Fast, deterministic computation of determinants and Hermite forms for polynomial matrices

George Labahn

Symbolic Computation Group
Cheriton School of Computer Science
University of Waterloo, Canada

Joint work with V. Neiger and Wei Zhou

Ens-Lyon February 3, 2017

Outline

1. Preliminaries
2. Tools
3. Algorithm for Triangularization
4. Algorithm for Hermite Normal Form

Hermite Normal Form

Problem : Given nonsingular $\mathbf{A} \in \mathbb{K}[x]^{n\times n}$. Compute $\mathbf{U}$ and $\mathbf{H}$:

(i) $\mathbf{U}$ unimodular,

(ii) $\mathbf{H}$ in (column) Hermite form

(iii) $\mathbf{A} \cdot \mathbf{U} = \mathbf{H}$

Hermite Normal Form :

$$\mathbf{H} = \begin{bmatrix} h_{11} & 0 & \cdots & \cdots & 0 \\ h_{21} & h_{22} & 0 & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & 0 \\ h_{n1} & \cdots & & \cdots & h_{nn} \end{bmatrix} \qquad \text{deg } h_{ij} < \text{ deg } h_{ii}.$$

Results

Results:

- Fast, deterministic algorithms for $\mathbf{H}$
- Fast, deterministic algorithms for determinant of $(\mathbf{A})$
- Complexity : $O^{\sim}(n^{\omega}\lceil s\rceil)$ where s bounded by average
  : of row and column degrees of $\mathbf{A}$

Details in the paper :

- G. Labahn, V. Neiger and W. Zhou,
  Fast, deterministic computation of determinants and Hermite normal forms of polynomial matrices, Arxiv 2016.

References

Other relevant papers:

- W. Zhou, G. Labahn and A. Storjohann,
  Computing Minimal Nullspace Bases, *ISSAC 2012*,
- W. Zhou and G. Labahn,
  Computing Column Bases for polynomial matrices, *ISSAC 2013*
- S. Gupta, S. Sarkar, A. Storjohann, J. Valeriote, Triangular x-basis decompositions …, *ISSAC 2012*
- S. Gupta and A. Storjohann, Computing Hermite Forms of Polynomial Matrices, *ISSAC 2012*
- V. Neiger, Fast computation of shifted Popov forms, *ISSAC 2016*

Previous work : Hermite Form

- Polynomial-time over $\mathbb{Q}[x]$: Kannan 1985.
- $O\tilde{}(n^4 d)$: Hafner-McCurley 1991 deterministic
- $O\tilde{}(n^{\omega+1} d)$: Hafner-McCurley (1991), Villard (1996)
  Storjohann and Labahn (1996)
  deterministic
- $O\tilde{}(n^3 d^2)$: Mulders and Storjohann (2003) deterministic
- $O\tilde{}(n^3 d)$: Gupta and Storjohann (2012) probabilistic
- $O\tilde{}(n^\omega d)$: Gupta and Storjohann (2012) probabilistic
- $O\tilde{}(n^\omega s)$: Labahn-Neiger-Zhou (This talk) deterministic

Previous work : Determinants

- Storjohann (2000) $O^{\sim}(n^{\omega+1}d)$, deterministic
- Mulders and Storjohann (2003) $O(n^3d^2)$, deterministic
- Eberly-Giesbrecht-Villard (2000) $O^{\sim}(n^{2+\omega/2}d)$, probabilistic
- Storjohan (2003) $O^{\sim}(n^{\omega}d)$; probabilistic

- Giorgi-Jeannerod-Villard (2003) $O^{\sim}(n^{\omega}d)$,
- Kaltofen (1992)
- Kaltofen and Villard (2004)

Our Approach

- Triangularize $\mathbf{A}$
  - Gives diagonal entries of $\mathbf{H}$ which can be large
- Reduce remaining off-diagonal entries
  - First Try
  - Second Try

Our Approach

- Triangularize $\mathbf{A}$
  - Gives diagonal entries of $\mathbf{H}$ which can be large
- Reduce remaining off-diagonal entries
  - First Try
  - Second Try
- Need to avoid computing unimodular multiplier $\mathbf{U}$

  (since $\mathbf{U}$ can be too large)

Tools

- Shifted Degrees
- Kernel Bases
- Column Bases

Tool 1 : Shifted Degrees

- The column degree of a column vector $\mathbf{p}$ is

$$\operatorname{cdeg} \mathbf{p} \quad = \quad \max_{1 \le i \le n} \left[\deg p^{(i)} \right].$$

