Multiplication and Inversion Algorithms for Matrices with Displacement Structure

Claude-Pierre Jeannerod

Inria – Université de Lyon
Laboratoire LIP (CNRS, ENSL, Inria, UCBL)

05.2019

Based on joint work with A. Bostan, C. Mouilleron, and É. Schost.

Structured matrices

- $n \times n$ dense matrices with "patterns".

Structured matrices

- $n \times n$ dense matrices with "patterns".
- Examples:

$$\text{Toeplitz} \begin{bmatrix} a_0 & a_{-1} & a_{-2} \\ a_1 & a_0 & a_{-1} \\ a_2 & a_1 & a_0 \end{bmatrix} \qquad \begin{bmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \\ a_2 & a_3 & a_4 \end{bmatrix} \text{Hankel}$$

$$\text{Vandermonde} \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{bmatrix} \qquad \begin{bmatrix} \frac{1}{x_1-y_1} & \frac{1}{x_1-y_2} & \frac{1}{x_1-y_3} \\ \frac{1}{x_2-y_1} & \frac{1}{x_2-y_2} & \frac{1}{x_2-y_3} \\ \frac{1}{x_3-y_1} & \frac{1}{x_3-y_2} & \frac{1}{x_3-y_3} \end{bmatrix} \text{Cauchy}$$

Structured matrices

- $n \times n$ dense matrices with "patterns".
- Examples:

$$\text{Toeplitz} \begin{bmatrix} a_0 & a_{-1} & a_{-2} \\ a_1 & a_0 & a_{-1} \\ a_2 & a_1 & a_0 \end{bmatrix} \qquad \begin{bmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \\ a_2 & a_3 & a_4 \end{bmatrix} \text{Hankel}$$

$$\text{Vandermonde} \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{bmatrix} \qquad \begin{bmatrix} \frac{1}{x_1-y_1} & \frac{1}{x_1-y_2} & \frac{1}{x_1-y_3} \\ \frac{1}{x_2-y_1} & \frac{1}{x_2-y_2} & \frac{1}{x_2-y_3} \\ \frac{1}{x_3-y_1} & \frac{1}{x_3-y_2} & \frac{1}{x_3-y_3} \end{bmatrix} \text{Cauchy}$$

- Features:
  - can be represented by a few parameters: here, $O(n)$.
  - fast multiplication by a vector: $\tilde{O}(n)$ by FFT.
  - related to univariate polynomial arithmetic.

Structured matrices

- Generalizations $\rightarrow$ "still structured but maybe a bit less" $(\star)$
  - $T = [T_{ij}]$ with T_{ij} Toeplitz [block, mosaic, ...]
  - $T_1 T_2$, T^{-1}, Schur complement, ... [Toeplitz-like]
  - ...

Structured matrices

- Generalizations $\rightarrow$ "still structured but maybe a bit less" $(\star)$
  - $T = [T_{ij}]$ with T_{ij} Toeplitz [block, mosaic, ...]
  - $T_1 T_2$, T^{-1}, Schur complement, ... [Toeplitz-like]
  - ...
- **Displacement structure** = a way to formalize $(\star)$:
  - Parametrization size: $O(\alpha n)$ for some $1 \leqslant \alpha \leqslant n$

    $\hookrightarrow$ **a nice continuum of structures.**
  - Algorithms designed in 1980–2000's for AB, A^{-1}, $A^{-1}b$, ... use $\tilde{O}(\alpha^2 n)$ field operations [Morf, Bitmead–Anderson, Gohberg, Kaltofen, Olshevsky, Pan, Shokrollahi, ...]

Structured matrices

- Generalizations $\rightarrow$ "still structured but maybe a bit less" $(\star)$
  - $T = [T_{ij}]$ with T_{ij} Toeplitz [block, mosaic, ...]
  - $T_1 T_2$, T^{-1}, Schur complement, ... [Toeplitz-like]
  - ...
- **Displacement structure** = a way to formalize $(\star)$:
  - Parametrization size: $O(\alpha n)$ for some $1 \leqslant \alpha \leqslant n$

    $\hookrightarrow$ **a nice continuum of structures.**

  - Algorithms designed in 1980–2000's for AB, A^{-1}, $A^{-1}b$, ... use $\tilde{O}(\alpha^2 n)$ field operations [Morf, Bitmead–Anderson, Gohberg, Kaltofen, Olshevsky, Pan, Shokrollahi, ...]

    Great for $\alpha \rightarrow 1$, but not so nice for $\alpha \rightarrow n$, since dense unstructured linear algebra has cost $O(n^\omega)$ with $\omega < 2.38$.

Our objective: replace α^2 by $\alpha^{\omega-1}$ in these cost bounds.

