On the truncated conjugate gradient method for linear matrix equations

Valeria Simoncini

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

Multiterm linear matrix equation

$$A_1 \boldsymbol{X} B_1 + A_2 \boldsymbol{X} B_2 + \ldots + A_\ell \boldsymbol{X} B_\ell = C$$

$A_i \in \mathbb{R}^{n\times n}$, $B_i \in \mathbb{R}^{m\times m}$, $\boldsymbol{X}$ unknown matrix

Possibly large dimensions, structured coefficient matrices

> *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 available for the solution of such problems.*
>
> *Peter Lancaster, SIAM Rev. 1970*

Multiterm linear matrix equation

$$A_1 \boldsymbol{X} B_1 + A_2 \boldsymbol{X} B_2 + \ldots + A_\ell \boldsymbol{X} B_\ell = C$$

- Kronecker form and back on track
- Fixed point iterations (an "evergreen"...)
- Projection-type methods ⇒ low rank approximation
- Ad-hoc problem-dependent procedures
- etc.

A sample of these methodologies on different problems:

- Stochastic PDEs
- PDEs on polygonal domains, IGA, spectral methods, etc
- Space-time PDEs
- All-at-once PDE-constrained optimization problem
- Bilinear control problems
- ....

Multiterm linear matrix equation

$$A_1 \boldsymbol{X} B_1 + A_2 \boldsymbol{X} B_2 + \ldots + A_\ell \boldsymbol{X} B_\ell = C$$

- Kronecker form and back on track
- Fixed point iterations (an "evergreen"...)
- Projection-type methods $\Rightarrow$ low rank approximation
- Ad-hoc problem-dependent procedures
- etc.

A sample of these methodologies on different problems:

- Stochastic PDEs
- PDEs on polygonal domains, IGA, spectral methods, etc
- Space-time PDEs
- All-at-once PDE-constrained optimization problem
- Bilinear control problems
- ....

Multiterm linear matrix equation. Classical device

$$A_1 \boldsymbol{X} B_1 + A_2 \boldsymbol{X} B_2 + \ldots + A_\ell \boldsymbol{X} B_\ell = C$$

Kronecker formulation $\boxed{\left(B_1^\top \otimes A_1 + \ldots + B_\ell^\top \otimes A_\ell\right) \boldsymbol{x} = c \;\Leftrightarrow\; \mathcal{A}\boldsymbol{x} = c}$

Iterative methods: matrix-matrix multiplications and rank truncation

(Benner, Breiten, Bouhamidi, Chehab, Damm, Grasedyck, Jbilou, Kressner, Kuerschner, Matthies, Nagy, Palitta, Onwunta, Raydan, Stoll, Tobler, Wedderburn, Zander, ...)

$$\text{Kronecker product}: \quad M \otimes P = \begin{bmatrix} m_{11}P & \ldots & m_{1n}P \\ \vdots & \ddots & \vdots \\ m_{n1}P & \ldots & m_{nn}P \end{bmatrix} \qquad \text{and} \qquad \mathrm{vec}(AXB) = (B^\top \otimes A)\mathrm{vec}(X)$$

Current very active area of research

Multiterm linear matrix equation. Classical device

$$A_1 \boldsymbol{X} B_1 + A_2 \boldsymbol{X} B_2 + \ldots + A_\ell \boldsymbol{X} B_\ell = C$$

Kronecker formulation $\boxed{\left(B_1^\top \otimes A_1 + \ldots + B_\ell^\top \otimes A_\ell\right) \boldsymbol{x} = c \;\Leftrightarrow\; \mathcal{A}\boldsymbol{x} = c}$

Iterative methods: matrix-matrix multiplications and rank truncation

(Benner, Breiten, Bouhamidi, Chehab, Damm, Grasedyck, Jbilou, Kressner, Kuerschner, Matthies, Nagy, Palitta, Onwunta, Raydan, Stoll, Tobler, Wedderburn, Zander, ...)

$$\text{Kronecker product}: \quad M \otimes P = \begin{bmatrix} m_{11}P & \ldots & m_{1n}P \\ \vdots & \ddots & \vdots \\ m_{n1}P & \ldots & m_{nn}P \end{bmatrix} \quad \text{and} \quad \mathrm{vec}(AXB) = (B^\top \otimes A)\mathrm{vec}(X)$$

Current very active area of research

CG matricization and 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:

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

(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 SVD of $X^{(k+1)}$

Alternative truncation criteria:

♣ Fix lower threshold tolerance ♣ Fix maximum allowed rank

CG matricization and 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:

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

  (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 SVD of $X^{(k+1)}$

Alternative truncation criteria:

♣ Fix lower threshold tolerance ♣ Fix maximum allowed rank

CG matricization and 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:

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

(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 SVD of $X^{(k+1)}$

Alternative truncation criteria:

♣ Fix lower threshold tolerance ♣ Fix maximum allowed rank

CG matricization and 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:

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

  (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 SVD of $X^{(k+1)}$

Alternative truncation criteria:

♣ Fix lower threshold tolerance ♣ Fix maximum allowed rank

CG matricization and 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:

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

(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 SVD of $X^{(k+1)}$

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. $\|\mathcal{L}(X) - C\|_F / \|C\|_F \leq tol$

$X_0 = 0,\ R_0 = C,\ P_0 = R_0,\ Q_0 = \mathcal{L}(P_0)$

$\xi_0 = \langle P_0, Q_0 \rangle,\ k = 0$ $\qquad\qquad \langle X, Y \rangle = \mathrm{tr}(X^\top Y)$

While $\|R_k\|_F > tol$

  $\omega_k = \langle R_k, P_k \rangle / \xi_k$

  $X_{k+1} = X_k + \omega_k P_k,$ $\qquad X_{k+1} \leftarrow \mathcal{T}(X_{k+1})$

  $R_{k+1} = C - \mathcal{L}(X_{k+1}),$ $\qquad$ Optionally: $R_{k+1} \leftarrow \mathcal{T}(R_{k+1})$

  $\beta_k = -\langle R_{k+1}, Q_k \rangle / \xi_k$

  $P_{k+1} = R_{k+1} + \beta_k P_k,$ $\qquad P_{k+1} \leftarrow \mathcal{T}(P_{k+1})$

  $Q_{k+1} = \mathcal{L}(P_{k+1}),$ $\qquad$ 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! $\qquad$ Kressner and Tobler, '11

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. $||\mathcal{L}(X) - C||_F/||C||_F \leq tol$

1. $X_0 = 0,\ R_0 = C,\ P_0 = R_0,\ Q_0 = \mathcal{L}(P_0)$
2. $\xi_0 = \langle P_0, Q_0\rangle,\ k = 0$ $\qquad\qquad \langle X, Y\rangle = \mathrm{tr}(X^\top Y)$
3. While $||R_k||_F > tol$
4. $\qquad \omega_k = \langle R_k, P_k\rangle/\xi_k$
5. $\qquad X_{k+1} = X_k + \omega_k P_k,$ $\qquad X_{k+1} \leftarrow \mathcal{T}(X_{k+1})$
6. $\qquad R_{k+1} = C - \mathcal{L}(X_{k+1}),$ $\qquad$ Optionally: $R_{k+1} \leftarrow \mathcal{T}(R_{k+1})$
7. $\qquad \beta_k = -\langle R_{k+1}, Q_k\rangle/\xi_k$
8. $\qquad P_{k+1} = R_{k+1} + \beta_k P_k,$ $\qquad P_{k+1} \leftarrow \mathcal{T}(P_{k+1})$
9. $\qquad Q_{k+1} = \mathcal{L}(P_{k+1}),$ $\qquad$ Optionally: $Q_{k+1} \leftarrow \mathcal{T}(Q_{k+1})$
10. $\qquad \xi_{k+1} = \langle P_{k+1}, Q_{k+1}\rangle$
11. $\qquad k = k + 1$
12. end while

♣ Iterates kept in factored form! $\qquad$ Kressner and Tobler, '11

Exploring the iteration matrices

CG minimizes error in the energy norm $\Rightarrow$ in the $\|X\|_{\mathcal{L}}$ norm

$\|X\|_{\mathcal{L}}^2 := \mathrm{trace}\left(X^\top \mathcal{L}(X)\right)$ with $\mathcal{L}(\boldsymbol{X}) = A_1 \boldsymbol{X} B_1 + A_2 \boldsymbol{X} B_2 + \ldots + A_\ell \boldsymbol{X} B_\ell$

Numerical evidence: As TCG iterations proceed

- Singular triplets of $X^{(k)}$ seem to converge in an orderly fashion to those of $X^\star$
- The numerical rank of $X^{(k)}$ increases up to some point, then it decreases

