

Vincent Neiger

Laboratoire LIP6, Sorbonne Université

`vincent.neiger@lip6.fr`

polynomial matrices: fast approximation and applications

Algorithmes Efficaces en Calcul Formel
Master Parisien de Recherche en Informatique
22 October 2025

outline

- ▶ introduction
- ▶ shifted reduced forms
- ▶ fast algorithms
- ▶ applications

outline

▶ introduction

- ▶ rational approximation and interpolation
- ▶ the vector case
- ▶ pol. matrices: reminders and motivation

▶ shifted reduced forms

▶ fast algorithms

▶ applications

introduction

⇓ earlier in the course ⇓

⇓ in this lecture ⇓

introduction

⇓ earlier in the course ⇓

- ▶ addition $f + g$, multiplication $f * g$
- ▶ division with remainder $f = qg + r$
- ▶ truncated inverse $f^{-1} \bmod X^d$
- ▶ extended GCD $uf + vg = \gcd(f, g)$
- ▶ multipoint eval. $f \mapsto f(\alpha_1), \dots, f(\alpha_d)$
- ▶ interpolation $f(\alpha_1), \dots, f(\alpha_d) \mapsto f$
- ▶ Padé approximation $f = \frac{p}{q} \bmod X^d$
- ▶ minpoly of linearly recurrent sequence

⇓ in this lecture ⇓

introduction

⇓ earlier in the course ⇓

$O(M(d))$

- ▶ addition $f + g$, multiplication $f * g$
- ▶ division with remainder $f = qg + r$
- ▶ truncated inverse $f^{-1} \bmod X^d$
- ▶ extended GCD $uf + vg = \gcd(f, g)$

$O(M(d) \log(d))$

- ▶ multipoint eval. $f \mapsto f(\alpha_1), \dots, f(\alpha_d)$
- ▶ interpolation $f(\alpha_1), \dots, f(\alpha_d) \mapsto f$
- ▶ Padé approximation $f = \frac{p}{q} \bmod X^d$
- ▶ minpoly of linearly recurrent sequence

⇓ in this lecture ⇓

introduction

⇓ earlier in the course ⇓

$O(M(d))$

- ▶ addition $f + g$, multiplication $f * g$
- ▶ division with remainder $f = qg + r$
- ▶ truncated inverse $f^{-1} \bmod X^d$
- ▶ extended GCD $uf + vg = \gcd(f, g)$

$O(M(d) \log(d))$

- ▶ multipoint eval. $f \mapsto f(\alpha_1), \dots, f(\alpha_d)$
- ▶ interpolation $f(\alpha_1), \dots, f(\alpha_d) \mapsto f$
- ▶ Padé approximation $f = \frac{p}{q} \bmod X^d$
- ▶ minpoly of linearly recurrent sequence

⇓ in this lecture ⇓

Padé approximation, sequence minpoly, extended GCD

$O(M(d) \log(d))$ operations in \mathbb{K}

matrix versions of these problems

$O(m^\omega M(d) \log(d))$ operations in \mathbb{K}

or a tiny bit more for matrix-GCD

introduction

rational approximation and interpolation

given **power series** $p(X)$ and $q(X)$ over \mathbb{K} at precision d ,
with $q(X)$ invertible,

→ compute $\frac{p(X)}{q(X)} \bmod X^d$

algo?? $O(??)$

introduction

rational approximation and interpolation

given power series $p(X)$ and $q(X)$ over \mathbb{K} at precision d ,
with $q(X)$ invertible,

→ compute $\frac{p(X)}{q(X)} \bmod X^d$

algo?? $O(??)$
inv+mul: $O(M(d))$

introduction

rational approximation and interpolation

given **power series** $p(X)$ and $q(X)$ over \mathbb{K} at precision d ,
with $q(X)$ invertible,

→ compute $\frac{p(X)}{q(X)} \bmod X^d$

algo?? $O(??)$

inv+mul: $O(M(d))$

given $M(X) \in \mathbb{K}[X]$ of degree $d > 0$,

given **polynomials** $p(X)$ and $q(X)$ over \mathbb{K} of degree $< d$,

with $q(X)$ invertible modulo $M(X)$,

what does that mean?

→ compute $\frac{p(X)}{q(X)} \bmod M(X)$

algo?? $O(??)$

introduction

rational approximation and interpolation

given **power series** $p(X)$ and $q(X)$ over \mathbb{K} at precision d ,
with $q(X)$ invertible,

→ compute $\frac{p(X)}{q(X)} \bmod X^d$

algo?? $O(??)$

inv+mul: $O(M(d))$

given $M(X) \in \mathbb{K}[X]$ of degree $d > 0$,

given **polynomials** $p(X)$ and $q(X)$ over \mathbb{K} of degree $< d$,

with $q(X)$ invertible modulo $M(X)$,

what does that mean?

→ compute $\frac{p(X)}{q(X)} \bmod M(X)$

algo?? $O(??)$

xgcd+mul+rem $O(M(d) \log(d))$

introduction

rational approximation and interpolation

given **power series** $p(X)$ and $q(X)$ over \mathbb{K} at precision d ,
with $q(X)$ invertible,

→ compute $\frac{p(X)}{q(X)} \bmod X^d$

algo?? $O(??)$
inv+mul: $O(M(d))$

given $M(X) \in \mathbb{K}[X]$ of degree $d > 0$,

given **polynomials** $p(X)$ and $q(X)$ over \mathbb{K} of degree $< d$,

with $q(X)$ invertible modulo $M(X)$,

what does that mean?

→ compute $\frac{p(X)}{q(X)} \bmod M(X)$

algo?? $O(??)$
xgcd+mul+rem $O(M(d) \log(d))$

given $M(X) = (X - \alpha_1) \cdots (X - \alpha_d) \in \mathbb{K}[X]$,

for pairwise distinct $\alpha_1, \dots, \alpha_d \in \mathbb{K}$,

given **polynomials** $p(X)$ and $q(X)$ over \mathbb{K} of degree $< d$,

with $q(X)$ invertible modulo $M(X)$,

what does that mean?

→ compute $\frac{p(X)}{q(X)} \bmod M(X)$

algo?? $O(??)$

introduction

rational approximation and interpolation

given **power series** $p(X)$ and $q(X)$ over \mathbb{K} at precision d ,
with $q(X)$ invertible,

→ compute $\frac{p(X)}{q(X)} \bmod X^d$

algo?? $O(??)$
inv+mul: $O(M(d))$

given $M(X) \in \mathbb{K}[X]$ of degree $d > 0$,

given **polynomials** $p(X)$ and $q(X)$ over \mathbb{K} of degree $< d$,

with $q(X)$ invertible modulo $M(X)$,

what does that mean?

→ compute $\frac{p(X)}{q(X)} \bmod M(X)$

algo?? $O(??)$
xgcd+mul+rem $O(M(d) \log(d))$

given $M(X) = (X - \alpha_1) \cdots (X - \alpha_d) \in \mathbb{K}[X]$,

for pairwise distinct $\alpha_1, \dots, \alpha_d \in \mathbb{K}$,

given **polynomials** $p(X)$ and $q(X)$ over \mathbb{K} of degree $< d$,

with $q(X)$ invertible modulo $M(X)$,

what does that mean?

→ compute $\frac{p(X)}{q(X)} \bmod M(X)$

algo?? $O(??)$
eval+div+interp $O(M(d) \log(d))$

introduction

rational approximation and interpolation

rational fractions \longleftrightarrow linearly recurrent sequences
reminders from lecture 4

introduction

rational approximation and interpolation

rational fractions \longleftrightarrow linearly recurrent sequences
reminders from lecture 4

Duality lemma (link between C-recursive sequences and rational functions)

Let $A(x) = \sum_{n \geq 0} u_n x^n \in \mathbb{K}[[x]]$ be the generating function of $(u_n)_{n \geq 0}$.

The following assertions are equivalent:

- (i) $(u_n)_{n \geq 0}$ is **C-recursive**, with characteristic polynomial Γ of degree d ;
- (ii) **$A(x)$ is rational**, $A = \frac{P}{Q}$ with $P \in \mathbb{K}[x]_{<d}$ and $Q = \text{rev}(\Gamma) := \Gamma(\frac{1}{x})x^d$.

▷ The *denominator* of A encodes a *recurrence* for $(u_n)_{n \geq 0}$; the *numerator* encodes *initial conditions*.

introduction

rational approximation and interpolation

rational fractions \longleftrightarrow linearly recurrent sequences
reminders from lecture 4

Duality lemma (link between C-recursive sequences and rational functions)

Let $A(x) = \sum_{n \geq 0} u_n x^n \in \mathbb{K}[[x]]$ be the generating function of $(u_n)_{n \geq 0}$.

The following assertions are equivalent:

- (i) $(u_n)_{n \geq 0}$ is **C-recursive**, with characteristic polynomial Γ of degree d ;
- (ii) **$A(x)$ is rational**, $A = \frac{P}{Q}$ with $P \in \mathbb{K}[x]_{<d}$ and $Q = \text{rev}(\Gamma) := \Gamma(\frac{1}{x})x^d$.

▷ The *denominator* of A encodes a *recurrence* for $(u_n)_{n \geq 0}$; the *numerator* encodes *initial conditions*.

↪ compute N first terms of a linearly recurrent sequence

↪ see slides 23 and 24, including Shoup's algorithm

introduction

rational approximation and interpolation

rational fractions \longleftrightarrow linearly recurrent sequences
reminders from lecture 4

expand $\frac{P}{\text{rev}(\Gamma)} \bmod X^N$

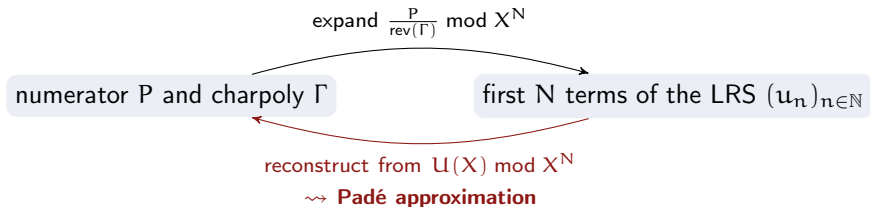
numerator P and charpoly Γ

first N terms of the LRS $(u_n)_{n \in \mathbb{N}}$

introduction

rational approximation and interpolation

rational fractions \longleftrightarrow linearly recurrent sequences
reminders from lecture 4



introduction

rational approximation and interpolation

Padé approximation:

given **power series** $f(X)$ at precision d ,

→ compute $p(X), q(X)$ such that $f = \frac{p}{q} \bmod X^d$

introduction

rational approximation and interpolation

Padé approximation:

given **power series** $f(X)$ at precision d ,

→ compute $p(X), q(X)$ such that $f = \frac{p}{q} \bmod X^d$

opinions on this algorithmic problem?

introduction

rational approximation and interpolation

Padé approximation:

given **power series** $f(X)$ at precision d ,

given **degree constraints** $d_1, d_2 > 0$,

→ compute **polynomials** $(p(X), q(X))$ of **degrees** $< (d_1, d_2)$

and such that $f = \frac{p}{q} \bmod X^d$

introduction

rational approximation and interpolation

Padé approximation:

given **power series** $f(X)$ at precision d ,

given **degree constraints** $d_1, d_2 > 0$,

→ compute **polynomials** $(p(X), q(X))$ of **degrees** $< (d_1, d_2)$

and such that $f = \frac{p}{q} \bmod X^d$

Cauchy interpolation:

given $M(X) = (X - \alpha_1) \cdots (X - \alpha_d) \in \mathbb{K}[X]$,

for pairwise distinct $\alpha_1, \dots, \alpha_d \in \mathbb{K}$,

given **degree constraints** $d_1, d_2 > 0$,

→ compute **polynomials** $(p(X), q(X))$ of **degrees** $< (d_1, d_2)$

and such that $f = \frac{p}{q} \bmod M(X)$

introduction

rational approximation and interpolation

Padé approximation:

given **power series** $f(X)$ at precision d ,

given **degree constraints** $d_1, d_2 > 0$,

→ compute **polynomials** $(p(X), q(X))$ of **degrees** $< (d_1, d_2)$

and such that $f = \frac{p}{q} \bmod X^d$

Cauchy interpolation:

given $M(X) = (X - \alpha_1) \cdots (X - \alpha_d) \in \mathbb{K}[X]$,

for pairwise distinct $\alpha_1, \dots, \alpha_d \in \mathbb{K}$,

given **degree constraints** $d_1, d_2 > 0$,

→ compute **polynomials** $(p(X), q(X))$ of **degrees** $< (d_1, d_2)$

and such that $f = \frac{p}{q} \bmod M(X)$

- ▶ degree constraints specified by the context
- ▶ usual choices have $d_1 + d_2 \approx d$ and existence of a solution

Sur la généralisation des fractions continues algébriques;

PAR M. H. PADÉ,

Docteur ès Sciences mathématiques,
Professeur au lycée de Lille.

[1894, Journal de mathématiques pures et appliquées]

INTRODUCTION.

M. Hermite s'est, dans un travail récemment paru (1), occupé de la généralisation des fractions continues algébriques. La question est de déterminer les polynomes X_1, X_2, \dots, X_n , de degrés $\mu_1, \mu_2, \dots, \mu_n$, qui satisfont à l'équation

$$S_1 X_1 + S_2 X_2 + \dots + S_n X_n = S x^{\mu_1 + \mu_2 + \dots + \mu_n + n - 1},$$

S_1, S_2, \dots, S_n étant des séries entières données, et S une série également entière. Ou plutôt, il s'agit d'obtenir un algorithme qui permette le calcul de proche en proche de ces systèmes de n polynomes, et qui soit analogue à l'algorithme par lequel le numérateur et le dénominateur d'une réduite d'une fraction continue se déduisent des numérateurs et dénominateurs des réduites précédentes. D'élégantes considé-

Hermite-Padé approximation

[Hermite 1893, Padé 1894]

input:

- ▶ polynomials $f_1, \dots, f_m \in \mathbb{K}[X]$
- ▶ precision $d \in \mathbb{Z}_{>0}$
- ▶ degree bounds $d_1, \dots, d_m \in \mathbb{Z}_{>0}$

output:

polynomials $p_1, \dots, p_m \in \mathbb{K}[X]$ such that

- ▶ $p_1 f_1 + \dots + p_m f_m = 0 \bmod X^d$
- ▶ $\text{cdeg}([p_1 \dots p_m]) < (d_1, \dots, d_m)$

(Padé approximation: particular case $m = 2$ and $f_2 = -1$)

M-Padé approximation / vector rational interpolation

[Cauchy 1821, Mahler 1968]

input:

- ▶ polynomials $f_1, \dots, f_m \in \mathbb{K}[X]$
- ▶ pairwise distinct points $\alpha_1, \dots, \alpha_d \in \mathbb{K}$
- ▶ degree bounds $d_1, \dots, d_m \in \mathbb{Z}_{>0}$

output:

polynomials $p_1, \dots, p_m \in \mathbb{K}[X]$ such that

- ▶ $p_1(\alpha_i)f_1(\alpha_i) + \dots + p_m(\alpha_i)f_m(\alpha_i) = 0$ for all $1 \leq i \leq d$
- ▶ $\text{cdeg}([p_1 \cdots p_m]) < (d_1, \dots, d_m)$

(rational interpolation: particular case $m = 2$ and $f_2 = -1$)

introduction

approximation and interpolation: the vector case

in this lecture: modular reconstruction and fast algorithms

[van Barel-Bultheel 1992; Beckermann-Labahn 1994, 1997, 2000; Giorgi-Jeannerod-Villard 2003; Storjohann 2006; Zhou-Labahn 2012; Jeannerod-Neiger-Schost-Villard 2017, 2020]

input:

- ▶ polynomials $f_1, \dots, f_m \in \mathbb{K}[X]$
- ▶ field elements $\alpha_1, \dots, \alpha_d \in \mathbb{K}$ \rightsquigarrow not necessarily distinct
- ▶ degree bounds $d_1, \dots, d_m \in \mathbb{Z}_{>0}$ \rightsquigarrow general “shift” $s \in \mathbb{Z}^m$

output:

polynomials $p_1, \dots, p_m \in \mathbb{K}[X]$ such that

- ▶ $p_1 f_1 + \dots + p_m f_m = 0 \bmod \prod_{1 \leq i \leq d} (X - \alpha_i)$
- ▶ $\text{cdeg}([p_1 \cdots p_m]) < (d_1, \dots, d_m)$ \rightsquigarrow minimal s -row degree

(Hermite-Padé: $\alpha_1 = \dots = \alpha_d = 0$; interpolation: pairwise distinct points)

introduction

approximation and interpolation: the vector case

applications:

- univariate polynomials and linearly recurrent sequences

XGCD, rational reconstruction, “fast Berlekamp-Massey”, ...

- sparse \mathbb{K} -linear systems

Coppersmith's block-Wiedemann approach

- structured \mathbb{K} -matrices

Hankel/Toeplitz/Vandermonde, block structures, displacement rank, ...

- computations with \mathbb{K} -matrices

Krylov iterates, minimal/characteristic polynomial, Frobenius form, ...

- computations with $\mathbb{K}[X]$ -matrices

determinant, nullspace/kernel, inversion, Hermite normal form, ...

- computations with multivariate polynomials

multivariate interpolation, syzygy modules, Gröbner bases, ...

introduction

approximation and structured linear system

$$\mathbb{K} = \mathbb{F}_7$$

$$f = 2X^7 + 2X^6 + 5X^4 + 2X^2 + 4$$

$$d = 8, d_1 = 3, d_2 = 6$$

→ look for (p, q) of degree $< (3, 6)$ such that $f = \frac{p}{q} \bmod X^8$

$$\begin{bmatrix} q & p \end{bmatrix} \begin{bmatrix} f \\ -1 \end{bmatrix} = 0 \bmod X^8$$

introduction

approximation and structured linear system

$$\mathbb{K} = \mathbb{F}_7$$

$$f = 2X^7 + 2X^6 + 5X^4 + 2X^2 + 4$$

$$d = 8, d_1 = 3, d_2 = 6$$

→ look for (p, q) of degree $< (3, 6)$ such that $f = \frac{p}{q} \bmod X^8$

$$\begin{bmatrix} q & p \end{bmatrix} \begin{bmatrix} f \\ -1 \end{bmatrix} = 0 \bmod X^8$$

$$\begin{bmatrix} q_0 & q_1 & q_2 & q_3 & q_4 & q_5 & | & p_0 & p_1 & p_2 \end{bmatrix} \begin{bmatrix} 4 & 0 & 2 & 0 & 5 & 0 & 2 & 2 \\ & 4 & 0 & 2 & 0 & 5 & 0 & 2 \\ & & 4 & 0 & 2 & 0 & 5 & 0 \\ & & & 4 & 0 & 2 & 0 & 5 \\ & & & & 4 & 0 & 2 & 0 \\ & & & & & 4 & 0 & 2 \\ \hline 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 6 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & 6 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} = 0$$

introduction

approximation and structured linear system

$$\mathbb{K} = \mathbb{F}_7$$

$$f = 2X^7 + 2X^6 + 5X^4 + 2X^2 + 4$$

$$d = 8, d_1 = 3, d_2 = 6$$

→ look for (p, q) of degree $< (3, 6)$ such that $f = \frac{p}{q} \bmod X^8$

$$\begin{bmatrix} q & p \end{bmatrix} \begin{bmatrix} f \\ -1 \end{bmatrix} = 0 \bmod X^8$$

$$[q_0 \ q_1 \ q_2 \ q_3 \ q_4 \ q_5 \mid p_0 \ p_1 \ p_2]$$

$$\begin{bmatrix} 4 & 0 & 2 & 0 & 5 & 0 & 2 & 2 \\ & 4 & 0 & 2 & 0 & 5 & 0 & 2 \\ & & 4 & 0 & 2 & 0 & 5 & 0 \\ & & & 4 & 0 & 2 & 0 & 5 \\ & & & & 4 & 0 & 2 & 0 \\ & & & & & 4 & 0 & 2 \\ & & & & & & 4 & 0 \\ & & & & & & & 4 \end{bmatrix}$$

$$= 0$$

introduction

interpolation and structured linear system

application of vector rational interpolation:

given pairwise distinct points $\{(\alpha_i, \beta_i), 1 \leq i \leq 8\}$
 $= \{(24, 80), (31, 73), (15, 73), (32, 35), (83, 66), (27, 46), (20, 91), (59, 64)\}$,
compute a **bivariate** polynomial $p(X, Y) \in \mathbb{K}[X, Y]$
such that $p(\alpha_i, \beta_i) = 0$ for $1 \leq i \leq 8$

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

solutions of smaller X-degree: $p(X, Y) = p_0(X) + p_1(X)Y + p_2(X)Y^2$

$$p(X, L(X)) = \begin{bmatrix} p_0 & p_1 & p_2 \end{bmatrix} \begin{bmatrix} 1 \\ L \\ L^2 \end{bmatrix} = 0 \bmod M(X)$$

- instance of **univariate** rational vector interpolation
- with a **structured** input equation (powers of $L \bmod M$)

introduction

interpolation and structured linear system

application of vector rational interpolation:

given pairwise distinct points $\{(\alpha_i, \beta_i), 1 \leq i \leq 8\}$

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

compute a **bivariate** polynomial $p(X, Y) \in \mathbb{K}[X, Y]$

such that $p(\alpha_i, \beta_i) = 0$ for $1 \leq i \leq 8$

add **degree constraints**: seek $p(X, Y)$ of the form

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

$$\begin{bmatrix} p_{00} & p_{01} & p_{02} & p_{03} & p_{04} & \vdots & p_{10} & p_{11} & p_{12} & \vdots & p_{20} \end{bmatrix} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_8 \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_8^2 \\ \alpha_1^3 & \alpha_2^3 & \cdots & \alpha_8^3 \\ \alpha_1^4 & \alpha_2^4 & \cdots & \alpha_8^4 \\ \hline \beta_1 & \beta_2 & \cdots & \beta_8 \\ \alpha_1\beta_1 & \alpha_2\beta_2 & \cdots & \alpha_8\beta_8 \\ \alpha_1^2\beta_1 & \alpha_2^2\beta_2 & \cdots & \alpha_8^2\beta_8 \\ \hline \beta_1^2 & \beta_2^2 & \cdots & \beta_8^2 \end{bmatrix} = 0$$

- **\mathbb{K} -linear** system
- **two levels** of structure

$$p(X, Y) = (2X^4 + 56X^3 + 42X^2 + 48X + 15) + (72X^2 + 12X + 30)Y + Y^2$$

introduction

polynomial matrices: reminder and motivation

why polynomial matrices here?

