
Optimization over Polynomials with Moment Matrices and SOS: Algebraic Properties of Moment Matrices

Monique Laurent

CWI, Amsterdam

IMA Tutorial: Algebraic Algorithms in Optimization

Polynomial Ideals and Varieties

$I \subseteq \mathbb{R}[x]$ **ideal** if $f, g \in I, h \in \mathbb{R}[x] \implies f + g \in I, fh \in I$

Finite Basis theorem: I has a **finite set of generators** :

$$I = \langle h_1, \dots, h_m \rangle := \left\{ \sum_{j=1}^m u_j h_j \mid u_j \in \mathbb{R}[x] \right\}$$

• For $V \subseteq \mathbb{C}^n$, its **vanishing ideal**:

$$I(V) := \{ f \in \mathbb{R}[x] \mid f(x) = 0 \forall x \in V \}$$

• For $I \subseteq \mathbb{R}[x]$

$V(I) := \{ x \in \mathbb{C}^n \mid f(x) = 0 \forall f \in I \}$: **complex variety** of I

$V_{\mathbb{R}}(I) := V(I) \cap \mathbb{R}^n$: **real variety** of I

Note: $V(I(V)) = V$, but $I \subseteq I(V(I))$ may be strict

The (Real) Nullstellensatz

radical ideal of I :

$$\sqrt{I} := \{f \in \mathbb{R}[x] \mid \exists m \in \mathbb{N} \text{ s.t. } f^m \in I\}$$

real radical ideal of I :

$$\sqrt{\mathbb{R}}I := \{f \mid \exists m \in \mathbb{N} \exists s \in \text{SOS}_n \text{ s.t. } f^{2m} + s \in I\}$$

$$I \subseteq \sqrt{I} \subseteq I(V(I)), \quad I \subseteq \sqrt{\mathbb{R}}I \subseteq I(V_{\mathbb{R}}(I))$$

Hilbert Nullstellensatz:

$$I(V(I)) = \sqrt{I}$$

Real Nullstellensatz:

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

(Real) Radical Ideal

Definition: Let $I \subseteq \mathbb{R}[x]$ be an ideal .

• I is **radical** if $I = \sqrt{I}$

Equivalently if $f^2 \in I \implies f \in I$

Equivalently if the only polynomials vanishing at $V(I)$ are the polynomials in I

Ex: $I = \langle x^2 \rangle$ is **not** radical, since $x \in \sqrt{I} \setminus I$

• I is **real radical** if $I = \sqrt{\mathbb{R}}I$

Equivalently if $\sum_i f_i^2 \in I \implies f_i \in I$

As $I \subseteq I(V(I)) \subseteq I(V_{\mathbb{R}}(I))$,

I is real radical $\iff I$ is radical and $V(I) \subseteq \mathbb{R}^n$

The quotient space $\mathbb{R}[x]/I$, Gröbner bases, standard monomials

$\mathbb{R}[x]/I :=$ set of equivalence classes $p \bmod I$ (for $p \in \mathbb{R}[x]$)

- Fix a monomial ordering on $\mathbb{R}[x]$.
- $G = \{g_1, \dots, g_k\} \subseteq I$ is a **Gröbner basis** of I if $\forall f \in I$ $\text{LT}(f)$ is divisible by $\text{LT}(g)$ for some $g \in G$
- x^α is a **standard monomial** if x^α is not divisible by $\text{LT}(f)$ $\forall f \in I$ (equivalently, $\forall f \in G$)
- The set \mathcal{B} of standard monomials is a **linear basis of $\mathbb{R}[x]/I$**

Division algorithm: Any $f \in \mathbb{R}[x]$ can uniquely written as

$$f = \sum_{x^\alpha \in \mathcal{B}} \lambda_\alpha x^\alpha + \sum_{h=1}^k u_h g_h, \quad \text{with } \text{LT}(u_h g_h) < \text{LT}(f)$$

Zero-Dimensional Ideal

Definition: I is zero-dimensional if $|V(I)| < \infty$

Theorem:

$$|V(I)| < \infty \iff \dim \mathbb{R}[x]/I < \infty$$

Moreover, $|V(I)| \leq \dim \mathbb{R}[x]/I$, with equality if and only if I is radical

Prove: $\dim \mathbb{R}[x]/I < \infty \implies |V(I)| < \infty$

Set $k := \dim \mathbb{R}[x]/I$.

