg(h_j)$, $d = \lceil D/2 \rceil$, $t \geq D$

Theorem (Stopping criterion)

Let Λ be a generic *element of $\mathcal{L}_{t,\succeq}$.*

Assume one of the following two **flatness conditions** *holds:*

(F1) $\mathrm{rank}\ M_s(\Lambda) = \mathrm{rank}\ M_{s-1}(\Lambda)$ *for some* $s \in [D, \lfloor t/2 \rfloor]$

(Fd) $\mathrm{rank}\ M_s(\Lambda) = \mathrm{rank}\ M_{s-d}(\Lambda)$ *for some* $s \in [d, \lfloor t/2 \rfloor]$

Then:

- $\sqrt[\mathbb{R}]{I} = (\mathrm{Ker} M_s(\Lambda))$
- *Any column base $\mathcal{B}$ of $M_{s-1}(\Lambda)$ is a base of $R/\sqrt[\mathbb{R}]{I}$*
- *The multiplication matrices can be constructed from $M_s(\Lambda)$*

Key tool: Use the **Flat extension thm** of [Curto-Fialkow 1996]:
$\mathrm{rank}\ M_s(\Lambda) = \mathrm{rank}\ M_{s-1}(\Lambda) \Longrightarrow \Lambda$ has a **flat extension** $\tilde{\Lambda} \in R^*$,
i.e., $\mathrm{rank}\ M(\tilde{\Lambda}) = \mathrm{rank}\ M_s(\Lambda)$; thus $\mathrm{Ker}\ M(\tilde{\Lambda}) = (\mathrm{Ker}\ M_s(\Lambda))$.

Termination when $|V_{\mathbb{R}}(I)| < \infty$

Theorem (Termination)

1. *If $V_{\mathbb{R}}(I) = \emptyset$, then $1 \in \mathcal{N}_t$ for some t.*
2. *If $1 \leq |V_{\mathbb{R}}(I)| < \infty$, then the stopping criterion (F1) (or (Fd)) holds for some t.*

Sketch of proof:

1. For $t \geq t_0$, $\mathcal{N}_t = \mathrm{Ker}\ M_{\lfloor t/2 \rfloor}(\Lambda)$ contains a Gröbner base $\{g_1, \ldots, g_L\}$ of $\sqrt[\mathbb{R}]{I}$ (for total degree ordering)
   $\mathcal{B} = \{b_1, \ldots, b_N\}$: standard monomials $\rightsquigarrow$ base of $R/\sqrt[\mathbb{R}]{I}$
2. Let $s := 1 + \max_{b \in \mathcal{B}} \deg(b)$ and $t \geq \max\{t_0, 2s\}$
3. $x^\alpha = \underbrace{\sum_{i=1}^{N} \lambda_i b_i}_{\in \mathrm{Span}(\mathcal{B}),\ \deg \leq s-1} + \underbrace{\sum_{l=1}^{L} u_l g_l}_{\in \sqrt[\mathbb{R}]{I},\ \deg \leq |\alpha| \leq s < \lfloor t/2 \rfloor}$ if $\deg(x^\alpha) \leq s$

   $\rightsquigarrow x^\alpha - \sum^{N} \lambda_i b_i \in \mathrm{Ker} M_{\lfloor t/2 \rfloor}(\Lambda) \rightsquigarrow rk\ M_s(\Lambda) = rk\ M_{s-1}(\Lambda)$