Tool 1 : Shifted Degrees

- The column degree of a column vector $\mathbf{p}$ is

$$\operatorname{cdeg} \mathbf{p} \quad = \quad \max_{1 \le i \le n} \left[\deg p^{(i)} \right].$$

- The $\vec{s}$-column degree of $\mathbf{p}$ is

$$\operatorname{cdeg}_{\vec{s}} \mathbf{p} \quad = \quad \max_{1 \le i \le n} \left[\deg p^{(i)} + s_i \right] \quad = \quad \operatorname{cdeg} x^{\vec{s}} \cdot \mathbf{p}.$$

Tool 1 : Shifted Degrees

- The column degree of a column vector $\mathbf{p}$ is

$$\operatorname{cdeg} \mathbf{p} \quad = \quad \max_{1 \le i \le n} \left[\deg p^{(i)} \right].$$

- The $\vec{s}$-column degree of $\mathbf{p}$ is

$$\operatorname{cdeg}_{\vec{s}} \mathbf{p} \quad = \quad \max_{1 \le i \le n} \left[\deg p^{(i)} + s_i \right] \quad = \quad \operatorname{cdeg} x^{\vec{s}} \cdot \mathbf{p}.$$

- e.g. $\operatorname{cdeg} \begin{bmatrix} x \\ x^2 \end{bmatrix} = 2, \quad \operatorname{cdeg}_{[3,1]} \begin{bmatrix} x \\ x^2 \end{bmatrix} = \operatorname{cdeg} \begin{bmatrix} x^4 \\ x^3 \end{bmatrix} = 4$

- For any matrix $\mathbf{A}$: $\operatorname{cdeg}_{-\vec{s}} \mathbf{A} \le 0$ same as $\operatorname{rdeg} \mathbf{A} \le \vec{s}$

Tool 2 : Minimal Kernel Bases

Given $\mathbf{F} \in \mathbb{K}[z]^{m \times n}$, $\quad m \leq n$:

A *Kernel Basis* for $\mathbf{F}$ is a $\mathbb{K}[z]$ module basis for

$$\{ \ \mathbf{p} \in \mathbb{K}[x]^n \ \mid \ \mathbf{F} \cdot \mathbf{p} = 0 \ \}$$

Can represent basis as matrix $\mathbf{M} \in \mathbb{K}[z]^{n \times *}$.

Tool 2 : Minimal Kernel Bases

Given $\mathbf{F} \in \mathbb{K}[z]^{m \times n}, \quad m \leq n$:

A *Kernel Basis* for $\mathbf{F}$ is a $\mathbb{K}[z]$ module basis for

$$\{ \ \mathbf{p} \in \mathbb{K}[x]^n \ \mid \ \mathbf{F} \cdot \mathbf{p} = 0 \ \}$$

Can represent basis as matrix $\mathbf{M} \in \mathbb{K}[z]^{n \times *}$.

Minimal Kernel Basis if matrix $\mathbf{M}$ is column reduced,

Shifted $\vec{s}$-Minimal Kernel Basis if $\quad z^{\vec{s}} \cdot \mathbf{M} \quad$ is column reduced.

Tool 3 : Column Bases

Given $\mathbf{F} \in \mathbb{K}[x]^{m \times n}$ with $m \le n$.

A *Column Basis* for $\mathbf{F}$ is a $\mathbb{K}[x]$ module basis for

$$\{ \ \mathbf{q} \in \mathbb{K}[x]^{m} \ \mid \ \exists \ \mathbf{p} \in \mathbb{K}[x]^{n} \quad \text{with} \quad \mathbf{q} = \mathbf{F} \cdot \mathbf{p} \ \}$$

Tool 3 : Column Bases

Given $\mathbf{F} \in \mathbb{K}[x]^{m \times n}$ with $m \leq n$.

A *Column Basis* for $\mathbf{F}$ is a $\mathbb{K}[x]$ module basis for

$$\{ \ \mathbf{q} \in \mathbb{K}[x]^{m} \ \mid \ \exists \ \mathbf{p} \in \mathbb{K}[x]^{n} \quad \text{with} \quad \mathbf{q} = \mathbf{F} \cdot \mathbf{p} \ \}$$

Again

(i) Represent column basis as full rank matrix $\mathbf{T} \in \mathbb{K}[x]^{m \times r}$.