(Kressner, Plesinger& Tobler, '14)

Exploring the iteration matrices

CG minimizes error in the energy norm $\Rightarrow$ in the $\|X\|_{\mathcal{L}}$ norm

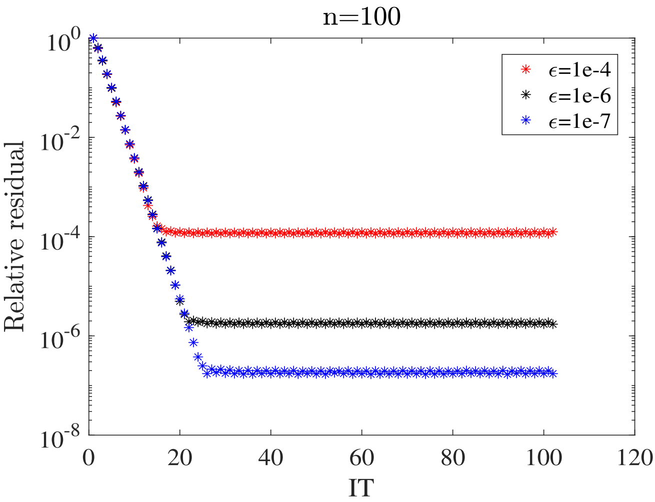
$\|X\|^2_{\mathcal{L}} := \mathrm{trace}\left(X^\top \mathcal{L}(X)\right)$ with $\mathcal{L}(\boldsymbol{X}) = A_1\boldsymbol{X}B_1 + A_2\boldsymbol{X}B_2 + \ldots + A_\ell\boldsymbol{X}B_\ell$

Numerical evidence: As TCG iterations proceed

- Singular triplets of $X^{(k)}$ seem to converge in an orderly fashion to those of $X^\star$
- The numerical rank of $X^{(k)}$ increases up to some point, then it decreases

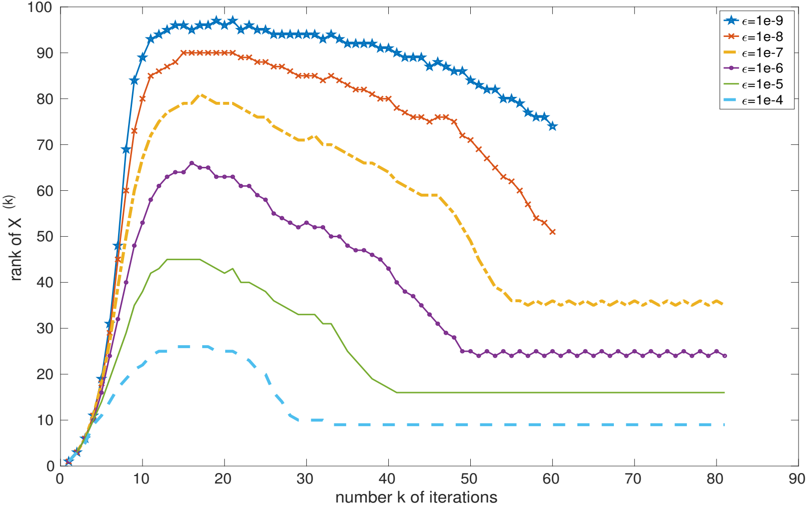
(Kressner, Plesinger& Tobler, '14)

Typical convergence behavior



(Hao, '20, personal comm.)

Typical iterate rank behavior



(Simoncini & Hao, '22, also main reference for the following)

Towards an understanding.

Outline of our findings:

- No theoretical ground for a low rank throughout
- Approximation space wildly affected by truncation
- Loss of orthogonality properties leads to “relaxed Krylov” framework

Towards an understanding. A skinny setting. 1

Consider

$$AX + XA + MXM = C$$

with $A, M \in \mathbb{R}^{n\times n}$ spd, $C \in \mathbb{R}^{n\times n}$ sym. low rank.

