Lanczos Method for Complex Symmetric Eigenproblems

and

Shift-and-Invert Spectral Transformation

Z. Bai

T. Kowalski and R. Li

Workshop on the Use of Iterative Methods for Large-Scale Eigenvalue Problem

Argonne National Laboratory, May 14-16, 1997

Outline

Part I: Complex Symmetric Lanczos

1. Variants of Lanczos for complex symmetric
2. Numerical stability
3. Complex symmetric ABLE

Outline

Part II: Paradoxes in Shift-and-invert

1. Numerics: examples,
2. More on numerics: phenomena
3. Conditioning of shift-and-invert
4. Second shift choice:
   a global approach and a local approach
5. Final remarks and project goal

Complex Symmetric ABLE Program

- ABLE = **A**daptive **B**lock **L**anczos method for **E**igenproblems (B., Day and Ye'95)
- Error analysis of variants of Lanczos procedure for complex symmetric
- Block version

  complex orthogonal modified Gram-Schmidt

  Tagaki SVD
- Look-ahead and adaptive blocking
- Optional no, partial or full reorthogonalization
- Monitoring the loss of orthogonality, correction (*a new restart scheme*)
- Enhanced design of stopping criterion
- Shift-and-invert spectral transformation for $Ax = \lambda x$ and $Ax = \lambda Bx$

Notation

$\mathbf{u}$:	order of machine precision
cond(λ):	Wilkinson's condition number of eigenvalue λ
$\kappa(A)$	$= \|A\| \, \|A^{-1}\|$
$\Delta^2(A)$	$= \|A\|_F^2 - \Sigma \, |\lambda_i|^2$
δ_i	$= |\lambda_i - \mu|$, $\delta_1 \geq \delta_2 \geq \cdots \geq \delta_n > 0$
$\Lambda_\epsilon(A)$	ϵ-pseudospectra of A

Shift-and-Invert

$$
\begin{array}{rcl}
AX & = & X\Lambda \\
 & \updownarrow & \\
(A-\mu I)^{-1}X & = & X(\Lambda-\mu)^{-1}
\end{array}
$$

Condition numbers $\mathrm{cond}(\lambda)$ and $\kappa(X)$ are invariant (in exact arithmetic).

Arnoldi factorization

$$(A-\mu I)^{-1}V_m = V_m H_m + g_m e_m^T + F_m$$

where

$$\|F_m\| \leq \kappa((A-\mu I)^{-1})\mathbf{u}$$

$$\|V_m^H V_m - I\| = O(\mathbf{u}) \qquad \text{(MGS-IR)}$$

Rayleigh-Ritz step:

$$H_m Z = Z\Theta + E_m, \qquad \|E_m\| \leq \|H_m\|\mathbf{u}.$$

Residues

$$R_m^{(e)} = R_m^{(r)} - (A - \mu I)(V_m E_m + F_m Z)\Theta^{-1}$$

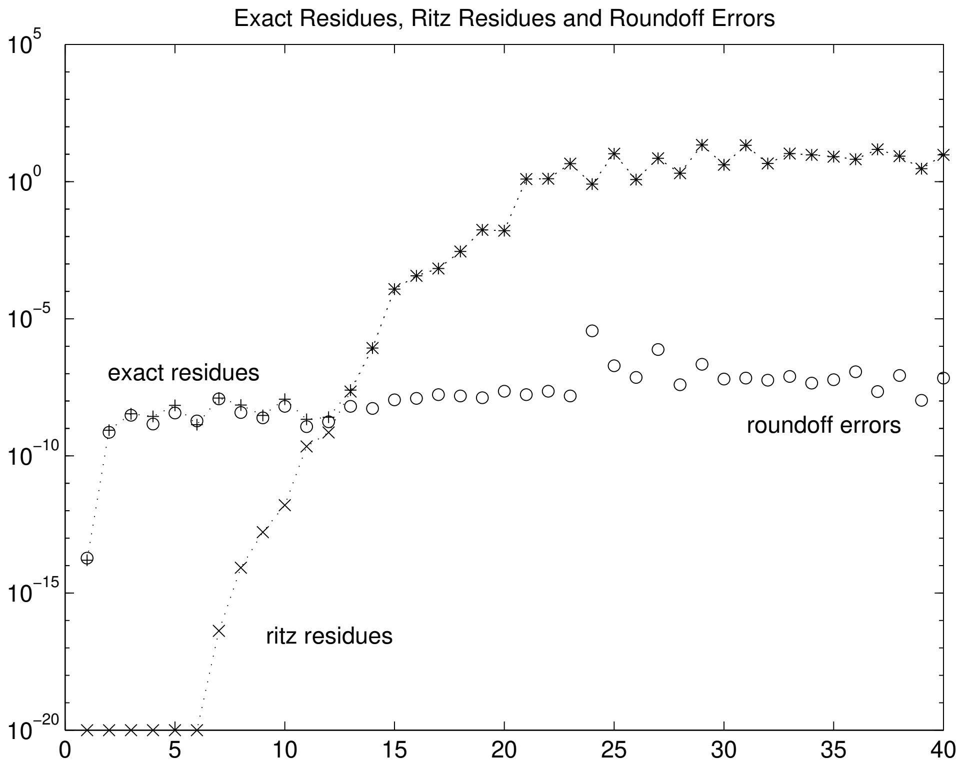
where

Exact: $R_m^{(e)} = A(V_m Z) - (V_m Z)(\mu I + \Theta^{-1})$

Ritz: $R_m^{(r)} = (A - \mu I) g_m (e_m^T Z)\Theta^{-1}$

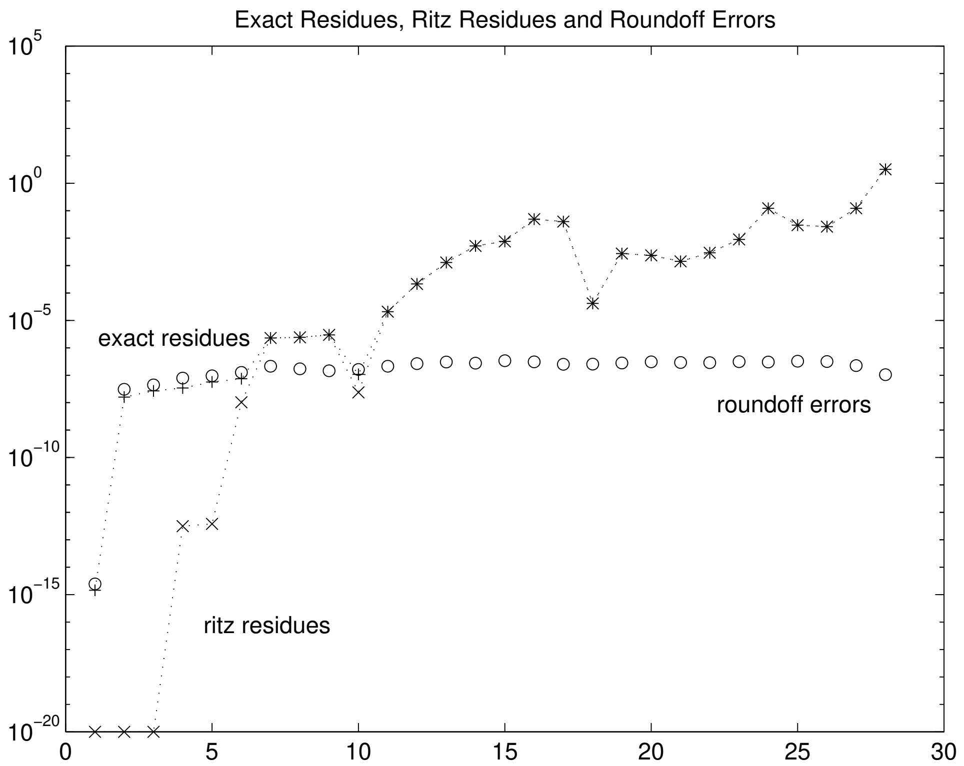
Mass matrix for bilinear quadrilateral element on a simple mesh