introduction

polynomial matrices: reminder and motivation

why polynomial matrices here?

omitting degree constraints, the set of solutions is

$$\mathcal{S} = \{(p_1, \dots, p_m) \in \mathbb{K}[X]^m \mid p_1 f_1 + \dots + p_m f_m = 0 \bmod M\}$$

recall $M(X) = \prod_{1 \leq i \leq d} (X - \alpha_i)$

introduction

polynomial matrices: reminder and motivation

why polynomial matrices here?

omitting degree constraints, the set of solutions is

$$\mathcal{S} = \{(p_1, \dots, p_m) \in \mathbb{K}[X]^m \mid p_1 f_1 + \dots + p_m f_m = 0 \bmod M\}$$

recall $M(X) = \prod_{1 \leq i \leq d} (X - \alpha_i)$

\mathcal{S} is a “free $\mathbb{K}[X]$ -module of rank m ”, meaning:

- ▶ stable under $\mathbb{K}[X]$ -linear combinations
- ▶ admits a basis consisting of m elements
- ▶ basis = $\mathbb{K}[X]$ -linear independence + generates all solutions

introduction

polynomial matrices: reminder and motivation

why polynomial matrices here?

omitting degree constraints, the set of solutions is

$$\mathcal{S} = \{(p_1, \dots, p_m) \in \mathbb{K}[X]^m \mid p_1 f_1 + \dots + p_m f_m = 0 \bmod M\}$$

$$\text{recall } M(X) = \prod_{1 \leq i \leq d} (X - \alpha_i)$$

\mathcal{S} is a “free $\mathbb{K}[X]$ -module of rank m ”, meaning:

- ▶ stable under $\mathbb{K}[X]$ -linear combinations
- ▶ admits a basis consisting of m elements
- ▶ basis = $\mathbb{K}[X]$ -linear independence + generates all solutions

$$\triangleright \mathcal{S} \subset \mathbb{K}[X]^m \Rightarrow \mathcal{S} \text{ has rank } \leq m$$

$$\triangleright M(X)\mathbb{K}[X]^m \subset \mathcal{S} \Rightarrow \mathcal{S} \text{ has rank } \geq m$$

remark: solutions are not considered modulo M

e.g. $(M, 0, \dots, 0)$ is in \mathcal{S} and may appear in a basis

introduction

polynomial matrices: reminder and motivation

why polynomial matrices here?

omitting degree constraints, the set of solutions is

$$\mathcal{S} = \{(p_1, \dots, p_m) \in \mathbb{K}[X]^m \mid p_1 f_1 + \dots + p_m f_m = 0 \bmod M\}$$

recall $M(X) = \prod_{1 \leq i \leq d} (X - \alpha_i)$

basis of solutions:

- ▶ square nonsingular matrix \mathbf{P} in $\mathbb{K}[X]^{m \times m}$
- ▶ each row of \mathbf{P} is a solution
- ▶ any solution is a $\mathbb{K}[X]$ -combination \mathbf{uP} , $\mathbf{u} \in \mathbb{K}[X]^{1 \times m}$

i.e. \mathcal{S} is the $\mathbb{K}[X]$ -row space of \mathbf{P}

introduction

polynomial matrices: reminder and motivation

why polynomial matrices here?

omitting degree constraints, the set of solutions is

$$\mathcal{S} = \{(p_1, \dots, p_m) \in \mathbb{K}[X]^m \mid p_1 f_1 + \dots + p_m f_m = 0 \bmod M\}$$

recall $M(X) = \prod_{1 \leq i \leq d} (X - \alpha_i)$

basis of solutions:

- ▶ square nonsingular matrix \mathbf{P} in $\mathbb{K}[X]^{m \times m}$
- ▶ each row of \mathbf{P} is a solution
- ▶ any solution is a $\mathbb{K}[X]$ -combination \mathbf{uP} , $\mathbf{u} \in \mathbb{K}[X]^{1 \times m}$

i.e. \mathcal{S} is the $\mathbb{K}[X]$ -row space of \mathbf{P}

prove: $\det(\mathbf{P})$ is a divisor of $M(X)^m$

introduction

polynomial matrices: reminder and motivation

why polynomial matrices here?

omitting degree constraints, the set of solutions is

$$\mathcal{S} = \{(p_1, \dots, p_m) \in \mathbb{K}[X]^m \mid p_1 f_1 + \dots + p_m f_m = 0 \bmod M\}$$

$$\text{recall } M(X) = \prod_{1 \leq i \leq d} (X - \alpha_i)$$

basis of solutions:

- ▶ square nonsingular matrix \mathbf{P} in $\mathbb{K}[X]^{m \times m}$
- ▶ each row of \mathbf{P} is a solution
- ▶ any solution is a $\mathbb{K}[X]$ -combination \mathbf{uP} , $\mathbf{u} \in \mathbb{K}[X]^{1 \times m}$

i.e. \mathcal{S} is the $\mathbb{K}[X]$ -row space of \mathbf{P}

prove: $\det(\mathbf{P})$ is a divisor of $M(X)^m$

prove: any other basis is \mathbf{UP} for $\mathbf{U} \in \mathbb{K}[X]^{m \times m}$ with $\det(\mathbf{U}) \in \mathbb{K} \setminus \{0\}$

introduction

polynomial matrices: reminder and motivation

why polynomial matrices here?

omitting degree constraints, the set of solutions is

$$\mathcal{S} = \{(p_1, \dots, p_m) \in \mathbb{K}[X]^m \mid p_1 f_1 + \dots + p_m f_m = 0 \bmod M\}$$

recall $M(X) = \prod_{1 \leq i \leq d} (X - \alpha_i)$

basis of solutions:

- ▶ square nonsingular matrix \mathbf{P} in $\mathbb{K}[X]^{m \times m}$
- ▶ each row of \mathbf{P} is a solution
- ▶ any solution is a $\mathbb{K}[X]$ -combination \mathbf{uP} , $\mathbf{u} \in \mathbb{K}[X]^{1 \times m}$

i.e. \mathcal{S} is the $\mathbb{K}[X]$ -row space of \mathbf{P}

computing a **basis** of \mathcal{S} with “**minimal degrees**”

- ▶ has many more applications than a single small-degree solution
- ▶ is in most cases the fastest known strategy anyway(!)

\rightsquigarrow degree minimality ensured via **shifted reduced forms**

introduction

polynomial matrices: reminder and motivation

$$A = \begin{bmatrix} 3X + 4 & X^3 + 4X + 1 & 4X^2 + 3 \\ 5 & 5X^2 + 3X + 1 & 5X + 3 \\ 3X^3 + X^2 + 5X + 3 & 6X + 5 & 2X + 1 \end{bmatrix} \in \mathbb{K}[X]^{3 \times 3}$$

3×3 matrix of degree 3
with entries in $\mathbb{K}[X] = \mathbb{F}_7[X]$

operations in $\mathbb{K}[X]_{<d}^{m \times m}$:

- combination of matrix and polynomial computations
- addition in $O(m^2 d)$, naive multiplication in $O(m^3 d^2)$
- some tools shared with \mathbb{K} -matrices, others specific to $\mathbb{K}[X]$ -matrices

[Cantor-Kaltofen'91]

multiplication in $O(m^\omega d \log(d) + m^2 d \log(d) \log \log(d))$

$\in O(m^\omega M(d)) \subset \tilde{O}(m^\omega d)$

introduction

polynomial matrices: reminder and motivation

$$A = \begin{bmatrix} 3X + 4 & X^3 + 4X + 1 & 4X^2 + 3 \\ 5 & 5X^2 + 3X + 1 & 5X + 3 \\ 3X^3 + X^2 + 5X + 3 & 6X + 5 & 2X + 1 \end{bmatrix} \in \mathbb{K}[X]^{3 \times 3}$$

3×3 matrix of degree 3
with entries in $\mathbb{K}[X] = \mathbb{F}_7[X]$

operations in $\mathbb{K}[X]_{<d}^{m \times m}$:

- ▶ combination of matrix and polynomial computations
- ▶ addition in $O(m^2 d)$, naive multiplication in $O(m^3 d^2)$
- ▶ some tools shared with \mathbb{K} -matrices, others specific to $\mathbb{K}[X]$ -matrices

[Cantor-Kaltofen'91]

multiplication in $O(m^\omega d \log(d) + m^2 d \log(d) \log \log(d))$

$\in O(m^\omega M(d)) \subset \tilde{O}(m^\omega d)$

- ▶ Newton truncated inversion, matrix-QuoRem
- ▶ inversion and determinant via evaluation-interpolation
- ▶ vector rational approximation & interpolation

→ fast $\tilde{O}(m^\omega d)$

→ medium $\tilde{O}(m^{\omega+1} d)$

→ ???

introduction

polynomial matrices: reminder and motivation

reductions of most problems to polynomial matrix multiplication

matrix $m \times m$ of degree d $\rightarrow O(m^\omega d)$
of “average” degree $\frac{D}{m}$ $\rightarrow O(m^\omega \frac{D}{m})$

classical matrix operations

- ▶ multiplication
- ▶ kernel, system solving
- ▶ rank, determinant
- ▶ inversion $O(m^3 d)$

univariate specific operations

- ▶ truncated inverse, QuoRem
- ▶ Hermite-Padé approximation
- ▶ vector rational interpolation
- ▶ syzygies / modular equations

transformation to normal forms

- ▶ triangularization: Hermite form
- ▶ row reduction: Popov form
- ▶ diagonalization: Smith form

introduction

polynomial matrices: reminder and motivation

reductions of most problems to polynomial matrix multiplication

matrix $m \times m$ of degree d $\rightarrow O(m^\omega d)$
of “average” degree $\frac{D}{m}$ $\rightarrow O(m^\omega \frac{D}{m})$

classical matrix operations

- ▶ multiplication
- ▶ kernel, system solving
- ▶ rank, determinant
- ▶ inversion $O(m^3 d)$

univariate specific operations

- ▶ truncated inverse, QuoRem
- ▶ Hermite-Padé approximation
- ▶ vector rational interpolation
- ▶ syzygies / modular equations

transformation to normal forms

- ▶ triangularization: Hermite form
- ▶ row reduction: Popov form
- ▶ diagonalization: Smith form

introduction

polynomial matrices: reminder and motivation

reductions of most problems to polynomial matrix multiplication

matrix $m \times m$ of degree d $\rightarrow O^{\sim}(m^{\omega} d)$
of “average” degree $\frac{D}{m}$ $\rightarrow O^{\sim}(m^{\omega} \frac{D}{m})$

classical matrix operations

- ▶ multiplication
- ▶ kernel, system solving
- ▶ rank, determinant
- ▶ inversion $O^{\sim}(m^3 d)$

univariate specific operations

- ▶ truncated inverse, QuoRem
- ▶ Hermite-Padé approximation
- ▶ vector rational interpolation
- ▶ syzygies / modular equations

transformation to normal forms

- ▶ triangularization: Hermite form
- ▶ row reduction: Popov form
- ▶ diagonalization: Smith form

introduction

polynomial matrices: reminder and motivation

reductions of most problems to polynomial matrix multiplication

matrix $m \times m$ of degree d $\rightarrow O(m^\omega d)$
of “average” degree $\frac{D}{m}$ $\rightarrow O(m^\omega \frac{D}{m})$

classical matrix operations

- ▶ multiplication
- ▶ kernel, system solving
- ▶ rank, determinant
- ▶ inversion $O(m^3 d)$

univariate specific operations

- ▶ truncated inverse, QuoRem
- ▶ Hermite-Padé approximation
- ▶ vector rational interpolation
- ▶ syzygies / modular equations

transformation to normal forms

- ▶ triangularization: Hermite form
- ▶ row reduction: Popov form
- ▶ diagonalization: Smith form

outline

▶ introduction

- ▶ rational approximation and interpolation
- ▶ the vector case
- ▶ pol. matrices: reminders and motivation

▶ shifted reduced forms

▶ fast algorithms

▶ applications

outline

▶ introduction

- ▶ rational approximation and interpolation
- ▶ the vector case
- ▶ pol. matrices: reminders and motivation

▶ shifted reduced forms

- ▶ reducedness: examples and properties
- ▶ shifted forms and degree constraints
- ▶ stability under multiplication

▶ fast algorithms

▶ applications

shifted reduced forms

reducedness: examples and properties

notation:

let $\mathbf{A} \in \mathbb{K}[X]^{m \times n}$ with no zero row,

define $\mathbf{d} = (d_1, \dots, d_m) = \text{rdeg}(\mathbf{A})$

$$\text{and } \mathbf{X}^{\mathbf{d}} = \begin{bmatrix} X^{d_1} & & \\ & \ddots & \\ & & X^{d_m} \end{bmatrix} \in \mathbb{K}[X]^{m \times m}$$

definition: (row-wise) leading matrix

the **leading matrix of \mathbf{A}** is the unique matrix $\text{lm}(\mathbf{A}) \in \mathbb{K}^{m \times n}$ such that $\mathbf{A} = \mathbf{X}^{\mathbf{d}} \text{lm}(\mathbf{A}) + \mathbf{R}$ with $\text{rdeg}(\mathbf{R}) < \mathbf{d}$ entry-wise

equivalently, $\mathbf{X}^{-\mathbf{d}} \mathbf{A} = \text{lm}(\mathbf{A}) + \text{terms of strictly negative degree}$

shifted reduced forms

reducedness: examples and properties

notation:

let $\mathbf{A} \in \mathbb{K}[X]^{m \times n}$ with no zero row,
define $\mathbf{d} = (d_1, \dots, d_m) = \text{rdeg}(\mathbf{A})$

$$\text{and } \mathbf{X}^{\mathbf{d}} = \begin{bmatrix} X^{d_1} & & \\ & \ddots & \\ & & X^{d_m} \end{bmatrix} \in \mathbb{K}[X]^{m \times m}$$

definition: (row-wise) leading matrix

the **leading matrix of \mathbf{A}** is the unique matrix $\text{lm}(\mathbf{A}) \in \mathbb{K}^{m \times n}$ such that $\mathbf{A} = \mathbf{X}^{\mathbf{d}} \text{lm}(\mathbf{A}) + \mathbf{R}$ with $\text{rdeg}(\mathbf{R}) < \mathbf{d}$ entry-wise

equivalently, $\mathbf{X}^{-\mathbf{d}} \mathbf{A} = \text{lm}(\mathbf{A}) + \text{terms of strictly negative degree}$

definition: (row-wise) reduced matrix

$\mathbf{A} \in \mathbb{K}[X]^{m \times n}$ is said to be **reduced**
if $\text{lm}(\mathbf{A})$ has full row rank

shifted reduced forms

reducedness: examples and properties

consider the following matrices, with $\mathbb{K} = \mathbb{F}_7$:

$$\mathbf{A}_1 = \begin{bmatrix} 3X + 4 & X^3 + 4X + 1 & 4X^2 + 3 \\ 5 & 5X^2 + 3X + 1 & 5X + 3 \end{bmatrix}$$

$$\mathbf{A}_2 = \begin{bmatrix} 3X + 1 & 4X + 3 & 5X + 5 \\ 0 & 4X^2 + 6X & 5 \\ 4X^2 + 5X + 2 & 5 & 6X^2 + 1 \end{bmatrix}$$

$\mathbf{A}_3 = \text{transpose of } \mathbf{A}_1$

$\mathbf{A}_4 = \text{transpose of } \mathbf{A}_2$

answer the following, for $i \in \{1, 2, 3, 4\}$:

1. what is $\text{rdeg}(\mathbf{A}_i)$?
2. what is $\text{Im}(\mathbf{A}_i)$?
3. is \mathbf{A}_i reduced?

polynomial matrices in reduced form

reducedness: examples and properties

let $\mathbf{A} \in \mathbb{K}[X]^{m \times n}$ with $m \leq n$,
the following are equivalent:

- (i) \mathbf{A} is reduced (i.e. $\text{Im}(\mathbf{A})$ has full rank)

polynomial matrices in reduced form

reducedness: examples and properties

let $\mathbf{A} \in \mathbb{K}[X]^{m \times n}$ with $m \leq n$,
the following are equivalent:

(i) \mathbf{A} is reduced (i.e. $\text{Im}(\mathbf{A})$ has full rank)

(ii) for any vector $\mathbf{u} = [\mathbf{u}_1 \ 1 \ \mathbf{u}_2] \in \mathbb{K}[X]^{1 \times m}$ with 1 at index i ,
 $\text{rdeg}(\mathbf{u}\mathbf{A}) \geq \text{rdeg}(\mathbf{A}_{i,*})$

polynomial matrices in reduced form

reducedness: examples and properties

let $\mathbf{A} \in \mathbb{K}[X]^{m \times n}$ with $m \leq n$,
the following are equivalent:

- (i) \mathbf{A} is reduced (i.e. $\text{Im}(\mathbf{A})$ has full rank)
- (ii) for any vector $\mathbf{u} = [\mathbf{u}_1 \ 1 \ \mathbf{u}_2] \in \mathbb{K}[X]^{1 \times m}$ with 1 at index i ,
 $\text{rdeg}(\mathbf{uA}) \geq \text{rdeg}(\mathbf{A}_{i,*})$
- (iii) **predictable degree**: for any vector $\mathbf{u} = [\mathbf{u}_1 \cdots \mathbf{u}_m] \in \mathbb{K}[X]^{1 \times m}$,
 $\text{rdeg}(\mathbf{uA}) = \max_{1 \leq i \leq m} (\deg(\mathbf{u}_i) + \text{rdeg}(\mathbf{A}_{i,*}))$

polynomial matrices in reduced form

reducedness: examples and properties

let $\mathbf{A} \in \mathbb{K}[X]^{m \times n}$ with $m \leq n$,
the following are equivalent:

(i) \mathbf{A} is reduced (i.e. $\text{Im}(\mathbf{A})$ has full rank)

(ii) for any vector $\mathbf{u} = [\mathbf{u}_1 \ 1 \ \mathbf{u}_2] \in \mathbb{K}[X]^{1 \times m}$ with 1 at index i ,
 $\text{rdeg}(\mathbf{uA}) \geq \text{rdeg}(\mathbf{A}_{i,*})$

(iii) **predictable degree**: for any vector $\mathbf{u} = [\mathbf{u}_1 \cdots \mathbf{u}_m] \in \mathbb{K}[X]^{1 \times m}$,
 $\text{rdeg}(\mathbf{uA}) = \max_{1 \leq i \leq m} (\deg(\mathbf{u}_i) + \text{rdeg}(\mathbf{A}_{i,*}))$

(iv) **degree minimality**: $\text{rdeg}(\mathbf{A}) \preceq \text{rdeg}(\mathbf{UA})$ holds for any nonsingular matrix $\mathbf{U} \in \mathbb{K}[X]^{m \times m}$, where \preceq sorts the tuples in nondecreasing order and then uses lexicographic comparison

polynomial matrices in reduced form

reducedness: examples and properties

let $\mathbf{A} \in \mathbb{K}[X]^{m \times n}$ with $m \leq n$,
the following are equivalent:

(i) \mathbf{A} is reduced (i.e. $\text{Im}(\mathbf{A})$ has full rank)

(ii) for any vector $\mathbf{u} = [\mathbf{u}_1 \ 1 \ \mathbf{u}_2] \in \mathbb{K}[X]^{1 \times m}$ with 1 at index i ,
 $\text{rdeg}(\mathbf{uA}) \geq \text{rdeg}(\mathbf{A}_{i,*})$

(iii) **predictable degree**: for any vector $\mathbf{u} = [\mathbf{u}_1 \cdots \mathbf{u}_m] \in \mathbb{K}[X]^{1 \times m}$,
 $\text{rdeg}(\mathbf{uA}) = \max_{1 \leq i \leq m} (\deg(\mathbf{u}_i) + \text{rdeg}(\mathbf{A}_{i,*}))$

(iv) **degree minimality**: $\text{rdeg}(\mathbf{A}) \preceq \text{rdeg}(\mathbf{UA})$ holds for any nonsingular matrix $\mathbf{U} \in \mathbb{K}[X]^{m \times m}$, where \preceq sorts the tuples in nondecreasing order and then uses lexicographic comparison

(v) **predictable determinantal degree**: $\deg \det(\mathbf{A}) = |\text{rdeg}(\mathbf{A})|$
(only when $m = n$)

shifted reduced forms

reducedness: examples and properties

recall the matrix, with $\mathbb{K} = \mathbb{F}_7$,

$$\mathbf{A} = \begin{bmatrix} 3X + 1 & 4X + 3 & 5X + 5 \\ 0 & 4X^2 + 6X & 5 \\ 4X^2 + 5X + 2 & 5 & 6X^2 + 1 \end{bmatrix}$$

1. what is $\deg \det(\mathbf{A})$?
2. what is $\text{rdeg}([4X^2 + 1 \quad 2X \quad 4X + 5] \mathbf{A})$?
3. is it possible to find a matrix

$$\mathbf{P} = \begin{bmatrix} p_{00} & p_{01} & p_{02} \\ p_{10} & p_{11} & p_{12} \end{bmatrix}$$

whose rank is 2, whose degree is 1, and which is a left-multiple of \mathbf{A} ?

shifted reduced forms

reducedness: examples and properties

recall the matrix, with $\mathbb{K} = \mathbb{F}_7$,

$$\mathbf{A} = \begin{bmatrix} 3X + 1 & 4X + 3 & 5X + 5 \\ 0 & 4X^2 + 6X & 5 \\ 4X^2 + 5X + 2 & 5 & 6X^2 + 1 \end{bmatrix}$$

1. what is $\deg \det(\mathbf{A})$?
2. what is $\text{rdeg}([4X^2 + 1 \quad 2X \quad 4X + 5] \mathbf{A})$?
3. is it possible to find a matrix

$$\mathbf{P} = \begin{bmatrix} p_{00} & p_{01} & p_{02} \\ p_{10} & p_{11} & p_{12} \end{bmatrix}$$

whose rank is 2, whose degree is 1, and which is a left-multiple of \mathbf{A} ?

find a row vector \mathbf{u} of degree 1 such that \mathbf{uA} has degree 2, where

$$\mathbf{A} = \begin{bmatrix} 3X + 4 & X^3 + 4X + 1 & 4X^2 + 3 \\ 5 & 5X^2 + 3X + 1 & 5X + 3 \end{bmatrix}$$

shifted reduced forms

shifted forms and degree constraints

keeping our problem in mind:

