A Tuned Preconditioner for Inexact Inverse Iteration Applied to Hermitian Eigenvalue Problems

Melina Freitag

Department of Mathematical Sciences
University of Bath, United Kingdom

IWASEP VI
May 22-25, 2006
Pennsylvania State University, University Park

Joint work with: Alastair Spence
1 Motivation

2 Inexact Inverse Iteration
   - Convergence rates - independent of inner solver
   - MINRES - inner solves

3 Hermitian problems and preconditioning
   - Preconditioning
   - Tuning the preconditioner
   - Numerical Results
   - Perturbation theory
   - Another approach

4 Hermitian generalised eigenproblems
Outline

1 Motivation

2 Inexact Inverse Iteration
   - Convergence rates - independent of inner solver
   - MINRES - inner solves

3 Hermitian problems and preconditioning
   - Preconditioning
   - Tuning the preconditioner
   - Numerical Results
   - Perturbation theory
   - Another approach

4 Hermitian generalised eigenproblems
Problem and Inverse Iteration

- Find an eigenvalue and eigenvector of a Hermitian positive definite $A$:

$$Ax = \lambda x,$$

- Inverse Iteration:

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

$A$ large, sparse.

- Inverse iteration with preconditioned iterative solves
Outline

1 Motivation

2 Inexact Inverse Iteration
   - Convergence rates - independent of inner solver
   - MINRES - inner solves

3 Hermitian problems and preconditioning
   - Preconditioning
   - Tuning the preconditioner
   - Numerical Results
   - Perturbation theory
   - Another approach

4 Hermitian generalised eigenproblems
Inexact Inverse Iteration

\begin{align*}
\textbf{for } i = 1 \text{ to } i_{\text{max}} \textbf{ do} \\
\quad \text{choose } \tau(i), \sigma(i) \\
\quad \text{solve} \\
\quad \| (A - \sigma(i) I) y(i) - x(i) \| \leq \tau(i), \\
\quad \text{Rescale } x(i+1) = \frac{y(i)}{\| y(i) \|}, \\
\quad \text{Update } \lambda(i+1) = x(i+1)^T A x(i+1), \\
\quad \text{possibly: update the shift } \sigma(i) \\
\quad \text{Test: eigenvalue residual } r^{(i+1)} = (A - \lambda(i+1) I) x(i+1). \\
\textbf{end for}
\end{align*}
Error indicator

Error indicator (Orthogonal decomposition for symmetric $A$, Parlett)

$$Q \mathbf{x}^{(i)} = O \left( \sin \theta^{(i)} \right)$$

measure for the error

$$\mathbf{x}^{(i)} = \cos \theta^{(i)} \mathbf{x}_1 + \sin \theta^{(i)} \mathbf{x}_\perp, \quad \mathbf{x}_\perp \perp \mathbf{x}_1.$$ 

Eigenvalue residual

$$| \sin \theta^{(i)} | \| \lambda_2 - \lambda^{(i)} \| \leq \| r^{(i)} \| \leq | \sin \theta^{(i)} | \| \lambda_n - \lambda_1 \|$$
Convergence rates of inexact inverse iteration

Decreasing tolerance $\tau^{(i)} = C\|r^{(i)}\| = O(\sin \theta^{(i)})$

1. For decreasing tolerance $\tau^{(i)} \leq C\|r^{(i)}\| = O(\sin \theta^{(i)})$ the inexact method recovers the rate of convergence achieved by exact solves.

2. **Fixed shift $\sigma$: linear convergence.**

3. Rayleigh quotient shift $\sigma^{(i)} = \rho(x^{(i)}) = \frac{x^{(i)^T}Ax^{(i)}}{x^{(i)^T}x^{(i)}}$: cubic convergence for $A = A^*$. 

Fixed tolerance $\tau^{(i)} = \tau$

1. **Rayleigh quotient shift**: quadratic convergence
MINRES \((A - \sigma I)y = x\) when \(A\) is symmetric

Solving a linear system \((A - \sigma I)y = x\)

- standard MINRES theory for \(y_0 = 0\):

\[
\|x - (A - \sigma I)y_k\| \leq 2 \left( \sqrt{\frac{\kappa - 1}{\kappa + 1}} \right)^{k-1} \|x\|.
\]

where \(\kappa\) is the condition number of \(A - \sigma I\).