(ii) Can find unimodular matrix $\mathbf{U}$ with $\mathbf{F} \cdot \mathbf{U} = [\mathbf{0}, \mathbf{T}]$.

Cost of Tools

$\mathbf{F} \in \mathbb{K}[x]^{n\times n}$, $\vec{s} \in \mathbb{Z}^n$ bounds column degrees, $\sum \vec{s} \leq \xi$

From (Zhou-Labahn-Storjohann, ISSAC 2012)

Theorem

$\vec{s}$-Minimal kernel basis computation costs $O^{\sim}(n^{\omega} s)$.

From (Zhou-Labahn, ISSAC 2013)

Theorem

Column basis computation costs $O^{\sim}(n^{\omega} s)$.

Triangularization

- Finding Diagonals
- Complexity
- Computing determinant

Finding Diagonal Elements

Partition $\mathbf{A}$ and $\mathbf{U}$ and reduce via

$$\mathbf{A} \cdot \mathbf{U} \quad = \quad \begin{bmatrix} \mathbf{A}_u \\ \mathbf{A}_d \end{bmatrix} \begin{bmatrix} \mathbf{U}_\ell & \mathbf{U}_r \end{bmatrix} \quad = \quad \begin{bmatrix} \mathbf{B}_1 & 0 \\ * & \mathbf{B}_2 \end{bmatrix}.$$

Finding Diagonal Elements

Partition $\mathbf{A}$ and $\mathbf{U}$ and reduce via

$$\mathbf{A} \cdot \mathbf{U} \;=\; \begin{bmatrix} \mathbf{A}_u \\ \mathbf{A}_d \end{bmatrix} \begin{bmatrix} \mathbf{U}_\ell & \mathbf{U}_r \end{bmatrix} \;=\; \begin{bmatrix} \mathbf{B}_1 & 0 \\ * & \mathbf{B}_2 \end{bmatrix}.$$

Here

(i) $\mathbf{B}_1$ is nonsingular and a column basis of $\mathbf{A}_u$.

(ii) $\mathbf{U}_r$ a right kernel basis of $\mathbf{A}_u$

(iii) $\mathbf{B}_2 = \mathbf{A}_d \cdot \mathbf{U}_r$,

Finding Diagonal Elements

Partition $\mathbf{A}$ and $\mathbf{U}$ and reduce via

$$\mathbf{A} \cdot \mathbf{U} \;=\; \begin{bmatrix} \mathbf{A}_u \\ \mathbf{A}_d \end{bmatrix} \begin{bmatrix} \mathbf{U}_\ell & \mathbf{U}_r \end{bmatrix} \;=\; \begin{bmatrix} \mathbf{B}_1 & 0 \\ * & \mathbf{B}_2 \end{bmatrix}.$$

Here

(i) $\mathbf{B}_1$ is nonsingular and a column basis of $\mathbf{A}_u$.

(ii) $\mathbf{U}_r$ a right kernel basis of $\mathbf{A}_u$

(iii) $\mathbf{B}_2 = \mathbf{A}_d \cdot \mathbf{U}_r$,

Recurse on $\mathbf{B}_1$ and $\mathbf{B}_2$ to get diagonal elements

Finding Diagonal Elements

Partition $\mathbf{A}$ and $\mathbf{U}$ and reduce via

$$\mathbf{A} \cdot \mathbf{U} = \begin{bmatrix} \mathbf{A}_u \\ \mathbf{A}_d \end{bmatrix} \begin{bmatrix} \mathbf{U}_\ell & \mathbf{U}_r \end{bmatrix} = \begin{bmatrix} \mathbf{B}_1 & 0 \\ * & \mathbf{B}_2 \end{bmatrix}.$$

Here

- (i) $\mathbf{B}_1$ is nonsingular and a column basis of $\mathbf{A}_u$.
- (ii) $\mathbf{U}_r$ a right kernel basis of $\mathbf{A}_u$
- (iii) $\mathbf{B}_2 = \mathbf{A}_d \cdot \mathbf{U}_r$,

Recurse on $\mathbf{B}_1$ and $\mathbf{B}_2$ to get diagonal elements

Important to control size (measured by column degrees).

