Exploiting fast linear algebra in the computation of multivariate relations

Éric Schost .... U. Waterloo, Canada

Vincent Neiger .... **U. Limoges, France**

Structured Matrix Days 2019, XLIM, Limoges, France

May 23, 2019

Outline

- Multivariate relations and linear algebra
- Computing relations (known multiplication matrices)
- Computing the multiplication matrices

Outline

- Multivariate relations and linear algebra
- Computing relations (known multiplication matrices)
- Computing the multiplication matrices

Relations

$$\begin{bmatrix} \mathfrak{p}_1 & \cdots & \mathfrak{p}_m \end{bmatrix} \begin{bmatrix} \mathfrak{f}_1 \\ \vdots \\ \mathfrak{f}_m \end{bmatrix} = 0 \mod \mathcal{M}$$

Relations

polynomials $\in \mathbb{K}[\mathbf{X}] = \mathbb{K}[X_1, \ldots, X_r]$

ideal, module, ...

$$\begin{bmatrix} p_1 & \cdots & p_m \end{bmatrix} \begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix} = 0 \mod \mathcal{M}$$

a relation
(or syzygy)

polynomials modulo $\mathcal{M}$

Univariate Hermite-Padé approximation

Over $\mathbb{K} = \mathbb{Z}/7\mathbb{Z}$, $m = 4$, $\mathcal{M} = \langle X^4 \rangle$:

$$\begin{bmatrix} p_1 & p_2 & p_3 & p_4 \end{bmatrix} \begin{bmatrix} 5X^3 + 4X^2 + 6X + 4 \\ 2X^3 + X^2 + X + 3 \\ 2X + 1 \\ 4X^3 + X^2 + 4X \end{bmatrix} = 0 \bmod X^4$$

trivial relation $\rightsquigarrow$ $\mathbf{p} = \begin{bmatrix} X^4 & 0 & 0 & 0 \end{bmatrix}$

relation of small degree $\rightsquigarrow$ $\mathbf{p} = \begin{bmatrix} X+5 & 1 & 5 & 1 \end{bmatrix}$

basis of relations $\rightsquigarrow$ $\mathcal{B} = \left\{ \begin{matrix} [X+2 & 0 & 6 & 0], \\ [X^2 & X^2 & 0 & 0], \\ [X+2 & 3X+2 & X & 0], \\ [X+5 & 1 & 5 & 1] \end{matrix} \right\}$

Bivariate interpolation

$\mathcal{M}$ = set of polynomials $p(X, Y)$ vanishing at points in $\mathbb{K}^2$:
$\{(24, 80), (31, 73), (15, 73), (32, 35), (83, 66), (27, 46), (20, 91), (59, 64)\}$

All interpolants are relations:

$$p(X, Y) \in \mathcal{M} \quad \Leftrightarrow \quad p(X, Y)1 = 0 \bmod \mathcal{M}$$

$\rightsquigarrow$ “matrices” over $\mathbb{K}[X, Y]$

Bivariate interpolation

$\mathcal{M}$ = set of polynomials $p(X,Y)$ vanishing at points in $\mathbb{K}^2$:

$$\{(24,80),(31,73),(15,73),(32,35),(83,66),(27,46),(20,91),(59,64)\}$$

All interpolants are relations:

$$p(X,Y) \in \mathcal{M} \quad \Leftrightarrow \quad p(X,Y)1 = 0 \bmod \mathcal{M}$$

$\rightsquigarrow$ "matrices" over $\mathbb{K}[X,Y]$

$$\left.\begin{array}{l} G = (X-24)\cdots(X-59) \\ L = \text{Lagrange interpolant} \end{array}\right\} \longrightarrow \mathcal{M} = \langle G(X), Y - L(X) \rangle$$

Interpolants $p(X,Y) = p_0(X) + p_1(X)Y + p_2(X)Y^2$:

$$\begin{bmatrix} p_0 & p_1 & p_2 \end{bmatrix} \begin{bmatrix} 1 \\ L \\ L^2 \end{bmatrix} = 0 \bmod G$$

