Lecture 3: Inexact inverse iteration with preconditioning

Alastair Spence

Department of Mathematical Sciences
University of Bath

CLAPDE, Durham, July 2008

Joint work with M. Freitag (Bath), and M. Robbé & M. Sadkane (Brest)

Outline

Introduction

$$Ax = \lambda x, \quad \lambda \in \mathbb{C}, x \in \mathbb{C}^n$$

- Lecture 2: Detect pure imaginary eigevalues of large sparse matrices
- Seek λ near a given shift σ (good estimate eg. continuation).
- A is large, sparse, nonsymmetric (discretised PDE: $Ax = \lambda Mx$)

Introduction

$$Ax = \lambda x, \quad \lambda \in \mathbb{C}, x \in \mathbb{C}^n$$

- Lecture 2: Detect pure imaginary eigevalues of large sparse matrices
- Seek λ near a given shift σ (good estimate eg. continuation).
- A is large, sparse, nonsymmetric (discretised PDE: $Ax = \lambda Mx$)
- Inverse Iteration:
  - $y = (A - \sigma I)^{-1} x$
  - Solve $(A - \sigma I)y = x$
- Preconditioned iterative solves

Introduction

$$Ax = \lambda x, \quad \lambda \in \mathbb{C}, x \in \mathbb{C}^n$$

- Lecture 2: Detect pure imaginary eigevalues of large sparse matrices
- Seek λ near a given shift σ (good estimate eg. continuation).
- A is large, sparse, nonsymmetric (discretised PDE: $Ax = \lambda Mx$)
- Inverse Iteration:
  - $y = (A - \sigma I)^{-1} x$
  - Solve $(A - \sigma I)y = x$
- Preconditioned iterative solves
- Extensions
  - Inverse Subspace Iteration
  - Jacobi-Davidson method
  - Shift-invert Arnoldi method (Melina Freitag: Tuesday lecture)

Inexact inverse iteration

- Assume $x^{(i)}$ is an approximate normalised eigenvector
- Iterative solves (e.g. GMRES) of

$$(A - \sigma I)y = x^{(i)}$$

Inexact inverse iteration

- Assume $x^{(i)}$ is an approximate normalised eigenvector
- Iterative solves (e.g. GMRES) of

$$(A - \sigma I)y = x^{(i)}$$

- inner-outer
- $\|x^{(i)} - (A - \sigma I)y_k\| \leq \tau^{(i)}$, ($\tau^{(i)}$ = solve tolerance)
- Rescale y_k to get $x^{(i+1)}$
- Update shift?

Inexact inverse iteration

- Assume $x^{(i)}$ is an approximate normalised eigenvector
- Iterative solves (e.g. GMRES) of

$$(A - \sigma I)y = x^{(i)}$$

- inner-outer
- $\|x^{(i)} - (A - \sigma I)y_k\| \leq \tau^{(i)}$, ($\tau^{(i)}$ = solve tolerance)
- Rescale y_k to get $x^{(i+1)}$
- Update shift?
- (Right) preconditioned solves
  1. P^{-1} "known"
  2. $$(A - \sigma I)P^{-1}\tilde{y} = x^{(i)} \quad , P^{-1}\tilde{y} = y.$$

Convergence of inexact inverse iteration

- Given $x^{(i)}$ and $\lambda^{(i)}$

$$r^{(i)} = Ax^{(i)} - \lambda^{(i)} x^{(i)} \quad \text{Eigenvalue residual}$$

Theorem (Convergence)

If the solve tolerance, $\tau^{(i)}$, is chosen to reduce proportional to the norm of the eigenvalue residual $\|r^{(i)}\|$ then we recover the rate of convergence achieved when using direct solves.

- Other options/strategies possible: For example Rayleigh quotient iteration with a fixed tolerance converges linearly.

