Inverting integer and polynomial matrices

Jo60

Arne Storjohann
University of Waterloo

Integer Matrix Inverse

Input: An $n \times n$ matrix A filled with entries of size d digits.

Output: The matrix inverse A^{-1}.

Example: 6×6 with $d = 2$

$$\begin{bmatrix} 67 & -81 & -77 & -2 & 69 & 10 \\ 29 & -9 & -18 & 27 & -74 & 94 \\ 44 & -50 & 87 & -93 & -4 & 12 \\ 92 & -22 & 33 & -76 & 27 & -2 \\ -31 & 45 & -98 & -72 & 8 & 50 \\ 99 & -16 & -38 & 57 & -32 & 25 \end{bmatrix}^{-1} = \begin{bmatrix} -\frac{613719389}{436045910232} & \frac{20118095}{72674318372} & -\frac{851335927}{218022955116} & \frac{3893593471}{436045910232} & -\frac{433417321}{436045910232} & \frac{297871357}{72674318372} \\ -\frac{886053851}{145348636744} & \frac{620810971}{72674318372} & -\frac{1522814569}{72674318372} & \frac{3362614441}{145348636744} & -\frac{453009351}{145348636744} & -\frac{838573559}{72674318372} \\ -\frac{99479911}{109011477558} & \frac{218826900}{18168579593} & -\frac{674660030}{54505738779} & \frac{1856662385}{109011477558} & -\frac{1148992613}{109011477558} & -\frac{300458509}{18168579593} \\ \frac{447817619}{218022955116} & \frac{340592137}{36337159186} & -\frac{1594931579}{109011477558} & \frac{2114306231}{218022955116} & -\frac{2037856205}{218022955116} & -\frac{347815739}{36337159186} \\ \frac{770731325}{109011477558} & \frac{356811254}{18168579593} & -\frac{1721772194}{54505738779} & \frac{3697142975}{109011477558} & -\frac{1450539425}{109011477558} & -\frac{584700471}{18168579593} \\ \frac{2028363569}{436045910232} & \frac{1921892393}{72674318372} & -\frac{5197032317}{218022955116} & \frac{11614232501}{436045910232} & -\frac{4273458011}{436045910232} & -\frac{2043699293}{72674318372} \end{bmatrix}$$

- input A has total size $O(n^2 d)$ words
- output A^{-1} has total size $\tilde{O}(n^3 d)$ words

Polynomial Matrix Inverse

Input: An $n \times n$ matrix $A \in \mathsf{K}[x]$ of degree bounded by d.

Output: The matrix inverse A^{-1}.

Example: 3×3 over $\mathbb{Z}_7[x]$ with $d = 3$

$$\begin{bmatrix} 6x^2+4x+2 & 2x^2+5x+3 & 2x^2+x \\ 5x^2+3x+2 & 6x^2+6x+4 & 4x^2+2 \\ 5x^2+3x & 5x^2+4x+3 & 5x^2+4 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{2x^4+3x^2+6x+2}{x^6+4x^5+5x^4+6x^3+6x^2+4x+2} & \frac{6x^3+3x^2+5x+6}{x^6+4x^5+5x^4+6x^3+6x^2+4x+2} & \frac{2x^4+6x^3+6x^2+4x+4}{x^6+4x^5+5x^4+6x^3+6x^2+4x+2} \\ \frac{6x^3+4x^2+3}{x^5+5x^4+3x^3+2x^2+x+5} & \frac{4x^3+3x^2+5x+4}{x^5+5x^4+3x^3+2x^2+x+5} & \frac{6x^2+2x+5}{x^5+5x^4+3x^3+2x^2+x+5} \\ \frac{6x^4+3x^3+4x^2+x+4}{x^6+4x^5+5x^4+6x^3+6x^2+4x+2} & \frac{3x^4+3x^3+2x+3}{x^6+4x^5+5x^4+6x^3+6x^2+4x+2} & \frac{x^4+3x^3+x^2+6x+6}{x^6+4x^5+5x^4+6x^3+6x^2+4x+2} \end{bmatrix}$$

