On efficient numerical approximation of the scattering amplitude

Petr Tichý

joint work with

Zdeněk Strakoš

Institute of Computer Science AS CR

June 23, 2009, Modelling 2009,
Rožnov pod Radhoštěm, Czech Republic

Formulation of the problem

Given a nonsingular matrix $\mathbf{A}$ and vectors b and c.

We want to approximate

$$c^*\mathbf{A}^{-1}b\,.$$

Equivalently, we look for an approximation to

$$c^*x \qquad \text{such that} \qquad \mathbf{A}x = b\,.$$

Motivation

- **Approximation of the jth component of the solution**
  - i.e., we want to approximate $e_j^T \mathbf{A}^{-1} b$.

Motivation

- **Approximation of the jth component of the solution**
  - i.e., we want to approximate $e_j^T \mathbf{A}^{-1} b$.
- **Signal processing (the scattering amplitude)**
  - b and c represent incoming and outgoing waves, respectively, and the operator $\mathbf{A}$ relates the incoming and scattered fields on the surface of an object,
  - $\mathbf{A}x = b$ determines the field x from the signal b. The signal is received on an antenna c. The signal received by the antenna is then $c^* x$. The value $c^* x$ is called *the scattering amplitude*.

Motivation

- **Approximation of the jth component of the solution**
  - i.e., we want to approximate $e_j^T \mathbf{A}^{-1} b$.
- **Signal processing (the scattering amplitude)**
  - b and c represent incoming and outgoing waves, respectively, and the operator $\mathbf{A}$ relates the incoming and scattered fields on the surface of an object,
  - $\mathbf{A}x = b$ determines the field x from the signal b. The signal is received on an antenna c. The signal received by the antenna is then $c^* x$. The value $c^* x$ is called *the scattering amplitude*.
- **Optimization (the primal linear output)**

Motivation

- **Approximation of the jth component of the solution**
  - i.e., we want to approximate $e_j^T \mathbf{A}^{-1} b$.
- **Signal processing (the scattering amplitude)**
  - b and c represent incoming and outgoing waves, respectively, and the operator $\mathbf{A}$ relates the incoming and scattered fields on the surface of an object,
  - $\mathbf{A}x = b$ determines the field x from the signal b. The signal is received on an antenna c. The signal received by the antenna is then $c^* x$. The value $c^* x$ is called *the scattering amplitude*.
- **Optimization (the primal linear output)**
- **Nuclear physics, quantum mechanics, other disciplines**

Krylov subspace methods approach

Projection of the original problem onto Krylov subspaces

$$\mathcal{K}_n(\mathbf{A}, b) = \mathrm{span}\{b, \mathbf{A}b, \ldots \mathbf{A}^{n-1}b\}.$$

Krylov subspace methods approach

Projection of the original problem onto Krylov subspaces

$$\mathcal{K}_n(\mathbf{A}, b) = \mathrm{span}\{b, \mathbf{A}b, \ldots \mathbf{A}^{n-1}b\}.$$

A possible approach: compute x_n using a Krylov subspace method,

$$c^*\mathbf{A}^{-1}b \;=\; c^*x \;\approx\; c^*x_n\,.$$

Krylov subspace methods approach

Projection of the original problem onto Krylov subspaces

$$\mathcal{K}_n(\mathbf{A}, b) = \mathrm{span}\{b, \mathbf{A}b, \ldots \mathbf{A}^{n-1}b\}.$$

A possible approach: compute x_n using a Krylov subspace method,

$$c^*\mathbf{A}^{-1}b \;=\; c^*x \;\approx\; c^*x_n\,.$$

- Due to numerical instabilities, the explicit numerical computation of c^*x_n can be *highly inefficient*.
  [HPD case: Strakoš & T. '02, '05]

Krylov subspace methods approach

Projection of the original problem onto Krylov subspaces

$$\mathcal{K}_n(\mathbf{A}, b) = \mathrm{span}\{b, \mathbf{A}b, \dots \mathbf{A}^{n-1}b\}.$$