$\rightsquigarrow$ structured matrices over $\mathbb{K}[X]$

Bivariate interpolation

$\mathcal{M}$ = set of polynomials $p(X,Y)$ vanishing at points in $\mathbb{K}^2$:

$$\{(24,80),(31,73),(15,73),(32,35),(83,66),(27,46),(20,91),(59,64)\}$$
$$=\{(x_1,y_1),(x_2,y_2),(x_3,y_3),(x_4,y_4),(x_5,y_5),(x_6,y_6),(x_7,y_7),(x_8,y_8)\}$$

Interpolants $p_{00}+p_{01}X+p_{02}X^2+p_{03}X^3+p_{04}X^4+(p_{10}+p_{11}X+p_{12}X^2)Y+p_{20}Y^2$:

$$\begin{bmatrix} p_{00} & p_{01} & p_{02} & p_{03} & p_{04} & p_{10} & p_{11} & p_{12} & p_{20} \end{bmatrix} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_8 \\ x_1^2 & x_2^2 & \cdots & x_8^2 \\ x_1^3 & x_2^3 & \cdots & x_8^3 \\ x_1^4 & x_2^4 & \cdots & x_8^4 \\ \hline y_1 & y_2 & \cdots & y_8 \\ x_1y_1 & x_2y_2 & \cdots & x_8y_8 \\ x_1^2y_1 & x_2^2y_2 & \cdots & x_8^2y_8 \\ \hline y_1^2 & y_2^2 & \cdots & y_8^2 \end{bmatrix} = 0$$

$\rightsquigarrow$ 2-level structured matrices over $\mathbb{K}$

Finite-dimensional vector spaces

polynomials $\in \mathbb{K}[\mathbf{X}] = \mathbb{K}[X_1, \ldots, X_r]$

submodule of $\mathbb{K}[\mathbf{X}]^n$

$$\begin{bmatrix} p_1 & \cdots & p_m \end{bmatrix} \begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix} = 0 \mod \mathcal{M}$$

a relation

elements of $\mathbb{K}[\mathbf{X}]^n/\mathcal{M}$
finite dimension D as
a $\mathbb{K}$-vector space

$\rightsquigarrow$ these relations form a submodule of $\mathbb{K}[\mathbf{X}]^m$
which has co-dimension $\leqslant D$

Using linear algebra?

often, handling structured matrices = incorporating polynomial operations...

why

interpreting approximation/interpolation as linear algebra?

how

can this be done for relations in general?

Using linear algebra?

often, handling structured matrices = incorporating polynomial operations...

why
interpreting approximation/interpolation as linear algebra?

- fastest known approach for $m \geqslant D$
  (roughly: large matrix dimensions, small polynomial degrees)
- fastest known approach for any parameters for general relations

how
can this be done for relations in general?

using multiplication matrices
$\rightsquigarrow$ operations on polynomials translated into linear algebra

- elements f of $\mathbb{K}[\mathbf{X}]^n/\mathcal{M} \longleftrightarrow$ vectors $[v_1 \;\; \cdots \;\; v_D] \in \mathbb{K}^{1\times D}$
- multiplication by variable $X_i \longleftrightarrow$ multiplication by matrix $\mathbf{M}_i \in \mathbb{K}^{D\times D}$

Multiplication matrices

Working in $\mathbb{K}[X]/\langle X^4 \rangle$, with monomial basis $(1, X, X^2, X^3)$,
polynomial $p_0 + p_1 X + p_2 X^2 + p_3 X^3 \longleftrightarrow$ vector $[p_0 \;\; p_1 \;\; p_2 \;\; p_3]$

$$\text{Multiplication by } X = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Working in $\mathbb{K}[X, Y]/\langle G, Y - L \rangle$, with monomial basis $(1, X, X^2, \ldots, X^7)$

$$\mathbf{M} = \text{Multiplication by } X = \begin{bmatrix} & 1 & & & & & & \\ & & 1 & & & & & \\ & & & 1 & & & & \\ & & & & 1 & & & \\ & & & & & 1 & & \\ & & & & & & 1 & \\ & & & & & & & 1 \\ g_0 & g_1 & g_2 & g_3 & g_4 & g_5 & g_6 & g_7 \end{bmatrix}$$

