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

Problem and Inverse Iteration

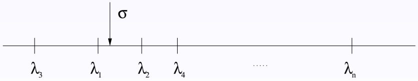
- Find an eigenvalue and eigenvector of s.p.d. A:

$$Ax = \lambda x,$$

- Inverse Iteration:

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

  A large, sparse.



Inexact Inverse Iteration

for $i = 1$ to $\dots$ **do**

choose $\tau^{(i)}$

solve

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

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

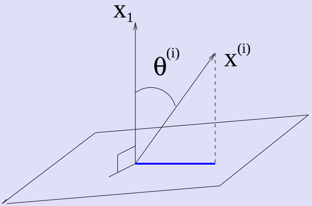
Update $\lambda^{(i+1)} = x^{(i+1)^T} A x^{(i+1)}$,

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

end for

Error indicator

Error indicator (orthogonal decomposition, Parlett)



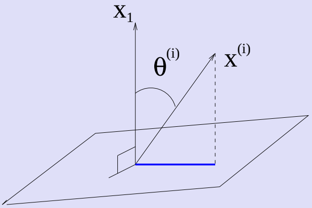
$P_{\perp} x^{(i)} = O(\sin \theta^{(i)})$ measure for the error

Eigenvalue residual

$$C|\sin\theta^{(i)}| \leq \|r^{(i)}\| \leq C'|\sin\theta^{(i)}|$$

Error indicator

Error indicator (orthogonal decomposition, Parlett)



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

Eigenvalue residual

$$C|\sin\theta^{(i)}| \leq \|r^{(i)}\| \leq C'|\sin\theta^{(i)}|$$

Convergence rates of inexact inverse iteration

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

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. [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]

Unpreconditioned solves with MINRES

Convergence rates for solves with MINRES for simple eigenvalue

If A is positive definite and has a simple eigenvalue then

$$\|x^{(i)} - (A - \sigma I)y_k^{(i)}\|_2 \leq 2\frac{|\lambda_1 - \lambda_n|}{|\lambda_1 - \sigma|}\left(\sqrt{\frac{\kappa_1 - 1}{\kappa_1 + 1}}\right)^{k-1} \|\mathcal{P}_1^{\perp} x^{(i)}\|_2.$$

Number of inner solves for each i

$$k^{(i)} \geq C_1 + C_2 \log\left(\frac{\|\mathcal{P}_1^{\perp} x^{(i)}\|_2}{|\lambda_1 - \sigma|\tau^{(i)}}\right)$$

Unpreconditioned solves with MINRES

Convergence rates for solves with MINRES for simple eigenvalue

If A is positive definite and has a simple eigenvalue then

$$\|x^{(i)} - (A - \sigma I)y_k^{(i)}\|_2 \leq 2\frac{|\lambda_1 - \lambda_n|}{|\lambda_1 - \sigma|}\left(\sqrt{\frac{\kappa_1 - 1}{\kappa_1 + 1}}\right)^{k-1} \|\mathcal{P}_1^{\perp} x^{(i)}\|_2.$$

Number of inner solves for each i

$$k^{(i)} \geq C_1 + C_2 \log\left(\frac{\|\mathcal{P}_1^{\perp} x^{(i)}\|_2}{|\lambda_1 - \sigma|\tau^{(i)}}\right)$$

Unpreconditioned solves with MINRES

Convergence rates for solves with MINRES for simple eigenvalue

If A is positive definite and has a simple eigenvalue then

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

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

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

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

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 $L \longrightarrow \mathbb{L}$

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

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 $L \longrightarrow \mathbb{L}$

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

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 $L \longrightarrow \mathbb{L}$

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

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 $L \to \mathbb{L}$

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

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

Condition

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

Choice of $\mathbb{L}$

Theorem

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

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

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

$\mathbb{L}$ is a rank-one update of L.

Convergence rates

The tuned preconditioner

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

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

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

Fixed shift solves

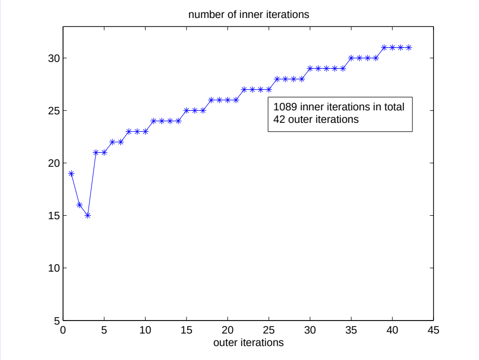
SPD matrix from the Matrix Market library `bscsstk10.mtx`

`Setup:`

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

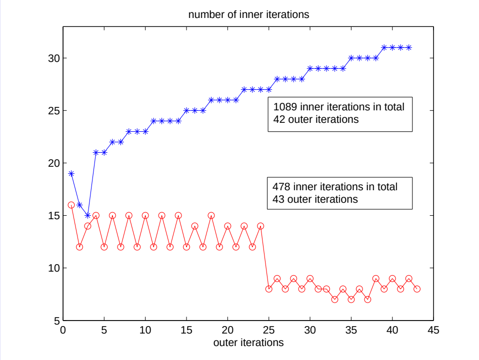
Fixed shift solves

Preconditioning with standard incomplete Cholesky



Fixed shift solves

Preconditioning with tuned incomplete Cholesky



Solves with Rayleigh quotient shifts

Central finite difference approximation of the self-adjoint elliptic operator

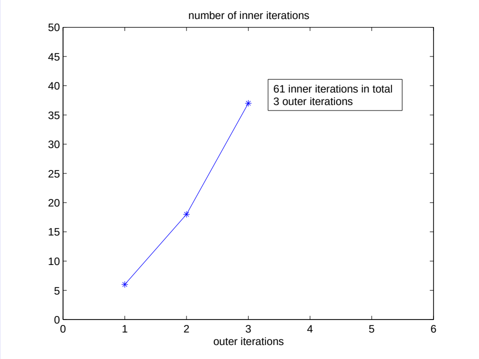
$$\mathcal{A}(t)u = ((1+tx)u_x)_x + ((1+ty)u_y)_y$$

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: $\|r^{(i)}\| \leq 10^{-14}$.

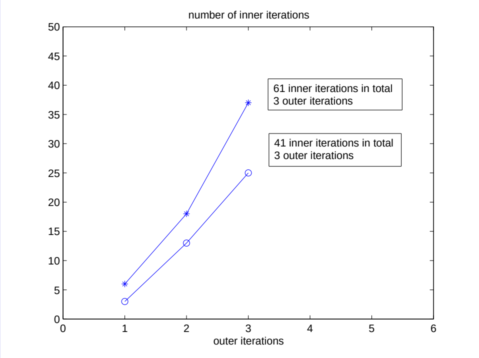
Solves with Rayleigh quotient shifts

Preconditioning with standard incomplete Cholesky



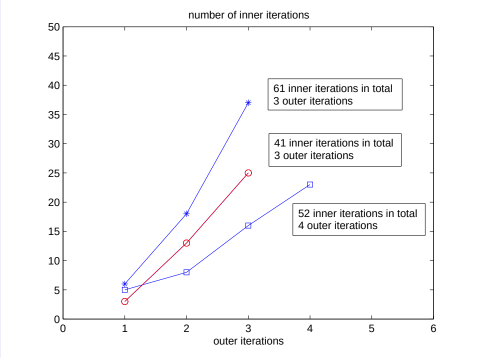
Solves with Rayleigh quotient shifts

Preconditioning with tuned incomplete Cholesky



Solves with Rayleigh quotient shifts

Preconditioning with Simoncini & Eldén incomplete Cholesky



Solves with Rayleigh quotient shifts

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

	Incompl. Cholesky		*Tuned precond.*		*Simoncini & Eldén*	
	Drop Tolerances					
	It.		It.		It.	
		1.0e-2		1.0e-2		1.0e-2
1	6	1.2e-3	3	2.8e-3	5	2.1e-3
2	18	1.9e-6	13	1.9e-6	8	2.5e-5
3	37	8.6e-15	25	5.1e-15	16	2.6e-9
4					23	2.6e-15
total	61		41		52	

M. A. Freitag and A. Spence, *Convergence rates for inexact inverse iteration with application to preconditioned iterative solves*, 2005.
Submitted to BIT.

———, *A tuned preconditioner for inexact inverse iteration applied to Hermitian eigenvalue problems*, 2005.
In preparation.