Università di Bologna

The Nullspace–free eigenvalue problem and the inexact Shift–and–invert Lanczos method

V. Simoncini

Dipartimento di Matematica, Università di Bologna
and CIRSA, Ravenna, Italy
valeria@dm.unibo.it

The generalized eigenvalue problem

Given

$$Ax = \lambda Mx$$

A sym. pos. semidef.

M sym. positive def.

C sparse basis for large null space of A

Find

$$\min_{\lambda_i \neq 0} \lambda_i$$

Difficulty: zero eigenvalues pollute approximation

Equivalent formulation for smallest eigenvalue

$$\min_{\substack{C^T M x = 0 \\ 0 \neq x \in \mathbb{R}^n}} \frac{x^T A x}{x^T M x}$$

$C^T M x = 0$ constraint $\quad \Rightarrow \quad$ Nullspace free eigenvectors

Outline

- General Spectral Transformation
- Inexact Shift-and-Invert Lanczos Method
- Inexactness vs. Constraint
- Alternative Problem Formulations
  - Augmented Formulation
  - Modified Formulation
- A numerical example
- Computational enhancements

General Spectral Transformation

Original Problem: Solve

$$Ax = \lambda Mx$$

Transformed Problem: Fix $\sigma \in \mathbb{R}$ and rewrite as

$$(A - \sigma M)^{-1} Mx = \eta x \qquad \eta = (\lambda - \sigma)^{-1}$$

σ close to eigenvalues of interest $\Rightarrow$ $\quad \eta$ large $\qquad \lambda = \sigma + \frac{1}{\eta}$

- Fast convergence of Lanczos method is expected

Other methods: Jacobi-Davidson, Subspace iteration, Preconditioned Inverse Iteration, etc.

General Shift–and–Invert Lanczos (SI(σ) Lanczos)

Given $v_1 \in \mathbb{R}$,

for $j = 1, 2, \ldots$

$$\begin{aligned} \tilde{v} &= (A - \sigma M)^{-1} M v_j \\ v_{j+1} t_{j+1,j} &= \tilde{v} - V_j T_{:,j} \qquad T_{:,j} = V_j^T M \tilde{v} \end{aligned}$$

$V_j = [v_1, \ldots, v_j]$

Yielding

$$(A - \sigma M)^{-1} M V_j = V_j T_j + [0, \ldots, 0, r_{j+1}]$$

$$(A - \sigma M)^{-1} M V_j = V_j T_j + [0, \ldots, 0, r_{j+1}]$$

If $\boxed{T_j s_j^{(i)} = \eta_j^{(i)} s_j^{(i)}}$ $\quad i = 1, \ldots, j$ then

$$\left(\sigma + \frac{1}{\eta_j^{(i)}},\ V_j s_j^{(i)} \right), \qquad i = 1, \ldots, j$$

are approximate eigenpairs of $Ax = \lambda M x$

Warning: In our setting A is singular. If σ is close to zero, then many zero eigenvalues may be detected

Take v_1 s.t. $C^T M v_1 = 0 \qquad \Rightarrow \qquad C^T M V_j = 0 \quad \forall j$

(exact arithmetic)

General Inexact Shift–and–Invert Lanczos

Take v_1 s.t. $C^T M v_1 = 0$,

for $j = 1, 2, \ldots$

$w \leftarrow M v_j$

$\hat{v}$ approx. solves $(A - \sigma M)\hat{v} = w$

M-orthogonalize $\hat{v}$ w.r. to $\{v_1, \ldots, v_j\} \rightarrow v_{j+1}$

$V_j = [v_1, v_2, \ldots, v_j]$

$$C^T M V_j \stackrel{?}{=} 0$$

⋆ Krylov subspace inner solvers

Natural fixes

[1] Null space is available. Explicitly deflate (purge) null space

Starting vector $v_1 \perp_M \mathcal{N}(A) \Rightarrow \mathrm{span}(V_j) \perp_M \mathcal{N}(A)$