$$\text{Multiplication by } Y = \begin{bmatrix} \text{coeff}(L) \\ \text{coeff}(XL \bmod G) \\ \text{coeff}(X^2L \bmod G) \\ \text{coeff}(X^3L \bmod G) \\ \text{coeff}(X^4L \bmod G) \\ \text{coeff}(X^5L \bmod G) \\ \text{coeff}(X^6L \bmod G) \\ \text{coeff}(X^7L \bmod G) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\ell} \\ \boldsymbol{\ell}\mathbf{M} \\ \boldsymbol{\ell}\mathbf{M}^2 \\ \boldsymbol{\ell}\mathbf{M}^3 \\ \boldsymbol{\ell}\mathbf{M}^4 \\ \boldsymbol{\ell}\mathbf{M}^5 \\ \boldsymbol{\ell}\mathbf{M}^6 \\ \boldsymbol{\ell}\mathbf{M}^7 \end{bmatrix}$$

Outline

- Multivariate relations and linear algebra
- Computing relations (known multiplication matrices)
- Computing the multiplication matrices

Problem

Input:

- submodule $\mathcal{M}$ of $\mathbb{K}[\mathbf{X}]^n$, of finite codimension D
- equation $\mathfrak{f} = [\mathfrak{f}_1 \ \cdots \ \mathfrak{f}_m]^T$ with entries in $\mathcal{M}/\mathbb{K}[\mathbf{X}]^n$
- a monomial order $\prec$ on $\mathbb{K}[\mathbf{X}]^m$

Represented as:

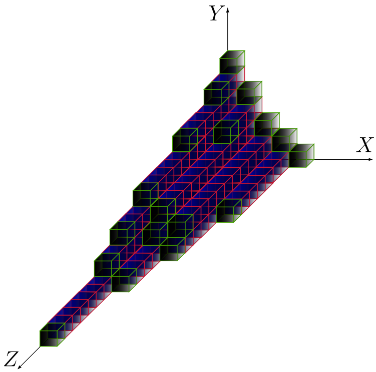
- multiplication matrices $\mathbf{M}_1, \ldots, \mathbf{M}_r$ in $\mathbb{K}^{D\times D}$
- vectors $\mathbf{e}_1, \ldots, \mathbf{e}_m$ in $\mathbb{K}^{1\times D}$

Output:

the $\prec$-Gröbner basis of the module of relations

$\mathcal{R} = \{\mathbf{p} \in \mathbb{K}[\mathbf{X}]^m \mid \mathbf{p}\mathfrak{f} = 0 \bmod \mathcal{M}\}$

$\rightsquigarrow$ **nice properties**: unique, minimal degrees, computing modulo $\mathcal{R}$, ...



Relations and multi-Krylov matrices

Notation: $\mathcal{V} = \mathbb{K}[X_1, \ldots, X_r]^n / \mathcal{M}$ is a $\mathbb{K}$-vector space of dimension D

Relations are vectors in the nullspace of a matrix over $\mathbb{K}$

- matrix $\mathbf{E} = \begin{bmatrix} \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_m \end{bmatrix} \in \mathbb{K}^{m \times D}$ (equation $\begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix} \in \mathcal{V}^{m \times 1}$)
- matrix $\mathbf{M}_i \in \mathbb{K}^{D \times D}$, $1 \leqslant i \leqslant r$ (multiplying by X_i in $\mathcal{V}$)

$$[p_1 \; \cdots \; p_m] \begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix} = \sum_{1 \leqslant i \leqslant m} \sum_{\mathbf{j}} \underbrace{\alpha_{i,\mathbf{j}}}_{\in \mathbb{K}} X_1^{j_1} \cdots X_r^{j_r} f_i$$

relation $=$ $\mathbb{K}$-linear relation between vectors $\{\mathbf{e}_i \mathbf{M}_1^{j_1} \cdots \mathbf{M}_r^{j_r}\}_{\mathbf{j},i}$ $\in \mathbb{K}^{1 \times D}$

