Inexact preconditioned shift-and-invert Arnoldi's method and implicit restarts for eigencomputations

Melina Freitag

Department of Mathematical Sciences
University of Bath

Householder Symposium XVII
Zeuthen, Germany
June 2008

joint work with Alastair Spence (Bath)

Problem and iterative methods

Find a small number of eigenvalues close to a shift σ and corresponding eigenvectors of:

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

- A is large, sparse, nonsymmetric $\Rightarrow$ iterative solves (e.g. Arnoldi method)

Problem and iterative methods

Find a small number of eigenvalues close to a shift σ and corresponding eigenvectors of:

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

- A is large, sparse, nonsymmetric $\Rightarrow$ iterative solves (e.g. Arnoldi method)
- 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

Problem and iterative methods

Find a small number of eigenvalues close to a shift σ and corresponding eigenvectors of:

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

- A is large, sparse, nonsymmetric $\Rightarrow$ iterative solves (e.g. Arnoldi method)
- 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 or Galerkin-Krylov method for linear systems.
- leading to inner-outer iterative method.

The algorithm

Arnoldi's method

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

$$\mathcal{K}_k(\mathcal{A}, q_1) = \text{span}\{q_1, \mathcal{A}q_1, \mathcal{A}^2 q_1, \dots, \mathcal{A}^{k-1} q_1\},$$

$$\mathcal{A}Q_k = Q_k H_k + q_{k+1} h_{k+1,k} e_k^H = Q_{k+1} \begin{bmatrix} H_k \\ h_{k+1,k} e_k^H \end{bmatrix}$$

$$Q_k^H Q_k = I.$$

- Eigenvalues of the upper Hessenberg matrix H_k are eigenvalue approximations of ("outlying") eigenvalues of $\mathcal{A}$

$$\|r_k\| = \|\mathcal{A}x - \theta x\| = \|(\mathcal{A}Q_k - Q_k H_k)u\| = |h_{k+1,k}||e_k^H u|,$$

- at each step: application of $\mathcal{A}$ to q_k:

$$\mathcal{A}q_k = \tilde{q}_{k+1}$$

Enhancements: Shift-Invert Arnoldi and IRA

Shift-Invert Arnoldi's method $\mathcal{A} := A^{-1}$ $(\sigma = 0)$

- Arnoldi factorisation

$$A^{-1}Q_k = Q_k H_k + q_{k+1} h_{k+1,k} e_k^H = Q_{k+1} \begin{bmatrix} H_k \\ h_{k+1,k} e_k^H \end{bmatrix}$$

- Eigenvalues of the upper Hessenberg matrix H_k are eigenvalue approximations of ("outlying") eigenvalues of A^{-1}
- at each step: application of A^{-1} to q_k: $A^{-1} q_k = \tilde{q}_{k+1}$.

Enhancements: Shift-Invert Arnoldi and IRA

Shift-Invert Arnoldi's method $\mathcal{A} := A^{-1}$ $(\sigma = 0)$

- Arnoldi factorisation

$$A^{-1}Q_k = Q_k H_k + q_{k+1} h_{k+1,k} e_k^H = Q_{k+1} \begin{bmatrix} H_k \\ h_{k+1,k} e_k^H \end{bmatrix}$$

- Eigenvalues of the upper Hessenberg matrix H_k are eigenvalue approximations of ("outlying") eigenvalues of A^{-1}
- at each step: application of A^{-1} to q_k: $A^{-1}q_k = \tilde{q}_{k+1}$.

Implicitly Restarted Arnoldi

- perform $m = k + p$ Arnoldi iterations
- IRA: restart from step k: $\mathcal{A}Q_k = Q_k H_k + q_{k+1} \underbrace{h_{k+1,k}}_{\to 0} e_k^H$

This talk

Extend the results by Simoncini (2005) for Arnoldi to IRA

Extend the idea of tuning (previous talk) to Arnoldi and IRA

Inexact solves (Simoncini 2005), Bouras and Frayssé (2000)

- Wish to solve

$$\|q_k - A\tilde{q}_{k+1}\| = \|\tilde{d}_k\| \leq \tau_k$$

Inexact solves (Simoncini 2005), Bouras and Frayssé (2000)