Small example: $I = (h)$, $h = (x_1^2 + x_2^2)^2$

$$
\begin{array}{cc}
 & \begin{array}{cccccc} 1 & x_1 & x_2 & \mathbf{x_1^2} & x_1x_2 & \mathbf{x_2^2} \end{array} \\
\begin{array}{c} 1 \\ x_1 \\ x_2 \\ \mathbf{x_1^2} \\ x_1x_2 \\ \mathbf{x_2^2} \end{array} &
\begin{pmatrix}
\Lambda(1) & \Lambda(x_1) & \Lambda(x_2) & \mathbf{\Lambda(x_1^2)} & \Lambda(x_1x_2) & \mathbf{\Lambda(x_2^2)} \\
\Lambda(x_1) & \mathbf{\Lambda(x_1^2)} & \Lambda(x_1x_2) & \Lambda(x_1^3) & \Lambda(x_1^2x_2) & \Lambda(x_1x_2^2) \\
\Lambda(x_2) & \Lambda(x_1x_2) & \mathbf{\Lambda(x_2^2)} & \Lambda(x_1^2x_2) & \Lambda(x_1x_2^2) & \mathbf{\Lambda(x_2^3)} \\
\mathbf{\Lambda(x_1^2)} & \Lambda(x_1^3) & \Lambda(x_1^2x_2) & \Lambda(x_1^4) & \Lambda(x_1^3x_2) & \mathbf{\Lambda(x_1^2x_2^2)} \\
\Lambda(x_1x_2) & \Lambda(x_1^2x_2) & \Lambda(x_1x_2^2) & \Lambda(x_1^3x_2) & \Lambda(x_1^2x_2^2) & \Lambda(x_1x_2^3) \\
\mathbf{\Lambda(x_2^2)} & \Lambda(x_1x_2^2) & \mathbf{\Lambda(x_2^3)} & \mathbf{\Lambda(x_1^2x_2^2)} & \Lambda(x_1x_2^3) & \Lambda(x_2^4)
\end{pmatrix}
\end{array}
\succeq 0
$$

$\Lambda \in \mathcal{L}_{4,\succeq}$ if $\mathbf{M_2(\Lambda)} \succeq \mathbf{0}$ and $\mathbf{0} = \mathbf{\Lambda(h)} = \underbrace{\Lambda(x_1^4)}_{\geq 0} + \underbrace{\Lambda(x_2^4)}_{\geq 0} + 2\underbrace{\Lambda(x_1^2x_2^2)}_{\geq 0}$

Hence: $\Lambda(x_1^4) = \Lambda(x_2^4) = \Lambda(x_1^2x_2^2) = 0$

Small example: $I = (h)$, $h = (x_1^2 + x_2^2)^2$

$$\begin{array}{cc} & \begin{array}{cccccc} 1 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 \end{array} \\ \begin{array}{c} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1x_2 \\ x_2^2 \end{array} & \begin{pmatrix} \Lambda(1) & \Lambda(x_1) & \Lambda(x_2) & \Lambda(x_1^2) & \Lambda(x_1x_2) & \Lambda(x_2^2) \\ \Lambda(x_1) & \Lambda(x_1^2) & \Lambda(x_1x_2) & \Lambda(x_1^3) & \Lambda(x_1^2x_2) & \Lambda(x_1x_2^2) \\ \Lambda(x_2) & \Lambda(x_1x_2) & \Lambda(x_2^2) & \Lambda(x_1^2x_2) & \Lambda(x_1x_2^2) & \Lambda(x_2^3) \\ \Lambda(x_1^2) & \Lambda(x_1^3) & \Lambda(x_1^2x_2) & 0 & \Lambda(x_1^3x_2) & \Lambda(x_1^2x_2^2) \\ \Lambda(x_1x_2) & \Lambda(x_1^2x_2) & \Lambda(x_1x_2^2) & \Lambda(x_1^3x_2) & 0 & \Lambda(x_1x_2^3) \\ \Lambda(x_2^2) & \Lambda(x_1x_2^2) & \Lambda(x_2^3) & \Lambda(x_1^2x_2^2) & \Lambda(x_1x_2^3) & 0 \end{pmatrix} \succeq 0 \end{array}$$

Hence: $\Lambda(x^\alpha) = 0$ for all $x^\alpha \neq 1$

$\rightsquigarrow \operatorname{rank} M_2(\Lambda) = \operatorname{rank} M_0(\Lambda) = 1$

$\rightsquigarrow \sqrt[\mathbb{R}]{I} = (\operatorname{Ker} M_2(\Lambda)) = (x_1, x_2)$

Moment matrix algorithm for $\sqrt[\mathbb{R}]{I}$

Input: $h_1, \ldots, h_m \in R$ [Assume $|V_{\mathbb{R}}(I)| < \infty$]

Output: base $\mathcal{B}$ of $R/\sqrt[\mathbb{R}]{I}$, multiplication matrices M_{x_i} of $R/\sqrt[\mathbb{R}]{I}$
or an infeasibility certificate if $V_{\mathbb{R}}(I) = \emptyset$