- input A has total size $O(n^2 d)$ field elements from K
- output A^{-1} has total size $O(n^3 d)$

Notation

- nonsingular input A has dimension n
- $\mathsf{K}[x]$: size of entries is $d \geq \deg A$
  $\Rightarrow$ count field operations from K
- $\mathbb{Z}$: size of entries is $d \geq \log ||A||_\infty$
  $\Rightarrow$ count word operations
- ω is a feasible exponent for matrix multiplication: $2 \leq \omega \leq 3$

Classical inversion algorithms

- Hensel-Newton iteration: p-adic lifting
- Homomorphic imaging and Chinese remaindering

Classical		Goal
$O^{\sim}(n^{\omega+1}d)$	$\longrightarrow$	$O^{\sim}(n^3 d)$

Outline: Three completely different approaches for inversion

A1 Sparse inverse expansion

- useful representation of $A^{-1} \bmod p^n$
- applicable for integer and polynomial inputs

A2 Balanced kernel basis

- nearly optimal inversion algorithm for generic polynomials inputs
- [*Jeannerod & Villard, 2005*]

A3 Outer product adjoint formula

- nearly optimal inversion algorithm for
  - all polynomial inputs
  - well conditioned integer inputs

Alternative representations for the inverse of $A = \begin{bmatrix} 3 & 2 \\ x & 5 \end{bmatrix}$

$$A^{-1} = \begin{bmatrix} \frac{2}{6+2x} & \frac{2}{6+2x} \\ \frac{x}{6+2x} & \frac{4}{6+2x} \end{bmatrix}$$

$$= \left[\begin{array}{c|c} 1+x+x^2+x^3+\cdots & -1-x-x^2-x^3+\cdots \\ \hline -x-x^2-x^3+\cdots & 1+x+x^2+x^3+\cdots \end{array}\right]$$

$$= \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} x + \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} x^2 + \cdots .$$

- A^{-1} over $\mathsf{K}[[x]]$ modulo x^{2nd} sufficient to reconstruct A^{-1} over $\mathsf{K}(x)$
- compute expansion over $\mathsf{K}[[x]]$ using quadratic Newton iteration

The sparse inverse formula for $A = \left[\, 1 - cx \,\right]$

Explicit inverse modulo x^{32}

$$(1-cx)^{-1} \bmod x^{32} \equiv 1 + cx + c^2x^2 + c^3x^3 + c^4x^4 + c^5x^5 + \cdots + c^{31}x^{31}$$

The sparse inverse formula for $A = \left[\, 1 - cx \,\right]$

Explicit inverse modulo x^{32}

$$(1-cx)^{-1} \bmod x^{32} \equiv 1 + cx + c^2x^2 + c^3x^3 + c^4x^4 + c^5x^5 + \cdots + c^{31}x^{31}$$

$$\equiv (1+cx)(1+c^2x^2)(1+c^4x^4)(1+c^8x^8)(1+c^{16}x^{16})$$

The sparse inverse formula for $A = \left[\, 1 - cx \,\right]$

Explicit inverse modulo x^{32}

$$(1-cx)^{-1} \bmod x^{32} = \overbrace{1 + cx + c^2x^2 + c^3x^3 + c^4x^4 + c^5x^5 + \cdots + c^{31}x^{31}}^{\text{32 coefficients}}$$

$$= \underbrace{(1+cx)(1+c^2x^2)(1+c^4x^4)(1+c^8x^8)(1+c^{16}x^{16})}_{\text{5 coefficients}}$$

$$\underbrace{(1+\overbrace{c}^{R_1}x)(1+\overbrace{c^2}^{R_2}x^2)(1+\overbrace{c^4}^{R_4}x^4)}_{(1-cx)^{-1} \bmod x^8}(1+\overbrace{c^8}^{R_8}x^8)(1+\overbrace{c^{16}}^{R_{16}}x^{16})$$

