Moment Matrices, Trace Matrices and the Radical of Ideals

Agnes Szanto
North Carolina State University

In collaboration with

Itnuit Janovitz-Freireich (North Carolina State University)

Bernard Mourrain (GALAAD, INRIA),

Lajos Rónyai (Hungarian Academy of Sciences and Budapest University of Technology and Economics)

The problem

Given: $f_1, \ldots, f_s \in \mathbb{C}[\mathbf{x}]$ polynomials in $\mathbf{x} = (x_1, \ldots, x_m)$ generating an ideal I.
Assume that I has finitely many roots in $\mathbb{C}^m$.
Suppose I either has roots with multiplicities or form clusters with radius $\varepsilon > 0$.

The problem

Given: $f_1, \ldots, f_s \in \mathbb{C}[\mathbf{x}]$ polynomials in $\mathbf{x} = (x_1, \ldots, x_m)$ generating an ideal I.
Assume that I has finitely many roots in $\mathbb{C}^m$.
Suppose I either has roots with multiplicities or form clusters with radius $\varepsilon > 0$.

We compute an **approximate radical** of I, an ideal which has exactly one root for each cluster, corresponding to the arithmetic mean of the cluster, up to an error term asymptotically bound by ε^2.

The problem

Given: $f_1, \ldots, f_s \in \mathbb{C}[\mathbf{x}]$ polynomials in $\mathbf{x} = (x_1, \ldots, x_m)$ generating an ideal I.
Assume that I has finitely many roots in $\mathbb{C}^m$.
Suppose I either has roots with multiplicities or form clusters with radius $\varepsilon > 0$.

We compute an **approximate radical** of I, an ideal which has exactly one root for each cluster, corresponding to the arithmetic mean of the cluster, up to an error term asymptotically bound by ε^2.

The method's computationally most expensive part is computing a **matrix of traces**.

The problem

Given: $f_1, \ldots, f_s \in \mathbb{C}[\mathbf{x}]$ polynomials in $\mathbf{x} = (x_1, \ldots, x_m)$ generating an ideal I.
Assume that I has finitely many roots in $\mathbb{C}^m$.
Suppose I either has roots with multiplicities or form clusters with radius $\varepsilon > 0$.

We compute an **approximate radical** of I, an ideal which has exactly one root for each cluster, corresponding to the arithmetic mean of the cluster, up to an error term asymptotically bound by ε^2.

The method's computationally most expensive part is computing a **matrix of traces**.

We propose a simple method using **Sylvester matrices** to compute matrices of traces.

Related previous work

- Global methods for approximate square-free factorization (univariate case): Sasaki and Noda (1989), Hribernig and Stetter (1997), Kaltofen and May (2003), Zeng (2003), Corless, Watt and Zhi (2004).
- Exact radical computation using trace matrices: Dickson (1923), González-Vega and Trujillo (1994,1995), Armendáriz and Solernó (1995), Becker and Wörmann (1996)
- Local methods to handle near root multiplicities
  - Using eigenvalue computations: Manocha and Demmel (1995), Corless, Gianni and Trager (1997).
  - Using Newton method or deflation: Ojica, Watanabe and Mitsui (1983), Ojica (1987), Lecerf (2002), Giusti, Lecerf, Salvy and Yakoubsohn (2004), Leykin, Verschelde and Zhao (2005).
  - Using dual bases: Stetter (1996) and (2004), Dayton and Zeng (2005), Zhi (2008).

Multiplication matrices

Definition

Let $I = \langle f_1, \ldots, f_s \rangle$ be and ideal for which $A = \mathbb{C}[\mathbf{x}]/I$ is finite dimensional. Let $B = [b_1, \ldots, b_n]$ be a basis of A. The multiplication matrix M_h is the transpose of the matrix of the map

$$m_h \ : \ A \to A, \quad [g] \mapsto [hg]$$

written in the basis B.

Expressions in the roots

Let $\mathbf{z}_1, \ldots, \mathbf{z}_n \in \mathbb{C}^m$ be the roots of I and $B = [b_1, \ldots, b_n]$ be a basis of $A = \mathbb{C}[\mathbf{x}]/I$. Define the Vandermonde matrix

$$V := [b_j(\mathbf{z}_i)]_{i,j=1}^{n} \in \mathbb{C}^{n \times n}.$$

Expressions in the roots

Let $\mathbf{z}_1, \ldots, \mathbf{z}_n \in \mathbb{C}^m$ be the roots of I and $B = [b_1, \ldots, b_n]$ be a basis of $A = \mathbb{C}[\mathbf{x}]/I$. Define the Vandermonde matrix

$$V := [b_j(\mathbf{z}_i)]_{i,j=1}^n \in \mathbb{C}^{n \times n}.$$

Fact

If V is invertible then

$$M_h = V \ \operatorname{diag}(h(\mathbf{z}_1), \ldots, h(\mathbf{z}_n)) \ V^{-1},$$

i.e. he multiplication matrices M_h are simultaneously diagonalizable with $h(\mathbf{z}_1), \ldots, h(\mathbf{z}_n)$ eigenvalues.

