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

1. Preconditioning
   - Motivation
   - Algorithms
   - Finite precision and successful preconditioning

2. The Gram-Schmidt algorithm
   - Generalized Gram-Schmidt algorithm
   - Finite precision
   - Our results
   - Influence of initialization of the algorithm

3. Test problems
   - Test problem 1
   - Test problem 2

4. Conclusion and open questions
Motivation of preconditioning

\[ \mathbf{Ax} = \mathbf{b} \]

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

General strategy: multiply equation \( \mathbf{Ax} = \mathbf{b} \) 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: $P = ZZ^T$

$$\|I - Z^T AZ\| \rightarrow 0$$

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 $\|I - PA\|_F$, decomposition to n-independent problems, TU Munich
- AINV, SAINV - based on the generalized Gram-Schmidt algorithm

Preconditioning matrix $Z$ should be sparse, in order to have $Z^TAZ$ sparse as well $\rightarrow$ prescribing pattern, dropping...
Successful preconditioning

\[ \text{Accuracy: } \| I - Z^T A Z \| \quad \text{Stability: } \| A - (Z^T Z)^{-1} \| \]

- Low
- Controlled (via dropping)

NOTE: Currently no general theory of preconditioning (stability + accuracy) → convergence
Generalized Gram-Schmidt algorithm

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

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

Properties in exact arithmetic

- $Z^T A Z = I$
- $A^{-1} = Z Z^T$
- $Z^{(0)} = Z U$
- $(Z^{(0)})^T A Z^{(0)} = U^T U$
In finite precision arithmetic $\|Z^TAZ - I\| < ???$

Question: How to estimate the loss of orthogonality $\|Z^TAZ - I\|$ 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 - $\tau$)
Our results

\[ \| Z^T AZ - I \| \leq O(u) \kappa(A) \] (CGS2, EIG)

\[ \leq \frac{O(u) \kappa^{3/2}(A)}{1 - O(u) \kappa^{3/2}(A)} \] (MGS - SAINV)

\[ \leq \frac{O(u) \kappa^2(A)}{1 - O(u) \kappa(A)} \] (CGS, AINV)
Choice of $Z^{(0)}$

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

Loss of orthogonality (MGS):

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

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

\[ A = \text{diagonal matrix}, \quad \kappa(A) \approx 10^i, \ i = 0, \ldots, 15 \]

\[ Z^{(0)} = H^{1/2}, \text{ where } H \text{ is the inverse Hilbert matrix}, \quad \kappa(Z^{(0)}) \approx 10^5 \]
A, \(Z^{(0)}\) 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 $\|Z^T A Z - I\|$ and dropping strategy
- analyze behavior of preconditioned iterative methods
- analyze other related algorithms and preconditioners
Thank you for your attention!!!