Fast Hermite Form Computation for polynomial matrices

(also fast determinant)

George Labahn and Arne Storjohann

Cheriton School of Computer Science
University of Waterloo, Canada

Recent Trends in Computer Algebra, Institut Henri Poincaré,

September 2023

Outline

Recall

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 :

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

Goal:

We describe:

- Fast, deterministic algorithms for $\mathbf{H}$
- Fast, deterministic algorithms for determinant of $(\mathbf{A})$
- Complexity : $O^{\sim}\left(n^{\omega}\lceil s\rceil\right)$ where s bounded by average
  : of row and column degrees of $\mathbf{A}$
  : here $O^{\sim}$ is big O without log factors.

Goal:

We describe:

- Fast, deterministic algorithms for $\mathbf{H}$
- Fast, deterministic algorithms for determinant of $(\mathbf{A})$
- Complexity : $O^{\sim}\left(n^{\omega}\lceil s\rceil\right)$ where s bounded by average
  : of row and column degrees of $\mathbf{A}$
  : here $O^{\sim}$ is big O without log factors.

Details :

- G. Labahn, V. Neiger and W. Zhou,
  Fast, deterministic computation of determinants and Hermite normal forms of polynomial matrices, J. of Complexity 2018.

Recall: Previous work

- Polynomial-time over $\mathbb{Q}[x]$: Kannan 1985.
- $O^{\sim}(n^4 d)$: Hafner-McCurley 1991 deterministic
- $O^{\sim}(n^{\omega+1} d)$: Hafner-McCurley (1991), Villard (1996)
  Storjohann and Labahn (1996)
  deterministic
- $O^{\sim}(n^3 d^2)$: Mulders and Storjohann (2003) deterministic
- $O^{\sim}(n^3 d)$: Gupta and Storjohann (2012) probabilistic
- $O^{\sim}(n^{\omega} d)$: Gupta and Storjohann (2012) probabilistic
- $O^{\sim}(n^{\omega} s)$: Labahn-Neiger-Zhou (2018) deterministic

Approach used today for polynomial matrices

- Triangularize $\mathbf{A}$
  - Gives diagonal entries of $\mathbf{H}$ which can be large
  - Best a-priori bound $\deg h_{ii} \leq id$ with $\sum_i id \in O(n^2 d)$.

    Too large since we know $\sum_i \deg h_{ii} = \deg \det \mathbf{A} \leq nd$

Approach used today for polynomial matrices

- Triangularize $\mathbf{A}$
  - Gives diagonal entries of $\mathbf{H}$ which can be large
  - Best a-priori bound $\deg h_{ii} \leq id$ with $\sum_i id \in O(n^2 d)$.

    Too large since we know $\sum_i \deg h_{ii} = \deg \det \mathbf{A} \leq nd$
- Reduce remaining off-diagonal entries

Approach used today for polynomial matrices

- Triangularize $\mathbf{A}$
  - Gives diagonal entries of $\mathbf{H}$ which can be large
  - Best a-priori bound $\deg h_{ii} \leq id$ with $\sum_i id \in O(n^2 d)$.

    Too large since we know $\sum_i \deg h_{ii} = \deg \det \mathbf{A} \leq nd$
- Reduce remaining off-diagonal entries
  - Need to avoid actually computing unimodular multiplier $\mathbf{U}$

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

Matrix Polynomial Algebra

Leading coefficients

Shifted Degrees

Reduced forms

Leading coefficients and degrees and shifts

$$A = 9z^8 + z^7 - 3z^3 + 2z = 9z^8 + \cdots$$

$$\mathrm{lcoeff}_z(A) = 9, \qquad \mathrm{degree}_z(A) = 8$$

Leading coefficients and degrees and shifts

$$A = \begin{bmatrix} 1 & z^2+3 & 7z^8+z \\ 1 & 2z^7+1 & 3 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 7 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} z^8 + \cdots$$

$$\mathsf{lcoeff}_z(A) = \begin{bmatrix} 0 & 0 & 7 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \mathsf{deg}_z(A) = 8$$

Leading coefficients and degrees and shifts

$$A = \begin{bmatrix} 1 & z^2+3 & 7z^8+z \\ 1 & 2z^7+1 & 3 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 7 \\ 1 & 2 & 0 \\ 1 & 0 & 0 \end{bmatrix} \mathit{diag}(1, z^7, z^8) + \cdots$$

$$\mathsf{c\ lcoeff}(A) = \begin{bmatrix} 1 & 0 & 7 \\ 1 & 2 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \qquad \mathsf{cdeg}(A) = (0, 7, 8)$$

