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

Melina Freitag and Alastair Spence

Department of Mathematical Sciences
University of Bath

21st Biennial Conference on Numerical Analysis
University of Dundee, Scotland, UK
30th June 2005
Outline

1. Outline
2. Inverse Iteration (large sparse matrices)
3. Inexact Inverse Iteration
4. Preconditioned Inexact Inverse Iteration
5. Tuning the preconditioner
6. Numerical Results
Problem and Inverse Iteration

- Find an eigenvalue and eigenvector of s.p.d. $A$:
  
  $$Ax = \lambda x,$$

- Inverse Iteration:
  
  $$(A - \sigma I)y = x$$

$A$ large, sparse.
for $i = 1$ to $\ldots$ do
  choose $\tau^{(i)}$
  solve
  \[ \|(A - \sigma I)y^{(i)} - x^{(i)}\| \leq \tau^{(i)}, \]
  Rescale $x^{(i+1)} = \frac{y^{(i)}}{\|y^{(i)}\|}$,
  Update $\lambda^{(i+1)} = x^{(i+1)T}A x^{(i+1)}$,
  Eigenvalue residual $r^{(i+1)} = (A - \lambda^{(i+1)} I)x^{(i+1)}$.
end for
Error indicator (orthogonal decomposition, Parlett)

\[ P_{\perp} x^{(i)} = O(\sin \theta^{(i)}) \] measure for the error

Eigenvalue residual

\[ C | \sin \theta^{(i)} | \leq \| r^{(i)} \| \leq C' | \sin \theta^{(i)} | \]
Error indicator (orthogonal decomposition, Parlett)

\[ P_\perp x^{(i)} = O(\sin \theta^{(i)}) \]

measure for the error

\[ C | \sin \theta^{(i)} | \leq \| r^{(i)} \| \leq C' | \sin \theta^{(i)} | \]
Convergence rates of inexact inverse iteration

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

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.** [see Golub/Ye 2000, Berns-Müller/Graham/Spence 2005]

3. **Rayleigh quotient shift** $\sigma^{(i)} = \rho(x^{(i)}) = \frac{x^{(i)^T} Ax^{(i)}}{x^{(i)^T} x^{(i)}}$: cubic convergence. [see Smit/Paardekooper 1999, Berns-Müller/Graham/Spence 2005]
If $A$ is positive definite and has a simple eigenvalue then

$$
\|x^{(i)} - (A - \sigma I)y_k^{(i)}\|_2 \leq 2\frac{|\lambda_1 - \lambda_n|}{|\lambda_1 - \sigma|} \left( \sqrt{\frac{\kappa_1 - 1}{\kappa_1 + 1}} \right)^{k-1} \|P_1^\perp x^{(i)}\|_2.
$$

Number of inner solves for each $i$

$$
k^{(i)} \geq C_1 + C_2 \log \left( \frac{\|P_1^\perp x^{(i)}\|_2}{|\lambda_1 - \sigma| \tau^{(i)}} \right)
$$
Unpreconditioned solves with MINRES

Convergence rates for solves with MINRES for simple eigenvalue

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

$$\|x^{(i)} - (A - \sigma I)y_k^{(i)}\|_2 \leq 2\frac{|\lambda_1 - \lambda_n|}{|\lambda_1 - \sigma|} \left(\sqrt{\frac{\kappa_1 - 1}{\kappa_1 + 1}}\right)^{k-1} \|P_1^\bot x^{(i)}\|_2.$$

Number of inner solves for each $i$

$$k^{(i)} \geq C_1 + C_2 \log \left(\frac{\|P_1^\bot x^{(i)}\|_2}{|\lambda_1 - \sigma|\tau^{(i)}}\right)$$
Unpreconditioned solves with MINRES

Convergence rates for solves with MINRES for simple eigenvalue

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

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

Number of inner solves for each $i$

$$k^{(i)} \geq C_1 + C_2 \log \left( \frac{|\sin \theta^{(i)}|}{|\lambda_1 - \sigma| \tau^{(i)}} \right).$$
Unpreconditioned solves with MINRES

Number of inner solves for each $i$

$$k^{(i)} \geq C_1 + C_2 \log \left( \frac{|\sin \theta^{(i)}|}{|\lambda_1 - \sigma| \tau^{(i)}} \right)$$

Example

For fixed shift $\sigma$, and $\tau^{(i)} \leq C \| r^{(i)} \| = O(\sin \theta^{(i)})$ the number of inner solves $k^{(i)}$ for each $i$ does not increase with $i$
## 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

\[
k^{(i)} \geq C_1 + C_2 \log \left( \frac{\|L^{-1}\|}{|\lambda_1 - \sigma| r^{(i)}} \right)
\]

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

---

Melina Freitag and Alastair Spence  
A Tuned Preconditioner
**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

\[ k^{(i)} \geq C_1 + C_2 \log \left( \frac{\|L^{-1}\|}{|\lambda_1 - \sigma|\tau(i)} \right) \]

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

\[ k^{(i)} \geq C_1 + C_2 \log \left( \frac{\|L^{-1}\|}{|\lambda_1 - \sigma|\tau^{(i)}} \right) \]