Computing the residues for $A = 1 - cx$

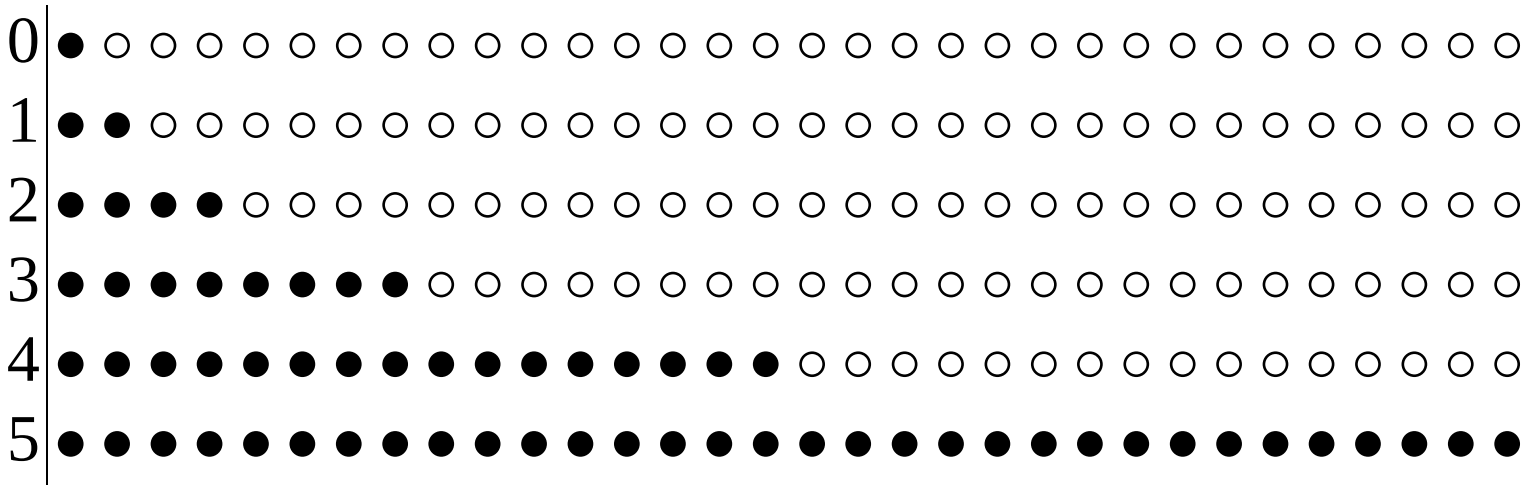
$$
\begin{aligned}
R_2 x^2 &= I - A(1 + cx) \\
&= c^2 x^2 \\
R_4 x^4 &= I - A(\overbrace{1 + cx + c^2x^2 + c^3x^3}^{A^{-1} \bmod x^4}) \\
&= c^4 x^4 \\
R_8 x^8 &= I - A(\overbrace{1 + cx + c^2x^2 + c^3x^3 + c^4x^4 + c^5x^5 + c^6x^6 + c^7x^7}^{A^{-1} \bmod x^8}) \\
&= c^8 x^8
\end{aligned}
$$