Leading coefficients and degrees and shifts

$$\mathit{diag}(1, z, z^8)A = \begin{bmatrix} 1 & z^2+3 & 7z^8+z \\ z & 2z^8+z & 3z \\ z^8 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 7 \\ 0 & 2 & 0 \\ 1 & 0 & 0 \end{bmatrix} z^8 + \cdots$$

$$\mathsf{c\ lcoeff}_v(A) = \begin{bmatrix} 0 & 0 & 7 \\ 0 & 2 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \qquad \mathsf{cdeg}_v(A) = (8, 8, 8)$$

Leading coefficients and degrees and shifts

$$\mathit{diag}\,(1, z, z^8)A = \begin{bmatrix} 1 & z^2+3 & 7z^8+z \\ z & 2z^8+z & 3z \\ z^8 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 7 \\ 0 & 2 & 0 \\ 1 & 0 & 0 \end{bmatrix} z^8 + \cdots$$

$$\text{c lcoeff}_v(A) = \begin{bmatrix} 0 & 0 & 7 \\ 0 & 2 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \qquad \text{cdeg}_v(A) = (8, 8, 8)$$

Note in last two cases the leading coefficient was nonsingular.

A is column reduced and shifted column reduced (with shift $(0, 1, 8)$), respectively.

Shifted Degrees

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

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

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

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

Examples of Shifted degree

$$A = \begin{bmatrix} 1 & z^2+3 & 7z^8+z \\ 1 & 2z^7+1 & 3 \\ 1 & 0 & 0 \end{bmatrix}$$

- $\operatorname{cdeg}(A) = (0, 7, 8)$
- $\operatorname{cdeg}_{(0,1,8)}(A) = (8, 8, 8)$
- $\operatorname{cdeg}_{(1,0,0)}(A) = (1, 7, 9)$

Examples of Shifted degree

$$A = \begin{bmatrix} 1 & z^2+3 & 7z^8+z \\ 1 & 2z^7+1 & 3 \\ 1 & 0 & 0 \end{bmatrix}$$

- $\operatorname{cdeg}(A) = (0, 7, 8)$
- $\operatorname{cdeg}_{(0,1,8)}(A) = (8, 8, 8)$
- $\operatorname{cdeg}_{(1,0,0)}(A) = (1, 7, 9)$
- For any matrix $\mathbf{A}$: $\operatorname{cdeg}_{-\vec{s}} \mathbf{A} \leq 0$ same as $\operatorname{rdeg} \mathbf{A} \leq \vec{s}$

Why are shifts useful?

Example:

Given $\mathbf{A}$ and an algorithm for column reduced kernel.

Find unimodular $\mathbf{U}$ such that $\mathbf{AU}$ is column reduced.

$$\mathbf{A} \cdot \mathbf{U} = \mathbf{B} \quad \text{same as} \quad [\mathbf{A}, \quad -\mathbf{I}_n] \cdot \begin{bmatrix} \mathbf{U} \\ \mathbf{B} \end{bmatrix} = 0$$

Why are shifts useful?

Example:

Given $\mathbf{A}$ and an algorithm for column reduced kernel.

Find unimodular $\mathbf{U}$ such that $\mathbf{AU}$ is column reduced.

$$\mathbf{A} \cdot \mathbf{U} = \mathbf{B} \quad \text{same as} \quad [\mathbf{A}, \quad -\mathbf{I}_n] \cdot \begin{bmatrix} \mathbf{U} \\ \mathbf{B} \end{bmatrix} = 0$$

Can we use column reduced kernel algorithm?

Why are shifts useful?

Example:

Given $\mathbf{A}$ and an algorithm for column reduced kernel.

Find unimodular $\mathbf{U}$ such that $\mathbf{AU}$ is column reduced.

$$z^k\mathbf{A} \cdot \mathbf{U} = z^k\mathbf{B} \quad \text{same as} \quad [z^k\mathbf{A}, \quad -\mathbf{I}_n] \cdot \begin{bmatrix} \mathbf{U} \\ z^k\mathbf{B} \end{bmatrix} = 0$$

Can we use column reduced kernel algorithm? Not really

Why are shifts useful?

Example:

Given $\mathbf{A}$ and an algorithm for column reduced kernel.

