Solution of a constraint generalised eigenvalue problem using the inexact Shift-and-Invert Lanczos method
on a paper by V. Simoncini

Melina Freitag

Department of Mathematical Sciences
University of Bath

Numerical Analysis Seminar
9th December 2005

Outline

Problem

- Eigenproblem for $A \in \mathbb{C}^{n,n}$, $A = A^T$:

$$Ax = \lambda x.$$

- let the eigenvalues be

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$$

- associated eigenvectors $x_1, x_2, \ldots, x_n$
- A is large and sparse, need iterative methods.

Idea behind Lanczos

- keep iterates from Power method $v, Av, \ldots, A^{k-1}v$ which form a Krylov subspace associated with A and v

$$\mathcal{K}_j(A, v) = \mathrm{span}\{v, Av, \ldots, A^{j-1}v\}.$$

- $v, Av, \ldots, A^{k-1}v$ are usually ill-conditioned
- orthogonalise the vectors $v, Av, \ldots, A^{k-1}v$ in the Krylov space using a modified Gram-Schmidt process

Idea behind Lanczos

- keep iterates from Power method $v, Av, \ldots, A^{k-1}v$ which form a Krylov subspace associated with A and v

$$\mathcal{K}_j(A, v) = \text{span}\{v, Av, \ldots, A^{j-1}v\}.$$

- $v, Av, \ldots, A^{k-1}v$ are usually ill-conditioned
- orthogonalise the vectors $v, Av, \ldots, A^{k-1}v$ in the Krylov space using a modified Gram-Schmidt process

Lanczos algorithm

- choose initial vector v and normalise $v_1 = \dfrac{v}{\|v\|_2}$
- On subsequent steps $k = 1, 2, \dots$ take

$$\tilde{v}_{k+1} = Av_k - \sum_{j=1}^{k} v_j t_{jk}$$

  where t_{jk} is the Gram-Schmidt coefficient $t_{jk} =< Av_k, v_j >$.
- normalise

$$v_{k+1} = \frac{\tilde{v}_{k+1}}{t_{k+1,k}} \quad \text{where} \quad t_{k+1,k} = \|\tilde{q}_{k+1}\|_2$$

Matrix formulation and calculation of eigenvalues

Lanczos in matrix form

The Lanczos process can be written in the form

$$AV_m = V_m T_m + v_{m+1}\beta_m e_m^T \quad \text{where} \quad T_m = \begin{bmatrix} \alpha_1 & \beta_1 & & \\ \beta_1 & \alpha_2 & \ddots & \\ & \ddots & \ddots & \beta_{m-1} \\ & & \beta_{m-1} & \alpha_m \end{bmatrix}$$

Theorem

Let V_m, T_m and β_m generated by the Lanczos process and

$$T_m s = \mu s, \quad \|s\|_2 = 1.$$

Let $y = V_m s \in \mathbb{C}^n$, then

Matrix formulation and calculation of eigenvalues

Lanczos in matrix form

The Lanczos process can be written in the form

$$AV_m = V_m T_m + v_{m+1}\beta_m e_m^T \quad \text{where} \quad T_m = \begin{bmatrix} \alpha_1 & \beta_1 & & \\ \beta_1 & \alpha_2 & \ddots & \\ & \ddots & \ddots & \beta_{m-1} \\ & & \beta_{m-1} & \alpha_m \end{bmatrix}$$

Theorem

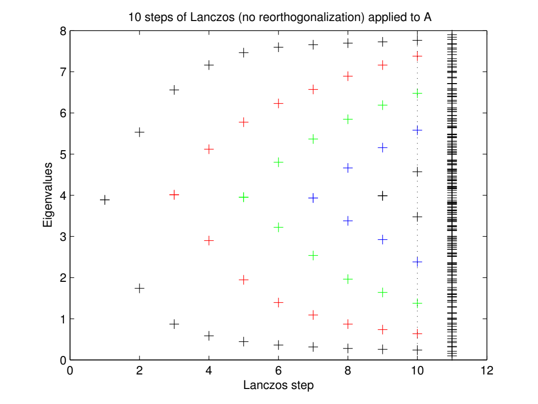
Let V_m, T_m and β_m generated by the Lanczos process and

$$T_m s = \mu s, \quad \|s\|_2 = 1.$$

Let $y = V_m s \in \mathbb{C}^n$, then

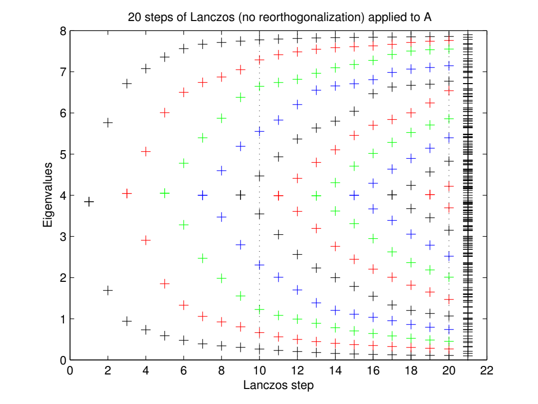
An example

first 10 Lanczos steps



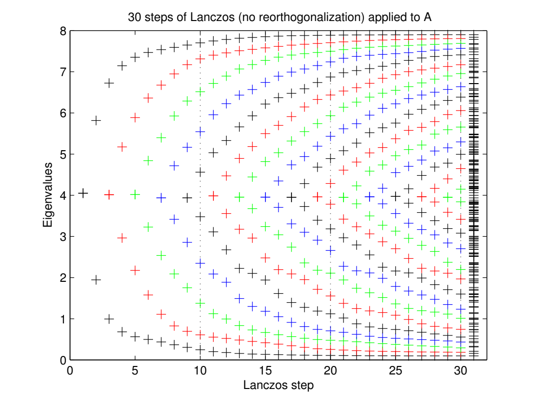
An example

first 20 Lanczos steps



An example

first 30 Lanczos steps



The constraint eigenvalue problem

- Computation of the smallest non-zero eigenvalues and corresponding eigenvectors of
$$Ax = \lambda Mx$$
where $M = M^T$ positive definite and $A = A^T$ positive semidefinite.
- assume sparse basis C for null-space of A is available
- dimension of the null-space is high compared with the problem dimension
- constraint in terms of the null-space orthogonality, for smallest non-zero eigenvalue:
$$\min_{\substack{C^T Mx=0 \\ 0\neq x\in\mathbb{R}}} \frac{x^T Ax}{x^T Mx}$$