Example

$$\mathbf{A} \quad = \quad \begin{bmatrix} x & -x^3 & -2\,x^4 & 2x & -x^2 \\ 1 & -1 & -2\,x & 2 & -x \\ -3 & 3\,x^2+x & 2\,x^2 & -x^4+1 & 3\,x \\ 0 & 1 & x^2+2\,x-2 & x^3+2x-2 & 0 \\ 1 & -x^2+2 & -2\,x^3-3\,x+3 & 2x+2 & 0 \end{bmatrix} \in \mathbb{Z}_7[x]^{5\times 5}.$$

Example

$$\mathbf{A} = \begin{bmatrix} x & -x^3 & -2x^4 & 2x & -x^2 \\ 1 & -1 & -2x & 2 & -x \\ -3 & 3x^2+x & 2x^2 & -x^4+1 & 3x \\ 0 & 1 & x^2+2x-2 & x^3+2x-2 & 0 \\ 1 & -x^2+2 & -2x^3-3x+3 & 2x+2 & 0 \end{bmatrix} \in \mathbb{Z}_7[x]^{5\times 5}.$$

$$\begin{bmatrix} \mathbf{A}_u \\ \mathbf{A}_d \end{bmatrix} \cdot [\mathbf{U}_\ell, \mathbf{U}_r] = \begin{bmatrix} x & -x^3 & -2x^4 & & \\ 1 & -1 & -2x & & \\ -3 & 3x^2+x & 2x^2 & & \\ * & * & * & x^3-1 & 0 \\ * & * & * & -x & x \end{bmatrix}$$

Example

$$\mathbf{A} = \begin{bmatrix} x & -x^3 & -2x^4 & 2x & -x^2 \\ 1 & -1 & -2x & 2 & -x \\ -3 & 3x^2+x & 2x^2 & -x^4+1 & 3x \\ 0 & 1 & x^2+2x-2 & x^3+2x-2 & 0 \\ 1 & -x^2+2 & -2x^3-3x+3 & 2x+2 & 0 \end{bmatrix} \in \mathbb{Z}_7[x]^{5\times 5}.$$

$$\begin{bmatrix} \mathbf{A}_u \\ \mathbf{A}_d \end{bmatrix} \cdot [\mathbf{U}_\ell^{(2)}, \mathbf{U}_r^{(2)}] = \begin{bmatrix} x & 0 & & & \\ 1 & x^2-1 & & & \\ * & * & x^3 & & \\ * & * & * & x^3-1 & 0 \\ * & * & * & -x & x \end{bmatrix}$$

Example

$$\mathbf{A} = \begin{bmatrix} x & -x^3 & -2x^4 & 2x & -x^2 \\ 1 & -1 & -2x & 2 & -x \\ -3 & 3x^2 + x & 2x^2 & -x^4 + 1 & 3x \\ 0 & 1 & x^2 + 2x - 2 & x^3 + 2x - 2 & 0 \\ 1 & -x^2 + 2 & -2x^3 - 3x + 3 & 2x + 2 & 0 \end{bmatrix} \in \mathbb{Z}_7[x]^{5\times 5}.$$

$$\begin{bmatrix} \mathbf{A}_u \\ \mathbf{A}_d \end{bmatrix} \cdot [\mathbf{U}_\ell^{(2)}, \mathbf{U}_r^{(2)}] = \begin{bmatrix} x & & & & \\ * & x^2 - 1 & & & \\ * & * & x^3 & & \\ * & * & * & x^3 - 1 & \\ * & * & * & * & x \end{bmatrix}$$

Costs?

- Compute (shifted) Kernel Basis : $\mathbf{U}_r$
- Compute Column Basis : $\mathbf{B}_1$
- Multiply two polynomial matrices : $\mathbf{B}_2 = \mathbf{A}_d \cdot \mathbf{U}_r$

Important Properties (ZLS ISSAC 2012)

$\mathbf{F} \in \mathbb{K}[x]^{m\times n}$, $\vec{s} \in \mathbb{Z}^n$ bounds column degrees, $\sum \vec{s} \le \xi$

Theorem

For $\mathbf{M}$ *a* $\vec{s}$*-minimal kernel basis of* $\mathbf{F}$*:* $\sum \operatorname{cdeg}_{\vec{s}} \mathbf{M} \le \sum \vec{s}$

Theorem

(i) $\mathbf{A} \in \mathbb{K}[x]^{m\times n}$, $m \le n$, $\vec{s} \in \mathbb{Z}^n$ *bounding column degrees of* $\mathbf{A}$

(ii) $\mathbf{B} \in \mathbb{K}[x]^{n\times k}$ *with* $k \in O(m)$, $\sum \operatorname{cdeg}_{\vec{s}} \mathbf{B} \le \sum \vec{s} \in O(\xi)$

Multiply $\mathbf{A}$ *and* $\mathbf{B}$ *:* $O^{\sim}(n^2 m^{\omega-2} s) \subset O^{\sim}(n^{\omega} s)$, $s = \xi/n$.

