Preconditioned inverse iteration and shift-invert Arnoldi method

Melina Freitag

Department of Mathematical Sciences
University of Bath

CSC Seminar
Max-Planck-Institute for Dynamics of Complex Technical Systems
Magdeburg
3rd May 2011

joint work with Alastair Spence (Bath)
1 Introduction

2 Inexact inverse iteration

3 Inexact Shift-invert Arnoldi method

4 Inexact Shift-invert Arnoldi method with implicit restarts

5 Conclusions
Outline

1 Introduction

2 Inexact inverse iteration

3 Inexact Shift-invert Arnoldi method

4 Inexact Shift-invert Arnoldi method with implicit restarts

5 Conclusions
Problem and iterative methods

Find a small number of eigenvalues and eigenvectors of:

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

- \( A \) 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

- repeated application of the matrix \( A \) to a vector
- Generally **convergence to largest/outlying eigenvector**
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
- The first three of these involve repeated application of the matrix \( A \) to a vector
- Generally **convergence to largest/outlying eigenvector**
Shift-invert strategy

- Wish to find a few eigenvalues close to a shift $\sigma$
Shift-invert strategy

- Wish to find a few eigenvalues close to a shift $\sigma$

```
\sigma
```

- Problem becomes

$$\begin{align*}
(A - \sigma I)^{-1} x &= \frac{1}{\lambda - \sigma} x \\
(A - \sigma I) y &= x
\end{align*}$$

- each step of the iterative method involves repeated application of $A = (A - \sigma I)^{-1}$ 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

Inverse iteration and Arnoldi method
Outline

1 Introduction

2 Inexact inverse iteration

3 Inexact Shift-invert Arnoldi method

4 Inexact Shift-invert Arnoldi method with implicit restarts

5 Conclusions
The algorithm

Inexact inverse iteration

\[
\text{for } i = 1 \text{ to } \ldots \text{ do}
\]
\begin{align*}
\text{choose } &\tau^{(i)} \\
\text{solve } &\| (A - \sigma I) y^{(i)} - x^{(i)} \| = \| d^{(i)} \| \leq \tau^{(i)}, \\
\text{Rescale } &x^{(i+1)} = \frac{y^{(i)}}{\| y^{(i)} \|}, \\
\text{Update } &\lambda^{(i+1)} = x^{(i+1)^H} A x^{(i+1)}, \\
\text{Test: eigenvalue residual } &r^{(i+1)} = (A - \lambda^{(i+1)} I) x^{(i+1)}. \\
\text{end for}
\]
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 \( x^{(i+1)} = \frac{y^{(i)}}{\| y^{(i)} \|} \),
  Update \( \lambda^{(i+1)} = x^{(i+1)H}Ax^{(i+1)} \),
  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 \textbf{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 $\lambda_j$ 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

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

where \(\lambda_j\) 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 \(y_0 = 0\)

\[
\|x - (A - \sigma I)y_k\| \leq \tilde{\kappa}(W) \frac{|\lambda_2 - \lambda_1|}{\lambda_1} \min_{p \in P_{k-1}} \max_{j=2,\ldots,n} |p(\lambda_j)| \|Qx\|
\]

where \(\lambda_j\) are eigenvalues of \(A - \sigma I\).

### Number of inner iterations

\[
k \geq C'_1 + C'_2 \log \frac{\|Qx\|}{\tau},
\]

where \(Q\) 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 $(A - \sigma I)y = x$

<table>
<thead>
<tr>
<th>Good news!</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k^{(i)} \geq C'_1 + C'_2 \log \frac{C_3 |r^{(i)}|}{\tau^{(i)}}$,</td>
</tr>
<tr>
<td>where $\tau^{(i)} = C|r^{(i)}|$. Iteration number approximately constant!</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bad news :-(</th>
</tr>
</thead>
<tbody>
<tr>
<td>For a standard preconditioner $P$</td>
</tr>
<tr>
<td>$(A - \sigma I)P^{-1}\tilde{y}^{(i)} = x^{(i)}$ $\quad$ $P^{-1}\tilde{y}^{(i)} = y^{(i)}$</td>
</tr>
<tr>
<td>$k^{(i)} \geq C''_1 + C''_2 \log \frac{|\tilde{Q}x^{(i)}|}{\tau^{(i)}} = C'_1 + C'_2 \log \frac{C}{\tau^{(i)}}$,</td>
</tr>
<tr>
<td>where $\tau^{(i)} = C|r^{(i)}|$. Iteration number increases!</td>
</tr>
</tbody>
</table>
Finite difference discretisation on a $32 \times 32$ 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 $\times$ 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
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
  \[
  P_i x^{(i)} = Ax^{(i)}, \quad P_i := P + (A - P)x^{(i)}x^{(i)H}
  \]
- minor modification and **minor extra computational cost**,  
- \([A P_i^{-1}]Ax^{(i)} = Ax^{(i)}\).
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,
- \([A\mathbb{P}_i^{-1}]Ax^{(i)} = Ax^{(i)}\).

