A tuned preconditioner for Hermitian eigenproblems

Melina Freitag

Department of Mathematical Sciences, University of Bath, UK†

email: m.freitag@maths.bath.ac.uk

Introduction

Consider the computation of a simple eigenvalue and corresponding eigenvector of a large sparse Hermitian positive definite matrix using either inexact inverse iteration with a fixed or Rayleigh quotient shift.

- large sparse linear systems solved approximately by means of symmetrically preconditioned MINRES,
- preconditioners (incomplete Cholesky factorisation)
- derivation of a new tuned Cholesky preconditioner,
- analysis using the convergence theory for MINRES,
- comparison of spectral properties of the tuned with those of the standard preconditioned matrix,
- perturbation and interlacing results.

Inexact inverse iteration (III) with fixed shift

Given σ and $\mathbf{x}^{(0)}$ with $\|\mathbf{x}^{(0)}\| = 1$. For $i = 0, 1, 2, \ldots$

- Choose $\tau^{(i)}$,
- Solve $(\mathbf{A} - \sigma\mathbf{I})\mathbf{y}^{(i)} = \mathbf{x}^{(i)}$ inexactly, that is,

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

- Compute $\mathbf{x}^{(i+1)} = \dfrac{\mathbf{y}^{(i)}}{\|\mathbf{y}^{(i)}\|}$,
- Compute $\lambda^{(i+1)} = \mathbf{x}^{(i+1)^*}\mathbf{A}\mathbf{x}^{(i+1)}$,
- Evaluate $\mathbf{r}^{(i+1)} = (\mathbf{A} - \lambda^{(i+1)}\mathbf{I})\mathbf{x}^{(i+1)}$,
- Test for convergence.

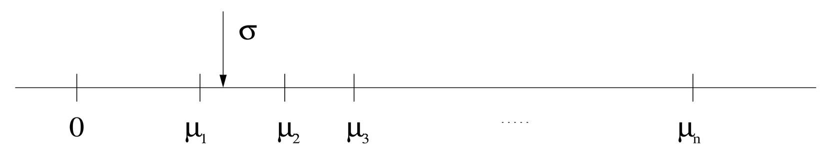
Convergence rates

For a decreasing tolerance $\tau^{(i)} = C\|\mathbf{r}^{(i)}\| = \mathcal{O}(\sin\theta^{(i)})$ and close enough starting guesses the inexact method recovers the rate of convergence achieved by exact solves.

- Fixed shift: linear convergence ([4], [1]).
- Rayleigh quotient shift: cubic convergence ([1], [5]).

Convergence theory of MINRES

- symmetric $\mathbf{B}$ has eigenvalues $\mu_1, \ldots, \mu_n$ and eigenvectors $\mathbf{w}_1, \ldots, \mathbf{w}_n$, $\kappa^1 = \dfrac{|\mu_n|}{|\mu_1|}$ reduced condition number,



- $\mathcal{P}^{\perp}$ orthogonal projection along $\mathbf{w}_1$ onto $\text{span}\{\mathbf{w}_2, \ldots, \mathbf{w}_n\}$.

If $\mathbf{z}_k$ is the result of applying MINRES to $\mathbf{B}\mathbf{z} = \mathbf{b}$ with starting value $\mathbf{z}_0 = \mathbf{0}$ then

$$\|\mathbf{b} - \mathbf{B}\mathbf{z}_k\| \leq 2 \max_{j=2,\ldots,n} \frac{|\mu_1 - \mu_j|}{|\mu_1|} \left(\frac{\sqrt{\kappa^1} - 1}{\sqrt{\kappa^1} + 1}\right)^{k-1} \|\mathcal{P}^{\perp}\mathbf{b}\|, \quad (1)$$

If, using $\|\mathcal{P}^{\perp}\mathbf{b}\| = |\sin\theta^{(i)}|$, the number of inner iterations satisfies

$$k^{(i)} \geq 1 + \frac{\sqrt{\kappa^1}}{2}\left(\log 2\frac{|\lambda_1 - \lambda_n|}{|\lambda_1 - \sigma|} + \log\frac{|\sin\theta^{(i)}|}{\tau^{(i)}}\right). \quad (2)$$

then $\|\mathbf{b} - \mathbf{B}\mathbf{z}_k\| \leq \tau^{(i)}$. The number of inner iterations does not increase with i, if $|\lambda_1 - \sigma|$ is fixed and $\tau^{(i)} = \mathcal{O}(\sin\theta^{(i)})$.