- Wish to solve
$$\|q_k - A\tilde{q}_{k+1}\| = \|\tilde{d}_k\| \leq \tau_k$$
- after m steps leads to inexact Arnoldi relation

$$A^{-1}Q_m = Q_{m+1}\begin{bmatrix} H_m \\ h_{m+1,m}e_k^H \end{bmatrix} + D_m = Q_{m+1}\begin{bmatrix} H_m \\ h_{m+1,m}e_m^H \end{bmatrix} + [d_1|\dots|d_m]$$

UNIVERSITY OF BATH

Inexact solves (Simoncini 2005), Bouras and Frayssé (2000)

- Wish to solve

$$\|q_k - A\tilde{q}_{k+1}\| = \|\tilde{d}_k\| \leq \tau_k$$

- after m steps leads to inexact Arnoldi relation

$$A^{-1}Q_m = Q_{m+1}\begin{bmatrix} H_m \\ h_{m+1,m}e_k^H \end{bmatrix} + D_m = Q_{m+1}\begin{bmatrix} H_m \\ h_{m+1,m}e_m^H \end{bmatrix} + [d_1|\dots|d_m]$$

- u eigenvector of H_m:

$$\|r_m\| = \|(A^{-1}Q_m - Q_mH_m)u\| = |h_{m+1,m}||e_m^Hu| + D_mu,$$

$$D_mu = \sum_{k=1}^{m} d_ku_k, \quad \text{if} \quad |u_k| \quad \text{small, then} \quad \|d_k\| \quad \text{allowed to be large!}$$

Inexact solves (Simoncini 2005), Bouras and Frayssé (2000)

- Wish to solve
$$\|q_k - A\tilde{q}_{k+1}\| = \|\tilde{d}_k\| \leq \tau_k$$
- after m steps leads to inexact Arnoldi relation
$$A^{-1}Q_m = Q_{m+1}\begin{bmatrix} H_m \\ h_{m+1,m}e_k^H \end{bmatrix} + D_m = Q_{m+1}\begin{bmatrix} H_m \\ h_{m+1,m}e_m^H \end{bmatrix} + [d_1|\dots|d_m]$$
- u eigenvector of H_m:
$$\|r_m\| = \|(A^{-1}Q_m - Q_mH_m)u\| = |h_{m+1,m}||e_m^Hu| + D_mu,$$
$$D_mu = \sum_{k=1}^{m} d_ku_k, \quad \text{if} \quad |u_k| \quad \text{small, then} \quad \|d_k\| \quad \text{allowed to be large!}$$
- Simoncini (2005) has shown
$$|u_k| \leq C(k,m)\|r_{k-1}\|$$
which leads to
$$\|\tilde{d}_k\| = C(k,m)\frac{1}{\|r_{k-1}\|}\varepsilon$$
for $\|D_mu\| < \varepsilon$.

Numerical Examples

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

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

Fixed tolerance

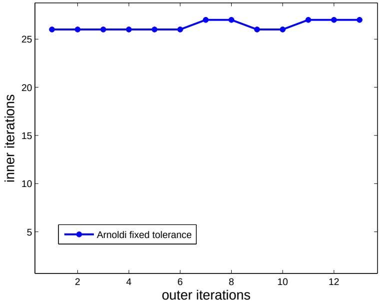

Figure: Inner iterations vs outer iterations

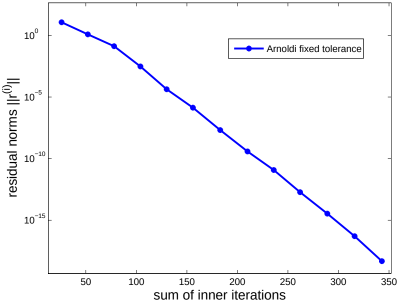

Figure: Eigenvalue residual norms vs total number of inner iterations

Relaxation (Simoncini 2005)