Why does that work?

Assume we have found eigenvector \(x_1\)

\[
Ax_1 = \mathbb{P}x_1 = \lambda_1 x_1 \quad \Rightarrow \quad (A - \sigma I)\mathbb{P}^{-1}x_1 = \frac{\lambda_1 - \sigma}{\lambda_1} x_1
\]

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

\[
k^{(i)} \geq C_1'' + C_2'' \log \frac{C_3\|r^{(i)}\|}{\tau^{(i)}}, \quad \text{where} \quad \tau^{(i)} = C\|r^{(i)}\|.
\]
Finite difference discretisation on a $32 \times 32$ 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
Figure: Inner iterations vs outer iterations

Figure: Eigenvalue residual norms vs total number of inner iterations
Outline

1 Introduction
2 Inexact inverse iteration
3 Inexact Shift-invert Arnoldi method
4 Inexact Shift-invert Arnoldi method with implicit restarts
5 Conclusions
Arnoldi’s method

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

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

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

$$
Q_k^H Q_k = I.
$$
Arnoldi’s method

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

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

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

$$Q_k^H Q_k = I.$$ 

Eigenvalues of $H_k$ are eigenvalue approximations of (outlying) eigenvalues of $\mathcal{A}$

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

**Arnoldi’s method**

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

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

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

$$
Q_k^H Q_k = I.
$$

- Eigenvalues of $H_k$ are eigenvalue approximations of (outlying) eigenvalues of $\mathcal{A}$

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

- at each step, application of $\mathcal{A}$ to $q_k$: $\mathcal{A}q_k = \tilde{q}_{k+1}$
random complex matrix of dimension \( n = 144 \) generated in MATLAB:
\[ G = \text{numgrid}('N', 14); B = \text{delsq}(G); A = \text{sprandn}(B) + i \times \text{sprandn}(B) \]
after 5 Arnoldi steps
after 10 Arnoldi steps
after 15 Arnoldi steps
after 20 Arnoldi steps
after 25 Arnoldi steps
after 30 Arnoldi steps
The algorithm: take $\sigma = 0$

**Shift-Invert Arnoldi’s method** $\mathbf{A} := \mathbf{A}^{-1}$

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

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

$$
\mathbf{A}^{-1}\mathbf{Q}_k = \mathbf{Q}_k\mathbf{H}_k + q_{k+1}h_{k+1,k}\mathbf{e}_k^H = \mathbf{Q}_{k+1} \begin{bmatrix} H_k \\ h_{k+1,k}\mathbf{e}_k^H \end{bmatrix}
$$

$$
\mathbf{Q}_k^H\mathbf{Q}_k = \mathbf{I}.
$$

- Eigenvalues of $\mathbf{H}_k$ are eigenvalue approximations of (outlying) eigenvalues of $\mathbf{A}^{-1}$

$$
\| \mathbf{r}_k \| = \| \mathbf{A}^{-1} \mathbf{x} - \theta \mathbf{x} \| = \| (\mathbf{A}^{-1}\mathbf{Q}_k - \mathbf{Q}_k\mathbf{H}_k)\mathbf{u} \| = | h_{k+1,k} | \| \mathbf{e}_k^H \mathbf{u} |,
$$

- at each step, application of $\mathbf{A}^{-1}$ to $\mathbf{q}_k$: $\mathbf{A}^{-1}\mathbf{q}_k = \tilde{\mathbf{q}}_{k+1}$
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

Wish to solve

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

leads to inexact Arnoldi relation

\[ A^{-1}Q_k = Q_{k+1} \begin{bmatrix} H_k \\ h_{k+1,k}e_k^H \end{bmatrix} + D_k = Q_{k+1} \begin{bmatrix} H_k \\ h_{k+1,k}e_k^H \end{bmatrix} + [d_1 | \ldots | d_k] \]
Inexact solves (Simoncini 2005), Bouras and Frayssé (2000)

- Wish to solve
  \[ \| q_k - A \tilde{q}_{k+1} \| = \| \tilde{d}_k \| \leq \tau_k \]