- ▶ input: f_i 's and α_i 's and degree constraints $d_1, \dots, d_m \in \mathbb{Z}_{>0}$
- ▶ output: a solution \mathbf{p} satisfying the constraints $\text{cdeg}(\mathbf{p}) < (d_1, \dots, d_m)$

obstacle:

computing a reduced basis of solutions ignores the constraints

exercise: suppose we have a reduced basis $\mathbf{P} \in \mathbb{K}[X]^{m \times m}$ of solutions

- ▶ think of particular constraints (d_1, \dots, d_m) that can be handled via \mathbf{P}
- ▶ give constraints (d_1, \dots, d_m) for which \mathbf{P} is “typically” not satisfactory

shifted reduced forms

shifted forms and degree constraints

keeping our problem in mind:

- ▶ input: f_i 's and α_i 's and degree constraints $d_1, \dots, d_m \in \mathbb{Z}_{>0}$
- ▶ output: a solution \mathbf{p} satisfying the constraints $\text{cdeg}(\mathbf{p}) < (d_1, \dots, d_m)$

obstacle:

computing a reduced basis of solutions ignores the constraints

exercise: suppose we have a reduced basis $\mathbf{P} \in \mathbb{K}[X]^{m \times m}$ of solutions

- ▶ think of particular constraints (d_1, \dots, d_m) that can be handled via \mathbf{P}
- ▶ give constraints (d_1, \dots, d_m) for which \mathbf{P} is “typically” not satisfactory

solution: compute \mathbf{P} in **shifted** reduced form

shifted reduced forms

shifted forms and degree constraints

$$\mathbf{A} = \begin{bmatrix} 3X + 4 & X^3 + 4X + 1 & 4X^2 + 3 \\ 5 & 5X^2 + 3X + 1 & 5X + 3 \\ 3X^3 + X^2 + 5X + 3 & 6X + 5 & 2X + 1 \end{bmatrix}$$

using [elementary row operations](#), transform \mathbf{A} into...

$$\text{Hermite form } \mathbf{H} = \begin{bmatrix} X^6 + 6X^4 + X^3 + X + 4 & 0 & 0 \\ 5X^5 + 5X^4 + 6X^3 + 2X^2 + 6X + 3 & X & 0 \\ 3X^4 + 5X^3 + 4X^2 + 6X + 1 & 5 & 1 \end{bmatrix}$$

$$\text{Popov form } \mathbf{P} = \begin{bmatrix} X^3 + 5X^2 + 4X + 1 & 2X + 4 & 3X + 5 \\ 1 & X^2 + 2X + 3 & X + 2 \\ 3X + 2 & 4X & X^2 \end{bmatrix}$$

shifted reduced forms

shifted forms and degree constraints

nonsingular $\mathbf{A} \in \mathbb{K}[X]^{m \times m}$

elementary row transformations

Hermite form [Hermite, 1851]

- ▶ triangular
- ▶ column normalized

$$\begin{bmatrix} 16 & & & \\ 15 & 0 & & \\ 15 & & 0 & \\ 15 & & & 0 \end{bmatrix} \quad \begin{bmatrix} 4 & & & \\ 3 & 7 & & \\ 1 & 5 & 3 & \\ 3 & 6 & 1 & 2 \end{bmatrix}$$

shifted reduced forms

shifted forms and degree constraints

nonsingular $\mathbf{A} \in \mathbb{K}[X]^{m \times m}$

elementary row transformations

Hermite form [Hermite, 1851]

- ▶ triangular
- ▶ column normalized

$$\begin{bmatrix} 16 & & & \\ 15 & 0 & & \\ 15 & & 0 & \\ 15 & & & 0 \end{bmatrix} \quad \begin{bmatrix} 4 & & & \\ 3 & 7 & & \\ 1 & 5 & 3 & \\ 3 & 6 & 1 & 2 \end{bmatrix}$$

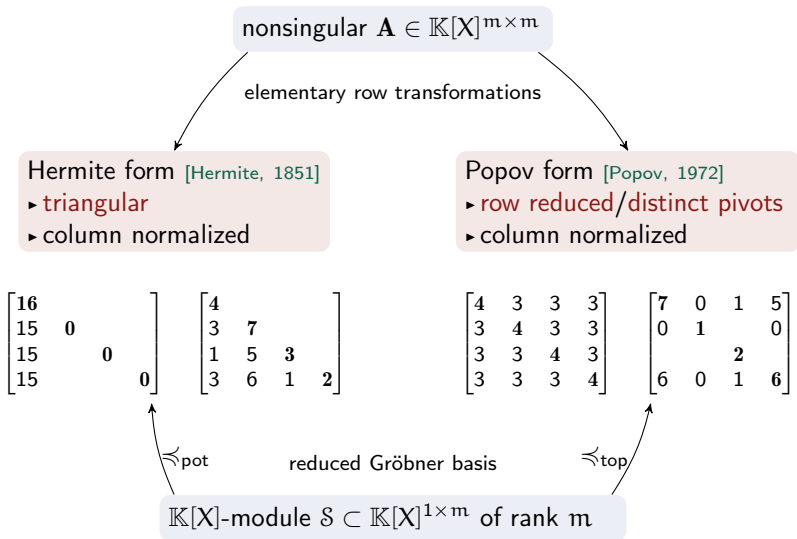
Popov form [Popov, 1972]

- ▶ row reduced/distinct pivots
- ▶ column normalized

$$\begin{bmatrix} 4 & 3 & 3 & 3 \\ 3 & 4 & 3 & 3 \\ 3 & 3 & 4 & 3 \\ 3 & 3 & 3 & 4 \end{bmatrix} \quad \begin{bmatrix} 7 & 0 & 1 & 5 \\ 0 & 1 & & 0 \\ & & 2 & \\ 6 & 0 & 1 & 6 \end{bmatrix}$$

shifted reduced forms

shifted forms and degree constraints



shifted reduced forms

shifted forms and degree constraints

nonsingular $\mathbf{A} \in \mathbb{K}[X]^{m \times m}$

elementary row transformations

Hermite form [Hermite, 1851]

- ▶ triangular
- ▶ column normalized

$$\begin{bmatrix} 16 & & & \\ 15 & 0 & & \\ 15 & & 0 & \\ 15 & & & 0 \end{bmatrix} \quad \begin{bmatrix} 4 & & & \\ 3 & 7 & & \\ 1 & 5 & 3 & \\ 3 & 6 & 1 & 2 \end{bmatrix}$$

Popov form [Popov, 1972]

- ▶ row reduced/distinct pivots
- ▶ column normalized

$$\begin{bmatrix} 4 & 3 & 3 & 3 \\ 3 & 4 & 3 & 3 \\ 3 & 3 & 4 & 3 \\ 3 & 3 & 3 & 4 \end{bmatrix} \quad \begin{bmatrix} 7 & 0 & 1 & 5 \\ 0 & 1 & & 0 \\ & & 2 & \\ 6 & 0 & 1 & 6 \end{bmatrix}$$

invariant: $D = \deg(\det(\mathbf{A})) = 4 + 7 + 3 + 2 = 7 + 1 + 2 + 6$

- ▶ average column degree is $\frac{D}{m}$
- ▶ size of object is $mD + m^2 = m^2(\frac{D}{m} + 1)$

shifted reduced forms

shifted forms and degree constraints

nonsingular $\mathbf{A} \in \mathbb{K}[X]^{m \times m}$

elementary row transformations

Hermite form [Hermite, 1851]

- ▶ triangular
- ▶ column normalized

Popov form [Popov, 1972]

- ▶ row reduced/distinct pivots
- ▶ column normalized

$$\begin{bmatrix} 16 & & & \\ 15 & 0 & & \\ 15 & & 0 & \\ 15 & & & 0 \end{bmatrix} \quad \begin{bmatrix} 4 & & & \\ 3 & 7 & & \\ 1 & 5 & 3 & \\ 3 & 6 & 1 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 3 & 3 & 3 \\ 3 & 4 & 3 & 3 \\ 3 & 3 & 4 & 3 \\ 3 & 3 & 3 & 4 \end{bmatrix} \quad \begin{bmatrix} 7 & 0 & 1 & 5 \\ 0 & 1 & & 0 \\ & & 2 & \\ 6 & 0 & 1 & 6 \end{bmatrix}$$

[Beckermann-Labahn-Villard, 1999; Mulders-Storjohann, 2003]

shifted reduced form:
arbitrary degree constraints + no column normalization

\approx minimal, non-reduced, \prec -Gröbner basis

shifted reduced forms

shift: integer tuple $\mathbf{s} = (s_1, \dots, s_m)$ acting as **column weights**

→ connects Popov and Hermite forms

$\mathbf{s} = (0, 0, 0, 0)$ Popov	$\begin{bmatrix} 4 & 3 & 3 & 3 \\ 3 & 4 & 3 & 3 \\ 3 & 3 & 4 & 3 \\ 3 & 3 & 3 & 4 \end{bmatrix}$	$\begin{bmatrix} 7 & 0 & 1 & 5 \\ 0 & 1 & & 0 \\ & & 2 & \\ 6 & 0 & 1 & 6 \end{bmatrix}$
$\mathbf{s} = (0, 2, 4, 6)$ s-Popov	$\begin{bmatrix} 7 & 4 & 2 & 0 \\ 6 & 5 & 2 & 0 \\ 6 & 4 & 3 & 0 \\ 6 & 4 & 2 & 1 \end{bmatrix}$	$\begin{bmatrix} 8 & 5 & 1 & \\ 7 & 6 & 1 & \\ & & 2 & \\ 0 & 1 & & 0 \end{bmatrix}$
$\mathbf{s} = (0, D, 2D, 3D)$ Hermite	$\begin{bmatrix} 16 & & & \\ 15 & 0 & & \\ 15 & & 0 & \\ 15 & & & 0 \end{bmatrix}$	$\begin{bmatrix} 4 & & & \\ 3 & 7 & & \\ 1 & 5 & 3 & \\ 3 & 6 & 1 & 2 \end{bmatrix}$

- ▶ **normal** form, **average** column degree D/m
- ▶ shifted reduced form: same without normalization
- ▶ shifts arise naturally in algorithms (approximants, kernel, ...)

shifted reduced forms

shifted forms and degree constraints

shifted row degree of a polynomial matrix
= the list of the maximum **shifted** degree in each of its rows

for $\mathbf{A} = (a_{i,j}) \in \mathbb{K}[X]^{m \times n}$, and $\mathbf{s} = (s_1, \dots, s_n) \in \mathbb{Z}^n$,

$$\begin{aligned} \text{rdeg}_{\mathbf{s}}(\mathbf{A}) &= (\text{rdeg}_{\mathbf{s}}(\mathbf{A}_{1,*}), \dots, \text{rdeg}_{\mathbf{s}}(\mathbf{A}_{m,*})) \\ &= \left(\max_{1 \leq j \leq n} (\deg(\mathbf{A}_{1,j}) + s_j), \dots, \max_{1 \leq j \leq n} (\deg(\mathbf{A}_{m,j}) + s_j) \right) \in \mathbb{Z}^m \end{aligned}$$

example: for the matrix $\mathbf{A} = \begin{bmatrix} 3X + 4 & X^3 + 4X + 1 & 4X^2 + 3 \\ 5 & 5X^2 + 3X + 1 & 5X + 3 \end{bmatrix}$,
describe $\text{rdeg}_{(0,0,0)}(\mathbf{A})$, $\text{rdeg}_{(0,1,2)}(\mathbf{A})$, and $\text{rdeg}_{(-1,-3,-2)}(\mathbf{A})$

shifted reduced forms

shifted forms and degree constraints

shifted row degree of a polynomial matrix
= the list of the maximum **shifted** degree in each of its rows

for $\mathbf{A} = (a_{i,j}) \in \mathbb{K}[X]^{m \times n}$, and $\mathbf{s} = (s_1, \dots, s_n) \in \mathbb{Z}^n$,

$$\begin{aligned} \text{rdeg}_{\mathbf{s}}(\mathbf{A}) &= (\text{rdeg}_{\mathbf{s}}(\mathbf{A}_{1,*}), \dots, \text{rdeg}_{\mathbf{s}}(\mathbf{A}_{m,*})) \\ &= \left(\max_{1 \leq j \leq n} (\deg(\mathbf{A}_{1,j}) + s_j), \dots, \max_{1 \leq j \leq n} (\deg(\mathbf{A}_{m,j}) + s_j) \right) \in \mathbb{Z}^m \end{aligned}$$

example: for the matrix $\mathbf{A} = \begin{bmatrix} 3X + 4 & X^3 + 4X + 1 & 4X^2 + 3 \\ 5 & 5X^2 + 3X + 1 & 5X + 3 \end{bmatrix}$,
describe $\text{rdeg}_{(0,0,0)}(\mathbf{A})$, $\text{rdeg}_{(0,1,2)}(\mathbf{A})$, and $\text{rdeg}_{(-1,-3,-2)}(\mathbf{A})$

- ▶ $\text{rdeg}_{\mathbf{s}}(\mathbf{A}) = \text{rdeg}(\mathbf{A}\mathbf{X}^{\mathbf{s}})$
- ▶ $\text{rdeg}_{\mathbf{s}}(\mathbf{A})$ only depends on \mathbf{s} and the degrees in \mathbf{A}
- ▶ $\text{rdeg}_{\mathbf{s}+(c,\dots,c)}(\mathbf{A}) = \text{rdeg}_{\mathbf{s}}(\mathbf{A}) + c$

shifted reduced forms

shifted forms and degree constraints

notation:

let $\mathbf{A} \in \mathbb{K}[X]^{m \times n}$ with no zero row, and $\mathbf{s} \in \mathbb{Z}^n$,
define $\mathbf{d} = (d_1, \dots, d_m) = \text{rdeg}_s(\mathbf{A})$

$$\text{and } \mathbf{X}^{\mathbf{d}} = \begin{bmatrix} X^{d_1} & & \\ & \ddots & \\ & & X^{d_m} \end{bmatrix} \in \mathbb{K}[X, X^{-1}]^{m \times m}$$

definition: s -leading matrix / s -reduced matrix

assuming $s \geq 0$,

- ▶ the s -leading matrix of \mathbf{A} is $\text{lm}_s(\mathbf{A}) = \text{lm}(\mathbf{A}\mathbf{X}^{\mathbf{s}}) \in \mathbb{K}^{m \times n}$
- ▶ $\mathbf{A} \in \mathbb{K}[X]^{m \times n}$ is s -reduced if $\text{lm}_s(\mathbf{A})$ has full row rank

shifted reduced forms

shifted forms and degree constraints

notation:

let $\mathbf{A} \in \mathbb{K}[X]^{m \times n}$ with no zero row, and $\mathbf{s} \in \mathbb{Z}^n$,
define $\mathbf{d} = (d_1, \dots, d_m) = \text{rdeg}_s(\mathbf{A})$

$$\text{and } \mathbf{X}^{\mathbf{d}} = \begin{bmatrix} X^{d_1} & & \\ & \ddots & \\ & & X^{d_m} \end{bmatrix} \in \mathbb{K}[X, X^{-1}]^{m \times m}$$

definition: s -leading matrix / s -reduced matrix

assuming $\mathbf{s} \geq 0$,

- ▶ the s -leading matrix of \mathbf{A} is $\text{lm}_s(\mathbf{A}) = \text{lm}(\mathbf{A}\mathbf{X}^{\mathbf{s}}) \in \mathbb{K}^{m \times n}$
- ▶ $\mathbf{A} \in \mathbb{K}[X]^{m \times n}$ is s -reduced if $\text{lm}_s(\mathbf{A})$ has full row rank

- ▶ these notions are invariant under $\mathbf{s} \rightarrow \mathbf{s} + (c, \dots, c)$
- ▶ they coincide with the non-shifted case when $\mathbf{s} = (0, \dots, 0)$
- ▶ $\mathbf{X}^{-\mathbf{d}}\mathbf{A}\mathbf{X}^{\mathbf{s}} = \text{lm}_s(\mathbf{A}) + \text{terms of strictly negative degree}$

shifted reduced forms

shifted forms and degree constraints

exercise: for each of the matrices below, and each shift \mathbf{s} ,

1. give the \mathbf{s} -leading matrix
2. deduce whether the matrix is \mathbf{s} -reduced

$$\mathbf{A} = \begin{bmatrix} 3X + 4 & X^3 + 4X + 1 & 4X^2 + 3 \\ 5 & 5X^2 + 3X + 1 & 5X + 3 \\ 3X^3 + X^2 + 5X + 3 & 6X + 5 & 2X + 1 \end{bmatrix}$$

$$\mathbf{H} = \begin{bmatrix} X^6 + 6X^4 + X^3 + X + 4 & 0 & 0 \\ 5X^5 + 5X^4 + 6X^3 + 2X^2 + 6X + 3 & X & 0 \\ 3X^4 + 5X^3 + 4X^2 + 6X + 1 & 5 & 1 \end{bmatrix}$$

$$\mathbf{P} = \begin{bmatrix} X^3 + 5X^2 + 4X + 1 & 2X + 4 & 3X + 5 \\ 1 & X^2 + 2X + 3 & X + 2 \\ 3X + 2 & 4X & X^2 \end{bmatrix}$$

$$\mathbf{s} = (0, 0, 0), \mathbf{s} = (0, 5, 6), \mathbf{s} = (-3, -2, -2)$$

shifted reduced forms

shifted forms and degree constraints

the characterizations generalize to the s -shifted case,
using s -row degrees and s -leading matrices where appropriate

(proofs: direct, with: \mathbf{A} is s -reduced $\Leftrightarrow \mathbf{A}\mathbf{X}^s$ is reduced)

for example recall the [predictable degree property](#):

\mathbf{A} is reduced if and only if for any $\mathbf{u} = [u_1 \cdots u_m] \in \mathbb{K}[X]^{1 \times m}$,
$$\text{rdeg}(\mathbf{u}\mathbf{A}) = \max_{1 \leq i \leq m} (\deg(u_i) + \text{rdeg}(\mathbf{A}_{i,*}))$$

shifted reduced forms

shifted forms and degree constraints

the characterizations generalize to the s -shifted case,
using s -row degrees and s -leading matrices where appropriate

(proofs: direct, with: \mathbf{A} is s -reduced $\Leftrightarrow \mathbf{A}\mathbf{X}^s$ is reduced)

for example recall the [predictable degree property](#):

\mathbf{A} is reduced if and only if for any $\mathbf{u} = [u_1 \cdots u_m] \in \mathbb{K}[X]^{1 \times m}$,
$$\text{rdeg}(\mathbf{uA}) = \max_{1 \leq i \leq m} (\deg(u_i) + \text{rdeg}(\mathbf{A}_{i,*}))$$

- ▶ this means $\text{rdeg}(\mathbf{uA}) = \text{rdeg}_{\mathbf{t}}(\mathbf{u})$ where $\mathbf{t} = \text{rdeg}(\mathbf{A})$
- ▶ i.e. $\text{rdeg}(\mathbf{uA}) = \text{rdeg}(\mathbf{uX}^{\text{rdeg}(\mathbf{A})})$, “no surprising cancellation”
- ▶ proof: let $\delta = \text{rdeg}_{\mathbf{t}}(\mathbf{u})$, our goal is to show $\text{rdeg}(\mathbf{uA}) = \delta$
terms of $X^{-\delta}\mathbf{uA}$ have degree ≤ 0 ,
and $X^{-\delta}\mathbf{uA} = (X^{-\delta}\mathbf{uX}^{\mathbf{t}})(\mathbf{X}^{-\mathbf{t}}\mathbf{A})$;
the term of degree 0 is $\text{lm}_{\mathbf{t}}(\mathbf{u})\text{lm}(\mathbf{A})$,
it is nonzero since $\text{lm}(\mathbf{A})$ has full rank and $\text{lm}_{\mathbf{t}}(\mathbf{u}) \neq 0$
(the case $\mathbf{u} = \mathbf{0}$ is trivial)

shifted reduced forms

shifted forms and degree constraints

the characterizations generalize to the \mathbf{s} -shifted case,
using \mathbf{s} -row degrees and \mathbf{s} -leading matrices where appropriate

(proofs: direct, with: \mathbf{A} is \mathbf{s} -reduced $\Leftrightarrow \mathbf{A}\mathbf{X}^{\mathbf{s}}$ is reduced)

for example recall the [predictable degree property](#):

\mathbf{A} is reduced if and only if for any $\mathbf{u} = [u_1 \cdots u_m] \in \mathbb{K}[X]^{1 \times m}$,
$$\text{rdeg}(\mathbf{uA}) = \max_{1 \leq i \leq m} (\deg(u_i) + \text{rdeg}(\mathbf{A}_{i,*}))$$

\mathbf{A} is \mathbf{s} -reduced if and only if for any $\mathbf{u} = [u_1 \cdots u_m] \in \mathbb{K}[X]^{1 \times m}$,
$$\text{rdeg}_{\mathbf{s}}(\mathbf{uA}) = \max_{1 \leq i \leq m} (\deg(u_i) + \text{rdeg}_{\mathbf{s}}(\mathbf{A}_{i,*}))$$

this means $\text{rdeg}_{\mathbf{s}}(\mathbf{uA}) = \text{rdeg}_{\mathbf{t}}(\mathbf{u})$, where $\mathbf{t} = \text{rdeg}_{\mathbf{s}}(\mathbf{A})$

shifted reduced forms

shifted forms and degree constraints

the characterizations generalize to the **s**-shifted case,
using **s**-row degrees and **s**-leading matrices where appropriate

(proofs: direct, with: \mathbf{A} is **s**-reduced $\Leftrightarrow \mathbf{A}\mathbf{X}^{\mathbf{s}}$ is reduced)

for example recall the **predictable degree property**:

\mathbf{A} is reduced if and only if for any $\mathbf{u} = [u_1 \cdots u_m] \in \mathbb{K}[X]^{1 \times m}$,
$$\text{rdeg}(\mathbf{u}\mathbf{A}) = \max_{1 \leq i \leq m} (\deg(u_i) + \text{rdeg}(\mathbf{A}_{i,*}))$$

\mathbf{A} is **s**-reduced if and only if for any $\mathbf{u} = [u_1 \cdots u_m] \in \mathbb{K}[X]^{1 \times m}$,
$$\text{rdeg}_{\mathbf{s}}(\mathbf{u}\mathbf{A}) = \max_{1 \leq i \leq m} (\deg(u_i) + \text{rdeg}_{\mathbf{s}}(\mathbf{A}_{i,*}))$$

this means $\text{rdeg}_{\mathbf{s}}(\mathbf{u}\mathbf{A}) = \text{rdeg}_{\mathbf{t}}(\mathbf{u})$, where $\mathbf{t} = \text{rdeg}_{\mathbf{s}}(\mathbf{A})$