Computing the residues for $A = 1 - cx$

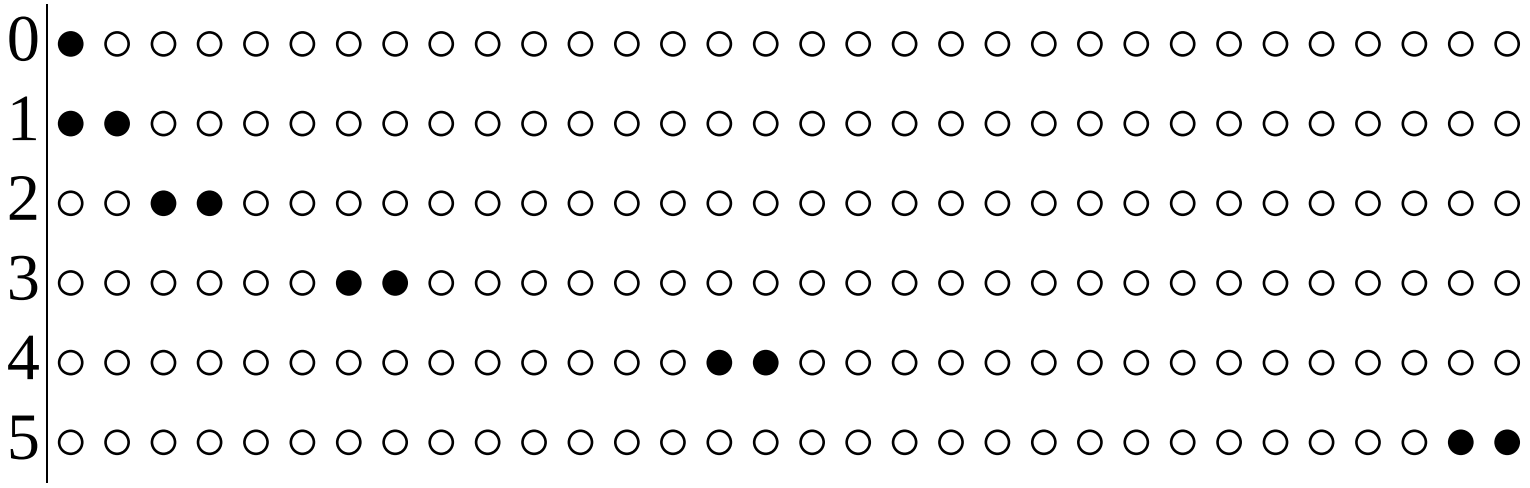
$$
\begin{aligned}
R_2 x^2 &= I - A(1 + cx) \\
&= c^2 x^2 \\
R_4 x^4 &= I - A(\overbrace{1 + cx + c^2x^2 + c^3x^3}^{A^{-1} \bmod x^4}) \\
&= c^4 x^4 \\
R_8 x^8 &= I - A(\overbrace{1 + cx + c^2x^2 + c^3x^3 + c^4x^4 + c^5x^5 + c^6x^6 + c^7x^7}^{A^{-1} \bmod x^8}) \\
&= c^8 x^8
\end{aligned}
$$

- If $\deg A = d$ then $R_2, R_4, R_8, \cdots$ have degree $\leq d - 1$.
- Can use reverse short product
  $\Rightarrow$ need only top d coefficients of $A^{-1} \bmod x^*$

Standard quadratic lifting



High-order component lifting



Sparse inverse formula for integer and polynomial matrices

Example for $7^{-1} \bmod x^{32}$, $x = 11$

$$7^{-1} \bmod x^{32}$$

$$= 904876146229395122376692251804252$$

$$= (((52(1-3x^2)+2x^4)(1-5x^4)+4x^8)(1-3x^8)+2x^{16})(1-5x^{16})+4x^{32}$$

General case

- formula can be computed in time $O\tilde{}(n^\omega d)$
- sparse inverse formula for $A^{-1} \bmod x^n$ has size $O(n^2(\log n)d)$
- can be used to compute $A^{-1}b$ in time $O\tilde{}(n^\omega d)$
- $\Rightarrow$ fast algorithms for linear algebra over $\mathbb{Z}$ and $\mathsf{K}[x]$
- randomized Las Vegas
- algorithm for explicit inverse?