- leads to inexact Arnoldi relation
  \[
  A^{-1} Q_k = Q_{k+1} \begin{bmatrix}
  H_k \\
  h_{k+1,1} k e_k^H
  \end{bmatrix}
  + D_k = Q_{k+1} \begin{bmatrix}
  H_k \\
  h_{k+1,1} k e_k^H
  \end{bmatrix}
  + [d_1| \ldots | d_k]
  \]

- \( u \) eigenvector of \( H_k \):
  \[
  \| r_k \| = \| (A^{-1} Q_k - Q_k H_k) u \| = | h_{k+1,1} k \| e_k^H u \| + D_k u,
  \]
Inexact solves (Simoncini 2005), Bouras and Frayssé (2000)

- Wish to solve
  \[ \|q_k - A\tilde{q}_{k+1}\| = \|\tilde{d}_k\| \leq \tau_k \]
- leads to inexact Arnoldi relation
  \[
  A^{-1}Q_k = Q_{k+1} \begin{bmatrix}
  H_k \\
  h_{k+1,k}e_k^H
  \end{bmatrix}
  + D_k = Q_{k+1} \begin{bmatrix}
  H_k \\
  h_{k+1,k}e_k^H
  \end{bmatrix}
  + [d_1| \ldots |d_k]
  \]
- \(u\) eigenvector of \(H_k\):
  \[ \|r_k\| = \|(A^{-1}Q_k - Q_kH_k)u\| = |h_{k+1,k}|e_k^Hu + D_ku, \]
- Linear combination of the columns of \(D_k\)
  \[
  D_ku = \sum_{l=1}^{k} d_lu_l, \quad \text{if \ } u_l \text{ small, then } \|d_l\| \text{ allowed to be large!} \]
Inexact solves (Simoncini 2005), Bouras and Frayssé (2000)

Linear combination of the columns of $D_k$

\[
D_k u = \sum_{l=1}^{k} d_l u_l, \quad \text{if } u_l \text{ small, then } ||d_l|| \text{ allowed to be large!}
\]

\[
||d_l u_l|| \leq \frac{1}{k} \varepsilon \Rightarrow ||D_k u|| < \varepsilon
\]

and

\[
|u_l| \leq C(l, k) ||r_{l-1}||
\]

leads to

\[
||q_k - A\tilde{q}_{k+1}|| = ||\tilde{d}_k|| = C \frac{1}{||r_{k-1}||}
\]

Solve tolerance can be relaxed.
The inner iteration for $AP^{-1}\tilde{q}_{k+1} = q_k$

Preconditioning

GMRES convergence bound

$$\|q_k - AP^{-1}\tilde{q}_{k+1}\| = \kappa \min_{p \in \Pi_i} \max_{i=1,...,n} |p(\mu_i)||q_k|$$

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

### Preconditioning

GMRES convergence bound

$$
\|q_k - AP^{-1}\tilde{q}_{k+1}\| = \kappa \min_{p \in \Pi_i} \max_{i=1, \ldots, n} |p(\mu_i)| \|q_k\|
$$

depending on

- the eigenvalue clustering of $AP^{-1}$
- the condition number
- the right hand side (initial guess)
The inner iteration for $AP^{-1}\tilde{q}_{k+1} = q_k$

Preconditioning

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

### Preconditioning

- 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

### Tuned Preconditioner

using a tuned preconditioner for Arnoldi’s method

\[
P_k Q_k = AQ_k; \quad \text{given by} \quad P_k = P + (A - P)Q_k Q_k^H
\]
The inner iteration for $A\tilde{q} = q$

Theorem (Properties of the tuned preconditioner)

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 $AP^{-1}^k$ are equal to one:

$$[AP^{-1}_k]AQ_k = AQ_k$$

and $n - k$ eigenvalues are close to the corresponding eigenvalues of $AP^{-1}$. 
The inner iteration for $A\tilde{q} = q$

<table>
<thead>
<tr>
<th>Theorem (Properties of the tuned preconditioner)</th>
</tr>
</thead>
<tbody>
<tr>
<td>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 $A P_k^{-1}$ are equal to one:</td>
</tr>
<tr>
<td>[ [A P_k^{-1}] A Q_k = A Q_k ]</td>
</tr>
<tr>
<td>and $n - k$ eigenvalues are close to the corresponding eigenvalues of $A P^{-1}$.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>■ Sherman-Morrison-Woodbury.</td>
</tr>
<tr>
<td>■ Only minor extra costs (one back substitution per outer iteration)</td>
</tr>
</tbody>
</table>
Numerical Example

**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).
No tuning and standard preconditioner

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


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
### Table: Ritz values of exact Arnoldi’s method and inexact Arnoldi’s method with the tuning strategy compared to exact eigenvalues closest to zero after 14 shift-invert Arnoldi steps.