- ▶ **s**-reduced forms provide vectors of **minimal s-degree** in the module
- ▶ satisfying **degree constraints** $(d_1, \dots, d_m) \Rightarrow$ taking $\mathbf{s} = (-d_1, \dots, -d_m)$
- ▶ indeed $\text{cdeg}([p_1 \cdots p_m]) < (d_1, \dots, d_m)$
if and only if $\text{rdeg}_{(-d_1, \dots, -d_m)}([p_1 \cdots p_m]) < 0$

shifted reduced forms

stability under multiplication

algorithms based on polynomial matrix multiplication

[iterative: van Barel-Bultheel 1991, Beckermann-Labahn 2000]

[divide and conquer: Beckermann-Labahn 1994, Giorgi-Jeannerod-Villard 2003]

- ▶ compute a first basis \mathbf{P}_1 for a subproblem
- ▶ update the input instance to get the second subproblem
- ▶ compute a second basis \mathbf{P}_2 for this second subproblem
- ▶ the output basis of solutions is $\mathbf{P}_2\mathbf{P}_1$

we want $\mathbf{P}_2\mathbf{P}_1$ to be reduced:

1. is it implied by " \mathbf{P}_1 reduced and \mathbf{P}_2 reduced"?
2. any idea of how to fix this?

shifted reduced forms

stability under multiplication

algorithms based on polynomial matrix multiplication

[iterative: van Barel-Bultheel 1991, Beckermann-Labahn 2000]

[divide and conquer: Beckermann-Labahn 1994, Giorgi-Jeannerod-Villard 2003]

- ▶ compute a first basis \mathbf{P}_1 for a subproblem
- ▶ update the input instance to get the second subproblem
- ▶ compute a second basis \mathbf{P}_2 for this second subproblem
- ▶ the output basis of solutions is $\mathbf{P}_2\mathbf{P}_1$

we want $\mathbf{P}_2\mathbf{P}_1$ to be reduced:

1. is it implied by “ \mathbf{P}_1 reduced and \mathbf{P}_2 reduced”?
2. any idea of how to fix this?

we want $\mathbf{P}_2\mathbf{P}_1$ to be reduced

theorem: implied by “ \mathbf{P}_1 is reduced and \mathbf{P}_2 is \mathbf{t} -reduced”

where $\mathbf{t} = \text{rdeg}(\mathbf{P}_1)$

shifted reduced forms

stability under multiplication

algorithms based on polynomial matrix multiplication

[iterative: van Barel-Bultheel 1991, Beckermann-Labahn 2000]

[divide and conquer: Beckermann-Labahn 1994, Giorgi-Jeannerod-Villard 2003]

- ▶ compute a first basis \mathbf{P}_1 for a subproblem
- ▶ update the input instance to get the second subproblem
- ▶ compute a second basis \mathbf{P}_2 for this second subproblem
- ▶ the output basis of solutions is $\mathbf{P}_2\mathbf{P}_1$

we want $\mathbf{P}_2\mathbf{P}_1$ to be reduced:

1. is it implied by “ \mathbf{P}_1 reduced and \mathbf{P}_2 reduced”?
2. any idea of how to fix this?

we want $\mathbf{P}_2\mathbf{P}_1$ to be **s**-reduced

theorem: implied by “ \mathbf{P}_1 is **s**-reduced and \mathbf{P}_2 is **t**-reduced”

where $\mathbf{t} = \text{rdeg}_{\mathbf{s}}(\mathbf{P}_1)$

shifted reduced forms

stability under multiplication

let $\mathcal{M} \subseteq \mathcal{M}_1$ be two $\mathbb{K}[X]$ -submodules of $\mathbb{K}[X]^m$ of rank m ,
let $\mathbf{P}_1 \in \mathbb{K}[X]^{m \times m}$ be a basis of \mathcal{M}_1 ,
let $\mathbf{s} \in \mathbb{Z}^m$ and $\mathbf{t} = \text{rdeg}_s(\mathbf{P}_1)$,

- ▶ the rank of the module $\mathcal{M}_2 = \{\boldsymbol{\lambda} \in \mathbb{K}[X]^{1 \times m} \mid \boldsymbol{\lambda} \mathbf{P}_1 \in \mathcal{M}\}$ is m
and for any basis $\mathbf{P}_2 \in \mathbb{K}[X]^{m \times m}$ of \mathcal{M}_2 ,
the product $\mathbf{P}_2 \mathbf{P}_1$ is a basis of \mathcal{M}
- ▶ if \mathbf{P}_1 is \mathbf{s} -reduced and \mathbf{P}_2 is \mathbf{t} -reduced,
then $\mathbf{P}_2 \mathbf{P}_1$ is \mathbf{s} -reduced

shifted reduced forms

stability under multiplication

let $\mathcal{M} \subseteq \mathcal{M}_1$ be two $\mathbb{K}[X]$ -submodules of $\mathbb{K}[X]^m$ of rank m ,
let $\mathbf{P}_1 \in \mathbb{K}[X]^{m \times m}$ be a basis of \mathcal{M}_1 ,
let $\mathbf{s} \in \mathbb{Z}^m$ and $\mathbf{t} = \text{rdeg}_s(\mathbf{P}_1)$,
► the rank of the module $\mathcal{M}_2 = \{\lambda \in \mathbb{K}[X]^{1 \times m} \mid \lambda \mathbf{P}_1 \in \mathcal{M}\}$ is m
and for any basis $\mathbf{P}_2 \in \mathbb{K}[X]^{m \times m}$ of \mathcal{M}_2 ,
the product $\mathbf{P}_2 \mathbf{P}_1$ is a basis of \mathcal{M}
► if \mathbf{P}_1 is \mathbf{s} -reduced and \mathbf{P}_2 is \mathbf{t} -reduced,
then $\mathbf{P}_2 \mathbf{P}_1$ is \mathbf{s} -reduced

Let $\mathbf{A} \in \mathbb{K}[X]^{m \times m}$ denote the adjugate of \mathbf{P}_1 . Then, we have $\mathbf{A} \mathbf{P}_1 = \det(\mathbf{P}_1) \mathbf{I}_m$.
Thus, $\mathbf{p} \mathbf{A} \mathbf{P}_1 = \det(\mathbf{P}_1) \mathbf{p} \in \mathcal{M}$ for all $\mathbf{p} \in \mathcal{M}$, and therefore $\mathcal{M} \mathbf{A} \subseteq \mathcal{M}_2$. Now,
the nonsingularity of \mathbf{A} ensures that $\mathcal{M} \mathbf{A}$ has rank m ; this implies that \mathcal{M}_2 has
rank m as well (see e.g. [Dummit-Foote 2004, Sec. 12.1, Thm. 4]). The matrix $\mathbf{P}_2 \mathbf{P}_1$
is nonsingular since $\det(\mathbf{P}_2 \mathbf{P}_1) \neq 0$. Now let $\mathbf{p} \in \mathcal{M}$; we want to prove that \mathbf{p}
is a $\mathbb{K}[X]$ -linear combination of the rows of $\mathbf{P}_2 \mathbf{P}_1$. First, $\mathbf{p} \in \mathcal{M}_1$, so there exists
 $\lambda \in \mathbb{K}[X]^{1 \times m}$ such that $\mathbf{p} = \lambda \mathbf{P}_1$. But then $\lambda \in \mathcal{M}_2$, and thus there exists $\mu \in$
 $\mathbb{K}[X]^{1 \times m}$ such that $\lambda = \mu \mathbf{P}_2$. This yields the combination $\mathbf{p} = \mu \mathbf{P}_2 \mathbf{P}_1$.

shifted reduced forms

stability under multiplication

let $\mathcal{M} \subseteq \mathcal{M}_1$ be two $\mathbb{K}[X]$ -submodules of $\mathbb{K}[X]^m$ of rank m ,
let $\mathbf{P}_1 \in \mathbb{K}[X]^{m \times m}$ be a basis of \mathcal{M}_1 ,
let $\mathbf{s} \in \mathbb{Z}^m$ and $\mathbf{t} = \text{rdeg}_s(\mathbf{P}_1)$,
► the rank of the module $\mathcal{M}_2 = \{\lambda \in \mathbb{K}[X]^{1 \times m} \mid \lambda \mathbf{P}_1 \in \mathcal{M}\}$ is m
and for any basis $\mathbf{P}_2 \in \mathbb{K}[X]^{m \times m}$ of \mathcal{M}_2 ,
the product $\mathbf{P}_2 \mathbf{P}_1$ is a basis of \mathcal{M}
► if \mathbf{P}_1 is \mathbf{s} -reduced and \mathbf{P}_2 is \mathbf{t} -reduced,
then $\mathbf{P}_2 \mathbf{P}_1$ is \mathbf{s} -reduced

Let $\mathbf{d} = \text{rdeg}_t(\mathbf{P}_2)$; we have $\mathbf{d} = \text{rdeg}_s(\mathbf{P}_2 \mathbf{P}_1)$ by the predictable degree property. Using $\mathbf{X}^{-\mathbf{d}} \mathbf{P}_2 \mathbf{P}_1 \mathbf{X}^{\mathbf{s}} = \mathbf{X}^{-\mathbf{d}} \mathbf{P}_2 \mathbf{X}^{\mathbf{t}} \mathbf{X}^{-\mathbf{t}} \mathbf{P}_1 \mathbf{X}^{\mathbf{s}}$, we obtain that $\text{Im}_s(\mathbf{P}_2 \mathbf{P}_1) = \text{Im}_t(\mathbf{P}_2) \text{Im}_s(\mathbf{P}_1)$. By assumption, $\text{Im}_t(\mathbf{P}_2)$ and $\text{Im}_s(\mathbf{P}_1)$ are invertible, and therefore $\text{Im}_s(\mathbf{P}_2 \mathbf{P}_1)$ is invertible as well; thus $\mathbf{P}_2 \mathbf{P}_1$ is \mathbf{s} -reduced.

outline

▶ introduction

- ▶ rational approximation and interpolation
- ▶ the vector case
- ▶ pol. matrices: reminders and motivation

▶ shifted reduced forms

- ▶ reducedness: examples and properties
- ▶ shifted forms and degree constraints
- ▶ stability under multiplication

▶ fast algorithms

▶ applications

outline

▶ introduction

- ▶ rational approximation and interpolation
- ▶ the vector case
- ▶ pol. matrices: reminders and motivation

▶ shifted reduced forms

- ▶ reducedness: examples and properties
- ▶ shifted forms and degree constraints
- ▶ stability under multiplication

▶ fast algorithms

- ▶ iterative algorithm and output size
- ▶ base case: modulus of degree 1
- ▶ recursion: residual and basis multiplication

▶ applications

fast algorithms

iterative algorithm [van Barel-Bultheel / Beckermann-Labahn]

input: vector $\mathbf{F} = \begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix}$, points $\alpha_1, \dots, \alpha_d \in \mathbb{K}$, shift $\mathbf{s} = (s_1, \dots, s_m) \in \mathbb{Z}^m$

1. $\mathbf{P} = \begin{bmatrix} -\mathbf{p}_1 \\ \vdots \\ -\mathbf{p}_m \end{bmatrix}$ = identity matrix in $\mathbb{K}[X]^{m \times m}$

2. for i from 1 to d :

a. evaluate updated vector $\begin{bmatrix} (\mathbf{p}_1 \cdot \mathbf{F})(\alpha_i) \\ \vdots \\ (\mathbf{p}_m \cdot \mathbf{F})(\alpha_i) \end{bmatrix} = (\mathbf{P} \cdot \mathbf{F})(\alpha_i)$

b. choose pivot π with smallest s_π such that $(\mathbf{p}_\pi \cdot \mathbf{F})(\alpha_i) \neq 0$
update pivot shift $s_\pi = s_\pi + 1$

c. eliminate: /* after this, $\forall j \neq \pi, (\mathbf{p}_j \cdot \mathbf{F})(\alpha_i) = 0$ */
for $j \neq \pi$ do $\mathbf{p}_j \leftarrow \mathbf{p}_j - \frac{(\mathbf{p}_j \cdot \mathbf{F})(\alpha_i)}{(\mathbf{p}_\pi \cdot \mathbf{F})(\alpha_i)} \mathbf{p}_\pi$; $\mathbf{p}_\pi \leftarrow (X - \alpha_i) \mathbf{p}_\pi$

after i iterations: \mathbf{P} is an \mathbf{s} -reduced basis of solutions for $(\alpha_1, \dots, \alpha_i)$

fast algorithms

iterative algorithm [van Barel-Bultheel / Beckermann-Labahn]

parameters: $d = 8$ $m = 4$ $s = (0, 2, 4, 6)$, base field \mathbb{F}_{97}

input: $(24, 31, 15, 32, 83, 27, 20, 59)$ and $\mathbf{F} = [1 \ L \ L^2 \ L^3]^\top$

iteration: $i = 1$ point: 24, 31, 15, 32, 83, 27, 20, 59

shift

$[0 \ 2 \ 4 \ 6]$

basis

$$\begin{bmatrix} 1 & & & & 0 & & 0 & 0 \\ 0 & & & & 1 & & 0 & 0 \\ 0 & & & & 0 & & 1 & 0 \\ 0 & & & & 0 & & 0 & 1 \end{bmatrix}$$

values

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 80 & 73 & 73 & 35 & 66 & 46 & 91 & 64 \\ 95 & 91 & 91 & 61 & 88 & 79 & 36 & 22 \\ 34 & 47 & 47 & 1 & 85 & 45 & 75 & 50 \end{bmatrix}$$

fast algorithms

iterative algorithm [van Barel-Bultheel / Beckermann-Labahn]

parameters: $d = 8$ $m = 4$ $s = (0, 2, 4, 6)$, base field \mathbb{F}_{97}

input: (24, 31, 15, 32, 83, 27, 20, 59) and $\mathbf{F} = [1 \quad \mathbf{L} \quad \mathbf{L}^2 \quad \mathbf{L}^3]^T$

iteration: $i = 1$ point: 24, 31, 15, 32, 83, 27, 20, 59

shift

[0 2 4 6]

basis
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

values

1	1	1	1	1	1	1	1
80	73	73	35	66	46	91	64
95	91	91	61	88	79	36	22
34	47	47	1	85	45	75	50

fast algorithms

iterative algorithm [van Barel-Bultheel / Beckermann-Labahn]

parameters: $d = 8$ $m = 4$ $s = (0, 2, 4, 6)$, base field \mathbb{F}_{97}

input: $(24, 31, 15, 32, 83, 27, 20, 59)$ and $\mathbf{F} = [1 \ L \ L^2 \ L^3]^\top$

iteration: $i = 1$ point: 24, 31, 15, 32, 83, 27, 20, 59

shift

[0 2 4 6]

basis

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 17 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 63 & 0 & 0 & 1 \end{bmatrix}$$

values

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 90 & 90 & 52 & 83 & 63 & 11 & 81 \\ 0 & 93 & 93 & 63 & 90 & 81 & 38 & 24 \\ 0 & 13 & 13 & 64 & 51 & 11 & 41 & 16 \end{bmatrix}$$

fast algorithms

iterative algorithm [van Barel-Bultheel / Beckermann-Labahn]

parameters: $d = 8$ $m = 4$ $s = (0, 2, 4, 6)$, base field \mathbb{F}_{97}

input: $(24, 31, 15, 32, 83, 27, 20, 59)$ and $\mathbf{F} = [1 \ L \ L^2 \ L^3]^\top$

iteration: $i = 1$ point: 24, 31, 15, 32, 83, 27, 20, 59

shift

[1 2 4 6]

basis

$$\begin{bmatrix} X + 73 & 0 & 0 & 0 \\ 17 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 63 & 0 & 0 & 1 \end{bmatrix}$$

values

$$\begin{bmatrix} 0 & 7 & 88 & 8 & 59 & 3 & 93 & 35 \\ 0 & 90 & 90 & 52 & 83 & 63 & 11 & 81 \\ 0 & 93 & 93 & 63 & 90 & 81 & 38 & 24 \\ 0 & 13 & 13 & 64 & 51 & 11 & 41 & 16 \end{bmatrix}$$

fast algorithms

iterative algorithm [van Barel-Bultheel / Beckermann-Labahn]

parameters: $d = 8$ $m = 4$ $s = (0, 2, 4, 6)$, base field \mathbb{F}_{97}

input: $(24, 31, 15, 32, 83, 27, 20, 59)$ and $\mathbf{F} = [1 \ L \ L^2 \ L^3]^\top$

iteration: $i = 2$ point: 24, 31, 15, 32, 83, 27, 20, 59

shift

$[1 \ 2 \ 4 \ 6]$

basis

$$\begin{bmatrix} X + 73 & 0 & 0 & 0 \\ 17 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 63 & 0 & 0 & 1 \end{bmatrix}$$

values

$$\begin{bmatrix} 0 & 7 & 88 & 8 & 59 & 3 & 93 & 35 \\ 0 & 90 & 90 & 52 & 83 & 63 & 11 & 81 \\ 0 & 93 & 93 & 63 & 90 & 81 & 38 & 24 \\ 0 & 13 & 13 & 64 & 51 & 11 & 41 & 16 \end{bmatrix}$$

fast algorithms

iterative algorithm [van Barel-Bultheel / Beckermann-Labahn]

parameters: $d = 8$ $m = 4$ $s = (0, 2, 4, 6)$, base field \mathbb{F}_{97}

input: $(24, 31, 15, 32, 83, 27, 20, 59)$ and $F = [1 \ L \ L^2 \ L^3]^T$

iteration: $i = 2$ point: 24, 31, 15, 32, 83, 27, 20, 59

shift

[1 2 4 6]

basis

$$\begin{bmatrix} X + 73 & 0 & 0 & 0 \\ X + 90 & 1 & 0 & 0 \\ 56X + 16 & 0 & 1 & 0 \\ 12X + 66 & 0 & 0 & 1 \end{bmatrix}$$

values

$$\begin{bmatrix} 0 & 7 & 88 & 8 & 59 & 3 & 93 & 35 \\ 0 & 0 & 81 & 60 & 45 & 66 & 7 & 19 \\ 0 & 0 & 74 & 26 & 96 & 55 & 8 & 44 \\ 0 & 0 & 2 & 63 & 80 & 47 & 90 & 48 \end{bmatrix}$$

fast algorithms

iterative algorithm [van Barel-Bultheel / Beckermann-Labahn]

parameters: $d = 8$ $m = 4$ $s = (0, 2, 4, 6)$, base field \mathbb{F}_{97}

input: $(24, 31, 15, 32, 83, 27, 20, 59)$ and $F = [1 \ L \ L^2 \ L^3]^T$

iteration: $i = 2$ point: 24, 31, 15, 32, 83, 27, 20, 59

shift

$[2 \ 2 \ 4 \ 6]$

basis

$$\begin{bmatrix} X^2 + 42X + 65 & 0 & 0 & 0 \\ X + 90 & 1 & 0 & 0 \\ 56X + 16 & 0 & 1 & 0 \\ 12X + 66 & 0 & 0 & 1 \end{bmatrix}$$

values

$$\begin{bmatrix} 0 & 0 & 47 & 8 & 61 & 85 & 44 & 10 \\ 0 & 0 & 81 & 60 & 45 & 66 & 7 & 19 \\ 0 & 0 & 74 & 26 & 96 & 55 & 8 & 44 \\ 0 & 0 & 2 & 63 & 80 & 47 & 90 & 48 \end{bmatrix}$$

fast algorithms

iterative algorithm [van Barel-Bultheel / Beckermann-Labahn]

parameters: $d = 8$ $m = 4$ $s = (0, 2, 4, 6)$, base field \mathbb{F}_{97}

input: $(24, 31, 15, 32, 83, 27, 20, 59)$ and $F = [1 \ L \ L^2 \ L^3]^T$

iteration: $i = 3$ point: 24, 31, 15, 32, 83, 27, 20, 59

shift

[2 2 4 6]

basis

$$\begin{bmatrix} X^2 + 42X + 65 & 0 & 0 & 0 \\ X + 90 & 1 & 0 & 0 \\ 56X + 16 & 0 & 1 & 0 \\ 12X + 66 & 0 & 0 & 1 \end{bmatrix}$$

values

$$\begin{bmatrix} 0 & 0 & 47 & 8 & 61 & 85 & 44 & 10 \\ 0 & 0 & 81 & 60 & 45 & 66 & 7 & 19 \\ 0 & 0 & 74 & 26 & 96 & 55 & 8 & 44 \\ 0 & 0 & 2 & 63 & 80 & 47 & 90 & 48 \end{bmatrix}$$

fast algorithms

iterative algorithm [van Barel-Bultheel / Beckermann-Labahn]

parameters: $d = 8$ $m = 4$ $s = (0, 2, 4, 6)$, base field \mathbb{F}_{97}

input: $(24, 31, 15, 32, 83, 27, 20, 59)$ and $\mathbf{F} = [1 \ L \ L^2 \ L^3]^\top$

iteration: $i = 3$ point: 24, 31, 15, 32, 83, 27, 20, 59

shift

[3 2 4 6]

basis

$$\begin{bmatrix} X^3 + 27X^2 + 17X + 92 & 0 & 0 & 0 \\ 54X^2 + 38X + 11 & 1 & 0 & 0 \\ 17X^2 + 91X + 54 & 0 & 1 & 0 \\ 66X^2 + 68X + 88 & 0 & 0 & 1 \end{bmatrix}$$

values

$$\begin{bmatrix} 0 & 0 & 0 & 39 & 74 & 50 & 26 & 52 \\ 0 & 0 & 0 & 7 & 41 & 0 & 55 & 74 \\ 0 & 0 & 0 & 65 & 66 & 45 & 77 & 20 \\ 0 & 0 & 0 & 9 & 32 & 31 & 84 & 29 \end{bmatrix}$$

fast algorithms

iterative algorithm [van Barel-Bultheel / Beckermann-Labahn]

parameters: $d = 8$ $m = 4$ $s = (0, 2, 4, 6)$, base field \mathbb{F}_{97}

input: $(24, 31, 15, 32, 83, 27, 20, 59)$ and $F = [1 \ L \ L^2 \ L^3]^T$

iteration: $i = 4$ point: 24, 31, 15, 32, 83, 27, 20, 59

shift

$[3 \ 2 \ 4 \ 6]$

basis

$$\begin{bmatrix} X^3 + 27X^2 + 17X + 92 & 0 & 0 & 0 \\ 54X^2 + 38X + 11 & 1 & 0 & 0 \\ 17X^2 + 91X + 54 & 0 & 1 & 0 \\ 66X^2 + 68X + 88 & 0 & 0 & 1 \end{bmatrix}$$