Find unimodular $\mathbf{U}$ such that $\mathbf{AU}$ is column reduced.

$$z^k\mathbf{A} \cdot \mathbf{U} = z^k\mathbf{B} \quad \text{same as} \quad [z^k\mathbf{A}, \quad -\mathbf{I}_n] \cdot \begin{bmatrix} \mathbf{U} \\ z^k\mathbf{B} \end{bmatrix} = 0$$

But we can use column reduced kernel algorithm if k big enough!

Shifts allows us to focus on specific parts of a polynomial matrix.

Matrix Polynomial Domains

Kernel Bases

Order Bases

Column Bases

Minimal Kernel Bases

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

A (right) *Kernel Basis* for $\mathbf{A}$ is a $\mathbb{K}[z]$-module basis for

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

Minimal Kernel Bases

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

A (right) *Kernel Basis* for $\mathbf{A}$ is a $\mathbb{K}[z]$-module basis for

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

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

Minimal Kernel Bases

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

A (right) *Kernel Basis* for $\mathbf{A}$ is a $\mathbb{K}[z]$-module basis for

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

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

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

Minimal Kernel Bases

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

A (right) *Kernel Basis* for $\mathbf{A}$ is a $\mathbb{K}[z]$-module basis for

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

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

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

- i.e. leading coeff matrix has full column rank

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

Order Bases

An *Order Basis* for $\mathbf{A}$ and order $\vec{\sigma}$ is a $\mathbb{K}[z]$-module basis for

$$\{\ \mathbf{p} \in \mathbb{K}[z]^n \ \mid\ \mathbf{A} \cdot \mathbf{p} = O(z^{\vec{\sigma}})\ \}$$

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

Column Bases

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

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

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

Column Bases

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

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

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

Again

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

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

Compute Kernel Basis via Order Basis

- We have a fast order basis algorithm

Compute Kernel Basis via Order Basis

- We have a fast order basis algorithm

  An order basis of $\mathbf{A}$ with shift $\vec{s}$ and high enough order σ contains an $\vec{s}$-minimal kernel basis.

Compute Kernel Basis via Order Basis

- We have a fast order basis algorithm

  An order basis of $\mathbf{A}$ with shift $\vec{s}$ and high enough order σ contains an $\vec{s}$-minimal kernel basis.

- But high σ is required, hence inefficient

Compute Kernel Basis via Order Basis

- We have a fast order basis algorithm

  An order basis of $\mathbf{A}$ with shift $\vec{s}$ and high enough order σ contains an $\vec{s}$-minimal kernel basis.

- But high σ is required, hence inefficient
  - $\Theta(md)$ for uniform shift, cost $O^{\sim}\left(n^{\omega}\left\lceil m^2 d/n\right\rceil\right)$.

Compute Kernel Basis via Order Basis

- We have a fast order basis algorithm

  An order basis of $\mathbf{A}$ with shift $\vec{s}$ and high enough order σ contains an $\vec{s}$-minimal kernel basis.

- But high σ is required, hence inefficient
  - $\Theta(md)$ for uniform shift, cost $O^{\sim}\left(n^{\omega}\left\lceil m^2 d/n\right\rceil\right)$.
- Instead we compute a partial kernel basis

Compute Kernel Basis via Order Basis

- We have a fast order basis algorithm

  An order basis of $\mathbf{A}$ with shift $\vec{s}$ and high enough order σ contains an $\vec{s}$-minimal kernel basis.

- But high σ is required, hence inefficient
  - $\Theta(md)$ for uniform shift, cost $O^{\sim}\left(n^{\omega}\left\lceil m^2 d/n\right\rceil\right)$.
- Instead we compute a partial kernel basis
  - let σ be 3 times the average of the highest m entries of $\vec{s}$.