Numerical Example

$Ax = \lambda x$

- discretisation of convection-diffusion operator

$$-\Delta u + 5u_x + 5u_y = \lambda u \quad \text{on} \quad (0,1)^2,$$

- 3 experiments:
  1. Rayleigh quotient shift; exact solves
  2. Rayleigh quotient shift; with decreasing solve tolerance in GMRES

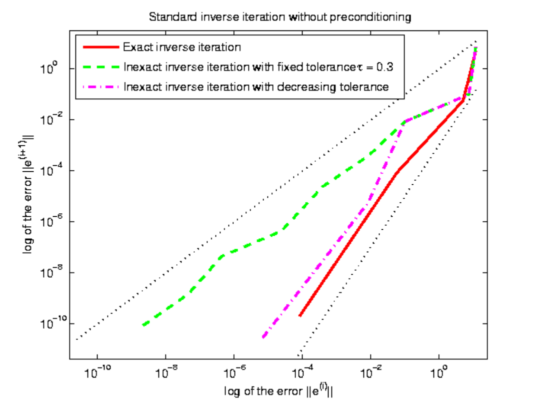
$$\tau^{(i)} = \min\{\tau, \tau\|r^{(i)}\|\}, \quad \text{with} \quad \tau = 0.3$$

  3. Rayleigh quotient shift; with fixed tolerance $\tau = 0.3$
- In all cases solve till

$$\left\| \frac{r^{(i)}}{\lambda^{(i)}} \right\| < 10^{-10}$$

Numerical Example

Linear and Quadratic convergence



Outline

Inverse Power Method with and without preconditioned solves

- From now on, assume $\sigma = 0$. So: $Ay = x^{(i)}$

Inverse Power Method with and without preconditioned solves

- From now on, assume $\sigma = 0$. So: $Ay = x^{(i)}$
- $AP^{-1}\tilde{y} = x^{(i)} \quad , P^{-1}\tilde{y} = y$.
- Always assume decreasing tolerance: $\tau^{(i)} = C\|Ax^{(i)} - \lambda^{(i)}x^{(i)}\|$
- Convection-Diffusion Example;
  1. smallest eigenvalue: $\lambda_1 \approx 32.18560954$,
  2. Preconditioned GMRES with tolerance $\tau^{(i)} = 0.01\|r^{(i)}\|$,
  3. ILU based preconditioners.

Convection-Diffusion problem: No Preconditioning - $\|Ay_k - x^{(i)}\| \leq \tau^{(i)}$

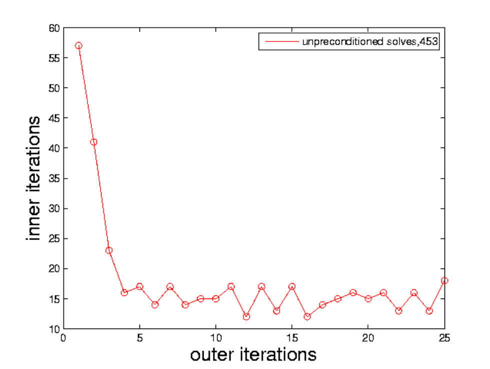


Figure: Inner iterations vs outer iterations

Question

Why is there no increase in inner iterations as i increases?

Convection-Diffusion problem: Preconditioning - $\|AP^{-1}\tilde{y}_k - x^{(i)}\| \leq \tau^{(i)}$

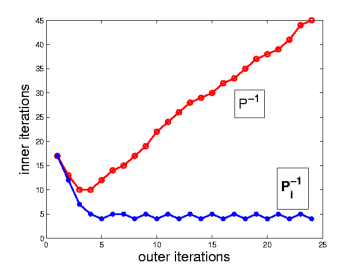


Figure: Inner iterations vs outer iterations

Question

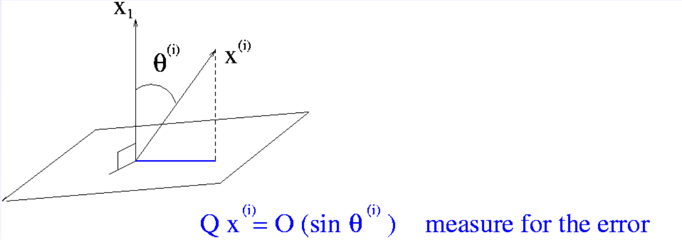
Why is $\mathbb{P}_i^{-1}$ better than P^{-1}?