Structured matrices

- Generalizations $\rightarrow$ "still structured but maybe a bit less" $(\star)$
  - $T = [T_{ij}]$ with T_{ij} Toeplitz [block, mosaic, ...]
  - $T_1 T_2$, T^{-1}, Schur complement, ... [Toeplitz-like]
  - ...
- **Displacement structure** = a way to formalize $(\star)$:
  - Parametrization size: $O(\alpha n)$ for some $1 \leqslant \alpha \leqslant n$

    $\hookrightarrow$ **a nice continuum of structures.**

  - Algorithms designed in 1980–2000's for AB, A^{-1}, $A^{-1}b$, ... use $\tilde{O}(\alpha^2 n)$ field operations [Morf, Bitmead–Anderson, Gohberg, Kaltofen, Olshevsky, Pan, Shokrollahi, ...]

    Great for $\alpha \rightarrow 1$, but not so nice for $\alpha \rightarrow n$, since dense unstructured linear algebra has cost $O(n^\omega)$ with $\omega < 2.38$.

Our objective: replace α^2 by $\alpha^{\omega-1}$ in these cost bounds.

Remark: in some important cases, costs in $\tilde{O}(\alpha^{\omega-1} n)$ via σ-bases [Beckermann–Labahn'94, ..., Neiger'16, ...].

Continuums of complexity: basic examples

Block dot products:

- $A, B \in \mathbb{K}^{n\times\alpha} \longrightarrow A^T B$ in $\frac{n}{\alpha}\cdot\alpha^\omega + (\frac{n}{\alpha}-1)\cdot\alpha^2 = O(\alpha^{\omega-1}n)$.
- Cost ranges from $O(n)$ to $O(n^\omega)$.

$$\left[\begin{array}{cc|cc} * & * & * & * \\ * & * & * & * \end{array}\right]\left[\begin{array}{cc} * & * \\ * & * \\ \hline * & * \\ * & * \end{array}\right]$$

Continuums of complexity: basic examples

Block dot products:

- $A, B \in \mathbb{K}^{n\times\alpha} \longrightarrow A^T B$ in $\frac{n}{\alpha}\cdot\alpha^\omega + (\frac{n}{\alpha}-1)\cdot\alpha^2 = O(\alpha^{\omega-1}n)$.
- Cost ranges from $O(n)$ to $O(n^\omega)$.

$$\left[\begin{array}{cc|cc} * & * & * & * \\ * & * & * & * \end{array}\right]\left[\begin{array}{cc} * & * \\ * & * \\ \hline * & * \\ * & * \end{array}\right]$$

Polynomial matrix products:

- $A, B \in \mathbb{K}[X]^{\alpha\times\alpha}_{<n/\alpha} \longrightarrow AB$ in $O\big(\alpha^\omega\cdot \mathsf{M}(n/\alpha)\big) \subset \tilde{O}\big(\alpha^{\omega-1}n\big)$.
- Cost ranges from $\tilde{O}(n)$ to $O(n^\omega)$.

Compression of rank-r matrices:

- $A \in \mathbb{K}^{n\times n} \longrightarrow G, H \in \mathbb{K}^{n\times r}$ s.t. $A = GH^T$ in $O(r^{\omega-2}n^2)$.
- Cost ranges from $O(n^2)$ to $O(n^\omega)$.

$$\begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} = \begin{bmatrix} * & * \\ * & * \\ * & * \\ * & * \end{bmatrix}\begin{bmatrix} * & * & * & * \\ * & * & * & * \end{bmatrix}$$

1. Displacement rank

2. Transformation techniques

3. Structured matrix multiplication

4. Structured matrix inversion

1. Displacement rank

2. Transformation techniques

3. Structured matrix multiplication

4. Structured matrix inversion

Example: Toeplitz structure

$$\underbrace{\begin{bmatrix} a_0 & a_{-1} & a_{-2} \\ a_1 & \color{blue}{a_0} & \color{blue}{a_{-1}} \\ a_2 & \color{blue}{a_1} & \color{blue}{a_0} \end{bmatrix}}_{A} - \underbrace{\begin{bmatrix} & & \\ 1 & & \\ & 1 & \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} \color{blue}{a_0} & \color{blue}{a_{-1}} & a_{-2} \\ \color{blue}{a_1} & \color{blue}{a_0} & a_{-1} \\ a_2 & a_1 & a_0 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} & 1 & \\ & & 1 \\ & & \end{bmatrix}}_{N}$$