Consider variable x_1 .

$\{1, x_1, \dots, x_1^k\}$ is linearly dependent in $\mathbb{R}[x]/I$

$\implies \exists \lambda_0, \dots, \lambda_k \in \mathbb{R}$ (not all zero) s.t. $f := \sum_{h=0}^k \lambda_h x_1^h \in I$

$\implies f(v) = 0 \quad \forall v \in V(I)$

$\implies |\{v_1 \mid v \in V(I)\}| < \infty$

Because a univariate nonzero polynomial has *finitely* many roots

Prove: $|V(I)| < \infty \implies \dim \mathbb{R}[x]/I < \infty$

The set \mathcal{B} of standard monomials (w.r.t. some monomial ordering) is a basis of $\mathbb{R}[x]/I$.

Show: $|\mathcal{B}| < \infty$

$$\{v_1 \mid v \in V(I)\} =: \{a_1, \dots, a_k\}$$

$$f(x) := \prod_{j=1}^k (x_1 - a_j)$$

Then, $f \in I(V(I)) = \sqrt{I} \implies f^m \in I$ for some $m \in \mathbb{N}$
 $\implies x_1^{km}$ is the **leading term** of a polynomial in I

Thus, $\alpha_1 < km$ for all $x^\alpha \in \mathcal{B} \implies |\mathcal{B}| < \infty$

Moreover, $|V(I)| \leq \dim \mathbb{R}[x]/I$, with equality if I is radical

$\mathcal{B} = \{p_v \mid v \in V(I)\}$: set of interpolation polynomials at the points of $V(I)$. That is, $p_v(v') = \delta_{v,v'} \quad \forall v, v' \in V(I)$.

- \mathcal{B} is linearly independent in $\mathbb{C}[x]/I$:

$$\sum_{v \in V(I)} \lambda_v p_v \in I \implies \lambda_v = 0 \quad \forall v \in V(I)$$

- If I is radical, \mathcal{B} generates $\mathbb{C}[x]/I$:

$$\forall f \in \mathbb{C}[x], f - \sum_{v \in V(I)} f(v) p_v \in I(V(I)) = I$$

Hence: $\dim \mathbb{C}[x]/I = |\mathcal{B}| = |V(I)|$

Finally: $\dim \mathbb{R}[x]/I = \dim \mathbb{C}[x]/I$

Application: Solving Systems of Polynomial Equations

Problem: Find the (complex) roots to the polynomial equations:

$$h_1(x) = 0, \dots, h_m(x) = 0$$

i.e., find $V(I)$, assuming $I = \langle h_1, \dots, h_m \rangle$ is 0-dimensional

Basic idea:

Reduce to linear algebra in the finite dimensional vector space
 $\mathbb{R}[x]/I$

The eigenvalue method

For $h \in \mathbb{R}[x]$, consider the **multiplication operator**:

$$\begin{aligned} m_h : \mathbb{R}[x]/I &\rightarrow \mathbb{R}[x]/I \\ f \bmod I &\mapsto fh \bmod I \end{aligned}$$

Let M_h be the matrix of m_h w.r.t. a basis \mathcal{B} of $\mathbb{R}[x]/I$, of size $|\mathcal{B}| = \dim \mathbb{R}[x]/I (\geq |V(I)|)$