- Number of inner iterations:

\[
k \geq 1 + \kappa \left\{ \log 2 + \log \frac{\|x\|}{\tau} \right\}
\]

then

\[
\|x - (A - \sigma I)y_k\| \leq \tau.
\]
Unpreconditioned solves with MINRES

Convergence rates for solves with MINRES for simple eigenvalue

If $A$ is positive definite and has a simple eigenvalue then

$$
\|x^{(i)} - (A - \sigma I)y_k^{(i)}\|_2 \leq 2 \frac{|\lambda_1 - \lambda_n|}{|\lambda_1 - \sigma|} \left( \sqrt{\frac{\kappa_1 - 1}{\kappa_1 + 1}} \right)^{k-2} \|Qx^{(i)}\|_2.
$$

where $Q$ is the orthogonal projection onto span$\{x_2, \ldots, x_n\}$ and $\kappa_1$ is the reduced condition number $\kappa_1 = \max_{i=2, \ldots, n} |\lambda_i - \sigma| / \min_{i=2, \ldots, n} |\lambda_i - \sigma|$.

Number of inner solves for each $i$ for $\|x^{(i)} - (A - \sigma^{(i)} I)y^{(i)}\| \leq \tau^{(i)}$,

$$
k^{(i)} \geq 2 + \kappa_1 \left( \log 2|\lambda_1 - \lambda_n| + \log \frac{\|Qx^{(i)}\|_2}{|\lambda_1 - \sigma|\tau^{(i)}} \right).
$$
Unpreconditioned solves with MINRES

Convergence rates for solves with MINRES for simple eigenvalue

If $A$ is positive definite and has a simple eigenvalue then

$$\|x^{(i)} - (A - \sigma I)y_k^{(i)}\|_2 \leq 2 \frac{|\lambda_1 - \lambda_n|}{|\lambda_1 - \sigma|} \left(\sqrt{\frac{\kappa_1 - 1}{\kappa_1 + 1}}\right)^{k-2} |\sin \theta^{(i)}|.$$ 

where $Q$ is the orthogonal projection onto span$\{x_2, \ldots, x_n\}$ and $\kappa_1$ is the reduced condition number $\kappa_1 = \frac{\max_{i=2,\ldots,n} |\lambda_i - \sigma|}{\min_{i=2,\ldots,n} |\lambda_i - \sigma|}$.

Number of inner solves for each $i$ for $\|x^{(i)} - (A - \sigma_1 I)y^{(i)}\| \leq \tau^{(i)}$

$$k^{(i)} \geq 2 + \kappa_1 \left(\log 2|\lambda_1 - \lambda_n| + \log \frac{|\sin \theta^{(i)}|}{|\lambda_1 - \sigma|\tau^{(i)}}\right)$$
Outline

1. Motivation

2. Inexact Inverse Iteration
   - Convergence rates - independent of inner solver
   - MINRES - inner solves

3. Hermitian problems and preconditioning
   - Preconditioning
   - Tuning the preconditioner
   - Numerical Results
   - Perturbation theory
   - Another approach

4. Hermitian generalised eigenproblems
Preconditioning

Incomplete Cholesky preconditioning

\[ A = LL^T + E \]

symmetric preconditioning of \((A - \sigma I)y^{(i)} = x^{(i)}:\)

\[ L^{-1}(A - \sigma I)L^{-T}\tilde{y}^{(i)} = L^{-1}x^{(i)}, \quad y^{(i)} = L^{-T}\tilde{y}^{(i)} \]

Remarks

1. \(k^{(i)}\) changes number of inner iterations

\[ k^{(i)} \geq 2 + \kappa_1 \left( \log 2|\lambda_1 - \lambda_n| + \log \frac{\|L^{-1}\|}{|\lambda_1 - \sigma|\tau^{(i)}} \right) \]

2. \(k^{(i)}\) increases with \(i\) for \(\tau^{(i)} = C\|r^{(i)}\|\).
Inexact inverse iteration and tuned preconditioning

Melina Freitag

Outline
Motivation
Inexact Inverse Iteration
Convergence rates - independent of inner solver
MINRES - inner solves
Hermitian problems and preconditioning

