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

Outline

Problem and Inverse Iteration

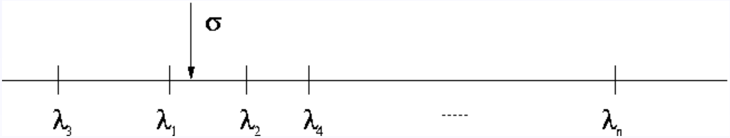
- Find an eigenvalue and eigenvector of a Hermitian positive definite A:

$$Ax = \lambda x,$$

- Inverse Iteration:

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

A large, sparse.



- Inverse iteration with preconditioned iterative solves

Outline

Inexact Inverse Iteration

for $i = 1$ to i_{max} **do**

choose $\tau^{(i)}, \sigma^{(i)}$

solve

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

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

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

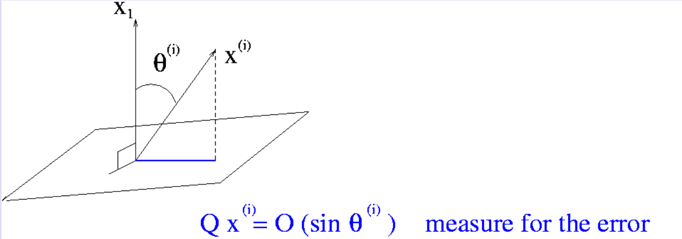
possibly: update the shift $\sigma^{(i)}$

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

end for

Error indicator

Error indicator (Orthogonal decomposition for symmetric A, Parlett)



$$x^{(i)} = \cos\theta^{(i)} x_1 + \sin\theta^{(i)} x_\perp^{(i)}, \quad x_\perp^{(i)} \perp x_1.$$

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

1. For decreasing tolerance $\tau^{(i)} \leq C\|r^{(i)}\| = \mathcal{O}(\sin\theta^{(i)})$ the inexact method recovers the rate of convergence achieved by exact solves.
2. Fixed shift σ: linear convergence.
3. Rayleigh quotient shift $\sigma^{(i)} = \rho(x^{(i)}) = \dfrac{{x^{(i)}}^T A x^{(i)}}{{x^{(i)}}^T x^{(i)}}$: 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 κ 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

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

where $\mathcal{Q}$ is the orthogonal projection onto span$\{x_2, \ldots, x_n\}$ and κ_1 is the reduced condition number $\kappa_1 = \dfrac{\max_{i=2,\ldots,n}|\lambda_i - \sigma|}{\min_{i=2,\ldots,n}|\lambda_i - \sigma|}$.

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

$$k^{(i)} \geq 2 + \kappa_1\left(\log 2|\lambda_1 - \lambda_n| + \log\frac{\|\mathcal{Q}x^{(i)}\|_2}{|\lambda_1 - \sigma|\tau^{(i)}}\right)$$

Unpreconditioned solves with MINRES

Convergence rates for solves with MINRES for simple eigenvalue

If A is positive definite and has a simple eigenvalue then

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

where $\mathcal{Q}$ is the orthogonal projection onto span$\{x_2, \ldots, x_n\}$ and κ_1 is the reduced condition number $\kappa_1 = \dfrac{\max_{i=2,\ldots,n}|\lambda_i - \sigma|}{\min_{i=2,\ldots,n}|\lambda_i - \sigma|}$.

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

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

Outline

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

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

Derivation

Aims

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

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

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)

Derivation

Aims

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

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

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 $\mathbb{L}$

Condition

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

$$\mathbb{L}\mathbb{L}^T x^{(i)} = Ax^{(i)}$$

Justification of $\mathbb{L}\mathbb{L}^T x^{(i)} = Ax^{(i)}$

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

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

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

Choice of $\mathbb{L}$

Condition

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

$$\mathbb{L}\mathbb{L}^T x^{(i)} = Ax^{(i)}$$

Justification of $\mathbb{L}\mathbb{L}^T x^{(i)} = Ax^{(i)}$

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

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

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

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.

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 $\mathbb{P}^{-1}$ is necessary

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

Sherman-Morrison formula

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

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

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

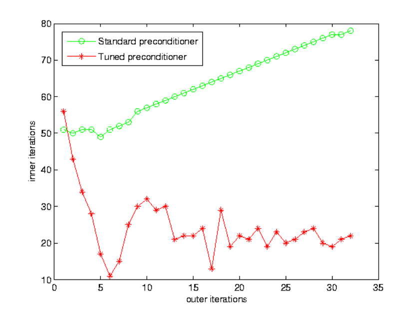
3. only a single extra back substitution with $P = LL^T$ per outer iteration needed

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

Fixed shift solves

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 $\mathbb{L}\mathbb{L}^T$

Spectral properties of preconditioned matrix

Let

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

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

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 $\mathbb{L}\mathbb{L}^T$

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 $\mathbb{L}\mathbb{L}^T$

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

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

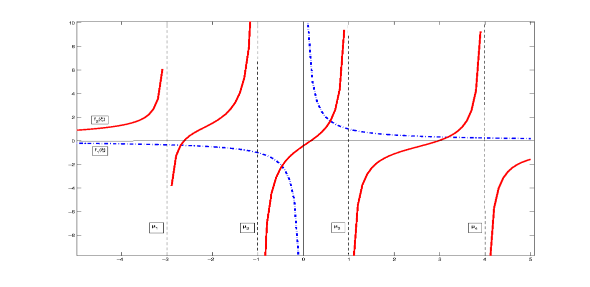
where $L^{-1}(A - \sigma I)L^{-T} = QDQ^T$, $z = Q^T v$, $(I + \alpha v v^T)Qt = \hat{w}$.