Note: This is equivalent to $B^T YD + D^T YB + Y = F$ in the unknown Y

Some properties:

- Solution $X^\star \in \mathbb{R}^{n\times n}$ is symmetric
- Assuming $|\lambda_1| \geq \ldots \geq |\lambda_n|$,

$$|\lambda_{m+1}| = \min_{\substack{X\in\mathbb{R}^{n\times n}\\ \mathrm{rank}(X)=m}} \|X^\star - X\| \leq \|X^\star - \widetilde{X}\|$$

for any $\widetilde{X}$ rank-m symmetric approximation to $X^\star$

Towards an understanding. A skinny setting. 1

Consider

$$AX + XA + MXM = C$$

with $A, M \in \mathbb{R}^{n\times n}$ spd, $C \in \mathbb{R}^{n\times n}$ sym. low rank.

Note: This is equivalent to $B^T YD + D^T YB + Y = F$ in the unknown Y

Some properties:

- Solution $X^\star \in \mathbb{R}^{n\times n}$ is symmetric
- Assuming $|\lambda_1| \geq \ldots \geq |\lambda_n|$,

$$|\lambda_{m+1}| = \min_{\substack{X\in\mathbb{R}^{n\times n}\\ \mathrm{rank}(X)=m}} \|X^\star - X\| \leq \|X^\star - \widetilde{X}\|$$

for any $\widetilde{X}$ rank-m symmetric approximation to $X^\star$

Towards an understanding. A skinny setting. 2

$$|\lambda_{m+1}| = \min_{\substack{X\in\mathbb{R}^{n\times n}\\ \mathrm{rank}(X)=m}} \|X^\star - X\| \le \|X^\star - \widetilde{X}\| \qquad (*)$$

- The error norm $\|X^\star - \widetilde{X}\|$ in $(*)$ provides a, not necessarily sharp, upper bound for the $(m+1)$st singular value of $X^\star$
  (Penzl, '00)

- Given a rank-m matrix $\widetilde{X}$, $(*)$ indicates that the error norm $\|X^\star - \widetilde{X}\|$ cannot go below $|\lambda_{m+1}|$

We assume $X^\star$ can be well approximated by a low rank matrix

(see, e.g., Benner & Breiten, '13)

Towards an understanding. A skinny setting. 2

$$|\lambda_{m+1}| = \min_{\substack{X \in \mathbb{R}^{n\times n} \\ \mathrm{rank}(X)=m}} \|X^\star - X\| \le \|X^\star - \widetilde{X}\| \qquad (*)$$

- The error norm $\|X^\star - \widetilde{X}\|$ in $(*)$ provides a, not necessarily sharp, upper bound for the $(m+1)$st singular value of $X^\star$
  (Penzl, '00)

- Given a rank-m matrix $\widetilde{X}$, $(*)$ indicates that the error norm $\|X^\star - \widetilde{X}\|$ cannot go below $|\lambda_{m+1}|$

> We assume $X^\star$ can be well approximated by a low rank matrix

(see, e.g., Benner & Breiten, '13)

Singular values and error matrix

Let $E^{(k)} = X^\star - X^{(k)}$. We first note

$$\|E^{(k)}\|_2 \leq \|E^{(k)}\|_F \leq \lambda_{\min}(\mathcal{A})^{-\frac{1}{2}}\|X^\star - X^{(k)}\|_{\mathcal{L}}$$

$\Rightarrow$ the approximation of $X^{(k)}$ to $X^\star$ occurs in terms of singular values

That is,

As convergence takes place (i.e., $\|X^\star - X^{(k)}\|$ decreases) the **leading** singular triplets of $X^{(k)}$ tend to match those of X

However, **below the level of the error norm the singular values of the two matrices $X_\star$ and $X^{(k)}$ can vary significantly**

(Formalization by using Wedin result on singular vector subspace angle)

Singular values and error matrix

Let $E^{(k)} = X^\star - X^{(k)}$. We first note

$$\|E^{(k)}\|_2 \leq \|E^{(k)}\|_F \leq \lambda_{\min}(\mathcal{A})^{-\frac{1}{2}} \|X^\star - X^{(k)}\|_{\mathcal{L}}$$

$\Rightarrow$ the approximation of $X^{(k)}$ to $X^\star$ occurs in terms of singular values