Expressions in the roots

Let $\mathbf{z}_1, \ldots, \mathbf{z}_n \in \mathbb{C}^m$ be the roots of I and $B = [b_1, \ldots, b_n]$ be a basis of $A = \mathbb{C}[\mathbf{x}]/I$. Define the Vandermonde matrix

$$V := [b_j(\mathbf{z}_i)]_{i,j=1}^n \in \mathbb{C}^{n\times n}.$$

Fact

If V is invertible then

$$M_h = V \ \mathrm{diag}(h(\mathbf{z}_1), \ldots, h(\mathbf{z}_n)) \ V^{-1},$$

i.e. he multiplication matrices M_h are simultaneously diagonalizable with $h(\mathbf{z}_1), \ldots, h(\mathbf{z}_n)$ eigenvalues.

Note: If I has multiple roots then M_h is not diagonalizable. Also, its entries are not continuous near root multiplicites.

Expressions in the roots

Let $\mathbf{z}_1, \ldots, \mathbf{z}_n \in \mathbb{C}^m$ be the roots of I and $B = [b_1, \ldots, b_n]$ be a basis of $A = \mathbb{C}[\mathbf{x}]/I$. Define the Vandermonde matrix

$$V := [b_j(\mathbf{z}_i)]_{i,j=1}^n \in \mathbb{C}^{n\times n}.$$

Fact

If V is invertible then

$$M_h = V \ \operatorname{diag}(h(\mathbf{z}_1), \ldots, h(\mathbf{z}_n)) \ V^{-1},$$

i.e. he multiplication matrices M_h are simultaneously diagonalizable with $h(\mathbf{z}_1), \ldots, h(\mathbf{z}_n)$ eigenvalues.

Note: If I has multiple roots then M_h is not diagonalizable. Also, its entries are not continuous near root multiplicites.

Goal: Compute multiplication matrices for the radical $\sqrt{I}$.

Matrix of traces

Definition

Let $B = [b_1, \ldots, b_n]$ be a basis of $A = \mathbb{C}[\mathbf{x}]/I$. The matrix of traces is the $n \times n$ symmetric matrix:

$$R = [Tr(b_i b_j)]_{i,j=1}^{n}$$

where $Tr(b_i b_j)$ is the trace of the multiplication matrix $M_{b_i b_j}$.

Matrix of traces

Definition

Let $B = [b_1, \ldots, b_n]$ be a basis of $A = \mathbb{C}[\mathbf{x}]/I$. The matrix of traces is the $n \times n$ symmetric matrix:

$$R = [Tr(b_i b_j)]_{i,j=1}^n$$

where $Tr(b_i b_j)$ is the trace of the multiplication matrix $M_{b_i b_j}$.

Fact

$$R = V \cdot V^T,$$

where $V := [b_i(\mathbf{z}_j)]_{i,j=1}^n$ is the Vandermonde matrix for the roots $\mathbf{z}_1, \ldots, \mathbf{z}_n \in \mathbb{C}^m$ of I. Moreover

$$\mathrm{rank}(R) = \#\{\text{ distinct roots of } I\} = \dim \mathbb{C}[\mathbf{x}]/\sqrt{I}.$$

Matrix of traces

Definition

Let $B = [b_1, \ldots, b_n]$ be a basis of $A = \mathbb{C}[\mathbf{x}]/I$. The matrix of traces is the $n \times n$ symmetric matrix:

$$R = [Tr(b_i b_j)]_{i,j=1}^n$$

where $Tr(b_i b_j)$ is the trace of the multiplication matrix $M_{b_i b_j}$.

Fact

$$R = V \cdot V^T,$$

where $V := [b_i(\mathbf{z}_j)]_{i,j=1}^n$ is the Vandermonde matrix for the roots $\mathbf{z}_1, \ldots, \mathbf{z}_n \in \mathbb{C}^m$ of I. Moreover

$$\mathrm{rank}(R) = \#\{\text{ distinct roots of } I\} = \dim \mathbb{C}[\mathbf{x}]/\sqrt{I}.$$

Note: R is continuous around root multiplicities. We will use a maximal non-singular submatrix of R to eliminate multiplicities.