Preconditioning
Tuning the preconditioner
Numerical Results
Perturbation theory
Another approach
Hermitian generalised eigenproblems

Preconditioning

Incomplete Cholesky preconditioning

\[ A = LL^T + E \]

symmetric preconditioning of \((A - \sigma I)y^{(i)} = x^{(i)}:\)

\[ L^{-1}(A - \sigma I)L^{-T}\tilde{y}^{(i)} = L^{-1}x^{(i)}, \quad y^{(i)} = L^{-T}\tilde{y}^{(i)} \]

Remarks

1. changes number of inner iterations

\[ k^{(i)} \geq 2 + \kappa_1 \left( \log 2|\lambda_1 - \lambda_n| + \log \frac{||L^{-1}||}{|\lambda_1 - \sigma|\tau^{(i)}} \right) \]

2. \( k^{(i)} \) increases with \( i \) for \( \tau^{(i)} = C||r^{(i)}||. \)
Aims

1. modify $L \rightarrow \mathbb{L}$

\[ \mathbb{L}^{-1}(A - \sigma I)\mathbb{L}^{-T}\tilde{y}^{(i)} = \mathbb{L}^{-1}x^{(i)}, \quad y^{(i)} = \mathbb{L}^{-T}\tilde{y}^{(i)} \]

2. minor extra computation cost for $\mathbb{L}$

3. "nice" right hand side $\mathbb{L}^{-1}x^{(i)}$ (same behaviour as unpreconditioned solves, e.g. for fixed shifts $k^{(i)}$ does not increase with $i$)
Aims

1. modify $L \rightarrow \mathbb{L}$

$$\mathbb{L}^{-1}(A - \sigma I)\mathbb{L}^{-T}\tilde{y}^{(i)} = \mathbb{L}^{-1}x^{(i)}, \quad y^{(i)} = \mathbb{L}^{-T}\tilde{y}^{(i)}$$

2. minor extra computation cost for $\mathbb{L}$

3. "nice" right hand side $\mathbb{L}^{-1}x^{(i)}$ (same behaviour as unpreconditioned solves, e.g. for fixed shifts $k^{(i)}$ does not increase with $i$)
Choice of $LL$

**Condition**

- MINRES theory indicates that $LL^{-1}x^{(i)}$ should be close to eigenvector of $LL^{-1}(A - \sigma I)L^{-T}$
- Holds if

$$LLL^T x^{(i)} = Ax^{(i)}$$

**Justification of $LLL^T x^{(i)} = Ax^{(i)}$**

If $x^{(i)} = x_1$ then $LLL^T x_1 = \lambda_1 x_1$

$$LL^{-1}(A - \sigma I)L^{-T}L^{-1}x_1 = \frac{\lambda_1 - \sigma}{\lambda_1}LL^{-1}x_1$$

$$LL^{-1}(A - \sigma I)L^{-T}L^{-1}x^{(i)} = \frac{\lambda_1 - \sigma}{\lambda_1}LL^{-1}x^{(i)} + C\|r^{(i)}\|$$
Choice of $LL$

Condition

- MINRES theory indicates that $L^{-1}x(i)$ should be close to eigenvector of $L^{-1}(A - \sigma I)L^{-T}$
- Holds if
  \[ LLL^T x(i) = Ax(i) \]

Justification of $LL^T x(i) = Ax(i)$

If $x(i) = x_1$ then $LL^T x_1 = \lambda_1 x_1$

\[ L^{-1}(A - \sigma I)L^{-T}L^{-1}x_1 = \frac{\lambda_1 - \sigma}{\lambda_1} L^{-1}x_1 \]

\[ L^{-1}(A - \sigma I)L^{-T}L^{-1}x(i) = \frac{\lambda_1 - \sigma}{\lambda_1} L^{-1}x(i) + C\|r(i)\| \]
How do we achieve $LL^T x^{(i)} = Ax^{(i)}$?

**Theorem**

Let $x^{(i)}$ current eigenvector approximation, $e^{(i)} = Ax^{(i)} - LL^T x^{(i)}$ (known) and $L$ chosen such that

$$L = L + \alpha^{(i)} e^{(i)} (L^{-1} e^{(i)})^T$$

with $\alpha^{(i)}$ root of quadratic function we get $LL^T x^{(i)} = Ax^{(i)}$.