for $j = 1, 2, \ldots$

$w \leftarrow M v_j$

$\hat{v}$ approx. solves $(A - \sigma M)\hat{v} = w$

$\hat{v} \leftarrow (I - \pi)\hat{v}$

M-orthogonalize $\hat{v}$ w.r. to $\{v_1, \ldots, v_j\} \rightarrow v_{j+1}$

π projection onto $\mathcal{N}(A)$ $\quad \pi = C(C^T M C)^{-1} C^T M$

see e.g. Golub, Zhang, Zha 2000

[2] Use $\sigma = 0$ and generate approximation space in Range(A)

$$A^{\dagger} M V_j = V_j T_j + r_{j+1} e_j^T$$

At each iteration i, solve consistent system

$$Ay = Mv_i$$

Arbenz, Drmac 2000

Preconditioning techniques for singular A: Notay 1989-1990, Hiptmair, Neymeyr 2001, ...

Enforcing the constraint

Given v s.t. $C^T M v = 0$, approximately solve

$$(A - \sigma M)x = Mv \qquad (1)$$

with the constraint $\quad C^T M x = 0$

This problem is equivalent to

$$\begin{pmatrix} A - \sigma M & MC \\ (MC)^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} Mv \\ 0 \end{pmatrix} \Leftrightarrow (\mathcal{A} - \sigma \mathcal{M})z = \mathcal{M}b \quad (2)$$

Saddle–point linear system

Exploit work on preconditioning

$$(\mathcal{A} - \sigma \mathcal{M})P^{-1}\hat{z} = \mathcal{M}b \qquad (*)$$

$\hat{z}_m$ approx. solution to (*) $\Rightarrow \quad z_m = P^{-1}\hat{z}_m$ approx. solution to (2)

Structured preconditioning

Given the linear system

$$\begin{pmatrix} K & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}$$

Effective preconditioners include

$$\mathcal{P} = \begin{pmatrix} K_1 & 0 \\ 0 & S \end{pmatrix} \qquad \mathcal{Q} = \begin{pmatrix} K_1 & B \\ B^T & 0 \end{pmatrix} \qquad \mathcal{R} = \begin{pmatrix} K_1 & B \\ 0 & S \end{pmatrix}$$

$S \approx B^T K^{-1} B$

Large bibliography. See Benzi, Golub, Liesen, Acta Numerica 2005

Definite preconditioning

$$(A - \sigma M)x = Mv \qquad (1)$$

If

$$\begin{pmatrix} A - \sigma M & MC \\ (MC)^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} Mv \\ 0 \end{pmatrix} \qquad (2)$$

is preconditioned by

$$\mathcal{P} = \begin{bmatrix} K_1 & 0 \\ 0 & (MC)^T K_1^{-1} (MC) \end{bmatrix}, \qquad K_1 = A_1 - \tau M \quad A_1 C = 0$$

then MINRES soln of (2) is given by $z_m = \begin{bmatrix} x_m \\ 0 \end{bmatrix}$, where x_m is MINRES soln. of (1) preconditioned by K_1

Indefinite preconditioning

$$(A - \sigma M)x = Mv \qquad (1)$$

If

$$\begin{pmatrix} A - \sigma M & MC \\ (MC)^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} Mv \\ 0 \end{pmatrix} \qquad (2)$$

is preconditioned by

$$\mathcal{Q} = \begin{bmatrix} K_1 & MC \\ (MC)^T & 0 \end{bmatrix}, \qquad K_1 = A_1 - \tau M \quad A_1 C = 0$$

then GMRES soln of (2) is given by $z_m = \begin{bmatrix} x_m \\ 0 \end{bmatrix}$, where x_m is GMRES soln. of (1) preconditioned by K_1