A possible approach: compute x_n using a Krylov subspace method,

$$c^*\mathbf{A}^{-1}b \;=\; c^*x \;\approx\; c^*x_n\,.$$

- Due to numerical instabilities, the explicit numerical computation of c^*x_n can be *highly inefficient*.
  [HPD case: Strakoš & T. '02, '05]
- If $\mathbf{A}$ is HPD and $c = b$, there are several efficient methods.
  [Golub & Meurant '94, '97, Axelsson & Kaporin '01, Strakoš & T. '02, '05]

Krylov subspace methods approach

Projection of the original problem onto Krylov subspaces

$$\mathcal{K}_n(\mathbf{A}, b) = \mathrm{span}\{b, \mathbf{A}b, \dots \mathbf{A}^{n-1}b\}.$$

A possible approach: compute x_n using a Krylov subspace method,

$$c^*\mathbf{A}^{-1}b \;=\; c^*x \;\approx\; c^*x_n\,.$$

- Due to numerical instabilities, the explicit numerical computation of c^*x_n can be *highly inefficient*.
  [HPD case: Strakoš & T. '02, '05]
- If $\mathbf{A}$ is HPD and $c = b$, there are several efficient methods.
  [Golub & Meurant '94, '97, Axelsson & Kaporin '01, Strakoš & T. '02, '05]
- How to generalize ideas from the HPD case to a general case?

Outline

Outline

Vorobyev moment problem, Vorobyev '58, '65

Popularized by Brezinski '97, Strakoš '08

Find a linear operator $\mathbf{A}_n$ on $\mathcal{K}_n(\mathbf{A}, v)$ such that

$$\begin{aligned}
\mathbf{A}_n\, v &= \mathbf{A}\, v, \\
\mathbf{A}_n^2\, v &= \mathbf{A}^2\, v, \\
&\vdots \\
\mathbf{A}_n^{n-1}\, v &= \mathbf{A}^{n-1}\, v, \\
\mathbf{A}_n^{n}\, v &= \mathbf{Q}_n \mathbf{A}^{n}\, v\,,
\end{aligned}$$

where $\mathbf{Q}_n$ is a given linear projection operator.

Vorobyev moment problem, Vorobyev ’58, ’65

Popularized by Brezinski ’97, Strakoš ’08

Find a linear operator $\mathbf{A}_n$ on $\mathcal{K}_n(\mathbf{A}, v)$ such that

$$\begin{aligned} \mathbf{A}_n\, v &= \mathbf{A}\, v, \\ \mathbf{A}_n^2\, v &= \mathbf{A}^2\, v, \\ &\vdots \\ \mathbf{A}_n^{n-1}\, v &= \mathbf{A}^{n-1}\, v, \\ \mathbf{A}_n^n\, v &= \mathbf{Q}_n \mathbf{A}^n\, v\,, \end{aligned}$$

where $\mathbf{Q}_n$ is a given linear projection operator.

- Some Krylov subspace methods can be interpreted as methods that solve the Vorobyev moment problem.
- Useful formulation for understanding approximation properties of Krylov subspace methods.

Non-Hermitian Lanczos

Given a nonsingular $\mathbf{A}$, v and w.

Non-Hermitian Lanczos algorithm is represented by

$$\begin{aligned} \mathbf{A}\mathbf{V}_n &= \mathbf{V}_n\mathbf{T}_n + \delta_{n+1}v_{n+1}e_n^T, \\ \mathbf{A}^*\mathbf{W}_n &= \mathbf{W}_n\mathbf{T}_n^* + \eta_{n+1}^* w_{n+1}e_n^T, \end{aligned}$$

where $\mathbf{W}_n^*\mathbf{V}_n = \mathbf{I}$ and $\mathbf{T}_n = \mathbf{W}_n^*\mathbf{A}\mathbf{V}_n$ is tridiagonal,

$$\mathbf{T}_n = \begin{bmatrix} \gamma_1 & \eta_2 & & \\ \delta_2 & \gamma_2 & \ddots & \\ & \ddots & \ddots & \eta_n \\ & & \delta_n & \gamma_n \end{bmatrix}.$$

Arnoldi algorithm

Given a nonsingular $\mathbf{A}$ and v.