Note

$\mathbb{P}_i$ is a rank-one change to P

Theory: Unpreconditioned solves to find λ_1, x_1

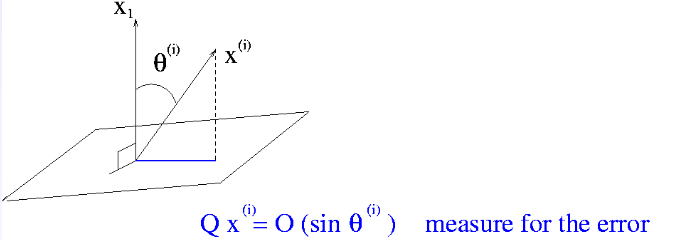
- $x^{(i)}$ is approximation to x_1



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

Theory: Unpreconditioned solves to find λ_1, x_1

- $x^{(i)}$ is approximation to x_1



- $x^{(i)} = \cos\theta^{(i)} x_1 + \sin\theta^{(i)} x_\perp$
- $r^{(i)} = Ax^{(i)} - \lambda^{(i)} x^{(i)}, \quad \|r^{(i)}\| \leq C|\sin\theta^{(i)}|$
- Parlett (1998) - ideas extend to nonsymmetric problems.

GMRES applied to $Ay = x^{(i)}$

- y_k after k steps
- $\|x^{(i)} - Ay_k\| \leq \tau^{(i)} = C\|r^{(i)}\|$

GMRES applied to $Ay = x^{(i)}$

- y_k after k steps
- $\|x^{(i)} - Ay_k\| \le \tau^{(i)} = C\|r^{(i)}\|$
- $$\begin{aligned} \|x^{(i)} - Ay_k\| &= \min \|p_k(A)x^{(i)}\| \\ &\le \min \|q_{k-1}(A)(I - \frac{1}{\lambda_1}A)(\cos\theta^{(i)}x_1 + \sin\theta^{(i)}x_\perp)\| \\ &\le C\rho^{k-1}|\sin\theta^{(i)}|, \quad 0 < \rho < 1. \end{aligned}$$
- $k \ge 1 + C_1\left(\log C_2 + \log \frac{|\sin\theta^{(i)}|}{\tau^{(i)}}\right)$
- bound on k does not increase with i.

GMRES applied to $Ay = x^{(i)}$

- y_k after k steps
- $\|x^{(i)} - Ay_k\| \leq \tau^{(i)} = C\|r^{(i)}\|$
-

$$\begin{aligned} \|x^{(i)} - Ay_k\| &= \min \|p_k(A)x^{(i)}\| \\ &\leq \min \|q_{k-1}(A)(I - \frac{1}{\lambda_1}A)(\cos\theta^{(i)}x_1 + \sin\theta^{(i)}x_\perp)\| \\ &\leq C\rho^{k-1}|\sin\theta^{(i)}|, \quad 0 < \rho < 1. \end{aligned}$$

- $k \geq 1 + C_1\left(\log C_2 + \log \frac{|\sin\theta^{(i)}|}{\tau^{(i)}}\right)$
- bound on k does not increase with i.
- Reason for no increase? $x^{(i)} = \cos\theta^{(i)}x_1 + \sin\theta^{(i)}x_\perp$

$$x^{(i)} = \boxed{\text{eigenvector of } A \quad + \quad \text{“term”} \quad \rightarrow 0}$$

GMRES applied to $AP^{-1}\tilde{y} = x^{(i)}$

- $AP^{-1}u_1 = \mu_1 u_1$: (μ_1, u_1) eigenpair nearest zero of AP^{-1}
- $x^{(i)} = \cos\tilde{\theta}^{(i)} u_1 + \sin\tilde{\theta}^{(i)} u_\perp$

GMRES applied to $AP^{-1}\tilde{y} = x^{(i)}$

- $AP^{-1}u_1 = \mu_1 u_1$: (μ_1, u_1) eigenpair nearest zero of AP^{-1}
- $x^{(i)} = \cos\tilde{\theta}^{(i)} u_1 + \sin\tilde{\theta}^{(i)} u_\perp$
- $k \geq 1 + \tilde{C}_1 \left(\log \tilde{C}_2 + \log \frac{|\sin\tilde{\theta}^{(i)}|}{\tau^{(i)}} \right)$
- BUT $\sin\tilde{\theta}^{(i)} \to 0$ only if $u_1 \in \mathrm{span}\{x_1\}$ generally won't hold