Complexity

Theorem

$\mathbf{A} \in \mathbb{K}[x]^{n\times n}$. *Diagonals costs* $O^{\sim}(n^{\omega}\lceil s\rceil)$ *where* $s = \frac{\sum \operatorname{cdeg} \mathbf{A}}{n}$.

Complexity

Theorem

$\mathbf{A} \in \mathbb{K}[x]^{n\times n}$. *Diagonals costs* $O^{\sim}(n^{\omega}\lceil s \rceil)$ *where* $s = \frac{\sum \operatorname{cdeg} \mathbf{A}}{n}$.

Proof.

If cost : $g(n)$ then recurrence relation: (with $s = \frac{\xi}{n}$)

$$g(n) \quad \in \quad O^{\sim}(n^{\omega}\lceil s \rceil) + g(\lceil n/2 \rceil) + g(\lfloor n/2 \rfloor)$$

Complexity

Theorem

$\mathbf{A} \in \mathbb{K}[x]^{n\times n}$. *Diagonals costs* $O^{\sim}(n^{\omega}\lceil s\rceil)$ *where* $s = \frac{\sum \operatorname{cdeg} \mathbf{A}}{n}$.

Proof.

If cost : $g(n)$ then recurrence relation: (with $s = \frac{\xi}{n}$)

$$\begin{aligned} g(n) &\in O^{\sim}(n^{\omega}\lceil s\rceil) + g(\lceil n/2\rceil) + g(\lfloor n/2\rfloor) \\ &\in O^{\sim}(n^{\omega-1}\xi + n^{\omega}) + g(\lceil n/2\rceil) + g(\lfloor n/2\rfloor) \end{aligned}$$

Complexity

Theorem

$\mathbf{A} \in \mathbb{K}[x]^{n\times n}$. *Diagonals costs* $O^{\sim}(n^{\omega}\lceil s \rceil)$ *where* $s = \frac{\sum \operatorname{cdeg} \mathbf{A}}{n}$.

Proof.

If cost : $g(n)$ then recurrence relation: (with $s = \frac{\xi}{n}$)

$$\begin{aligned} g(n) &\in O^{\sim}(n^{\omega}\lceil s\rceil) + g(\lceil n/2\rceil) + g(\lfloor n/2\rfloor) \\ &\in O^{\sim}(n^{\omega-1}\xi + n^{\omega}) + g(\lceil n/2\rceil) + g(\lfloor n/2\rfloor) \\ &\in O^{\sim}(n^{\omega-1}\xi + n^{\omega}) + 2g(\lceil n/2\rceil) \\ &\in O^{\sim}(n^{\omega-1}\xi + n^{\omega}) = O^{\sim}(n^{\omega}\lceil s\rceil). \end{aligned}$$

□

Determinants

Diagonals not enough - need to worry about unimodular part.

Determinants

Diagonals not enough - need to worry about unimodular part.

$$\det \mathbf{A} = \frac{\det \mathbf{B}_1 \cdot \det \mathbf{B}_2}{\det \mathbf{U}}$$

Determinants

Diagonals not enough - need to worry about unimodular part.

$$\det \mathbf{A} = \frac{\det \mathbf{B}_1 \cdot \det \mathbf{B}_2}{\det \mathbf{U}}$$

For $\det \mathbf{U} = \det [\, \mathbf{U}_\ell \ \mathbf{U}_r]$ we do:

1. $\det \mathbf{U} = \det \mathbf{U} \bmod z = \det U = \det [U_\ell, U_r]$
2. $\mathbf{V} = \mathbf{U}^{-1} = \begin{bmatrix} \mathbf{V}_u \\ \mathbf{V}_d \end{bmatrix}$
3. $\mathbf{U}_r$ and $\mathbf{V}_u$ determined in column bases computation
4. Find $U_\ell{}^*$ such that $U^* = [U_\ell{}^*, U_r]$ is unimodular
5. Let $V_u = \mathbf{V}_u \bmod z$. Then $\det \mathbf{U} = \frac{\det U^*}{\det V_u U_\ell{}^*}$

Hermite Normal Form

- First Try
- Second Try
- Complexity

Finding Rest of $\mathbf{H}$ (First Try)

Use method of Gupta and Storjohann (2012) to get rest of $\mathbf{H}$.