Dickson's Lemma

Theorem (Dickson (1923))

Let $B = [b_1, \ldots, b_n]$ *be a basis of* $A = \mathbb{C}[\mathbf{x}]/I$. *An element*

$$r = \sum_{k=1}^{n} c_k b_k$$

is in $\mathrm{Rad}(A) = \sqrt{I}/I$ *if and only if* $[c_1, \ldots, c_n]$ *is in the nullspace of the matrix of traces* R.

Dickson's Lemma

Theorem (Dickson (1923))

Let $B = [b_1, \ldots, b_n]$ be a basis of $A = \mathbb{C}[\mathbf{x}]/I$. An element

$$r = \sum_{k=1}^{n} c_k b_k$$

is in $\mathrm{Rad}(A) = \sqrt{I}/I$ if and only if $[c_1, \ldots, c_n]$ is in the nullspace of the matrix of traces R.

Corollary

Let $R = [Tr(b_i b_j)]_{i,j=1}^{n}$ and define $R_{x_k} := [Tr(x_k b_i b_j)]_{i,j=1}^{n}$ for $k = 1, \ldots, m$.
If $\tilde{R}$ is a maximal non-singular submatrix of R, and $\tilde{R}_{x_k}$ is the submatrix of R_{x_k} with the same row and column indices as in $\tilde{R}$, then the solution $\tilde{M}_{x_k}$ of the linear matrix equation

$$\tilde{R}\tilde{M}_{x_k} = \tilde{R}_{x_k}$$

is a multiplication matrix of x_k for the radical of $\sqrt{I}$.

Clusters of roots

We consider systems for which the common roots form clusters of roots.

Clusters of roots

We consider systems for which the common roots form clusters of roots.

Definition

Let $\mathbf{z}_i \in \mathbb{C}^m$ for $i = 1, \ldots, k$, and consider k clusters $C_1, \ldots, C_k$ of size $|C_i| = n_i$ such that $\sum_{i=1}^k n_i = n$, each of radius proportional to the parameter ε around $\mathbf{z}_1, \ldots, \mathbf{z}_k$:

$$C_i = \{\mathbf{z}_i + \delta_{i,1}\varepsilon, \ldots, \mathbf{z}_i + \delta_{i,n_i}\varepsilon\},$$

where all the coordinates of $\delta_{i,j}$ are less than 1 for all i, j.

Clusters of roots

We consider systems for which the common roots form clusters of roots.

Definition

Let $\mathbf{z}_i \in \mathbb{C}^m$ for $i = 1, \ldots, k$, and consider k clusters $C_1, \ldots, C_k$ of size $|C_i| = n_i$ such that $\sum_{i=1}^k n_i = n$, each of radius proportional to the parameter ε around $\mathbf{z}_1, \ldots, \mathbf{z}_k$:

$$C_i = \{\mathbf{z}_i + \delta_{i,1}\varepsilon, \ldots, \mathbf{z}_i + \delta_{i,n_i}\varepsilon\},$$

where all the coordinates of $\delta_{i,j}$ are less than 1 for all i, j.

In this setting we will use trace matrices to define an approximate radical.

GECP and SVD for the matrix of traces

GECP and SVD for the matrix of traces

Proposition

The U_k be the matrix obtained after k steps of the Gaussian Elimination with Complete Pivoting (GECP) on R for a system with k clusters is of the form

$$\begin{bmatrix} [U_k]_{1,1} & & \cdots & \cdots & \cdots & [U_k]_{1,n} \\ 0 & \ddots & \cdots & \cdots & \cdots & \vdots \\ & & [U_k]_{k,k} & \cdots & \cdots & [U_k]_{k,n} \\ \vdots & & 0 & c_{k+1,k+1}\varepsilon^2 & \cdots & c_{k+1,n}\varepsilon^2 \\ & & \vdots & \vdots & \ddots & \vdots \\ 0 & & 0 & c_{n,k+1}\varepsilon^2 & \cdots & c_{n,n}\varepsilon^2 \end{bmatrix} + h.o.t.(\varepsilon).$$

GECP and SVD for the matrix of traces

Proposition

The U_k be the matrix obtained after k steps of the Gaussian Elimination with Complete Pivoting (GECP) on R for a system with k clusters is of the form

$$\begin{bmatrix} [U_k]_{1,1} & \cdots & \cdots & \cdots & [U_k]_{1,n} \\ 0 & \ddots & \cdots & \cdots & \vdots \\ & [U_k]_{k,k} & \cdots & \cdots & [U_k]_{k,n} \\ \vdots & 0 & c_{k+1,k+1}\varepsilon^2 & \cdots & c_{k+1,n}\varepsilon^2 \\ & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & c_{n,k+1}\varepsilon^2 & \cdots & c_{n,n}\varepsilon^2 \end{bmatrix} + h.o.t.(\varepsilon).$$