\rightsquigarrow **companion matrix**

Example: the univariate case

$$I := \langle h_1 := x^3 - 6x^2 + 11x - 6 \rangle$$

$$\mathcal{B} = \{1, x, x^2\}: \text{basis of } \mathbb{R}[x]/I$$

$$\text{multiplication matrix: } M_x = \begin{matrix} & \begin{matrix} x & x^2 & x^3 \end{matrix} \\ \begin{matrix} 1 \\ x \\ x^2 \end{matrix} & \begin{pmatrix} 0 & 0 & 6 \\ 1 & 0 & -11 \\ 0 & 1 & 6 \end{pmatrix} \end{matrix}$$

$$\det(M_x - tI) = -h_1(t)$$

Hence: eigenvalues of M_x = roots of $\det(M_x - tI)$
= roots of h_1

~> **This generalizes to the multivariate case**

Stickelberger Theorem

Say the basis \mathcal{B} of $\mathbb{R}[x]/I$ is a monomial basis, i.e., has the form $\mathcal{B} = \{x^\alpha \mid \alpha \in \mathcal{A}\}$

• For $v \in V(I)$, set $\zeta_v = (v^\alpha)_{\alpha \in \mathcal{A}}$

• Let p_v be interpolation polynomials at the points of $V(I)$ s.t. *each p_v is a linear combination of monomials in the basis \mathcal{B}*

Theorem: The eigenvalues of M_h are $h(v)$ ($v \in V(I)$).

(i) $M_h^T \zeta_v = h(v) \zeta_v$ for all $v \in V(I)$.

(ii) If I is radical, $M_h p_v = h(v) p_v$ for all $v \in V(I)$.

Prove: $M_h^T \zeta_v = h(v) \zeta_v$

$$x^\beta \begin{pmatrix} h(x)x^\alpha \\ \vdots \\ \lambda_\beta \\ \vdots \end{pmatrix} = M_h$$

For $x^\alpha \in \mathcal{B}$,

$$h(x)x^\alpha \equiv \sum_{\beta \in \mathcal{A}} \lambda_\beta x^\beta \pmod{I}$$

The α -th coordinate of $M_h^T \zeta_v$ is equal to

$$\sum_{\beta \in \mathcal{A}} \lambda_\beta v^\beta = h(v)v^\alpha$$

thus equal to the α -th coordinate of $h(v)\zeta_v$

Find the points $v \in V(I)$ from the right eigenvectors ζ_v

[Corless-Gianni-Trager 1997]

1. Pick $h \in \mathbb{R}[\mathbf{x}]$ for which all values $h(v)$ ($v \in V(I)$) are distinct. True with high probability for

$$h(x) = c_1 x_1 + \dots + c_n x_n$$

with random scalars c_1, \dots, c_n

2. Then (if I is radical) the right eigenspaces of M_h are 1-dimensional, thus spanned by ζ_v ($v \in V(I)$)

\rightsquigarrow can compute ζ_v ($v \in V(I)$)

3. Easy to recover $v = (v_1, \dots, v_n)$ from ζ_v

Either $x_i \in \mathcal{B}$; else express x_i in terms of the basis \mathcal{B} in $\mathbb{R}[\mathbf{x}]/I$

Back to Moment Matrices

Objective: Prove the following result of Curto-Fialkow:

Theorem CF0: Assume $M_t(\mathbf{y}) \succeq 0$ and
 $\text{rank}M_t(\mathbf{y}) = \text{rank}M_{t-1}(\mathbf{y}) =: r$.

Then \mathbf{y} has a r -atomic representing measure μ with
 $\text{supp}(\mu) = V(\text{Ker}M_t(\mathbf{y}))$.

Motivation: This gives the Optimality Certificate theorem of Henrion-Lasserre.