Arnoldi algorithm is represented by

$$\mathbf{A}\mathbf{V}_n = \mathbf{V}_n\mathbf{H}_n + h_{n+1,n}v_{n+1}e_n^T,$$

where $\mathbf{V}_n^*\mathbf{V}_n = \mathbf{I}$, and $\mathbf{H}_n =\mathbf{V}_n^*\mathbf{A}\mathbf{V}_n$ is upper Hessenberg,

$$\mathbf{H}_n = \begin{bmatrix} h_{1,1} & h_{1,2} & \dots & h_{1,n} \\ h_{2,1} & h_{2,2} & \ddots & \vdots \\ & \ddots & \ddots & h_{n-n,n} \\ & & h_{n,n-1} & h_{n.n} \end{bmatrix}.$$

Define $\mathbf{Q}_n$: it projects onto $\mathcal{K}_n(\mathbf{A}, v)$ orthogonally to $\mathcal{K}_n(\mathbf{A}^*, w)$.

Define $\mathbf{Q}_n$: it projects onto $\mathcal{K}_n(\mathbf{A}, v)$ orthogonally to $\mathcal{K}_n(\mathbf{A}^*, w)$.

- Then

$$\begin{aligned} \mathbf{Q}_n &= \mathbf{V}_n \mathbf{W}_n^* \,, \\ \mathbf{A}_n &= \mathbf{V}_n \mathbf{T}_n \mathbf{W}_n^* \,. \end{aligned}$$

Non-Hermitian Lanczos

Vorobyev moment problem, matching moments, model reduction

Define $\mathbf{Q}_n$: it projects onto $\mathcal{K}_n(\mathbf{A}, v)$ orthogonally to $\mathcal{K}_n(\mathbf{A}^*, w)$.

- Then

$$\begin{aligned} \mathbf{Q}_n &= \mathbf{V}_n \mathbf{W}_n^*, \\ \mathbf{A}_n &= \mathbf{V}_n \mathbf{T}_n \mathbf{W}_n^*. \end{aligned}$$

- Matching moments property of Non-Hermitian Lanczos:

  [Gragg & Lindquist '83, Villemagne & Skelton '87]

  [Gallivan & Grimme & Van Dooren '94, Antoulas '05]

  [a simple proof using the Vorobyev moment problem - Strakoš '08]

$$w^* \mathbf{A}^k v = w^* \mathbf{A}_n^k v = e_1^* \mathbf{T}_n^k e_1, \qquad k = 0, \ldots, 2n-1.$$

Non-Hermitian Lanczos

Vorobyev moment problem, matching moments, model reduction

Define $\mathbf{Q}_n$: it projects onto $\mathcal{K}_n(\mathbf{A}, v)$ orthogonally to $\mathcal{K}_n(\mathbf{A}^*, w)$.

- Then

$$\begin{aligned} \mathbf{Q}_n &= \mathbf{V}_n \mathbf{W}_n^*\,, \\ \mathbf{A}_n &= \mathbf{V}_n \mathbf{T}_n \mathbf{W}_n^*\,. \end{aligned}$$

- Matching moments property of Non-Hermitian Lanczos:

  [Gragg & Lindquist '83, Villemagne & Skelton '87]

  [Gallivan & Grimme & Van Dooren '94, Antoulas '05]

  [a simple proof using the Vorobyev moment problem - Strakoš '08]

$$w^* \mathbf{A}^k v = w^* \mathbf{A}_n^k v = e_1^* \mathbf{T}_n^k e_1, \qquad k = 0, \ldots, 2n-1\,.$$

- Model reduction

$$\mathbf{A}, v, w \quad \rightarrow \quad \mathbf{T}_n, e_1, e_1\,.$$

Arnoldi algorithm

Vorobyev moment problem, matching moments, model reduction

Define $\mathbf{Q}_n$: it projects onto $\mathcal{K}_n(\mathbf{A}, v)$ orthogonally to $\mathcal{K}_n(\mathbf{A}, v)$.