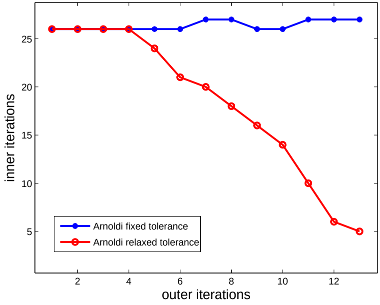

Figure: Inner iterations vs outer iterations

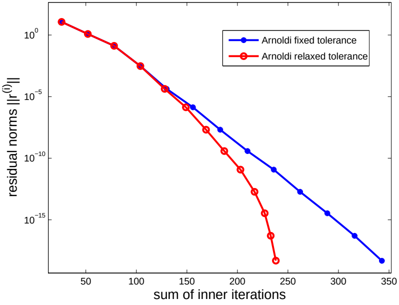

Figure: Eigenvalue residual norms vs total number of inner iterations

Relaxation strategy for invariant subspaces (F./Spence 2008)

- $m = k + p$ steps of the Arnoldi factorisation

$$\mathcal{A}Q_{k+p} = Q_{k+p}H_{k+p} + q_{k+p+1}h_{k+p+1,k+p}e_{k+p}^H$$

- let H_m have Schur decomposition

$$H_m = H_{k+p} = \begin{bmatrix} U & W_2 \end{bmatrix} \begin{bmatrix} \Theta & \star \\ 0 & T_{22} \end{bmatrix} \begin{bmatrix} U & W_2 \end{bmatrix}^H$$

Relaxation strategy for invariant subspaces (F./Spence 2008)

- $m = k + p$ steps of the Arnoldi factorisation

$$\mathcal{A}Q_{k+p} = Q_{k+p}H_{k+p} + q_{k+p+1}h_{k+p+1,k+p}e_{k+p}^H$$

- let H_m have Schur decomposition

$$H_m = H_{k+p} = \begin{bmatrix} U & W_2 \end{bmatrix} \begin{bmatrix} \Theta & \star \\ 0 & T_{22} \end{bmatrix} \begin{bmatrix} U & W_2 \end{bmatrix}^H$$

- let H_k be decomposed as $\Theta_k = {U_k}^H H_k U_k$
- let $R_k = q_{k+1}h_{k+1,k}e_k^H U_k$ be the residual after k Arnoldi steps.

Relaxation strategy for invariant subspaces (F./Spence 2008)

- $m = k + p$ steps of the Arnoldi factorisation

$$\mathcal{A}Q_{k+p} = Q_{k+p}H_{k+p} + q_{k+p+1}h_{k+p+1,k+p}e_{k+p}^H$$

- let H_m have Schur decomposition

$$H_m = H_{k+p} = \begin{bmatrix} U & W_2 \end{bmatrix} \begin{bmatrix} \Theta & \star \\ 0 & T_{22} \end{bmatrix} \begin{bmatrix} U & W_2 \end{bmatrix}^H$$

- let H_k be decomposed as $\Theta_k = {U_k}^H H_k U_k$
- let $R_k = q_{k+1}h_{k+1,k}e_k^H U_k$ be the residual after k Arnoldi steps.
- Then $U = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$ with $U^H U = I$, such that

$$\|U_2\| \leq \frac{\|R_k\|}{\text{sep}(T_{22}, \Theta_k)}$$

where $\text{sep}(T_{22}, \Theta_k) := \min_{\|V\|=1} \|T_{22}V - V\Theta_k\|$.

Relaxation strategy for IRA (F./Spence 2008)

Theorem

For any given $\varepsilon \in \mathbb{R}$ *with* $\varepsilon > 0$ *assume that*

$$\|d_l\| \leq \begin{cases} \dfrac{\varepsilon}{2(m-k)} \dfrac{sep(T_{22}, \Theta_k)}{\|R_k\|} & \text{if} \quad l > k, \\ \dfrac{\varepsilon}{2k} & \text{otherwise.} \end{cases}$$

Then

$$\|\mathcal{A}Q_m U - Q_m U \Theta - R_m\| \leq \varepsilon.$$

Relaxation strategy for IRA (F./Spence 2008)

Theorem

For any given $\varepsilon \in \mathbb{R}$ *with* $\varepsilon > 0$ *assume that*

$$\|d_l\| \leq \begin{cases} \dfrac{\varepsilon}{2(m-k)} \dfrac{sep(T_{22}, \Theta_k)}{\|R_k\|} & \text{if} \quad l > k, \\ \dfrac{\varepsilon}{2k} & \text{otherwise.} \end{cases}$$

Then

$$\|\mathcal{A} Q_m U - Q_m U \Theta - R_m\| \leq \varepsilon.$$

- In practice: perform $m = k + p$ initial steps and then relax the tolerance from the first restart.