Proposition

Let $\sigma_1 \geq \cdots \geq \sigma_n$ be the singular values of R. Then

$$\sigma_{k+1} = C\,\varepsilon^2 + h.o.t.(\varepsilon).$$

Multiplication matrices for the approximate radical

Definition

Let $\tilde{R}$ be a maximal numerically non-singular submatrix of R, and $\tilde{R}_{x_i}$ is the submatrix of R_{x_i} with the same row and column indices as in $\tilde{R}$. Then the solution $\tilde{M}_{x_i}$ of the linear matrix equation

$$\tilde{R}\tilde{M}_{x_i} = \tilde{R}_{x_i}$$

is the multiplication matrix of x_i defining the approximate radical.

Multiplication matrices for the approximate radical

Definition

Let $\tilde{R}$ be a maximal numerically non-singular submatrix of R, and $\tilde{R}_{x_i}$ is the submatrix of R_{x_i} with the same row and column indices as in $\tilde{R}$. Then the solution $\tilde{M}_{x_i}$ of the linear matrix equation

$$\tilde{R}\tilde{M}_{x_i} = \tilde{R}_{x_i}$$

is the multiplication matrix of x_i defining the approximate radical.

Theorem

Modulo ε^2 the multiplication matrices $\tilde{M}_{x_1}, \ldots, \tilde{M}_{x_m}$ form a pairwise commuting system of matrices for the roots $\xi_1, \ldots, \xi_k$ satisfying

$$\xi_s = \mathbf{z}_s + \frac{\sum_{r=1}^{n_s} \delta_{s,r}}{n_s}\varepsilon \pmod{\varepsilon^2}.$$

Example

Consider the polynomial system:

$$
\begin{aligned}
f_1 &= x_1^2 + 3.99980 x_1 x_2 - 5.89970 x_1 + 3.81765 x_2^2 - 11.25296 x_2 \\
&\quad + 8.33521 \\
f_2 &= x_1^3 + 12.68721 x_1^2 x_2 - 2.36353 x_1^2 + 81.54846 x_1 x_2^2 - 177.31082 x_1 x_2 \\
&\quad + 73.43867 x_1 - x_2^3 + 6 x_2^2 + x_2 + 5 \\
f_3 &= x_1^3 + 8.04041 x_1^2 x_2 - 2.16167 x_1^2 + 48.83937 x_1 x_2^2 - 106.72022 x_1 x_2 \\
&\quad + 44.00210 x_1 - x_2^3 + 4 x_2^2 + x_2 + 3
\end{aligned}
$$

Example

Consider the polynomial system:

$$
\begin{aligned}
f_1 &= x_1^2 + 3.99980x_1x_2 - 5.89970x_1 + 3.81765x_2^2 - 11.25296x_2 \\
&\quad + 8.33521 \\
f_2 &= x_1^3 + 12.68721x_1^2x_2 - 2.36353x_1^2 + 81.54846x_1x_2^2 - 177.31082x_1x_2 \\
&\quad + 73.43867x_1 - x_2^3 + 6x_2^2 + x_2 + 5 \\
f_3 &= x_1^3 + 8.04041x_1^2x_2 - 2.16167x_1^2 + 48.83937x_1x_2^2 - 106.72022x_1x_2 \\
&\quad + 44.00210x_1 - x_2^3 + 4x_2^2 + x_2 + 3
\end{aligned}
$$

Roots: $[0.8999, 1], [1, 1], [1, 0.8999]$ and $[-1, 2], [-1.0999, 2]$.
$\varepsilon = 0.1$.

Example

Basis: $[1, x_1, x_2, x_1x_2, x_1^2]$.

Example

Basis: $[1, x_1, x_2, x_1x_2, x_1^2]$.

The matrix of traces:

$$R = \begin{bmatrix} 5 & 0.79999 & 6.89990 & -1.40000 & 5.01960 \\ 0.79999 & 5.01960 & -1.40000 & 7.12928 & 0.39812 \\ 6.89990 & -1.40000 & 10.80982 & -5.68988 & 7.12928 \\ -1.40000 & 7.12928 & -5.68988 & 11.45876 & -2.03262 \\ 5.01960 & 0.39812 & 7.12928 & -2.03262 & 5.11937 \end{bmatrix}.$$

Example

Basis: $[1, x_1, x_2, x_1x_2, x_1^2]$.