values

$$\begin{bmatrix} 0 & 0 & 0 & 39 & 74 & 50 & 26 & 52 \\ 0 & 0 & 0 & 7 & 41 & 0 & 55 & 74 \\ 0 & 0 & 0 & 65 & 66 & 45 & 77 & 20 \\ 0 & 0 & 0 & 9 & 32 & 31 & 84 & 29 \end{bmatrix}$$

fast algorithms

iterative algorithm [van Barel-Bultheel / Beckermann-Labahn]

parameters: $d = 8$ $m = 4$ $s = (0, 2, 4, 6)$, base field \mathbb{F}_{97}

input: $(24, 31, 15, 32, 83, 27, 20, 59)$ and $F = [1 \ L \ L^2 \ L^3]^T$

iteration: $i = 4$ point: 24, 31, 15, 32, 83, 27, 20, 59

shift

$[3 \ 3 \ 4 \ 6]$

basis

$$\begin{bmatrix} X^3 + 31X^2 + 27X + 3 & 36 & 0 & 0 \\ 54X^3 + 56X^2 + 56X + 36 & X + 65 & 0 & 0 \\ 56X^2 + 43X + 35 & 60 & 1 & 0 \\ 52X^2 + 33X + 60 & 68 & 0 & 1 \end{bmatrix}$$

values

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 95 & 50 & 66 & 0 \\ 0 & 0 & 0 & 0 & 54 & 0 & 19 & 58 \\ 0 & 0 & 0 & 0 & 4 & 45 & 79 & 95 \\ 0 & 0 & 0 & 0 & 7 & 31 & 41 & 17 \end{bmatrix}$$

fast algorithms

iterative algorithm [van Barel-Bultheel / Beckermann-Labahn]

parameters: $d = 8$ $m = 4$ $s = (0, 2, 4, 6)$, base field \mathbb{F}_{97}

input: $(24, 31, 15, 32, 83, 27, 20, 59)$ and $F = [1 \ L \ L^2 \ L^3]^T$

iteration: $i = 5$ point: 24, 31, 15, 32, 83, 27, 20, 59

shift

[4 3 4 6]

basis

$$\begin{bmatrix} X^4 + 45X^3 + 73X^2 + 90X + 42 & 36X + 19 & 0 & 0 \\ 81X^3 + 20X^2 + 9X + 20 & X + 67 & 0 & 0 \\ 2X^3 + 21X^2 + 41 & 35 & 1 & 0 \\ 52X^3 + 15X^2 + 79X + 22 & 0 & 0 & 1 \end{bmatrix}$$

values

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 13 & 13 & 0 \\ 0 & 0 & 0 & 0 & 0 & 89 & 55 & 58 \\ 0 & 0 & 0 & 0 & 0 & 48 & 17 & 95 \\ 0 & 0 & 0 & 0 & 0 & 12 & 78 & 17 \end{bmatrix}$$

fast algorithms

iterative algorithm [van Barel-Bultheel / Beckermann-Labahn]

parameters: $d = 8$ $m = 4$ $s = (0, 2, 4, 6)$, base field \mathbb{F}_{97}

input: $(24, 31, 15, 32, 83, 27, 20, 59)$ and $F = [1 \ L \ L^2 \ L^3]^T$

iteration: $i = 6$ point: 24, 31, 15, 32, 83, 27, 20, 59

shift

$[4 \ 4 \ 4 \ 6]$

basis

$$\begin{bmatrix} X^4 + 19X^3 + 57X^2 + 44X + 26 & 74X + 43 & 0 & 0 \\ 81X^4 + 64X^3 + 51X^2 + 68X + 42 & X^2 + 40X + 34 & 0 & 0 \\ 3X^3 + 44X^2 + 54X + 64 & 6X + 49 & 1 & 0 \\ 28X^3 + 45X^2 + 44X + 52 & 50X + 52 & 0 & 1 \end{bmatrix}$$

values

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 66 & 70 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & 13 \\ 0 & 0 & 0 & 0 & 0 & 0 & 56 & 55 \\ 0 & 0 & 0 & 0 & 0 & 0 & 15 & 7 \end{bmatrix}$$

fast algorithms

iterative algorithm [van Barel-Bultheel / Beckermann-Labahn]

parameters: $d = 8$ $m = 4$ $s = (0, 2, 4, 6)$, base field \mathbb{F}_{97}

input: $(24, 31, 15, 32, 83, 27, 20, 59)$ and $F = [1 \ L \ L^2 \ L^3]^T$

iteration: $i = 7$ point: 24, 31, 15, 32, 83, 27, 20, 59

shift

[5 4 4 6]

basis

$$\begin{bmatrix} X^5 + 96X^4 + 65X^3 + 68X^2 + 19X + 62 & 74X^2 + 18X + 13 & 0 & 0 \\ 6X^4 + 94X^3 + 44X^2 + 66X + 32 & X^2 + 19X + 10 & 0 & 0 \\ 55X^4 + 78X^3 + 75X^2 + 49X + 39 & 2X + 86 & 1 & 0 \\ 13X^4 + 81X^3 + 10X^2 + 34X + 2 & 42X + 29 & 0 & 1 \end{bmatrix}$$

values

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 14 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 25 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 44 \end{bmatrix}$$

fast algorithms

iterative algorithm [van Barel-Bultheel / Beckermann-Labahn]

parameters: $d = 8$ $m = 4$ $s = (0, 2, 4, 6)$, base field \mathbb{F}_{97}

input: $(24, 31, 15, 32, 83, 27, 20, 59)$ and $\mathbf{F} = [1 \ L \ L^2 \ L^3]^T$

iteration: $i = 8$ point: 24, 31, 15, 32, 83, 27, 20, 59

shift

[5 5 4 6]

basis

$$\begin{bmatrix} X^5 + 12X^4 + 10X^3 + 34X^2 + 65X + 2 & 60X^2 + 43X + 67 & 0 & 0 \\ 6X^5 + 31X^4 + 27X^3 + 89X^2 + 18X + 52 & X^3 + 57X^2 + 53X + 89 & 0 & 0 \\ 2X^4 + 56X^3 + 42X^2 + 48X + 15 & 72X^2 + 12X + 30 & 1 & 0 \\ 40X^4 + 19X^3 + 14X^2 + 40X + 49 & 53X^2 + 79X + 74 & 0 & 1 \end{bmatrix}$$

values

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

modular vector reconstruction

input:

- ▶ vector $\mathbf{F} = [f_1 \ \cdots \ f_m]^T \in \mathbb{K}[X]^{m \times 1}$ of degree $< d$
- ▶ field elements $(\alpha_1, \dots, \alpha_d) \in \mathbb{K}^d$
- ▶ shift $\mathbf{s} = (s_1, \dots, s_m) \in \mathbb{Z}^m$

output:

matrix $\mathbf{P} \in \mathbb{K}[X]^{m \times m}$ such that

- ▶ $\mathbf{P}\mathbf{F} = 0 \bmod \prod_{1 \leq i \leq d} (X - \alpha_i)$
- ▶ \mathbf{P} generates all vectors \mathbf{p} such that $\mathbf{p}\mathbf{F} = 0 \bmod \prod_{1 \leq i \leq d} (X - \alpha_i)$
- ▶ \mathbf{P} is \mathbf{s} -reduced

notation: $\mathcal{J}(\boldsymbol{\alpha}, \mathbf{F}) = \{\mathbf{p} \in \mathbb{K}[X]^{1 \times m} \mid \mathbf{p}\mathbf{F} = 0 \bmod \prod_{1 \leq i \leq d} (X - \alpha_i)\}$

fast algorithms

base case: modulus of degree 1

modular vector reconstruction: base case

input:

- ▶ vector $\mathbf{F} = [f_1 \ \cdots \ f_m]^T \in \mathbb{K}[X]^{m \times 1}$ of degree < 1
- ▶ field element $\alpha \in \mathbb{K}$
- ▶ shift $\mathbf{s} = (s_1, \dots, s_m) \in \mathbb{Z}^m$

output:

matrix $\mathbf{P} \in \mathbb{K}[X]^{m \times m}$ such that

- ▶ $\mathbf{P}\mathbf{F} = 0 \bmod (X - \alpha)$
- ▶ \mathbf{P} generates all vectors \mathbf{p} such that $\mathbf{p}\mathbf{F} = 0 \bmod (X - \alpha)$
- ▶ \mathbf{P} is \mathbf{s} -reduced

fast algorithms

base case: modulus of degree 1

modular vector reconstruction: base case

input:

- ▶ vector $\mathbf{F} = [f_1 \ \cdots \ f_m]^T \in \mathbb{K}[X]^{m \times 1}$ of degree < 1
- ▶ field element $\alpha \in \mathbb{K}$
- ▶ shift $\mathbf{s} = (s_1, \dots, s_m) \in \mathbb{Z}^m$

$$\mathbf{F} \in \mathbb{K}^{m \times 1}$$

output:

matrix $\mathbf{P} \in \mathbb{K}[X]^{m \times m}$ such that

- ▶ $\mathbf{P}\mathbf{F} = 0 \bmod (X - \alpha)$
- ▶ \mathbf{P} generates all vectors \mathbf{p} such that $\mathbf{p}\mathbf{F} = 0 \bmod (X - \alpha)$
- ▶ \mathbf{P} is \mathbf{s} -reduced

$$(\mathbf{P}\mathbf{F})(\alpha) = \mathbf{P}(\alpha)\mathbf{F} = 0$$

fast algorithms

base case: modulus of degree 1

modular vector reconstruction: base case

iterative algorithm:
$$\mathbf{P} = \begin{bmatrix} \mathbf{I}_{\pi-1} & \lambda_1 & \mathbf{0} \\ \mathbf{0} & X - \alpha & \mathbf{0} \\ \mathbf{0} & \lambda_2 & \mathbf{I}_{m-\pi} \end{bmatrix}$$

where

- ▶ π **minimizes** s_π among indices such that $(\mathbf{p}_\pi \mathbf{F})(\alpha_i) \neq 0$
- ▶ the vectors $\lambda_1 \in \mathbb{K}^{(\pi-1) \times 1}$ and $\lambda_2 \in \mathbb{K}^{(m-\pi) \times 1}$ are constant

fast algorithms

base case: modulus of degree 1

modular vector reconstruction: base case

iterative algorithm:
$$\mathbf{P} = \begin{bmatrix} \mathbf{I}_{\pi-1} & \lambda_1 & \mathbf{0} \\ \mathbf{0} & X - \alpha & \mathbf{0} \\ \mathbf{0} & \lambda_2 & \mathbf{I}_{m-\pi} \end{bmatrix}$$

where

- ▶ π **minimizes** s_π among indices such that $(\mathbf{p}_\pi \mathbf{F})(\alpha_i) \neq 0$
- ▶ the vectors $\lambda_1 \in \mathbb{K}^{(\pi-1) \times 1}$ and $\lambda_2 \in \mathbb{K}^{(m-\pi) \times 1}$ are constant

iterative algorithm:

- ▶ \mathbf{P} = identity matrix in $\mathbb{K}[X]^{m \times m}$
- ▶ for i from 1 to d :
 - from the evaluation $\mathbf{F}(\alpha_i)$, find \mathbf{P}_i as above
 - update shift $s_\pi \leftarrow s_\pi + 1$
 - update $\mathbf{P} \leftarrow \mathbf{P}_i \mathbf{P}$ as well as $\mathbf{F} \leftarrow \frac{\mathbf{P}_i \mathbf{F}}{X - \alpha_i} \bmod \prod_{i+1 \leq j \leq d} (X - \alpha_j)$
called **residual vector**

fast algorithms

base case: modulus of degree 1

modular vector reconstruction: base case

iterative algorithm:
$$\mathbf{P} = \begin{bmatrix} \mathbf{I}_{\pi-1} & \lambda_1 & \mathbf{0} \\ \mathbf{0} & X - \alpha & \mathbf{0} \\ \mathbf{0} & \lambda_2 & \mathbf{I}_{m-\pi} \end{bmatrix}$$

where

- ▶ π minimizes s_π among indices such that $(\mathbf{p}_\pi \mathbf{F})(\alpha_i) \neq 0$
- ▶ the vectors $\lambda_1 \in \mathbb{K}^{(\pi-1) \times 1}$ and $\lambda_2 \in \mathbb{K}^{(m-\pi) \times 1}$ are constant

complexity $O(m^2 d^2)$:

- ▶ iteration with d steps
- ▶ each step: evaluation of \mathbf{F} + multiplications $\mathbf{P}_i \mathbf{F}$ and $\mathbf{P}_i \mathbf{P}$
- ▶ at any stage \mathbf{F} has degree $< d$ and size $m \times 1$
- ▶ at any stage \mathbf{P} has degree $\leq d$ and size $m \times m$

normalizing at each step + refined analysis yields $O(md^2)$

fast algorithms

base case: modulus of degree 1

modular vector reconstruction: base case

iterative algorithm:
$$\mathbf{P} = \begin{bmatrix} \mathbf{I}_{\pi-1} & \lambda_1 & \mathbf{0} \\ \mathbf{0} & X - \alpha & \mathbf{0} \\ \mathbf{0} & \lambda_2 & \mathbf{I}_{m-\pi} \end{bmatrix}$$

where

- ▶ π **minimizes** s_π among indices such that $(\mathbf{p}_\pi \mathbf{F})(\alpha_i) \neq 0$
- ▶ the vectors $\lambda_1 \in \mathbb{K}^{(\pi-1) \times 1}$ and $\lambda_2 \in \mathbb{K}^{(m-\pi) \times 1}$ are constant

correctness:

- ▶ the main task is to prove the base case with \mathbf{P}_i
- ▶ then, direct consequence of the “basis multiplication theorem”

fast algorithms

iterative algorithm – complexity aspects

- ▶ **input size:** $md + d$ elements from \mathbb{K}
 - . md coefficients of \mathbf{F} , assumed reduced modulo $M(X)$
 - . d points $\alpha_1, \dots, \alpha_d$
- ▶ **output size:** $\leq m^2(d + 1)$ elements from \mathbb{K}
 - . $m \times m$ matrix \mathbf{P} of degree at most i at step i

is this output size bound tight?

fast algorithms

iterative algorithm – complexity aspects

- ▶ **input size:** $md + d$ elements from \mathbb{K}
 - . md coefficients of \mathbf{F} , assumed reduced modulo $M(X)$
 - . d points $\alpha_1, \dots, \alpha_d$
- ▶ **output size:** $\leq m^2(d + 1)$ elements from \mathbb{K}
 - . $m \times m$ matrix \mathbf{P} of degree at most i at step i

is this output size bound tight?

- ▶ one can prove $\deg(\det(\mathbf{P})) \leq d$
 - . \mathbf{P} is a basis of $\mathcal{J}(\alpha, \mathbf{F})$, which is the kernel of $\mathbb{K}[X]^m \rightarrow \mathbb{K}[X]/\langle M(X) \rangle, \mathbf{p} \mapsto \mathbf{p}\mathbf{F}$
 - . $\mathbb{K}[X]^m / \mathcal{J}(\alpha, \mathbf{F})$ has \mathbb{K} -dimension at most $\dim_{\mathbb{K}}(\mathbb{K}[X]/\langle M(X) \rangle) = d$
- ▶ **normalized bases** have average column degree $\leq d$, and size $\leq m(d + 1)$
- ▶ yet **the bound $\Theta(m^2(d + 1))$ is tight for this algorithm**
 - . normalizing at each step is feasible for the iterative version
 - . but is much harder to incorporate in fast divide and conquer versions

fast algorithms

iterative algorithm – complexity aspects

example instance of Hermite-Padé approximation
where the output size is in $\Omega(m^2d)$

parameters: $\mathbb{K} = \mathbf{F}_{97}$, $m = 4$, $\alpha = \mathbf{0}$, $d = 128$, $s = (0, \dots, 0)$

choose random polynomial $R(X)$ of degree < 128

$$\mathbf{F} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix} = \begin{bmatrix} R \\ R + XR \\ XR + X^2R \\ X^2R + X^3R \end{bmatrix}$$

- ▶ approximants are \mathbf{p} such that $\mathbf{pF} = 0 \bmod X^{128}$
- ▶ \mathbf{F} has small vectors in its left kernel
 \Rightarrow reduced approximant basis has **unbalanced** row degrees $(1, 1, 1, 125)$
- ▶ will help to build an example with **output size** $\Omega(m^2d)$

fast algorithms

iterative algorithm – complexity aspects

running the iterative algorithm:

i 1

s (0, 0, 0, 0)

f₁ R

f₂ R + XR

f₃ XR + X²R

f₄ X²R + X³R

P

fast algorithms

iterative algorithm – complexity aspects

running the iterative algorithm:

i	1	2
s	($0, 0, 0, 0$)	($1, 0, 0, 0$)
f_1	R	XR
f_2	$R + XR$	XR
f_3	$XR + X^2R$	$XR + X^2R$
f_4	$X^2R + X^3R$	$X^2R + X^3R$
P	$\begin{bmatrix} 1 & & & \\ 0 & 0 & & \\ & & 0 & \\ & & & 0 \end{bmatrix}$	

fast algorithms

iterative algorithm – complexity aspects

running the iterative algorithm:

i	1	2	3
s	(0 , 0, 0, 0)	(1, 0 , 0, 0)	(1, 1, 0 , 0)
f ₁	R	XR	0
f ₂	R + XR	XR	X²R
f ₃	XR + X ² R	XR + X ² R	X²R
f ₄	X ² R + X ³ R	X ² R + X ³ R	X²R + X ³ R
P	$\begin{bmatrix} \mathbf{1} & & & \\ 0 & \mathbf{0} & & \\ & & \mathbf{0} & \\ & & & \mathbf{0} \end{bmatrix}$	$\begin{bmatrix} \mathbf{1} & 0 & & \\ 1 & \mathbf{1} & & \\ 0 & 0 & \mathbf{0} & \\ & & & \mathbf{0} \end{bmatrix}$	

fast algorithms

iterative algorithm – complexity aspects

running the iterative algorithm:

i	1	2	3	4
s	(0 , 0, 0, 0)	(1, 0 , 0, 0)	(1, 1, 0 , 0)	(1, 1, 1, 0)
f ₁	R	XR	0	0
f ₂	R + XR	XR	X²R	0
f ₃	XR + X ² R	XR + X ² R	X²R	X³R
f ₄	X ² R + X ³ R	X ² R + X ³ R	X²R + X ³ R	X³R
P	$\begin{bmatrix} \mathbf{1} & & & \\ 0 & \mathbf{0} & & \\ & & \mathbf{0} & \\ & & & \mathbf{0} \end{bmatrix}$	$\begin{bmatrix} \mathbf{1} & \mathbf{0} & & \\ \mathbf{1} & \mathbf{1} & & \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \\ & & & \mathbf{0} \end{bmatrix}$	$\begin{bmatrix} \mathbf{1} & \mathbf{0} & & \\ \mathbf{1} & \mathbf{1} & \mathbf{0} & \\ \mathbf{1} & \mathbf{1} & \mathbf{1} & \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}$	

fast algorithms

iterative algorithm – complexity aspects

running the iterative algorithm:

i	1	2	3	4	...
s	(0 , 0, 0, 0)	(1, 0 , 0, 0)	(1, 1, 0 , 0)	(1, 1, 1, 0)	...
f ₁	R	XR	0	0	0
f ₂	R + XR	XR	X²R	0	0
f ₃	XR + X ² R	XR + X ² R	X²R	X³R	0
f ₄	X ² R + X ³ R	X ² R + X ³ R	X²R + X ³ R	X³R	X⁴R
P	$\begin{bmatrix} \mathbf{1} & & & \\ 0 & \mathbf{0} & & \\ & & \mathbf{0} & \\ & & & \mathbf{0} \end{bmatrix}$	$\begin{bmatrix} \mathbf{1} & 0 & & \\ 1 & \mathbf{1} & & \\ 0 & 0 & \mathbf{0} & \\ & & & \mathbf{0} \end{bmatrix}$	$\begin{bmatrix} \mathbf{1} & 0 & & \\ 1 & \mathbf{1} & 0 & \\ 1 & 1 & \mathbf{1} & \\ 0 & 0 & 0 & \mathbf{0} \end{bmatrix}$	$\begin{bmatrix} \mathbf{1} & 0 & & \\ 1 & \mathbf{1} & 0 & \\ 1 & 1 & \mathbf{1} & 0 \\ 1 & 1 & 1 & \mathbf{1} \end{bmatrix}$...

fast algorithms

iterative algorithm – complexity aspects

running the iterative algorithm:

i	1	2	3	4	...
s	(0, 0, 0, 0)	(1, 0, 0, 0)	(1, 1, 0, 0)	(1, 1, 1, 0)	...
f ₁	R	XR	0	0	0
f ₂	R + XR	XR	X ² R	0	0
f ₃	XR + X ² R	XR + X ² R	X ² R	X ³ R	0
f ₄	X ² R + X ³ R	X ² R + X ³ R	X ² R + X ³ R	X ³ R	X ⁴ R
P	$\begin{bmatrix} 1 & & & \\ 0 & 0 & & \\ & & 0 & \\ & & & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & & \\ 1 & 1 & & \\ 0 & 0 & 0 & \\ & & & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & & \\ 1 & 1 & 0 & \\ 1 & 1 & 1 & \\ 0 & 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & & \\ 1 & 1 & 0 & \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$...

degrees and “pivots” in final basis **P**:

$$\begin{bmatrix} 1 & 0 & & \\ 1 & 1 & 0 & \\ 1 & 1 & 1 & 0 \\ 125 & 125 & 125 & 125 \end{bmatrix}$$

fast algorithms

iterative algorithm – complexity aspects

parameters: $m = 8$, $d = 128$, $s = (0, 0, 0, 0, d, d, d, d)$

input \mathbf{F} : same f_1, f_2, f_3, f_4 / random f_5, f_6, f_7, f_8

$i = 4$

$$\begin{bmatrix} 1 & 0 & & & & & & \\ 1 & 1 & 0 & & & & & \\ 1 & 1 & 1 & 0 & & & & \\ 1 & 1 & 1 & 1 & & & & \\ 0 & 0 & 0 & 0 & 0 & & & \\ 0 & 0 & 0 & 0 & & 0 & & \\ 0 & 0 & 0 & 0 & & & 0 & \\ 0 & 0 & 0 & 0 & & & & 0 \end{bmatrix}$$

fast algorithms

iterative algorithm – complexity aspects

parameters: $m = 8$, $d = 128$, $s = (0, 0, 0, 0, d, d, d, d)$

input \mathbf{F} : same f_1, f_2, f_3, f_4 / random f_5, f_6, f_7, f_8

$i = 4$

$$\begin{bmatrix} 1 & 0 & & & & & & \\ 1 & 1 & 0 & & & & & \\ 1 & 1 & 1 & 0 & & & & \\ 1 & 1 & 1 & 1 & & & & \\ 0 & 0 & 0 & 0 & 0 & & & \\ 0 & 0 & 0 & 0 & & 0 & & \\ 0 & 0 & 0 & 0 & & & 0 & \\ 0 & 0 & 0 & 0 & & & & 0 \end{bmatrix}$$