Outline: Three completely different approaches for inversion

A1 Sparse inverse expansion

- useful representation of $A^{-1} \bmod p^n$
- applicable for integer and polynomial inputs

A2 Balanced kernel basis

- nearly optimal inversion algorithm for generic polynomials inputs
- [*Jeannerod & Villard, 2005*]

A3 Outer product adjoint formula

- nearly optimal inversion algorithm for
  - all polynomial inputs
  - well conditioned integer inputs

Minimal kernel basis of generic polynomial matrices

Example 1: $A \in \mathsf{K}[x]^{2\times 1}$, $\mathsf{K} = \mathbb{Z}_{97}$

$$\overset{N}{\begin{bmatrix} g & -f \end{bmatrix}} \overset{A}{\begin{bmatrix} f \\ g \end{bmatrix}} = 0_{1\times 1}$$

Example 2: $A \in \mathsf{K}[x]^{8\times 4}$, $\mathsf{K} = \mathbb{Z}_{97}$

$$\overset{N}{\begin{bmatrix} x+19 & 81 & 69 & 39 & 87x & 45x+51 & 51x+44 & 33x+85 \\ 49 & x+9 & 75 & 15 & 95x+92 & 67x+90 & 83x+29 & 47x+35 \\ 3 & 5 & x+48 & 7 & 54x+64 & 42x+38 & 12x+67 & 67x+79 \\ 63 & 56 & 34 & x+78 & 31x+50 & 4x+52 & 93x+28 & 17x+52 \end{bmatrix}} \overset{A}{\begin{bmatrix} -60x+16 & 52x-20 & -4x-89 & -77x+69 \\ -64x+89 & -16x+59 & -69x-46 & -33x+87 \\ 18 & 52x+36 & 91x-22 & 51x-27 \\ 31x-27 & 65x+88 & 10x-6 & 80x-84 \\ -26x-51 & 88x+97 & -67x+58 & 29x+37 \\ 9x-91 & 81x+65 & -12x+78 & 5x-63 \\ 42x+9 & -21x-27 & -79x-22 & -51x+16 \\ -95x+86 & -97x-14 & 83x-96 & -8x-54 \end{bmatrix}} = 0_{4\times 4}$$

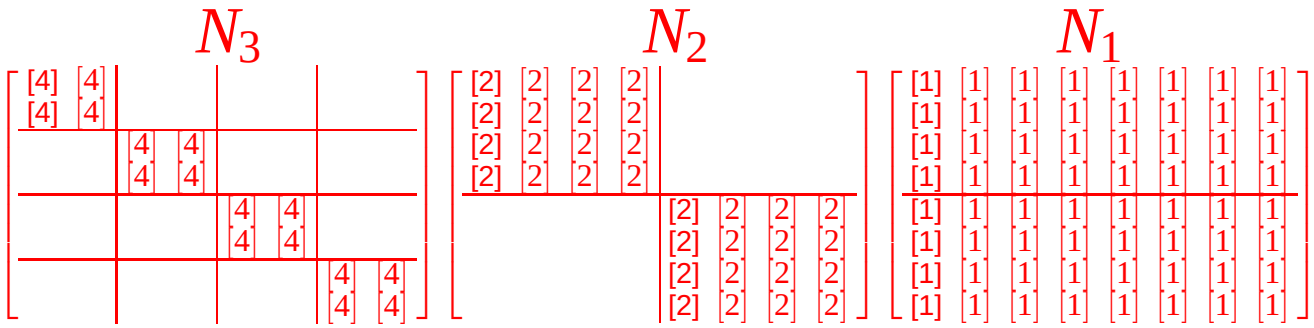
- generic $A \in \mathsf{K}[x]^{2n\times n}$ of degree d has kernel of degree d
- effective algorithms, e.g., [Giorgi, Jeannerod & Villard, ISSAC 03]