We need two results of Curto-Fialkow:

- Finite Rank Theorem [for ∞ moment matrices]
- Flat Extension Theorem [for truncated moment matrices]

Recap on moment matrices

• $\mathbf{y} \in \mathbb{R}^{\mathbb{N}^n} \rightsquigarrow \mathbf{M}(\mathbf{y}) = (\mathbf{y}_{\alpha+\beta})_{\alpha,\beta \in \mathbb{N}^n}$: moment matrix of \mathbf{y}

• If \mathbf{y} has a r -atomic representing measure $\mu = \sum_{i=1}^r \lambda_i \delta_{\mathbf{v}_i}$ ($\lambda_i > 0$), then $\mathbf{y} = \sum_{i=1}^r \lambda_i \zeta_{\mathbf{v}_i}$

$$\implies \mathbf{M}(\mathbf{y}) = \sum_{i=1}^r \lambda_i \zeta_{\mathbf{v}_i} \zeta_{\mathbf{v}_i}^T \succeq \mathbf{0}, \text{ rank } \mathbf{M}(\mathbf{y}) = r$$

Finite Rank Positive Semidefinite Moment Matrices

Finite Rank Theorem: [Curto-Fialkow 1996]

Assume $M(\mathbf{y}) \succeq 0$ and $r := \text{rank } M(\mathbf{y}) < \infty$.

- (i) \mathbf{y} has a (unique) r -atomic nonnegative representing measure μ .
- (ii) $\text{supp}(\mu) = V(\text{Ker } M(\mathbf{y}))$
- (iii) If $\text{rank } M_{t-1}(\mathbf{y}) = r$, then $\text{Ker } M(\mathbf{y}) = \langle \text{Ker } M_t(\mathbf{y}) \rangle$.

Elementary (algebraic) proof following [Laurent 2005]

The kernel of $M(\mathbf{y}) \succeq 0$ is a real radical ideal

Lemma 1: If $M(\mathbf{y}) \succeq 0$, then $I := \text{Ker } M(\mathbf{y})$ is a real radical ideal. Hence, $V(I) \subseteq \mathbb{R}^n$.

Proof: If $f \in I$, $g \in \mathbb{R}[x]$, then

$$(fg)^T M(\mathbf{y})(fg) = (fg^2)^T M(\mathbf{y})f = 0,$$

implying $fg \in I$. Thus I is an ideal.

If $\sum_i f_i^2 \in I$, then

$$\sum_i f_i^T M(\mathbf{y}) f_i = \mathbf{1}^T M(\mathbf{y}) \left(\sum_i f_i^2 \right) = 0,$$

implying $f_i \in I$. Thus I is real radical.

The kernel of $M(\mathbf{y})$ is a 0-dimensional ideal

Let $\mathcal{A} \subseteq \mathbb{N}^n$ indexing a maximum nonsingular principal submatrix of $M(\mathbf{y})$. Thus, $|\mathcal{A}| = r = \text{rank } M(\mathbf{y})$.

Lemma 2: $\mathcal{B} := \{x^\alpha \mid \alpha \in \mathcal{A}\}$ is a basis of $\mathbb{R}[\mathbf{x}]/I$

Proof:

- \mathcal{B} is linearly independent in $\mathbb{R}[\mathbf{x}]/I$ since the columns of $M(\mathbf{y})$ indexed by \mathcal{A} are linearly independent
- \mathcal{B} spans $\mathbb{R}[\mathbf{x}]/I$ since any column of $M(\mathbf{y})$ is a linear combination of its columns indexed by \mathcal{A}

Corollary: $r = \dim \mathbb{R}[\mathbf{x}]/I = |V(I)|$

Constructing the representing measure

$p_v :=$ interpolation polynomials at $v \in V(I)$.

$\{p_v \mid v \in V(I)\}$ is a basis of $\mathbb{R}[x]/I$

$$\textbf{Lemma 3: } M(y) = \sum_{v \in V(I)} \underbrace{p_v^T M(y) p_v}_{\lambda_v} \zeta_v \zeta_v^T.$$

That is, $\mu = \sum_{v \in V(I)} \lambda_v \delta_v$ is a representing measure for y with $\text{supp}(\mu) = V(I)$.

End of proof of the Finite Rank Theorem

Lemma 4: Assume $\text{rank}M_{t-1}(y) = \text{rank}M(y)$. Then $\text{Ker}M(y) = \langle \text{Ker}M_t(y) \rangle$.