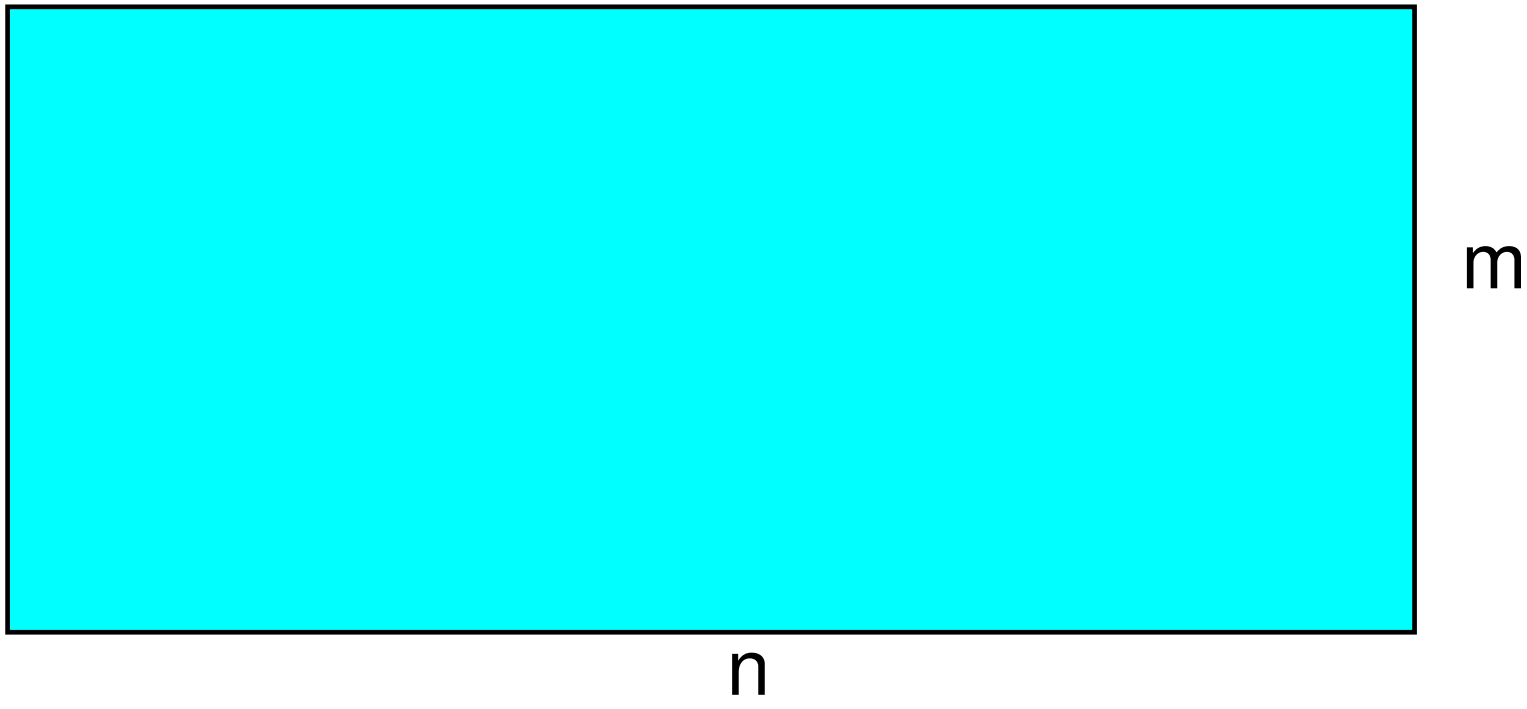
Compute Kernel Basis via Order Basis

- We have a fast order basis algorithm

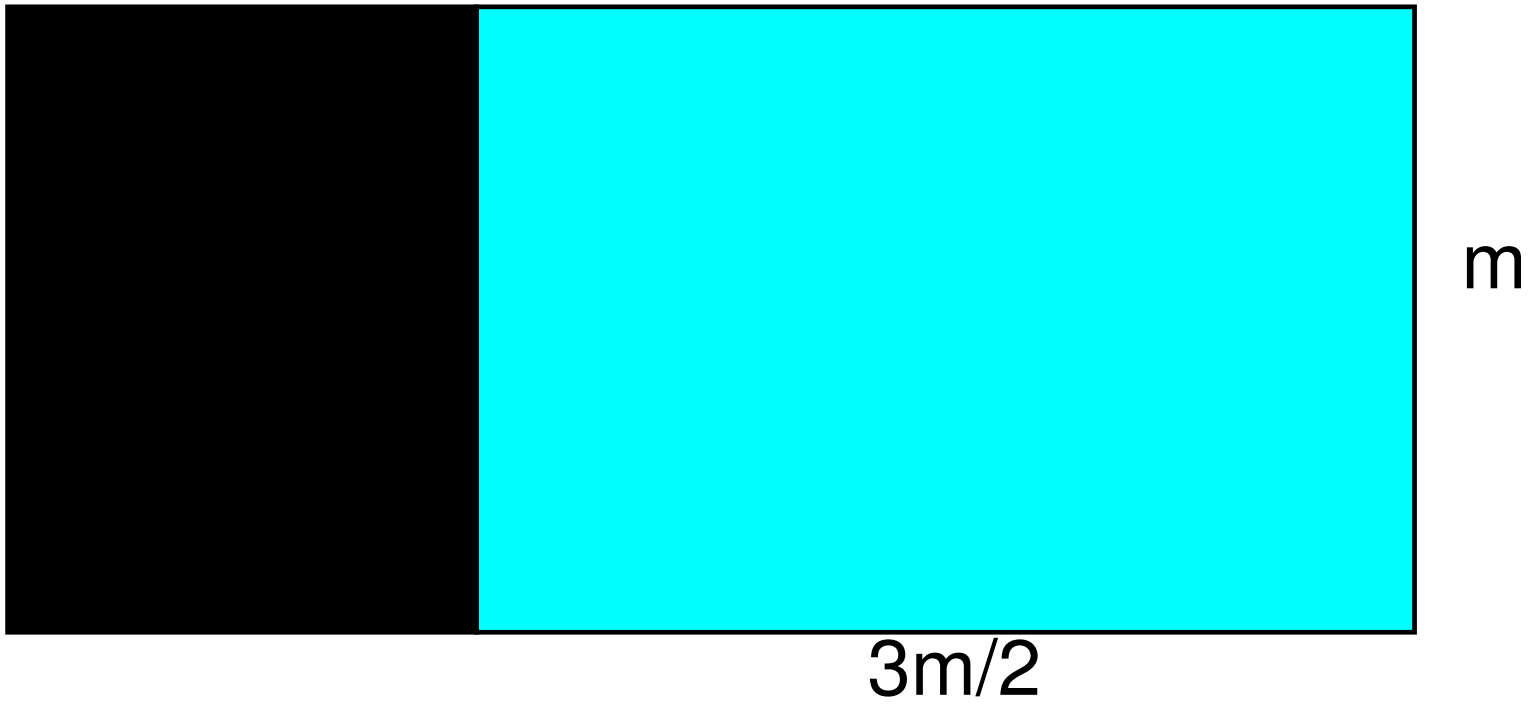
  An order basis of $\mathbf{A}$ with shift $\vec{s}$ and high enough order σ contains an $\vec{s}$-minimal kernel basis.

- But high σ is required, hence inefficient
  - $\Theta(md)$ for uniform shift, cost $O^{\sim}\left(n^{\omega}\left\lceil m^2 d/n\right\rceil\right)$.
- Instead we compute a partial kernel basis
  - let σ be 3 times the average of the highest m entries of $\vec{s}$.
  - implies only $3m/2$ kernel basis columns remaining