From generic kernel to generic inversion

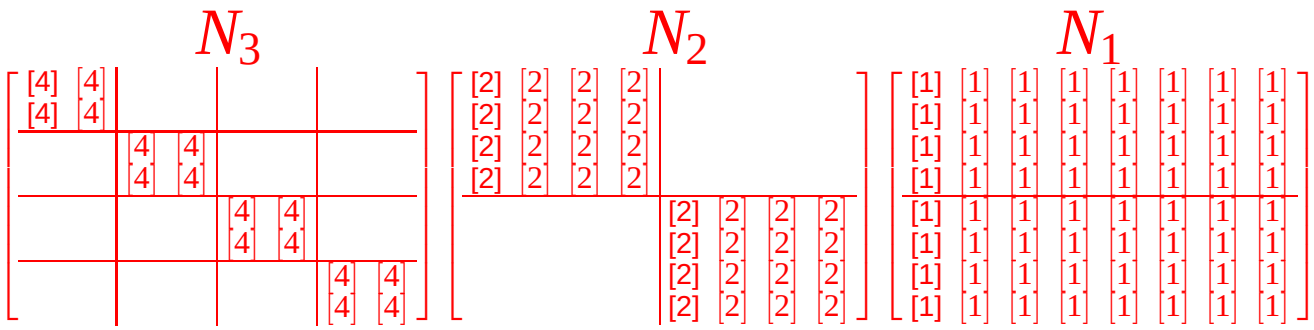
- split input $A \in \mathsf{K}[x]^{n \times n}$ of degree d into two halves: $A = \left[\, A_L \,\middle|\, A_R \,\right]$
- compute minimal kernel $\underline{U}$ for A_L and $\overline{U}$ for A_R

$$\begin{bmatrix} \overline{U} \\ \underline{U} \end{bmatrix} \left[\, A_L \,\middle|\, A_R \,\right] = \left[\begin{array}{c|c} \overline{U}A_L & \\ \hline & \underline{U}A_R \end{array}\right]$$

- recurse on degree $2d$ matrices $\overline{U}A_L$ and $\underline{U}A_R$
- Example: $n = 8$ and $d = 1$



From generic kernel to generic inversion



- cost to compute $N_1, N_2, \ldots, N_{\log_2 n}$ is $O\tilde{}(n^\omega d))$
- cost to compute $(\det A)A^{-1} = N_{\log n} \cdot N_2 \cdots N_1$ is $O\tilde{}(n^3 d)$

Notes

- requires generic input
- seems to require n to be a power of 2

Outline: Three completely different approaches for inversion

A1 Sparse inverse expansion

- useful representation of $A^{-1} \bmod p^n$
- applicable for integer and polynomial inputs

A2 Balanced kernel basis

- nearly optimal inversion algorithm for generic polynomials inputs
- [*Jeannerod & Villard, 2005*]

A3 Outer product adjoint formula

- nearly optimal inversion algorithm for
  - all polynomial inputs
  - well conditioned integer inputs

Special case of the outer product adjoint formula

Example input matrix

$$A = \begin{bmatrix} 67 & -93 & -74 & 99 & 33 \\ 44 & -77 & -72 & 27 & -38 \\ 92 & 57 & -2 & 8 & -18 \\ -31 & 27 & -32 & 69 & 87 \\ 29 & -98 & -76 & -4 & -81 \end{bmatrix}, \qquad \det A = 319347140$$

The adjoint

$$A^{\text{adj}} = \begin{bmatrix} 6117164 & -33407370 & 2491416 & 5361562 & 23369838 \\ -1592156 & 17039940 & -4579084 & -4579688 & -12544052 \\ -12481954 & 43671360 & -4016196 & -154642 & -24846648 \\ -23859948 & 97379940 & -12597172 & -10758524 & -64161216 \\ 17006140 & -78361370 & 10822480 & 8136810 & 53967650 \end{bmatrix}$$

$$\equiv \frac{1}{6117164} \begin{bmatrix} 6117164 \\ -1592156 \\ -12481954 \\ -23859948 \\ 17006140 \end{bmatrix} \begin{bmatrix} 6117164 & -33407370 & 2491416 & 5361562 & 23369838 \end{bmatrix} \bmod 319347140$$

General case of the output product adjoint formla

- suppose Smith form of A is $\mathrm{Diag}(s_1, s_2, \ldots, s_n)$

$$A^{\mathsf{adj}} \equiv \frac{\det A}{s_n} v_n u_n + \frac{\det A}{s_{n-1}} v_{n-1} u_{n-1} + \cdots + \frac{\det A}{s_1} v_1 u_1 \pmod{\det A}$$