**Implementation**

1. Note: $LL^T = LL^T + \frac{1}{e^{(i)^T x^{(i)}}} e^{(i)} e^{(i)^T}$
2. $L$ is a rank-one update of $L$. 

Outline
- Motivation
- Inexact Inverse Iteration
  - Convergence rates - independent of inner solver
  - MINRES - inner solves
- Hermitian problems and preconditioning
- Preconditioning
  - Tuning the preconditioner
  - Numerical Results
  - Perturbation theory
  - Another approach
- Hermitian generalised eigenproblems
How do we achieve $\mathbb{L}\mathbb{L}^T x^{(i)} = Ax^{(i)}$?

**Theorem**

Let $x^{(i)}$ current eigenvector approximation, $e^{(i)} = Ax^{(i)} - LL^T x^{(i)}$ (known) and $\mathbb{L}$ chosen such that

$$\mathbb{L} = L + \alpha^{(i)} e^{(i)} (L^{-1} e^{(i)})^T$$

with $\alpha^{(i)}$ root of quadratic function we get $\mathbb{L}\mathbb{L}^T x^{(i)} = Ax^{(i)}$.

**Implementation**

1. **Note:** $\mathbb{L}\mathbb{L}^T = LL^T + \frac{1}{e^{(i)^T} x^{(i)}} e^{(i)} e^{(i)^T}$
2. $\mathbb{L}$ is a rank-one update of $L$. 
Implementation

General positive definite preconditioner

For MINRES implementation only the evaluation of $P^{-1}$ is necessary

$$P = P + \gamma^{(i)} e^{(i)} e^{(i)\top}$$

Sherman-Morrison formula

$$P^{-1} = P^{-1} - \frac{(z^{(i)} - x^{(i)})(z^{(i)} - x^{(i)})^\top}{(z^{(i)} - x^{(i)})^\top A x^{(i)}}$$

where $z^{(i)} = P^{-1} A x^{(i)}$. 
Convergence rates

The tuned preconditioner

1. outer convergence rate is retained
2. cheap inner solves are provided
   \[ k^{(i)} \geq C_1 + C_2 \log \left( \frac{|\sin \theta^{(i)}|}{|\lambda_1 - \sigma| \tau^{(i)}} \right) \]
3. only a single extra back substitution with \( P = LL^T \) per outer iteration needed
Example

- SPD matrix from the Matrix Market library (nos5: 3 story building with attached tower)
- seek eigenvalue near fixed shift $\sigma = 100$
- $A \approx LL^T$, incomplete Cholesky factorisation (drop tol. $= 0.1$)
- compare standard and tuned preconditioner
Preconditioning with standard incomplete Cholesky

- total number of inner iterations using standard preconditioner: 2026
- total number of inner iterations using tuned preconditioner: 779
Comparison of $LL^T$ with $LLLT$

**Spectral properties of preconditioned matrix**

Let

$$L^{-1}(A - \sigma I)L^{-T}w = \mu w$$

$$LL^{-1}(A - \sigma I)L^{-T}\hat{w} = \xi \hat{w}$$

**Theorem**

*If* $\sigma \notin \Lambda(A)$ *then* $\mu, \xi \neq 0$ *and*

$$\min_{\mu \in \Lambda(L^{-1}(A-\sigma I)L^{-T})} \left| \frac{\mu - \xi}{\xi} \right| \leq |\gamma v^* v|,$$

*where* $\gamma = 1/(e^T x)$ *and* $v = L^{-1}e.$
Comparison of $LL^T$ with $LLT$

Spectral properties of preconditioned matrix

Let

$$L^{-1}(A - \sigma I)L^{-T}w = \mu w$$

$$LL^{-1}(A - \sigma I)L^{-T}\hat{w} = \xi \hat{w}$$

Interlacing property

Rewrite second equation

$$Dt = \xi (I + \gamma zz^T)t$$

where $L^{-1}(A - \sigma I)L^{-T} = QDQ^T$, $z = Q^Tv$, $(I + \alpha vv^T)Qt = \hat{w}$.