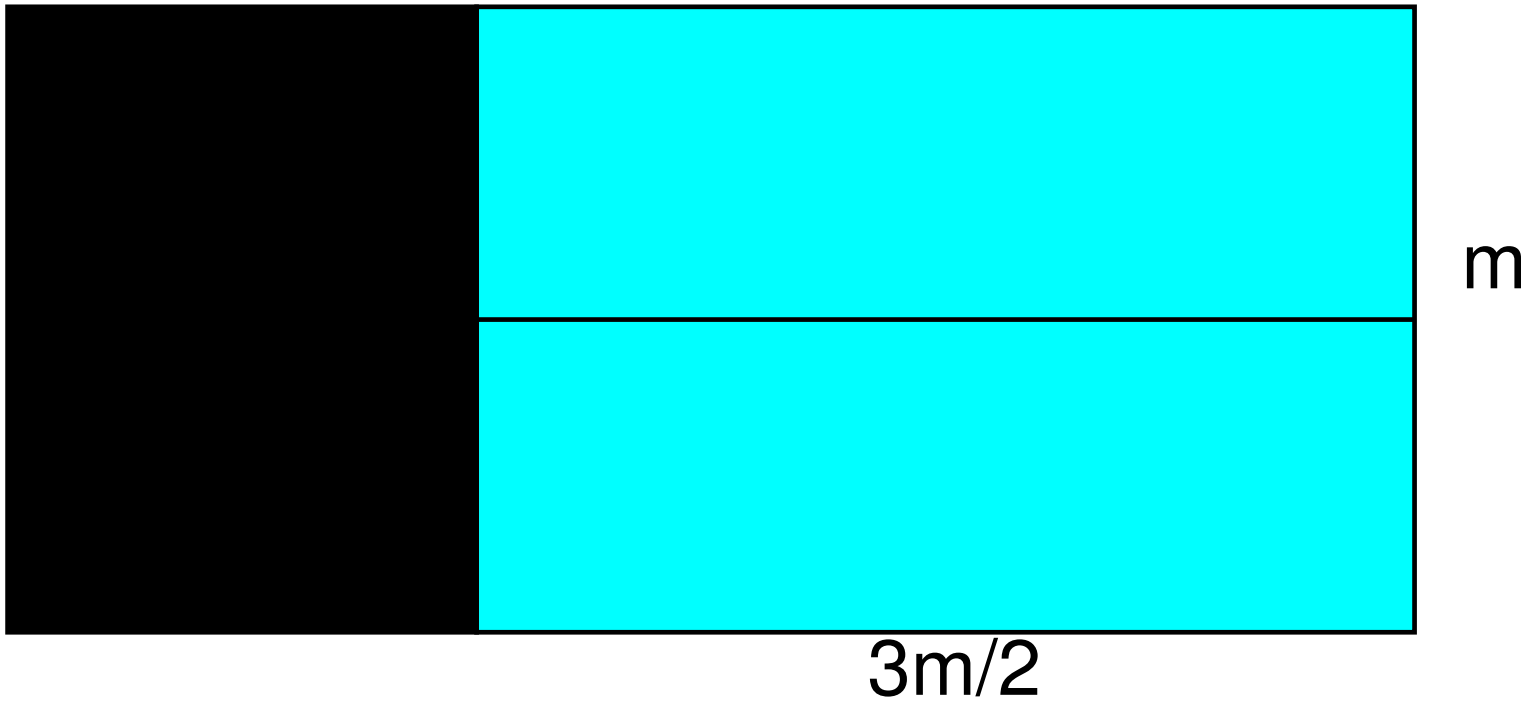
Process



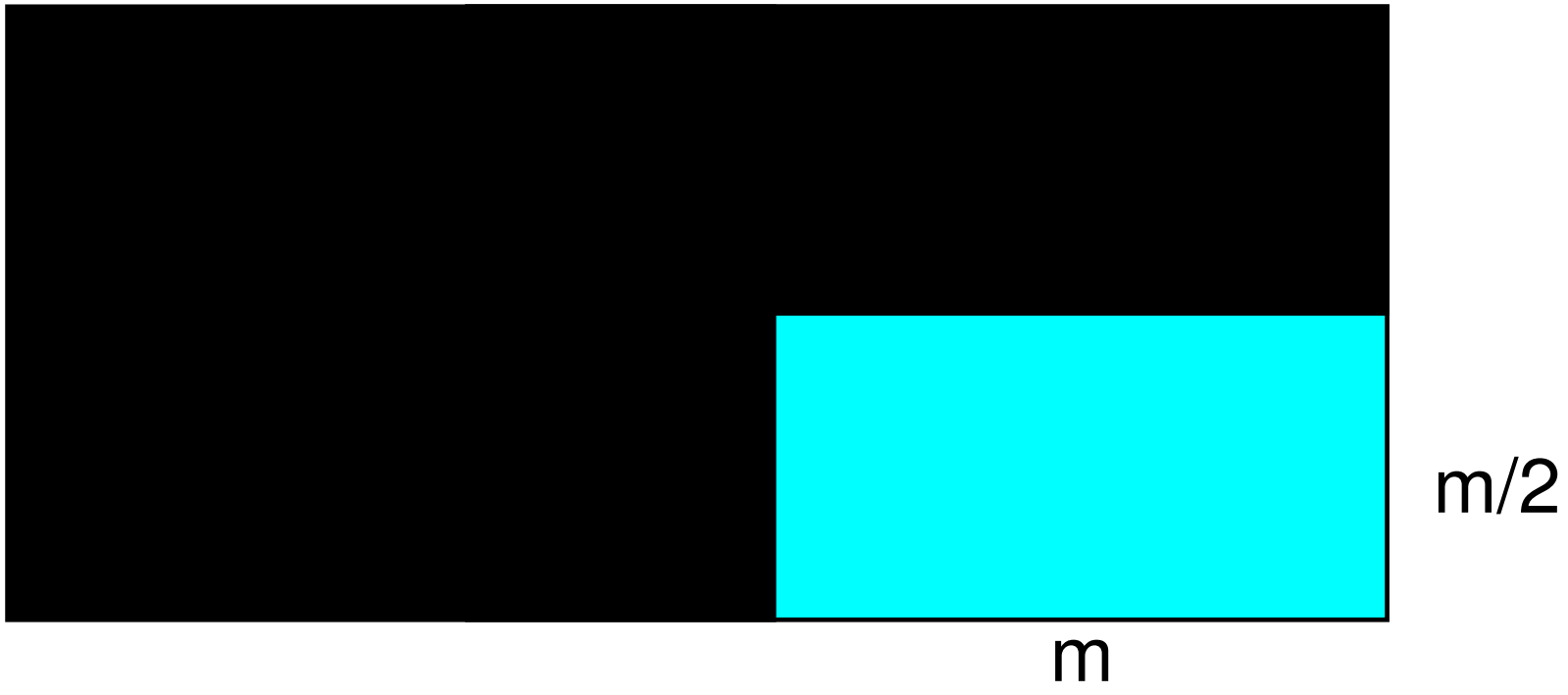
Process



Process



Process



Some Gory Details

To compute an $\vec{s}$-minimal kernel basis of $\mathbf{A}$:

Some Gory Details

To compute an $\vec{s}$-minimal kernel basis of $\mathbf{A}$:

1. Compute an $(\mathbf{A}, \sigma, \vec{s})$-basis $\mathbf{P}$
   - Value of σ?

Some Gory Details

To compute an $\vec{s}$-minimal kernel basis of $\mathbf{A}$:

1. Compute an $(\mathbf{A}, \sigma, \vec{s})$-basis $\mathbf{P}$
   - Value of σ? 3 times the average of the largest m entries of $\vec{s}$

Some Gory Details

To compute an $\vec{s}$-minimal kernel basis of $\mathbf{A}$:

1. Compute an $(\mathbf{A}, \sigma, \vec{s})$-basis $\mathbf{P}$
   - Value of σ? 3 times the average of the largest m entries of $\vec{s}$
2. Let $[\mathbf{P}_1, \mathbf{P}_2] = \mathbf{P}$ such that $\mathbf{A}\mathbf{P}_1 = 0$,

$$\begin{bmatrix}\mathbf{G}_1\\ \mathbf{G}_2\end{bmatrix} = \mathbf{G} = \mathbf{A}\mathbf{P}_2/z^{\sigma}, \text{ and } \vec{t} = \operatorname{cdeg}_{\vec{s}} \mathbf{P}_2 - [\sigma, \ldots, \sigma].$$

Some Gory Details

To compute an $\vec{s}$-minimal kernel basis of $\mathbf{A}$:

1. Compute an $(\mathbf{A}, \sigma, \vec{s})$-basis $\mathbf{P}$
   - Value of σ? 3 times the average of the largest m entries of $\vec{s}$
2. Let $[\mathbf{P}_1, \mathbf{P}_2] = \mathbf{P}$ such that $\mathbf{A}\mathbf{P}_1 = 0$,

$$\begin{bmatrix}\mathbf{G}_1\\ \mathbf{G}_2\end{bmatrix} = \mathbf{G} = \mathbf{A}\mathbf{P}_2/z^{\sigma}, \text{ and } \vec{t} = \operatorname{cdeg}_{\vec{s}} \mathbf{P}_2 - [\sigma, \ldots, \sigma].$$

3. Comp. $\vec{t}$-minimal kernel basis $\mathbf{N}_1$ of $\mathbf{G}_1$ with $\vec{t}$-column degrees $\vec{u}$,

Some Gory Details

To compute an $\vec{s}$-minimal kernel basis of $\mathbf{A}$:

1. Compute an $(\mathbf{A}, \sigma, \vec{s})$-basis $\mathbf{P}$
   - Value of σ? 3 times the average of the largest m entries of $\vec{s}$
2. Let $[\mathbf{P}_1, \mathbf{P}_2] = \mathbf{P}$ such that $\mathbf{A}\mathbf{P}_1 = 0$,

$$\begin{bmatrix}\mathbf{G}_1\\ \mathbf{G}_2\end{bmatrix} = \mathbf{G} = \mathbf{A}\mathbf{P}_2/z^{\sigma}, \text{ and } \vec{t} = \operatorname{cdeg}_{\vec{s}} \mathbf{P}_2 - [\sigma, \ldots, \sigma].$$

3. Comp. $\vec{t}$-minimal kernel basis $\mathbf{N}_1$ of $\mathbf{G}_1$ with $\vec{t}$-column degrees $\vec{u}$,
4. Compute a $\vec{u}$-minimal kernel basis $\mathbf{N}_2$ of $\mathbf{G}_2\mathbf{N}_1$,

Some Gory Details

To compute an $\vec{s}$-minimal kernel basis of $\mathbf{A}$:

1. Compute an $(\mathbf{A}, \sigma, \vec{s})$-basis $\mathbf{P}$
   - Value of σ? 3 times the average of the largest m entries of $\vec{s}$
2. Let $[\mathbf{P}_1, \mathbf{P}_2] = \mathbf{P}$ such that $\mathbf{A}\mathbf{P}_1 = 0$,

   $$\begin{bmatrix}\mathbf{G}_1\\ \mathbf{G}_2\end{bmatrix} = \mathbf{G} = \mathbf{A}\mathbf{P}_2/z^{\sigma}, \text{ and } \vec{t} = \operatorname{cdeg}_{\vec{s}} \mathbf{P}_2 - [\sigma, \ldots, \sigma].$$

3. Comp. $\vec{t}$-minimal kernel basis $\mathbf{N}_1$ of $\mathbf{G}_1$ with $\vec{t}$-column degrees $\vec{u}$,
4. Compute a $\vec{u}$-minimal kernel basis $\mathbf{N}_2$ of $\mathbf{G}_2\mathbf{N}_1$,
5. Return $[\mathbf{P}_1, \mathbf{P}_2\mathbf{N}_1\mathbf{N}_2]$.

Cost of Tools

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

Theorem (Zhou-L-Storjohann, ISSAC 2012)

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

Cost of Tools

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

Theorem (Zhou-L-Storjohann, ISSAC 2012)

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

Theorem (Zhou-L, ISSAC 2013)

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

Triangularization

- Finding Diagonals
- Complexity
- Aside: Computing the determinant

Finding Diagonal Elements

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

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

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) $\mathbf{A}_u \mathbf{U}_r = 0$. So $\mathbf{U}_r$ a right kernel basis of $\mathbf{A}_u$.

(ii) $\mathbf{A}_u \mathbf{U}_\ell = \mathbf{B}_1$. So $\mathbf{B}_1$ is nonsingular and a column 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) $\mathbf{A}_u \mathbf{U}_r = 0$. So $\mathbf{U}_r$ a right kernel basis of $\mathbf{A}_u$.

(ii) $\mathbf{A}_u \mathbf{U}_\ell = \mathbf{B}_1$. So $\mathbf{B}_1$ is nonsingular and a column 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

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

Here

(i) $\mathbf{A}_u \mathbf{U}_r = 0$. So $\mathbf{U}_r$ a right kernel basis of $\mathbf{A}_u$.

(ii) $\mathbf{A}_u \mathbf{U}_\ell = \mathbf{B}_1$. So $\mathbf{B}_1$ is nonsingular and a column 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

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

Example

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

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

Example

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

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

Example

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

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

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)

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

Theorem

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

Theorem

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

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

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

Complexity

Theorem

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

Complexity

Theorem

$\mathbf{A} \in \mathbb{K}[z]^{n \times n}$. *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

$\mathbf{A} \in \mathbb{K}[z]^{n \times n}$. *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

$\mathbf{A} \in \mathbb{K}[z]^{n \times n}$. *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 $s = \frac{\xi}{n}$)

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

□

Finding Rest of $\mathbf{H}$

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 $\mathbf{H}$

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 $\mathbf{H}$

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): $\mathbf{A} \to \mathcal{L}(\mathbf{A})$

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 $\mathbf{H} = \mathcal{E}_{\vec{\delta}} \cdot \widetilde{\mathbf{H}}$ 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.

$$\begin{array}{ccccc}
\mathbf{x}^{\vec{\mu}-\vec{\delta}}\,\mathbf{A} & \xrightarrow{\text{reduce}} & \mathbf{x}^{\vec{\mu}-\vec{\delta}}\,\mathbf{R} & \xrightarrow{\text{normalize}} & \mathbf{H} = \mathbf{R}\,\mathrm{lc}_{-\vec{\delta}}(\mathbf{R})^{-1} \\
\Big| & & & & \Big| \\
\text{partial linearization} & & & & \text{partial linearization} \\
\downarrow & & & & \downarrow \\
\mathbf{x}^{\vec{m}-\vec{d}}\,\mathcal{L}_{\vec{\delta}}(\mathbf{A}) & \xrightarrow{\text{reduce}} & \mathbf{x}^{\vec{m}-\vec{d}}\,\hat{\mathbf{R}} & \xrightarrow{\text{normalize}} & \mathcal{L}_{\vec{\delta}}(\mathbf{H}) = \hat{\mathbf{R}}\,\mathrm{lc}_{-\vec{d}}(\hat{\mathbf{R}})^{-1}
\end{array} .$$

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:* $O^{\sim}(n^{\omega}\lceil s \rceil)$.

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 $\det \mathbf{U} = \frac{\det U^*}{\det V_u U_\ell{}^*}$

Some 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*