Linear matrix equations, a personal journey

Valeria Simoncini

Dipartimento di Matematica
Alma Mater Studiorum - Università di Bologna
`valeria.simoncini@unibo.it`

Let us start from where it all really began

University of Illinois Willard Airport

IATA: CMI · ICAO: KCMI · FAA LID: CMI

Summary	
Airport type	Public
Operator	University of Illinois Urbana-Champaign
Serves	Champaign–Urbana metropolitan area

The next couple of days

Some highlights:

- Yet another party the following night
- Sleeping on a mattress on the floor in Gallopoulos' apartment
- Going for errands during the day in a small ghost town

I started wondering ...

The next couple of days

Some highlights:

- Yet another party the following night
- Sleeping on a mattress on the floor in Gallopoulos' apartment
- Going for errands during the day in a small ghost town

I started wondering ...

The next couple of days

Some highlights:

- Yet another party the following night
- Sleeping on a mattress on the floor in Gallopoulos' apartment
- Going for errands during the day in a small ghost town

I started wondering ...

The next couple of days

Some highlights:

- Yet another party the following night
- Sleeping on a mattress on the floor in Gallopoulos' apartment
- Going for errands during the day in a small ghost town

I started wondering ...

What it really was - Talbot Lab time

We were at the time of CSRD at UIUC

(Center for Supercomputing Research and Development)

- Problems: Large scale
- People: G. Cybenko, D. Padua, C. Polychronopoulos, A. Sameh, etc
- Machines: HPC - CRAY Y-MP C90, Cedar system, CM-5 Connection Machine, etc
- Software: BLAS3, OpenMP

`https://en.wikipedia.org/wiki/University_of_Illinois_Center_for_Supercomputing_Research_and_Development`

Finally some maths

Stratis knew little of my (modest) math background

He came up with two possible research directions:

$$v = \exp(A)b \qquad \text{vs} \qquad AX = B, \quad B = [b_1, \ldots, b_s]$$

Matrix exponential eval'n vs Multiple rhs systems

Finally some maths

Stratis knew little of my (modest) math background

He came up with two possible research directions:

$$v = \exp(A)b \qquad \text{vs} \qquad AX = B, \quad B = [b_1, \ldots, b_s]$$

Matrix exponential eval'n vs Multiple rhs systems

A first trip

Society for Industrial and Applied Mathematics

Fourth SIAM ALA Conference
September 11–14, 1991
Minneapolis, Minnesota

A family ...

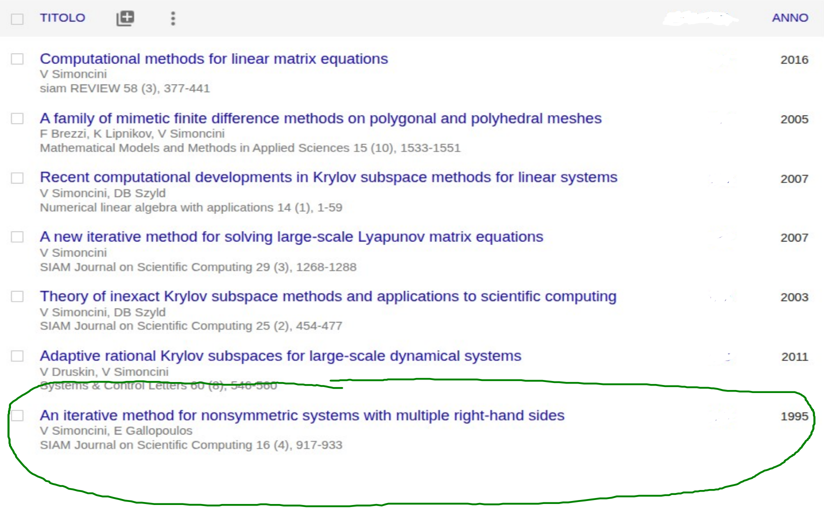
5 25 '92

4 18 '93

... and a warm (greek) community

.... (almost) always partying

But no jokes!



