MINRES-QLP: a Krylov subspace method for indefinite or singular symmetric systems

Sou-Cheng Choi — Univ of Chicago/Argonne Nat'l Lab
Chris Paige — School of CS, McGill University
Michael Saunders — ICME, Stanford University

2012 SIAM Conference on Applied Linear Algebra

Instituto de Matemática Multidisciplinar
Universitat Politècnica de València
Valencia, Spain

Abstract

CG, SYMMLQ, and MINRES are Krylóv subspace methods for solving symmetric systems of linear equations. When these methods are applied to an incompatible system (that is, a singular symmetric least-squares problem), CG could break down and SYMMLQ's solution could explode, while MINRES would give a least-squares solution but not necessarily the minimum-length (pseudoinverse) solution. This understanding motivates us to design a MINRES-like algorithm to compute minimum-length solutions to singular symmetric systems.

MINRES uses QR factors of the tridiagonal matrix from the Lanczos process (where R is upper-tridiagonal). MINRES-QLP uses a QLP decomposition (where rotations on the right reduce R to lower-tridiagonal form). On ill-conditioned systems (singular or not), MINRES-QLP can give more accurate solutions than MINRES. We derive preconditioned MINRES-QLP, new stopping rules, and better estimates of the solution and residual norms, the matrix norm, and the condition number.

Abstract

CG, SYMMLQ, and MINRES are Krylóv subspace methods for solving symmetric systems of linear equations. When these methods are applied to an incompatible system (that is, a singular symmetric least-squares problem), CG could break down and SYMMLQ's solution could explode, while MINRES would give a least-squares solution but not necessarily the minimum-length (pseudoinverse) solution. This understanding motivates us to design a MINRES-like algorithm to compute minimum-length solutions to singular symmetric systems.

MINRES uses QR factors of the tridiagonal matrix from the Lanczos process (where R is upper-tridiagonal). MINRES-QLP uses a QLP decomposition (where rotations on the right reduce R to lower-tridiagonal form). On ill-conditioned systems (singular or not), MINRES-QLP can give more accurate solutions than MINRES. We derive preconditioned MINRES-QLP, new stopping rules, and better estimates of the solution and residual norms, the matrix norm, and the condition number.

Krylov Крыло́в

Abstract

CG, SYMMLQ, and MINRES are Krylóv subspace methods for solving symmetric systems of linear equations. When these methods are applied to an incompatible system (that is, a singular symmetric least-squares problem), CG could break down and SYMMLQ's solution could explode, while MINRES would give a least-squares solution but not necessarily the minimum-length (pseudoinverse) solution. This understanding motivates us to design a MINRES-like algorithm to compute minimum-length solutions to singular symmetric systems.

MINRES uses QR factors of the tridiagonal matrix from the Lanczos process (where R is upper-tridiagonal). MINRES-QLP uses a QLP decomposition (where rotations on the right reduce R to lower-tridiagonal form). On ill-conditioned systems (singular or not), MINRES-QLP can give more accurate solutions than MINRES. We derive preconditioned MINRES-QLP, new stopping rules, and better estimates of the solution and residual norms, the matrix norm, and the condition number.

Krylov Крыло́в Chebyshev Чебышёв

Outline

- Symmetric Lanczos
- CG, SYMMLQ, MINRES
- Theorem
- Joke
- MINRES-QLP
- Numerical example

Tridiagonalization of symmetric A

Direct (product of Householder transformations):

$$\begin{pmatrix} 1 & \\ & V^T \end{pmatrix} \begin{pmatrix} 0 & b^T \\ b & A \end{pmatrix} \begin{pmatrix} 1 & \\ & V \end{pmatrix} = \begin{pmatrix} 0 & x & & & \\ x & x & x & & \\ & x & x & x & \\ & & x & x & x \\ & & & x & x \end{pmatrix}$$

Tridiagonalization of symmetric A

Direct (product of Householder transformations):

$$\begin{pmatrix} 1 & \\ & V^T \end{pmatrix} \begin{pmatrix} 0 & b^T \\ b & A \end{pmatrix} \begin{pmatrix} 1 & \\ & V \end{pmatrix} = \begin{pmatrix} 0 & x & & & \\ x & x & x & & \\ & x & x & x & \\ & & x & x & x \\ & & & x & x \end{pmatrix}$$

Iterative (symmetric Lanczos process):

$$\begin{pmatrix} b & AV_k \end{pmatrix} = V_{k+1} \begin{pmatrix} \beta e_1 & \underline{T_k} \end{pmatrix}$$

$$V_k = \begin{pmatrix} v_1 & \dots & v_k \end{pmatrix} \qquad \underline{T_k} = \begin{pmatrix} T_k \\ 0 \dots 0 \;\; \beta_{k+1} \end{pmatrix}$$