Numerical Example

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

- $k = 8$ eigenvalues closest to zero
- IRA with exact shifts $p = 4$
- Preconditioned GMRES as inner solver (fixed tolerance and relaxation strategy),
- standard and tuned preconditioner (incomplete LU).

Fixed tolerance

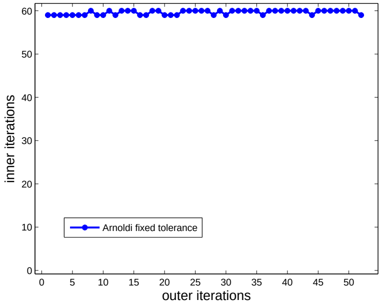

Figure: Inner iterations vs outer iterations

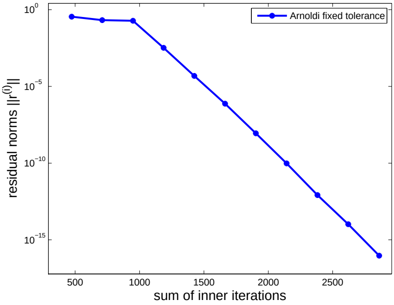

Figure: Eigenvalue residual norms vs total number of inner iterations

Relaxation

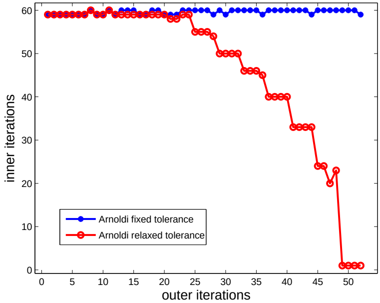

Figure: Inner iterations vs outer iterations

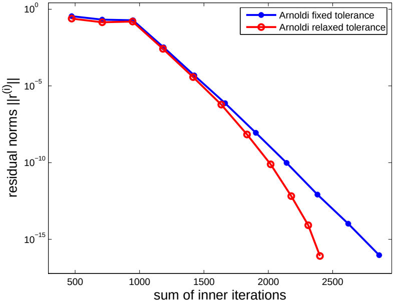

Figure: Eigenvalue residual norms vs total number of inner iterations

Tuning the preconditioner $AP^{-1}\tilde{q}_{k+1} = q_k$

- Introduce preconditioner P and solve

$$AP^{-1}\tilde{q}_{k+1} = q_k, \quad P^{-1}\tilde{q}_{k+1} = q_{k+1}$$

using GMRES:

$$\|d_l\| = \kappa \min_{p \in \Pi_l} \max_{i=1,\dots,n} |p(\mu_i)| \|d_0\|$$

Tuning the preconditioner $AP^{-1}\tilde{q}_{k+1} = q_k$

- Introduce preconditioner P and solve

$$AP^{-1}\tilde{q}_{k+1} = q_k, \quad P^{-1}\tilde{q}_{k+1} = q_{k+1}$$

  using GMRES:

$$\|d_l\| = \kappa \min_{p \in \Pi_l} \max_{i=1,\dots,n} |p(\mu_i)| \|d_0\|$$

- use a tuned preconditioner for Arnoldi's method

$$\mathbb{P}_k Q_k = AQ_k; \quad \text{given by} \quad \mathbb{P}_k = P + (A - P)Q_k Q_k^H$$

The inner iteration for $AP^{-1}\tilde{q}_{k+1} = q_k$

Theorem (Properties of the tuned preconditioner $\mathbb{P}_k Q_k = AQ_k$)