☐	TITOLO	[redacted]	ANNO
☐	Computational methods for linear matrix equations V Simoncini siam REVIEW 58 (3), 377-441	[redacted]	2016
☐	A family of mimetic finite difference methods on polygonal and polyhedral meshes F Brezzi, K Lipnikov, V Simoncini Mathematical Models and Methods in Applied Sciences 15 (10), 1533-1551	[redacted]	2005
☐	Recent computational developments in Krylov subspace methods for linear systems V Simoncini, DB Szyld Numerical linear algebra with applications 14 (1), 1-59	[redacted]	2007
☐	A new iterative method for solving large-scale Lyapunov matrix equations V Simoncini SIAM Journal on Scientific Computing 29 (3), 1268-1288	[redacted]	2007
☐	Theory of inexact Krylov subspace methods and applications to scientific computing V Simoncini, DB Szyld SIAM Journal on Scientific Computing 25 (2), 454-477	[redacted]	2003
☐	Adaptive rational Krylov subspaces for large-scale dynamical systems V Druskin, V Simoncini Systems & Control Letters 60 (8), 546-560	[redacted]	2011
☐	An iterative method for nonsymmetric systems with multiple right-hand sides V Simoncini, E Gallopoulos SIAM Journal on Scientific Computing 16 (4), 917-933	[redacted]	1995

My linear matrix equation journey in a nutshell

a) Linear systems with multiple right-hand sides:

$$A\boldsymbol{X} = B$$

b) Sylvester and Lyapunov equations:

$$A\boldsymbol{X} + \boldsymbol{X}B = C$$

c) Multiterm linear matrix equations:

$$A_1\boldsymbol{X}B_1 + \ldots + A_\ell\boldsymbol{X}B_\ell = C$$

My linear matrix equation journey in a nutshell

a) Linear systems with multiple right-hand sides:

$$A\boldsymbol{X} = B$$

b) Sylvester and Lyapunov equations:

$$A\boldsymbol{X} + \boldsymbol{X}B = C$$

c) Multiterm linear matrix equations:

$$A_1\boldsymbol{X}B_1 + \ldots + A_\ell\boldsymbol{X}B_\ell = C$$

My linear matrix equation journey in a nutshell

a) Linear systems with multiple right-hand sides:

$$A\boldsymbol{X} = B$$

b) Sylvester and Lyapunov equations:

$$A\boldsymbol{X} + \boldsymbol{X}B = C$$

c) Multiterm linear matrix equations:

$$A_1\boldsymbol{X}B_1 + \ldots + A_\ell\boldsymbol{X}B_\ell = C$$

Going from a) to b) to c)

a) to b)

$$A\boldsymbol{X} = B \qquad \Rightarrow \qquad A\boldsymbol{X} + \boldsymbol{X}B = C$$

Challenges of the two-term matrix equation:

- A, B large and sparse $\Rightarrow$ $\boldsymbol{X}$ large and *dense*
- $\boldsymbol{X}$ has entry decay and low-rank properties (to be exploited!)
- No preconditioning strategies

  $\Rightarrow$ Don't change the operator, but enrich the approximation space

Going from a) to b) to c)

b) to c)

$$AX + XB = C \qquad \Rightarrow \qquad A_1 \boldsymbol{X} B_1 + \ldots + A_\ell \boldsymbol{X} B_\ell = C$$

Two quotations:

> ♡ *The problem in its full generality is far from tractable, although the transformation to a matrix-vector equation [...] allows us to use the considerable arsenal of numerical weapons currently avaiable for the solution of such problems.*

Peter Lancaster, SIAM Rev. 1970

> ♡ *Another generalization of the Sylvester equation, mainly of theoretical interest, is* $\sum_i A_i \boldsymbol{X} B_i = C$

Nick Higham, "Accuracy and Stability of Numerical Algorithms", SIAM 1996 (page 322)