Lanczos for solving $Ax = b$

$$\beta v_1 = b$$

$$V_k = \begin{pmatrix} v_1 & \dots & v_k \end{pmatrix} \qquad n \times k$$

$$x_k = V_k y_k \qquad \text{for some } y_k$$

Lanczos for solving $Ax = b$

$$\beta v_1 = b$$

$$V_k = \begin{pmatrix} v_1 & \dots & v_k \end{pmatrix} \qquad n \times k$$

$$x_k = V_k y_k \qquad \text{for some } y_k$$

$$\begin{pmatrix} b & AV_k \end{pmatrix} = V_{k+1} \begin{pmatrix} \beta e_1 & \underline{T_k} \end{pmatrix}$$

$$b - AV_k y_k = V_{k+1} \left(\beta e_1 - \underline{T_k} y_k \right)$$

$$\|b - Ax_k\| \le \|V_{k+1}\| \underbrace{\|\beta e_1 - \underline{T_k} y_k\|}_{\text{make small}}$$

Lanczos properties

For most iterations, $AV_k = V_{k+1}\underline{T_k}$

Theorem

$\underline{T_k}$ has full column rank for all $k < \ell$
(so the MINRES subproblem min $\|\beta e_1 - \underline{T_k} y_k\|$ is well defined)

Lanczos properties

For most iterations, $AV_k = V_{k+1}\underline{T_k}$

Theorem

$\underline{T_k}$ has full column rank for all $k < \ell$
(so the MINRES subproblem min $\|\beta e_1 - \underline{T_k} y_k\|$ is well defined)

At the last iteration, $AV_\ell = V_\ell T_\ell$

Theorem

T_ℓ is nonsingular iff $b \in \text{range}(A)$, and $\text{rank}\, T_\ell = \ell$ or $\ell - 1$
(so MINRES is ok only if $Ax = b$)

Four ways to make $\underline{T_k} y_k \approx \beta e_1$

$$\begin{pmatrix} \alpha_1 & \beta_2 & & & & \\ \beta_2 & \alpha_2 & \beta_3 & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \beta_{k-1} & \alpha_{k-1} & \beta_k \\ & & & & & \end{pmatrix} y_k = \begin{pmatrix} \beta \\ 0 \\ \vdots \\ \vdots \\ 0 \\ \end{pmatrix}$$

SYMMLQ $\quad \min \|y_k\| \;\text{ st }\; \underline{T_{k-1}}^T y_k = \beta e_1$

Four ways to make $\underline{T_k} y_k \approx \beta e_1$

$$\begin{pmatrix} \alpha_1 & \beta_2 & & & & \\ \beta_2 & \alpha_2 & \beta_3 & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \beta_{k-1} & \alpha_{k-1} & \beta_k \\ & & & & \beta_k & \alpha_k \\ & & & & & \end{pmatrix} y_k = \begin{pmatrix} \beta \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 0 \\ \end{pmatrix}$$

CG $\qquad T_k y_k = \beta e_1$

Four ways to make $\underline{T_k} y_k \approx \beta e_1$

$$\begin{pmatrix} \alpha_1 & \beta_2 & & & & \\ \beta_2 & \alpha_2 & \beta_3 & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \beta_{k-1} & \alpha_{k-1} & \beta_k \\ & & & & \beta_k & \alpha_k \\ & & & & & \beta_{k+1} \end{pmatrix} y_k \approx \begin{pmatrix} \beta \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

MINRES $\quad \min \| \underline{T_k} y_k - \beta e_1 \|$

Four ways to make $\underline{T_k} y_k \approx \beta e_1$

$$\begin{pmatrix} \alpha_1 & \beta_2 & & & & \\ \beta_2 & \alpha_2 & \beta_3 & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \beta_{k-1} & \alpha_{k-1} & \beta_k \\ & & & & \beta_k & \alpha_k \\ & & & & & \beta_{k+1} \end{pmatrix} y_k \approx \begin{pmatrix} \beta \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

MINRES $\min \|\underline{T_k} y_k - \beta e_1\|$

MINRES-QLP $\min \|y_k\|$ st $\min \|\underline{T_k} y_k - \beta e_1\|$

QLP decomposition of $\underline{T_k}$:

$$Q_k \underline{T_k} = \begin{pmatrix} R_k \\ 0 \end{pmatrix}, \quad R_k P_k = L_k \quad \Rightarrow \quad Q_k \underline{T_k} P_k = \begin{pmatrix} L_k \\ 0 \end{pmatrix}$$

$$y = P_k u \quad \Rightarrow \quad Q_k(\underline{T_k} y - \beta e_1) = \begin{pmatrix} L_k \\ 0 \end{pmatrix} u - \begin{pmatrix} t_k \\ \phi_k \end{pmatrix}$$

