New Preconditioning Strategies for Eigenvalue Problems

Melina Freitag

Department of Mathematical Sciences
University of Bath

22nd Biennial Conference on Numerical Analysis, Dundee
Scotland, UK
26th June 2007

joint work with Howard Elman (Maryland) and Alastair Spence (Bath)

Outline

Problem and iterative methods

Find a small number of eigenvalues and eigenvectors of:

$$Ax = \lambda x, \quad \lambda \in \mathbb{C}, x \in \mathbb{C}^n$$

- A is large, sparse, nonsymmetric

Problem and iterative methods

Find a small number of eigenvalues and eigenvectors of:

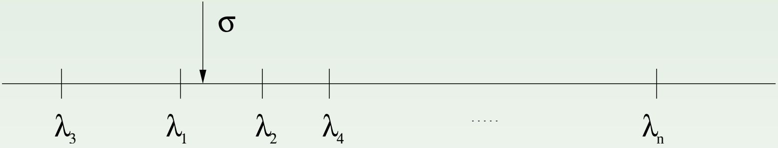
$$Ax = \lambda x, \quad \lambda \in \mathbb{C}, x \in \mathbb{C}^n$$

- A is large, sparse, nonsymmetric
- Iterative solves
  - Power method
  - Simultaneous iteration
  - Arnoldi method
  - Jacobi-Davidson method
- Usually involves repeated application of the matrix A to a vector
- Generally convergence to largest/outlying eigenvector

Outline

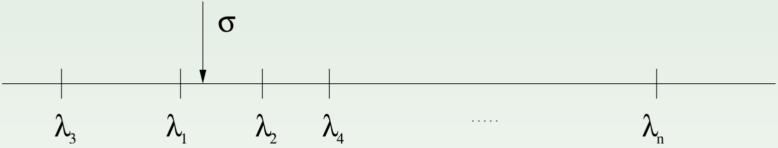
Shift-invert strategy

- Wish to find an eigenvalue close to a shift σ



Shift-invert strategy

- Wish to find an eigenvalue close to a shift σ



- Problem becomes
$$(A-\sigma I)^{-1}x = \frac{1}{\lambda-\sigma}x$$
- each step of the iterative method involves repeated application of $(A-\sigma I)^{-1}$ to a vector
- Inner iterative solve:
$$(A-\sigma I)y = x$$
using Krylov method for linear systems.
- leading to inner-outer iterative method.

Shift-invert strategy

This talk:
Inner Iteration and preconditioning
fixed shifts only.

Outline

1. The problem
2. Inverse iteration and Shift-invert Arnoldi method
3. Inexact inverse iteration
4. Inexact Shift-invert Arnoldi method
5. Conclusions

The algorithm

Inexact inverse iteration

for $i = 1$ to $\ldots$ **do**

choose $\tau^{(i)}$

solve

$$\|(A-\sigma I)y^{(i)} - x^{(i)}\| = \|d^{(i)}\| \leq \tau^{(i)},$$

Rescale $x^{(i+1)} = \dfrac{y^{(i)}}{\|y^{(i)}\|}$,

Update $\lambda^{(i+1)} = {x^{(i+1)}}^H A x^{(i+1)}$,

Test: eigenvalue residual $r^{(i+1)} = (A - \lambda^{(i+1)} I)x^{(i+1)}$.

end for

Convergence rates

If

$$\tau^{(i)} = C\|r^{(i)}\|$$

then convergence rate is linear (same convergence rate as for exact solves).

The inner iteration for $(A - \sigma I)y = x$

Standard GMRES theory for $y_0 = 0$ and A diagonalisable

$$\|x - (A - \sigma I)y_k\| \leq \kappa(W) \min_{p \in \mathcal{P}_k} \max_{j=1,\ldots,n} |p(\lambda_j)| \|x\|$$

where λ_j are eigenvalues of $A - \sigma I$ and $(A - \sigma I) = W \Lambda W^{-1}$.