Going from a) to b) to c)

b) to c)

$$AX + XB = C \qquad \Rightarrow \qquad A_1 X B_1 + \ldots + A_\ell X B_\ell = C$$

Two quotations:

> ♡ *The problem in its full generality is far from tractable, although the transformation to a matrix-vector equation [...] allows us to use the considerable arsenal of numerical weapons currently avaiable for the solution of such problems.*

Peter Lancaster, SIAM Rev. 1970

> ♡ *Another generalization of the Sylvester equation, mainly of theoretical interest, is* $\sum_i A_i X B_i = C$

Nick Higham, "Accuracy and Stability of Numerical Algorithms", SIAM 1996 (page 322)

Applications

Growing number of different applications:

- Convection-diffusion equations with separable coefficients on polygonal domain
- Tensorized discretizations (spectral methods, IGA, etc.)
- Space-time discretizations of PDEs
- Parameterized PDEs (e.g., with random inputs)
- Constrained optimization
- Control problems
- ...

The problem is the next challenge for system solvers

What to do?

$$A_1 \boldsymbol{X} B_1 + \ldots + A_\ell \boldsymbol{X} B_\ell = C$$

State of the art:

- Kronecker form and back on track: $\mathcal{A}\boldsymbol{x} = c$,

$$(B_1^T \otimes A_1 + \ldots + B_\ell^T \otimes A_\ell)\boldsymbol{x} = c, \qquad B \otimes A = \begin{bmatrix} b_{11}A & b_{12}A & \cdots & b_{1m_B}A \\ b_{21}A & b_{22}A & \cdots & b_{2m_B}A \\ \vdots & & & \vdots \\ b_{n_B1}A & b_{n_B2}A & \cdots & b_{n_Bm_B}A \end{bmatrix} \in \mathbb{C}^{n_An_B \times m_Am_B}$$

  with $c = \mathrm{vec}(C)$
- Fixed point iterations (an “evergreen”...)
- Projection-type methods ⇒ low rank approximation
- Ad-hoc problem-dependent procedures
- ...

What to do?

$$A_1 \boldsymbol{X} B_1 + \ldots + A_\ell \boldsymbol{X} B_\ell = C$$

State of the art:

- Kronecker form and back on track: $\mathcal{A}\boldsymbol{x} = c$,

$$(B_1^T \otimes A_1 + \ldots + B_\ell^T \otimes A_\ell)\boldsymbol{x} = c, \qquad B \otimes A = \begin{bmatrix} b_{11}A & b_{12}A & \cdots & b_{1m_B}A \\ b_{21}A & b_{22}A & \cdots & b_{2m_B}A \\ \vdots & & & \vdots \\ b_{n_B1}A & b_{n_B2}A & \cdots & b_{n_Bm_B}A \end{bmatrix} \in \mathbb{C}^{n_An_B \times m_Am_B}$$

  with $c = \mathrm{vec}(C)$
- Fixed point iterations (an "evergreen"...)
- Projection-type methods $\Rightarrow$ low rank approximation
- Ad-hoc problem-dependent procedures
- ...

Matrix-oriented recurrences - CG with truncation

⋆ **Matricization.** Typically,

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)} \in \mathbb{R}^{n^2} \qquad \Rightarrow \qquad X^{(k+1)} = X^{(k)} + \alpha_k P^{(k)} \in \mathbb{R}^{n\times n}$$

⋆ **Truncation.** If $X^{(k)} = X_1^{(k)}(X_1^{(k)})^\top$ with $X_1^{(k)}$ low rank, and similarly for $P^{(k)}$, then

$$X^{(k+1)} = X_1^{(k)}(X_1^{(k)})^\top + \alpha_k P_1^{(k)}(P_1^{(k)})^\top$$

- $X^{(k+1)}$ low rank:

$$X^{(k+1)} = [X_1^{(k)}, \sqrt{\alpha_k}P_1^{(k)}]\,[X_1^{(k)}, \sqrt{\alpha_k}P_1^{(k)}]^\top \qquad (1)$$