- Then

$$\begin{aligned} \mathbf{Q}_n &= \mathbf{V}_n \mathbf{V}_n^* \,, \\ \mathbf{A}_n &= \mathbf{V}_n \mathbf{H}_n \mathbf{V}_n^* \,. \end{aligned}$$

- Matching moments property of Arnoldi:

$$w^* \mathbf{A}^k v = w^* \mathbf{A}_n^k v = t_n^* \mathbf{H}_n^k e_1, \qquad k = 0, \ldots, n-1 \,,$$

  w is given, $t_n = \mathbf{V}_n^* w$.

- Model reduction

$$\mathbf{A}, v, w \quad \rightarrow \quad \mathbf{H}_n, e_1, t_n \,.$$

Outline

Approximation of $c^*\mathbf{A}^{-1}b$

General framework, Strakoš & T. '09

Vorobyev moment problem: $\mathbf{A} \quad \rightarrow \quad \mathbf{A}_n$

Define approximation: $c^*\mathbf{A}^{-1}b \quad \approx \quad c^*\mathbf{A}_n^{-1}b$

$\mathbf{A}_n^{-1}$ is the matrix representation of the inverse of the reduced order operator $\mathbf{A}_n$ which is restricted onto $\mathcal{K}_n(\mathbf{A}, b)$.

Approximation of $c^* \mathbf{A}^{-1} b$

General framework, Strakoš & T. '09

Vorobyev moment problem: $\mathbf{A} \quad \longrightarrow \quad \mathbf{A}_n$

Define approximation: $c^* \mathbf{A}^{-1} b \quad \approx \quad c^* \mathbf{A}_n^{-1} b$

$\mathbf{A}_n^{-1}$ is the matrix representation of the inverse of the reduced order operator $\mathbf{A}_n$ which is restricted onto $\mathcal{K}_n(\mathbf{A}, b)$.

Examples:

- $\mathbf{A}_n^{-1} \; = \; \mathbf{V}_n \mathbf{T}_n^{-1} \mathbf{W}_n^*$ (Non-Hermitian Lanczos)
- $\mathbf{A}_n^{-1} \; = \; \mathbf{V}_n \mathbf{H}_n^{-1} \mathbf{V}_n^*$ (Arnoldi)

Approximation of $c^*\mathbf{A}^{-1}b$

General framework, Strakoš & T. '09

Vorobyev moment problem: $\mathbf{A} \quad \longrightarrow \quad \mathbf{A}_n$

Define approximation: $c^*\mathbf{A}^{-1}b \quad \approx \quad c^*\mathbf{A}_n^{-1}b$

$\mathbf{A}_n^{-1}$ is the matrix representation of the inverse of the reduced order operator $\mathbf{A}_n$ which is restricted onto $\mathcal{K}_n(\mathbf{A}, b)$.

Examples:

- $\mathbf{A}_n^{-1} = \mathbf{V}_n \mathbf{T}_n^{-1} \mathbf{W}_n^*$ (Non-Hermitian Lanczos)
- $\mathbf{A}_n^{-1} = \mathbf{V}_n \mathbf{H}_n^{-1} \mathbf{V}_n^*$ (Arnoldi)

Questions:

- How to compute $c^*\mathbf{A}_n^{-1}b$ efficiently?
- Relationship to the existing approximations?

