Nečas Center for Mathematical Modeling
MOdelling REvisited + MOdel REduction
ERC-CZ project LL1202 - MORE
MINISTRY OF EDUCATION, YOUTH AND SPORTS

On the properties of Krylov subspaces in finite precision CG computations

Tomáš Gergelits, Zdeněk Strakoš

Department of Numerical Mathematics, Faculty of Mathematics and Physics,
Charles University in Prague

4th IMA Conference, Birmingham

3rd September, 2014

Content of the talk

1. The essence of the CG method
2. Krylov subspaces in practical computations
3. Idea of shift
4. Comparison of trajectories of approximation vectors
5. Comparison of generated Krylov subspaces
6. Concluding remarks

The essence of the CG method

Consider preconditioned system

$$Ax = b, \quad A \in \mathbb{F}^{N\times N} \text{ HPD}, \quad b \in \mathbb{F}^{N}, \quad \mathbb{F} \text{ is } \mathbb{R} \text{ or } \mathbb{C}$$

CG is the projection method which minimizes the energy norm of the error

$$x_k \in x_0 + \mathcal{K}_k(A, r_0), \quad r_k \perp \mathcal{K}_k(A, r_0), \quad k = 1, 2, \ldots$$
$$\mathcal{K}_k(A, r_0) = \text{span}\{r_0, Ar_0, A^2 r_0, \ldots, A^{k-1} r_0\}$$

$$\|x - x_k\|_A = \min\left\{\|x - y\|_A : \ y \in x_0 + \mathcal{K}_k(A, r_0)\right\}.$$

Krylov subspaces in practical computations

Krylov subspace

$$\mathcal{K}_k(B, v) = \mathsf{span}\{v, Bv, \ldots, B^{k-1}v\}$$

is built up by powering the matrix.

Krylov subspaces in practical computations

Krylov subspace

$$\mathcal{K}_k(B, v) = \mathsf{span}\{v, Bv, \ldots, B^{k-1}v\}$$

is built up by powering the matrix.

Important question arising in numerical computations

- What is the difference between

$$\mathcal{K}_k(B, v) \quad \text{and} \quad \overline{\mathcal{K}}_k(B, v)\,?$$

Krylov subspaces in practical computations

Krylov subspace

$$\mathcal{K}_k(B,v) = \mathsf{span}\{v, Bv, \ldots, B^{k-1}v\}$$

is built up by powering the matrix.

Important question arising in numerical computations

- What is the difference between

$$\mathcal{K}_k(B,v) \quad \text{and} \quad \overline{\mathcal{K}}_k(B,v)\,?$$

- Related question of sensitivity of Krylov subspaces

$$\mathcal{K}_k(B+\Delta B, v+\delta v).$$

Krylov subspaces in practical computations

Krylov subspace

$$\mathcal{K}_k(B, v) = \mathsf{span}\{v, Bv, \ldots, B^{k-1}v\}$$

is built up by powering the matrix.

Important question arising in numerical computations

- What is the difference between

$$\mathcal{K}_k(B, v) \quad \text{and} \quad \overline{\mathcal{K}}_k(B, v)\,?$$

- Related question of sensitivity of Krylov subspaces

$$\mathcal{K}_k(B + \Delta B, v + \delta v).$$

- Perturbation analysis, condition number of Krylov subspaces.
  [Carproux, Godunov, Kuznetsov (1997); Paige, Van Dooren (1998)]

Krylov subspaces in practical computations

Krylov subspace

$$\mathcal{K}_k(B, v) = \mathrm{span}\{v, Bv, \ldots, B^{k-1}v\}$$

is built up by powering the matrix.

Important question arising in numerical computations

- What is the difference between

$$\mathcal{K}_k(B, v) \quad \text{and} \quad \overline{\mathcal{K}}_k(B, v)\,?$$

- Related question of sensitivity of Krylov subspaces

$$\mathcal{K}_k(B + \Delta B, v + \delta v).$$

- Perturbation analysis, condition number of Krylov subspaces. [Carproux, Godunov, Kuznetsov (1997); Paige, Van Dooren (1998)]

Short recurrences $\Longrightarrow$ significant delay of convergence.

CG in finite precision computations

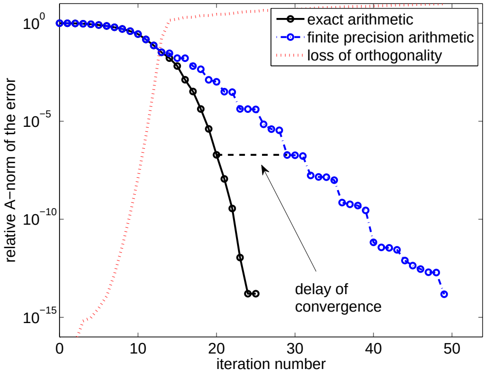
Short recurrences

CG in finite precision computations

Short recurrences $\Longrightarrow$ loss of orthogonality

CG in finite precision computations

Short recurrences $\Longrightarrow$ loss of orthogonality $\Longrightarrow$ delay of convergence & rank deficiency



Idea of shift

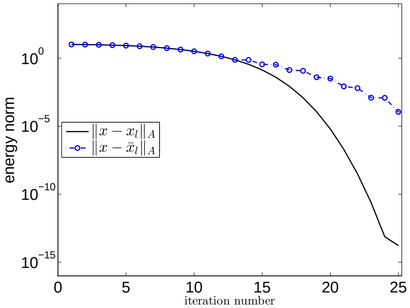
We relate: k-th iteration of FP CG $\Longleftrightarrow$ l-th iteration of exact CG

- $k - l \quad \approx \quad$ delay of convergence
- $k - l \quad \approx \quad$ rank-deficiency of computed Krylov subspace

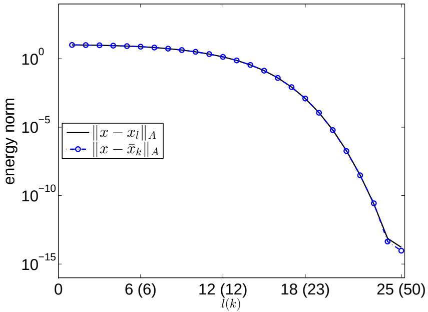
We want to study:

$$\begin{aligned} \|x - \overline{x}_k\|_A \quad &\times \quad \|x - x_l\|_A \\ \overline{x}_k \quad &\times \quad x_l \\ \overline{\mathcal{K}}_k(A, r_0) \quad &\times \quad \mathcal{K}_l(A, r_0) \end{aligned}$$

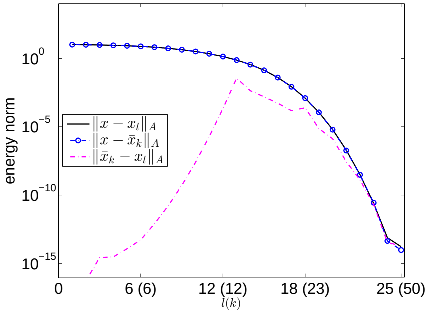
Comparison of trajectory of approximation vectors



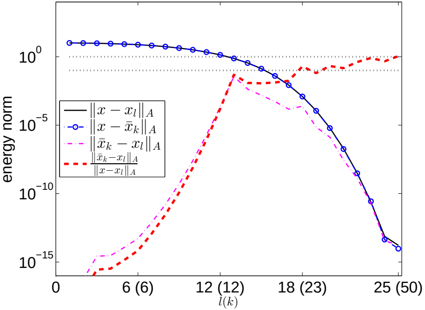
Comparison of trajectory of approximation vectors



Comparison of trajectory of approximation vectors



Comparison of trajectory of approximation vectors

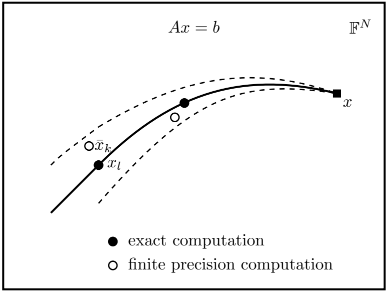


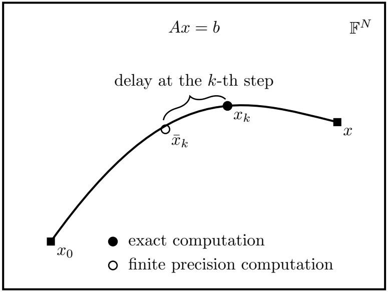
Observation

$$\frac{\|\bar{x}_k - x_l\|_A}{\|x - x_l\|_A} \ll 1$$

Trajectories of approximation vectors are very similar in space $\mathbb{F}^N$.

Comparison of trajectory of approximation vectors





Trajectory of approximations $\bar{x}_k$ generated by FP CG computations follows closely the trajectory of the exact CG approximations x_l.

Comparison of Krylov subspaces

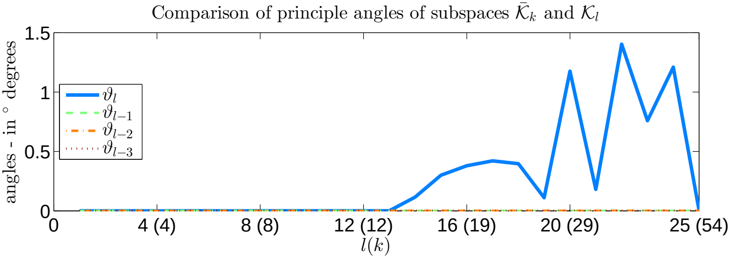
Canonical angles and vectors

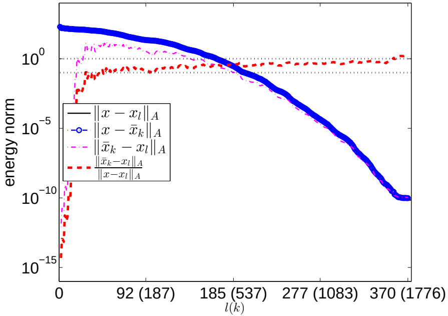
$$\vartheta_j = \min_{\substack{p \in \mathcal{F}_j \\ \|p\|=1}} \min_{\substack{q \in \mathcal{G}_j \\ \|q\|=1}} \arccos(p^* q) \equiv \arccos(p_j{}^* q_j), \quad j = 1, 2, \ldots, l$$

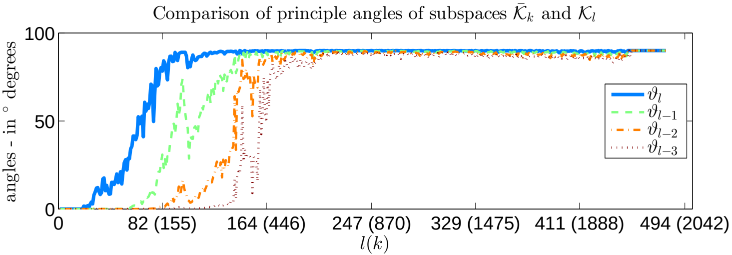
where

$$\mathcal{F}_j \equiv \mathcal{F} \cap \{p_1, \ldots, p_{j-1}\}^{\perp}, \qquad \mathcal{G}_j \equiv \mathcal{G} \cap \{q_1, \ldots, q_{j-1}\}^{\perp},$$

$$\mathcal{F} = \overline{\mathcal{K}}_k(A, r_0), \qquad \mathcal{G} = \mathcal{K}_l(A, r_0).$$

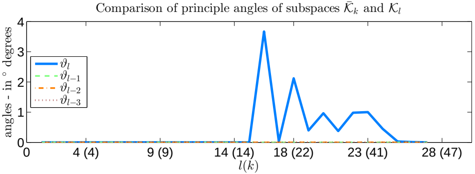



energy norm
10^{0}
10^{-5}
10^{-10}
10^{-15}
0
92 (187)
185 (537)
277 (1083)
370 (1776)
$l(k)$
$\|x - x_l\|_A$
$\|x - \bar{x}_k\|_A$
$\|\bar{x}_k - x_l\|_A$
$\frac{\|\bar{x}_k - x_l\|_A}{\|x - x_l\|_A}$

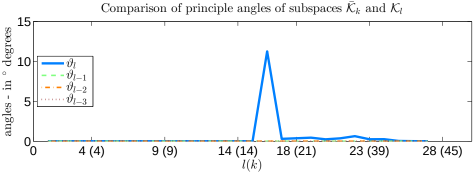

Comparison of principle angles of subspaces $\bar{\mathcal{K}}_k$ and $\mathcal{K}_l$
angles - in ° degrees
100
50
0
0
82 (155)
164 (446)
247 (870)
329 (1475)
411 (1888)
494 (2042)
$l(k)$
ϑ_l
ϑ_{l-1}
ϑ_{l-2}
ϑ_{l-3}

Cluster of 3 largest eigenvalues with width Δ.

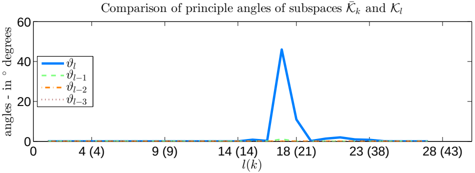
$\Delta = 10^{-3}$



$\Delta = 10^{-8}$



$\Delta = 10^{-12}$



Summary and outlook

- The trajectories of computed approximations are enclosed in a shrinking "cone".
- Observed "stability" (or inertia?) of computed Krylov subspaces represents phenomenon which needs to be further studied.

- How to determine pairs (l, k).
- Effect of clustered eigenvalues.
- Theoretical proofs, relationship to the structure of invariant subspaces.
- Principle difference between long and short recurrences.

Summary and outlook

- ☞ The trajectories of computed approximations are enclosed in a shrinking "cone".
- ☞ Observed "stability" (or inertia?) of computed Krylov subspaces represents phenomenon which needs to be further studied.

- How to determine pairs (l, k).
- Effect of clustered eigenvalues.
- Theoretical proofs, relationship to the structure of invariant subspaces.
- Principle difference between long and short recurrences.

Acknowledgement

This work has been supported by the ERC-CZ project LL1202, by the GACR grant 201/13-06684S and by the GAUK grant 695612.