- 65×65, symmetric positive definite
- $\|A\| = 3.09 \times 10^2$
- $\delta_n = \min|\lambda_i - \mu| = O(10^{-8})$,
  rest $\delta_i = O(1) \sim O(10^2)$.
- $\kappa((A - \mu I)^{-1}) = O(10^{10})$



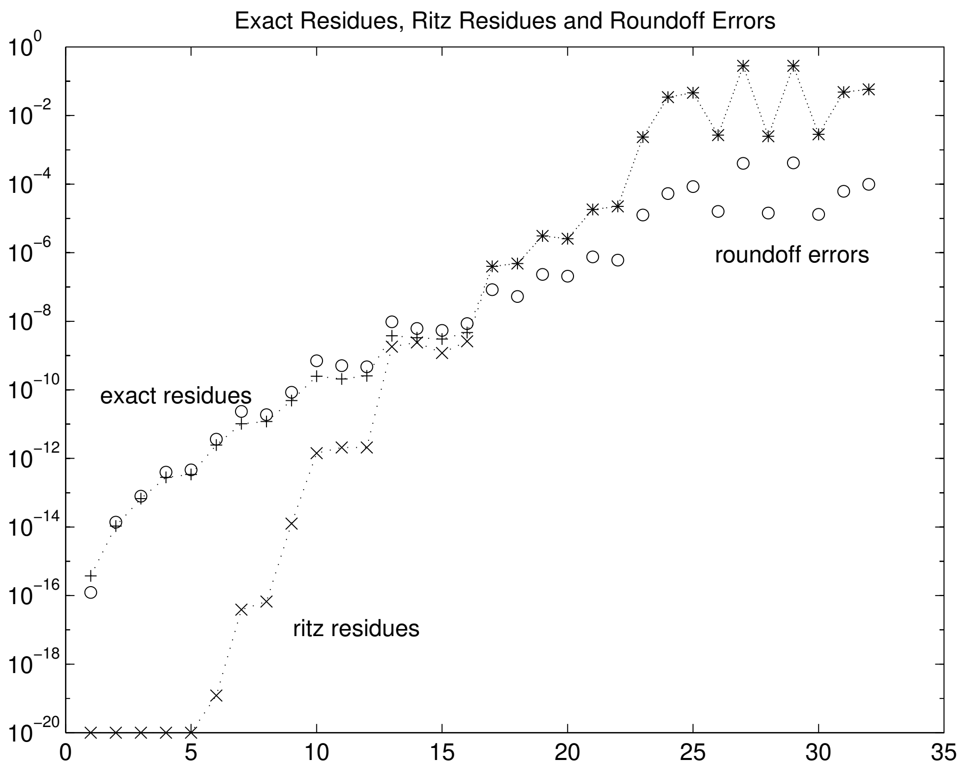
Grcar matrix:

- 30×30, nonsymmetric Toeplitz
- $\|A\| = 3.22$, $\Delta(A) = 5.8002$
- $\mathrm{cond}(\lambda_i) = O(10) \sim O(10^3)$
- $\delta_n = \min |\lambda_i - \mu| = O(10^{-8})$
- $\Delta((A - \mu I)^{-1}) = 4.03 \times 10^9$
- $\kappa((A - \mu I)^{-1}) = 3.505 \times 10^{10}$