Important remark

Solve Preconditioned system

$$(A - \sigma M)K_1^{-1}\hat{x} = Mv \qquad K_1 = A_1 - \tau M$$

so that

$$x_m = K_1^{-1}\hat{x}_m$$

Then it holds

$$\boxed{C^T M x_m = 0}$$

Alternative Problem Formulations

$$Ax = \lambda Mx \qquad C^T Mx = 0$$

[1] Augmented FE formulation (Kikuchi, 1987)

$$\begin{pmatrix} A & MC \\ (MC)^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \tilde{\lambda} \begin{pmatrix} M & 0 \\ 0^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$\min\{\tilde{\lambda}_i\} = \min_{\lambda_i \neq 0}\{\lambda_i\}$$

Computational aspects in Arbenz Geus (1999), Arbenz Geus Adam (2001)

[2] Modified FE formulation (Bespalov, 1988)

Given a sym. nonsingular $H \in \mathbb{R}^{n_c \times n_c}$, solve

$$(A + MCH^{-1}C^T M)x = \eta M x$$

for suitable H, $\min_i\{\eta_i\} = \min_{\lambda_i \neq 0}\{\lambda_i\}$

Augmented formulation

$$\underbrace{\begin{pmatrix} A & MC \\ (MC)^T & 0 \end{pmatrix}}_{\mathcal{A}} \begin{pmatrix} x \\ y \end{pmatrix} = \widetilde{\lambda} \underbrace{\begin{pmatrix} M & 0 \\ 0^T & 0 \end{pmatrix}}_{\mathcal{M}} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$\mathcal{A}z = \widetilde{\lambda}\mathcal{M}z$$

Spectral transformation:

$$(\mathcal{A} - \sigma\mathcal{M})^{-1}\mathcal{M}z = \eta z \qquad \eta = \frac{1}{\widetilde{\lambda} - \sigma}$$

$\Rightarrow$ Apply inexact SI(σ) Lanczos

Augmented formulation

At each iteration, solve Saddle-point linear system

$$\begin{pmatrix} A - \sigma M & MC \\ (MC)^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} Mv \\ 0 \end{pmatrix}$$

Natural inner preconditioners

$$\mathcal{P} = \begin{bmatrix} K_1 & 0 \\ 0 & (MC)^T K_1^{-1} (MC) \end{bmatrix}, \qquad \begin{array}{l} K_1 = A_1 - \tau M \\ A_1 C = 0 \end{array}$$

$$\mathcal{Q} = \begin{bmatrix} K_1 & MC \\ (MC)^T & 0 \end{bmatrix}$$

but

Inexact SI(σ)-Lanczos applied to augmented formulation with inner preconditioner $\mathcal{P}$ or $\mathcal{Q}$

generates the same approximate eigenpairs as

Inexact SI(σ)-Lanczos applied to

$$Ax = \lambda Mx$$

with inner preconditioner K_1

The modified formulation

Original problem

$$Ax = \lambda Mx$$

Given a sym. nonsingular $H \in \mathbb{R}^{n_c \times n_c}$, solve

$$(A + MCH^{-1}C^T M)x = \eta Mx$$

- $\lambda \neq 0 \quad \Rightarrow \exists \mu$ s.t. $\mu = \lambda$
- $\lambda = 0 \quad \Rightarrow \mu$ eigenvalue of $(C^T MC, H)$

Remark. No practical (numerical) advantages over solving $Ax = \lambda Mx$

A numerical example: Electromagnetic cavity resonator

Variational formulation for the 2D computational model

Find $w_h \in \mathbb{R}$ *s.t.* $\exists 0 \not\equiv \underline{u}_h \in \Sigma_h \subset \Sigma$: $\quad (rot(v_1, v_2) = (v_2)_x - (v_1)_y)$

$$(rot\ \underline{u}_h, rot\ \underline{v}_h) = \omega_h^2(\underline{u}_h, \underline{v}_h) \qquad \forall \underline{v}_h \in \Sigma_h,$$