Algorithm: For $t \geq D$

1. Compute a generic element $\Lambda \in \mathcal{L}_{t,\succeq}$
2. Check if (F1) or (Fd) holds
3. If **yes**, return a column base $\mathcal{B}$ of $M_{s-1}(\Lambda)$, $M_{x_i} = {M_{\mathcal{B}}}^{-1} P_i$
   - $M_{\mathcal{B}}$: principal submatrix of $M_{s-1}(\Lambda)$ indexed by $\mathcal{B}$
   - P_i: submatrix of $M_s(\Lambda)$ with rows in $\mathcal{B}$ and columns in $x_i\mathcal{B}$
4. If **no**, go to Step 1 with $t \to t+1$

Theorem (Termination)

If $|V_{\mathbb{R}}(I)| < \infty$, the algorithm terminates.

Compute a generic element in $\mathcal{L}_{t,\succeq}$ with SDP

Solve the SDP: $\min\ 0\ \text{ s.t. } \begin{cases} M_{\lfloor t/2 \rfloor}(\Lambda) \succeq 0 & \text{(PSD)} \\ \Lambda(f) = 0\ \forall f \in \langle H|t \rangle & \text{(L)} \\ \Lambda(1) = 1 & \text{(N)} \end{cases}$

Introduce variables $y_\alpha = \Lambda(x^\alpha)$ for $|\alpha| \leq t$

- Positive semidefinite condition: $\sum_{|\alpha| \leq t} B_\alpha y_\alpha \succeq 0$

  B_α are matrices indexed by monomials of degree $\leq \lfloor t/2 \rfloor$

- Linear condition: $A_t y = 0$

  The coefficient vectors of polynomials in $\langle H|t \rangle$ are rows of A_t

- Normalization condition: $y_0 = 1$

Interior-point algorithms (with self-dual embedding technique) return an optimal solution in the relative interior of the optimal face $\rightsquigarrow$ **generic element** *of* $\mathcal{L}_{t,\succeq}$

Example (from Bini-Mourrain list)

$$I = (5x_1^9 - 6x_1^5x_2 + x_1x_2^4 + 2x_1x_3, -2x_1^6x_2 + 2x_1^2x_2^3 + 2x_2x_3, x_1^2 + x_2^2 - 0.265625)$$

$$D = 9,\ d = 5,\ |V_{\mathbb{R}}(I)| = 8,\ |V_{\mathbb{C}}(I)| = 20$$

order t	rank sequence of $M_s(\Lambda)$ $(0 \le s \le \lfloor t/2 \rfloor)$	extract. order s	accuracy	comm. error
10	1 4 8 16 25 34	—	—	—
12	1 3 9 15 22 26 32	—	—	—
14	1 3 8 10 12 16 20 24	3	0.12786	0.00019754
16	1 4 8 8 8 12 16 20 24	4	4.6789e-5	4.7073e-5

$\mathcal{B} = \{1, x_1, x_2, x_3, x_1^2, x_1x_2, x_1x_3, x_2x_3\} \rightsquigarrow$ border basis G of size 10

Real solutions:

$$\begin{cases} x_1 = (-0.515, -0.000153, -0.0124) & x_2 = (-0.502, 0.119, 0.0124) \\ x_3 = (0.502, 0.119, 0.0124) & x_4 = (0.515, -0.000185, -0.0125) \\ x_5 = (0.262, 0.444, -0.0132) & x_6 = (-2.07\text{e-}5, 0.515, -1.27\text{e-}6) \\ x_7 = (-0.262, 0.444, -0.0132) & x_8 = (-1.05\text{e-}5, -0.515, -7.56\text{e-}7) \end{cases}$$

The moment matrix approach for $\sqrt{I}$

Omit the PSD condition and work with the **linear** *space:*

$$\mathcal{L}_t := \{\Lambda \in R^* \mid \Lambda(f) = 0 \ \forall f \in \langle H|t\rangle\}$$