The matrix of traces:

$$R = \begin{bmatrix} 5 & 0.79999 & 6.89990 & -1.40000 & 5.01960 \\ 0.79999 & 5.01960 & -1.40000 & 7.12928 & 0.39812 \\ 6.89990 & -1.40000 & 10.80982 & -5.68988 & 7.12928 \\ -1.40000 & 7.12928 & -5.68988 & 11.45876 & -2.03262 \\ 5.01960 & 0.39812 & 7.12928 & -2.03262 & 5.11937 \end{bmatrix}.$$

After 2 steps of GECP:

$$U_2 = \begin{bmatrix} 11.45876 & -5.68988 & 7.12928 & -1.40000 & -2.03262 \\ 0 & 7.98449 & 2.14006 & 6.20472 & 6.11998 \\ 0 & 0 & 0.01039 & 0.00799 & 0.02243 \\ 0 & 0 & 0.00799 & 0.00728 & 0.01544 \\ 0 & 0 & 0.02243 & 0.01544 & 0.06796 \end{bmatrix}.$$

Example

From the matrix of traces R we compute the matrix $\tilde{R}$, with columns indexed by 1 and x_1 and rows indexed by 1 and $x2$:

$$\tilde{R} := \begin{bmatrix} 5 & 0.79999 \\ 6.89990 & -1.40000 \end{bmatrix}.$$

Example

From the matrix of traces R we compute the matrix $\tilde{R}$, with columns indexed by 1 and x_1 and rows indexed by 1 and $x2$:

$$\tilde{R} := \begin{bmatrix} 5 & 0.79999 \\ 6.89990 & -1.40000 \end{bmatrix}.$$

We now solve the system:

$$\begin{aligned} \tilde{R}\tilde{M}_{x_i} &= \tilde{R}_{x_i}, \quad \text{with} \\ \tilde{R}_{x_1} &= \begin{bmatrix} 0.79999 & 5.01960002 \\ -1.40000 & 7.12928003 \end{bmatrix}, \\ \tilde{R}_{x_2} &= \begin{bmatrix} 6.8999 & -1.4000 \\ 10.80982 & -5.68988 \end{bmatrix} \end{aligned}$$

Example

We obtain the *approximate multiplication matrices*, in the basis $\{1, x_1\}$:

$$\tilde{M}_{x_1} = \begin{bmatrix} 0 & 1.01685 \\ 1 & -0.08080 \end{bmatrix}, \quad \text{with eigenvalues } 0.96880 \text{ and } -1.04960,$$

$$\tilde{M}_{x_2} = \begin{bmatrix} 1.46229 & -0.52012 \\ -0.51442 & 1.50078 \end{bmatrix}, \quad \text{with eigenvalues } 0.96391 \text{ and } 1.99915.$$

Example

We obtain the *approximate multiplication matrices*, in the basis $\{1, x_1\}$:

$$\tilde{M}_{x_1} = \begin{bmatrix} 0 & 1.01685 \\ 1 & -0.08080 \end{bmatrix}, \quad \text{with eigenvalues } 0.96880 \text{ and } -1.04960,$$

$$\tilde{M}_{x_2} = \begin{bmatrix} 1.46229 & -0.52012 \\ -0.51442 & 1.50078 \end{bmatrix}, \quad \text{with eigenvalues } 0.96391 \text{ and } 1.99915.$$

The roots of the approximate radical are then $[0.96880, 0.96391]$ and $[-1.0460, 1.99915]$.

Note: the arithmetic means of the roots of the clusters are $[0.96663, 0.96663]$ and $[-1.04995, 2]$.

Example

We obtain the *approximate multiplication matrices*, in the basis $\{1, x_1\}$:

$$\tilde{M}_{x_1} = \begin{bmatrix} 0 & 1.01685 \\ 1 & -0.08080 \end{bmatrix}, \quad \text{with eigenvalues } 0.96880 \text{ and } -1.04960,$$

$$\tilde{M}_{x_2} = \begin{bmatrix} 1.46229 & -0.52012 \\ -0.51442 & 1.50078 \end{bmatrix}, \quad \text{with eigenvalues } 0.96391 \text{ and } 1.99915.$$

The roots of the approximate radical are then $[0.96880, 0.96391]$ and

$[-1.0460, 1.99915]$.

Note: the arithmetic means of the roots of the clusters are $[0.96663, 0.96663]$ and $[-1.04995, 2]$.