$\Sigma = \{\underline{v} \in [L^2(\Omega)]^2 : \text{rot } \underline{v} \in L^2(\Omega), \underline{v} \cdot \underline{t} = 0 \text{ on } \partial\Omega\}$

$\underline{t}$ counterclockwise oriented tangent versor to the boundary

$\Omega =]0, \pi[^2 \qquad \Rightarrow \omega^2 = 1, 1, 2, 4, 4, 5, 5, 8, 9, 9, 10, 10, \ldots$

****

FE method using edge elements

Problem dimension $= 3229$ Number of zero eigenvalues $= 1036$

Boffi, Fernandes, Gastaldi, Perugia, SIAM J. Numer. Anal., v.36 (1999)

Norm of residual $$\frac{\|Ax_j^{(1)} - \lambda_j^{(1)} M x_j^{(1)}\|}{\lambda_j^{(1)}}$$

m	$(A-\sigma M)^{-1}Mx = \eta x$ K_1	$(\mathcal{A}-\sigma\mathcal{M})^{-1}\mathcal{M}z = \eta z$ $\mathcal{P}$	$(\mathcal{A}-\sigma\mathcal{M})^{-1}\mathcal{M}z = \eta z$ $\mathcal{Q}$
4	0.02426393067395	0.02426393067727	0.02426393066981
6	0.02898748221567	0.02898746782699	0.02898748572682
8	0.01156203523797	0.01156203705189	0.01156203467534
10	0.00000041284501	0.00000041284501	0.00000041283893
12	0.00000000158821	0.00000000158844	0.00000000158891
14	0.00000000158802	0.00000000158827	0.00000000158882

First wrap-up

Original Formulation (n) $\Leftrightarrow$ Augmented Formulation $(n + n_c)$

$\Downarrow$

Same Approximation iterates

Computational considerations: Inexact Lanczos

Starting vector $v_1 \perp_M \mathcal{N}(A) \Rightarrow \text{span}(V_j) \perp_M \mathcal{N}(A)$

for $j = 1, 2, \ldots$

$w \leftarrow M v_j$

$\hat{v}$ approx. solves $(A - \sigma M)\hat{v} = w$

$\hat{v} \leftarrow (I - \pi)\hat{v}$

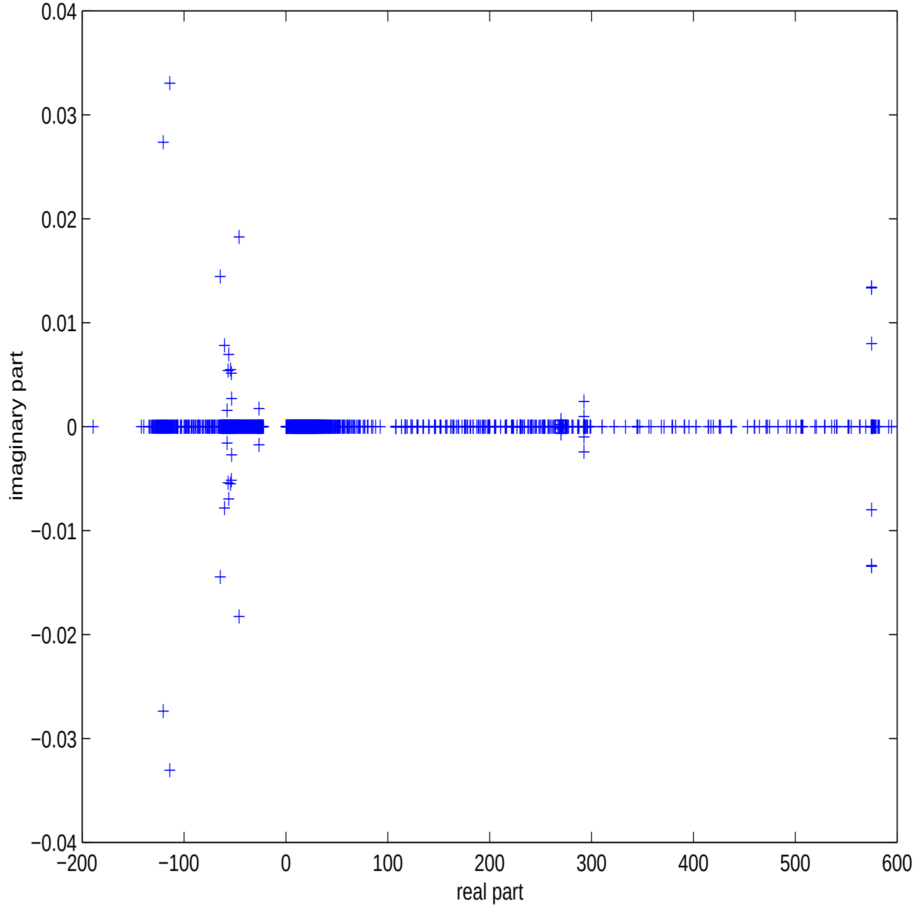
M-orthogonalize $\hat{v}$ w.r. to $\{v_1, \ldots, v_j\} \rightarrow v_{j+1}$