The same algorithm works: For $t \geq d$

1. Compute a **generic** element $\Lambda \in \mathcal{L}_t$
   [$\mathrm{rank}\ M_s(\Lambda)$ maximum $\forall s \leq \lfloor t/2 \rfloor$]
   [choose random element $\Lambda \in \mathcal{L}_t$]
2. Check if (F1) or (Fd) holds
3. If **yes**, return a base $\mathcal{B}$ of R/J, where $J = (\mathrm{Ker}\ M_s(\Lambda))$ satisfies $I \subseteq J \subseteq \sqrt{I}$ so that $\sqrt{J} = \sqrt{I}$
4. If **no**, go to Step 1 with $t \rightarrow t+1$

Computing the radical ideal $\sqrt{I}$

Given: Generic $\Lambda \in \mathcal{L}_t$ with $\mathrm{rank}\ M_s(\Lambda) = \mathrm{rank}\ M_{s-1}(\Lambda)$,
$J = (\mathrm{Ker}\ M_s(\Lambda))$, $\mathcal{B} = \{b_1, \ldots, b_N\}$ column base of $M_{s-1}(\Lambda)$

Goal: Find the Hermite matrix H of J $\rightsquigarrow$ description of $\sqrt{J}$.

Inspired by [Janovitz-Freireich, Szántó, Mourrain, Rónyai 2008]:

1. Compute the dual base $\{b_1^*, \ldots, b_N^*\}$ of $\mathcal{B}$ s.t. $\Lambda(b_i b_j^*) = \delta_{ij}$
   $[b_1^*, \ldots, b_N^*]^T = M_{\mathcal{B}}^{-1}[b_1, \ldots, b_N]^T$
2. $\Delta :=$ residue of $\sum_{i=1}^N b_i b_i^*$ in $\mathrm{Span}(\mathcal{B})$ modulo J
3. **Claim:** Hermite matrix $H = ST$
   - S: $N \times \mathcal{M}_{2s-2}$ matrix with rows the coefficient vectors of $\Delta b_1, \ldots, \Delta b_N$
   - T: submatrix of $M_s(\Lambda)$ with row indices $\mathcal{M}_{2s-2}$ and column indices $\mathcal{B}$
   - **Using the fact:** $Tr(M_h) = \Lambda(h\Delta)$ for $h \in R$

Example: the real/complex moment matrix algorithm

$$I = (x_1^2 - 2x_1x_3 + 5,\ x_1x_2^2 + x_2x_3 + 1,\ 3x_2^2 - 8x_1x_3),\ \ D = 3,\ d = 2$$

Ranks of $M_s(\Lambda)$ for generic $\Lambda \in \mathcal{L}_t,\ \mathcal{L}_{t,\succeq}$:

	$t=2$	3	4	5	6	7	8	9
$s=0$	1	1	1	1	1	1	1	1
$s=1$	4	4	4	4	4	4	4	4
$s=2$			8	8	8	8	8	**8**
$s=3$					11	10	9	**8**
$s=4$							12	10

no PSD $\rightsquigarrow$ 8 complex roots

	$t=2$	3	4	5	6
$s=0$	1	1	1	1	1
$s=1$	4	4	4	2	**2**
$s=2$			8	8	**2**
$s=3$					10

with PSD $\rightsquigarrow$ 2 real roots

Link to the elimination algorithm of Zhi-Reid

$$I(h_1, \dots, h_m),\ D = \max_j \deg(h_j)$$

Theorem (Zhi-Reid 2004)

If the following **dimension condition** *holds for* $s \in [D, t]$*:*

(D) $\dim \pi_s(\mathcal{L}_t) = \dim_{s-1}(\mathcal{L}_t) = \dim \pi_s(\mathcal{L}_{t+1})$

then one can construct the multiplication matrices of R/I*.*

Theorem (Link to the flatness condition)

The flatness condition for generic $\Lambda \in \mathcal{L}_t$*,* $s \in [D, \lfloor t/2 \rfloor]$*:*

(F1) $\mathrm{rank}\ M_s(\Lambda) = \mathrm{rank} M_{s-1}(\Lambda)$

implies the dimension condition at order $(t, 2s)$*:*

$$\dim \pi_{2s}(\mathcal{L}_t) = \dim_{2s-1}(\mathcal{L}_t) = \dim \pi_{2s}(\mathcal{L}_{t+1})$$