Proof: $I := \text{Ker}M(y) \stackrel{?}{\subseteq} \langle \text{Ker}M_t(y) \rangle =: J$.

- \exists basis $\mathcal{B} = \{x^\alpha \mid \alpha \in \mathcal{A}\}$ of $\mathbb{R}[x]/I$ with $\mathcal{A} \subseteq \mathbb{N}_{t-1}^n$.
- \mathcal{B} generates $\mathbb{R}[x]/J$? For this, show:

$$x^\gamma = \sum_{\alpha \in \mathcal{A}} \lambda_\alpha x^\alpha + q \quad \text{with } q \in J$$

by induction on $|\gamma|$. Obvious for $|\gamma| \leq t$.

For $|\gamma| \geq t + 1$, write $x^\gamma = x_1 x^\delta$. Then

$$x^\gamma = x_1 \left(\sum_{\alpha \in \mathcal{A}} \lambda_\alpha x^\alpha + q \right) = \sum_{\alpha \in \mathcal{A}} \lambda_\alpha \underbrace{x_1 x^\alpha}_{\text{deg} \leq t} + \underbrace{x_1 q}_{\in J}$$

The Flat Extension Theorem for Truncated Moment Matrices

The Flat Extension Theorem: [Curto-Fialkow 1996]

Assume $M_t(\mathbf{y}) \succeq 0$ and $\text{rank}M_t(\mathbf{y}) = \text{rank}M_{t-1}(\mathbf{y})$.

Then one can extend \mathbf{y} to a (unique) vector $\tilde{\mathbf{y}} \in \mathbb{R}^{\mathbb{N}_{2t+2}^n}$ in such a way that $\text{rank}M_{t+1}(\tilde{\mathbf{y}}) = \text{rank}M_t(\mathbf{y})$.

Say that $\mathbf{X} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{pmatrix}$ is a **flat extension** of \mathbf{A} if $\text{rank}\mathbf{X} = \text{rank}\mathbf{A}$. Then, $\mathbf{X} \succeq 0 \iff \mathbf{A} \succeq 0$.

Main tool: ‘Ideal-like’ property of $M_t(\mathbf{y})$:

$$f \in \text{Ker}M_t(\mathbf{y}), g \in \mathbb{R}[\mathbf{x}], \deg(fg) \leq t \implies fg \in \text{Ker}M_t(\mathbf{y})$$

Application 1: Theorem CF0

Theorem CF0: Assume $M_t(\mathbf{y}) \succeq 0$ and
 $\text{rank}M_t(\mathbf{y}) = \text{rank}M_{t-1}(\mathbf{y}) =: r$.

Then \mathbf{y} has a r -atomic representing measure μ with
 $\text{supp}(\mu) = V(\text{Ker}M_t(\mathbf{y}))$.

Proof:

- Repeatedly apply the **Flat Extension Theorem** to extend \mathbf{y} to $\tilde{\mathbf{y}} \in \mathbb{R}^{N^n}$ with $\text{rank}M(\tilde{\mathbf{y}}) = \text{rank}M_t(\mathbf{y})$.
- As $\text{rank}M(\tilde{\mathbf{y}}) < \infty$, apply the **Finite Rank Theorem**.

Application 2: Optimization with Constraints

$$K = \{x \in \mathbb{R}^n \mid h_1(x) \geq 0, \dots, h_m(x) \geq 0\}$$

$$d_j = \lceil \deg(h_j)/2 \rceil, d = \max_j d_j$$

Theorem CF1: [Curto-Fialkow 2000]

Assume $M_t(\mathbf{y}) \succeq 0$, $M_{t-d_j}(h_j \mathbf{y}) \succeq 0 \quad \forall j \leq m$ and
 $\text{rank } M_t(\mathbf{y}) = \text{rank } M_{t-d}(\mathbf{y}) =: r$.