The commutator of the multiplication matrices is

$$\begin{bmatrix} -0.00296 & -0.00289 \\ 0.00307 & 0.00296 \end{bmatrix}.$$

Computation of Matrices of Traces

Computation of Matrices of Traces

From the definition

Compute a basis $[b_1, \ldots, b_n]$ for $\mathbb{C}[\mathbf{x}]/I$ and the multiplication matrices $M_{b_i b_j}$ of I to compute the traces $Tr(M_{b_i b_j})$ for all $b_i, b_j \in B$.

Computation of Matrices of Traces

From the definition

Compute a basis $[b_1, \ldots, b_n]$ for $\mathbb{C}[\mathbf{x}]/I$ and the multiplication matrices $M_{b_i b_j}$ of I to compute the traces $Tr(M_{b_i b_j})$ for all $b_i, b_j \in B$.

Newton Sums

Let $f(x) = x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n = \prod_{i=1}^n (x - \xi_i)$. We have $R = [s_{i+j}]_{i,j=0}^{n-1}$ where $s_k := \sum_{t=1}^n \xi_t^k$.

Computation of Matrices of Traces

From the definition

Compute a basis $[b_1, \ldots, b_n]$ for $\mathbb{C}[\mathbf{x}]/I$ and the multiplication matrices $M_{b_i b_j}$ of I to compute the traces $Tr(M_{b_i b_j})$ for all $b_i, b_j \in B$.

Newton Sums

Let $f(x) = x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n = \prod_{i=1}^n (x - \xi_i)$. We have $R = [s_{i+j}]_{i,j=0}^{n-1}$ where $s_k := \sum_{t=1}^n \xi_t^k$. We find $s_1, \ldots, s_{2n-2}$ from:

$$\begin{aligned} s_1 + a_1 &= 0 \\ s_2 + a_1 s_1 + 2a_2 &= 0 \\ &\vdots \\ s_{2n-2} + a_1 s_{2n-3} + \cdots + a_n s_{n-3} &= 0. \end{aligned}$$

Note that this has generalizations to the multivariate case, but complicated.

Computing Matrices of Traces

Computation of multiplication matrices (and a basis of $\mathbb{C}[\mathbf{x}]/I$):

- resultant and subresultant matrices: Manocha and Demmel (1995), Chardin (1995), Szanto (2001),
- Gröbner bases: Corless (1996),
- Lazard's Algorithm: Lazard (1981), Corless, Gianni and Trager (1995),
- methods combining the above: Mourrain and Trébuchet (2005)
- moment matrices: Lasserre, Laurent and Rostalski (2007).

Computing Matrices of Traces

Computation of multiplication matrices (and a basis of $\mathbb{C}[\mathbf{x}]/I$):

- resultant and subresultant matrices: Manocha and Demmel (1995), Chardin (1995), Szanto (2001),
- Gröbner bases: Corless (1996),
- Lazard's Algorithm: Lazard (1981), Corless, Gianni and Trager (1995),
- methods combining the above: Mourrain and Trébuchet (2005)
- moment matrices: Lasserre, Laurent and Rostalski (2007).

It is however also possible to compute matrices of traces directly

- using Newton sums: Díaz-Toca and Gónzalez-Vega (2001), Briand and Gónzalez-Vega (2001)
- using residues: Becker, Cardinal, Roy, Szafraniec (1996), Cardinal and Mourrain (1996), Cattani, Dickenstein and Sturmfels (1996) and (1998)
- using resultants: D'Andrea and Jeronimo (2005)
- using reduced Bezoutians: Mourrain and Pan (2000), Mourrain (2005)

Sylvester Matrix

Let $\mathbf{f} = \{f_1, \ldots, f_s\} \subset \mathbb{C}[\mathbf{x}]$ generating an ideal I and $A = \mathbb{C}[\mathbf{x}]/I$.

Sylvester Matrix

Let $\mathbf{f} = \{f_1, \ldots, f_s\} \subset \mathbb{C}[\mathbf{x}]$ generating an ideal I and $A = \mathbb{C}[\mathbf{x}]/I$.

Definition

We define the Sylvester matrix $\mathrm{Syl}_\Delta(\mathbf{f})$ of degree Δ as the transpose of the matrix of the map

$$\begin{array}{rcl} \bigoplus_{i=1}^{s} \mathbb{C}[\mathbf{x}]_{\Delta - d_i} & \longrightarrow & \mathbb{C}[\mathbf{x}]_\Delta \\ (g_1, \ldots, g_s) & \mapsto & \sum_{i=1}^{s} f_i g_i \end{array}$$

Sylvester Matrix

Let $\mathbf{f} = \{f_1, \ldots, f_s\} \subset \mathbb{C}[\mathbf{x}]$ generating an ideal I and $A = \mathbb{C}[\mathbf{x}]/I$.