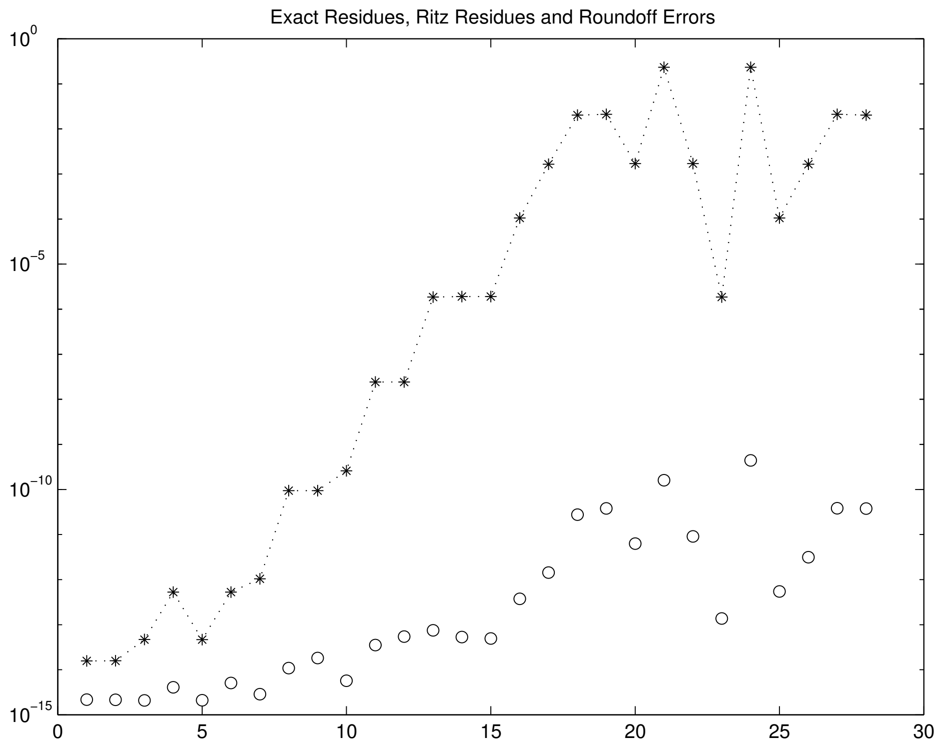
Orr-Sommerfeld Model Operator in Hydrodynamics I

- 64×64 ($\alpha = 1, R = 1500$), complex symmetric
- $\kappa(A) = 44.42$, $\Delta(A) = 4.92$
- $\mathrm{cond}(\lambda_i) = O(10^9)$ for the desired λ_i
- $\delta_n = \min |\lambda_i - \mu| = O(10^{-2})$.
- $\kappa((A - \mu I)^{-1}) = 3.12 \times 10^{14}$
- $\Delta((A - \mu I)^{-1}) = 1.15 \times 10^{13}$



Orr-Sommerfeld Model Operator in Hydrodynamics II

- $\delta_n = \min|\lambda_i - \mu| = O(10^{-1})$.
- $\kappa((A-\mu I)^{-1}) = 2.32 \times 10^7$
- $\Delta((A-\mu I)^{-1}) = 5.87 \times 10^6$



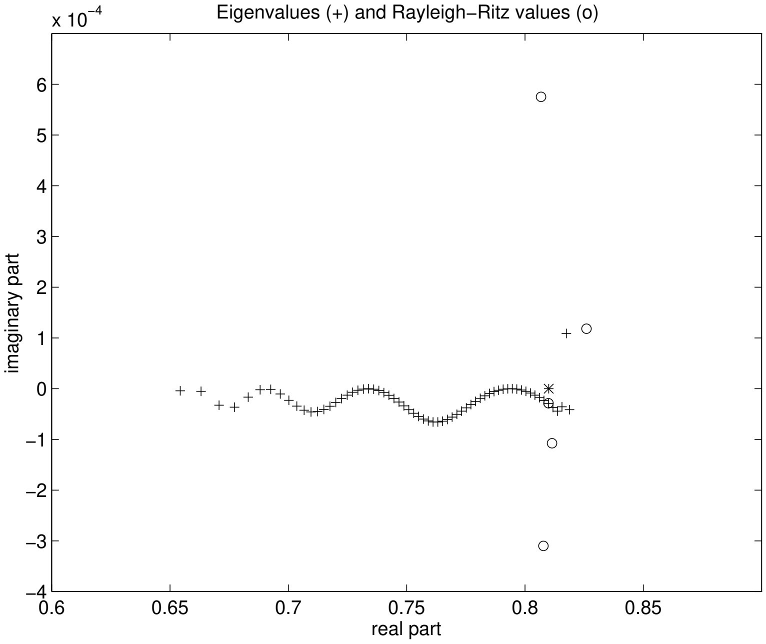
QC2534-sub: a complex scaled Schördinger operator