Integer matrices

- only obtain inverse modulo $\det A$
- inverse of $A \in \mathbb{Z}^{n \times n}$ in time $O\tilde{}(n^3 d + n^3 \log \kappa(A))$
- $\Rightarrow$ nearly optimal for well conditioned input
- randomized Las Vegas

Polynomial matrices

- use randomization and reversion to extend to arbitrary input
- inverse of any $A \in \mathsf{K}[x]^{n \times n}$ in time $O\tilde{}(n^3 d)$ operations from K
- randomized Las Vegas

Applications of the nearly optimal inversion formulas

Example: Linear solving over K: $\mathrm{K} = \mathbb{Z}_7$, $n = 5$

$$A = \begin{bmatrix} 5 & 1 & 0 & 4 & 5 \\ 5 & 2 & 3 & 1 & 0 \\ 4 & 6 & 0 & 4 & 0 \\ 5 & 2 & 0 & 2 & 4 \\ 5 & 1 & 1 & 0 & 2 \end{bmatrix} \qquad b = \begin{bmatrix} 1 \\ 6 \\ 2 \\ 4 \\ 3 \end{bmatrix}$$

$$\det A = 3$$

$$A^{-1}b = \begin{bmatrix} 0 \\ 5 \\ 1 \\ 0 \\ 2 \end{bmatrix}$$

Note: no expression swell over $\mathbb{Z}_7$

Example input over $\mathsf{K}[x]$: $n = 5$, $d = 2$ $\Longrightarrow$ expression swell

$$A = \begin{bmatrix} 6x^2+x+5 & 6x+1 & 2x^2 & x+4 & 3x^2+6x+5 \\ 3x+5 & 2x^2+x+2 & 6x^2+2x+3 & 3x^2+2x+1 & x^2+2x \\ 5x^2+x+4 & 6x^2+3x+6 & 5x^2+2x & 5x^2+6x+4 & 3x^2+x \\ 6x^2+6x+5 & 2x^2+3x+2 & 3x^2+6x & 3x^2+2x+2 & 5x^2+5x+4 \\ 3x^2+2x+5 & 5x^2+3x+1 & 4x+1 & 5x^2+4x & 2x^2+3x+2 \end{bmatrix} \quad b = \begin{bmatrix} x^2+x+1 \\ 2x^2+2x+6 \\ x^2+3x+2 \\ 2x^2+x+4 \\ x^2+4x+3 \end{bmatrix}$$

$$\det A = 6x^{10}+6x^9+x^8+3x^6+x^5+x^4+4x^2+2x+3$$

Degree is $n \times d$ where d is degrees in input matrix.

$$A^{-1}b = \begin{bmatrix} x^{10}+2x^9+2x^8+x^7+3x^6+x^4+6x^3+3x \\ 5x^9+x^8+6x^7+x^5+4x^4+5x^3+x^2+3x+1 \\ 6x^{10}+x^9+3x^8+6x^7+3x^6+2x^5+2x^3+5x^2+3 \\ 5x^{10}+3x^9+4x^8+3x^7+2x^6+3x^5+5x^4+6x \\ 3x^{10}+3x^9+4x^7+5x^6+2x^5+x+6 \end{bmatrix} (1/\det A)$$

Nearly optimal linear system solving with precomputation

Recall situation over K

Input: Nonsingular $A \in \mathsf{K}^{n \times n}$

Precompute: Inverse A^{-1}

Compute: $A^{-1}v$ for any given $v \in \mathsf{K}^{n \times 1}$

$$\underset{A^{-1}}{\begin{bmatrix} 3 & 0 & 1 & 3 & 1 \\ 4 & 0 & 5 & 6 & 2 \\ 3 & 1 & 5 & 1 & 1 \\ 3 & 6 & 3 & 6 & 4 \\ 5 & 5 & 3 & 1 & 4 \end{bmatrix}} \underset{b}{\begin{bmatrix} 5 \\ 5 \\ 6 \\ 6 \\ 3 \end{bmatrix}} = \underset{A^{-1}b}{\begin{bmatrix} 6 \\ 3 \\ 0 \\ 1 \\ 5 \end{bmatrix}}$$