Preconditioned inexact inverse iteration

Let $\mathbf{A}$ be Hermitian positive definite and consider the incomplete Cholesky factorisation $\mathbf{L}\mathbf{L}^*$, that is,

$$\mathbf{A} = \mathbf{L}\mathbf{L}^* + \mathbf{E}. \quad (3)$$

Solve

$$\mathbf{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbf{L}^{-*}\tilde{\mathbf{y}}^{(i)} = \mathbf{L}^{-1}\mathbf{x}^{(i)}, \quad \mathbf{y}^{(i)} = \mathbf{L}^{-*}\tilde{\mathbf{y}}^{(i)} \quad (4)$$

to a tolerance $\tau^{(i)}\|\mathbf{L}\|^{-1}$ so that $\|\mathbf{x}^{(i)} - (\mathbf{A} - \sigma\mathbf{I})\mathbf{y}^{(i)}\| \leq \tau^{(i)}$.

- does not change the linear outer rate of convergence
- number of iterations

$$k_{\mathbf{L}}^{(i)} \geq 1 + \frac{\sqrt{\kappa_{\mathbf{L}}^1}}{2}\left(\log 2\frac{|\mu_1 - \mu_n|\|\mathbf{L}\|\|\mathbf{L}^{-1}\|}{|\mu_1|} + \log\frac{1}{\tau^{(i)}}\right). \quad (5)$$

increases for $\tau^{(i)} \to 0$.

The tuned preconditioner

Solve the preconditioned Hermitian system

$$\mathbb{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbb{L}^{-*}\tilde{\mathbf{y}}^{(i)} = \mathbb{L}^{-1}\mathbf{x}^{(i)}, \quad \mathbf{y}^{(i)} = \mathbb{L}^{-*}\tilde{\mathbf{y}}^{(i)}, \quad (6)$$

inexactly, where $\mathbb{L}$ is chosen such that the right hand side of (6) is close to the eigenvector of $\mathbb{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbb{L}^{-*}$ corresponding to the eigenvalue closest to zero.

- reproduces the inner iteration behaviour observed for unpreconditioned solves,
- requires the preconditioner $\mathbb{L}\mathbb{L}^*$ to satisfy

$$\mathbb{L}\mathbb{L}^*\mathbf{x}^{(i)} = \mathbf{A}\mathbf{x}^{(i)}. \quad (7)$$

Then

$$\|\mathcal{P}^{\perp}\mathbb{L}^{-1}\mathbf{x}^{(i)}\| \leq C_2\|\mathbf{r}^{(i)}\|, \quad (8)$$

and with $\tau^{(i)} = C_1\|\mathbf{r}^{(i)}\|$ we obtain

$$k_{\mathbb{L}}^{(i)} \geq 1 + \frac{\sqrt{\kappa_{\mathbb{L}}^1}}{2}\left(\log 2\frac{|\xi_1 - \xi_n|}{|\xi_1|} + \log\frac{C_2}{C_1}\right), \quad (9)$$

that is no increase with i.

Implementation

Let $\mathbf{A}$ be Hermitian positive definite and consider its incomplete Cholesky factorisation $\mathbf{L}\mathbf{L}^*$, $\mathbf{A} = \mathbf{L}\mathbf{L}^* + \mathbf{E}$.

- $\mathbf{x}^{(i)}$ approximate eigenvector from the ith iteration,
- $\mathbf{u}^{(i)} = \mathbf{E}\mathbf{x}^{(i)} = \mathbf{A}\mathbf{x}^{(i)} - \mathbf{L}\mathbf{L}^*\mathbf{x}^{(i)}$ and $\mathbf{v}^{(i)} = \mathbf{L}^{-1}\mathbf{u}^{(i)}$,
- assume $\mathbf{u}^{(i)^*}\mathbf{x}^{(i)} \neq 0$, $\gamma^{(i)} := \dfrac{1}{\mathbf{u}^{(i)^*}\mathbf{x}^{(i)}}$,
- assume

$$1 + \gamma^{(i)}\mathbf{v}^{(i)^*}\mathbf{v}^{(i)} \geq 0 \quad (10)$$

and set

$$\alpha^{(i)} = \frac{-1 \pm \sqrt{1 + \gamma^{(i)}\mathbf{v}^{(i)^*}\mathbf{v}^{(i)}}}{\mathbf{v}^{(i)^*}\mathbf{v}^{(i)}}. \quad (11)$$

If $\mathbb{L}$ in (6) is chosen such that

$$\mathbb{L} = \mathbf{L} + \alpha^{(i)}\mathbf{u}^{(i)}\mathbf{v}^{(i)^*}, \quad (12)$$

then $\mathbb{L}\mathbb{L}^*\mathbf{x}^{(i)} = \mathbf{A}\mathbf{x}^{(i)}$.