Approximation of $c^* \mathbf{A}^{-1} b$

General framework, Strakoš & T. '09

Vorobyev moment problem: $\quad \mathbf{A} \quad \rightarrow \quad \mathbf{A}_n$

Define approximation: $\quad c^* \mathbf{A}^{-1} b \quad \approx \quad c^* \mathbf{A}_n^{-1} b$

$\mathbf{A}_n^{-1}$ is the matrix representation of the inverse of the reduced order operator $\mathbf{A}_n$ which is restricted onto $\mathcal{K}_n(\mathbf{A}, b)$.

Examples:

- $\mathbf{A}_n^{-1} \; = \; \mathbf{V}_n \mathbf{T}_n^{-1} \mathbf{W}_n^*$ (Non-Hermitian Lanczos)
- $\mathbf{A}_n^{-1} \; = \; \mathbf{V}_n \mathbf{H}_n^{-1} \mathbf{V}_n^*$ (Arnoldi)

Questions:

- How to compute $c^* \mathbf{A}_n^{-1} b$ efficiently?
- Relationship to the existing approximations?

We concentrate only to non-Hermitian Lanczos approach.

Non-Hermitian Lanczos approach

Define

$$v_1 = \frac{b}{\|b\|}, \qquad w_1 = \frac{c}{c^* v_1}, \qquad \text{i.e.} \qquad w_1^* v_1 = 1 .$$

Then

$$c^* \mathbf{A}_n^{-1} b \;=\; c^* \mathbf{V}_n \mathbf{T}_n^{-1} \mathbf{W}_n^* b \;=\; (c^* v_1)\, \|b\|\, (\mathbf{T}_n^{-1})_{1,1} .$$

Non-Hermitian Lanczos approach

Define

$$v_1 = \frac{b}{\|b\|}\,, \qquad w_1 = \frac{c}{c^* v_1}\,, \qquad \text{i.e.} \qquad w_1^* v_1 = 1\,.$$

Then

$$c^* \mathbf{A}_n^{-1} b \;=\; c^* \mathbf{V}_n \mathbf{T}_n^{-1} \mathbf{W}_n^* b \;=\; (c^* v_1)\, \|b\|\, (\mathbf{T}_n^{-1})_{1,1}\,.$$

Let $x_0 = 0$. We also know that $x_n \;=\; \|b\| \mathbf{V}_n \mathbf{T}_n^{-1} e_1$ is the approximate solution computed via BiCG. Therefore,

$$c^* \mathbf{A}_n^{-1} b \;=\; c^* \|b\| \mathbf{V}_n \mathbf{T}_n^{-1} \mathbf{W}_n^* \mathbf{V}_n e_1 \;=\; c^* x_n\,.$$

Non-Hermitian Lanczos approach

Define

$$v_1 = \frac{b}{\|b\|}\,, \qquad w_1 = \frac{c}{c^* v_1}\,, \qquad \text{i.e.} \qquad w_1^* v_1 = 1\,.$$

Then

$$c^* \mathbf{A}_n^{-1} b \;=\; c^* \mathbf{V}_n \mathbf{T}_n^{-1} \mathbf{W}_n^* b \;=\; (c^* v_1)\,\|b\|\,(\mathbf{T}_n^{-1})_{1,1}\,.$$

Let $x_0 = 0$. We also know that $x_n \;=\; \|b\| \mathbf{V}_n \mathbf{T}_n^{-1} e_1$ is the approximate solution computed via BiCG. Therefore,

$$c^* \mathbf{A}_n^{-1} b \;=\; c^* \|b\| \mathbf{V}_n \mathbf{T}_n^{-1} \mathbf{W}_n^* \mathbf{V}_n e_1 \;=\; c^* x_n\,.$$

- BiCG can be used for computing $c^* \mathbf{A}_n^{-1} b$!