$\rightsquigarrow$ *The stopping criterion (D) might hold earlier than (F1)*

Extension to the real case

Complex case: Dimensions of projections of $\mathcal{L}_t = \langle H|t\rangle^{\perp}$

Real case: Dimension of cone $\mathcal{L}_{t,\succeq} = \dim G_t^{\perp}$

$$G_t := \langle H|t\rangle \cup \{fx^{\alpha} \mid f \in \mathcal{N}_t,\ |\alpha| \leq \lfloor t/2 \rfloor\}$$

Theorem

If the following **dimension condition** *holds for $s \in [D, t]$:*

(D_+) $\quad \dim \pi_s(G_t^{\perp}) = \dim \pi_{s-1}(G_t^{\perp}) = \dim \pi_s((G_t^{+})^{\perp})$

then one can construct the multiplication matrices of R/J, where $I \subseteq J \subseteq \sqrt[\mathbb{R}]{I}$, with equality: $J = \sqrt[\mathbb{R}]{I}$ if $\dim \pi_s(G_t^{\perp}) = |V_{\mathbb{R}}(I)|$.

Theorem (Link to the flatness condition)

(F1) $\operatorname{rank} M_s(\Lambda) = \operatorname{rank} M_{s-1}(\Lambda)$ *for generic $\Lambda \in \mathcal{L}_{t,\succeq}$*

is equivalent *to the dimension condition at order $(t, 2s)$:*

(D_{++}) $\quad \dim \pi_{2s}(G_t^{\perp}) = \dim \pi_{2s-1}(G_t^{\perp}) = \dim \pi_{2s}((G_t^{+})^{\perp})$

Example

$$I = (x_1^2 - 2x_1x_3 + 5, x_1x_2^2 + x_2x_3 + 1, 3x_2^2 - 8x_1x_3)$$

	$t=3$	4	5	6
$s=0$	1	1	1	1
$s=1$	4	4	2	2
$s=2$		8	8	2
$s=3$				10

Two real roots

$$\mathrm{rank} M_2(\Lambda) = \mathrm{rank} M_1(\Lambda) = 2$$
for $\Lambda \in \mathcal{L}_{6,\succeq}$

	G_3	G_3^+	G_4	G_4^+	G_5	G_5^+	G_6	G_6^+
$s=1$	4	4	4	4	2	2	2	2
$s=2$	8	8	8	8	2	2	2	2
$s=3$	11	10	10	9	2	2	2	2
$s=4$			12	10	3	3	2	2

$$\dim \pi_2(G_5^\perp) = \dim \pi_1(G_5^\perp) = \dim \pi_2((G_5^+)^\perp) = 2$$

Outlook

Bottleneck: Solve large SDP problems involving matrices indexed by all monomials up to degree t

Idea: *Combine the SDP based moment matrix approach with border base algorithms to obtain an iterative procedure, involving SDP computations on smaller matrices*

Theorem (Generalized flat extension theorem, La-Mourrain 2009)

Let $\Lambda : \mathcal{B}^+ \cdot \mathcal{B}^+ \to \mathbb{R}$, where $\mathcal{B}$ is connected to 1.
If $\mathrm{rank}\ M_{\mathcal{B}}(\Lambda) = \mathrm{rank}\ M_{\mathcal{B}^+}(\Lambda)$, then Λ has a flat extension to $\tilde{\Lambda} \in R^*$.

$\rightsquigarrow$ **Lecture of Bernard Mourrain this afternoon**

Some references

- J.B. Lasserre, M. Laurent and P. Rostalski.
  Semidefinite characterization and computation of real radical ideals. Foundations of Computational Mathematics, 8(5):607–647, 2008.
- J.B. Lasserre, M. Laurent and P. Rostalski.
  A unified approach for real and complex zeros of zero-dimensional ideals. In Emerging Applications of Algebraic Geometry, M. Putinar and S. Sullivant, eds., vol. 149, pp 125–156, 2009.
- J.B. Lasserre, M. Laurent and P. Rostalski.
  A prolongation-projection algorithm for computing the finite real variety of an ideal. Theoretical Computer Science, 410:2685–2700, 2009.
- M. Laurent and P. Rostalski.
  The approach of moments for polynomial equations. Survey paper, to appear as chapter of *Handbook on Semidefinite, Cone and Polynomial Optimization*.