Definition

We define the Sylvester matrix $\mathrm{Syl}_\Delta(\mathbf{f})$ of degree Δ as the transpose of the matrix of the map

$$\begin{array}{ccc} \bigoplus_{i=1}^{s} \mathbb{C}[\mathbf{x}]_{\Delta - d_i} & \longrightarrow & \mathbb{C}[\mathbf{x}]_\Delta \\ (g_1, \ldots, g_s) & \mapsto & \sum_{i=1}^{s} f_i g_i \end{array}$$

Fact: If Δ is large enough, a basis $B = [b_1, \ldots, b_n]$ for A can be computed using $\mathrm{Syl}_\Delta(\mathbf{f})$. Bounds for Δ given if I has finite *projective roots* using Lazard (1981).

Moment Matrix

We fix a random element of the *Nullspace* of the Sylvester matrix

$$\mathbf{y} = [y_\alpha : \alpha \in \mathbb{N}^m,\ |\alpha| \leq \Delta]^T \in \mathrm{Null}(\mathrm{Syl}_\Delta(\mathbf{f})).$$

Moment Matrix

We fix a random element of the *Nullspace* of the Sylvester matrix

$$\mathbf{y} = [y_\alpha : \alpha \in \mathbb{N}^m,\ |\alpha| \leq \Delta]^T \in \mathrm{Null}(\mathrm{Syl}_\Delta(\mathbf{f})).$$

Definition

Let $B = [b_1, \ldots, b_n]$ be a basis for A. The $n \times n$ moment matrix $\mathfrak{M}_B(\mathbf{y})$ is defined by

$$\mathfrak{M}_B(\mathbf{y}) = [y_{b_i b_j}]_{i,j=1}^n.$$

Moment Matrix

We fix a random element of the *Nullspace* of the Sylvester matrix

$$\mathbf{y} = [y_\alpha : \alpha \in \mathbb{N}^m,\ |\alpha| \leq \Delta]^T \in \mathrm{Null}(\mathrm{Syl}_\Delta(\mathbf{f})).$$

Definition

Let $B = [b_1, \ldots, b_n]$ be a basis for A. The $n \times n$ moment matrix $\mathfrak{M}_B(\mathbf{y})$ is defined by

$$\mathfrak{M}_B(\mathbf{y}) = [y_{b_i b_j}]_{i,j=1}^n.$$

Note: We have that

$$\max_{\mathbf{y} \in \mathrm{Null}(\mathrm{Syl}_\Delta(\mathbf{f}))} \mathrm{rank}(\mathfrak{M}_\mathrm{B}(\mathbf{y})) = \begin{cases} n & \text{if } A \text{ is Gorenstein} \\ \leq n & \text{if } A \text{ is non-Gorenstein} \end{cases}$$

and the maximum is attained with high probability by taking a random element in $\mathrm{Null}(\mathrm{Syl}_\Delta(\mathbf{f}))$.

Generalized Jacobian

Definition

The dual basis for B is defined by $b_i^* := \sum_{j=1}^n c_{ji} b_j$ where $\mathfrak{M}_B^{-1}(\mathbf{y}) =: [c_{ij}]_{i,j=1}^n$.

Generalized Jacobian

Definition

The dual basis for B is defined by $b_i^* := \sum_{j=1}^n c_{ji} b_j$ where $\mathfrak{M}_B^{-1}(\mathbf{y}) =: [c_{ij}]_{i,j=1}^n$.

Definition

We define the generalized Jacobian by

$$J := \sum_{i=1}^n b_i b_i^* \mod I.$$

Generalized Jacobian

Definition

The dual basis for B is defined by $b_i^* := \sum_{j=1}^n c_{ji} b_j$ where $\mathfrak{M}_B^{-1}(\mathbf{y}) =: [c_{ij}]_{i,j=1}^n$.