Example: Toeplitz structure

$$\underbrace{\begin{bmatrix} a_0 & a_{-1} & a_{-2} \\ a_1 & a_0 & a_{-1} \\ a_2 & a_1 & a_0 \end{bmatrix}}_{A} - \underbrace{\begin{bmatrix} & & \\ 1 & & \\ & 1 & \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} a_0 & a_{-1} & a_{-2} \\ a_1 & a_0 & a_{-1} \\ a_2 & a_1 & a_0 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} & 1 & \\ & & 1 \\ & & \end{bmatrix}}_{N}$$

$$= \begin{bmatrix} a_0 & a_{-1} & a_{-2} \\ a_1 & a_0 & a_{-1} \\ a_2 & a_1 & a_0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & a_0 & a_{-1} \\ 0 & a_1 & a_0 \end{bmatrix}$$

Example: Toeplitz structure

$$\underbrace{\begin{bmatrix} a_0 & a_{-1} & a_{-2} \\ a_1 & a_0 & a_{-1} \\ a_2 & a_1 & a_0 \end{bmatrix}}_{A} - \underbrace{\begin{bmatrix} & & \\ 1 & & \\ & 1 & \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} a_0 & a_{-1} & a_{-2} \\ a_1 & a_0 & a_{-1} \\ a_2 & a_1 & a_0 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} & 1 & \\ & & 1 \\ & & \end{bmatrix}}_{N}$$

$$= \begin{bmatrix} a_0 & a_{-1} & a_{-2} \\ a_1 & a_0 & a_{-1} \\ a_2 & a_1 & a_0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & a_0 & a_{-1} \\ 0 & a_1 & a_0 \end{bmatrix}$$

$$= \begin{bmatrix} a_0 & a_{-1} & a_{-2} \\ a_1 & 0 & 0 \\ a_2 & 0 & 0 \end{bmatrix}$$

Example: Toeplitz structure

$$\underbrace{\begin{bmatrix} a_0 & a_{-1} & a_{-2} \\ a_1 & a_0 & a_{-1} \\ a_2 & a_1 & a_0 \end{bmatrix}}_{A} - \underbrace{\begin{bmatrix} & & \\ 1 & & \\ & 1 & \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} a_0 & a_{-1} & a_{-2} \\ a_1 & a_0 & a_{-1} \\ a_2 & a_1 & a_0 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} & 1 & \\ & & 1 \\ & & \end{bmatrix}}_{N}$$

$$= \begin{bmatrix} a_0 & a_{-1} & a_{-2} \\ a_1 & a_0 & a_{-1} \\ a_2 & a_1 & a_0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & a_0 & a_{-1} \\ 0 & a_1 & a_0 \end{bmatrix}$$

$$= \begin{bmatrix} a_0 & a_{-1} & a_{-2} \\ a_1 & 0 & 0 \\ a_2 & 0 & 0 \end{bmatrix}$$

$\hookrightarrow$ rank $\leqslant 2$ for all n.

$\hookrightarrow$ Stein's displacement operator: $A \mapsto A - MAN$.

Example: Toeplitz structure

$$\underbrace{\begin{bmatrix} a_0 & a_{-1} & a_{-2} \\ a_1 & a_0 & a_{-1} \\ a_2 & a_1 & a_0 \end{bmatrix}}_{A}$$

Example: Toeplitz structure

$$\underbrace{\begin{bmatrix} & & \\ 1 & & \\ & 1 & \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} a_0 & a_{-1} & a_{-2} \\ a_1 & a_0 & a_{-1} \\ a_2 & a_1 & a_0 \end{bmatrix}}_{A}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ a_0 & a_{-1} & a_{-2} \\ a_1 & a_0 & a_{-1} \end{bmatrix}$$

Example: Toeplitz structure

$$\underbrace{\begin{bmatrix} & & \\ 1 & & \\ & 1 & \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} a_0 & a_{-1} & a_{-2} \\ a_1 & a_0 & a_{-1} \\ a_2 & a_1 & a_0 \end{bmatrix}}_{A} - \underbrace{\begin{bmatrix} a_0 & a_{-1} & a_{-2} \\ a_1 & a_0 & a_{-1} \\ a_2 & a_1 & a_0 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} & & 1 \\ 1 & & \\ & 1 & \end{bmatrix}}_{N}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ a_0 & a_{-1} & a_{-2} \\ a_1 & a_0 & a_{-1} \end{bmatrix} - \begin{bmatrix} a_{-1} & a_{-2} & a_0 \\ a_0 & a_{-1} & a_1 \\ a_1 & a_0 & a_2 \end{bmatrix}$$

Example: Toeplitz structure

$$\underbrace{\begin{bmatrix} & & \\ 1 & & \\ & 1 & \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} a_0 & a_{-1} & a_{-2} \\ a_1 & a_0 & a_{-1} \\ a_2 & a_1 & a_0 \end{bmatrix}}_{A} - \underbrace{\begin{bmatrix} a_0 & a_{-1} & a_{-2} \\ a_1 & a_0 & a_{-1} \\ a_2 & a_1 & a_0 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} & & 1 \\ 1 & & \\ & 1 & \end{bmatrix}}_{N}$$

$$= \begin{bmatrix} -a_{-1} & -a_{-2} & -a_0 \\ 0 & 0 & a_{-2} - a_1 \\ 0 & 0 & a_{-1} - a_2 \end{bmatrix}$$

↪ rank $\leqslant 2$ for all n.