The constraint eigenvalue problem

- Computation of the smallest non-zero eigenvalues and corresponding eigenvectors of
$$Ax = \lambda Mx$$
where $M = M^T$ positive definite and $A = A^T$ positive semidefinite.
- assume sparse basis C for null-space of A is available
- dimension of the null-space is high compared with the problem dimension
- constraint in terms of the null-space orthogonality, for smallest non-zero eigenvalue:
$$\min_{\substack{C^T Mx=0 \\ 0\neq x\in\mathbb{R}}} \frac{x^T Ax}{x^T Mx}$$

UNIVERSITY OF BATH

The constraint eigenvalue problem

- Computation of the smallest non-zero eigenvalues and corresponding eigenvectors of

$$Ax = \lambda Mx$$

where $M = M^T$ positive definite and $A = A^T$ positive semidefinite.

- assume sparse basis C for null-space of A is available
- dimension of the null-space is high compared with the problem dimension
- constraint in terms of the null-space orthogonality, for smallest non-zero eigenvalue:

$$\min_{\substack{C^T Mx = 0 \\ 0 \neq x \in \mathbb{R}}} \frac{x^T Ax}{x^T Mx}$$

Application areas

Electromagnetic cavity resonator

$$\begin{aligned} \mathrm{curl}(\mu^{-1}\mathrm{curl}\mathbf{u}) = \omega^2\mathbf{u} \quad & \text{in} \quad \Omega \\ \mathrm{div}(\varepsilon\mathbf{u}) = 0 \quad & \text{in} \quad \Omega \\ \mathbf{u}\times\mathbf{n} = \mathbf{0} \quad & \text{on} \quad \partial\Omega \end{aligned}$$

where $\mathbf{u}$ is the electric field, $\mathbf{n}$ denotes the outward normal vector, μ the magnetic permeability, ε the electric permittivity.

Network problems

$$Ax = \lambda x, \quad \text{with} \quad Ac = 0$$

where $A = A^T$ SPD, $M = I$ and the eigenpair $(0, c)$ is known, looking for second smallest eigenvalue λ_2 with the constraint $c^T x = 0$.

Application areas

Electromagnetic cavity resonator

$$\begin{aligned} \text{curl}(\mu^{-1}\text{curl}\mathbf{u}) &= \omega^2\mathbf{u} && \text{in} && \Omega \\ \text{div}(\varepsilon\mathbf{u}) &= 0 && \text{in} && \Omega \\ \mathbf{u}\times\mathbf{n} &= \mathbf{0} && \text{on} && \partial\Omega \end{aligned}$$

where $\mathbf{u}$ is the electric field, $\mathbf{n}$ denotes the outward normal vector, μ the magnetic permeability, ε the electric permittivity.

Network problems

$$Ax = \lambda x, \quad \text{with} \quad Ac = 0$$

where $A = A^T$ SPD, $M = I$ and the eigenpair $(0, c)$ is known, looking for second smallest eigenvalue λ_2 with the constraint $c^T x = 0$.

Simplify the problem

- Consider smallest non-zero eigenvalues and corresponding eigenvectors of
$$Ax = \lambda x$$
where $M = I$ positive definite and $A = A^T$ positive semidefinite.
- the null-space is one-dimensional
$$Ac = 0$$
- Constraint in terms of the null-space orthogonality, for smallest non-zero eigenvalue:
$$\min_{\substack{c^T x = 0 \\ 0 \neq x \in \mathbb{R}}} \frac{x^T A x}{x^T x}$$

Different formulations of the problem

- $Ax = \lambda x, \quad c^T x = 0$
- Shifting the null eigenvalue

$$(A + cH^{-1}c^T)x = \eta x,$$

$H = \frac{1}{\gamma} c^T c$ shifts zero eigenvalue to γ.
Smallest eigenvalues coincide.

- Enforce constraint with augmented system

$$\begin{bmatrix} A & c \\ c^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Smallest eigenvalues coincide.

Different formulations of the problem

- $Ax = \lambda x, \quad c^T x = 0$
- Shifting the null eigenvalue

$$(A + cH^{-1}c^T)x = \eta x,$$

$H = \frac{1}{\gamma} c^T c$ shifts zero eigenvalue to γ.
Smallest eigenvalues coincide.

- Enforce constraint with augmented system

$$\begin{bmatrix} A & c \\ c^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Smallest eigenvalues coincide.

Different formulations of the problem

- $Ax = \lambda x, \quad c^T x = 0$
- Shifting the null eigenvalue

$$(A + cH^{-1}c^T)x = \eta x,$$

$H = \frac{1}{\gamma} c^T c$ shifts zero eigenvalue to γ.
Smallest eigenvalues coincide.

- Enforce constraint with augmented system

$$\begin{bmatrix} A & c \\ c^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Smallest eigenvalues coincide.