The inner iteration for $(A - \sigma I)y = x$

Standard GMRES theory for $y_0 = 0$ and A diagonalisable

$$\|x - (A - \sigma I)y_k\| \leq \kappa(W) \min_{p \in \mathcal{P}_k} \max_{j=1,\dots,n} |p(\lambda_j)| \|x\|$$

where λ_j are eigenvalues of $A - \sigma I$ and $(A - \sigma I) = W \Lambda W^{-1}$.

Number of inner iterations

$$k \geq C_1 + C_2 \log \frac{\|x\|}{\tau}$$

for $\|x - (A - \sigma I)y_k\| \leq \tau$.

The inner iteration for $(A - \sigma I)y = x$

More detailed GMRES theory for $y_0 = 0$

$$\|x - (A - \sigma I)y_k\| \leq \tilde{\kappa}(W)\frac{|\lambda_2 - \lambda_1|}{\lambda_1} \min_{p \in \mathcal{P}_{k-1}} \max_{j=2,\dots,n} |p(\lambda_j)| \|\mathcal{Q}x\|$$

where λ_j are eigenvalues of $A - \sigma I$.

Number of inner iterations

$$k \geq C_1' + C_2' \log \frac{\|\mathcal{Q}x\|}{\tau},$$

where $\mathcal{Q}$ projects onto the space *not* spanned by the eigenvector.

The inner iteration for $(A - \sigma I)y = x$

Good news!

$$k^{(i)} \geq C_1' + C_2' \log \frac{C_3 \|r^{(i)}\|}{\tau^{(i)}},$$

where $\tau^{(i)} = C\|r^{(i)}\|$. Iteration number approximately constant!

The inner iteration for $(A-\sigma I)y = x$

Good news!

$$k^{(i)} \geq C_1' + C_2' \log \frac{C_3 \|r^{(i)}\|}{\tau^{(i)}},$$

where $\tau^{(i)} = C\|r^{(i)}\|$. Iteration number approximately constant!

Bad news :-(

For a standard preconditioner P

$$(A-\sigma I)P^{-1}\tilde{y}^{(i)} = x^{(i)} \qquad P^{-1}\tilde{y}^{(i)} = y^{(i)}$$

$$k^{(i)} \geq C_1'' + C_2'' \log \frac{\|\tilde{\mathcal{Q}} x^{(i)}\|}{\tau^{(i)}} = C_1'' + C_2'' \log \frac{C}{\tau^{(i)}},$$

where $\tau^{(i)} = C\|r^{(i)}\|$. Iteration number increases!

The inner iteration for $(A - \sigma I)P^{-1}\tilde{y} = x$

How to overcome this problem

- Use a different preconditioner, namely one that satisfies

$$\mathbb{P}_i x^{(i)} = Ax^{(i)}, \quad \mathbb{P}_i := P + (A - P)x^{(i)}{x^{(i)}}^H$$

- minor modification and minor extra computational cost,
- $[A\mathbb{P}_i^{-1}]Ax^{(i)} = Ax^{(i)}$.

The inner iteration for $(A - \sigma I)P^{-1}\tilde{y} = x$

How to overcome this problem

- Use a different preconditioner, namely one that satisfies

$$\mathbb{P}_i x^{(i)} = A x^{(i)}, \quad \mathbb{P}_i := P + (A - P) x^{(i)} {x^{(i)}}^H$$

- minor modification and minor extra computational cost,
- $[A\mathbb{P}_i^{-1}] A x^{(i)} = A x^{(i)}$.

Why does that work?