↪ Sylvester's displacement operator: $A \mapsto MA - AN$.

Example: Vandermonde structure

$$\underbrace{\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{bmatrix}}_{A}$$

Example: Vandermonde structure

$$\underbrace{\begin{bmatrix} \frac{1}{x_1} & & \\ & \frac{1}{x_2} & \\ & & \frac{1}{x_3} \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{bmatrix}}_{A}$$

$$= \begin{bmatrix} x_1^{-1} & 1 & x_1 \\ x_2^{-1} & 1 & x_2 \\ x_3^{-1} & 1 & x_3 \end{bmatrix}$$

Example: Vandermonde structure

$$\underbrace{\begin{bmatrix} \frac{1}{x_1} & & \\ & \frac{1}{x_2} & \\ & & \frac{1}{x_3} \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{bmatrix}}_{A} - \underbrace{\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} & 1 & \\ & & 1 \\ & & \end{bmatrix}}_{N}$$

$$= \begin{bmatrix} x_1^{-1} & 1 & x_1 \\ x_2^{-1} & 1 & x_2 \\ x_3^{-1} & 1 & x_3 \end{bmatrix} - \begin{bmatrix} 0 & 1 & x_1 \\ 0 & 1 & x_2 \\ 0 & 1 & x_3 \end{bmatrix}$$

Example: Vandermonde structure

$$\underbrace{\begin{bmatrix} \frac{1}{x_1} & & \\ & \frac{1}{x_2} & \\ & & \frac{1}{x_3} \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{bmatrix}}_{A} - \underbrace{\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} & 1 & \\ & & 1 \\ & & \end{bmatrix}}_{N}$$

$$= \begin{bmatrix} x_1^{-1} & 0 & 0 \\ x_2^{-1} & 0 & 0 \\ x_3^{-1} & 0 & 0 \end{bmatrix}$$

$\hookrightarrow$ rank 1 for all n.

General framework [Kailath–Kung–Morf'79]

Definitions:

- displacement: $\nabla_{\mathsf{M},\mathsf{N}}(\mathsf{A}) = \mathsf{MA} - \mathsf{AN}$,
- displacement rank: $\alpha = \mathrm{rank}\,(\nabla_{\mathsf{M},\mathsf{N}}(\mathsf{A}))$,
- generator: G, H such that $\nabla_{\mathsf{M},\mathsf{N}}(\mathsf{A}) = \mathsf{G}\mathsf{H}^T$.

General framework [Kailath–Kung–Morf'79]

Definitions:

- displacement: $\nabla_{\mathsf{M},\mathsf{N}}(\mathsf{A}) = \mathsf{MA} - \mathsf{AN}$,
- displacement rank: $\alpha = \mathrm{rank}\,(\nabla_{\mathsf{M},\mathsf{N}}(\mathsf{A}))$,
- generator: G, H such that $\nabla_{\mathsf{M},\mathsf{N}}(\mathsf{A}) = \mathsf{GH}^T$.

Remarks:

- there exist G and H of dimensions $n \times \alpha$,
- $\mathrm{vec}(\mathsf{GH}^T) = \underbrace{\left(I \otimes \mathsf{M} - \mathsf{N}^T \otimes I\right)}_{\in \mathbb{K}^{n^2 \times n^2}} \mathrm{vec}(\mathsf{A})$

→ compact representation

→ compress, operate, recover

Choices for M and N

- Classically, cyclic shifts and diagonals:

$$Z_\varphi = \begin{bmatrix} & & & \varphi \\ 1 & & & \\ & \ddots & & \\ & & 1 & \end{bmatrix}, \qquad Z_\varphi^T, \qquad D(x) = \begin{bmatrix} x_1 & & & \\ & x_2 & & \\ & & \ddots & \\ & & & x_n \end{bmatrix}$$

Toeplitz-like and Hankel-like when two shifts, Cauchy-like when two diagonals, Vandermonde-like when one of each.

Choices for M and N

- Classically, cyclic shifts and diagonals:

$$\mathsf{Z}_\varphi = \begin{bmatrix} & & & \varphi \\ 1 & & & \\ & \ddots & & \\ & & 1 & \end{bmatrix}, \qquad \mathsf{Z}_\varphi^T, \qquad \mathsf{D}(\mathsf{x}) = \begin{bmatrix} x_1 & & & \\ & x_2 & & \\ & & \ddots & \\ & & & x_n \end{bmatrix}$$

  Toeplitz-like and Hankel-like when two shifts, Cauchy-like when two diagonals, Vandermonde-like when one of each.