Definition

We define the generalized Jacobian by

$$J := \sum_{i=1}^n b_i b_i^* \mod I.$$

$\mathrm{Syl}_B(J)$ is then constructed from the map

$$\sum_{i=1}^n c_i b_i \mapsto J \cdot \sum_{i=1}^n c_i b_i \in \mathbb{C}[x]_\Delta.$$

Main Theorem

Theorem

Let $B = [b_1, \ldots, b_n]$ *be a basis of* A *with* $\deg(b_i) \leq \Delta$*. With the generalized Jacobian* J *and* $\mathrm{Syl}_B(J)$ *defined before, we have*

$$[Tr(b_i b_j)]_{i,j=1}^n = \mathrm{Syl}_B(J) \cdot \mathfrak{M}'_B(\mathbf{y}),$$

where $\mathfrak{M}'_B(\mathbf{y})$ *is the unique extension of the square moment matrix* $\mathfrak{M}_B(\mathbf{y})$ *such that* $\mathrm{Syl}_\Delta(\mathbf{f}) \cdot \mathfrak{M}'_B(\mathbf{y}) = 0$.

Univariate example

Let $n = 3$ and $f = x^3 + a_1x^2 + a_2x + a_3$.

Univariate example

Let $n = 3$ and $f = x^3 + a_1x^2 + a_2x + a_3$. Then $\Delta = 4$, $B = [1, x, x^2]$ and

$$\mathrm{Syl}_4(f) := \begin{bmatrix} a_3 & a_2 & a_1 & 1 & 0 \\ 0 & a_3 & a_2 & a_1 & 1 \end{bmatrix},$$

Univariate example

Let $n = 3$ and $f = x^3 + a_1 x^2 + a_2 x + a_3$. Then $\Delta = 4$, $B = [1, x, x^2]$ and

$$\mathrm{Syl}_4(f) := \begin{bmatrix} a_3 & a_2 & a_1 & 1 & 0 \\ 0 & a_3 & a_2 & a_1 & 1 \end{bmatrix},$$

We take $\mathbf{y} := [0, 0, 1, -a_1, {a_1}^2 - a_2]^T \in \mathrm{Null}(\mathrm{Syl}_4(f))$.

Univariate example

Let $n = 3$ and $f = x^3 + a_1x^2 + a_2x + a_3$. Then $\Delta = 4$, $B = [1, x, x^2]$ and

$$\mathrm{Syl}_4(f) := \begin{bmatrix} a_3 & a_2 & a_1 & 1 & 0 \\ 0 & a_3 & a_2 & a_1 & 1 \end{bmatrix},$$

We take $\mathbf{y} := [0, 0, 1, -a_1, {a_1}^2 - a_2]^T \in \mathrm{Null}(\mathrm{Syl}_4(f))$.

The resulting moment matrices $\mathfrak{M}_B(\mathbf{y})$ and $\mathfrak{M}'_B(\mathbf{y})$ are:

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -a_1 \\ 1 & -a_1 & {a_1}^2 - a_2 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -a_1 \\ 1 & -a_1 & {a_1}^2 - a_2 \\ -a_1 & {a_1}^2 - a_2 & -{a_1}^3 + 2\,a_2a_1 - a_3 \\ {a_1}^2 - a_2 & -{a_1}^3 + 2\,a_2a_1 - a_3 & {a_1}^4 - 3\,a_2{a_1}^2 + 2\,a_3a_1 + {a_2}^2 \end{bmatrix}.$$

Univariate example cont.

The generalized Jacobian in this case is $J := f' = 3x^2 + 2a_1x + a_2$, and its Sylvester matrix is

$$\mathrm{Syl}_B(f') = \begin{bmatrix} a_2 & 2\,a_1 & 3 & 0 & 0 \\ 0 & a_2 & 2\,a_1 & 3 & 0 \\ 0 & 0 & a_2 & 2\,a_1 & 3 \end{bmatrix}.$$

Univariate example cont.

The generalized Jacobian in this case is $J := f' = 3x^2 + 2a_1x + a_2$, and its Sylvester matrix is

$$\mathrm{Syl}_B(f') = \begin{bmatrix} a_2 & 2\,a_1 & 3 & 0 & 0 \\ 0 & a_2 & 2\,a_1 & 3 & 0 \\ 0 & 0 & a_2 & 2\,a_1 & 3 \end{bmatrix}.$$

Finally, we get that $\mathrm{Syl}_B(f') \cdot \mathfrak{M}'_B(\mathbf{y})$ is the matrix of traces R:

$$\begin{bmatrix} 3 & -a_1 & -2\,a_2 + {a_1}^2 \\ -a_1 & -2\,a_2 + {a_1}^2 & -3\,a_3 + 3\,a_2a_1 - {a_1}^3 \\ -2\,a_2 + {a_1}^2 & -3\,a_3 + 3\,a_2a_1 - {a_1}^3 & -4\,a_2{a_1}^2 + 2\,{a_2}^2 + {a_1}^4 + 4\,a_3a_1 \end{bmatrix}.$$

THANK YOU!