Assume we have found eigenvector x_1

$$A x_1 = \mathbb{P} x_1 = \lambda_1 x_1 \quad \Rightarrow \quad (A - \sigma I)\mathbb{P}^{-1} x_1 = \frac{\lambda_1 - \sigma}{\lambda_1} x_1$$

and convergence of Krylov method applied to $(A - \sigma I)\mathbb{P}^{-1}\tilde{y} = x_1$ in one iteration. For general $x^{(i)}$

$$k^{(i)} \geq C_1'' + C_2'' \log \frac{C_3 \| r^{(i)} \|}{\tau^{(i)}}, \quad \text{where} \quad \tau^{(i)} = C \| r^{(i)} \|.$$

Convection-Diffusion operator

Finite difference discretisation on a 32×32 grid of the convection-diffusion operator

$$-\Delta u + 5u_x + 5u_y = \lambda u \quad \text{on} \quad (0,1)^2,$$

with homogeneous Dirichlet boundary conditions (961×961 matrix).

- smallest eigenvalue: $\lambda_1 \approx 32.18560954$,
- Preconditioned GMRES with tolerance $\tau^{(i)} = 0.01\|r^{(i)}\|$,
- standard and tuned preconditioner (incomplete LU).

Convection-Diffusion operator

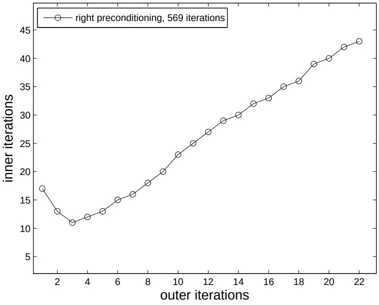


Figure: Inner iterations vs outer iterations

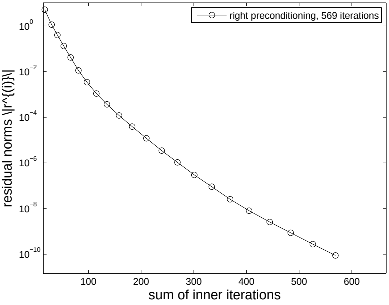


Figure: Eigenvalue residual norms vs total number of inner iterations

Convection-Diffusion operator

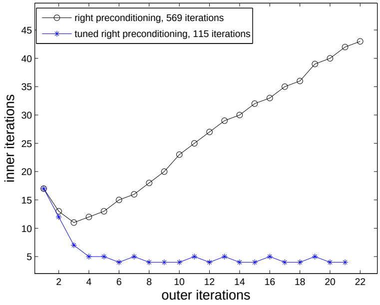


Figure: Inner iterations vs outer iterations

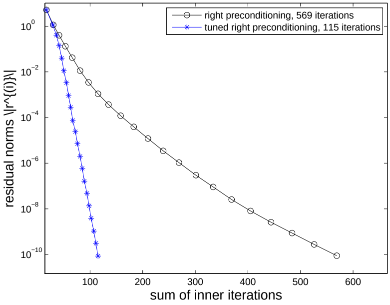


Figure: Eigenvalue residual norms vs total number of inner iterations

Outline

The algorithm

Shift-invert Arnoldi-method

- Arnoldi method constructs an orthogonal basis of k-dimensional Krylov subspace

$$\mathcal{K}^{(i)}(\mathcal{A}, q^{(1)}) = \text{span}\{q^{(1)}, \mathcal{A}q^{(1)}, \mathcal{A}^2 q^{(1)}, \ldots, \mathcal{A}^{i-1} q^{(1)}\},$$

$$\mathcal{A}Q^{(i)} = Q^{(i)} H^{(i)} + q^{(i+1)} h^{(i+1,i)} e_i^H, \quad (Q^{(i)})^H Q^{(i)} = I.$$

- at each step, application of $\mathcal{A}$ to $q^{(i)}$
- Eigenvalues of $H^{(i)}$ are eigenvalue approximation of (outlying) eigenvalues of $\mathcal{A}$
- Shift-Invert $\mathcal{A} := (A - \sigma I)^{-1}$