- More generally, Jordan forms [Olshevsky–Shokrollahi'99-00] or block-diagonal companion forms [Bostan, J., Mouilleron, Schost'17]:

$$\mathsf{M}_{\mathbf{P}} = \mathrm{diag}(\mathsf{C}_{P_i}), \qquad \mathsf{C}_{P_i} := \text{companion matrix of } P_i \in \mathbb{K}[X],$$

  with the P_i pairwise coprime.

Recovering A from a generator (G, H)

- Let $I \otimes M - N^T \otimes I$ be nonsingular

  $\hookrightarrow$ for example, M invertible and N nilpotent.

- Pre- and post-multiply $MA - AN = GH^T$ by powers of M and N to recover A as

$$A = M^{-1} \cdot \sum_{j=1}^{\alpha} \mathrm{Krylov}(M, G_{*,j})\, \mathrm{Krylov}(N, H_{*,j})^T.$$

[Pan–Wang'03]

Recovering A from a generator (G, H)

- Let $I \otimes \mathsf{M} - \mathsf{N}^T \otimes I$ be nonsingular

  $\hookrightarrow$ for example, M invertible and N nilpotent.

- Pre- and post-multiply $\mathsf{MA} - \mathsf{AN} = \mathsf{GH}^T$ by powers of M and N to recover A as

$$\mathsf{A} = \mathsf{M}^{-1} \cdot \sum_{j=1}^{\alpha} \mathsf{Krylov}(\mathsf{M}, \mathsf{G}_{*,j})\, \mathsf{Krylov}(\mathsf{N}, \mathsf{H}_{*,j})^T.$$

[Pan–Wang'03]

$\Rightarrow$ **fast matrix-vector multiplication**: for our choices of M, N, these Krylov matrices are special enough to allow for Av in $\tilde{O}(\alpha n)$.

Displacement structure after basic operations

Structure can be preserved by

- transposition, inversion: $\alpha \rightsquigarrow \alpha$
- addition, multiplication: $\alpha, \alpha' \rightsquigarrow \leqslant \alpha + \alpha'$
- submatrix extraction and Schur complement: $\alpha \rightsquigarrow \alpha + \varepsilon$, with $\varepsilon \leqslant \mathrm{rank}(\mathsf{M}_{12}) + \mathrm{rank}(\mathsf{N}_{21})$.

Compression of generators: moving from width $O(\alpha)$ to minimal width α can be done in time $O(\alpha^{\omega-1} n)$ via fast LU decomposition.

Illustration: solving structured linear systems

$$\mathsf{A} = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \xrightarrow{\text{compress}} \quad \mathsf{GH}^T = \begin{bmatrix} * & * \\ * & * \\ * & * \\ * & * \\ * & * \end{bmatrix} \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \end{bmatrix}$$

Illustration: solving structured linear systems

$$\mathsf{A} = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \xrightarrow{\text{compress}} \mathsf{GH}^T = \begin{bmatrix} * & * \\ * & * \\ * & * \\ * & * \\ * & * \end{bmatrix} \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \end{bmatrix}$$

$$\Big\downarrow \text{ inversion}$$

$$\mathsf{YZ}^T = \begin{bmatrix} * & * \\ * & * \\ * & * \\ * & * \\ * & * \end{bmatrix} \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \end{bmatrix}$$

Illustration: solving structured linear systems

$$\mathsf{A} = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \xrightarrow{\text{compress}} \mathsf{GH}^T = \begin{bmatrix} * & * \\ * & * \\ * & * \\ * & * \\ * & * \end{bmatrix} \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \end{bmatrix}$$

$$\Big\downarrow \text{ inversion}$$

$$\mathsf{A}^{-1} = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \xleftarrow{\text{recover}} \mathsf{YZ}^T = \begin{bmatrix} * & * \\ * & * \\ * & * \\ * & * \\ * & * \end{bmatrix} \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \end{bmatrix}$$

Illustration: solving structured linear systems

$$\mathsf{A} = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \xrightarrow{\text{compress}} \quad \mathsf{GH}^T = \begin{bmatrix} * & * \\ * & * \\ * & * \\ * & * \\ * & * \end{bmatrix} \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \end{bmatrix}$$

$$\Big\downarrow \text{inversion}$$

$$\mathsf{A}^{-1}b = \begin{bmatrix} * \\ * \\ * \\ * \\ * \end{bmatrix} \quad \xleftarrow{\text{recover}} \quad \mathsf{YZ}^T = \begin{bmatrix} * & * \\ * & * \\ * & * \\ * & * \\ * & * \end{bmatrix} \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \end{bmatrix}$$