Compute approximate $\eta_j^{(1)}, \ldots, \eta_j^{(j)}$ to $\eta_1, \eta_2, \ldots$

Check convergence with Lanczos residuals

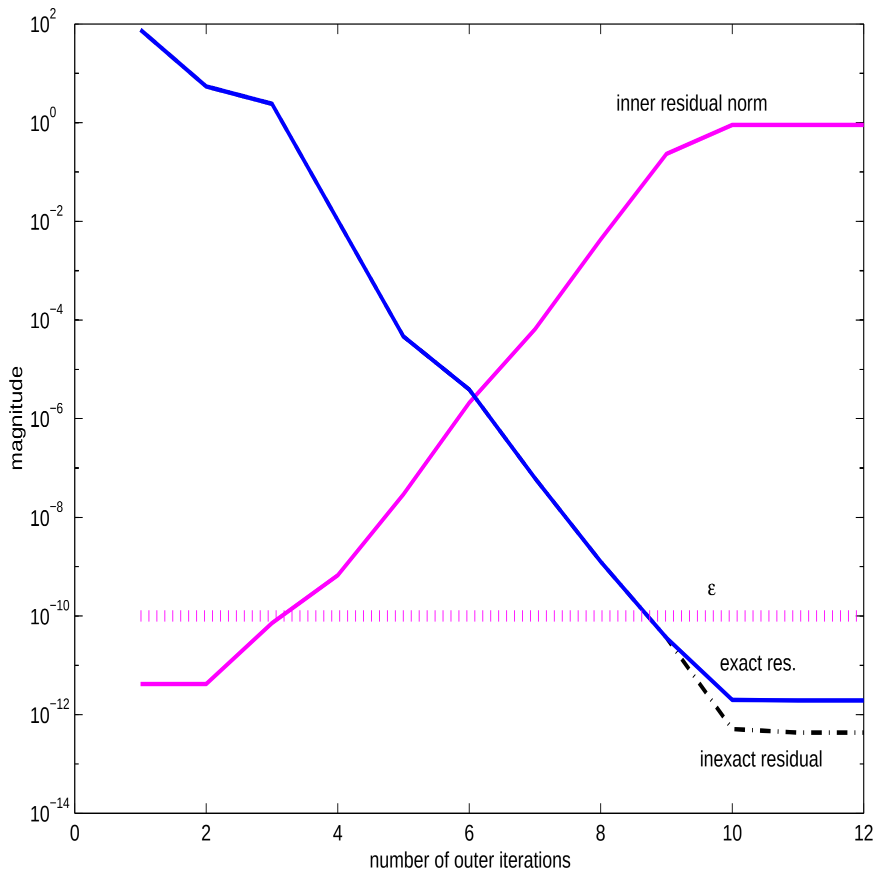
Question: How accurate should be the solution of $(A - \sigma M)\hat{v} = w$?