- 362×362, complex symmetric
- $\Delta(A) = 0.059$, $\kappa(A) = 8.10$

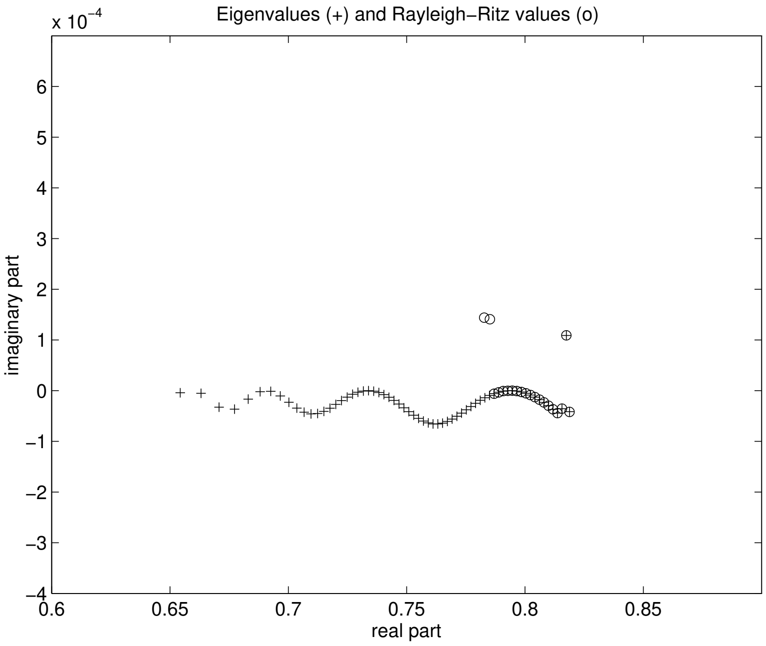
- With shift $\mu = 0.81$
  $\delta_n = \min |\lambda_i - \mu| = O(10^{-4})$
  $\kappa((A - \mu I)^{-1}) = O(10^{12})$
  $\Delta((A - \mu I)^{-1}) = O(10^{13})$

- With shift $\mu = 0.81 + 0.02i$
  $\delta_n = \min |\lambda_i - \mu| = O(10^{-2})$
  $\kappa((A - \mu I)^{-1}) = O(10^{2})$
  $\Delta((A - \mu I)^{-1}) = O(10^{2})$

$\mu = 0.81$



$\mu = 0.81 + 0.02i$



Phenomena

Upper Hessenberg matrix H_m

(a) If $(A - \mu I)^{-1}$ is ill-conditioned, the first two rows of H_m are large

$$\begin{bmatrix} \ell & \ell & \ell & \ell & \ell & \ell \\ \ell & \ell & \ell & \ell & \ell & \ell \\ & \times & \times & \times & \times & \times \\ & & \times & \times & \times & \times \\ & & & \times & \times & \times \\ & & & & \times & \times \end{bmatrix}$$

(b) Let $H_m z = z\theta + e$, then

$$|e_m^T z| \ll \mathbf{u}$$

for some z.

Exact residues $R_m^{(e)} = (r_i^{(e)})$ vs. Ritz $R_m^{(r)} = (r_i^{(r)})$

For some i

(a) $\|r_i^{(r)}\| \ll \mathbf{u}$

(b) $\|r_i^{(e)}\| \approx \|F_m\| + \|E_m\| \approx (\kappa(A - \mu I) + \|H_m\|)\mathbf{u}$

(APRACK's function `EIGS` in Matlab 5)

Not aware of very satisfactory explanation (analysis)

Accuracy

Residual:

$$\begin{aligned} Ax - (\mu + \frac{1}{\theta})x &= (A - \mu I) g_m (e_m^T z)\frac{1}{\theta} + \\ & \quad (A - \mu I)(V_m e + F_m z)\frac{1}{\theta}. \end{aligned}$$

Then there is an eigenvalue λ of A, up to the first order,

$$\begin{aligned} \left|\lambda - \left(\mu + \frac{1}{\theta}\right)\right| &\leq \frac{\mathrm{cond}(\lambda)}{|\theta|} \|(A - \mu I)\| \left[\|g_m\| \, |e_m^T z| + \right. \\ & \quad \left. \left(\|H_m\| + \kappa((A - \mu I)^{-1})\right) \mathbf{u}\right] \end{aligned}$$

Therefore, if $(A - \mu I)^{-1}$ is ill conditioned,

$$\left|\lambda - \left(\mu + \frac{1}{\theta}\right)\right| \lesssim \mathrm{cond}(\lambda) \frac{\|A - \mu I\|}{|\theta|} \kappa((A - \mu I)^{-1})\mathbf{u}$$

However, the accuracy of the eigenvalue nearest to the shift μ seems unaffected $\Leftarrow$ Inverse iteration ?