Then, \mathbf{y} has a r -atomic representing measure μ on K
with $\text{supp}(\mu) = V(\text{Ker } M_t(\mathbf{y}))$.

Proof of Theorem CF1 [Laurent 2005]

By Theorem CF0: y has a representing measure

$$\mu = \sum_{i=1}^r \lambda_i \delta_{v_i} \text{ with } \lambda_i > 0.$$

To show: $v_i \in K$, i.e., $h_j(v_i) \geq 0 \quad \forall j \leq m$

- \exists interpolation polynomials p_1, \dots, p_r at v_1, \dots, v_r with $\deg(p_i) \leq t - d$

Indeed, if \mathcal{B} is a basis of $\mathbb{R}[x]/\text{Ker } M(y)$ contained in $\mathbb{R}[x]_{t-d}$, may replace p_i by its residue mod. I w.r.t. \mathcal{B} .

$$\begin{aligned} \mathbf{0} \leq p_i^T M_{t-d}(h_j y) p_i &= \sum_{k=1}^r \lambda_k (p_i(v_k))^2 h_j(v_k) \\ &= \lambda_i h_j(v_i) \end{aligned}$$

$$\implies h_j(v_i) \geq 0$$

Back to the Optimality Certificate of Henrion-Lasserre

$$p_{\min} := \inf_{x \in K} p(x), \quad K = \{x \in \mathbb{R}^n \mid h_1(x) \geq 0 \dots h_m(x) \geq 0\}$$

Moment relaxation (MOMt):

$$p_t^* := \inf p^T y \quad \text{s.t.} \quad y_0 = 1, \quad M_t(y) \succeq 0 \\ M_{t-d_j}(h_j y) \succeq 0 \quad (j = 1, \dots, m)$$

for $t \geq \max(\lceil \deg(p)/2 \rceil, d)$, $d = \max_j d_j$

Optimality Certificate Theorem:

Let y be an optimum solution to (MOMt).
If (RC) $\text{rank } M_t(y) = \text{rank } M_{t-d}(y)$, then $p_t^* = p_{\min}$
and $V(\text{Ker } M_t(y)) \subseteq \{\text{global minimizers}\}$.

Find $V(\text{Ker}M_t(\mathbf{y}))$ with the eigenvalue method

1. Construct a basis \mathcal{B} of $\mathbb{R}[x]/I$: Any basis of the column space of $M_{t-d}(\mathbf{y})$

2. Construct the multiplication matrices M_{x_i} in $\mathbb{R}[x]/I$ **directly** from $M_t(\mathbf{y})$:

$$M_{x_i} = M_{\mathcal{B}}^{-1} P_{x_i}$$

- $M_{\mathcal{B}}$: $\mathcal{B} \times \mathcal{B}$ submatrix of $M_t(\mathbf{y})$
- P_{x_i} : $\mathcal{B} \times x_i \mathcal{B}$ submatrix of $M_t(\mathbf{y})$

3. Compute the eigenvalues of $M_h = \sum_{i=1}^n c_i M_{x_i}$ for a random linear combination $h = \sum_{i=1}^n c_i x_i$. If all eigenspaces are 1-dimensional, deduce $V(\text{Ker}M_t(\mathbf{y}))$ from the eigenvectors.

Finite Convergence in the Finite Real Variety Case

Theorem: [La 2002] [Lasserre-La-Rostalski 2006]

Assume the equations: $h_1(x) = 0 \dots h_k(x) = 0$ are present in the description of K and have **finitely many real roots**.

For t large enough, **(RC) holds** (for any feasible solution to (MOMt) and thus $p_t^* = p_{\min}$).

Sketch of proof: Let y feasible for (MOMt), $I := \langle h_1, \dots, h_k \rangle$.