Non-Hermitian Lanczos approach

Define

$$v_1 = \frac{b}{\|b\|}\,, \qquad w_1 = \frac{c}{c^* v_1}\,, \qquad \text{i.e.} \qquad w_1^* v_1 = 1\,.$$

Then

$$c^* \mathbf{A}_n^{-1} b \;=\; c^* \mathbf{V}_n \mathbf{T}_n^{-1} \mathbf{W}_n^* b \;=\; (c^* v_1)\,\|b\|\,(\mathbf{T}_n^{-1})_{1,1}\,.$$

Let $x_0 = 0$. We also know that $x_n \;=\; \|b\| \mathbf{V}_n \mathbf{T}_n^{-1} e_1$ is the approximate solution computed via BiCG. Therefore,

$$c^* \mathbf{A}_n^{-1} b \;=\; c^* \|b\| \mathbf{V}_n \mathbf{T}_n^{-1} \mathbf{W}_n^* \mathbf{V}_n e_1 \;=\; c^* x_n\,.$$

- BiCG can be used for computing $c^* \mathbf{A}_n^{-1} b$!
- We used the global biorthogonality !
  Do the identities hold in finite precision computations?

The BiCG method

Simultaneous solving of

$$\mathbf{A}x = b, \qquad \mathbf{A}^* y = c.$$

input $\mathbf{A}$, b, c

$x_0 = y_0 = 0$
$r_0 = p_0 = b,\ s_0 = q_0 = c$

for $n = 0, 1, \ldots$

$\alpha_n = \frac{s_n^* r_n}{q_n^* \mathbf{A} p_n}$,

$x_{n+1} = x_n + \alpha_n\, p_n$, $\qquad y_{n+1} = y_n + {\alpha_n}^*\, q_n$,

$r_{n+1} = r_n - \alpha_n\, \mathbf{A} p_n$, $\qquad s_{n+1} = s_n - \alpha_n^*\, \mathbf{A}^* q_n$,

$\beta_{n+1} = \frac{s_{n+1}^* r_{n+1}}{s_n^* r_n}$,

$p_{n+1} = r_{n+1} + \beta_{n+1}\, p_n$, $\qquad q_{n+1} = s_{n+1} + \beta_{n+1}^*\, q_n$

end

An efficient approximation based on the BiCG method

How to compute $c^* \mathbf{A}_n^{-1} b$ in BiCG without using the global biorthogonality?

Using local biorthogonality we can show that

$$s_j^* \mathbf{A}^{-1} r_j \; - \; s_{j+1}^* \mathbf{A}^{-1} r_{j+1} \; = \; \alpha_j s_j^* r_j .$$

An efficient approximation based on the BiCG method

How to compute $c^*\mathbf{A}_n^{-1}b$ in BiCG without using the global biorthogonality?

Using local biorthogonality we can show that

$$s_j^*\mathbf{A}^{-1}r_j \;-\; s_{j+1}^*\mathbf{A}^{-1}r_{j+1} \;=\; \alpha_j s_j^* r_j.$$

Consequently,

$$c^*\mathbf{A}^{-1}b \;=\; \sum_{j=0}^{n-1}\alpha_j s_j^* r_j \;+\; s_n^*\mathbf{A}^{-1}r_n\,.$$

An efficient approximation based on the BiCG method

How to compute $c^* \mathbf{A}_n^{-1} b$ in BiCG without using the global biorthogonality?

Using local biorthogonality we can show that

$$s_j^* \mathbf{A}^{-1} r_j \; - \; s_{j+1}^* \mathbf{A}^{-1} r_{j+1} \; = \; \alpha_j s_j^* r_j .$$

Consequently,

$$c^* \mathbf{A}^{-1} b \; = \; \sum_{j=0}^{n-1} \alpha_j s_j^* r_j \; + \; s_n^* \mathbf{A}^{-1} r_n \, .$$

Moreover, it can be shown (using global biorthogonality) that

$$c^* \mathbf{A}^{-1} b \; = \; c^* x_n \; + \; s_n^* \mathbf{A}^{-1} r_n \, .$$

An efficient approximation based on the BiCG method

How to compute $c^* \mathbf{A}_n^{-1} b$ in BiCG without using the global biorthogonality?