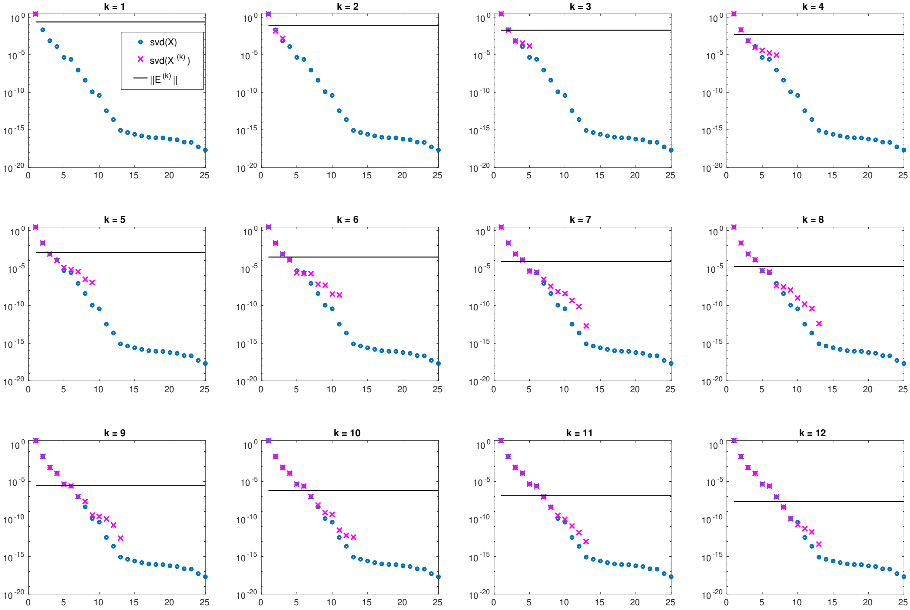
That is,

As convergence takes place (i.e., $\|X^\star - X^{(k)}\|$ decreases) the **leading** singular triplets of $X^{(k)}$ tend to match those of X

However, **below the level of the error norm the singular values of the two matrices $X_\star$ and $X^{(k)}$ can vary significantly**

(Formalization by using Wedin result on singular vector subspace angle)

An example



Singular values of $X^\star$ and of $X^{(k)}$ and error threshold for each of the first 12 iterations

Effect of truncation. Generated approximation subspace

$$AX + XA + MXM = c_1 c_1^\top$$

- The low rank iterates naturally lose rank: if $X^{(0)} := X_1^{(0)} X_1^{(0)\top} (= 0)$, $R^{(0)} := R_1^{(0)} S^{(0)} R_1^{(0)\top}$ with $R_1^{(0)} = c_1$ and $R^{(0)} = P^{(0)} =: P_1^{(0)} D^{(0)} P_1^{(0)\top}$ then

$$\mathbf{X}_1^{(1)} = [X_1^{(0)} \; P_1^{(0)}] = c_1; \qquad R_1^{(1)} = [c_1 \; AX_1^{(1)} \; X_1^{(1)} \; MX_1^{(1)}] = [c_1 \; Ac_1 \; c_1 \; Mc_1];$$

$$P_1^{(1)} = [R_1^{(1)} \; P_1^{(0)}] = [c_1 \; Ac_1 \; c_1 \; Mc_1 \; c_1] \qquad \mathbf{X}_1^{(2)} = [X_1^{(1)} \; P_1^{(1)}] = [c_1 \; c_1 \; Ac_1 \; c_1 \; Mc_1 \; c_1]$$

- The built space

$$\mathbb{Q} = \mathrm{span}\{c_1, \underbrace{Ac_1, Mc_1}, \underbrace{A^2c_1, AMc_1, MAc_1, M^2c_1},$$
$$\underbrace{A^3c_1, A^2Mc_1, AMAc_1, AM^2c_1, MA^2c_1, MAMc_1, M^2Ac_1, M^3c_1}, \cdots\},$$

Let $\mathbb{Q}_k$ be the smallest subspace of $\mathbb{Q}$ containing the range of $X_1^{(k)}$. Then

$$\boxed{\dim(\mathbb{Q}_{k+1}) \le \dim(\mathbb{Q}_k) + 2^k}$$

- CG iteration is unable to capture the underlying space $\mathbb{Q}_k$
$\Rightarrow$, any standard truncation strategy on the factor $X_1^{(k+1)}$ will lose part of the information contained in $\mathbb{Q}_k$