GMRES applied to $AP^{-1}\tilde{y} = x^{(i)}$

- $AP^{-1}u_1 = \mu_1 u_1$: (μ_1, u_1) eigenpair nearest zero of AP^{-1}
- $x^{(i)} = \cos\tilde{\theta}^{(i)} u_1 + \sin\tilde{\theta}^{(i)} u_\perp$
- $k \geq 1 + \tilde{C}_1 \left(\log \tilde{C}_2 + \log \frac{|\sin\tilde{\theta}^{(i)}|}{\tau^{(i)}} \right)$
- BUT $\sin\tilde{\theta}^{(i)} \to 0$ only if $u_1 \in \mathrm{span}\{x_1\}$ generally won't hold
- $\sin\tilde{\theta}^{(i)} \not\to 0$
- inner iteration costs increase with i.
- Reason: $x^{(i)} = \cos\tilde{\theta}^{(i)} u_1 + \sin\tilde{\theta}^{(i)} u_\perp$

$$x^{(i)} = \boxed{\text{eigenvector of} \quad AP^{-1} \quad + \quad \text{“term”} \quad \not\to 0}$$

Convection-Diffusion problem: Preconditioning - $\|AP^{-1}\tilde{y}_k - x^{(i)}\| \leq \tau^{(i)}$

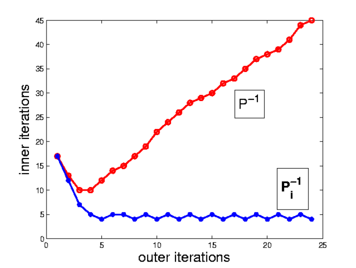


Figure: Inner iterations vs outer iterations

Question

Why is $\mathbb{P}_i^{-1}$ better than P^{-1}?

New "tuned" preconditioner $\mathbb{P}_i$

- Idea: recreate the good relationship between the right hand side and the iteration matrix
- $$x^{(i)} = \text{eigenvector of iteration matrix} \quad + \quad \text{"term"} \quad \to 0$$

New "tuned" preconditioner $\mathbb{P}_i$

- Idea: recreate the good relationship between the right hand side and the iteration matrix
- $$x^{(i)} = \text{eigenvector of iteration matrix} \quad + \quad \text{“term”} \quad \rightarrow 0$$
- Define
$$\mathbb{P}_i = P + (A - P)x^{(i)}{x^{(i)}}^H$$
- $\mathbb{P}_i$ is a rank one change to P (Sherman-Morrison)

New "tuned" preconditioner $\mathbb{P}_i$

- Idea: recreate the good relationship between the right hand side and the iteration matrix
- $$x^{(i)} = \text{eigenvector of iteration matrix} \quad + \quad \text{"term"} \quad \to 0$$
- Define
$$\mathbb{P}_i = P + (A - P)x^{(i)}x^{(i)H}$$
- $\mathbb{P}_i$ is a rank one change to P (Sherman-Morrison)
- $\mathbb{P}_i x^{(i)} = P x^{(i)} + (A - P)x^{(i)}x^{(i)H}x^{(i)}$
- $A x^{(i)} = \mathbb{P}_i x^{(i)}$

New "tuned" preconditioner $\mathbb{P}_i$

- Idea: recreate the good relationship between the right hand side and the iteration matrix
- $$x^{(i)} = \text{eigenvector of iteration matrix} \quad + \quad \text{“term”} \quad \to 0$$
- Define
$$\mathbb{P}_i = P + (A - P)x^{(i)}x^{(i)H}$$
- $\mathbb{P}_i$ is a rank one change to P (Sherman-Morrison)
- $\mathbb{P}_i x^{(i)} = Px^{(i)} + (A - P)x^{(i)}x^{(i)H}x^{(i)}$
- $Ax^{(i)} = \mathbb{P}_i x^{(i)}$
- Hence
$$A\mathbb{P}_i^{-1}Ax^{(i)} = Ax^{(i)}$$
- $Ax^{(i)}$ is an eigenvector of $A\mathbb{P}_i^{-1}$

GMRES with the tuned preconditioner

Recall

- $A\mathbb{P}_i^{-1}\tilde{y} = x^{(i)}$
- $A\mathbb{P}_i^{-1}Ax^{(i)} = Ax^{(i)}$

Is $x^{(i)}$ a "nice" RHS for $A\mathbb{P}_i^{-1}$?