Relations and multi-Krylov matrices

basis of relations = subset of nullspace of multi-Krylov matrix

$\prec_{\mathsf{lex}}^{\mathsf{top}}$ order:

$$\begin{bmatrix} \begin{bmatrix} \mathbf{E} \\ \mathbf{E}\mathbf{M_1} \\ \vdots \\ \mathbf{E}\mathbf{M_1}^{D} \end{bmatrix} \\ \begin{bmatrix} \mathbf{E} \\ \mathbf{E}\mathbf{M_1} \\ \vdots \\ \mathbf{E}\mathbf{M_1}^{D} \end{bmatrix} \mathbf{M}_2 \\ \vdots \\ \begin{bmatrix} \mathbf{E} \\ \mathbf{E}\mathbf{M_1} \\ \vdots \\ \mathbf{E}\mathbf{M_1}^{D} \end{bmatrix} \mathbf{M}_2^{D} \end{bmatrix}$$

Relations and multi-Krylov matrices

basis of relations = subset of nullspace of multi-Krylov matrix

$\prec_{\text{lex}}^{\text{top}}$ order:

$$\begin{bmatrix} \begin{bmatrix} \mathbf{E} \\ \mathbf{E}\mathbf{M_1} \\ \vdots \\ \mathbf{E}\mathbf{M_1}^{D} \end{bmatrix} \\ \begin{bmatrix} \mathbf{E} \\ \mathbf{E}\mathbf{M_1} \\ \vdots \\ \mathbf{E}\mathbf{M_1}^{D} \end{bmatrix} \mathbf{M}_2 \\ \vdots \\ \begin{bmatrix} \mathbf{E} \\ \mathbf{E}\mathbf{M_1} \\ \vdots \\ \mathbf{E}\mathbf{M_1}^{D} \end{bmatrix} \mathbf{M}_2^{D} \end{bmatrix}$$

ω: $D \times D$ matrix multiplication in $O(D^{\omega})$ operations

- [Keller-Gehrig, 1985]: charpoly$(\mathbf{M})$ in $O(D^{\omega} \log(D))$
  (one variable, $\mathbf{E} = \mathrm{Id}$, output = Hermite form)
- [FGLM, 1993]: general case in $O(rD^3)$
- [Beckermann&Labahn, 2000]: $O(mD^2)$ for structured $\mathbf{M}$
  (one variable, output = shifted Popov form)
- [Faugère et al., 2014]: for $\prec_{\text{lex}}$ and Shape position,
  $O(D^{\omega} \log(D) + rM(D) \log(D))$

General case with fast matrix multiplication?

Incorporating fast linear algebra

Size of dense representations:

input	multi-Krylov matrix	output
$rD^2 + mD$	mD^r	rD^2

Algorithm:

1. compute monomial basis = row rank profile
2. find $\prec$-Gröbner basis by nullspace computation

Incorporating fast linear algebra

Size of dense representations:

input	multi-Krylov matrix	output
$rD^2 + mD$	mD^r	rD^2

Algorithm:

1. compute monomial basis = row rank profile
2. find $\prec$-Gröbner basis by nullspace computation

Difficulty: incorporate fast multiplication in Step 1 for any $\prec$

- $X_1, \ldots, X_r$ $\rightsquigarrow$ gather operations involving M_i
- $X_i, X_i^2, X_i^4, \ldots$ $\rightsquigarrow$ gather operations involving $M_i^{2^j}$

} as if $\prec_{\text{lex}}^{\text{top}}$

- insert new rows according to the order $\prec$

Cost bound: $O(rD^\omega \log(D))$ operations in $\mathbb{K}$

Outline

- Multivariate relations and linear algebra
- Computing relations (known multiplication matrices)
- Computing the multiplication matrices

Change of monomial order

Arising in polynomial system solving:

Problem: $\prec_1$-GB of $\mathcal{M} \longrightarrow \prec_2$-GB of $\mathcal{M}$