- retains outer rate of convergence,
- provides cheap inner solves,
- only one single extra back substitution with $\mathbf{L}$ per outer iteration (Sherman-Morrison formula).

Numerical Example

Consider the matrix `nos5.mtx` from the Matrix Market.

- preconditioned III with decreasing tolerance
- fixed shift $\sigma = 100$, finds third smallest eigenvalue
- incomplete Cholesky factorisation with drop tol 0.1.

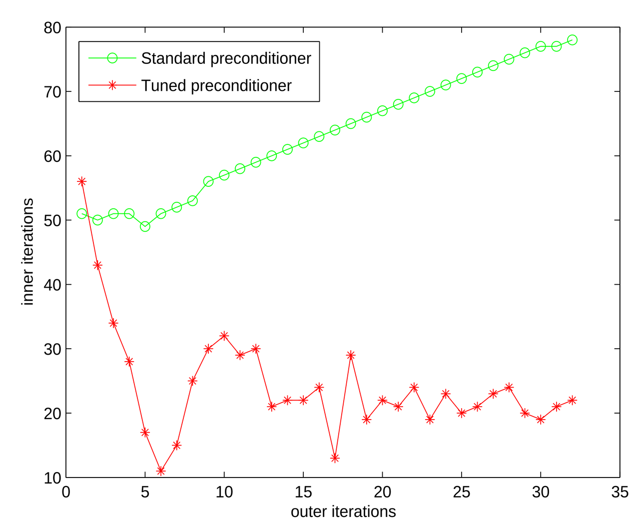


Figure 2: Inner iterations against outer iterations for the standard and tuned Cholesky preconditioner

Spectral analysis - Perturbation

Comparison of the spectral properties of

$$\mathbf{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbf{L}^{-*} \quad \text{and} \quad \mathbb{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbb{L}^{-*}.$$

Define $\mathbf{S} = \mathbf{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbf{L}^{-*}$ and consider the two eigenvalue problems

$$\mathbf{S}\mathbf{w} = \mu\mathbf{w} \quad (13)$$

and

$$\mathbf{S}\mathbf{w}' = \xi(\mathbf{I} + \gamma\mathbf{v}\mathbf{v}^*)\mathbf{w}'. \quad (14)$$

Then μ and ξ are nonzero and

$$\min_{\mu\in\Lambda(\mathbf{S})}\left|\frac{\mu - \xi}{\xi}\right| \leq |\gamma\mathbf{v}^*\mathbf{v}|. \quad (15)$$

Spectral analysis - Interlacing

Consider the two eigenvalue problems

$$\mathbf{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbf{L}^{-*}\mathbf{w} = \mu\mathbf{w} \quad (16)$$

and

$$\mathbb{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbb{L}^{-*}\hat{\mathbf{w}} = \xi\hat{\mathbf{w}}, \quad (17)$$

and assume condition (10) holds. Suppose $\mathbf{D} = \text{diag}(\mu_1 < \ldots < \mu_n) \in \mathbb{R}^{n\times n}$. Transform the problem to a generalised eigenproblem

$$\mathbf{D}\mathbf{t}_j = \xi_j(\mathbf{I} + \gamma\mathbf{z}\mathbf{z}^*)\mathbf{t}_j, \quad (18)$$

where ξ_j are the eigenvalues, with $\xi_1 \leq \ldots \leq \xi_n$ and $\mathbf{t}_j$ are the corresponding eigenvectors. Also, let $\mu_1 < \ldots < \mu_p < 0 < \mu_{p+1} < \ldots < \mu_n$, where p is the number of negative eigenvalues of $\mathbf{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbf{L}^{-*}$. Then

- The ξ_j are the n zeros of $f(\xi) = 1 - \xi\gamma\mathbf{z}^*(\mathbf{D} - \xi\mathbf{I})^{-1}\mathbf{z}$.
- If $\gamma > 0$, then

$$\mu_1 < \xi_1 < \mu_2 < \xi_2 < \ldots < \mu_p < \xi_p < 0$$

and

$$0 < \xi_{p+1} < \mu_{p+1} < \xi_{p+2} < \mu_{p+2} < \ldots < \xi_n < \mu_n,$$

that is the eigenvalues are shifted towards zero, while for $\gamma < 0$ the eigenvalues are shifted away from zero