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

$$[A\mathbb{P}_k^{-1}]AQ_k = AQ_k$$

and $n-k$ eigenvalues equivalent to eigenvalues of $L \in \mathbb{C}^{n-k\times n-k}$ with

$$\|L - I\| \leq C\|E\|.$$

The inner iteration for $AP^{-1}\tilde{q}_{k+1} = q_k$

Theorem (Properties of the tuned preconditioner $\mathbb{P}_k Q_k = AQ_k$)

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

$$[A\mathbb{P}_k^{-1}]AQ_k = AQ_k$$

and $n-k$ eigenvalues equivalent to eigenvalues of $L \in \mathbb{C}^{n-k \times n-k}$ with

$$\|L - I\| \leq C\|E\|.$$

Implementation

- Sherman-Morrison-Woodbury.
- Only minor extra costs (one back substitution per outer iteration)

Tuning

Why does tuning help?

- Arnoldi decomposition

$$A^{-1}Q_k = Q_k H_k + q_{k+1} h_{k+1,k} e_k^H$$

Tuning

Why does tuning help?

- Arnoldi decomposition

$$A^{-1}Q_k = Q_k H_k + q_{k+1} h_{k+1,k} e_k^H$$

- let A^{-1} be transformed into upper Hessenberg form

$$\begin{bmatrix} Q_k & Q_k{}^{\perp} \end{bmatrix}^H A^{-1} \begin{bmatrix} Q_k & Q_k{}^{\perp} \end{bmatrix} = \begin{bmatrix} H_k & T_{12} \\ h_{k+1,k} e_1 e_k{}^H & T_{22} \end{bmatrix},$$

where $\begin{bmatrix} Q_k & Q_k{}^{\perp} \end{bmatrix}$ is unitary and $H_k \in \mathbb{C}^{k,k}$ and $T_{22} \in \mathbb{C}^{n-k,n-k}$ are upper Hessenberg.

Tuning

Why does tuning help?

- Arnoldi decomposition

$$A^{-1}Q_k = Q_k H_k + q_{k+1} h_{k+1,k} e_k^H$$

- let A^{-1} be transformed into upper Hessenberg form

$$\begin{bmatrix} Q_k & Q_k{}^{\perp} \end{bmatrix}^H A^{-1} \begin{bmatrix} Q_k & Q_k{}^{\perp} \end{bmatrix} = \begin{bmatrix} H_k & T_{12} \\ h_{k+1,k} e_1 e_k{}^H & T_{22} \end{bmatrix},$$

where $\begin{bmatrix} Q_k & Q_k{}^{\perp} \end{bmatrix}$ is unitary and $H_k \in \mathbb{C}^{k,k}$ and $T_{22} \in \mathbb{C}^{n-k,n-k}$ are upper Hessenberg.

If $h_{k+1,k} \neq 0$ then

$$\begin{bmatrix} Q_k & Q_k{}^{\perp} \end{bmatrix}^H A\mathbb{P}_k^{-1} \begin{bmatrix} Q_k & Q_k{}^{\perp} \end{bmatrix} = \begin{bmatrix} I + \star & Q_k^H A\mathbb{P}_k^{-1} Q_k^{\perp} \\ \star & T_{22}^{-1}(Q_k^{\perp H} P Q_k^{\perp})^{-1} + \star \end{bmatrix}$$

Tuning

Why does tuning help?

- Assume we have found an approximate invariant subspace, that is

$$A^{-1}Q_k = Q_k H_k + \underbrace{q_{k+1}h_{k+1,k}e_k^H}_{\approx 0}$$

- let A^{-1} have the upper Hessenberg form

$$\left[\begin{array}{cc} Q_k & {Q_k}^{\perp} \end{array}\right]^H A^{-1} \left[\begin{array}{cc} Q_k & {Q_k}^{\perp} \end{array}\right] = \left[\begin{array}{cc} H_k & T_{12} \\ h_{k+1,k}e_1{e_k}^H & T_{22} \end{array}\right],$$

where $\left[\begin{array}{cc} Q_k & {Q_k}^{\perp} \end{array}\right]$ is unitary and $H_k \in \mathbb{C}^{k,k}$ and $T_{22} \in \mathbb{C}^{n-k,n-k}$ are upper Hessenberg.

If $h_{k+1,k} = 0$ then

$$\left[\begin{array}{cc} Q_k & {Q_k}^{\perp} \end{array}\right]^H A\mathbb{P}_k^{-1} \left[\begin{array}{cc} Q_k & {Q_k}^{\perp} \end{array}\right] = \left[\begin{array}{cc} I & Q_k^H A\mathbb{P}_k^{-1}Q_k^{\perp} \\ 0 & T_{22}^{-1}({Q_k^{\perp}}^H P Q_k^{\perp})^{-1} \end{array}\right]$$

Tuning

Another advantage of tuning

- System to be solved at each step of Arnoldi's method is

$$A\mathbb{P}_k^{-1}\tilde{q}_{k+1} = q_k, \quad \mathbb{P}_k^{-1}\tilde{q}_{k+1} = q_{k+1}$$

Tuning

Another advantage of tuning

- System to be solved at each step of Arnoldi's method is

$$A\mathbb{P}_k^{-1}\tilde{q}_{k+1} = q_k, \quad \mathbb{P}_k^{-1}\tilde{q}_{k+1} = q_{k+1}$$

- Assuming invariant subspace found then ($A^{-1}Q_k = Q_k H_k$):

$$A\mathbb{P}_k^{-1}q_k = q_k$$

Tuning

Another advantage of tuning

- System to be solved at each step of Arnoldi's method is

$$A\mathbb{P}_k^{-1}\tilde{q}_{k+1} = q_k, \quad \mathbb{P}_k^{-1}\tilde{q}_{k+1} = q_{k+1}$$

- Assuming invariant subspace found then ($A^{-1}Q_k = Q_k H_k$):

$$A\mathbb{P}_k^{-1}q_k = q_k$$

- the right hand side of the system matrix is an eigenvector of the system matrix!
- Krylov methods converge in one iteration

Numerical Example (Arnoldi)

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

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

Relaxation

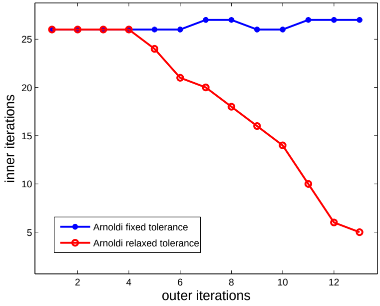

Figure: Inner iterations vs outer iterations

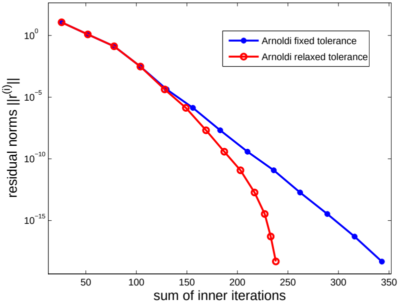

Figure: Eigenvalue residual norms vs total number of inner iterations

Tuning the preconditioner

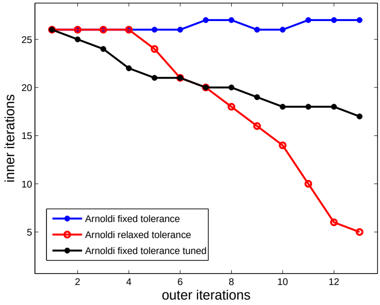

Figure: Inner iterations vs outer iterations

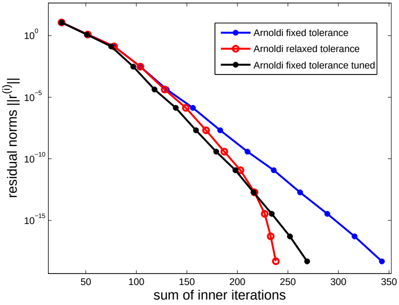

Figure: Eigenvalue residual norms vs total number of inner iterations

Tuning and relaxation strategy

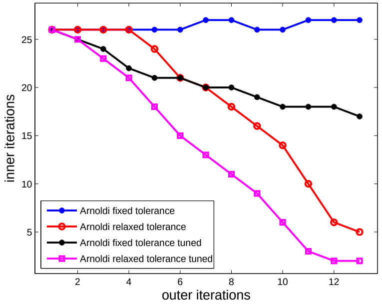

Figure: Inner iterations vs outer iterations

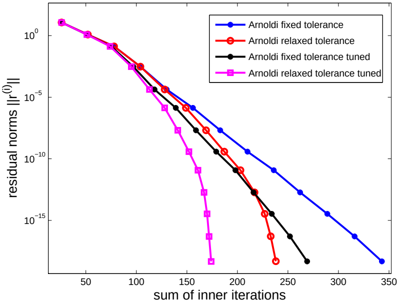

Figure: Eigenvalue residual norms vs total number of inner iterations

Numerical Example (IRA)

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

- $k = 8$ eigenvalues closest to zero
- IRA with exact shifts $p = 4$
- Preconditioned GMRES as inner solver (fixed tolerance and relaxation strategy),
- standard and tuned preconditioner (incomplete LU).

Relaxation

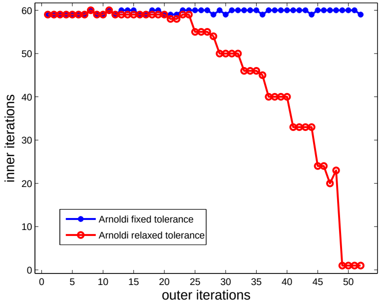

Figure: Inner iterations vs outer iterations

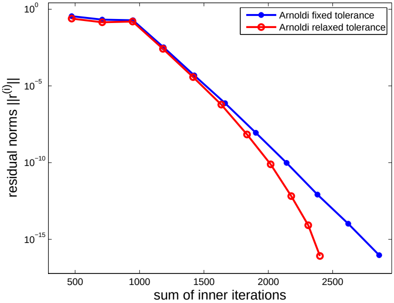

Figure: Eigenvalue residual norms vs total number of inner iterations

Tuning

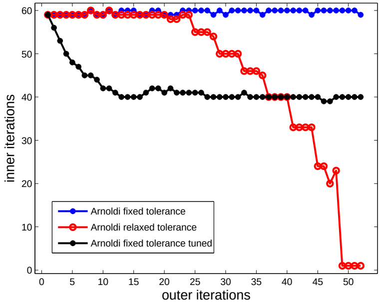

Figure: Inner iterations vs outer iterations

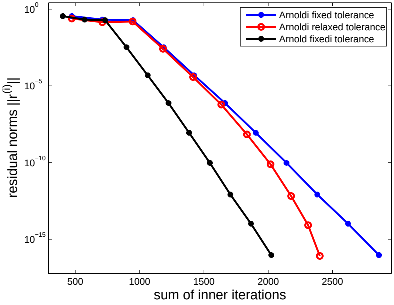

Figure: Eigenvalue residual norms vs total number of inner iterations

Tuning and relaxation strategy

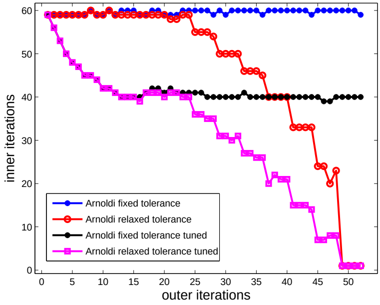

Figure: Inner iterations vs outer iterations

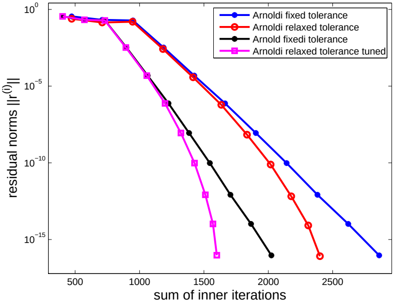

Figure: Eigenvalue residual norms vs total number of inner iterations

Conclusions

- For eigencomputations it is advantageous to consider small rank changes to the standard preconditioners (works for any preconditioner)
- Extension of the relaxation strategy to IRA
- Best results are obtained when relaxation and tuning are combined
- Link to Jacobi-Davidson method?