Answer: You can decrease the accuracy as the Lanczos method converges.



SHERMAN5 MatrixMarket. Approx. $\min|\lambda|$ with "inverted" Arnoldi

Example 2



SHERMAN5 MatrixMarket. Approx. $\min|\lambda|$ with "inverted" Arnoldi

At iteration j, solve $(A - \sigma M)\hat{v} = w$

$\eta_{j-1}^{(1)}$ approx eig. after $j-1$ Lanczos its.

r_{j-1} associated computed eigenvalue residual

$f_j = (A - \sigma M)\hat{v}_k - w$ residual after k inner iterations (linear system)

Stopping criterion for inner system solver

Empirically (Bouras and Frayssé, t.r. '00):

$$\|f_j\| \le \frac{10^{-\alpha}}{|\eta_{j-1}^{(1)}| \, \|r_{j-1}\|}\varepsilon, \qquad \alpha = 0, 1, 2$$

At iteration j, solve $(A - \sigma M)\hat{v} = w$

$\eta_{j-1}^{(1)}$ approx eig. after $j-1$ Lanczos its.

r_{j-1} associated computed eigenvalue residual

$f_j = (A - \sigma M)\hat{v}_k - w$ residual after k inner iterations (linear system)

Stopping criterion for inner system solver

New computable bound (Simoncini, t.r. '04):

$$\|f_j\| \le \frac{\min\{\|A - \sigma M\|, \delta^{(j-1)}\}}{2m|\eta_{j-1}^{(1)}| \, \|r_{j-1}\|}\varepsilon$$

where

$$\delta^{(j-1)} := \min_{\eta_{j-1}^{(k)} \in \Lambda(H_{j-1}) \setminus \{\eta_{j-1}^{(1)}\}} |\eta_{j-1}^{(k)} - \eta_{j-1}^{(1)}|$$

The whole story

If, for any $k = 1, \ldots, m$, $\quad \|r_k\| \leq \dfrac{\delta_{m,k-1}^2}{4\|s_m\|}$ $\quad$ and

$$\|f_k\| \leq \frac{\delta_{m,k-1}}{2m|\eta_{j-1}^{(1)}|\|r_{k-1}\|}\varepsilon$$

then after m iterations, $(\eta_m^{(1)}, s)$ satisfies

$$\|((A - \sigma M)^{-1} M V_m s - \eta_m^{(1)} V_m s) - r_m\| \leq \varepsilon$$

[1] P. Arbenz and R. Geus, *A comparison of solvers for large eigenvalue problems occuring in the design of resonant cavities*, Numer. Linear Alg. with Appl., 6 (1999), pp. 3–16.

[2] P. Arbenz, R. Geus, and S. Adam, *Solving Maxwell eigenvalue problems for accelerating cavities*, Physical Review special Topics - Accelerators and Beams, 4 (2001), pp. 1–10.

[3] A. N. Bespalov, *Finite element method for the eigenmode problem of a rf cavity resonator*, Sov. J. Numer. Anal. Math. Mod., 3 (1988), pp. 163–178.

[4] F. Kikuchi, *Mixed and penalty formulations for finite element analysis of an eigenvalue problem in electromagnetism*, Computer Methods in Appl. Mech. Eng., 64 (1987), pp. 509–521.

[5] V. Simoncini, *Algebraic formulations for the solution of the nullspace–free eigenvalue problem using the inexact Shift–and-Invert Lanczos method*, Numerical Linear Algebra w/Appl, 2003.

[6] V. Simoncini, *Variable accuracy of matrix-vector products in projection methods for eigencomputation.* March 2004. To appear in SIAM J. Numerical Analysis.

http://www.dm.unibo.it/~simoncin **e-mail:** valeria@dm.unibo.it