<table>
<thead>
<tr>
<th>Exact eigenvalues</th>
<th>Ritz values (exact Arnoldi)</th>
<th>Ritz values (inexact Arnoldi, tuning)</th>
</tr>
</thead>
<tbody>
<tr>
<td>+4.69249563e-02</td>
<td>+4.69249563e-02</td>
<td>+4.69249563e-02</td>
</tr>
<tr>
<td>+1.25445378e-01</td>
<td>+1.25445378e-01</td>
<td>+1.25445378e-01</td>
</tr>
<tr>
<td>+4.02658363e-01</td>
<td>+4.02658347e-01</td>
<td>+4.02658244e-01</td>
</tr>
<tr>
<td>+5.79574381e-01</td>
<td>+5.79625498e-01</td>
<td>+5.79817301e-01</td>
</tr>
<tr>
<td>+6.18836405e-01</td>
<td>+6.18798666e-01</td>
<td>+6.18650849e-01</td>
</tr>
</tbody>
</table>
Outline

1. Introduction
2. Inexact inverse iteration
3. Inexact Shift-invert Arnoldi method
4. Inexact Shift-invert Arnoldi method with implicit restarts
5. Conclusions
Implicitly restarted Arnoldi (Sorensen (1992))

Exact shifts

- take an $k + p$ step Arnoldi factorisation

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

- Compute $\Lambda(H_{k+p})$ and select $p$ shifts for an implicit QR iteration

- implicit restart with new starting vector $\hat{q}^{(1)} = \frac{p(A)q^{(1)}}{\|p(A)q^{(1)}\|}$
Implicitly restarted Arnoldi (Sorensen (1992))

Exact shifts

- take an $k + p$ step Arnoldi factorisation

\[ AQ_{k+p} = Q_{k+p}H_{k+p} + q_{k+p+1}h_{k+p+1,k+p}e_{k+p}^H \]

- Compute $\Lambda(H_{k+p})$ and select $p$ shifts for an implicit QR iteration

- implicit restart with new starting vector $\hat{q}^{(1)} = \frac{p(A)q^{(1)}}{\|p(A)q^{(1)}\|}$

Aim of IRA

\[ AQ_k = Q_kH_k + q_{k+1} h_{k+1,k} e_k^H \rightarrow 0 \]
**Theorem**

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

$$
\|d_l\| \leq \begin{cases} 
\varepsilon \frac{C}{\|R_k\|} & \text{if } l > k, \\
\varepsilon & \text{otherwise.}
\end{cases}
$$

Then

$$
\|AQ_mU - Q_mU\Theta - R_m\| \leq \varepsilon.
$$

- Very technical
- Relaxation strategy also works for IRA!
Tuning

Tuning for implicitly restarted Arnoldi’s method

- Introduce preconditioner $P$ and solve

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

using GMRES and a tuned preconditioner

\[ \mathbb{P}_k Q_k = A Q_k; \quad \text{given by} \quad \mathbb{P}_k = P + (A - P)Q_k Q_k^H \]
### Why does tuning help?

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

\[
A^{-1} Q_k = Q_k H_k + q_{k+1} h_{k+1,k} e_k^H 
\approx 0
\]
Tuning

Why does tuning help?

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

\[
A^{-1} Q_k = Q_k H_k + q_{k+1} h_{k+1,k} e_k^H \approx 0
\]

- let \( A^{-1} \) have the 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?

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

\[
A^{-1} Q_k = Q_k H_k + q_{k+1} h_{k+1,k} e_k^H \approx 0
\]

- let \( A^{-1} \) have the 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

\[
\begin{bmatrix}
    Q_k & Q_k^\perp
\end{bmatrix}^H A^P_k^{-1} \begin{bmatrix}
    Q_k & Q_k^\perp
\end{bmatrix} = \begin{bmatrix}
    I + \star & Q_k^H A^P_k^{-1} Q_k^\perp \\
    \star & T_{22}^{-1} (Q_k^\perp H P Q_k^{-1})^{-1} + \star
\end{bmatrix}
\]
Why does tuning help?