Effect of truncation. Generated approximation subspace

$$AX + XA + MXM = c_1 c_1^\top$$

- The low rank iterates naturally lose rank: if $X^{(0)} := X_1^{(0)} X_1^{(0)\top} (= 0)$, $R^{(0)} := R_1^{(0)} S^{(0)} R_1^{(0)\top}$ with $R_1^{(0)} = c_1$ and $R^{(0)} = P^{(0)} =: P_1^{(0)} D^{(0)} P_1^{(0)\top}$ then

$$\mathbf{X}_1^{(1)} = [X_1^{(0)} \ P_1^{(0)}] = c_1; \qquad R_1^{(1)} = [c_1 \ AX_1^{(1)} \ X_1^{(1)} \ MX_1^{(1)}] = [c_1 \ Ac_1 \ c_1 \ Mc_1];$$

$$P_1^{(1)} = [R_1^{(1)} \ P_1^{(0)}] = [c_1 \ Ac_1 \ c_1 \ Mc_1 \ c_1] \qquad \mathbf{X}_1^{(2)} = [X_1^{(1)} \ P_1^{(1)}] = [c_1 \ c_1 \ Ac_1 \ c_1 \ Mc_1 \ c_1]$$

- The built space

$$\mathbb{Q} = \text{span}\{c_1, \underbrace{Ac_1, Mc_1}, \underbrace{A^2c_1, AMc_1, MAc_1, M^2c_1},$$
$$\underbrace{A^3c_1, A^2Mc_1, AMAc_1, AM^2c_1, MA^2c_1, MAMc_1, M^2Ac_1, M^3c_1}, \cdots\},$$

Let $\mathbb{Q}_k$ be the smallest subspace of $\mathbb{Q}$ containing the range of $X_1^{(k)}$. Then

$$\boxed{\dim(\mathbb{Q}_{k+1}) \le \dim(\mathbb{Q}_k) + 2^k}$$

- CG iteration is unable to capture the underlying space $\mathbb{Q}_k$
$\Rightarrow$, any standard truncation strategy on the factor $X_1^{(k+1)}$ will lose part of the information contained in $\mathbb{Q}_k$

Effect of truncation. Generated approximation subspace

$$AX + XA + MXM = c_1 c_1^\top$$

- The low rank iterates naturally lose rank: if $X^{(0)} := X_1^{(0)} X_1^{(0)\top} (= 0)$, $R^{(0)} := R_1^{(0)} S^{(0)} R_1^{(0)\top}$ with $R_1^{(0)} = c_1$ and $R^{(0)} = P^{(0)} =: P_1^{(0)} D^{(0)} P_1^{(0)\top}$ then

$$\mathbf{X}_1^{(1)} = [X_1^{(0)} \; P_1^{(0)}] = c_1; \qquad R_1^{(1)} = [c_1 \; AX_1^{(1)} \; X_1^{(1)} \; MX_1^{(1)}] = [c_1 \; Ac_1 \; c_1 \; Mc_1];$$

$$P_1^{(1)} = [R_1^{(1)} \; P_1^{(0)}] = {\color{red}[c_1 \; Ac_1 \; c_1 \; Mc_1 \; c_1]} \qquad \mathbf{X}_1^{(2)} = [X_1^{(1)} \; P_1^{(1)}] = {\color{red}[c_1 \; c_1 \; Ac_1 \; c_1 \; Mc_1 \; c_1]}$$

- The built space

$$\mathbb{Q} = \mathrm{span}\{c_1, \underbrace{Ac_1, Mc_1}, \underbrace{A^2c_1, AMc_1, MAc_1, M^2c_1},$$
$$\underbrace{A^3c_1, A^2Mc_1, AMAc_1, AM^2c_1, MA^2c_1, MAMc_1, M^2Ac_1, M^3c_1}, \cdots\},$$

Let $\mathbb{Q}_k$ be the smallest subspace of $\mathbb{Q}$ containing the range of $X_1^{(k)}$. Then

$$\boxed{\dim(\mathbb{Q}_{k+1}) \le \dim(\mathbb{Q}_k) + 2^k}$$

- CG iteration is unable to capture the underlying space $\mathbb{Q}_k$
$\Rightarrow$, any standard truncation strategy on the factor $X_1^{(k+1)}$ will lose part of the information contained in $\mathbb{Q}_k$

Effect of truncation on the iterates

Let $x_k = \mathrm{vec}(X_k)$ (and similarly for the other variables). Truncation can be written as

$$x^{(k+1)} = x_{ex}^{(k+1)} + \boldsymbol{e}_X^{(k+1)}, \qquad p^{(k+1)} = p_{ex}^{(k+1)} + \boldsymbol{e}_P^{(k+1)}$$