Using local biorthogonality we can show that

$$s_j^* \mathbf{A}^{-1} r_j \; - \; s_{j+1}^* \mathbf{A}^{-1} r_{j+1} \; = \; \alpha_j s_j^* r_j .$$

Consequently,

$$c^* \mathbf{A}^{-1} b \; = \; \sum_{j=0}^{n-1} \alpha_j s_j^* r_j \; + \; s_n^* \mathbf{A}^{-1} r_n \, .$$

Moreover, it can be shown (using global biorthogonality) that

$$c^* \mathbf{A}^{-1} b \; = \; c^* x_n \; + \; s_n^* \mathbf{A}^{-1} r_n \, .$$

Finally,

$$c^* \mathbf{A}_n^{-1} b \; = \; (c^* v_1) \, \|b\| \, (\mathbf{T}_n^{-1})_{1,1} \; = \; c^* x_n \; = \; \sum_{j=0}^{n-1} \alpha_j s_j^* r_j \, .$$

Saylor-Smolarski approach

For diagonalizable matrices

[Saylor & Smolarski '01] introduce

- formally orthogonal polynomials,
- complex Gauss quadrature,

as a tool for approximating the quantity $c^* \mathbf{A}^{-1} b$. Motivated by [Freund & Hochbruck '93], [Golub & Meurant '94, '97].

Saylor-Smolarski approach

For diagonalizable matrices

[Saylor & Smolarski '01] introduce

- formally orthogonal polynomials,
- complex Gauss quadrature,

as a tool for approximating the quantity $c^* \mathbf{A}^{-1} b$. Motivated by [Freund & Hochbruck '93], [Golub & Meurant '94, '97].

Non-Hermitian Lanczos $\rightarrow \hat{\mathbf{T}}_n$ (complex) symmetric. Define

$$c^* \mathbf{A}^{-1} b \;\approx\; G\left(\lambda^{-1}\right) \;\equiv\; \sum_{j=1}^{n} \frac{\omega_j}{\zeta_j} \,,$$

ζ_j ...eigenvalues of $\hat{\mathbf{T}}_n$, ω_j ...scaled and squared first components of the normalized eigenvectors of $\hat{\mathbf{T}}_n$.

Saylor-Smolarski approach

For diagonalizable matrices

[Saylor & Smolarski '01] introduce

- formally orthogonal polynomials,
- complex Gauss quadrature,

as a tool for approximating the quantity $c^* \mathbf{A}^{-1} b$. Motivated by [Freund & Hochbruck '93], [Golub & Meurant '94, '97].

Non-Hermitian Lanczos $\rightarrow \hat{\mathbf{T}}_n$ (complex) symmetric. Define

$$c^* \mathbf{A}^{-1} b \;\approx\; G\left(\lambda^{-1}\right) \;\equiv\; \sum_{j=1}^{n} \frac{\omega_j}{\zeta_j}\,,$$

ζ_j ... eigenvalues of $\hat{\mathbf{T}}_n$, ω_j ... scaled and squared first components of the normalized eigenvectors of $\hat{\mathbf{T}}_n$.

[Warnick '00] showed:

$$G\left(\lambda^{-1}\right) = c^* x_n\,.$$

Yet another approach

Hybrid BiCG methods

We know that

$$c^* \mathbf{A}_n^{-1} b \;=\; \sum_{j=0}^{n-1} \alpha_j \, s_j^* r_j \quad \text{and} \quad s_j^* r_j \;=\; (c^* b) \prod_{k=0}^{j-1} \beta_k \,.$$

Yet another approach

Hybrid BiCG methods

We know that

$$c^*\mathbf{A}_n^{-1}b \;=\; \sum_{j=0}^{n-1} \alpha_j\, s_j^* r_j \quad \text{and} \quad s_j^* r_j \;=\; (c^*b)\prod_{k=0}^{j-1} \beta_k\,.$$

In hybrid BiCG methods like CGS, BiCGStab, BiCGStab(ℓ), the BiCG coefficients are available, i.e. we can compute the approximation $c^*\mathbf{A}_n^{-1}b$ during the run of these method.