(i) Convert HNF to shifted $\vec{s}$-minimal kernel basis problem

$$\mathbf{A}\mathbf{U} = \mathbf{H} \quad \text{same as} \quad [\mathbf{A} \quad -\mathbf{I}] \begin{bmatrix} \mathbf{U} \\ \mathbf{H} \end{bmatrix} = \mathbf{0}.$$

(ii) Adjust to alternative $\vec{s}'$-minimal kernel basis problem

$$[\mathbf{A} \quad -\mathbf{E}] \begin{bmatrix} \mathbf{U} \\ \mathbf{H}' \end{bmatrix} = \mathbf{0}.$$

Ease to construct $\mathbf{E}$. Easy to get $\mathbf{H}$ from $\mathbf{H}'$

(iii) Find $\mathbf{Q}$ and $\mathbf{R}$ such that $\mathbf{E} \; = \; \mathbf{A}\mathbf{Q} + \mathbf{R}$. Solve via HOL.

Then repeat (ii) but with $\mathbf{E}$ replaced by $\mathbf{R}$.

(iv) Complexity is $O^{\sim}(n^{\omega} d)$

Finding Rest of $\mathbf{H}$: Second Try

Know : $\vec{\delta}$ diagonal degrees of $\mathbf{H}$. Set $\mu = \max(\vec{\delta})$

$$\mathbf{x}^{\vec{\mu}-\vec{\delta}}\mathbf{A} \xrightarrow{\text{reduce}} \mathbf{x}^{\vec{\mu}-\vec{\delta}}\mathbf{R} \xrightarrow{\text{normalize}} \mathbf{H} = \mathbf{R} \cdot \mathrm{lc}_{-\vec{\delta}}(\mathbf{R})^{-1}$$

where $\mathbf{R}$ is any $-\vec{\delta}$-column reduced form of $\mathbf{A}$.

Finding Rest of $\mathbf{H}$: Second Try

Know : $\vec{\delta}$ diagonal degrees of $\mathbf{H}$. Set $\mu = \max(\vec{\delta})$

$$\mathbf{x}^{\vec{\mu}-\vec{\delta}}\mathbf{A} \xrightarrow{\text{reduce}} \mathbf{x}^{\vec{\mu}-\vec{\delta}}\mathbf{R} \xrightarrow{\text{normalize}} \mathbf{H} = \mathbf{R} \cdot \mathrm{lc}_{-\vec{\delta}}(\mathbf{R})^{-1}$$

where $\mathbf{R}$ is any $-\vec{\delta}$-column reduced form of $\mathbf{A}$.

Problem : Shift $\vec{\mu} - \vec{\delta}$ might be too large

Finding Rest of $\mathbf{H}$: Second Try

Know : $\vec{\delta}$ diagonal degrees of $\mathbf{H}$. Set $\mu = \max(\vec{\delta})$

$$\mathbf{x}^{\vec{\mu}-\vec{\delta}}\mathbf{A} \xrightarrow{\text{reduce}} \mathbf{x}^{\vec{\mu}-\vec{\delta}}\mathbf{R} \xrightarrow{\text{normalize}} \mathbf{H} = \mathbf{R} \cdot \mathrm{lc}_{-\vec{\delta}}(\mathbf{R})^{-1}$$

where $\mathbf{R}$ is any $-\vec{\delta}$-column reduced form of $\mathbf{A}$.

Problem : Shift $\vec{\mu} - \vec{\delta}$ might be too large

Answer : Partial linearization of Storjohann (2007): $\mathbf{A} \rightarrow \mathcal{L}(\mathbf{A})$

Smooths shifts, keeps properties of $\mathbf{A}$ while enlarging a bit.

Partial Linearization

Consider $\mathbf{H}$ with diagonal degrees $(2, 37, 7, 18)$.

$$\mathbf{H} = \begin{bmatrix} (2) & & & \\ [36] & (37) & & \\ [6] & [6] & (7) & \\ [17] & [17] & [17] & (18) \end{bmatrix},$$

$[d]$: degree at most d and (d) : monic , degree exactly d.