Interlacing property

- If $\gamma > 0$ eigenvalues are moved towards the origin.
- If $\gamma < 0$ eigenvalues are moved away from the origin.
Comparison of $LL^T$ with $LL^T$

**Spectral properties of preconditioned matrix**

Let

\[ L^{-1}(A - \sigma I)L^{-T}w = \mu w \]
\[ \mathbb{L}^{-1}(A - \sigma I)\mathbb{L}^{-T}\hat{w} = \xi \hat{w} \]

**Interlacing property**

Rewrite second equation

\[ Dt = \xi (I + \gamma zz^T)t \]

where $L^{-1}(A - \sigma I)L^{-T} = QDQ^T$, $z = Q^Tv$, $(I + \alpha vv^T)Qt = \hat{w}$.

**Interlacing property**

- If $\gamma > 0$ eigenvalues are moved towards the origin.
- If $\gamma < 0$ eigenvalues are moved away from the origin.
Comparison of $LL^T$ with $LL^T$

Interlacing property

- $\mu$ and $\xi$ interlace each other depending on the sign of $\gamma$
- Clustering properties are preserved
- reduced condition number $\kappa_L^1 \leq \kappa_L^1 \leq \kappa_L^1 (1 + \gamma v^T v)$
Comparison of $LL^T$ with $LL^T$

- $\mu$ and $\xi$ interlace each other depending on the sign of $\gamma$
- Clustering properties are preserved
- reduced condition number $\kappa_L^1 \leq \kappa_L^1 \leq \kappa_L^1(1 + \gamma v^T v)$
Changing the right hand side

**Approach by Simoncini/Eldén [3]**

Instead of solving

\[ L^{-1}(A - \sigma^{(i)}I)L^{-T}\tilde{y}^{(i)} = L^{-1}x^{(i)}, \quad y^{(i)} = L^{-T}\tilde{y}^{(i)} \]

change the right hand side

\[ L^{-1}(A - \sigma^{(i)}I)L^{-T}\tilde{y}^{(i)} = L^{T}x^{(i)}, \quad y^{(i)} = L^{-T}\tilde{y}^{(i)} \]
### Tuned preconditioner and Simoncini & Eldén approach

Example `nos5.mtx` from Matrix Market. Solves to fixed tolerance $\tau = 0.01$. Rayleigh quotient shift. Quadratic convergence for both methods.

<table>
<thead>
<tr>
<th>Outer Iteration</th>
<th>Simoncini &amp; Eldén Drop Tolerances</th>
<th>Tuned preconditioner Drop Tolerances</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.25</td>
<td>0.1</td>
</tr>
<tr>
<td>1</td>
<td>67</td>
<td>62</td>
</tr>
<tr>
<td>2</td>
<td>74</td>
<td>66</td>
</tr>
<tr>
<td>3</td>
<td>85</td>
<td>75</td>
</tr>
<tr>
<td>4</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>289</td>
<td>203</td>
</tr>
</tbody>
</table>
Outline

1 Motivation

2 Inexact Inverse Iteration
   - Convergence rates - independent of inner solver
   - MINRES - inner solves

3 Hermitian problems and preconditioning
   - Preconditioning
   - Tuning the preconditioner
   - Numerical Results
   - Perturbation theory
   - Another approach

4 Hermitian generalised eigenproblems
Inexact inverse iteration and tuned preconditioning

Melina Freitag

Outline
Motivation
Inexact Inverse Iteration
Convergence rates - independent of inner solver
MINRES - inner solves
Preconditioning
Tuning the preconditioner
Numerical Results
Perturbation theory
Another approach
Hermitian generalised eigenproblems

Ax = λMx with bcsstk08 (Structural engineering)

Figure: Fixed Shift
Figure: Rayleigh Quotient Shift
Inexact inverse iteration and tuned preconditioning
Melina Freitag

Outline
Motivation
Inexact Inverse Iteration
Convergence rates - independent of inner solver
MINRES - inner solves
Hermitian problems and preconditioning
Preconditioning Tuning the preconditioner
Numerical Results
Perturbation theory
Another approach
Hermitian generalised eigenproblems

Numerical example for the generalised eigenproblem

\[ Ax = \lambda M x \text{ with bcsstk08 (Structural engineering)} \]

**Figure: Fixed Shift**

**Figure: Rayleigh Quotient Shift**
Submitted to BIT.

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