GMRES with the tuned preconditioner

Recall

- $A\mathbb{P}_i^{-1}\tilde{y} = x^{(i)}$
- $A\mathbb{P}_i^{-1}Ax^{(i)} = Ax^{(i)}$

Is $x^{(i)}$ a "nice" RHS for $A\mathbb{P}_i^{-1}$?

- $r^{(i)} = Ax^{(i)} - \lambda^{(i)}x^{(i)} \quad \Rightarrow \quad x^{(i)} = \frac{1}{\lambda^{(i)}}Ax^{(i)} - \frac{1}{\lambda^{(i)}}r^{(i)}$
- Idea of tuning: change iteration matrix so that

$$x^{(i)} = \boxed{\text{eigenvector of } A\mathbb{P}_i^{-1} \quad + \quad \text{“term”} \quad \to 0}$$

GMRES with the tuned preconditioner

Recall

- $A\mathbb{P}_i^{-1}\tilde{y} = x^{(i)}$
- $A\mathbb{P}_i^{-1}Ax^{(i)} = Ax^{(i)}$

Is $x^{(i)}$ a "nice" RHS for $A\mathbb{P}_i^{-1}$?

- $r^{(i)} = Ax^{(i)} - \lambda^{(i)}x^{(i)} \quad \Rightarrow \quad x^{(i)} = \frac{1}{\lambda^{(i)}}Ax^{(i)} - \frac{1}{\lambda^{(i)}}r^{(i)}$
- Idea of tuning: change iteration matrix so that

$$x^{(i)} = \boxed{\text{eigenvector of } A\mathbb{P}_i^{-1} \quad + \quad \text{"term"} \quad \rightarrow 0}$$

- Analysis of GMRES is essentially the same as for unpreconditioned case
- No increase in inner iterations as i increases

Convection-Diffusion problem: Preconditioning - $\|AP^{-1}\tilde{y}_k - x^{(i)}\| \leq \tau^{(i)}$

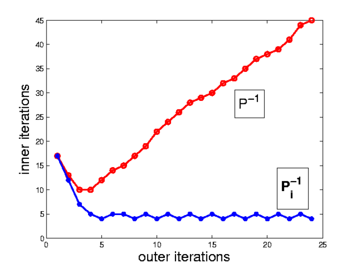


Figure: Inner iterations vs outer iterations

Question and Answer

Why is $\mathbb{P}_i^{-1}$ better than P^{-1}? $\mathbb{P}_i^{-1}$ is tuned so that the rhs of the preconditioned system is "good" for the iteration matrix $A\mathbb{P}_i^{-1}$

Numerical Example (Freitag/Sp./Vainikko)

- Linearised Stability on Navier-Stokes: Flow past a circular cylinder (Re=25)
- $Ax = \lambda Mx$
- Both Rayleigh Quotient and fixed shifts
- Mixed FEM $Q_2 - Q_1$ elements with $n = 6734, 27294, 61678$
- FGMRES with block preconditioner of Elman
- seek "dangerous" complex eigenvalue near imaginary axis ($\approx 10i$)
- stop when residual $\leq 10^{-11}$

Numerics for Navier-Stokes example, $Ax = \lambda Mx$

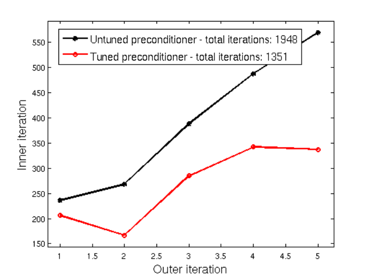


Figure: Rayleigh Quotient shift and decreasing tolerance

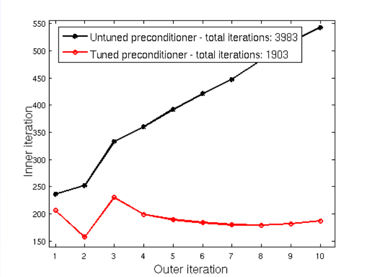


Figure: Fixed shift and decreasing tolerance

Conclusions

Savings of 30% for variable shift: over 50% for fixed shift

Outline

Inexact subspace iteration

- Repeated solve of p-dimensional block system

$$AY = X^{(i)},$$

which is preconditioned as

$$A\mathbb{P}_i\tilde{Y} = X^{(i)}$$

- The tuned preconditioner, $\mathbb{P}_i$ is a rank p update:

$$\mathbb{P}_i = P + (A - P)X^{(i)}{X^{(i)}}^H$$

Numerical Example

- matrix market library `qc2534`
- complex symmetric (non-Hermitian)
- $n = 2534$, $nz = 463360$
- ILU preconditioner
- subspace dimension 16
- seek first 10 eigenvalues

Preconditioned GMRES

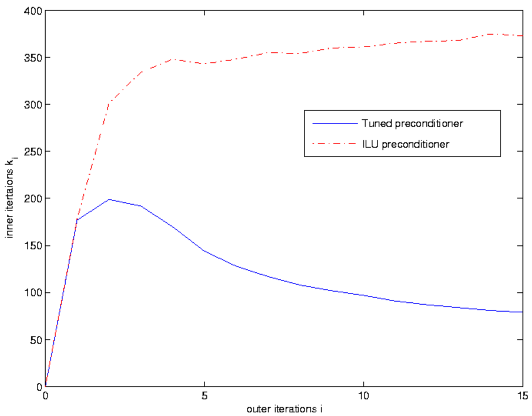


Figure: Inner iterations vs outer iterations

Preconditioned GMRES

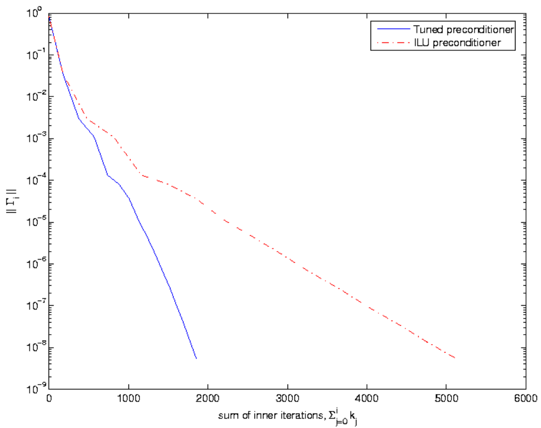


Figure: Residual norms vs total number of iterations

Outline

RQI and J-D: Exact solves ($x^{(i)} \rightarrow x$)

Rayleigh quotient iteration

At each iteration a system of the form

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

has to be solved.

Jacobi-Davidson method

At each iteration a system of the form

$$(I - xx^H)(A - \rho(x)I)(I - xx^H)s = -r$$

has to be solved, where $r = (A - \rho(x)I)x$ is the eigenvalue residual and $s \perp x$.

Exact solves

Sleijpen and van der Vorst (1996):

$$y = \alpha(x + s)$$

for some constant α

RQI and J-D: Inexact solves

Rayleigh quotient iteration

At each iteration a system of the form

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

has to be solved.

Jacobi-Davidson method

At each iteration a system of the form

$$(I - xx^H)(A - \rho(x)I)(I - xx^H)s = -r$$

has to be solved, where $r = (A - \rho(x)I)x$ is the eigenvalue residual and $s \perp x$.

Galerkin-Krylov Solver

- Simoncini and Eldén (2002):

$$y_{k+1} = \beta(x + s_k)$$

for some constant β if both systems are solved using a Galerkin-Krylov subspace method

RQI and J-D: Preconditioned Solves

Preconditioning for RQ iteration

At each iteration a system of the form

$$(A - \rho(x)I)P^{-1}\tilde{y} = x,$$

(with $y = P^{-1}\tilde{y}$) has to be solved.