The algorithm

Shift-invert Arnoldi-method

- Arnoldi method constructs an orthogonal basis of k-dimensional Krylov subspace

$$\mathcal{K}^{(i)}(\mathcal{A}, q^{(1)}) = \text{span}\{q^{(1)}, \mathcal{A}q^{(1)}, \mathcal{A}^2 q^{(1)}, \ldots, \mathcal{A}^{i-1} q^{(1)}\},$$

$$\mathcal{A}Q^{(i)} = Q^{(i)} H^{(i)} + q^{(i+1)} h^{(i+1,i)} e_i^H, \quad (Q^{(i)})^H Q^{(i)} = I.$$

- at each step, application of $\mathcal{A}$ to $q^{(i)}$
- Eigenvalues of $H^{(i)}$ are eigenvalue approximation of (outlying) eigenvalues of $\mathcal{A}$
- Shift-Invert $\mathcal{A} := (A - \sigma I)^{-1}$

Inexact solves

The solve tolerance can be relaxed! (Simoncini 2005)

$$\|q^{(i)} - (A - \sigma I)\tilde{q}^{(i+1)}\| = \|d^{(i)}\|, \quad \|d^{(i)}\| = C \frac{1}{\|r^{(i)}\|}.$$

The inner iteration for $A\tilde{q} = q$ $(\sigma = 0)$

Tuning for Arnoldi's method (consider $\sigma = 0$)

Solve

$$AP^{-1}\tilde{\tilde{q}} = q, \quad P^{-1}\tilde{\tilde{q}} = \tilde{q}$$

using a tuned preconditioner

$$\mathbb{P}_i Q^{(i)} = AQ^{(i)}; \quad \text{given by} \quad \mathbb{P}_i = P + (A - P)Q^{(i)}{Q^{(i)}}^H$$

The inner iteration for $A\tilde{q} = q$ $(\sigma = 0)$

Tuning for Arnoldi's method (consider $\sigma = 0$)

Solve

$$AP^{-1}\tilde{\tilde{q}} = q, \quad P^{-1}\tilde{\tilde{q}} = \tilde{q}$$

using a tuned preconditioner

$$\mathbb{P}_i Q^{(i)} = AQ^{(i)}; \quad \text{given by} \quad \mathbb{P}_i = P + (A - P)Q^{(i)}{Q^{(i)}}^H$$

Theorem (Properties of the tuned preconditioner)

Let P with $P = A + E$ be a preconditioner for A and assume i steps of Arnoldi's method have been carried out; then i eigenvalues of $A\mathbb{P}_i^{-1}$ are equal to one:

$$[A\mathbb{P}_i^{-1}]AQ^{(i)} = AQ^{(i)}$$

and $n - i$ eigenvalues are perturbations one, close to the eigenvalues of AP^{-1}.

Numerical Example

`sherman5.mtx` matrix from the Matrix Market library (3312×3312).

- smallest eigenvalue: $\lambda_1 \approx 4.69 \times 10^{-2}$,
- Preconditioned GMRES as inner solver (fixed tolerance and relaxation strategy),
- standard and tuned preconditioner (incomplete LU).