Claim 1: For $d \geq 2t$,

$$M_{t-d_j}(h_j y) = 0 \implies h_j \in \text{Ker} M_t(y) \quad (j = 1, \dots, k)$$

Sketch of proof (continued)

$\{g_1, \dots, g_L\}$: Gröbner basis of $I(V_{\mathbb{R}}(I))$ (for total degree monomial ordering)

\mathcal{B} : basis of $\mathbb{R}[x]/I(V_{\mathbb{R}}(I))$, $d_{\mathcal{B}} := \max_{b \in \mathcal{B}} \deg(b)$

Claim 2: $\{g_1, \dots, g_L\} \subseteq \text{Ker}M_t(\mathbf{y})$ for $t \geq t_0$

Proof: By the Real Nullstellensatz: $\exists m_l \in \mathbb{N}$, $s_l \in \text{SOS}_n$ s.t.

$$g_l^{2m_l} + s_l \in I, \text{ i.e., } g_l^{2m_l} + s_l = \sum_{j=1}^k u_j h_j$$

$$h_1, \dots, h_k \in \text{Ker}M_t(\mathbf{y}) \xrightarrow{t \text{ large}} \sum_{j=1}^k u_j h_j \in \text{Ker}M_t(\mathbf{y})$$

$$\implies g_l^{2m_l} + s_l \in \text{Ker}M_t(\mathbf{y}) \implies g_l \in \text{Ker}M_t(\mathbf{y})$$

End of proof

Claim 3: For $t \geq s + 1$, $s := \max(d_{\mathcal{B}} + d, t_0)$,
 $\text{rank}M_s(\mathbf{y}) = \text{rank}M_{s-d}(\mathbf{y})$

Proof: For $|\alpha| \leq s$, write:

$$x^\alpha = \underbrace{\sum_{\mathcal{B}} \lambda_\beta x^\beta}_{\text{deg} \leq d_{\mathcal{B}} \leq s-d} + \underbrace{\sum_{l=1}^L u_l g_l}_{\text{deg} \leq |\alpha| \leq s \leq t-1}$$

$$\implies \sum_{l=1}^L u_l g_l \in \text{Ker}M_t(\mathbf{y})$$

\implies each column of $M_s(\mathbf{y})$ is a linear combination of columns of $M_{s-d}(\mathbf{y})$

Application: Solving polynomial equations in \mathbb{R}^n

Problem: Solve: $h_1(\mathbf{x}) = 0 \dots h_k(\mathbf{x}) = 0$ in \mathbb{R}^n

Consider the ideal $I = \langle h_1, \dots, h_k \rangle$

- If $|V(I)| < \infty$, apply the eigenvalue method (in $\mathbb{R}[\mathbf{x}]/I$)

\rightsquigarrow needs a basis of $\mathbb{R}[\mathbf{x}]/I$

thus a Gröbner basis of I ...

- What if $|V(I)| = \infty$?

But $|V_{\mathbb{R}}(I)| < \infty \dots$

This is an instance of (POP)

$$\min 1 \text{ s.t. } h_1(x) = 0 \dots h_k(x) = 0$$

Let \mathbf{y} be a maximum rank optimum solution to (MOMt). Then,

- (RC) $\implies V_{\mathbb{R}}(I) = V(\text{Ker}M_t(\mathbf{y}))$, which can be found with the eigenvalue method (in $\mathbb{R}[\mathbf{x}]/I(V_{\mathbb{R}}(I))$)
- (RC) holds for t large enough.

Advantages:

1. Complex roots are filtered out!
1. A basis of $\mathbb{R}[\mathbf{x}]/I(V_{\mathbb{R}}(I))$ is read directly from $M_t(\mathbf{y})$
2. Can find a (border) basis of $I(V_{\mathbb{R}}(I))$ (from $M_t(\mathbf{y})$)

Katsura 5 example (extraction possible only with a SVD quotient basis)

$$I = \langle 2x_6^2 + 2x_5^2 + 2x_4^2 + 2x_3^2 + 2x_2^2 + x_1^2 - x_1, x_6x_5 + x_5x_4 + 2x_4x_3 + 2x_3x_2 + 2x_2x_1 - x_2, 2x_6x_4 + 2x_5x_3 + 2x_4x_2 + x_2^2 + 2x_3x_1 - x_3, 2x_6x_3 + 2x_5x_2 + 2x_3x_2 + 2x_4x_1 - x_4, x_3^2 + 2x_6x_1 + 2x_5x_1 + 2x_4x_1 - x_5, 2x_6 + 2x_5 + 2x_4 + 2x_3 + 2x_2 + x_1 - 1 \rangle$$