= $\prec_2$-GB of relations: $\mathbf{p}1 = 0 \bmod \mathcal{M}$

Approach: [FGLM, 1993]

1. compute $\mathbf{M}_1, \ldots, \mathbf{M}_r$ from $\prec_1$-GB — [FGLM, 1993] $\rightarrow$ $O(rD^3)$
2. compute the $\prec_2$-GB of relations — $O(rD^\omega \log(D))$

Result (case of ideals):
step **1.** in $O(rD^\omega \log(D))$

assuming the $\prec_1$-initial ideal is Borel-fixed

$\rightsquigarrow$ extends [Faugère et al., 2014]

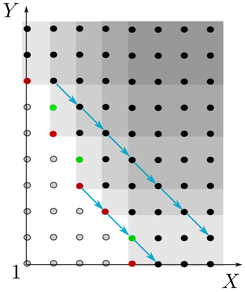
Borel-fixedness and multiplication matrices

Property of the ideal $\mathcal{J}$ of leading terms of $\mathcal{I}$:

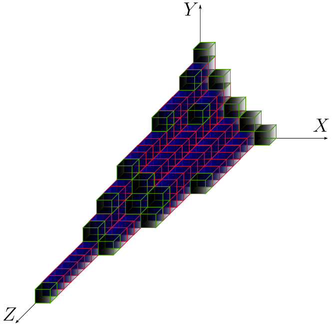
Borel-fixed monomial ideal $\mathcal{J}$ (in characteristic 0)

for all $\mu \in \mathcal{J}$, if X_j divides μ then $\frac{X_i}{X_j}\mu \in \mathcal{J}$ for all $i < j$.

Example in $\mathbb{K}[X, Y]$:



Example in $\mathbb{K}[X, Y, Z]$:



Main operation for obtaining the multiplication matrices:
computing parts of the multi-Krylov matrix, *à la* Keller-Gehrig

Borel-fixedness and multiplication matrices

Property of the ideal $\mathcal{J}$ of leading terms of $\mathcal{I}$:

Borel-fixed monomial ideal $\mathcal{J}$ (in characteristic 0)

for all $\mu \in \mathcal{J}$, if X_j divides μ then $\frac{X_i}{X_j}\mu \in \mathcal{J}$ for all $i < j$.

[Galligo 1974 & Bayer-Stillman 1987]:
existence and Borel-fixedness of the "GIN" of a *homogeneous* ideal $\mathcal{I}$
$\rightsquigarrow$ a random linear change of coordinates ensures Borel-fixedness w.h.p.

generalized to any ideal, for graded monomial orders

Perspectives (ranked by perceived difficulty):

- extension to the case of modules
- generalization to any monomial order (preliminary experiments with $\prec_{\text{lex}}$ revealed no counterexample)
- same cost $O(rD^\omega \log(D))$ without assumption on the ideal/module

Conclusion

Basis of relations

$$\mathbf{pf} = 0 \bmod \mathcal{M}$$

knowing multiplication matrices

Change of monomial order

$\rightsquigarrow$ polynomial system solving

$\prec_1$-GB of $\mathcal{M}$ $\longrightarrow$ $\prec_2$-GB of $\mathcal{M}$

- Computations with multi-Krylov matrices
- Incorporates fast dense linear algebra
- Cost bound: $O(rD^\omega \log(D))$
- For the second problem: assumptions on $\mathcal{M}$

Ongoing work (with Simone Naldi): incorporating polynomial multiplication in the computation of multivariate relations

$$\left[\begin{array}{ccc:ccc} a_1 & a_2 & a_3 & b_1 & b_2 & b_3 \\ a_2 & a_3 & a_4 & b_2 & b_3 & b_4 \\ a_3 & a_4 & a_5 & b_3 & b_4 & b_5 \\ \hdashline b_1 & b_2 & b_3 & d_1 & d_2 & d_3 \\ b_2 & b_3 & b_4 & d_2 & d_3 & d_4 \\ b_3 & b_4 & b_5 & d_3 & d_4 & d_5 \end{array}\right]$$