No tuning and standard preconditioner

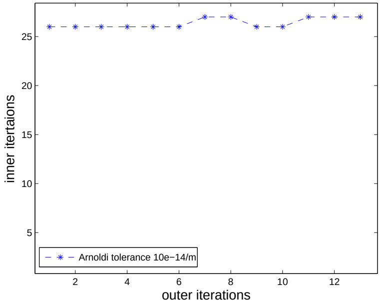


Figure: Inner iterations vs outer iterations

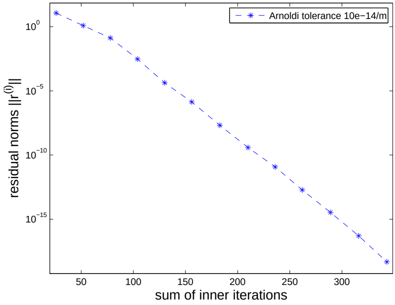


Figure: Eigenvalue residual norms vs total number of inner iterations

Tuning

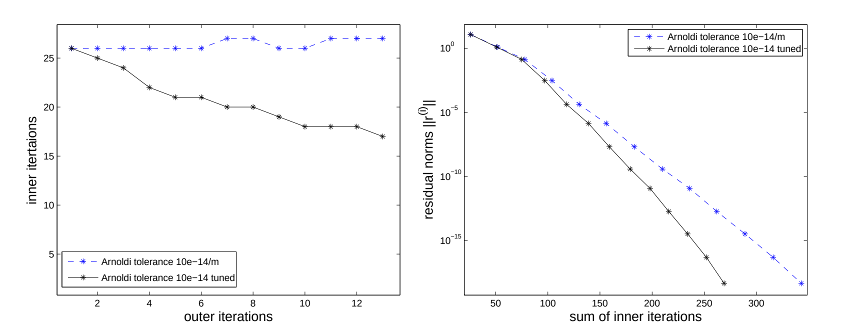


Figure: Inner iterations vs outer iterations

Figure: Eigenvalue residual norms vs total number of inner iterations

Relaxation

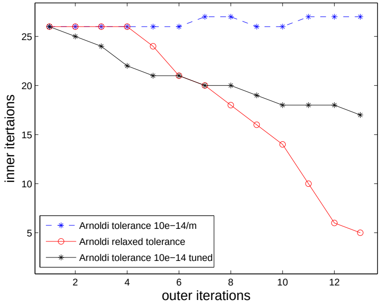


Figure: Inner iterations vs outer iterations

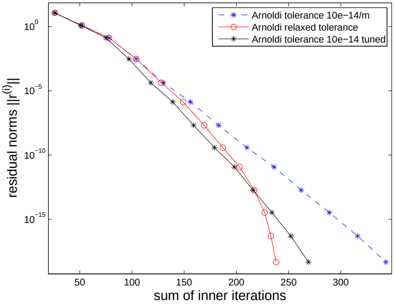


Figure: Eigenvalue residual norms vs total number of inner iterations

Tuning and relaxation strategy

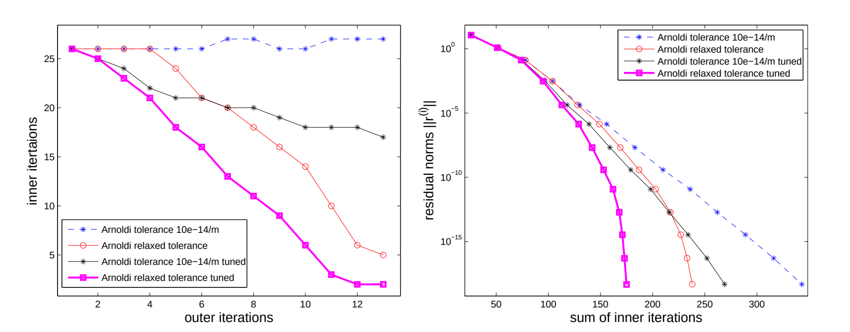


Figure: Inner iterations vs outer iterations

Figure: Eigenvalue residual norms vs total number of inner iterations

Outline

Conclusions

- Inexact solves for Inverse iteration and Krylov subspace methods
- Inner iteration depends on preconditioner
- For eigencomputations it is advantageous to consider modified preconditioners (works for any preconditioner)
- Extension to restarted methods is possible

M. A. Freitag and H. C. Elman, *Inexact preconditioned arnoldi's method and implicit restarts for eigenvalue computations*, 2007.
In preparation.

M. A. Freitag and A. Spence, *Convergence theory for inexact inverse iteration applied to the generalised nonsymmetric eigenproblem*, 2006.
To appear in ETNA.

——, *A tuned preconditioner for inexact inverse iteration applied to Hermitian eigenvalue problems*, 2006.
Submitted to IMAJNA.

——, *Convergence rates for inexact inverse iteration with application to preconditioned iterative solves*, BIT, 47 (2007), pp. 27–44.