  (but generally larger than at iteration k)

- Cure: Rank shrinking $[X_1^{(k)}, \sqrt{\alpha_k}P_1^{(k)}] \Rightarrow X_1^{(k+1)}$ $\quad X^{(k+1)} \approx X_1^{(k+1)}(X_1^{(k+1)})^\top$

  Implementation: $\mathcal{T}(X^{(k+1)})$ acts on the QR-SVD of factor in (1)

Alternative truncation criteria:
♣ Fix lower threshold tolerance ♣ Fix maximum allowed rank

Matrix-oriented recurrences - CG with truncation

⋆ **Matricization.** Typically,

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)} \in \mathbb{R}^{n^2} \qquad \Rightarrow \qquad X^{(k+1)} = X^{(k)} + \alpha_k P^{(k)} \in \mathbb{R}^{n\times n}$$

⋆ **Truncation.** If $X^{(k)} = X_1^{(k)}(X_1^{(k)})^\top$ with $X_1^{(k)}$ low rank, and similarly for $P^{(k)}$, then

$$X^{(k+1)} = X_1^{(k)}(X_1^{(k)})^\top + \alpha_k P_1^{(k)}(P_1^{(k)})^\top$$

- $X^{(k+1)}$ low rank:

  $$X^{(k+1)} = [X_1^{(k)}, \sqrt{\alpha_k}P_1^{(k)}]\,[X_1^{(k)}, \sqrt{\alpha_k}P_1^{(k)}]^\top \qquad (1)$$

  (but generally larger than at iteration k)
- Cure: Rank shrinking $[X_1^{(k)}, \sqrt{\alpha_k}P_1^{(k)}] \Rightarrow X_1^{(k+1)}$ $\quad X^{(k+1)} \approx X_1^{(k+1)}(X_1^{(k+1)})^\top$

  Implementation: $\mathcal{T}(X^{(k+1)})$ acts on the QR-SVD of factor in (1)

Alternative truncation criteria:

♣ Fix lower threshold tolerance ♣ Fix maximum allowed rank

Matrix-oriented recurrences - CG with truncation

⋆ **Matricization.** Typically,

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)} \in \mathbb{R}^{n^2} \qquad \Rightarrow \qquad X^{(k+1)} = X^{(k)} + \alpha_k P^{(k)} \in \mathbb{R}^{n\times n}$$

⋆ **Truncation.** If $X^{(k)} = X_1^{(k)}(X_1^{(k)})^\top$ with $X_1^{(k)}$ low rank, and similarly for $P^{(k)}$, then

$$X^{(k+1)} = X_1^{(k)}(X_1^{(k)})^\top + \alpha_k P_1^{(k)}(P_1^{(k)})^\top$$

- ▶ $X^{(k+1)}$ low rank:

$$X^{(k+1)} = [X_1^{(k)}, \sqrt{\alpha_k}P_1^{(k)}]\,[X_1^{(k)}, \sqrt{\alpha_k}P_1^{(k)}]^\top \tag{1}$$

(but generally larger than at iteration k)

- ▶ Cure: Rank shrinking $[X_1^{(k)}, \sqrt{\alpha_k}P_1^{(k)}] \Rightarrow X_1^{(k+1)}$ $\quad X^{(k+1)} \approx X_1^{(k+1)}(X_1^{(k+1)})^\top$

Implementation: $\mathcal{T}(X^{(k+1)})$ acts on the QR-SVD of factor in (1)

Alternative truncation criteria:

♣ Fix lower threshold tolerance ♣ Fix maximum allowed rank

Matrix-oriented recurrences - CG with truncation

⋆ **Matricization.** Typically,

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)} \in \mathbb{R}^{n^2} \qquad \Rightarrow \qquad X^{(k+1)} = X^{(k)} + \alpha_k P^{(k)} \in \mathbb{R}^{n\times n}$$