- Assume we have found an approximate invariant subspace, that is
  \[ A^{-1}Q_k = Q_kH_k + q_{k+1}h_{k+1,k}e_k^H \approx 0 \]

- Let \( A^{-1} \) have the upper Hessenberg form

  \[
  \begin{bmatrix}
  Q_k & Q_k^\perp \\
  \end{bmatrix}^H A^{-1} \begin{bmatrix}
  Q_k & Q_k^\perp \\
  \end{bmatrix} = \begin{bmatrix}
  H_k & T_{12} \\
  h_{k+1,k}e_1e_k^H & T_{22} \\
  \end{bmatrix},
  \]

  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 \( h_{k+1,k} = 0 \) then

\[
\begin{bmatrix}
  Q_k & Q_k^\perp \\
  \end{bmatrix}^H A_{\text{IP}}^{-1} \begin{bmatrix}
  Q_k & Q_k^\perp \\
  \end{bmatrix} = \begin{bmatrix}
  I & Q_k^H A_{\text{IP}}^{-1}Q_k^\perp \\
  0 & T_{22}^{-1}(Q_k^H P Q_k^\perp)^{-1} \\
  \end{bmatrix}
\]
Another advantage of tuning

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

\[ A \mathbb{P}_k^{-1} \tilde{q}_{k+1} = q_k, \quad \mathbb{P}_k^{-1} \tilde{q}_{k+1} = \tilde{q}_k \]
Another advantage of tuning

- System to be solved at each step of Arnoldi’s method is
  \[ A\mathbb{P}_k^{-1}\tilde{q}_{k+1} = q_k, \quad \mathbb{P}_k^{-1}\tilde{q}_{k+1} = \tilde{q}_k \]

- Assuming invariant subspace found then (\(A^{-1}Q_k = Q_kH_k\)):
  \[ A\mathbb{P}_k^{-1}q_k = q_k \]
Another advantage of tuning

- System to be solved at each step of Arnoldi’s method is
  \[ A \mathbb{P}_k^{-1} \tilde{q}_{k+1} = q_k, \quad \mathbb{P}_k^{-1} \tilde{q}_{k+1} = \tilde{q}_k \]

- Assuming invariant subspace found then \( (A^{-1}Q_k = Q_kH_k) \):
  \[ A \mathbb{P}_k^{-1} q_k = q_k \]

- the right hand side of the system matrix is an eigenvector of the system!
Another advantage of tuning

- System to be solved at each step of Arnoldi’s method is
  \[ AP_k^{-1} \tilde{q}_{k+1} = q_k, \quad P_k^{-1} \tilde{q}_{k+1} = \tilde{q}_k \]

- Assuming invariant subspace found then \((A^{-1}Q_k = Q_kH_k)\):
  \[ A P_k^{-1} q_k = q_k \]

- The right hand side of the system matrix is an eigenvector of the system!

- Krylov methods converge in one iteration
Another advantage of tuning

- In practice:

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

and

\[ \| A P_k^{-1} q_k - q_k \| = O(|h_{k+1,k}|) \]
Another advantage of tuning

- In practice:

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

and

\[ \| A^{-1} P_k q_k - q_k \| = \mathcal{O}(|h_{k+1,k}|) \]

- number of iterations decreases as the outer iteration proceeds
Another advantage of tuning

- In practice:
  \[ A^{-1}Q_k = Q_k H_k + q_{k+1} h_{k+1,k} e_k^H \]
  and
  \[ \| A_{P_k}^{-1} q_k - q_k \| = O(|h_{k+1,k}|) \]
- number of iterations decreases as the outer iteration proceeds
- Rigorous analysis quite technical.
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).
No tuning and standard preconditioner

Figure: Inner iterations vs outer iterations

Figure: Eigenvalue residual norms vs total number of inner iterations
Figure: Inner iterations vs outer iterations

Figure: Eigenvalue residual norms vs total number of inner iterations
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](image1)

![Figure: Eigenvalue residual norms vs total number of inner iterations](image2)
Numerical Example

\texttt{qc2534.mtx} matrix from the Matrix Market library.

- $k = 6$ 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).
Tuning and relaxation strategy

**Figure:** Inner iterations vs outer iterations

**Figure:** Eigenvalue residual norms vs total number of inner iterations
Outline

1 Introduction

2 Inexact inverse iteration

3 Inexact Shift-invert Arnoldi method

4 Inexact Shift-invert Arnoldi method with implicit restarts

5 Conclusions
Conclusions

- For eigenvalue computations it is advantageous to consider small rank changes to the standard preconditioners.
- Works for any preconditioner.
- Inexact inverse iteration with a special tuned preconditioner is equivalent to the Jacobi-Davidson method (without subspace expansion).
- For Arnoldi method best results are obtained when relaxation and tuning are combined.
M. A. Freitag and A. Spence, *Rayleigh quotient iteration and simplified Jacobi-Davidson method with preconditioned iterative solves.*

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