QLP decomposition of $\underline{T_k}$:

$$Q_k \underline{T_k} = \begin{pmatrix} R_k \\ 0 \end{pmatrix}, \quad R_k P_k = L_k \quad \Rightarrow \quad Q_k \underline{T_k} P_k = \begin{pmatrix} L_k \\ 0 \end{pmatrix}$$

$$y = P_k u \quad \Rightarrow \quad Q_k(\underline{T_k} y - \beta e_1) = \begin{pmatrix} L_k \\ 0 \end{pmatrix} u - \begin{pmatrix} t_k \\ \phi_k \end{pmatrix}$$

$k < \ell$:

$$L_k u_k = t_k, \qquad x_k = V_k P_k u_k$$

orthogonal steps
like SYMMLQ

QLP decomposition of $\underline{T_k}$:

$$Q_k \underline{T_k} = \begin{pmatrix} R_k \\ 0 \end{pmatrix}, \quad R_k P_k = L_k \qquad \Rightarrow \qquad Q_k \underline{T_k} P_k = \begin{pmatrix} L_k \\ 0 \end{pmatrix}$$

$$y = P_k u \qquad \Rightarrow \qquad Q_k(\underline{T_k} y - \beta e_1) = \begin{pmatrix} L_k \\ 0 \end{pmatrix} u - \begin{pmatrix} t_k \\ \phi_k \end{pmatrix}$$

$k < \ell$:

$$L_k u_k = t_k, \qquad x_k = V_k P_k u_k$$

orthogonal steps
like SYMMLQ

$k = \ell$:

$$L_\ell u_\ell = t_\ell \qquad \text{or} \qquad \min \|u_\ell\| \;\; \text{st} \;\; \min \|L_\ell u_\ell - t_\ell\|$$

Theorem

In MINRES-QLP, $x_\ell = V_\ell P_\ell u_\ell$ is the min-length solution of $Ax \approx b$

Theorem

In MINRES-QLP, $x_\ell = V_\ell P_\ell u_\ell$ is the min-length solution of $Ax \approx b$

Additional features:

- Two-sided spd preconditioner (reduce number of iterations)
- Transfer from MINRES to MINRES-QLP when $\underline{T_k}$ is moderately ill-conditioned

Theorem

In MINRES-QLP, $x_\ell = V_\ell P_\ell u_\ell$ is the min-length solution of $Ax \approx b$

Additional features:

- Two-sided spd preconditioner (reduce number of iterations)
- Transfer from MINRES to MINRES-QLP when $\underline{T_k}$ is moderately ill-conditioned

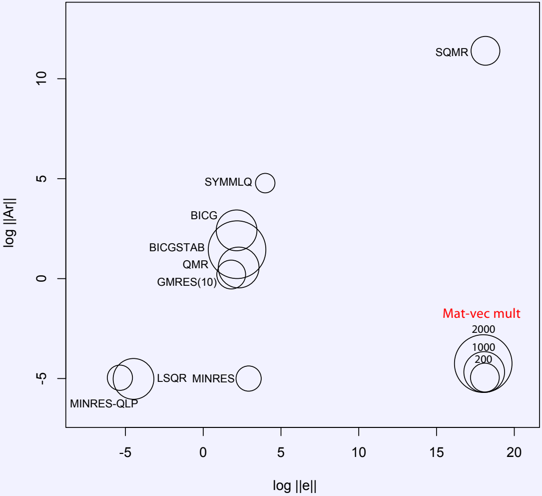
Per iteration costs:

- Storage: $7n$–$8n$ vectors
- Matrix-vector multiply: 1
- Work: $9n$–$14n$ flops
- (Solve a system with preconditioner)

Numerical example

$$A = \text{tridiag}\begin{pmatrix} T & T & T \end{pmatrix} \in \mathbb{R}^{400\times 400}, \quad T = \text{tridiag}\begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \in \mathbb{R}^{20\times 20}$$

$$|\lambda_1|, |\lambda_2| = O(\varepsilon), \quad |\lambda_3|, \dots, |\lambda_{400}| \in [0.2, 4.3], \quad b_i \sim i.i.d.\ U(0, 10)$$



Papers

- S.-C. T. Choi, C. C. Paige and M. A. Saunders, "MINRES-QLP: A Krylov subspace method for indefinite or singular symmetric systems," *SIAM J. Sci. Comput*, **33** (2011), no. 4, pp. 1810–1836.
- S.-C. T. Choi, C. C. Paige and M. A. Saunders, "ALGORITHM: MINRES-QLP for singular symmetric and Hermitian linear equations and least-squares problems," *ACM Trans. Math. Software*, to appear.
- S.-C. T. Choi, "CS-MINRES: a Krylov subspace method for Complex Symmetric Linear Equations and Least-Squares Problems," *preprint*, (2012).

Huge thanks

Research	NSF, NSERC, ONR, AHPCRC
Travel	SIAM, CI (U of Chicago/ANL), NSERC
Prize	SIAG/LA!

We dedicate MINRES-QLP
to the memory of Gene Golub

Gene's 75th + Stanford CS 50th
March 30, 2007