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

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” RHS $\mathbb{L}^{-1}x(i)$ (same behaviour as unpreconditioned solves, e.g. for fixed shifts $k(i)$ does not increase with $i$)

Condition

$$\mathbb{L}^{-T}\mathbb{L}^{-1}x(i) = x(i) \quad \text{or} \quad \mathbb{L}\mathbb{L}^{-T}x(i) = Ax(i)$$
Derivation

Aims

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

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

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

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

Condition

\[ \mathbb{L}^{-T}\mathbb{L}^{-1}x^{(i)} = x^{(i)} \quad \text{or} \quad \mathbb{L}\mathbb{L}^{-T}x^{(i)} = Ax^{(i)} \]
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" RHS $\mathbb{L}^{-1}x(i)$ (same behaviour as unpreconditioned solves, e.g. for fixed shifts $k(i)$ does not increase with $i$)

Condition

\[
\mathbb{L}^{-T}\mathbb{L}^{-1}x(i) = x(i) \quad \text{or} \quad \mathbb{L}\mathbb{L}^{-T}x(i) = Ax(i)
\]
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" RHS \( \mathbb{L}^{-1} x(i) \) (same behaviour as unpreconditioned solves, e.g. for fixed shifts \( k(i) \) does not increase with \( i \))

Condition

\[
\mathbb{L}^{-T} \mathbb{L}^{-1} x(i) = x(i) \quad \text{or} \quad \mathbb{L} \mathbb{L}^{-T} x(i) = A x(i)
\]
Choice of $L$

Theorem

With $e^{(i)} = Ax^{(i)} - LL^T x^{(i)}$ (known) and $L$ chosen such that

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

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

$L$ is a rank-one update of $L$. 
Convergence rates

The tuned preconditioner

1. retains outer convergence rates
2. provides cheap inner solves

\[ k^{(i)} \geq C_1 + C_2 \log \left( \frac{\sin \theta^{(i)}}{|\lambda_1 - \sigma|^{\tau(i)}} \right) \]

3. only a single extra back substitution with L per outer iteration needed
Fixed shift solves

SPD matrix from the Matrix Market library bscsstk10.mtx

Setup:

- decreasing tolerance $\tau^{(i)}$,
- drop tolerance $10^{-3}$,
- stopping condition: $\|r^{(i)}\| \leq 10^{-10}$. 
Fixed shift solves

Preconditioning with standard incomplete Cholesky

![Graph showing number of inner iterations vs. outer iterations]

- 1089 inner iterations in total
- 42 outer iterations
Fixed shift solves

Preconditioning with tuned incomplete Cholesky

- 42 outer iterations...
- 478 inner iterations in total...
- 1089 inner iterations in total...
Solves with Rayleigh quotient shifts

Central finite difference approximation of the self-adjoint elliptic operator

\[ A(t)u = ((1 + tx)u_x)_x + ((1 + ty)u_y)_y \]

on an equidistant grid on the unit square with Dirichlet boundary conditions and 50 nodes in each dimension. Setup:

- decreasing tolerance \( \tau^{(i)} \),
- drop tolerance \( 10^{-2} \),
- stopping condition: \( \| r^{(i)} \| \leq 10^{-14} \).
Solves with Rayleigh quotient shifts

Preconditioning with standard incomplete Cholesky

- 61 inner iterations in total
- 3 outer iterations
Solves with Rayleigh quotient shifts

Preconditioning with tuned incomplete Cholesky

The graph shows the number of inner iterations plotted against the number of outer iterations. There are two sets of data points:

- 61 inner iterations in total, with 3 outer iterations.
- 41 inner iterations in total, with 3 outer iterations.

The graph indicates that as the number of outer iterations increases, the number of inner iterations also increases linearly.
Solves with Rayleigh quotient shifts

Preconditioning with Simoncini & Eldén incomplete Cholesky

Melina Freitag and Alastair Spence
A Tuned Preconditioner
### Solves with Rayleigh quotient shifts

<table>
<thead>
<tr>
<th></th>
<th>Incompl. Cholesky</th>
<th>Tuned precond.</th>
<th>Simoncini &amp; Eldén</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DROP TOLERANCES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IT.</td>
<td>IT.</td>
<td>IT.</td>
<td>IT.</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>18</td>
<td>13</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>37</td>
<td>25</td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>total</td>
<td>61</td>
<td>41</td>
<td>52</td>
</tr>
</tbody>
</table>

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.0e-2</td>
<td>1.0e-2</td>
<td>1.0e-2</td>
</tr>
<tr>
<td></td>
<td>1.2e-3</td>
<td>2.8e-3</td>
<td>2.1e-3</td>
</tr>
<tr>
<td></td>
<td>1.9e-6</td>
<td>1.9e-6</td>
<td>2.5e-5</td>
</tr>
<tr>
<td></td>
<td>8.6e-15</td>
<td>5.1e-15</td>
<td>2.6e-9</td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>23</td>
<td>23</td>
</tr>
</tbody>
</table>

**Iteration numbers and error** $||Ax^{(i)} - \lambda^{(i)}x^{(i)}||_2$

Melina Freitag and Alastair Spence

A Tuned Preconditioner