($\boldsymbol{e}_X^{(k+1)}, \boldsymbol{e}_P^{(k+1)}$ local truncation errors)

TH: Let $\Delta_k = \max\{\|\boldsymbol{e}_P^{(k)}\|, \|\boldsymbol{e}_X^{(k)}\|, \|\boldsymbol{e}_P^{(k+1)}\|, \|\boldsymbol{e}_X^{(k+1)}\|\}$ and also $\delta_k = \min\{\|\boldsymbol{e}_P^{(k)}\|, \|\boldsymbol{e}_X^{(k)}\|, \|\boldsymbol{e}_P^{(k+1)}\|, \|\boldsymbol{e}_X^{(k+1)}\|\}$. Then there exists $\eta \in [0,1]$ such that

$$\eta \frac{1}{\|\mathcal{A}^{-1}\|} \frac{\delta_k}{\|r^{(k+1)}\|} \le \frac{|r^{(k+1)})^\top p^{(k)}|}{\|r^{(k+1)}\|\|p^{(k)}\|} \le \|\mathcal{A}\| \frac{\Delta_k}{\|r^{(k+1)}\|},$$

and

$$\beta_k = -\frac{(r_{ex}^{(k+1)})^\top \mathcal{A}p^{(k)} - (\mathcal{A}\boldsymbol{e}_X^{(k+1)})^\top \mathcal{A}p^{(k)}}{(p^{(k)})^\top \mathcal{A}p^{(k)}}$$

Moreover,

$$\frac{|r^{(k+1)})^\top r^{(k)}|}{\|r^{(k+1)}\|\|r^{(k)}\|} \le \gamma \frac{\Delta_k}{\|r^{(k+1)}\|} \qquad \gamma = \|\mathcal{A}p^{(k)}\| + (2|\beta_{k-1}| + |\beta_{k-1}\alpha_k|)\|\mathcal{A}p^{(k-1)}\| + \|r^{(k+1)}\|)/\|r^{(k)}\|$$

Effect of truncation on the iterates

Let $x_k = \mathrm{vec}(X_k)$ (and similarly for the other variables). Truncation can be written as

$$x^{(k+1)} = x_{ex}^{(k+1)} + \boldsymbol{e}_X^{(k+1)}, \qquad p^{(k+1)} = p_{ex}^{(k+1)} + \boldsymbol{e}_P^{(k+1)}$$

($\boldsymbol{e}_X^{(k+1)}, \boldsymbol{e}_P^{(k+1)}$ local truncation errors)

TH: Let $\Delta_k = \max\{\|\boldsymbol{e}_P^{(k)}\|, \|\boldsymbol{e}_X^{(k)}\|, \|\boldsymbol{e}_P^{(k+1)}\|, \|\boldsymbol{e}_X^{(k+1)}\|\}$ and also $\delta_k = \min\{\|\boldsymbol{e}_P^{(k)}\|, \|\boldsymbol{e}_X^{(k)}\|, \|\boldsymbol{e}_P^{(k+1)}\|, \|\boldsymbol{e}_X^{(k+1)}\|\}$. Then there exists $\eta \in [0, 1]$ such that

$$\eta \frac{1}{\|\mathcal{A}^{-1}\|} \frac{\delta_k}{\|r^{(k+1)}\|} \le \frac{|r^{(k+1)})^\top p^{(k)}|}{\|r^{(k+1)}\|\|p^{(k)}\|} \le \|\mathcal{A}\| \frac{\Delta_k}{\|r^{(k+1)}\|},$$

and

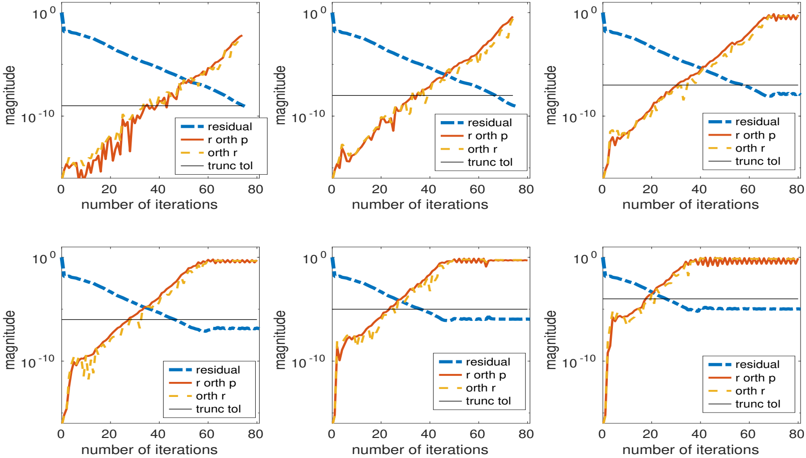
$$\beta_k = -\frac{(r_{ex}^{(k+1)})^\top \mathcal{A} p^{(k)} - (\mathcal{A}\boldsymbol{e}_X^{(k+1)})^\top \mathcal{A} p^{(k)}}{(p^{(k)})^\top \mathcal{A} p^{(k)}}$$