⋆ **Truncation.** If $X^{(k)} = X_1^{(k)}(X_1^{(k)})^\top$ with $X_1^{(k)}$ low rank, and similarly for $P^{(k)}$, then

$$X^{(k+1)} = X_1^{(k)}(X_1^{(k)})^\top + \alpha_k P_1^{(k)}(P_1^{(k)})^\top$$

- $X^{(k+1)}$ low rank:

$$X^{(k+1)} = [X_1^{(k)}, \sqrt{\alpha_k}P_1^{(k)}]\ [X_1^{(k)}, \sqrt{\alpha_k}P_1^{(k)}]^\top \tag{1}$$

  (but generally larger than at iteration k)

- Cure: Rank shrinking $[X_1^{(k)}, \sqrt{\alpha_k}P_1^{(k)}] \Rightarrow X_1^{(k+1)}$ $\quad X^{(k+1)} \approx X_1^{(k+1)}(X_1^{(k+1)})^\top$

  Implementation: $\mathcal{T}(X^{(k+1)})$ acts on the QR-SVD of factor in (1)

Alternative truncation criteria:

♣ Fix lower threshold tolerance ♣ Fix maximum allowed rank

Matrix-oriented recurrences - CG with truncation

⋆ **Matricization.** Typically,

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)} \in \mathbb{R}^{n^2} \qquad \Rightarrow \qquad X^{(k+1)} = X^{(k)} + \alpha_k P^{(k)} \in \mathbb{R}^{n\times n}$$

⋆ **Truncation.** If $X^{(k)} = X_1^{(k)}(X_1^{(k)})^\top$ with $X_1^{(k)}$ low rank, and similarly for $P^{(k)}$, then

$$X^{(k+1)} = X_1^{(k)}(X_1^{(k)})^\top + \alpha_k P_1^{(k)}(P_1^{(k)})^\top$$

- $X^{(k+1)}$ low rank:

$$X^{(k+1)} = [X_1^{(k)}, \sqrt{\alpha_k}P_1^{(k)}]\,[X_1^{(k)}, \sqrt{\alpha_k}P_1^{(k)}]^\top \qquad (1)$$

  (but generally larger than at iteration k)

- Cure: Rank shrinking $[X_1^{(k)}, \sqrt{\alpha_k}P_1^{(k)}] \Rightarrow X_1^{(k+1)}$ $\quad X^{(k+1)} \approx X_1^{(k+1)}(X_1^{(k+1)})^\top$

  Implementation: $\mathcal{T}(X^{(k+1)})$ acts on the QR-SVD of factor in (1)

Alternative truncation criteria:

♣ Fix lower threshold tolerance ♣ Fix maximum allowed rank

Truncated matrix-oriented CG (TCG) for Kronecker form

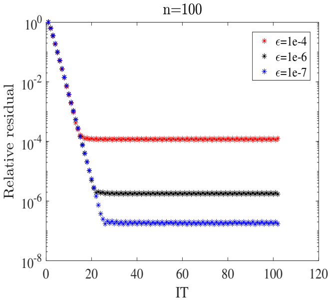
Input: $\mathcal{L}(\boldsymbol{X}) = A_1\boldsymbol{X}B_1 + A_2\boldsymbol{X}B_2 + \ldots + A_\ell\boldsymbol{X}B_\ell$, right-hand side $C \in \mathbb{R}^{n\times n}$ in low-rank format.
Truncation operator $\mathcal{T}$.
Output: Matrix $X \in \mathbb{R}^{n\times n}$ in low-rank format s.t. $||\mathcal{L}(X) - C||_F/||C||_F \le tol$

```
$X_0 = 0,\ R_0 = C,\ P_0 = R_0,\ Q_0 = \mathcal{L}(P_0)$
$\xi_0 = \langle P_0, Q_0\rangle,\ k = 0$                         $\langle X, Y\rangle = \mathrm{tr}(X^\top Y)$
While $||R_k||_F > tol$
    $\alpha_k = \langle R_k, P_k\rangle/\xi_k$
    $X_{k+1} = X_k + \alpha_k P_k,$              $X_{k+1} \leftarrow \mathcal{T}(X_{k+1})$
    $R_{k+1} = C - \mathcal{L}(X_{k+1}),$        Optionally:  $R_{k+1} \leftarrow \mathcal{T}(R_{k+1})$
    $\beta_k = -\langle R_{k+1}, Q_k\rangle/\xi_k$
    $P_{k+1} = R_{k+1} + \beta_k P_k,$           $P_{k+1} \leftarrow \mathcal{T}(P_{k+1})$
    $Q_{k+1} = \mathcal{L}(P_{k+1}),$            Optionally:  $Q_{k+1} \leftarrow \mathcal{T}(Q_{k+1})$
    $\xi_{k+1} = \langle P_{k+1}, Q_{k+1}\rangle$
    $k = k + 1$
end while
```

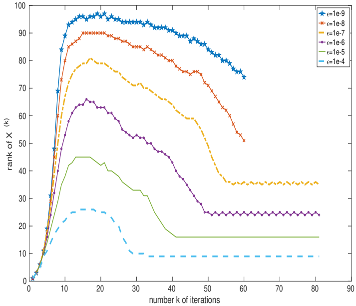
♣ Iterates kept in factored form!

Kressner and Tobler, '11

Typical convergence and rank behavior



(Hao, '20, personal comm.)



(Simoncini & Hao, '22)

Considerations:

1. At best, convergence as for Kronecker problem
2. Rank of iterates hard to control to maintain convergence
3. Coeffs α, β under exploited

Back to the roots

Consider

$$\mathcal{A}\boldsymbol{x} = c \qquad \mathcal{A} \quad \text{nonsing, nonsym}$$

One-dimensional projection method: Saad, "Iterative methods for sparse linear systems", SIAM, 2003

$$x_{k+1} = x_k + \alpha_k r_k, \quad r_{k+1} = b - \mathcal{A}x_{k+1},$$

For instance, $\alpha_k := \frac{(\mathcal{A}r_k)^T r_k}{(\mathcal{A}r_k)^T \mathcal{A}r_k}$ minimizes $\|c - \mathcal{A}(x_k + \alpha r_k)\|^2$

Consider $A_1 \boldsymbol{X} B_1 + \ldots + A_\ell \boldsymbol{X} B_\ell = C$ with

$$\mathcal{L}(V) := A_1 V B_1 + \ldots + A_\ell V B_\ell$$

"One-dimensional" projection method:

$$\begin{aligned} \boldsymbol{X}_{k+1} &= \boldsymbol{X}_k + R_k^{(l)} \boldsymbol{\alpha}_k (R_k^{(r)})^T \qquad && \boldsymbol{\alpha}_k \in \mathbb{R}^{q_k \times q_k} \\ \boldsymbol{R}_{k+1} &= CD^T - \mathcal{L}(\boldsymbol{X}_{k+1}), \qquad && \boldsymbol{R}_{k+1} =: R_{k+1}^{(l)} (R_{k+1}^{(r)})^T, \end{aligned}$$

At step k, $\boldsymbol{\alpha}_k$ is chosen to minimize the Frobenius norm of the residual, namely

$$\min_{\boldsymbol{\alpha} \in \mathbb{R}^{q_k \times q_k}} \|CD^T - \mathcal{L}(\boldsymbol{X}_k + R_k^{(l)} \boldsymbol{\alpha} (R_k^{(r)})^T)\|_F^2$$

Back to the roots

Consider

$$\mathcal{A}\boldsymbol{x} = c \qquad \mathcal{A} \quad \text{nonsing, nonsym}$$

One-dimensional projection method: Saad, "Iterative methods for sparse linear systems", SIAM, 2003

$$x_{k+1} = x_k + \alpha_k r_k, \quad r_{k+1} = b - \mathcal{A}x_{k+1},$$

For instance, $\alpha_k := \frac{(\mathcal{A}r_k)^T r_k}{(\mathcal{A}r_k)^T \mathcal{A}r_k}$ minimizes $\|c - \mathcal{A}(x_k + \alpha r_k)\|^2$

Consider $A_1 \boldsymbol{X} B_1 + \ldots + A_\ell \boldsymbol{X} B_\ell = C$ with

$$\mathcal{L}(V) := A_1 V B_1 + \ldots + A_\ell V B_\ell$$

"One-dimensional" projection method:

$$\begin{aligned} \boldsymbol{X}_{k+1} &= \boldsymbol{X}_k + R_k^{(l)} \boldsymbol{\alpha}_k (R_k^{(r)})^T \qquad && \boldsymbol{\alpha}_k \in \mathbb{R}^{q_k \times q_k} \\ \boldsymbol{R}_{k+1} &= CD^T - \mathcal{L}(\boldsymbol{X}_{k+1}), \qquad && \boldsymbol{R}_{k+1} =: R_{k+1}^{(l)} (R_{k+1}^{(r)})^T, \end{aligned}$$

At step k, α_k is chosen to minimize the Frobenius norm of the residual, namely

$$\min_{\boldsymbol{\alpha} \in \mathbb{R}^{q_k \times q_k}} \|CD^T - \mathcal{L}(\boldsymbol{X}_k + R_k^{(l)} \alpha (R_k^{(r)})^T)\|_F^2$$

Back to the roots

Consider

$$\mathcal{A}\boldsymbol{x} = c \qquad \mathcal{A} \quad \text{nonsing, nonsym}$$

One-dimensional projection method: Saad, "Iterative methods for sparse linear systems", SIAM, 2003

$$x_{k+1} = x_k + \alpha_k r_k, \quad r_{k+1} = b - \mathcal{A}x_{k+1},$$

For instance, $\alpha_k := \frac{(\mathcal{A}r_k)^T r_k}{(\mathcal{A}r_k)^T \mathcal{A}r_k}$ minimizes $\|c - \mathcal{A}(x_k + \alpha r_k)\|^2$

Consider $A_1 \boldsymbol{X} B_1 + \ldots + A_\ell \boldsymbol{X} B_\ell = C$ with

$$\mathcal{L}(V) := A_1 V B_1 + \ldots + A_\ell V B_\ell$$

"One-dimensional" projection method:

$$\begin{aligned} \boldsymbol{X}_{k+1} &= \boldsymbol{X}_k + R_k^{(l)} \boldsymbol{\alpha}_k (R_k^{(r)})^T \qquad \boldsymbol{\alpha}_k \in \mathbb{R}^{q_k \times q_k} \\ \boldsymbol{R}_{k+1} &= CD^T - \mathcal{L}(\boldsymbol{X}_{k+1}), \qquad \boldsymbol{R}_{k+1} =: R_{k+1}^{(l)} (R_{k+1}^{(r)})^T, \end{aligned}$$

At step k, $\boldsymbol{\alpha}_k$ is chosen to minimize the Frobenius norm of the residual, namely

$$\min_{\boldsymbol{\alpha} \in \mathbb{R}^{q_k \times q_k}} \|CD^T - \mathcal{L}(\boldsymbol{X}_k + R_k^{(l)} \boldsymbol{\alpha} (R_k^{(r)})^T)\|_F^2$$

Low-rank short matrix recurrences

$$\begin{aligned} \boldsymbol{X}_{k+1} &= \boldsymbol{X}_k + R_k^{(l)} \boldsymbol{\alpha}_k (R_k^{(r)})^T \\ \boldsymbol{R}_{k+1} &= CD^T - \mathcal{L}(\boldsymbol{X}_{k+1}), \qquad \boldsymbol{R}_{k+1} =: R_{k+1}^{(l)} (R_{k+1}^{(r)})^T, \end{aligned}$$

with $\boldsymbol{\alpha}_k$ such that $\qquad (\boldsymbol{R}_k = R_k^{(l)} (R_k^{(r)})^T)$

$$(R_k^{(l)})^T \mathcal{L}^* \left(\mathcal{L}(R_k^{(l)} \boldsymbol{\alpha} (R_k^{(r)})^T) \right) R_k^{(r)} = (R_k^{(l)})^T \mathcal{L}^* (\boldsymbol{R}_k) R_k^{(r)}.$$

Moreover, $\mathrm{vec}(\boldsymbol{R}_{k+1}) \perp \mathcal{A} \cdot \mathrm{Range}(R_k^{(r)} \otimes R_k^{(l)})$

Additional ingredients:

- Rank truncation
- Preconditioning
- Randomized strategies

Low-rank short matrix recurrences

$$\begin{aligned} \boldsymbol{X}_{k+1} &= \boldsymbol{X}_k + R_k^{(l)} \boldsymbol{\alpha}_k (R_k^{(r)})^T \\ \boldsymbol{R}_{k+1} &= CD^T - \mathcal{L}(\boldsymbol{X}_{k+1}), \qquad \boldsymbol{R}_{k+1} =: R_{k+1}^{(l)} (R_{k+1}^{(r)})^T, \end{aligned}$$

with $\boldsymbol{\alpha}_k$ such that $\qquad (\boldsymbol{R}_k = R_k^{(l)} (R_k^{(r)})^T)$

$$(R_k^{(l)})^T \mathcal{L}^* \left(\mathcal{L}(R_k^{(l)} \boldsymbol{\alpha} (R_k^{(r)})^T) \right) R_k^{(r)} = (R_k^{(l)})^T \mathcal{L}^*(\boldsymbol{R}_k) R_k^{(r)}.$$

Moreover, $\mathrm{vec}(\boldsymbol{R}_{k+1}) \perp \mathcal{A} \cdot \mathrm{Range}(R_k^{(r)} \otimes R_k^{(l)})$

Additional ingredients:

- Rank truncation
- Preconditioning
- Randomized strategies

Scientific Summary

$$A_1 \boldsymbol{X} B_1 + \ldots + A_\ell \boldsymbol{X} B_\ell = C$$

- Matrix short recurrences are a viable and effective strategy
- Caveat: Requires low-rank C and low-rank approximability of $\boldsymbol{X}$
- Tensor version is also possible

REFERENCES

[1] D. Palitta, M. Iannacito, and V. S., *A subspace-conjugate gradient method for linear matrix equations* SIMAX, 46 (2025)

[2] M. Iannacito, L. Piccinini, and V. S., *Subspace gradient descent methods for linear tensor equations*, ArXiv 2602.21974.

[3] D. Palitta, C. E. Powell, and V. S., *A class of low-rank short recurrences for nonsymmetric linear matrix equations*, ArXiv 2605.01276.

Scientific Summary

$$A_1 \boldsymbol{X} B_1 + \ldots + A_\ell \boldsymbol{X} B_\ell = C$$

- Matrix short recurrences are a viable and effective strategy
- Caveat: Requires low-rank C and low-rank approximability of $\boldsymbol{X}$
- Tensor version is also possible

REFERENCES

[1] D. Palitta, M. Iannacito, and V. S., *A subspace-conjugate gradient method for linear matrix equations* SIMAX, 46 (2025)

[2] M. Iannacito, L. Piccinini, and V. S., *Subspace gradient descent methods for linear tensor equations*, ArXiv 2602.21974.

[3] D. Palitta, C. E. Powell, and V. S., *A class of low-rank short recurrences for nonsymmetric linear matrix equations*, ArXiv 2605.01276.

My first Householder symposium

Thanks Strati