1. Displacement rank

2. Transformation techniques

3. Structured matrix multiplication

4. Structured matrix inversion

Example: from Toeplitz-like to Vandermonde-like

Let A be Toeplitz-like with

$$\mathsf{Z}_1\mathsf{A} - \mathsf{A}\mathsf{Z}_0 = \mathsf{G}\mathsf{H}^T, \qquad \mathsf{Z}_\varphi = \begin{bmatrix} & & & \varphi \\ 1 & & & \\ & \ddots & & \\ & & 1 & \end{bmatrix}.$$

- Fourier diagonalizes circulants: $\mathsf{Z}_1 = \mathsf{F}^{-1}\mathsf{D}\mathsf{F}, \quad \mathsf{D} = \mathrm{diag}(\mathsf{F}_2)$.
- Hence $\mathsf{D} \cdot \mathsf{F}\mathsf{A} - \mathsf{F}\mathsf{A} \cdot \mathsf{Z}_0 = (\mathsf{F}\mathsf{G}) \cdot \mathsf{H}^T$ and so FA is Vandermonde-like.

Example: from Toeplitz-like to Vandermonde-like

Let A be Toeplitz-like with

$$\mathsf{Z}_1\mathsf{A} - \mathsf{A}\mathsf{Z}_0 = \mathsf{G}\mathsf{H}^T, \qquad \mathsf{Z}_\varphi = \begin{bmatrix} & & & \varphi \\ 1 & & & \\ & \ddots & & \\ & & 1 & \end{bmatrix}.$$

- Fourier diagonalizes circulants: $\mathsf{Z}_1 = \mathsf{F}^{-1}\mathsf{D}\mathsf{F}$, $\quad \mathsf{D} = \mathrm{diag}(\mathsf{F}_2)$.
- Hence $\mathsf{D} \cdot \mathsf{F}\mathsf{A} - \mathsf{F}\mathsf{A} \cdot \mathsf{Z}_0 = (\mathsf{F}\mathsf{G}) \cdot \mathsf{H}^T$ and so FA is Vandermonde-like.

Overheads:

- to generate FA: left multiply $\mathsf{G} \in \mathbb{K}^{n\times\alpha}$ by F in $\tilde{O}(\alpha n)$;
- to recover Av, multiply the vector (FA)v by F^{-1}, in $\tilde{O}(n)$.

Fast reductions to ∇_{Z_0, Z_1^T}

More generally, multiplicative transforms

$$A \quad \to \quad L\,A\,R$$

to reduce a structure to another one, at negligible cost.

[Pan'90]

Theorem: reductions from ∇_{M_P, N_Q} to simpler ∇_{Z_0, Z_1^T} in

$\tilde{O}(\alpha n)$ for MUL and LINSOLVE,

$\tilde{O}(\alpha n) + O(\alpha^{\omega-1} n)$ for INV.

[Bostan, J., Mouilleron, Schost'17]

Multiplicative transforms for regularizing A

Toeplitz pre-conditioning: [Kaltofen–Saunders'91]

$\widetilde{A} := T_1 A T_2$ with T_1, T_2 unit upper/lower triangular Toeplitz

Multiplicative transforms for regularizing A

Toeplitz pre-conditioning: [Kaltofen–Saunders'91]

$\widetilde{A} := T_1 A T_2$ with T_1, T_2 unit upper/lower triangular Toeplitz

- If A is Toeplitz- or Hankel-like, then so is $\widetilde{A}$, with $\alpha + O(1)$.
- If entries of T_1, T_2 at random from $S \subset \mathbb{K}$, then

$$\widetilde{A} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix}$$ has generic rank profile with proba $\geqslant 1 - \frac{r(r+1)}{|S|}$.

Proof via Schwartz–Zippel.

- Again, low overhead: $\tilde{O}(\alpha n)$.

Multiplicative transforms for regularizing A

Toeplitz pre-conditioning: [Kaltofen–Saunders'91]

$\widetilde{A} := T_1 A T_2$ with T_1, T_2 unit upper/lower triangular Toeplitz

- If A is Toeplitz- or Hankel-like, then so is $\widetilde{A}$, with $\alpha + O(1)$.
- If entries of T_1, T_2 at random from $S \subset \mathbb{K}$, then

$$\widetilde{A} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix}$$ has generic rank profile with proba $\geqslant 1 - \frac{r(r+1)}{|S|}$.

Proof via Schwartz–Zippel.

- Again, low overhead: $\tilde{O}(\alpha n)$.

A variant: from Toeplitz-like to Cauchy-like via two random Vandermonde V_1, V_2. [Hyun–Lebreton–Schost'17]

1. Displacement rank

2. Transformation techniques

3. Structured matrix multiplication

4. Structured matrix inversion

Generating AB from generators of A and B

- Let $A, B \in \mathbb{K}^{n \times n}$ have displacement rank α for some compatible displacement operators:

$$MA - AN = GH^T, \qquad NB - BP = XY^T.$$

Then

$$\begin{aligned} M(AB) - (AB)P &= GH^T B + A\,XY^T \\ &= [G \,|\, A\,X]\,[B^T H \,|\, Y]^T \end{aligned}$$

and it suffices to be able to multiply A and B by α vectors.

Generating AB from generators of A and B

- Let $A, B \in \mathbb{K}^{n \times n}$ have displacement rank α for some compatible displacement operators:

$$MA - AN = GH^T, \qquad NB - BP = XY^T.$$

Then

$$\begin{aligned} M(AB) - (AB)P &= GH^T B + A\, XY^T \\ &= [G \,|\, A\, X]\, [B^T H \,|\, Y]^T \end{aligned}$$

and it suffices to be able to multiply A and B by α vectors.

- Obvious solution has cost $2\alpha \times \tilde{O}(\alpha\, n) \subset \tilde{O}(\alpha^2 n)$.

Incorporating polynomial matrix multiplication

[Bostan, J., Schost'07]

- Rewrite "(reconstruction formula of A) $\times$ (α vectors)" as

$$U^T\big(V\,W^T \bmod X^n\big), \quad U, V, W \in \mathbb{K}[X]^{\alpha}_{<n}.$$

Incorporating polynomial matrix multiplication

[Bostan, J., Schost'07]

- Rewrite "(reconstruction formula of A) $\times$ (α vectors)" as

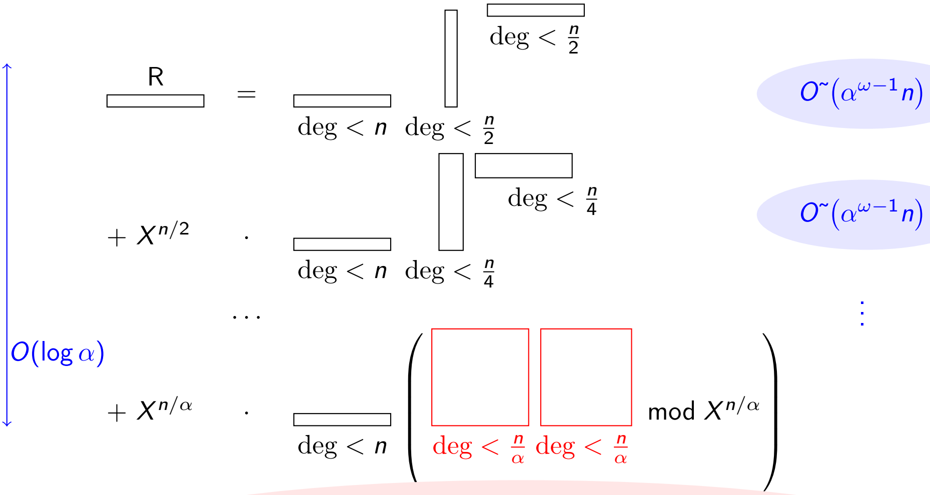
$$U^T\left(V\,W^T \bmod X^n\right), \quad U, V, W \in \mathbb{K}[X]^{\alpha}_{<n}.$$

- Split $V = V_0 + V_1 \cdot X^{n/2}$ and similarly for W:

$$VW^T \bmod X^n = V_0 W_0^T + \left([V_0\ \ V_1][W_1\ \ W_0]^T \bmod X^{n/2}\right) \cdot X^{n/2}$$

  - Outer product has width $\times 2$ and modulus degree $/\,2$.
  - We can continue down to $X^{n/\alpha}$.

Computing $R = U^T(VW^T \bmod X^n)$



+ one product of two polynomial matrices of dimensions $\alpha \times \alpha$ and degree $< \frac{n}{\alpha}$: $O\tilde{}(\alpha^{\omega-1}n)$

1. Displacement rank

2. Transformation techniques

3. Structured matrix multiplication

4. Structured matrix inversion

Block recursive inversion [Strassen'69]

Partition A, assume $\det(A_{11}) \neq 0$, and let $S = A_{22} - A_{21}A_{11}^{-1}A_{12}$:

$$\underbrace{\begin{bmatrix} I & \\ -A_{21}A_{11}^{-1} & I \end{bmatrix}}_{E} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \underbrace{\begin{bmatrix} I & -A_{11}^{-1}A_{12} \\ & I \end{bmatrix}}_{F} = \begin{bmatrix} A_{11} & \\ & S \end{bmatrix}$$

$$\Rightarrow \qquad A^{-1} = F \begin{bmatrix} A_{11}^{-1} & \\ & S^{-1} \end{bmatrix} E.$$

Block recursive inversion [Strassen'69]

Partition A, assume $\det(A_{11}) \neq 0$, and let $S = A_{22} - A_{21}A_{11}^{-1}A_{12}$:

$$\underbrace{\begin{bmatrix} I & \\ -A_{21}A_{11}^{-1} & I \end{bmatrix}}_{E} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \underbrace{\begin{bmatrix} I & -A_{11}^{-1}A_{12} \\ & I \end{bmatrix}}_{F} = \begin{bmatrix} A_{11} & \\ & S \end{bmatrix}$$

$$\Rightarrow \qquad A^{-1} = F \begin{bmatrix} A_{11}^{-1} & \\ & S^{-1} \end{bmatrix} E.$$

- If A is strongly regular, then so are A_{11} and S.
- Reduction to MUL: $C_{\mathrm{INV}}(n) = 2C_{\mathrm{INV}}(n/2) + O(n^\omega) = O(n^\omega)$.

Structured block recursive inversion (MBA)

[Morf/Bitmead–Anderson'80]

Same scheme as Strassen's:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \qquad \Rightarrow \qquad A^{-1} = F \begin{bmatrix} A_{11}^{-1} & \\ & S^{-1} \end{bmatrix} E,$$

but now with every matrix represented by a generator.

Structured block recursive inversion (MBA)

[Morf/Bitmead–Anderson'80]

Same scheme as Strassen's:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \qquad \Rightarrow \qquad A^{-1} = F \begin{bmatrix} A_{11}^{-1} & \\ & S^{-1} \end{bmatrix} E,$$

but now with every matrix represented by a generator.

- Trick: reveal same structure for A_{11} and S as for A.
- Recursion on $n \geqslant \alpha$:

$$C_{\mathsf{INV}}(\alpha, n) = 2C_{\mathsf{INV}}(\alpha, n/2) + O\big(C_{\mathsf{MUL}}(\alpha, n)\big).$$

- Total cost in $\tilde{O}(\alpha^{\omega-1} n)$ thanks to our fast structured MUL.
- Requirement: strongly regularity of A.

Handling arbitrary matrices A [Kaltofen'94-95], [Pan'99]

input: gen(A)
output: "A is singular" or gen(A^{-1})

0. Choose entries of $T_1, T_2 \in \mathbb{K}^{n\times n}$ at random in $S \subset \mathbb{K}$
1. Compute gen($\widetilde{A}$) for $\widetilde{A} = T_1 A T_2$
2. Compute r and gen($\widetilde{A}_r^{-1}$), where
   - $r := \mathrm{rank}(\widetilde{A}) = \mathrm{rank}(A)$,
   - $\widetilde{A}_r$:= largest leading principal submatrix being strongly regular
3. Precond. failed iff Schur complement nonzero.
4. If not failed then
   - 4.1 if $r < n$ then return "A is singular"
   - 4.2 else compute and return gen($T_2\,\widetilde{A}^{-1}\,T_1$).

Handling arbitrary matrices A [Kaltofen'94-95], [Pan'99]

input: gen(A)
output: "A is singular" or $\text{gen}(A^{-1})$

0. Choose entries of $T_1, T_2 \in \mathbb{K}^{n\times n}$ at random in $S \subset \mathbb{K}$
1. Compute $\text{gen}(\widetilde{A})$ for $\widetilde{A} = T_1 A T_2$
2. Compute r and $\text{gen}(\widetilde{A}_r^{-1})$, where
   - $r := \text{rank}(\widetilde{A}) = \text{rank}(A)$,
   - $\widetilde{A}_r :=$ largest leading principal submatrix being strongly regular
3. Precond. failed iff Schur complement nonzero.
4. If not failed then
   - 4.1 if $r < n$ then return "A is singular"
   - 4.2 else compute and return $\text{gen}(T_2\,\widetilde{A}^{-1}\,T_1)$.

- Adapts to LINSOLVE.
- Inherits our previous costs in $\tilde{O}(\alpha^{\omega-1} n)$.
- Probability of failure $< 1/2$ if $|S| \geqslant 2n(n+1)$.

Conclusion

Summary:

- MUL, INV, LINSOLVE in time $\tilde{O}(\alpha^{\omega-1}n)$ for $\nabla_{\mathsf{M_P},\mathsf{N_Q}}$.
- A continuum of cost bounds, from quasi-linear ones to $O(n^\omega)$.
- Generalized operators.

On-going and future work:

- Derandomization of INV and LINSOLVE for finite fields.
- Broader M and N for the same cost.
- Nullspace bases.
- More links with polynomial matrices (Beckermann–Labahn).
- Beyond one-level structures.