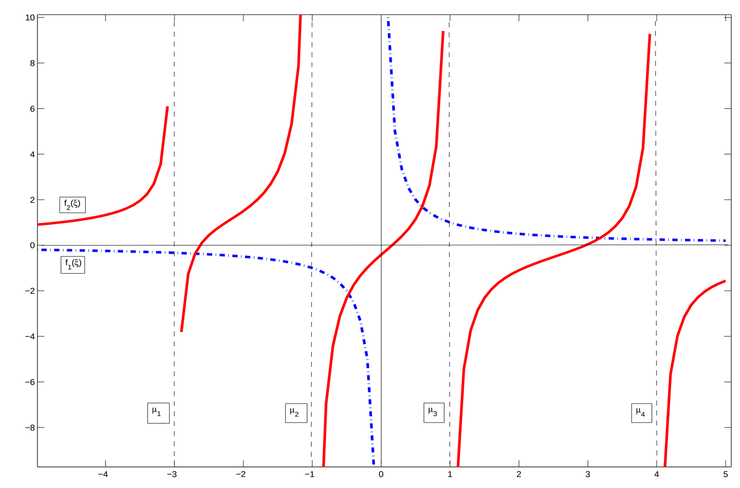


Figure 3: Interlacing property for $\gamma > 0$

Using perturbation and interlacing results we obtain that eigenvalue clustering properties of $\mathbf{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbf{L}^{-*}$ are preserved in $\mathbb{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbb{L}^{-*}$ and in particular

$$\kappa_{\mathbf{L}}^1 \leq \kappa_{\mathbb{L}}^1 \leq \kappa_{\mathbf{L}}^1(1 + |\gamma\mathbf{v}^*\mathbf{v}|). \quad (19)$$

Inexact RQ iteration

- preconditioned III with decreasing/fixed tolerance
- find third smallest eigenvalue, Rayleigh quotient shift
- incomplete Cholesky factorisation

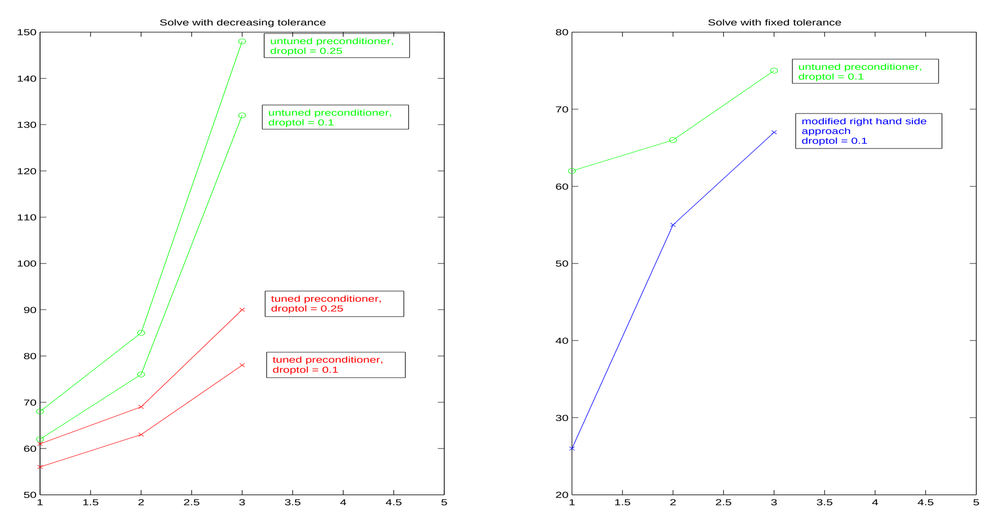


Figure 4: Inner iterations against outer iterations for the standard and tuned Cholesky preconditioner

Conclusions

For III the tuning of the preconditioner reduces the number of inner iterations for the iterative solves in each step.

References

[1] J. Berns-Müller, I. G. Graham, and A. Spence. Inexact inverse iteration for symmetric matrices, 2005. Submitted to Linear Algebra Appl.

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

[3] M. A. Freitag and A. Spence. A tuned preconditioner for inexact inverse iteration applied to Hermitian eigenvalue problems, 2005. Submitted to SIMAX.

[4] G. H. Golub and Q. Ye. Inexact inverse iteration for generalized eigenvalue problems. *BIT*, 40(4):671–684, 2000.

[5] P. Smit and M. H. C. Paardekooper. The effects of inexact solvers in algorithms for symmetric eigenvalue problems. *Linear Algebra and its Applications*, 287:337–357, 1999.

†Address for correspondence: Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY