An Analysis of algebraic preconditioning based on different variants of the Gram-Schmidt algorithm

Jiří Kopal[†]
Miroslav Tůma[‡]
Miroslav Rozložník[‡]

[†]NTI FM TUL
[‡]ÚI AV ČR

1. 10. 2010

Outline

Motivation of preconditioning

$$\mathbf{A}x = b$$

- improving of spectral properties of the matrix $\mathbf{A}$
- absolute necessity for increasing robustness of iterative solvers
- acceleration of convergence

General strategy: multiply equation $\mathbf{A}x = b$ by $\mathbf{P}$, where $\mathbf{P} \approx \mathbf{A}^{-1}$, so that $\|\mathbf{I} - \mathbf{PA}\| \to 0$

Preserving symmetry

Assume A is SPD

For preserving symmetry: $\mathbf{P} = \mathbf{Z}\mathbf{Z}^T$

$$\|\mathbf{I} - \mathbf{Z}^T\mathbf{A}\mathbf{Z}\| \rightarrow 0$$

Question: How to compute preconditioning **Z** efficiently, so that computation cost is much less $O(n^3)$ (Gaussian elimination) for dense case and/or parallelizable?

Algorithms for preconditioning

- ILU - classical type of preconditioning
- SPAI - minimization of functional $\|\mathbf{I} - \mathbf{PA}\|_F$, decomposition to n-independent problems, TU Munich
- AINV, SAINV - based on the generalized Gram-Schmidt algorithm
- ...

Preconditioning matrix $\mathbf{Z}$ should be sparse, in order to have $\mathbf{Z}^T\mathbf{AZ}$ sparse as well $\rightarrow$ prescribing pattern, dropping...

Successful preconditioning

Accuracy: $\|\mathbf{I} - \mathbf{Z}^T\mathbf{A}\mathbf{Z}\|$ Stability: $\|\mathbf{A} - (\mathbf{Z}^T\mathbf{Z})^{-1}\|$

- Low
- Controlled (via dropping)

NOTE: Currently no general theory of preconditioning (stability + accuracy) $\longrightarrow$ convergence

Generalized Gram-Schmidt algorithm

- $\mathbf{A}$ is a SPD matrix of dimension $n \times n$
- energetic dot product $\langle x, y \rangle_A = y^T \mathbf{A} x$
- $\mathbf{A}$-orthogonality (conjugate gradient method)
- basis $\mathbf{Z}^{(0)}$, which will be $\mathbf{A}$-orthogonalized against previously computed vectors
  (for simplicity and for real-world problems $\mathbf{Z}^{(0)} = \mathbf{I}$ or $\mathbf{Z}^{(0)} = diag(\mathbf{A})^{-1/2}$

Versions

- classical, modified, with iterative refinement
- right looking, left looking

Properties in exact arithmetic

- $\mathbf{Z}^{\mathbf{T}}\mathbf{A}\mathbf{Z} = \mathbf{I}$
- $\mathbf{A}^{-1} = \mathbf{Z}\mathbf{Z}^{\mathbf{T}}$
- $\mathbf{Z}^{(0)} = \mathbf{Z}\mathbf{U}$
- $(\mathbf{Z}^{(0)})^{\mathbf{T}}\mathbf{A}\mathbf{Z}^{(0)} = \mathbf{U}^{\mathbf{T}}\mathbf{U}$

Finite precision

- **In finite precision arithmetic** $\|\mathbf{Z}^{\mathbf{T}}\mathbf{A}\mathbf{Z} - \mathbf{I}\| < ???$

Question: How to estimate the loss of orthogonality $\|\mathbf{Z}^{\mathbf{T}}\mathbf{A}\mathbf{Z} - \mathbf{I}\|$ in case of incomplete Gram-Schmidt algorithm (with dropping)?

Idea: Björck,Å.,*Solving linear least squares problems by Gram-Schmidt orthogonalization*, BIT 7 (1967),1-21

Preconditioning theory

Analysis:
Behavior in finite precision arithmetic (rounding errors)

Practical algorithms:
Dropping and incomplete algorithm Gram-Schmidt

Theory: Move from (full) algorithm behavior (u) to preconditioning (incomplete algorithm - τ)

Our results

$$
\begin{aligned}
\|\mathbf{Z}^{\mathbf{T}}\mathbf{A}\mathbf{Z}-\mathbf{I}\| &\leq O(u)\kappa(A) && \text{(CGS2, EIG)}\\
&\leq \frac{O(u)\kappa^{3/2}(A)}{1-O(u)\kappa^{3/2}(A)} && \text{(MGS - SAINV)}\\
&\leq \frac{O(u)\kappa^{2}(A)}{1-O(u)\kappa(A)} && \text{(CGS, AINV)}
\end{aligned}
$$

Choice of $\mathbf{Z}^{(0)}$

$\mathbf{Z}^{(0)}$ can be any $n \times m$ matrix with full column rank and $m \leq n$ ($\mathbf{A}$ is an $n \times n$ matrix)

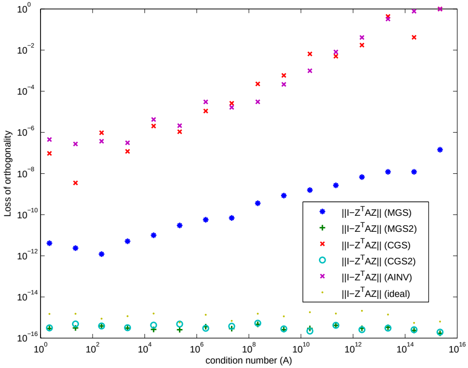
Loss of orthogonality (MGS):

$$\|\mathbf{Z}^T \mathbf{A} \mathbf{Z} - \mathbf{I}\| \leq \frac{O(u)\kappa(A)\kappa(A^{1/2} Z_k^{(0)})}{1 - O(u)\kappa(A)\kappa(A^{1/2} Z_k^{(0)})}$$

NOTE: If $\mathbf{Z}^{(0)}$ is an upper triangular matrix, then $\mathbf{Z} = \mathbf{L}^{-1}$, where $\mathbf{A} = \mathbf{L}\mathbf{L}^T$ (Cholesky factorization), otherwise $\mathbf{Z}$ is general matrix.

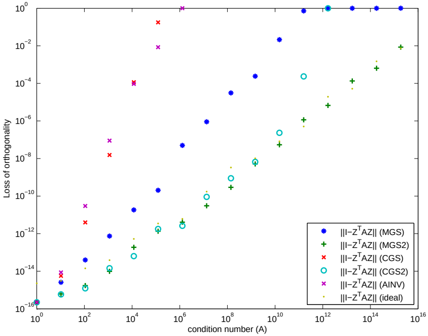
Test problem 1

A = diagonal matrix, $\kappa(A) \approx 10^i, i = 0, \ldots, 15$
$\mathbf{Z}^{(0)} = \mathbf{H}^{1/2}$, where **H** is the inverse Hilbert matrix, $\kappa(\mathbf{Z}^{(0)}) \approx 10^5$



Test problem 2

A, $\mathbf{Z}^{(0)}$ are the inverse Hilbert matrices (general powers)
$\kappa(A), \kappa(Z^{(0)}) \approx 10^i, i = 0, \ldots, 15$



Conclusion and open questions

Done

- (full) algorithm analysis

Future work

- find a connection between estimate of $\|\mathbf{Z}^T\mathbf{A}\mathbf{Z} - \mathbf{I}\|$ and dropping strategy
- analyze behavior of preconditioned iterative methods
- analyze other related algorithms and preconditioners

Thank you for your attention!!!