$\delta = 1 + \lfloor (2 + 37 + 7 + 18)/4 \rfloor = 17$. Construct by "expanding rows":

$$\widetilde{\mathbf{H}} = \begin{bmatrix} (2) & & & \\ [16] & [16] & & \\ [16] & [16] & & \\ [2] & (3) & & \\ [6] & [6] & (7) & \\ [16] & [16] & [16] & [16] \\ [0] & [0] & [0] & (1) \end{bmatrix}.$$

$\mathbf{H}$ and $\widetilde{\mathbf{H}}$ are related by $\quad \mathbf{H} = \mathcal{E}_{\vec{\delta}} \cdot \widetilde{\mathbf{H}}$ where

$$\mathcal{E}_{\vec{\delta}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & x^{17} & x^{34} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & x^{17} \end{bmatrix}.$$

Insert elementary columns in $\widetilde{\mathbf{H}}$ by

$$\mathcal{L}_{\vec{\delta}}(\mathbf{H}) = \begin{bmatrix} (2) & & & & & & \\ [16] & x^{17} & & [16] & & & \\ [16] & -1 & x^{17} & [16] & & & \\ [2] & & -1 & (3) & & & \\ [6] & & & [6] & (7) & & \\ [16] & & & [16] & [16] & x^{17} & [16] \\ [0] & & & [0] & [0] & -1 & (1) \end{bmatrix}$$

Row degrees $\vec{d} = (2, 17, 17, 3, 7, 17, 1)$ - maximum 17.

Main property kept : shifted column reduction.

$$\begin{tikzcd}
\mathbf{x}^{\vec{\mu}-\vec{\delta}}\,\mathbf{A} \arrow[r, "\text{reduce}"] \arrow[d, "\text{partial linearization}"] & \mathbf{x}^{\vec{\mu}-\vec{\delta}}\,\mathbf{R} \arrow[r, "\text{normalize}"] & \mathbf{H} = \mathbf{R}\,\mathrm{lc}_{-\vec{\delta}}(\mathbf{R})^{-1} \arrow[d, "\text{partial linearization}"] \\
\mathbf{x}^{\vec{m}-\vec{d}}\,\mathcal{L}_{\vec{\delta}}(\mathbf{A}) \arrow[r, "\text{reduce}"] & \mathbf{x}^{\vec{m}-\vec{d}}\,\hat{\mathbf{R}} \arrow[r, "\text{normalize}"] & \mathcal{L}_{\vec{\delta}}(\mathbf{H}) = \hat{\mathbf{R}}\,\mathrm{lc}_{-\vec{d}}(\hat{\mathbf{R}})^{-1}
\end{tikzcd} .$$

Theorem

Let $\mathbf{A} \in \mathbb{K}[x]^{n\times n}$ *nonsingular with* $\vec{\delta}$ *the degrees of the diagonal entries of the Hermite form.*

Then the Hermite form is computed using $O^{\sim}(n^{\omega} d)$ *field operations.*

Improving the Complexity

Repeat : partial linearization (this time with columns) :

(i) Enlarge : $\mathbf{A} \to \mathcal{L}^c(\mathbf{A})$
  - size of $\mathcal{L}^c(\mathbf{A})$ at most twice size of $\mathbf{A}$
  - degree $\mathcal{L}^c(\mathbf{A})$ at most average of $\mathbf{A}$

(ii) Compute Hermite form of $\mathcal{L}^c(\mathbf{A})$

(iii) $\mathbf{H}$ is found in lower right corner of Hermite form of $\mathcal{L}^c(\mathbf{A})$

Theorem

$\mathbf{A} \in \mathbb{K}[x]^{n\times n}$ *nonsingular. Hermite form computed:* $O^{\sim}(n^{\omega}\lceil s \rceil)$.

Future Work

We want to make progress with:

- Fast but with coefficient control (e.g. matrices over $Z[x]$)
  - Beckermann, Labahn, Villard (2006)

    $O\tilde{}((m+n)(m^2 d \min(m,n))^3 \log \|\mathbf{A}\|)$ bit operations

    $O\tilde{}(n^{10} d^3 \log \|\mathbf{A}\|)$ when $m = n$
- Fast Popov form. Fast shifted Popov form
- Fast Hermite and Popov for alternate domains
  (e.g. matrices over $\mathbb{K}(x)[D_x]$)