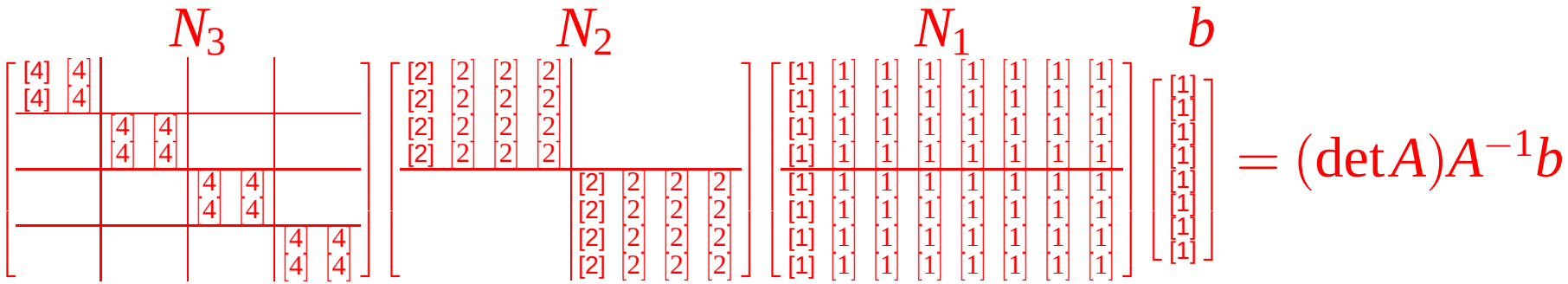
- size of $A^{-1}b$ is $\Omega(n)$
- time to compute $A^{-1}b$ is $O(n^2)$

Nearly optimal linear system solving with precomputation over $\mathsf{K}[x]$

Input: nonsingular $A \in \mathsf{K}[x]^{n \times n}$ of degree d

Precompute: adjoint decomposition $N_{\log n} \cdots N_2 N_1 = (\det A) A^{-1}$.

Compute: $A^{-1}v$ for any given $v \in \mathsf{K}^{n \times 1}$ of degree d.



- size of $A^{-1}b$ is $\Omega(n^2 d)$
- cost to compute $A^{-1}b$ is $O\tilde{}(n^2 d)$
- result for nongeneric input via outer product adjoint formula

Krylov basis and matrix powers

Input: matrix $A \in \mathsf{K}^{n \times n}$ and vector $v \in \mathsf{K}^{n \times 1}$
Output: $v, Av, A^2, A^3v, \ldots, A^{n-1}v$

Standard algorithm for matrix vector interates

- compute $Av, A(Av), A(A^2v), A(A^3v), \ldots, A(A^{n-2}v)$
- requires $n-1$ matrix-vector products
- total cost is $O(n^3)$ operations from K

Fast algorithm [*Keller-Gehrig, 1985]*

- precompute $A^2, (A^2)^2, (A^4)^2, \ldots,$
- compute iterates via matrix multiplication
- total cost is $O(n^w \times \log n)$ operations from K

Nearly optimal computation of matrix powers

Input: matrix $A \in \mathsf{K}^{n \times n}$

Output: $I, A, A^2, A^3, A^4, \ldots, A^{n-1}$

- consider matrix $I - xA \in \mathsf{K}[x]$
- compute $(I - xA)^{-1} \bmod x^n = I + xA + x^2A^2 + x^3A^3 + \cdots x^{n-1}A^{n-1}$
- total size of output is $\Theta(n^3)$ elements of K
- total cost is $\tilde{O}(n^3)$ using fast inversion algorithm