$d = 1, |V_{\mathbb{R}}(I)| = 12, |V(I)| = 32$

order t	rank sequence	extract. order	accuracy	comm. error
2	1 6 16	—	—	—
3	1 6 12 12	3(3)	1.1928e-5	2.3073e-7

Real solutions:

(0.277, 0.226, 0.162, 0.0858, 0.0115, -0.124)	(0.59, 0.0422, 0.327, -0.0642, -0.0874, -0.0874)
(1, -2.8e-7, 4.7e-7, 8.81e-7, -2.79e-6, -3.69e-6)	(0.239, 0.0608, -0.0622, -0.0233, 0.186, 0.186)
(0.441, 0.151, 0.0225, 0.219, 0.0935, -0.207)	(0.726, -0.0503, 0.122, 0.164, 0.11, -0.208)
(0.462, 0.309, 0.0553, -0.102, -0.0844, 0.0917)	(0.292, -0.101, 0.181, -0.0591, 0.193, 0.14)
(0.753, 0.0532, 0.191, -0.114, -0.146, 0.139)	(0.409, -0.0732, 0.0657, -0.127, 0.252, 0.1)
(0.68, 0.266, -0.154, 0.0323, 0.0897, -0.0735)	(0.136, 0.0428, 0.0417, 0.0404, 0.0964, 0.21)

An example from the Bini-Mourrain test suite: extraction possible without full (RC)

$$I = \langle 5x_1^9 - 6x_1^5x_2 + x_1x_2^4 + 2x_1x_3, -2x_1^6x_2 + 2x_1^2x_2^3 + 2x_2x_3, x_1^2 + x_2^2 - 0.265625 \rangle$$

$$d = 5, |\mathbb{V}_{\mathbb{R}}(I)| = 8, |\mathbb{V}(I)| = 20$$

order t	rank sequence	extract. order MON/SVD	accuracy MON/SVD	comm. error MON/SVD
5	1 4 8 16 25 34	—	—	—
6	1 3 9 15 22 26 32	—	—	—
7	1 3 8 10 12 16 20 24	3(3)/—(—)	0.12786/—	0.00019754/—
8	1 4 8 8 8 12 16 20 24	4(3)/3(3)	4.6789e-5/1.3406e-4	4.7073e-5/7.5005e-4

Real solutions:

$$\left\{ \begin{array}{ll} \mathbf{x}_1 = (-0.515, -0.000153, -0.0124) & \mathbf{x}_2 = (-0.502, 0.119, 0.0124) \\ \mathbf{x}_3 = (0.502, 0.119, 0.0124) & \mathbf{x}_4 = (0.515, -0.000185, -0.0125) \\ \mathbf{x}_5 = (0.262, 0.444, -0.0132) & \mathbf{x}_6 = (-2.07e-5, 0.515, -1.27e-6) \\ \mathbf{x}_7 = (-0.262, 0.444, -0.0132) & \mathbf{x}_8 = (-1.05e-5, -0.515, -7.56e-7) \end{array} \right.$$

Choosing another objective function

$$\min \text{Tr } M_t(\mathbf{y}) \quad \text{s.t. } M_t(\mathbf{y}) \succeq 0, \quad M_{t-d_j}(h_j \mathbf{y}) = 0 \quad (j = 1, \dots, m)$$

↪ optimum solution with **small** rank

order t	rank sequence	extracted points
5	1 3 4 4 4 4	not feasible
6	1 2 2 2 2 2 2	$(-0.2619, 0.4439, -0.0132)$ $(0.2619, 0.4439, -0.0132)$