Conditioning of a Shift-and-invert Matrix

Fact 1: "Equivalence"

$$\|(A-\mu I)^{-1}\|-\frac{1}{\delta_n} \leq \Delta((A-\mu I)^{-1}) \leq \|(A-\mu I)^{-1}\|+\frac{1}{\delta_n}$$

Fact 2: If $\Delta(A) \leq \delta_n$ then

$$\frac{1}{\delta_n} \leq \|(A-\mu I)^{-1}\| \leq \frac{n}{\delta_n}$$

and

$$\Delta((A-\mu I)^{-1}) \leq \frac{n-1}{\delta_n}.$$

Variants of Henrici's bounds (1962)

Q:

- What to do with a "bad" shift?
- How to assess the quality of Ritz values/vectors?

The Second Shift Choice – A Global Approach

Let the new shift

$$\hat{\mu} = \mu + \sigma e^{i\phi}$$

such that

$$\Delta(A) \leq \min_i |\lambda_i - \hat{\mu}|.$$

For example, $\sigma = \Delta(A)$ and ϕ is chosen so that

$$\begin{aligned} \min_i |\lambda_i - \hat{\mu}| &= \min_i |\lambda_i - \mu| + |\sigma| \\ &= \min_i |\lambda_i - \mu| + \Delta(A). \end{aligned}$$

Q: How to know $\Delta(A)$? Which direction ϕ to move?

Calculation of pseudospectra by Arnoldi iteration

Fact:

$$\Lambda_\epsilon(\tilde{H}_1) \subseteq \Lambda_\epsilon(\tilde{H}_2) \subseteq \cdots \subseteq \Lambda_\epsilon(A)$$

`Manteuffel and Starke'95, Ruhe'95, Toh and Trefethen'96, Braconnier and Higham'96`

For the assessment of computed quantities, one should look these plots anyway!

The assumption $\Delta(A) \leq \delta_n$ can be relaxed as far as the conditioning of the shift-and-invert under control.

The Second Shift Choice – A Local Approach

Idea: move away *a little bit* from the desired λ when discovered μ is *too close.*

Let $\Pi(A - \mu I) = LU$, then in a neighborhood of μ,

$$\Pi(A - \xi) = L(\xi)U(\xi)$$

and

$$u_{nn}(\xi) = u_{nn} - (\xi_n^* - x^* y)(\xi - \mu) + O(|\xi - \mu|^2),$$

where $(x^*, \xi_n^*) \equiv e_n^T L^{-1}\Pi$, $y \equiv U^{-1}{}_{(1:n-1,1:n-1)}U_{(1:n-1,n)}$

Rank-revealing LU and differential LU

Take a Newton step

$$\lambda - \mu \approx u_{nn}/(\xi_n^* - x^* y) \equiv \delta\mu,$$

Q: How small is "*a little bit*"?

A plausible new shift might be $\mu - \frac{\delta\mu}{\sqrt{|\delta\mu|}}$.

A by-product: estimate λ's condition number cond(λ)

Final Remarks

- Conditioning of a shift-and-invert eigenproblem could be very sensitive to the shift.
- Ritz residues might be misleading to the convergence of a Krylov subspace method.
- Devices toward a better second shift seems feasible.

Goal:

Develop a few output *parameters* for assessing the reliability of computed quantities (not just simply Ritz values/vectors and their residues).

Excursion to the Gaussian elimination for $Ax = b$: condition number estimator, growth factor, iterative refinement, ..., and $Ax = \lambda x$ (Expert driver routines in LAPACK)