Preconditioning for JD method

At each iteration a system of the form

$$(I - xx^H)(A - \rho(x)I)(I - xx^H)\tilde{P}^{\dagger}\tilde{s} = -r$$

(with $s = \tilde{P}^{\dagger}\tilde{s}$) has to be solved. Note the restricted preconditioner

$$\tilde{P} := (I - xx^H)P(I - xx^H).$$

Equivalence does not hold!

Example: `sherman5.mtx`

fixed shift; (preconditioned) FOM as inner solver

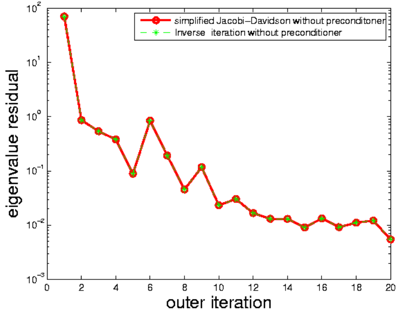


Figure: Convergence history of the eigenvalue residuals; no preconditioner

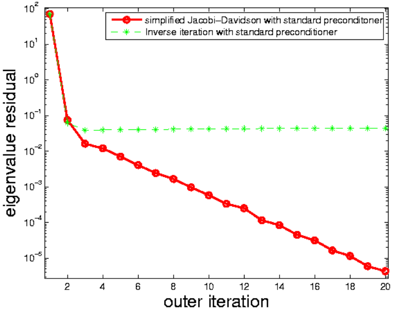


Figure: Convergence history of the eigenvalue residuals; standard preconditioner

Tuned RQI $\equiv$ preconditioned JD

Tuning condition:

$$\mathbb{P}x = x$$

- Implement tuning condition by:

$$\mathbb{P} = P + (I - P)xx^H$$

- Rethink as:

$$\mathbb{P} = xx^H + P(I - xx^H)$$

Equivalence for inexact solves

Theorem

Let both

$$(A-\rho(x)I)\mathbb{P}^{-1}\tilde{y} = x, \quad y = \mathbb{P}^{-1}\tilde{y}$$

and

$$(I-xx^H)(A-\rho(x)I)(I-xx^H)\tilde{P}^{\dagger}\tilde{s} = -r, \quad s = \tilde{P}^{\dagger}\tilde{s}$$

be solved with the same Galerkin-Krylov method. Then

$$y_{k+1}^{RQ} = \gamma(x + s_k^{JD}).$$

Proof.

Based on Simoncini and Eldén (2002). $\square$

Example: `sherman5.mtx`

fixed shift; (preconditioned) FOM as inner solver

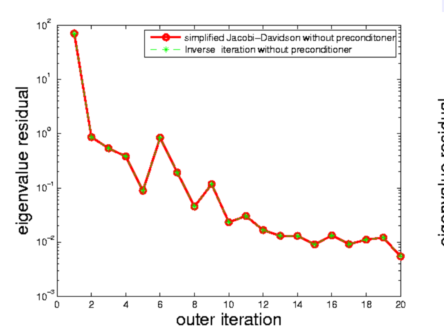


Figure: Convergence history of the eigenvalue residuals; no preconditioner

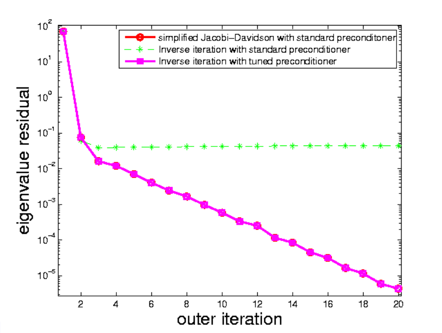


Figure: standard preconditioner for JD, tuned preconditioner for II

Outline

1. Introduction
2. Preconditioned GMRES for Inverse Power Method
3. Inexact Subspace iteration
4. Preconditioned Rayleigh Quotient Iteration and Jacobi-Davidson
5. Conclusions/Further Work

Conclusions

- When using Krylov solvers for shifted systems $(A - \sigma I)y = x^{(i)}$ in eigenvalue computations then one should “tune” the preconditioner so that the iteration matrix has a “good relationship” with the right hand side,
- For any preconditioner “tuning” is achieved by a small rank change,
- Plenty of unanswered questions arise from PDE eigenvalue problems