$i = 128$

$$\begin{bmatrix} 1 & 0 & & & & & & \\ 1 & 1 & 0 & & & & & \\ 1 & 1 & 1 & 0 & & & & \\ 125 & 125 & 125 & 125 & & & & \\ 124 & 124 & 124 & 124 & 0 & & & \\ 124 & 124 & 124 & 124 & & 0 & & \\ 124 & 124 & 124 & 124 & & & 0 & \\ 124 & 124 & 124 & 124 & & & & 0 \end{bmatrix}$$

fast algorithms

iterative algorithm – complexity aspects

parameters: $m = 8$, $d = 128$, $s = (0, 0, 0, 0, d, d, d, d)$

input \mathbf{F} : same f_1, f_2, f_3, f_4 / random f_5, f_6, f_7, f_8

$i = 4$

$$\begin{bmatrix} 1 & 0 & & & & & & \\ 1 & 1 & 0 & & & & & \\ 1 & 1 & 1 & 0 & & & & \\ 1 & 1 & 1 & 1 & & & & \\ 0 & 0 & 0 & 0 & 0 & & & \\ 0 & 0 & 0 & 0 & & 0 & & \\ 0 & 0 & 0 & 0 & & & 0 & \\ 0 & 0 & 0 & 0 & & & & 0 \end{bmatrix}$$

$i = 128$

$$\begin{bmatrix} 1 & 0 & & & & & & \\ 1 & 1 & 0 & & & & & \\ 1 & 1 & 1 & 0 & & & & \\ 125 & 125 & 125 & 125 & & & & \\ 124 & 124 & 124 & 124 & 0 & & & \\ 124 & 124 & 124 & 124 & & 0 & & \\ 124 & 124 & 124 & 124 & & & 0 & \\ 124 & 124 & 124 & 124 & & & & 0 \end{bmatrix}$$

- ▶ 1/4 of the entries have degree $\approx d$: size $\Theta(m^2 d)$
- ▶ remark: complexity of iterative algorithm is $O(m^2 d^2)$
→ improved to $O(m d^2)$ via normalization
- ▶ opinions on a “reasonable” target cost for fast algorithms?

fast algorithms

recursion: residual and basis multiplication

divide and conquer algorithm:

input: $\mathbf{F}, (\alpha_1, \dots, \alpha_d), \mathbf{s}$ | output: \mathbf{P}

► if $d = 1$, use the base case algorithm to find \mathbf{P} and return

► otherwise:

a. $M_1 \leftarrow (X - \alpha_1) \cdots (X - \alpha_{\lfloor d/2 \rfloor}); M_2 \leftarrow (X - \alpha_{\lfloor d/2 \rfloor + 1}) \cdots (X - \alpha_d)$

b. $\mathbf{P}_1 \leftarrow$ call the algorithm on $\mathbf{F} \bmod M_1, (\alpha_1, \dots, \alpha_{\lfloor d/2 \rfloor}), \mathbf{s}$

c. updated shift: $\mathbf{t} \leftarrow \text{rdeg}_s(\mathbf{P}_1)$

d. residual: $\mathbf{G} \leftarrow \frac{1}{M_1} \mathbf{P}_1 \mathbf{F}$

e. $\mathbf{P}_2 \leftarrow$ call the algorithm on $\mathbf{G} \bmod M_2, (\alpha_{\lfloor d/2 \rfloor + 1}, \dots, \alpha_d), \mathbf{t}$

f. return the product $\mathbf{P}_2 \mathbf{P}_1$

fast algorithms

recursion: residual and basis multiplication

divide and conquer algorithm:

input: $\mathbf{F}, (\alpha_1, \dots, \alpha_d), \mathbf{s}$ | output: \mathbf{P}

► if $d = 1$, use the base case algorithm to find \mathbf{P} and return

► otherwise:

a. $M_1 \leftarrow (X - \alpha_1) \cdots (X - \alpha_{\lfloor d/2 \rfloor}); M_2 \leftarrow (X - \alpha_{\lfloor d/2 \rfloor + 1}) \cdots (X - \alpha_d)$

b. $\mathbf{P}_1 \leftarrow$ call the algorithm on $\mathbf{F} \bmod M_1, (\alpha_1, \dots, \alpha_{\lfloor d/2 \rfloor}), \mathbf{s}$

c. updated shift: $\mathbf{t} \leftarrow \text{rdeg}_s(\mathbf{P}_1)$

d. residual: $\mathbf{G} \leftarrow \frac{1}{M_1} \mathbf{P}_1 \mathbf{F}$

e. $\mathbf{P}_2 \leftarrow$ call the algorithm on $\mathbf{G} \bmod M_2, (\alpha_{\lfloor d/2 \rfloor + 1}, \dots, \alpha_d), \mathbf{t}$

f. return the product $\mathbf{P}_2 \mathbf{P}_1$

correctness:

► correctness of base case

► then, direct consequence of the “basis multiplication theorem”

► about the residual: $\{\mathbf{p} \mid \mathbf{p} \mathbf{P}_1 \mathbf{F} = 0 \bmod M\} = \{\mathbf{p} \mid \mathbf{p} \mathbf{G} = 0 \bmod M_2\}$

fast algorithms

recursion: residual and basis multiplication

divide and conquer algorithm:

input: $\mathbf{F}, (\alpha_1, \dots, \alpha_d), \mathbf{s}$ | output: \mathbf{P}

- ▶ if $d = 1$, use the base case algorithm to find \mathbf{P} and return
- ▶ otherwise:

- $M_1 \leftarrow (X - \alpha_1) \cdots (X - \alpha_{\lfloor d/2 \rfloor}); M_2 \leftarrow (X - \alpha_{\lfloor d/2 \rfloor + 1}) \cdots (X - \alpha_d)$
- $\mathbf{P}_1 \leftarrow$ call the algorithm on $\mathbf{F} \bmod M_1, (\alpha_1, \dots, \alpha_{\lfloor d/2 \rfloor}), \mathbf{s}$
- updated shift: $\mathbf{t} \leftarrow \text{rdeg}_s(\mathbf{P}_1)$
- residual: $\mathbf{G} \leftarrow \frac{1}{M_1} \mathbf{P}_1 \mathbf{F}$
- $\mathbf{P}_2 \leftarrow$ call the algorithm on $\mathbf{G} \bmod M_2, (\alpha_{\lfloor d/2 \rfloor + 1}, \dots, \alpha_d), \mathbf{t}$
- return the product $\mathbf{P}_2 \mathbf{P}_1$

complexity $O(m^\omega M(d) \log(d))$:

- ▶ if $\omega = 2$, quasi-linear in worst-case output size
- ▶ most expensive step in the recursion is the product $\mathbf{P}_2 \mathbf{P}_1$
- ▶ equation $\mathcal{C}(m, d) = \mathcal{C}(m, \lfloor d/2 \rfloor) + \mathcal{C}(m, \lceil d/2 \rceil) + O(m^\omega M(d))$

fast algorithms

recursion: residual and basis multiplication

input: $\deg(\mathbf{F}) < d$

output: $\deg(\mathbf{P}) \leq d$

complexity of each step:

- | | |
|--|--|
| ▸ residual $\mathbf{G} \leftarrow \frac{1}{M_1} \mathbf{P}_1 \mathbf{F}$ | $O(m^2 M(d))$ |
| ▸ $\mathbf{F} \bmod M_1$ and $\mathbf{G} \bmod M_2$ | $O(m M(d))$ |
| ▸ product $\mathbf{P}_2 \mathbf{P}_1$ | $O(m^\omega M(d))$ |
| ▸ two recursive calls | $2\mathcal{C}(m, \lfloor d/2 \rfloor)$ |

fast algorithms

recursion: residual and basis multiplication

input: $\deg(\mathbf{F}) < d$

output: $\deg(\mathbf{P}) \leq d$

complexity of each step:

- | | |
|--|--|
| ► residual $\mathbf{G} \leftarrow \frac{1}{M_1} \mathbf{P}_1 \mathbf{F}$ | $O(m^2 M(d))$ |
| ► $\mathbf{F} \bmod M_1$ and $\mathbf{G} \bmod M_2$ | $O(m M(d))$ |
| ► product $\mathbf{P}_2 \mathbf{P}_1$ | $O(m^\omega M(d))$ |
| ► two recursive calls | $2\mathcal{C}(m, \lfloor d/2 \rfloor)$ |

$$\begin{cases} \mathcal{C}(m, d) = \mathcal{C}(m, \lfloor d/2 \rfloor) + \mathcal{C}(m, \lceil d/2 \rceil) + O(m^\omega M(d)) \\ d \text{ base cases, each one costs } \dots ?? \end{cases}$$

fast algorithms

recursion: residual and basis multiplication

input: $\deg(\mathbf{F}) < d$

output: $\deg(\mathbf{P}) \leq d$

complexity of each step:

► residual $\mathbf{G} \leftarrow \frac{1}{M_1} \mathbf{P}_1 \mathbf{F}$	$O(m^2 M(d))$
► $\mathbf{F} \bmod M_1$ and $\mathbf{G} \bmod M_2$	$O(m M(d))$
► product $\mathbf{P}_2 \mathbf{P}_1$	$O(m^\omega M(d))$
► two recursive calls	$2\mathcal{C}(m, \lfloor d/2 \rfloor)$

$$\begin{cases} \mathcal{C}(m, d) = \mathcal{C}(m, \lfloor d/2 \rfloor) + \mathcal{C}(m, \lceil d/2 \rceil) + O(m^\omega M(d)) \\ d \text{ base cases, each one costs } O(m) \end{cases}$$

$$\Rightarrow O(m^\omega M(d) \log(d))$$

$$\text{unrolling: } m^\omega \left(M(d) + 2M\left(\frac{d}{2}\right) + 4M\left(\frac{d}{4}\right) + \cdots + \frac{d}{2}M(2) \right) + dm$$

fast algorithms

recursion: residual and basis multiplication

input: $\deg(\mathbf{F}) < d$

output: $\deg(\mathbf{P}) \leq d$

output: $\deg(\mathbf{P}) \approx \lceil \frac{d}{m} \rceil$

complexity of each step:

- ▶ residual $\mathbf{G} \leftarrow \frac{1}{M_1} \mathbf{P}_1 \mathbf{F}$
- ▶ $\mathbf{F} \bmod M_1$ and $\mathbf{G} \bmod M_2$
- ▶ product $\mathbf{P}_2 \mathbf{P}_1$
- ▶ two recursive calls

$O(m^2 M(d))$
 $O(m M(d))$
 $O(m^\omega M(d))$
 $2\mathcal{C}(m, \lfloor d/2 \rfloor)$

$s = 0$ and generic \mathbf{F} :

$O(m^\omega M(\lceil \frac{d}{m} \rceil))$
unchanged
 $O(m^\omega M(\lceil \frac{d}{m} \rceil))$
unchanged

- ▶ partial linearization

$$\begin{cases} \mathcal{C}(m, d) = \mathcal{C}(m, \lfloor d/2 \rfloor) + \mathcal{C}(m, \lceil d/2 \rceil) + O(m^\omega M(d)) \\ d \text{ base cases, each one costs } O(m) \end{cases}$$

$$\Rightarrow O(m^\omega M(d) \log(d))$$

fast algorithms

recursion: residual and basis multiplication

input: $\deg(\mathbf{F}) < d$

output: $\deg(\mathbf{P}) \leq d$

complexity of each step:

- ▶ residual $\mathbf{G} \leftarrow \frac{1}{M_1} \mathbf{P}_1 \mathbf{F}$
- ▶ $\mathbf{F} \bmod M_1$ and $\mathbf{G} \bmod M_2$
- ▶ product $\mathbf{P}_2 \mathbf{P}_1$
- ▶ two recursive calls

$O(m^2 M(d))$
 $O(m M(d))$
 $O(m^\omega M(d))$
 $2\mathcal{C}(m, \lfloor d/2 \rfloor)$

output: $\deg(\mathbf{P}) \approx \lceil \frac{d}{m} \rceil$

$s = 0$ and generic \mathbf{F} :

$O(m^\omega M(\lceil \frac{d}{m} \rceil))$
unchanged
 $O(m^\omega M(\lceil \frac{d}{m} \rceil))$
unchanged

- ▶ partial linearization
- ▶ base case for $d \approx m$, costs $O(m^\omega)$

$$\begin{cases} \mathcal{C}(m, d) = \mathcal{C}(m, \lfloor d/2 \rfloor) + \mathcal{C}(m, \lceil d/2 \rceil) + O(m^\omega M(d)) \\ d \text{ base cases, each one costs } O(m) \end{cases}$$

$$\Rightarrow O(m^\omega M(d) \log(d))$$

$$O(m^\omega M(\lceil \frac{d}{m} \rceil) \log(\lceil \frac{d}{m} \rceil))$$

fast algorithms

recursion: residual and basis multiplication

input: $\deg(\mathbf{F}) < d$

output: $\deg(\mathbf{P}) \leq d$

complexity of each step:

- ▶ residual $\mathbf{G} \leftarrow \frac{1}{M_1} \mathbf{P}_1 \mathbf{F}$
- ▶ \mathbf{F} rem M_1 and \mathbf{G} rem M_2
- ▶ product $\mathbf{P}_2 \mathbf{P}_1$
- ▶ two recursive calls

$$\begin{aligned} &O(m^2 M(d)) \\ &O(m M(d)) \\ &O(m^\omega M(d)) \\ &2\mathcal{C}(m, \lfloor d/2 \rfloor) \end{aligned}$$

output: $\deg(\mathbf{P}) \approx \lceil \frac{d}{m} \rceil$

$s = 0$ and generic \mathbf{F} :

$$\begin{aligned} &O(m^\omega M(\lceil \frac{d}{m} \rceil)) \\ &\text{unchanged} \\ &O(m^\omega M(\lceil \frac{d}{m} \rceil)) \\ &\text{unchanged} \end{aligned}$$

- ▶ partial linearization
- ▶ base case for $d \approx m$, costs $O(m^\omega)$

$$\begin{cases} \mathcal{C}(m, d) = \mathcal{C}(m, \lfloor d/2 \rfloor) + \mathcal{C}(m, \lceil d/2 \rceil) + O(m^\omega M(d)) \\ d \text{ base cases, each one costs } O(m) \end{cases}$$

$$\Rightarrow O(m^\omega M(d) \log(d))$$

$$O(m^\omega M(\lceil \frac{d}{m} \rceil) \log(\lceil \frac{d}{m} \rceil))$$

m	n	d	PM-BASIS	PM-BASIS with linearization
4	1	65536	1.6693	1.26891
16	1	16384	1.8535	0.89652
64	1	2048	2.2865	0.14362
256	1	1024	36.620	0.20660

fast algorithms

recursion: residual and basis multiplication

state of the art:

- ▶ **recursive** algorithm: from [Beckermann-Labahn 1994] (for Hermite-Padé)
it also works for $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$ with $n > 1$
- ▶ [Giorgi-Jeannerod-Villard 2003] achieved $O(m^\omega M(d) \log(d))$
for $\mathbf{F} \bmod X^d$, with $n \geq 1$ and $n \in O(m)$
- ▶ for $s = 0$ and **generic** \mathbf{F} : $O^\sim(m^\omega \lceil \frac{nd}{m} \rceil)$ [Lecerf, ca 2001, unpublished]

fast algorithms

recursion: residual and basis multiplication

state of the art:

- ▶ **recursive** algorithm: from [Beckermann-Labahn 1994] (for Hermite-Padé)
it also works for $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$ with $n > 1$
- ▶ [Giorgi-Jeannerod-Villard 2003] achieved $O(m^\omega M(d) \log(d))$
for $\mathbf{F} \bmod X^d$, with $n \geq 1$ and $n \in O(m)$
- ▶ for $s = 0$ and **generic** \mathbf{F} : $\tilde{O}(m^\omega \lceil \frac{nd}{m} \rceil)$ [Lecerf, ca 2001, unpublished]
- ▶ more recently: $\tilde{O}(m^{\omega-1}nd)$ for $\mathbf{F} \bmod X^d$
[Storjohann 2006] [Zhou-Labahn 2012] [Jeannerod-Neiger-Villard 2020]
 \rightsquigarrow **any** s , **no genericity** assumption, returns the **canonical basis** “s-Popov”

fast algorithms

recursion: residual and basis multiplication

state of the art:

- ▶ **recursive** algorithm: from [Beckermann-Labahn 1994] (for Hermite-Padé) it also works for $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$ with $n > 1$
- ▶ [Giorgi-Jeannerod-Villard 2003] achieved $O(m^\omega M(d) \log(d))$ for $\mathbf{F} \bmod X^d$, with $n \geq 1$ and $n \in O(m)$
- ▶ for $s = 0$ and **generic** \mathbf{F} : $O^\sim(m^\omega \lceil \frac{nd}{m} \rceil)$ [Lecerf, ca 2001, unpublished]
- ▶ more recently: $O^\sim(m^{\omega-1}nd)$ for $\mathbf{F} \bmod X^d$
[Storjohann 2006] [Zhou-Labahn 2012] [Jeannerod-Neiger-Villard 2020]
 \rightsquigarrow **any** s , **no genericity** assumption, returns the **canonical basis** “s-Popov”
- ▶ $\mathbf{F} \bmod M$ and **general modular matrix equations** in similar complexity
[Beckermann-Labahn 1997] [Jeannerod-Neiger-Schost-Villard 2017]
[Neiger-Vu 2017] [Rosenkilde-Storjohann 2021]
 \rightsquigarrow **any** s , **no genericity** assumption, returns the canonical “s-Popov” basis

outline

▶ introduction

- ▶ rational approximation and interpolation
- ▶ the vector case
- ▶ pol. matrices: reminders and motivation

▶ shifted reduced forms

- ▶ reducedness: examples and properties
- ▶ shifted forms and degree constraints
- ▶ stability under multiplication

▶ fast algorithms

- ▶ iterative algorithm and output size
- ▶ base case: modulus of degree 1
- ▶ recursion: residual and basis multiplication

▶ applications

outline

introduction

- rational approximation and interpolation
- the vector case
- pol. matrices: reminders and motivation

shifted reduced forms

- reducedness: examples and properties
- shifted forms and degree constraints
- stability under multiplication

fast algorithms

- iterative algorithm and output size
- base case: modulus of degree 1
- recursion: residual and basis multiplication

applications

- minimal kernel bases and linear systems
- fast gcd and extended gcd
- perspectives

applications

minimal kernel bases and linear systems

for $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$, its **left kernel** is

$$\mathcal{K}(\mathbf{F}) = \{\mathbf{p} \in \mathbb{K}[X]^{1 \times m} \mid \mathbf{p}\mathbf{F} = \mathbf{0}\}$$

- ▶ $\mathcal{K}(\mathbf{F})$ is a $\mathbb{K}[X]$ -module
- ▶ it has rank $m - r$, where r is the rank of \mathbf{F}

\Rightarrow basis $\mathbf{K} \in \mathbb{K}[X]^{(m-r) \times m}$

applications

minimal kernel bases and linear systems

for $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$, its **left kernel** is

$$\mathcal{K}(\mathbf{F}) = \{\mathbf{p} \in \mathbb{K}[X]^{1 \times m} \mid \mathbf{p}\mathbf{F} = \mathbf{0}\}$$

- ▶ $\mathcal{K}(\mathbf{F})$ is a $\mathbb{K}[X]$ -module
- ▶ it has rank $m - r$, where r is the rank of \mathbf{F}

\Rightarrow basis $\mathbf{K} \in \mathbb{K}[X]^{(m-r) \times m}$

kernel basis for a **constant matrix**?

input matrix \mathbf{F}

$$\begin{bmatrix} 5 & 6 \\ 6 & 1 \\ 2 & 6 \\ 5 & 2 \\ 5 & 6 \end{bmatrix}$$

applications

minimal kernel bases and linear systems

for $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$, its **left kernel** is

$$\mathcal{K}(\mathbf{F}) = \{\mathbf{p} \in \mathbb{K}[X]^{1 \times m} \mid \mathbf{p}\mathbf{F} = \mathbf{0}\}$$

► $\mathcal{K}(\mathbf{F})$ is a $\mathbb{K}[X]$ -module

► it has rank $m - r$, where r is the rank of \mathbf{F}

\Rightarrow basis $\mathbf{K} \in \mathbb{K}[X]^{(m-r) \times m}$

kernel basis for a **constant matrix**? \rightarrow usual nullspace

kernel basis \mathbf{K}

$$\begin{bmatrix} 5 & 6 & 1 & 0 & 0 \\ 0 & 5 & 0 & 1 & 0 \\ 0 & 0 & 3 & 2 & 1 \end{bmatrix}$$

input matrix \mathbf{F}

$$\begin{bmatrix} 5 & 6 \\ 6 & 1 \\ 2 & 6 \\ 5 & 2 \\ 5 & 6 \end{bmatrix}$$

applications

minimal kernel bases and linear systems

for $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$, its **left kernel** is

$$\mathcal{K}(\mathbf{F}) = \{\mathbf{p} \in \mathbb{K}[X]^{1 \times m} \mid \mathbf{p}\mathbf{F} = \mathbf{0}\}$$

- ▶ $\mathcal{K}(\mathbf{F})$ is a $\mathbb{K}[X]$ -module
- ▶ it has rank $m - r$, where r is the rank of \mathbf{F}

\Rightarrow basis $\mathbf{K} \in \mathbb{K}[X]^{(m-r) \times m}$

kernel basis of the following matrix over \mathbb{F}_2 ?

input matrix \mathbf{F}

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ X^2 & X^2 + X + 1 & X^2 + X \\ X^2 + 1 & X^2 & X^2 + X + 1 \\ X^2 & X^2 + X & X^2 \end{bmatrix}$$

applications

minimal kernel bases and linear systems

for $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$, its **left kernel** is

$$\mathcal{K}(\mathbf{F}) = \{\mathbf{p} \in \mathbb{K}[X]^{1 \times m} \mid \mathbf{p}\mathbf{F} = \mathbf{0}\}$$

► $\mathcal{K}(\mathbf{F})$ is a $\mathbb{K}[X]$ -module

► it has rank $m - r$, where r is the rank of \mathbf{F}

\Rightarrow basis $\mathbf{K} \in \mathbb{K}[X]^{(m-r) \times m}$

kernel basis of the following matrix over \mathbb{F}_2 ?

kernel basis \mathbf{K}						input matrix \mathbf{F}		
$\begin{bmatrix} X^2 & X^2 + X + 1 & X^2 + X & 1 & 0 & 0 \\ X^2 + 1 & X^2 & X^2 + X + 1 & 0 & 1 & 0 \\ X^2 & X^2 + X & X^2 & 0 & 0 & 1 \end{bmatrix}$						$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ X^2 & X^2 + X + 1 & X^2 + X \\ X^2 + 1 & X^2 & X^2 + X + 1 \\ X^2 & X^2 + X & X^2 \end{bmatrix}$		

applications

minimal kernel bases and linear systems

for $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$, its **left kernel** is

$$\mathcal{K}(\mathbf{F}) = \{\mathbf{p} \in \mathbb{K}[X]^{1 \times m} \mid \mathbf{p}\mathbf{F} = \mathbf{0}\}$$

- ▶ $\mathcal{K}(\mathbf{F})$ is a $\mathbb{K}[X]$ -module
- ▶ it has rank $m - r$, where r is the rank of \mathbf{F}

\Rightarrow basis $\mathbf{K} \in \mathbb{K}[X]^{(m-r) \times m}$

kernel basis of the following block matrix with \mathbf{G} nonsingular?

input matrix \mathbf{F}

$$\begin{bmatrix} \mathbf{G} \\ \mathbf{H} \end{bmatrix} \in \mathbb{K}[X]^{(n+m) \times n}$$

applications

minimal kernel bases and linear systems

for $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$, its **left kernel** is

$$\mathcal{K}(\mathbf{F}) = \{\mathbf{p} \in \mathbb{K}[X]^{1 \times m} \mid \mathbf{p}\mathbf{F} = \mathbf{0}\}$$

► $\mathcal{K}(\mathbf{F})$ is a $\mathbb{K}[X]$ -module

► it has rank $m - r$, where r is the rank of \mathbf{F}

\Rightarrow basis $\mathbf{K} \in \mathbb{K}[X]^{(m-r) \times m}$

kernel basis of the following block matrix with \mathbf{G} nonsingular?

kernel basis \mathbf{K}

... is left multiple of $\begin{bmatrix} -\mathbf{H}\mathbf{G}^{-1} & \mathbf{I}_m \end{bmatrix}$
... $\det(\mathbf{G}) \begin{bmatrix} -\mathbf{H}\mathbf{G}^{-1} & \mathbf{I}_m \end{bmatrix}$ is left multiple of it

input matrix \mathbf{F}

$$\begin{bmatrix} \mathbf{G} \\ \mathbf{H} \end{bmatrix} \in \mathbb{K}[X]^{(n+m) \times n}$$

applications

minimal kernel bases and linear systems

for $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$, its **left kernel** is

$$\mathcal{K}(\mathbf{F}) = \{\mathbf{p} \in \mathbb{K}[X]^{1 \times m} \mid \mathbf{p}\mathbf{F} = \mathbf{0}\}$$

- ▶ $\mathcal{K}(\mathbf{F})$ is a $\mathbb{K}[X]$ -module
- ▶ it has rank $m - r$, where r is the rank of \mathbf{F}

\Rightarrow basis $\mathbf{K} \in \mathbb{K}[X]^{(m-r) \times m}$

kernel basis of the following 4×1 vector with $R \in \mathbb{K}[X] \setminus \{0\}$?

input matrix \mathbf{F}

$$\begin{bmatrix} R \\ R + XR \\ XR + X^2R \\ X^2R + X^3R \end{bmatrix}$$

applications

minimal kernel bases and linear systems

for $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$, its **left kernel** is

$$\mathcal{K}(\mathbf{F}) = \{\mathbf{p} \in \mathbb{K}[X]^{1 \times m} \mid \mathbf{p}\mathbf{F} = \mathbf{0}\}$$

- ▶ $\mathcal{K}(\mathbf{F})$ is a $\mathbb{K}[X]$ -module
- ▶ it has rank $m - r$, where r is the rank of \mathbf{F}

\Rightarrow basis $\mathbf{K} \in \mathbb{K}[X]^{(m-r) \times m}$

kernel basis of the following 4×1 vector with $R \in \mathbb{K}[X] \setminus \{0\}$?

kernel basis \mathbf{K}

$$\begin{bmatrix} 1+X & -1 & & \\ 0 & X & -1 & \\ 0 & 0 & X & -1 \end{bmatrix}$$

input matrix \mathbf{F}

$$\begin{bmatrix} R \\ R + XR \\ XR + X^2R \\ X^2R + X^3R \end{bmatrix}$$

applications

minimal kernel bases and linear systems

for $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$, its **left kernel** is

$$\mathcal{K}(\mathbf{F}) = \{\mathbf{p} \in \mathbb{K}[X]^{1 \times m} \mid \mathbf{p}\mathbf{F} = \mathbf{0}\}$$

► $\mathcal{K}(\mathbf{F})$ is a $\mathbb{K}[X]$ -module

► it has rank $m - r$, where r is the rank of \mathbf{F}

\Rightarrow basis $\mathbf{K} \in \mathbb{K}[X]^{(m-r) \times m}$

$$\text{inclusion } \mathcal{K}(\mathbf{F}) \subset \mathcal{I}(\mathbf{M}, \mathbf{F}) = \{\mathbf{p} \in \mathbb{K}[X]^{1 \times m} \mid \mathbf{p}\mathbf{F} = \mathbf{0} \bmod \mathbf{M}\}$$

\Rightarrow **recover kernel via interpolation** with suitable choices of \mathbf{M}

applications

minimal kernel bases and linear systems

input:

- ▶ matrix $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$
- ▶ $\delta \in \mathbb{Z}_{>0}$ such that there exists a basis of $\mathcal{K}(\mathbf{F})$ of degree $\leq \delta$

algorithm via interpolation at sufficiently many points

- ▶ $d \leftarrow \delta + \deg(\mathbf{F}) + 1$
- ▶ $\alpha \leftarrow$ choose some $(\alpha_1, \dots, \alpha_d)$ in \mathbb{K}^d (not necessarily distinct)
- ▶ $\mathbf{P} \in \mathbb{K}[X]^{m \times m} \leftarrow$ reduced basis of $\mathcal{I}(\alpha, \mathbf{F})$
- ▶ $\mathbf{K} \in \mathbb{K}[X]^{k \times m} \leftarrow$ rows of \mathbf{P} which have degree $\leq \delta$

applications

minimal kernel bases and linear systems

input:

- ▶ matrix $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$
- ▶ $\delta \in \mathbb{Z}_{>0}$ such that there exists a basis of $\mathcal{K}(\mathbf{F})$ of degree $\leq \delta$

algorithm via interpolation at sufficiently many points

- ▶ $d \leftarrow \delta + \deg(\mathbf{F}) + 1$
- ▶ $\alpha \leftarrow$ choose some $(\alpha_1, \dots, \alpha_d)$ in \mathbb{K}^d (not necessarily distinct)
- ▶ $\mathbf{P} \in \mathbb{K}[X]^{m \times m} \leftarrow$ reduced basis of $\mathcal{I}(\alpha, \mathbf{F})$
- ▶ $\mathbf{K} \in \mathbb{K}[X]^{k \times m} \leftarrow$ rows of \mathbf{P} which have degree $\leq \delta$

$\Rightarrow \mathbf{K}$ is a reduced basis of $\mathcal{K}(\mathbf{F})$

\Rightarrow complexity $O(m^{\omega} M(\lceil \frac{nd}{m} \rceil) \log(\lceil \frac{nd}{m} \rceil))$

applications

minimal kernel bases and linear systems

input:

- ▶ matrix $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$
- ▶ $\delta \in \mathbb{Z}_{>0}$ such that there exists a basis of $\mathcal{K}(\mathbf{F})$ of degree $\leq \delta$

algorithm via interpolation at sufficiently many points

- ▶ $d \leftarrow \delta + \deg(\mathbf{F}) + 1$
- ▶ $\alpha \leftarrow$ choose some $(\alpha_1, \dots, \alpha_d)$ in \mathbb{K}^d (not necessarily distinct)
- ▶ $\mathbf{P} \in \mathbb{K}[X]^{m \times m} \leftarrow$ reduced basis of $\mathcal{I}(\alpha, \mathbf{F})$
- ▶ $\mathbf{K} \in \mathbb{K}[X]^{k \times m} \leftarrow$ rows of \mathbf{P} which have degree $\leq \delta$

$\Rightarrow \mathbf{K}$ is a reduced basis of $\mathcal{K}(\mathbf{F})$

\Rightarrow complexity $O(m^\omega M(\lceil \frac{nd}{m} \rceil) \log(\lceil \frac{nd}{m} \rceil))$

how to find the degree bound δ ?

applications

minimal kernel bases and linear systems

knowing $\delta \in \mathbb{Z}_{>0}$ such that there exists a basis of $\mathcal{K}(\mathbf{F})$ of degree $\leq \delta$

- ▶ take $d \leftarrow \delta + \deg(\mathbf{F}) + 1$ and some $\alpha \leftarrow (\alpha_1, \dots, \alpha_d)$ in \mathbb{K}^d
- ▶ $\mathbf{P} \in \mathbb{K}[X]^{m \times m}$ reduced basis of $\mathcal{J}(\alpha, \mathbf{F})$
- ▶ $\mathbf{K} \in \mathbb{K}[X]^{k \times m}$ rows of \mathbf{P} which have degree $\leq \delta$

$\Rightarrow \mathbf{K}$ is a reduced basis of $\mathcal{K}(\mathbf{F})$

applications

minimal kernel bases and linear systems

knowing $\delta \in \mathbb{Z}_{>0}$ such that there exists a basis of $\mathcal{K}(\mathbf{F})$ of degree $\leq \delta$

- ▶ take $d \leftarrow \delta + \deg(\mathbf{F}) + 1$ and some $\alpha \leftarrow (\alpha_1, \dots, \alpha_d)$ in \mathbb{K}^d
- ▶ $\mathbf{P} \in \mathbb{K}[X]^{m \times m}$ reduced basis of $\mathcal{J}(\alpha, \mathbf{F})$
- ▶ $\mathbf{K} \in \mathbb{K}[X]^{k \times m}$ rows of \mathbf{P} which have degree $\leq \delta$

$\Rightarrow \mathbf{K}$ is a reduced basis of $\mathcal{K}(\mathbf{F})$

proof:

$\Rightarrow \mathbf{K}$ is reduced by construction

. \mathbf{K} satisfies $\mathbf{KF} = \mathbf{0} \bmod (X - \alpha_1) \cdots (X - \alpha_d)$

. and $\deg(\mathbf{K}) \leq \delta$, hence $\deg(\mathbf{KF}) \leq \delta + \deg(\mathbf{F}) < d$

$\Rightarrow \mathbf{KF} = \mathbf{0}$, i.e. the rows of \mathbf{K} are in $\mathcal{K}(\mathbf{F})$

. let $\mathbf{B} \in \mathbb{K}[X]^{(m-r) \times m}$ be a basis of $\mathcal{K}(\mathbf{F})$ of degree $\leq \delta$

. then $\mathbf{B} = \mathbf{UP}$ for some \mathbf{U}

. by the predictable degree property, in fact $\mathbf{B} = \mathbf{VK}$

\Rightarrow any vector in $\mathcal{K}(\mathbf{F})$ is generated by \mathbf{K}

applications

minimal kernel bases and linear systems

knowing $\delta \in \mathbb{Z}_{>0}$ such that there exists a basis of $\mathcal{K}(\mathbf{F})$ of degree $\leq \delta$

how to find the degree bound δ ?

a specific bound may be known from the context e.g. gcd, "row bases"

► a general bound is $\delta = n \deg(\mathbf{F})$

► yields complexity $O^{\sim}(m^{\omega \lceil \frac{n^2 \deg(\mathbf{F})}{m} \rceil})$

how far from "optimal"?

applications

minimal kernel bases and linear systems

knowing $\delta \in \mathbb{Z}_{>0}$ such that there exists a basis of $\mathcal{K}(\mathbf{F})$ of degree $\leq \delta$

how to find the degree bound δ ?

a specific bound may be known from the context

e.g. gcd, "row bases"

► a general bound is $\delta = n \deg(\mathbf{F})$

► yields complexity $O^{\sim}(m^{\omega \lceil \frac{n^2 \deg(\mathbf{F})}{m} \rceil})$

how far from "optimal"?

proof:

complexity $O^{\sim}(m^{\omega \lceil \frac{nd}{m} \rceil})$

with $d = \delta + \deg(\mathbf{F}) + 1 = (n + 1) \deg(\mathbf{F}) + 1$

applications

minimal kernel bases and linear systems

knowing $\delta \in \mathbb{Z}_{>0}$ such that there exists a basis of $\mathcal{K}(\mathbf{F})$ of degree $\leq \delta$

how to find the degree bound δ ?

a specific bound may be known from the context e.g. gcd, "row bases"

► a general bound is $\delta = n \deg(\mathbf{F})$

► yields complexity $O(m^{\omega \lceil \frac{n^2 \deg(\mathbf{F})}{m} \rceil})$

how far from "optimal"?

proof:

up to row and column permutation, $\mathbf{F} = \begin{bmatrix} \mathbf{G} & * \\ \mathbf{H} & * \end{bmatrix}$

with $\mathbf{G} \in \mathbb{K}[X]^{r \times r}$ nonsingular

then, $\mathcal{K}(\mathbf{F}) = \mathcal{K}(\begin{bmatrix} \mathbf{G} \\ \mathbf{H} \end{bmatrix})$

the matrix $[-\mathbf{H}(\det(\mathbf{G})\mathbf{G}^{-1}) \quad \det(\mathbf{G})\mathbf{I}_{m-r}]$ has polynomial entries,

it has rank $m - r$ and its rows are in $\mathcal{K}(\mathbf{F})$,

it has degree $\leq \max(\deg \det(\mathbf{G}), \deg(\mathbf{H}) + (r - 1) \deg(\mathbf{G})) \leq r \deg(\mathbf{F})$

by degree minimality of reduced matrices,

any reduced basis of $\mathcal{K}(\mathbf{F})$ must have degree $\leq r \deg(\mathbf{F})$

applications

minimal kernel bases and linear systems

knowing $\delta \in \mathbb{Z}_{>0}$ such that there exists a basis of $\mathcal{K}(\mathbf{F})$ of degree $\leq \delta$

how to find the degree bound δ ?

a specific bound may be known from the context

e.g. gcd, “row bases”

► a general bound is $\delta = n \deg(\mathbf{F})$

► yields complexity $O^{\sim}(m^{\omega \lceil \frac{n^2 \deg(\mathbf{F})}{m} \rceil})$

how far from “optimal”?

► rules of thumb, generically:

“quantity of information is preserved”

+

“degrees in reduced basis are uniform”

$$\rightsquigarrow (m - r)m \deg(\mathbf{K}) \approx mn \deg(\mathbf{F})$$

$$\Leftrightarrow \deg(\mathbf{K}) \approx \frac{n}{m-r} \deg(\mathbf{F}), \text{ which is } \leq \frac{n}{m-n} \deg(\mathbf{F})$$

example: if \mathbf{F} is $m \times \frac{m}{2}$, generically $\deg(\mathbf{K}) = \deg(\mathbf{F})$

$\Rightarrow d = 2 \deg(\mathbf{F}) + 1$ and complexity $O^{\sim}(m^{\omega \deg(\mathbf{F})})$ how far from optimal?

applications

minimal kernel bases and linear systems

breakthrough

[Zhou-Labahn-Storjohann 2012]

- complexity $\tilde{O}(m^{\omega \lceil \frac{n \deg(\mathbf{F})}{m} \rceil})$ without assumption
- computes s -reduced basis of $\mathcal{K}(\mathbf{F})$ for $s = \text{rdeg}(\mathbf{F})$

- n large: divide and conquer on n , via **residual + basis multiplication**
 - ↪ partial linearization for **multiplying matrices** with weakly unbalanced degrees
- n small: use fast approximation/interpolation algorithms
 - ↪ **well-chosen d yields at least half the kernel efficiently**

applications

minimal kernel bases and linear systems

breakthrough

[Zhou-Labahn-Storjohann 2012]

- complexity $\tilde{O}(m^\omega \lceil \frac{n \deg(\mathbf{F})}{m} \rceil)$ without assumption
- computes s -reduced basis of $\mathcal{K}(\mathbf{F})$ for $s = \text{rdeg}(\mathbf{F})$

- n large: divide and conquer on n , via **residual + basis multiplication**
 - ↪ partial linearization for **multiplying matrices** with weakly unbalanced degrees
- n small: use fast approximation/interpolation algorithms
 - ↪ **well-chosen d yields at least half the kernel efficiently**

if $n > \frac{m}{2}$:

$\mathbf{K}_1 \leftarrow$ recursive call on first $\frac{n}{2}$ columns of \mathbf{F} , and shift s

$\mathbf{G} \leftarrow$ multiply $\mathbf{K}_1 \cdot \mathbf{F}_{*, \frac{n}{2}..n}$ (last $\frac{n}{2}$ columns of \mathbf{F})

$\mathbf{K}_2 \leftarrow$ recursive call on \mathbf{G} , and shift $t = \text{rdeg}_s(\mathbf{K}_1)$

return $\mathbf{K}_2 \mathbf{K}_1$

applications

minimal kernel bases and linear systems

breakthrough

[Zhou-Labahn-Storjohann 2012]

- complexity $O^{\omega}(\mathfrak{m}^{\omega \lceil \frac{n \deg(\mathbf{F})}{\mathfrak{m}} \rceil})$ without assumption
- computes s -reduced basis of $\mathcal{K}(\mathbf{F})$ for $s = \text{rdeg}(\mathbf{F})$

- n large: divide and conquer on n , via **residual + basis multiplication**
 \rightsquigarrow partial linearization for **multiplying matrices** with weakly unbalanced degrees
- n small: use fast approximation/interpolation algorithms
 \rightsquigarrow **well-chosen d yields at least half the kernel efficiently**

if $n \leq \frac{\mathfrak{m}}{2}$:

$\delta \leftarrow$ degree of kernel basis expected generically

$d \leftarrow \delta + \deg(\mathbf{F}) + 1$ and take some $\alpha \leftarrow (\alpha_1, \dots, \alpha_d)$ in \mathbb{K}^d

$\mathbf{P} \in \mathbb{K}[X]^{\mathfrak{m} \times \mathfrak{m}} \leftarrow s$ -reduced basis of $\mathcal{I}(\alpha, \mathbf{F})$

$\mathbf{K}_1, \mathbf{Q} \leftarrow$ **rows of \mathbf{P} which are in $\mathcal{K}(\mathbf{F})$** / which are not in $\mathcal{K}(\mathbf{F})$

$\mathbf{K}_2 \leftarrow$ recursive call on $\frac{1}{(X-\alpha_1)\cdots(X-\alpha_d)}\mathbf{QF}$, return $\begin{bmatrix} \mathbf{K}_1 \\ \mathbf{K}_2 \end{bmatrix}$

linear system solving:

given $\mathbf{A} \in \mathbb{K}[X]^{m \times m}$ nonsingular and $\mathbf{v} \in \mathbb{K}[X]^{1 \times m}$

find $\mathbf{u} \in \mathbb{K}[X]^{1 \times m}$ and $g \in \mathbb{K}[X]$ such that

$$\mathbf{u}\mathbf{A} = g\mathbf{v} \quad \text{and} \quad g \text{ has minimal degree.}$$

- . the equation has a solution: $\mathbf{u} = g\mathbf{v}\mathbf{A}^{-1}$ with $g = \det(\mathbf{A})$
- . but there is often no polynomial solution with $g = 1$
- . **target complexity?** (recall that $\det(\mathbf{A})\mathbf{A}^{-1}$ can have degree $\approx m \deg(\mathbf{A})$)
- . **propose an algorithm based on a kernel computation**

applications

minimal kernel bases and linear systems

linear system solving:

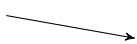
given $\mathbf{A} \in \mathbb{K}[X]^{m \times m}$ nonsingular and $\mathbf{v} \in \mathbb{K}[X]^{1 \times m}$

find $\mathbf{u} \in \mathbb{K}[X]^{1 \times m}$ and $g \in \mathbb{K}[X]$ such that

$$\mathbf{u}\mathbf{A} = g\mathbf{v} \quad \text{and} \quad g \text{ has minimal degree.}$$

- . the equation has a solution: $\mathbf{u} = g\mathbf{v}\mathbf{A}^{-1}$ with $g = \det(\mathbf{A})$
- . but there is often no polynomial solution with $g = 1$
- . **target complexity?** (recall that $\det(\mathbf{A})\mathbf{A}^{-1}$ can have degree $\approx m \deg(\mathbf{A})$)
- . **propose an algorithm based on a kernel computation**

compute $[\mathbf{u} \ g] \in \mathbb{K}[X]^{1 \times (m+1)}$ kernel basis of $\mathbf{F} = \begin{bmatrix} \mathbf{A} \\ -\mathbf{v} \end{bmatrix} \in \mathbb{K}[X]^{(m+1) \times m}$

- using the shift $\mathbf{s} = (\text{rdeg}(\mathbf{A}), \deg(\mathbf{v}))$
- complexity $\tilde{O}(m^\omega \max(\deg(\mathbf{A}), \deg(\mathbf{v})))$  in fact: $\max(\deg(\mathbf{A}), \frac{\deg(\mathbf{v})}{m})$
- \mathbf{u}, g is a solution to the equation $\mathbf{u}\mathbf{A} = g\mathbf{v}$
- minimality of $\deg(g)$ follows from *basis* of $\mathcal{K}(\mathbf{F})$

applications

fast gcd and extended gcd

gcd

input: f and g univariate polynomials in $\mathbb{K}[X]$

output: $h = \gcd(f, g)$

xgcd

input: f and g univariate polynomials in $\mathbb{K}[X]$

output: (u, v, h) where $h = \gcd(f, g) = uf + vg$

applications

fast gcd and extended gcd

gcd

input: f and g univariate polynomials in $\mathbb{K}[X]$
output: $h = \gcd(f, g)$

xgcd

input: f and g univariate polynomials in $\mathbb{K}[X]$
output: (u, v, h) where $h = \gcd(f, g) = uf + vg$

some notation:

. polynomials $\bar{f} = f/h$ and $\bar{g} = g/h$

. $m = \deg(f)$ and $n = \deg(g)$

. $\ell = \deg(h)$

\rightsquigarrow then $\deg(\bar{f}) = m - \ell$ and $\deg(\bar{g}) = n - \ell$

\bar{f} and \bar{g} are coprime

we assume $m, n > 0$

hence $\ell \leq \min(m, n)$

earlier in the course:

claim: gcd and xgcd are solved in $O(M(d) \log(d))$

where $d = \max(m, n)$

applications

fast gcd and extended gcd

gcd

input: f and g univariate polynomials in $\mathbb{K}[X]$

output: $h = \gcd(f, g)$

some notation:

. polynomials $\bar{f} = f/h$ and $\bar{g} = g/h$

\bar{f} and \bar{g} are coprime

. $m = \deg(f)$ and $n = \deg(g)$

we assume $m, n > 0$

result: gcd is solved in $O(M(\max(m, n)) \log(\max(m, n)))$

applications

fast gcd and extended gcd

gcd

input: f and g univariate polynomials in $\mathbb{K}[X]$

output: $h = \gcd(f, g)$

some notation:

. polynomials $\bar{f} = f/h$ and $\bar{g} = g/h$

\bar{f} and \bar{g} are coprime

. $m = \deg(f)$ and $n = \deg(g)$

we assume $m, n > 0$

result: gcd is solved in $O(M(\max(m, n)) \log(\max(m, n)))$

lemma: $[-\bar{g} \ \bar{f}]$ is a basis of the left kernel of $\begin{bmatrix} f \\ g \end{bmatrix}$

proof:

this kernel has rank 1 (f and g are nonzero)

let $[a \ b]$ be a basis of it; all other bases are $[ca \ cb]$ for some $c \in \mathbb{K} \setminus \{0\}$

since $[-\bar{g} \ \bar{f}]\begin{bmatrix} f \\ g \end{bmatrix} = -\frac{g}{h}f + \frac{f}{h}g = 0$, we get $[-\bar{g} \ \bar{f}] = [\lambda a \ \lambda b]$ for some $\lambda \in \mathbb{K}[X] \setminus \{0\}$

then λ divides \bar{f} and \bar{g} , so λ is a nonzero constant

applications

fast gcd and extended gcd

gcd

input: f and g univariate polynomials in $\mathbb{K}[X]$

output: $h = \gcd(f, g)$

some notation:

. polynomials $\bar{f} = f/h$ and $\bar{g} = g/h$

\bar{f} and \bar{g} are coprime

. $m = \deg(f)$ and $n = \deg(g)$

we assume $m, n > 0$

result: gcd is solved in $O(M(\max(m, n)) \log(\max(m, n)))$

lemma: $[-\bar{g} \ \bar{f}]$ is a basis of the left kernel of $\begin{bmatrix} f \\ g \end{bmatrix}$

algorithm: kernel basis via interpolation at sufficiently many points

- ▶ the input matrix $\mathbf{F} = \begin{bmatrix} f \\ g \end{bmatrix}$ has degree $\max(m, n)$
- ▶ the sought kernel basis has degree at most $\delta = \max(m, n)$

$\Rightarrow \begin{cases} 1. \text{ pick } \delta + \deg(\mathbf{F}) + 1 = 2\delta + 1 \text{ points } \alpha \in \mathbb{K}^{2\delta+1} \\ 2. \text{ find } [-\bar{g} \ \bar{f}] \text{ via a reduced basis of } \mathcal{I}(\alpha, \begin{bmatrix} f \\ g \end{bmatrix}) \\ 3. \text{ deduce } h = g/\bar{g} \end{cases}$

$O(1)$

$O(M(\delta) \log(\delta))$

$O(M(\delta))$

applications

fast gcd and extended gcd

xgcd

input: f and g univariate polynomials in $\mathbb{K}[X]$

output: (u, v, h) where $h = \gcd(f, g) = uf + vg$

some notation:

. polynomials $\bar{f} = f/h$ and $\bar{g} = g/h$

. $m = \deg(f)$, $n = \deg(g)$, $\ell = \deg(h)$

$\rightsquigarrow \deg(\bar{f}) = m - \ell$ and $\deg(\bar{g}) = n - \ell$

\bar{f} and \bar{g} are coprime

$m, n > 0$, $\ell \leq \min(m, n)$

applications

fast gcd and extended gcd

xgcd

input: f and g univariate polynomials in $\mathbb{K}[X]$

output: (u, v, h) where $h = \gcd(f, g) = uf + vg$

some notation:

. polynomials $\bar{f} = f/h$ and $\bar{g} = g/h$

\bar{f} and \bar{g} are coprime

. $m = \deg(f)$, $n = \deg(g)$, $\ell = \deg(h)$

$m, n > 0$, $\ell \leq \min(m, n)$

$\rightsquigarrow \deg(\bar{f}) = m - \ell$ and $\deg(\bar{g}) = n - \ell$

lemma:

. there exists a unique (u, v) in $\mathbb{K}[X]^2$ such that

$$\begin{cases} uf + vg = h, \\ \deg(u) < n - \ell \quad \text{and} \quad \deg(v) < m - \ell. \end{cases}$$

. for this $(u, v) \in \mathbb{K}[X]^2$ one has $\begin{bmatrix} u & v \\ -\bar{g} & \bar{f} \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix} = \begin{bmatrix} h \\ 0 \end{bmatrix}$,
and the leftmost matrix in this identity is unimodular

applications

fast gcd and extended gcd

xgcd

input: f and g univariate polynomials in $\mathbb{K}[X]$

output: (u, v, h) where $h = \gcd(f, g) = uf + vg$

some notation:

. polynomials $\bar{f} = f/h$ and $\bar{g} = g/h$

\bar{f} and \bar{g} are coprime

. $m = \deg(f)$, $n = \deg(g)$, $\ell = \deg(h)$

$m, n > 0$, $\ell \leq \min(m, n)$

$\rightsquigarrow \deg(\bar{f}) = m - \ell$ and $\deg(\bar{g}) = n - \ell$

theorem:

. defining $R = \begin{bmatrix} \text{rev}(u, n - \ell - 1) & \text{rev}(v, m - \ell - 1) \\ -\text{rev}(\bar{g}, n - \ell) & \text{rev}(\bar{f}, m - \ell) \end{bmatrix} \in \mathbb{K}[X]^{2 \times 2}$,

one has: $R \begin{bmatrix} \text{rev}(f, m) \\ \text{rev}(g, n) \end{bmatrix} = \begin{bmatrix} x^{m+n-2\ell-1} \text{rev}(h, \ell) \\ 0 \end{bmatrix}$

. the matrix R is a $(-n, -m)$ -reduced basis of $\mathcal{I}(\mathbf{0}, \begin{bmatrix} \text{rev}(f, m) \\ \text{rev}(g, n) \end{bmatrix})$

$$= \left\{ [p \ q] \in \mathbb{K}[X]^{1 \times 2} \mid [p \ q] \begin{bmatrix} \text{rev}(f, m) \\ \text{rev}(g, n) \end{bmatrix} = 0 \bmod x^{m+n-2\ell-1} \right\}$$

applications

fast gcd and extended gcd

xgcd

input: f and g univariate polynomials in $\mathbb{K}[X]$

output: (u, v, h) where $h = \gcd(f, g) = uf + vg$

some notation:

. polynomials $\bar{f} = f/h$ and $\bar{g} = g/h$

\bar{f} and \bar{g} are coprime

. $m = \deg(f)$, $n = \deg(g)$, $\ell = \deg(h)$

$m, n > 0$, $\ell \leq \min(m, n)$

$\rightsquigarrow \deg(\bar{f}) = m - \ell$ and $\deg(\bar{g}) = n - \ell$

theorem:

. defining $R = \begin{bmatrix} \text{rev}(u, n - \ell - 1) & \text{rev}(v, m - \ell - 1) \\ -\text{rev}(\bar{g}, n - \ell) & \text{rev}(\bar{f}, m - \ell) \end{bmatrix} \in \mathbb{K}[X]^{2 \times 2}$,

one has:
$$R \begin{bmatrix} \text{rev}(f, m) \\ \text{rev}(g, n) \end{bmatrix} = \begin{bmatrix} x^{m+n-2\ell-1} \text{rev}(h, \ell) \\ 0 \end{bmatrix}$$

ℓ is unknown!

. the matrix R is a $(-n, -m)$ -reduced basis of $\mathcal{I}(\mathbf{0}, \begin{bmatrix} \text{rev}(f, m) \\ \text{rev}(g, n) \end{bmatrix})$

$$= \left\{ [p \ q] \in \mathbb{K}[X]^{1 \times 2} \mid [p \ q] \begin{bmatrix} \text{rev}(f, m) \\ \text{rev}(g, n) \end{bmatrix} = 0 \bmod x^{m+n-2\ell-1} \right\}$$

applications

fast gcd and extended gcd

xgcd

input: f and g univariate polynomials in $\mathbb{K}[X]$

output: (u, v, h) where $h = \gcd(f, g) = uf + vg$

some notation:

. polynomials $\bar{f} = f/h$ and $\bar{g} = g/h$

\bar{f} and \bar{g} are coprime

. $m = \deg(f)$, $n = \deg(g)$, $\ell = \deg(h)$

$m, n > 0$, $\ell \leq \min(m, n)$

$\rightsquigarrow \deg(\bar{f}) = m - \ell$ and $\deg(\bar{g}) = n - \ell$

corollary: **xgcd** in $O(M(d) \log(d))$

for any $d \geq n + m - 2\ell - 1$

e.g. $d = n + m + 1$

let $e = d - (n + m - 2\ell - 1)$

hence $e = 2\ell$

$$\text{then } \begin{bmatrix} x^e & 0 \\ 0 & 1 \end{bmatrix} R = \begin{bmatrix} x^e \operatorname{rev}(u, n - \ell - 1) & x^e \operatorname{rev}(v, m - \ell - 1) \\ -\operatorname{rev}(\bar{g}, n - \ell) & \operatorname{rev}(\bar{f}, m - \ell) \end{bmatrix}$$

is a $(-n, -m)$ -reduced basis of

$$= \left\{ [p \ q] \in \mathbb{K}[X]^{1 \times 2} \mid [p \ q] \begin{bmatrix} \operatorname{rev}(f, m) \\ \operatorname{rev}(g, n) \end{bmatrix} = 0 \bmod x^d \right\}$$

applications

perspectives — row bases

a row basis of a matrix $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$
is a basis of its $\mathbb{K}[X]$ -row space $\rightarrow \{\mathbf{pF} \mid \mathbf{p} \in \mathbb{K}[X]^{1 \times m}\}$

\rightsquigarrow represented as $\mathbf{R} \in \mathbb{K}[X]^{r \times n}$, where r is the rank of \mathbf{F}

$\rightsquigarrow \mathbf{F} = \mathbf{UR}$ for some $\mathbf{U} \in \mathbb{K}[X]^{m \times r}$

applications

perspectives — row bases

a row basis of a matrix $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$
is a basis of its $\mathbb{K}[X]$ -row space $\rightarrow \{\mathbf{pF} \mid \mathbf{p} \in \mathbb{K}[X]^{1 \times m}\}$

\rightsquigarrow represented as $\mathbf{R} \in \mathbb{K}[X]^{r \times n}$, where r is the rank of \mathbf{F}

$\rightsquigarrow \mathbf{F} = \mathbf{UR}$ for some $\mathbf{U} \in \mathbb{K}[X]^{m \times r}$

examples:

► row basis for $\mathbf{F} \in \mathbb{K}[X]^{m \times m}$ nonsingular?

► row basis of $\begin{bmatrix} f \\ g \end{bmatrix}$ for f, g coprime polynomials?

► $\mathbf{K} \in \mathbb{K}[X]^{(m-r) \times m}$ a left kernel basis of $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$
row basis of \mathbf{K} ? column basis of \mathbf{K} ?

applications

perspectives — row bases

a row basis of a matrix $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$
is a basis of its $\mathbb{K}[X]$ -row space $\rightarrow \{\mathbf{pF} \mid \mathbf{p} \in \mathbb{K}[X]^{1 \times m}\}$

\rightsquigarrow represented as $\mathbf{R} \in \mathbb{K}[X]^{r \times n}$, where r is the rank of \mathbf{F}

$\rightsquigarrow \mathbf{F} = \mathbf{UR}$ for some $\mathbf{U} \in \mathbb{K}[X]^{m \times r}$

examples:

► row basis for $\mathbf{F} \in \mathbb{K}[X]^{m \times m}$ nonsingular?

$$\mathbf{R} = \mathbf{F}$$

► row basis of $\begin{bmatrix} f \\ g \end{bmatrix}$ for f, g coprime polynomials?

► $\mathbf{K} \in \mathbb{K}[X]^{(m-r) \times m}$ a left kernel basis of $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$
row basis of \mathbf{K} ? column basis of \mathbf{K} ?

applications

perspectives — row bases

a row basis of a matrix $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$
is a basis of its $\mathbb{K}[X]$ -row space $\rightarrow \{\mathbf{pF} \mid \mathbf{p} \in \mathbb{K}[X]^{1 \times m}\}$

\rightsquigarrow represented as $\mathbf{R} \in \mathbb{K}[X]^{r \times n}$, where r is the rank of \mathbf{F}

$\rightsquigarrow \mathbf{F} = \mathbf{UR}$ for some $\mathbf{U} \in \mathbb{K}[X]^{m \times r}$

examples:

► row basis for $\mathbf{F} \in \mathbb{K}[X]^{m \times m}$ nonsingular? $\mathbf{R} = \mathbf{F}$

► row basis of $\begin{bmatrix} f \\ g \end{bmatrix}$ for f, g coprime polynomials? $\mathbf{R} = [1]$

► $\mathbf{K} \in \mathbb{K}[X]^{(m-r) \times m}$ a left kernel basis of $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$
row basis of \mathbf{K} ? column basis of \mathbf{K} ?

applications

perspectives — row bases

a row basis of a matrix $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$
is a basis of its $\mathbb{K}[X]$ -row space $\rightarrow \{\mathbf{pF} \mid \mathbf{p} \in \mathbb{K}[X]^{1 \times m}\}$

\rightsquigarrow represented as $\mathbf{R} \in \mathbb{K}[X]^{r \times n}$, where r is the rank of \mathbf{F}

$\rightsquigarrow \mathbf{F} = \mathbf{UR}$ for some $\mathbf{U} \in \mathbb{K}[X]^{m \times r}$

examples:

► row basis for $\mathbf{F} \in \mathbb{K}[X]^{m \times m}$ nonsingular? $\mathbf{R} = \mathbf{F}$

► row basis of $\begin{bmatrix} f \\ g \end{bmatrix}$ for f, g coprime polynomials? $\mathbf{R} = [1]$

► $\mathbf{K} \in \mathbb{K}[X]^{(m-r) \times m}$ a left kernel basis of $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$
row basis of \mathbf{K} ? column basis of \mathbf{K} ? $\mathbf{R} = \mathbf{K}$ and $\mathbf{C} = \mathbf{I}_{m-r}$

\mathbf{K} has full rank so \mathbf{C} is $(m-r) \times (m-r)$ nonsingular

and by definition $\mathbf{K} = \mathbf{C}\bar{\mathbf{K}}$ for some $\bar{\mathbf{K}}$

so $\mathbf{KF} = \mathbf{0} \Rightarrow \bar{\mathbf{K}}\mathbf{F} = \mathbf{0}$, hence $\bar{\mathbf{K}} = \mathbf{VK}$

from $\mathbf{K} = \mathbf{CVK}$, with \mathbf{K} having full row rank, we deduce $\mathbf{CV} = \mathbf{I}_{m-r}$

applications

perspectives — row bases

a row basis of a matrix $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$
is a basis of its $\mathbb{K}[X]$ -row space $\rightarrow \{\mathbf{pF} \mid \mathbf{p} \in \mathbb{K}[X]^{1 \times m}\}$

\rightsquigarrow represented as $\mathbf{R} \in \mathbb{K}[X]^{r \times n}$, where r is the rank of \mathbf{F}

$\rightsquigarrow \mathbf{F} = \mathbf{UR}$ for some $\mathbf{U} \in \mathbb{K}[X]^{m \times r}$

applications:

- ▶ compute an s-reduced basis of the row space
- ▶ verify that a matrix is a kernel basis
- ▶ triangularization: Hermite normal form and determinant

applications

perspectives — row bases

a row basis of a matrix $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$
is a basis of its $\mathbb{K}[X]$ -row space $\rightarrow \{\mathbf{pF} \mid \mathbf{p} \in \mathbb{K}[X]^{1 \times m}\}$

\rightsquigarrow represented as $\mathbf{R} \in \mathbb{K}[X]^{r \times n}$, where r is the rank of \mathbf{F}

$\rightsquigarrow \mathbf{F} = \mathbf{UR}$ for some $\mathbf{U} \in \mathbb{K}[X]^{m \times r}$

applications:

- ▶ compute an s-reduced basis of the row space
- ▶ verify that a matrix is a kernel basis
- ▶ triangularization: Hermite normal form and determinant

algorithm:

- ▶ $\mathbf{K} \leftarrow$ left kernel basis for \mathbf{F}
- ▶ $\mathbf{G} \leftarrow$ right kernel basis for \mathbf{K}
- ▶ $\mathbf{R} \leftarrow$ matrix such that $\mathbf{F} = \mathbf{GR}$

complexity $\tilde{O}(mn^{\omega-1} \deg(\mathbf{F}))$, assuming $m \geq n$ [Zhou-Labahn, 2013]

applications

perspectives — triangularization

[Mulders-Storjohann 2003, Giorgi-Jeannerod-Villard 2003, Zhou 2012]

triangularization of $m \times m$ matrix \mathbf{A} using $\frac{m}{2} \times \frac{m}{2}$ blocks

not computed —

$$\begin{bmatrix} * & * \\ \mathbf{K}_1 & \mathbf{K}_2 \end{bmatrix} \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{A}_3 & \mathbf{A}_4 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & * \\ \mathbf{0} & \mathbf{B} \end{bmatrix}$$

kernel basis of $\begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_3 \end{bmatrix}$

$$\mathbf{K}_1 \mathbf{A}_2 + \mathbf{K}_2 \mathbf{A}_4$$

row basis of $\begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_3 \end{bmatrix}$

applications

perspectives — triangularization

[Mulders-Storjohann 2003, Giorgi-Jeannerod-Villard 2003, Zhou 2012]

triangularization of $m \times m$ matrix \mathbf{A} using $\frac{m}{2} \times \frac{m}{2}$ blocks

not computed $\begin{bmatrix} * & * \\ \mathbf{K}_1 & \mathbf{K}_2 \end{bmatrix} \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{A}_3 & \mathbf{A}_4 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & * \\ \mathbf{0} & \mathbf{B} \end{bmatrix}$

kernel basis of $\begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_3 \end{bmatrix}$

$$\mathbf{K}_1 \mathbf{A}_2 + \mathbf{K}_2 \mathbf{A}_4$$

row basis of $\begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_3 \end{bmatrix}$

main property: $\begin{bmatrix} * & * \\ \mathbf{K}_1 & \mathbf{K}_2 \end{bmatrix}$ is unimodular

- Hermite form of \mathbf{A} = Hermite form of $\begin{bmatrix} \mathbf{R} & * \\ \mathbf{0} & \mathbf{B} \end{bmatrix}$
- $\det(\mathbf{A}) = \det(\mathbf{R}) \det(\mathbf{B})$

Hermite normal form and determinant in $O^\sim(m^\omega \deg(\mathbf{A}))$

[Zhou, 2012] [Labahn-Neiger-Zhou, 2017]

applications

perspectives — block Wiedemann techniques

given a **sparse** matrix $\mathbf{A} \in \mathbb{K}^{n \times n}$:

- ▶ solve a **linear system** $\mathbf{A}\mathbf{u} = \mathbf{v}$
- ▶ compute the **minimal polynomial** of \mathbf{A}

- . sparse means that \mathbf{A} has a large proportion of zero entries
- . goal: exploit sparsity to do better than exponent ω

[Wiedemann 1986, Coppersmith 1994, Kaltofen 1995, Villard 1997]

block Wiedemann approach, for block dimension m :

1. choose random blocking matrices $\mathbf{U}, \mathbf{V} \in \mathbb{K}^{n \times m}$
2. compute **linearly recurrent sequence of matrices** in $\mathbb{K}^{m \times m}$
$$\mathbf{U}^T \mathbf{V}, \mathbf{U}^T \mathbf{A} \mathbf{V}, \dots, \mathbf{U}^T \mathbf{A}^k \mathbf{V}, \dots$$
3. find polynomial matrix generator $\mathbf{P} \in \mathbb{K}[X]^{m \times m}$ of this sequence

applications

perspectives — block Wiedemann techniques

given a **sparse** matrix $\mathbf{A} \in \mathbb{K}^{n \times n}$:

- ▶ solve a **linear system** $\mathbf{A}\mathbf{u} = \mathbf{v}$
- ▶ compute the **minimal polynomial** of \mathbf{A}

- . sparse means that \mathbf{A} has a large proportion of zero entries
- . goal: exploit sparsity to do better than exponent ω

[Wiedemann 1986, Coppersmith 1994, Kaltofen 1995, Villard 1997]

block Wiedemann approach, for block dimension m :

1. choose random blocking matrices $\mathbf{U}, \mathbf{V} \in \mathbb{K}^{n \times m}$
2. compute **linearly recurrent sequence of matrices** in $\mathbb{K}^{m \times m}$
$$\mathbf{U}^T \mathbf{V}, \mathbf{U}^T \mathbf{A} \mathbf{V}, \dots, \mathbf{U}^T \mathbf{A}^k \mathbf{V}, \dots$$
3. find polynomial matrix generator $\mathbf{P} \in \mathbb{K}[X]^{m \times m}$ of this sequence

- ▶ generically, $d = 2 \frac{n}{m} - 1$ terms of the sequence are sufficient
- ▶ step 3 is **matrix-Padé approx.**, in $O^\sim(m^\omega d) = O^\sim(m^{\omega-1} n)$
- ▶ often, m is taken as the **number of threads** available for parallel computation of the matrix sequence

summary

introduction

- rational approximation and interpolation
- the vector case
- pol. matrices: reminders and motivation

shifted reduced forms

- reducedness: examples and properties
- shifted forms and degree constraints
- stability under multiplication

fast algorithms

- iterative algorithm and output size
- base case: modulus of degree 1
- recursion: residual and basis multiplication

applications

- minimal kernel bases and linear systems
- fast gcd and extended gcd
- perspectives