Outline

Shift-and-Invert Lanczos for generalised eigenproblem

- Consider
$$\mathcal{A}x = \lambda \mathcal{I} x, \quad \mathcal{A} = \mathcal{A}^T,$$
- apply Lanczos to spectrally transformed problem
$$(\mathcal{A} - \sigma \mathcal{I})^{-1} \mathcal{I} x = \eta x, \quad \eta = (\lambda - \sigma)^{-1}$$
- basic recursion for SI-Lanczos
$$(\mathcal{A} - \sigma \mathcal{I})^{-1} V_j = V_j T_j + v_{j+1} t_{j+1,j} e_j^T,$$
where $V_j = [v_1, \ldots, v_j]$ is an orthogonal basis, T_j tridiagonal with $T_j = V_j^T (\mathcal{A} - \sigma \mathcal{I})^{-1} V_j$
- If $T_j s_j^{(i)} = \eta_j^{(i)} s_j^{(i)}$ we get eigenpairs for $\mathcal{A}$ by $(1/\eta_j^{(i)} + \sigma, V_j s_j^{(i)})$

Shift-and-Invert Lanczos for generalised eigenproblem

- Consider

$$\mathcal{A}x = \lambda \mathcal{I} x, \quad \mathcal{A} = \mathcal{A}^T,$$

- apply Lanczos to spectrally transformed problem

$$(\mathcal{A} - \sigma \mathcal{I})^{-1} \mathcal{I} x = \eta x, \quad \eta = (\lambda - \sigma)^{-1}$$

- basic recursion for SI-Lanczos

$$(\mathcal{A} - \sigma \mathcal{I})^{-1} V_j = V_j T_j + v_{j+1} t_{j+1,j} e_j^T,$$

where $V_j = [v_1, \ldots, v_j]$ is an orthogonal basis, T_j tridiagonal with $T_j = V_j^T (\mathcal{A} - \sigma \mathcal{I})^{-1} V_j$

- If $T_j s_j^{(i)} = \eta_j^{(i)} s_j^{(i)}$ we get eigenpairs for $\mathcal{A}$ by $(1/\eta_j^{(i)} + \sigma, V_j s_j^{(i)})$

Shift-and-Invert Lanczos for generalised eigenproblem

- Consider

$$\mathcal{A}x = \lambda \mathcal{I} x, \quad \mathcal{A} = \mathcal{A}^T,$$

- apply Lanczos to spectrally transformed problem

$$(\mathcal{A} - \sigma \mathcal{I})^{-1} \mathcal{I} x = \eta x, \quad \eta = (\lambda - \sigma)^{-1}$$

- basic recursion for SI-Lanczos

$$(\mathcal{A} - \sigma \mathcal{I})^{-1} V_j = V_j T_j + v_{j+1} t_{j+1,j} e_j^T,$$

  where $V_j = [v_1, \ldots, v_j]$ is an orthogonal basis, T_j tridiagonal with $T_j = V_j^T (\mathcal{A} - \sigma \mathcal{I})^{-1} V_j$

- If $T_j s_j^{(i)} = \eta_j^{(i)} s_j^{(i)}$ we get eigenpairs for $\mathcal{A}$ by $(1/\eta_j^{(i)} + \sigma, V_j s_j^{(i)})$

Shift-and-Invert Lanczos and constraints

- $(A - \sigma I)^{-1} x = \eta x, \quad c^T x = 0.$
- start iteration with v_1 such that $c^T v_1 = 0$ and $v_1^T v_1 = 1$
- $c^T x$ is automatically satisfied by exact eigenvectors
- finite precision arithmetic orthogonality constraint not satisfied
- let $\pi = c(c^T c)^{-1} c^T$, then $I - \pi$ projects onto $\mathbb{R}^n$ orthogonal to the null-space of A
- modify Lanczos algorithm to enforce orthogonality constraint $c^T v_j = 0$:

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

Shift-and-Invert Lanczos and constraints

- $(A-\sigma I)^{-1}x = \eta x, \quad c^T x = 0.$
- start iteration with v_1 such that $c^T v_1 = 0$ and $v_1^T v_1 = 1$
- $c^T x$ is automatically satisfied by exact eigenvectors
- finite precision arithmetic orthogonality constraint not satisfied
- let $\pi = c(c^T c)^{-1} c^T$, then $I - \pi$ projects onto $\mathbb{R}^n$ orthogonal to the null-space of A
- modify Lanczos algorithm to enforce orthogonality constraint $c^T v_j = 0$:

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

Shift-and-Invert Lanczos and constraints

- $(A-\sigma I)^{-1}x = \eta x, \quad c^T x = 0.$
- start iteration with v_1 such that $c^T v_1 = 0$ and $v_1^T v_1 = 1$
- $c^T x$ is automatically satisfied by exact eigenvectors
- finite precision arithmetic orthogonality constraint not satisfied
- let $\pi = c(c^T c)^{-1} c^T$, then $I - \pi$ projects onto $\mathbb{R}^n$ orthogonal to the null-space of A
- modify Lanczos algorithm to enforce orthogonality constraint $c^T v_j = 0$:

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

Inexact Shift-and-Invert Lanczos and constraints

- let z_j be approximate solution to the system

$$(A - \sigma I)z = v_j$$

- Set $Z_j = [z_1, \ldots, z_j]$

$$(A - \sigma I)^{-1} V_j = V_j T_j + v_{j+1} t_{j+1,j} e_j^T,$$

becomes

$$Z_j = V_j \bar{T}_j + v_{j+1} t_{j+1,j} e_j^T, \quad \bar{T}_j = V_j^T Z_j$$

- Problem: $c^T z_j = 0$??? depending on the iterative solver and preconditioning strategy
- enforce the constraint in the outer Lanczos iteration:

$$(I - \pi) Z_j = V_j T_j + v_{j+1} t_{j+1,j} e_j^T, \quad V_j = (I - \pi) V_j$$

- enforce the constraint during the solution of the inner system

Inexact Shift-and-Invert Lanczos and constraints

- let z_j be approximate solution to the system

$$(A - \sigma I)z = v_j$$

- Set $Z_j = [z_1, \ldots, z_j]$

$$(A - \sigma I)^{-1} V_j = V_j T_j + v_{j+1} t_{j+1,j} e_j^T,$$

becomes

$$Z_j = V_j \bar{T}_j + v_{j+1} t_{j+1,j} e_j^T, \quad \bar{T}_j = V_j^T Z_j$$

- Problem: $c^T z_j = 0$??? depending on the iterative solver and preconditioning strategy
- enforce the constraint in the outer Lanczos iteration:

$$(I - \pi)Z_j = V_j T_j + v_{j+1} t_{j+1,j} e_j^T, \quad V_j = (I - \pi)V_j$$

- enforce the constraint during the solution of the inner system

Inexact Shift-and-Invert Lanczos and constraints

- let z_j be approximate solution to the system

$$(A - \sigma I)z = v_j$$

- Set $Z_j = [z_1, \dots, z_j]$

$$(A - \sigma I)^{-1} V_j = V_j T_j + v_{j+1} t_{j+1,j} e_j^T,$$

becomes

$$Z_j = V_j \bar{T}_j + v_{j+1} t_{j+1,j} e_j^T, \quad \bar{T}_j = V_j^T Z_j$$

- Problem: $c^T z_j = 0$??? depending on the iterative solver and preconditioning strategy
- enforce the constraint in the outer Lanczos iteration:

$$(I - \pi) Z_j = V_j T_j + v_{j+1} t_{j+1,j} e_j^T, \quad V_j = (I - \pi) V_j$$

- enforce the constraint during the solution of the inner system

Inexact Shift-and-Invert Lanczos and constraints

- let z_j be approximate solution to the system

$$(A - \sigma I)z = v_j$$

- Set $Z_j = [z_1, \ldots, z_j]$

$$(A - \sigma I)^{-1} V_j = V_j T_j + v_{j+1} t_{j+1,j} e_j^T,$$

becomes

$$Z_j = V_j \bar{T}_j + v_{j+1} t_{j+1,j} e_j^T, \quad \bar{T}_j = V_j^T Z_j$$

- Problem: $c^T z_j = 0$??? depending on the iterative solver and preconditioning strategy
- enforce the constraint in the outer Lanczos iteration:

$$(I - \pi) Z_j = V_j T_j + v_{j+1} t_{j+1,j} e_j^T, \quad V_j = (I - \pi) V_j$$

- enforce the constraint during the solution of the inner system

Inexact Shift-and-Invert Lanczos and constraints

- let z_j be approximate solution to the system

$$(A - \sigma I)z = v_j$$

- Set $Z_j = [z_1, \ldots, z_j]$

$$(A - \sigma I)^{-1} V_j = V_j T_j + v_{j+1} t_{j+1,j} e_j^T,$$

becomes

$$Z_j = V_j \bar{T}_j + v_{j+1} t_{j+1,j} e_j^T, \quad \bar{T}_j = V_j^T Z_j$$

- Problem: $c^T z_j = 0$??? depending on the iterative solver and preconditioning strategy
- enforce the constraint in the outer Lanczos iteration:

$$(I - \pi) Z_j = V_j T_j + v_{j+1} t_{j+1,j} e_j^T, \quad V_j = (I - \pi) V_j$$

- enforce the constraint during the solution of the inner system

On Krylov subspace methods (for solving linear systems)

- want to solve

$$(A-\sigma I)z = v$$

- using right preconditioner P we obtain

$$(A-\sigma I)P^{-1}\hat{z} = v$$

- minimise the residual $v-(A-\sigma I)P^{-1}\hat{z}$ with zero starting guess

$$z^{(m)} = P^{-1}\hat{z}^{m} \quad \text{with} \quad \hat{z}^{(m)} \in \mathcal{K}_m((A-\sigma I)P^{-1}, v)$$

- examples: CG, MINRES, GMRES

On Krylov subspace methods (for solving linear systems)

- want to solve

$$(A-\sigma I)z = v$$

- using right preconditioner P we obtain

$$(A-\sigma I)P^{-1}\hat{z} = v$$

- minimise the residual $v-(A-\sigma I)P^{-1}\hat{z}$ with zero starting guess

$$z^{(m)} = P^{-1}\hat{z}^{m} \quad \text{with} \quad \hat{z}^{(m)} \in \mathcal{K}_m((A-\sigma I)P^{-1}, v)$$

- examples: CG, MINRES, GMRES

Original system - Augmented system

Augmented System

$$\begin{bmatrix} A - \sigma I & c \\ c^T & 0 \end{bmatrix} P^{-1} \hat{z} = \begin{bmatrix} v \\ 0 \end{bmatrix} \Leftrightarrow (\mathcal{A} - \sigma \mathcal{I}) P^{-1} \hat{z} = \mathcal{I} b$$

Vectors generating the subspace $\mathcal{K}_m((\mathcal{A} - \sigma\mathcal{I})P^{-1}, \mathcal{I}b)$

$$((\mathcal{A} - \sigma \mathcal{I})^k P^{-1})^k \mathcal{I} b = \begin{bmatrix} G^k v \\ 0 \end{bmatrix}$$

Minimisation procedure

$$\hat{z}^{(m)} \in \mathcal{K}_m((\mathcal{A} - \sigma\mathcal{I})P^{-1}, \mathcal{I}b) \quad \hat{z}^{(m)} = [\hat{x}^{(m)}; 0]$$

optimal approximate solution of $G\hat{x} = v$ in

Original system - Augmented system

Augmented System

$$\begin{bmatrix} A-\sigma I & c \\ c^T & 0 \end{bmatrix} P^{-1}\hat{z} = \begin{bmatrix} v \\ 0 \end{bmatrix} \Leftrightarrow (\mathcal{A}-\sigma\mathcal{I})P^{-1}\hat{z} = \mathcal{I}b$$

Vectors generating the subspace $\mathcal{K}_m((\mathcal{A}-\sigma\mathcal{I})P^{-1}, \mathcal{I}b)$

$$((\mathcal{A}-\sigma\mathcal{I})^k P^{-1})^k \mathcal{I}b = \begin{bmatrix} G^k v \\ 0 \end{bmatrix}$$

Minimisation procedure

$$\hat{z}^{(m)} \in \mathcal{K}_m((\mathcal{A}-\sigma\mathcal{I})P^{-1}, \mathcal{I}b) \quad \hat{z}^{(m)} = [\hat{x}^{(m)}; 0]$$

optimal approximate solution of $G\hat{x} = v$ in

Original system - Augmented system

Augmented System

$$\begin{bmatrix} A-\sigma I & c \\ c^T & 0 \end{bmatrix} P^{-1}\hat{z} = \begin{bmatrix} v \\ 0 \end{bmatrix} \Leftrightarrow (\mathcal{A}-\sigma\mathcal{I})P^{-1}\hat{z} = \mathcal{I}b$$

Vectors generating the subspace $\mathcal{K}_m((\mathcal{A}-\sigma\mathcal{I})P^{-1}, \mathcal{I}b)$

$$((\mathcal{A}-\sigma\mathcal{I})^k P^{-1})^k \mathcal{I}b = \begin{bmatrix} G^k v \\ 0 \end{bmatrix}$$

Minimisation procedure

$$\hat{z}^{(m)} \in \mathcal{K}_m((\mathcal{A}-\sigma\mathcal{I})P^{-1}, \mathcal{I}b) \quad \hat{z}^{(m)} = [\hat{x}^{(m)}; 0]$$

optimal approximate solution of $G\hat{x} = v$ in

Outline

Original system - Augmented system

- Original system:

$$(A - \sigma I)x = v \quad \text{with} \quad c^T x = 0$$

- Augmented system

$$\begin{bmatrix} A - \sigma I & c \\ c^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} v \\ 0 \end{bmatrix} \Leftrightarrow (\mathcal{A} - \sigma \mathcal{I})z = \mathcal{I}b$$

- show that augmented system is not better than original system
- analyse 2 preconditioning techniques

Original system - Augmented system

- Original system:

$$(A - \sigma I)x = v \quad \text{with} \quad c^T x = 0$$

- Augmented system

$$\begin{bmatrix} A - \sigma I & c \\ c^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} v \\ 0 \end{bmatrix} \Leftrightarrow (\mathcal{A} - \sigma \mathcal{I})z = \mathcal{I}b$$

- show that augmented system is not better than original system
- analyse 2 preconditioning techniques

Original system - Augmented system

- Original system:

$$(A - \sigma I)x = v \quad \text{with} \quad c^T x = 0$$

- Augmented system

$$\begin{bmatrix} A - \sigma I & c \\ c^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} v \\ 0 \end{bmatrix} \Leftrightarrow (\mathcal{A} - \sigma \mathcal{I})z = \mathcal{I}b$$

- show that augmented system is not better than original system
- analyse 2 preconditioning techniques

The preconditioner and its properties

- structured symmetric definite preconditioner

$$P_D = \begin{bmatrix} K_1 & 0 \\ 0 & c^T K_1^{-1} c \end{bmatrix}, \quad K_1 = A_1 - \tau I, \quad \tau \in \mathbb{R}$$

where $K_1 = K_1^T$ nonsingular and $A_1 c = 0$ so $\tau \neq 0$ $(A_1 = 0)$

- We have

$$c^T K_1^{-1} = -\frac{1}{\tau} c^T$$

$$c^T K_1^{-1} c = -\frac{1}{\tau} c^T c \quad \text{symplifies} \quad P_D$$

and with $K = A - \sigma I$

$$c^T (K K_1^{-1})^k = \frac{\sigma}{\tau} c^T \quad \text{for} \quad k \geq 0$$

The preconditioner and its properties

- structured symmetric definite preconditioner

$$P_D = \begin{bmatrix} K_1 & 0 \\ 0 & c^T K_1^{-1} c \end{bmatrix}, \quad K_1 = A_1 - \tau I, \quad \tau \in \mathbb{R}$$

where $K_1 = K_1^T$ nonsingular and $A_1 c = 0$ so $\tau \neq 0$ $(A_1 = 0)$

- We have

$$c^T K_1^{-1} = -\frac{1}{\tau} c^T$$

$$c^T K_1^{-1} c = -\frac{1}{\tau} c^T c \quad \text{symplifies} \quad P_D$$

and with $K = A - \sigma I$

$$c^T (K K_1^{-1})^k = \frac{\sigma}{\tau} c^T \quad \text{for} \quad k \geq 0$$

Equivalence of optimal solutions

Theorem

Let v satisfy $c^T v = 0$. The optimal Krylov subspace solution of the augmented system $z^{(m)}$ with right preconditioner P_D can be written as

$$z^{(m)} = [x^{(m)}; 0],$$

where $x^{(m)}$ is the optimal Krylov subspace solution of the original (non-augmented) system with preconditioner K_1.

Proof Idea

$$((\mathcal{A} - \sigma\mathcal{I})P_D^{-1})^k \begin{bmatrix} v \\ 0 \end{bmatrix} = \begin{bmatrix} (KK_1^{-1})^k v \\ 0 \end{bmatrix}$$

If $\hat{z}^{(m)} \in \mathcal{K}_m((\mathcal{A} - \sigma\mathcal{I})P_D^{-1}, \mathcal{I}b)$ then $\hat{x}^{(m)} \in \mathcal{K}_m((KK_1^{-1}, v)$ both optimal approximate solutions.

Equivalence of optimal solutions

Theorem

Let v satisfy $c^T v = 0$. The optimal Krylov subspace solution of the augmented system $z^{(m)}$ with right preconditioner P_D can be written as

$$z^{(m)} = [x^{(m)}; 0],$$

where $x^{(m)}$ is the optimal Krylov subspace solution of the original (non-augmented) system with preconditioner K_1.

Proof Idea

$$((\mathcal{A} - \sigma\mathcal{I})P_D^{-1})^k \begin{bmatrix} v \\ 0 \end{bmatrix} = \begin{bmatrix} (KK_1^{-1})^k v \\ 0 \end{bmatrix}$$

If $\hat{z}^{(m)} \in \mathcal{K}_m((\mathcal{A} - \sigma\mathcal{I})P_D^{-1}, \mathcal{I}b)$ then $\hat{x}^{(m)} \in \mathcal{K}_m((KK_1^{-1}, v)$ both optimal approximate solutions.

Orthogonality constraint satisfied

Remark

The solution $z^{(m)} = [x^{(m)}; 0]$, satisfies $c^T x^{(m)} = 0$.

Proof

- since $\hat{x}^{(m)}$ is optimal approximate solution in $\mathcal{K}_m((KK_1^{-1}, v)$

$$\hat{x}^{(m)} = \phi_{m-1}(KK_1^{-1})v$$

and $c^T(KK_1^{-1})^k = \frac{\sigma}{\tau}c^T$ we have $c^T\hat{x}^{(m)} = 0$.

- Then

$$c^T x^{(m)} = -\frac{1}{\tau}c^T K_1 x^{(m)} = -\frac{1}{\tau}c^T\hat{x}^{(m)} = 0$$

Orthogonality constraint satisfied

Remark

The solution $z^{(m)} = [x^{(m)}; 0]$, satisfies $c^T x^{(m)} = 0$.

Proof

- since $\hat{x}^{(m)}$ is optimal approximate solution in $\mathcal{K}_m((KK_1^{-1}, v)$

$$\hat{x}^{(m)} = \phi_{m-1}(KK_1^{-1})v$$

  and $c^T(KK_1^{-1})^k = \frac{\sigma}{\tau}c^T$ we have $c^T\hat{x}^{(m)} = 0$.
- Then

$$c^T x^{(m)} = -\frac{1}{\tau}c^T K_1 x^{(m)} = -\frac{1}{\tau}c^T\hat{x}^{(m)} = 0$$

Orthogonality constraint satisfied

Remark

The solution $z^{(m)} = [x^{(m)}; 0]$, satisfies $c^T x^{(m)} = 0$.

Proof

- since $\hat{x}^{(m)}$ is optimal approximate solution in $\mathcal{K}_m((KK_1^{-1}, v)$

$$\hat{x}^{(m)} = \phi_{m-1}(KK_1^{-1})v$$

  and $c^T(KK_1^{-1})^k = \frac{\sigma}{\tau}c^T$ we have $c^T\hat{x}^{(m)} = 0$.
- Then

$$c^T x^{(m)} = -\frac{1}{\tau}c^T K_1 x^{(m)} = -\frac{1}{\tau}c^T\hat{x}^{(m)} = 0$$

The preconditioner and its properties

- structured symmetric indefinite preconditioner

$$P_I = \begin{bmatrix} K_1 & c \\ c^T & 0 \end{bmatrix}, \quad K_1 = A_1 - \tau I, \quad \tau \in \mathbb{R}$$

  where $K_1 = K_1^T$ nonsingular and $A_1 c = 0$ so $\tau \neq 0$ ($A_1 = 0$ possible)

- also possible $K_1 = A_1 - \sigma M_1$ for general $M_1 = M_1^T$

The preconditioner and its properties

- structured symmetric indefinite preconditioner

$$P_I = \begin{bmatrix} K_1 & c \\ c^T & 0 \end{bmatrix}, \quad K_1 = A_1 - \tau I, \quad \tau \in \mathbb{R}$$

  where $K_1 = K_1^T$ nonsingular and $A_1 c = 0$ so $\tau \neq 0$ ($A_1 = 0$ possible)
- also possible $K_1 = A_1 - \sigma M_1$ for general $M_1 = M_1^T$

Equivalence of optimal solutions

Theorem

Let v satisfy $c^T v = 0$. The optimal Krylov subspace solution of the augmented system $z^{(m)}$ with right preconditioner P_I can be written as

$$z^{(m)} = [x^{(m)}; 0],$$

where $x^{(m)}$ is the optimal Krylov subspace solution of the original (non-augmented) system with preconditioner K_1.

Remark

The solution $z^{(m)} = [x^{(m)}; 0]$, satisfies $c^T x^{(m)} = 0$.

Equivalence of optimal solutions

Theorem

Let v satisfy $c^T v = 0$. The optimal Krylov subspace solution of the augmented system $z^{(m)}$ with right preconditioner P_I can be written as

$$z^{(m)} = [x^{(m)}; 0],$$

where $x^{(m)}$ is the optimal Krylov subspace solution of the original (non-augmented) system with preconditioner K_1.

Remark

The solution $z^{(m)} = [x^{(m)}; 0]$, satisfies $c^T x^{(m)} = 0$.

Outline

Back to the solve of the outer system

- augmented formulation of the problem

$$\begin{bmatrix} A & c \\ c^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} I & 0 \\ 0^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

- $n_c = 1$ zero eigenvalue of the original problem become infinite
- $n_c = 1$ more eigenvalues arise (corresponding) to the singular part of $\mathcal{I}$; infinite
- non-zero eigenvalues remain unchanged; find smallest eigenvalues of the augmented system; eigenvectors are of the form $[x; 0]$
- exact SI-Lanczos - inexact SI Lanczos

Back to the solve of the outer system

- augmented formulation of the problem

$$\begin{bmatrix} A & c \\ c^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} I & 0 \\ 0^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

- $n_c = 1$ zero eigenvalue of the original problem become infinite
- $n_c = 1$ more eigenvalues arise (corresponding) to the singular part of $\mathcal{I}$; infinite
- non-zero eigenvalues remain unchanged; find smallest eigenvalues of the augmented system; eigenvectors are of the form $[x; 0]$
- exact SI-Lanczos - inexact SI Lanczos

Back to the solve of the outer system

- augmented formulation of the problem

$$\begin{bmatrix} A & c \\ c^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} I & 0 \\ 0^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

- $n_c = 1$ zero eigenvalue of the original problem become infinite
- $n_c = 1$ more eigenvalues arise (corresponding) to the singular part of $\mathcal{I}$; infinite
- non-zero eigenvalues remain unchanged; find smallest eigenvalues of the augmented system; eigenvectors are of the form $[x; 0]$
- exact SI-Lanczos - inexact SI Lanczos

Back to the solve of the outer system

- augmented formulation of the problem

$$\begin{bmatrix} A & c \\ c^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} I & 0 \\ 0^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

- $n_c = 1$ zero eigenvalue of the original problem become infinite
- $n_c = 1$ more eigenvalues arise (corresponding) to the singular part of $\mathcal{I}$; infinite
- non-zero eigenvalues remain unchanged; find smallest eigenvalues of the augmented system; eigenvectors are of the form $[x; 0]$
- exact SI-Lanczos - inexact SI Lanczos

Back to the solve of the outer system

- augmented formulation of the problem

$$\begin{bmatrix} A & c \\ c^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} I & 0 \\ 0^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

- $n_c = 1$ zero eigenvalue of the original problem become infinite
- $n_c = 1$ more eigenvalues arise (corresponding) to the singular part of $\mathcal{I}$; infinite
- non-zero eigenvalues remain unchanged; find smallest eigenvalues of the augmented system; eigenvectors are of the form $[x; 0]$
- exact SI-Lanczos - inexact SI Lanczos

Preconditioning with P_D

Theorem

Let u_1 satisfy $c^T u_1 = 0$. Inexact SI-Lanczos with shift σ applied to the augmented formulation with staring vector $v_1 = [u_1; 0]$ and inner right preconditioner

$$P_D = \begin{bmatrix} K_1 & 0 \\ 0 & c^T K_1^{-1} c \end{bmatrix}, \quad K_1 = A_1 - \tau I, \quad A_1 c = 0, \quad A_1 \in \mathbb{R}^{n \times n}$$

with $K_1 = K_1^T$ nonsingular generates the same approximation iterates as inexact SI-Lanczos with shift σ applied to the original problem with starting vector u_1 and inner right preconditioner K_1.

Proof Idea

Uses results that optimal Krylov subspace approximate solution of inner systems are essentially the same. Induction.

Preconditioning with P_D

Theorem

Let u_1 satisfy $c^T u_1 = 0$. Inexact SI-Lanczos with shift σ applied to the augmented formulation with staring vector $v_1 = [u_1; 0]$ and inner right preconditioner

$$P_D = \begin{bmatrix} K_1 & 0 \\ 0 & c^T K_1^{-1} c \end{bmatrix}, \quad K_1 = A_1 - \tau I, \quad A_1 c = 0, \quad A_1 \in \mathbb{R}^{n \times n}$$

with $K_1 = K_1^T$ nonsingular generates the same approximation iterates as inexact SI-Lanczos with shift σ applied to the original problem with starting vector u_1 and inner right preconditioner K_1.

Proof Idea

Uses results that optimal Krylov subspace approximate solution of inner systems are essentially the same. Induction.

Preconditioning with P_I

Theorem

Let u_1 satisfy $c^T u_1 = 0$. Inexact SI-Lanczos with shift σ applied to the augmented formulation with staring vector $v_1 = [u_1; 0]$ and inner right preconditioner

$$P_I = \begin{bmatrix} K_1 & c \\ c^T & 0 \end{bmatrix}, \quad K_1 = A_1 - \tau I, \quad A_1 c = 0, \quad A_1 \in \mathbb{R}^{n \times n}$$

with $K_1 = K_1^T$ nonsingular generates the same approximation iterates as inexact SI-Lanczos with shift σ applied to the original problem with starting vector u_1 and inner right preconditioner K_1.

Proof Idea

Uses results that optimal Krylov subspace approximate solution of inner systems are essentially the same. Induction.

Remarks

- key condition is $c^T K_1 = \beta c^T$ for $\beta \neq 0$
- here: $K_1 = A_1 - \tau I$
- could use $K_1 = \alpha A_1 + cH^{-1}c^T$ with $H = c^T c$, $\alpha \in \mathbb{R}$, $A_1 c = 0$.

Remarks

- also possible: higher dimensional null-spaces of A, where C is a basis of the null-space such that $AC = 0$
- also possible: generalised eigenproblem $Ax = \lambda Mx$

Remarks

- also possible: higher dimensional null-spaces of A, where C is a basis of the null-space such that $AC = 0$
- also possible: generalised eigenproblem $Ax = \lambda Mx$

Outline

Regularisation of the problem

- other than augmented formulation, a so-called regularised formulation is available
- move zero eigenvalues away from the origin and also (hopefully) far away from the sought after eigenvalues
- let $H \in \mathbb{R}^{n_c \times n_c}$ (here H is just a scalar) be symmetric and nonsingular, then the transformed generalised eigenvalue problem is given by

$$(A + cH^{-1}c^T)x = \eta x$$

Regularisation of the problem

- other than augmented formulation, a so-called regularised formulation is available
- move zero eigenvalues away from the origin and also (hopefully) far away from the sought after eigenvalues
- let $H \in \mathbb{R}^{n_c \times n_c}$ (here H is just a scalar) be symmetric and nonsingular, then the transformed generalised eigenvalue problem is given by

$$(A + cH^{-1}c^T)x = \eta x$$

Regularisation of the problem

- other than augmented formulation, a so-called regularised formulation is available
- move zero eigenvalues away from the origin and also (hopefully) far away from the sought after eigenvalues
- let $H \in \mathbb{R}^{n_c \times n_c}$ (here H is just a scalar) be symmetric and nonsingular, then the transformed generalised eigenvalue problem is given by

$$(A + cH^{-1}c^T)x = \eta x$$

Shifting of the zero eigenvalue

Theorem

Let

$$Ax = \lambda x \quad \text{and} \quad (A + cH^{-1}c^T)x = \eta x$$

and λ_i, η_i be eigenvalues.

- *If $\lambda_i \neq 0$ there exists j such that $\lambda_i = \eta_j$.*
- *If $\lambda_i = 0$ there corresponds an eigenvalue η_j with $\eta_j \in \Lambda(c^T c, H)$*

Remarks

- no practical advantage
- inner solver $(A + cH^{-1}c^T)z = v$ produces the same Krylov subspace as $Az = v$

Shifting of the zero eigenvalue

Theorem

Let

$$Ax = \lambda x \quad \text{and} \quad (A + cH^{-1}c^T)x = \eta x$$

and λ_i, η_i be eigenvalues.

- *If $\lambda_i \neq 0$ there exists j such that $\lambda_i = \eta_j$.*
- *If $\lambda_i = 0$ there corresponds an eigenvalue η_j with $\eta_j \in \Lambda(c^T c, H)$*

Remarks

- no practical advantage
- inner solver $(A + cH^{-1}c^T)z = v$ produces the same Krylov subspace as $Az = v$

Shifting of the zero eigenvalue

Theorem

Let

$$Ax = \lambda x \quad \textit{and} \quad (A + cH^{-1}c^T)x = \eta x$$

and λ_i, η_i be eigenvalues.

- *If $\lambda_i \neq 0$ there exists j such that $\lambda_i = \eta_j$.*
- *If $\lambda_i = 0$ there corresponds an eigenvalue η_j with $\eta_j \in \Lambda(c^T c, H)$*

Remarks

- no practical advantage
- inner solver $(A + cH^{-1}c^T)z = v$ produces the same Krylov subspace as $Az = v$

Outline

2D computational model of an electromagnetic cavity resonator

- variational formulation: Find $\omega_h \in \mathbb{R}$ s.t. $\exists 0 \neq \mathbf{u}_h \in \Sigma_h \subset \Sigma$

$$(\text{rot}\mathbf{u}_h, \text{rot}\mathbf{v}_h) = \omega_h^2(\mathbf{u}_h, \mathbf{v}_h) \quad \forall \mathbf{v}_h \in \Sigma_h,$$

  where $\text{rot}(v_1, v_2) = (v_2)_x - (v_1)_y$,
  $\Sigma = \{\mathbf{v} \in L^2(\Omega)^2 : \text{rot}\mathbf{v} \in L^2(\Omega), \mathbf{v} \cdot \mathbf{t} = 0 \,\text{on}\, \partial\Omega\}$ and $\mathbf{t}$ is the counterclockwise oriented tangent unit vector to the boundary
- FEM discretisation
- size $n = 3229$, null-space dimension $n_c = 1036$
- solver: right preconditioned GMRES
- preconditioner $K_1 = A_1 - \sigma M$ with $A_1 = 0$ and $\sigma = 0.8$
- inner tolerance 10^{-8} for the solve of the inner system

Results

j	$(A-\sigma M)^{-1}Mx=\eta x$ K_1	$(\mathcal{A}-\sigma\mathcal{M})^{-1}\mathcal{M}z=\eta z$ P_D	P_I
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

Table: Relative eigenvalue residual norm $\frac{Ax_j - \lambda_j M x_j}{\lambda_j}$ of approximate smallest eigenpair in the inexact SI-Lanczos method applied to different formulations and different preconditioners

Outline

Conclusions

- different formulations for constraint eigenvalue problem (especially augmented formulation)
- augmented schemes are equivalent to original formulation if inexact SI-Lanczos is used (for natural choices of the preconditioner for the inner system)
- dependent on the fact that the constraint matrix C is a basis for the null-space of the problem
- approximation space is maintained M-orthogonal to the null-space without explicit orthogonalisation (constraint $C^T Mx = 0$ automatically satisfied)
- inner accuracy influences the performance of the method

Conclusions

- different formulations for constraint eigenvalue problem (especially augmented formulation)
- augmented schemes are equivalent to original formulation if inexact SI-Lanczos is used (for natural choices of the preconditioner for the inner system)
- dependent on the fact that the constraint matrix C is a basis for the null-space of the problem
- approximation space is maintained M-orthogonal to the null-space without explicit orthogonalisation (constraint $C^T Mx = 0$ automatically satisfied)
- inner accuracy influences the performance of the method

Conclusions

- different formulations for constraint eigenvalue problem (especially augmented formulation)
- augmented schemes are equivalent to original formulation if inexact SI-Lanczos is used (for natural choices of the preconditioner for the inner system)
- dependent on the fact that the constraint matrix C is a basis for the null-space of the problem
- approximation space is maintained M-orthogonal to the null-space without explicit orthogonalisation (constraint $C^T Mx = 0$ automatically satisfied)
- inner accuracy influences the performance of the method

K. A. Cliffe, T. J. Garrett, and A. Spence, *Eigenvalues of block matrices arising from problems in fluid mechanics*, SIAM Journal of Matrix Analysis and Applications, 15 (1994), pp. 1310–1318.

V. Simoncini, *Algebraic formulations for the solution of the nullspace-free eigenvalue problem using the inexact Shift-and-Invert Lanczos method*, Numer. Linear Algebra Appl., 10 (2003), pp. 357–375.