Moreover,

$$\frac{|r^{(k+1)})^\top r^{(k)}|}{\|r^{(k+1)}\|\|r^{(k)}\|} \le \gamma \frac{\Delta_k}{\|r^{(k+1)}\|} \qquad \gamma = \|\mathcal{A}p^{(k)}\| + (2|\beta_{k-1}| + |\beta_{k-1}\alpha_k|)\|\mathcal{A}p^{(k-1)}\| + \|r^{(k+1)}\|)/\|r^{(k)}\|$$

A: 2D Laplace operator, M =pentadiag$(-0.5, -1, 3.2, -1, -0.5)$, c_1 random entries

Truncated CG residual norm (blue line) for different truncation values



Also reported: Loss of orthogonality (cosine of the angles) between consecutive residuals and residual and directions

Wrap-up and Outlook

- Truncated CG in its youth (and happily behaves as such)
- Truncated CG behavior accepted in the "inexact" context (by necessity)

- Open problem: new truncation strategy that can capture the right information
- Open problem: new truncation strategy that better controls the rank

Visit: www.dm.unibo.it/~simoncin

Email address: valeria.simoncini@unibo.it

Reference:
V. Simoncini and Yue Hao
Analysis of the truncated conjugate gradient method for linear matrix equations
pp. 1-24, Dipartimento di Matematica, Universita' di Bologna, Feb. 2022. HAL archive hal-03579267

Wrap-up and Outlook

- Truncated CG in its youth (and happily behaves as such)
- Truncated CG behavior accepted in the "inexact" context (by necessity)

- Open problem: new truncation strategy that can capture the right information
- Open problem: new truncation strategy that better controls the rank

Visit: www.dm.unibo.it/~simoncin

Email address: valeria.simoncini@unibo.it

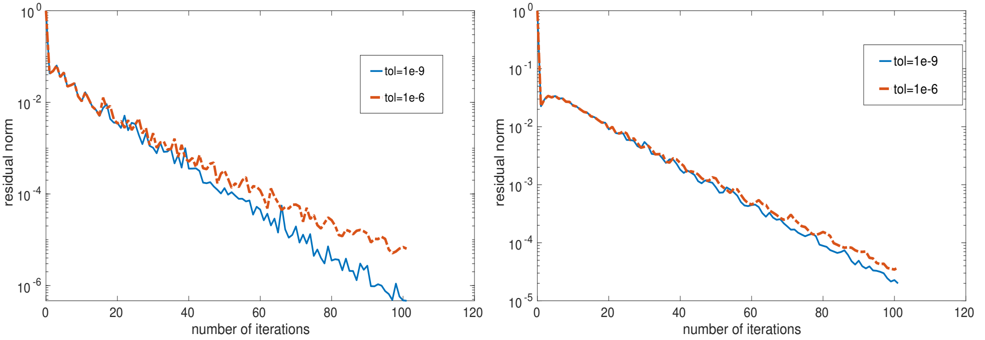
Reference:
V. Simoncini and Yue Hao
Analysis of the truncated conjugate gradient method for linear matrix equations
pp. 1-24, Dipartimento di Matematica, Universita' di Bologna, Feb. 2022. HAL archive hal-03579267

Another example

$A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ with $\lambda_i = \lambda_1 + \frac{(i-1)}{(n-1)}(\lambda_n - \lambda_1)\rho^{n-i}$, $\lambda_1 = 0.1, \quad \lambda_n = 100$

M: diagonal matrix with elements logarithmically distributed in $[10^{-2}, 10^{0}]$

Convergence history of TCG for two truncation tolerances:



Left: $\rho = 0.4$

Right: $\rho = 0.8$