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 $\mathbb{L}\mathbb{L}^T$

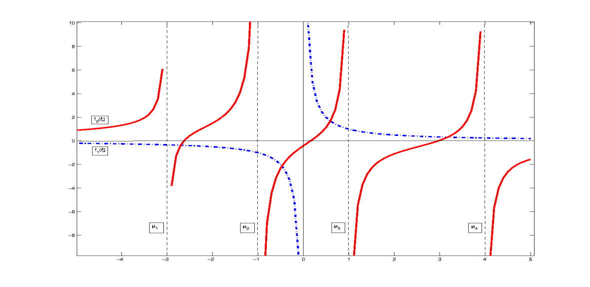
Interlacing property



- μ and ξ interlace each other depending on the sign of γ
- Clustering properties are preserved
- reduced condition number $\kappa_L^1 \leq \kappa_{\mathbb{L}}^1 \leq \kappa_L^1(1+\gamma v^T v)$

Comparison of LL^T with $\mathbb{L}\mathbb{L}^T$

Interlacing property



- μ and ξ interlace each other depending on the sign of γ
- Clustering properties are preserved
- reduced condition number $\kappa_L^1 \leq \kappa_{\mathbb{L}}^1 \leq \kappa_L^1(1 + \gamma v^T v)$

Changing the right hand side

Approach by Simoncini/Eldén [3]

Instead of solving

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

change the right hand side

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

Comparison

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.

	Simoncini & Eldén		*Tuned preconditioner*	
	DROP TOLERANCES			
OUTER ITERATION	0.25	0.1	0.25	0.1
1	67	62	29	26
2	74	66	56	55
3	85	75	71	67
4	63		18	
total	289	203	174	148

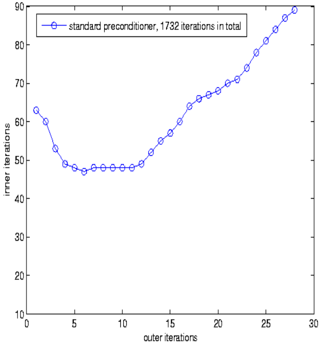
Outline

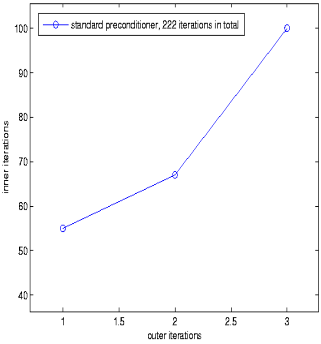
Numerical example

$Ax = \lambda Mx$ with `bcsstk08` (Structural engineering)

Figure: Fixed Shift

Figure: Rayleigh Quotient Shift

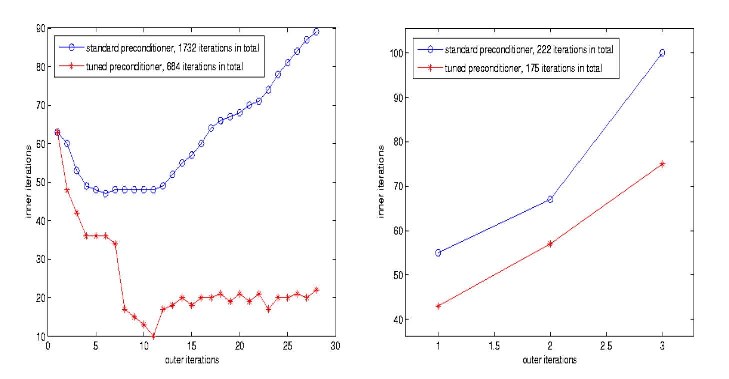




Numerical example for the generalised eigenproblem

$Ax = \lambda Mx$ with bcsstk08 (Structural engineering)

Figure: Fixed Shift Figure: Rayleigh Quotient Shift



- M. A. Freitag and A. Spence, *Convergence rates for inexact inverse iteration with application to preconditioned iterative solves*, 2006.
  Submitted to BIT.
- ——, *A tuned preconditioner for inexact inverse iteration applied to Hermitian eigenvalue problems*, 2006.
  Submitted to IMA J. Numer. Anal.
- V. Simoncini and L. Eldén, *Inexact Rayleigh quotient-type methods for eigenvalue computations*, BIT, 42 (2002), pp. 159–182.