Yet another approach
Hybrid BiCG methods

We know that

$$c^* \mathbf{A}_n^{-1} b \;=\; \sum_{j=0}^{n-1} \alpha_j \, s_j^* r_j \quad \text{and} \quad s_j^* r_j \;=\; (c^* b) \prod_{k=0}^{j-1} \beta_k \,.$$

In hybrid BiCG methods like CGS, BiCGStab, BiCGStab(ℓ), the BiCG coefficients are available, i.e. we can compute the approximation $c^* \mathbf{A}_n^{-1} b$ during the run of these method.

Question: Hybrid BiCG methods produce approximations $\mathbf{x}_n$, better than x_n produced by BiCG.

Is $c^* \mathbf{x}_n$ a better approximation of $c^* \mathbf{A}^{-1} b$ than $c^* x_n$?

Yet another approach

Hybrid BiCG methods

We know that

$$c^* \mathbf{A}_n^{-1} b \;=\; \sum_{j=0}^{n-1} \alpha_j \, s_j^* r_j \quad \text{and} \quad s_j^* r_j \;=\; (c^* b) \prod_{k=0}^{j-1} \beta_k \, .$$

In hybrid BiCG methods like CGS, BiCGStab, BiCGStab(ℓ), the BiCG coefficients are available, i.e. we can compute the approximation $c^* \mathbf{A}_n^{-1} b$ during the run of these method.

Question: Hybrid BiCG methods produce approximations $\mathbf{x}_n$, better than x_n produced by BiCG.

Is $c^* \mathbf{x}_n$ a better approximation of $c^* \mathbf{A}^{-1} b$ than $c^* x_n$?

No. We showed that mathematically [Strakoš & T. '09],

$$c^* \mathbf{x}_n = c^* x_n .$$

Summary (non-Hermitian Lanczos approach)

How to compute $c^* \mathbf{A}_n^{-1} b$?

Algorithm of choice:

- non-Hermitian Lanczos
- BiCG
- hybrid BiCG methods

Way of computing the approximation:

- $c^* x_n$
- $(c^* v_1) \, \|b\| \, (\mathbf{T}_n^{-1})_{1,1}$
- complex Gauss quadrature
- from the BiCG coefficients, or, in BiCG using

$$\varepsilon_n^B \equiv \sum_{j=0}^{n-1} \alpha_j \, s_j^* r_j \, .$$

Outline

Numerical experiments

Diffraction of light on periodic structures, RCWA method

[Hench & Strakoš '08]

$$\mathbf{A}\,x \equiv \begin{bmatrix} -\mathbf{I} & \mathbf{I} & e^{\mathbf{i}\sqrt{\mathbf{C}}\varrho} & 0 \\ \mathbf{Y}_I & \sqrt{\mathbf{C}} & -\sqrt{\mathbf{C}}e^{\mathbf{i}\sqrt{C}\varrho} & 0 \\ 0 & e^{\mathbf{i}\sqrt{\mathbf{C}}\varrho} & I & -\mathbf{I} \\ 0 & \sqrt{\mathbf{C}}e^{\mathbf{i}\sqrt{\mathbf{C}}\varrho} & -\sqrt{\mathbf{C}} & -\mathbf{Y}_{\mathrm{II}} \end{bmatrix} x \;=\; b\,,$$

$\mathbf{Y}_{\mathrm{I}},\ \mathbf{Y}_{\mathrm{II}}, \mathbf{C} \in \mathbb{C}^{(2M+1)\times(2M+1)}$, $\varrho > 0$, M is the discretization parameter representing the number of Fourier nodes used for approximation of the electric and magnetic fields as well as the material properties.

Numerical experiments

Diffraction of light on periodic structures, RCWA method

[Hench & Strakoš '08]

$$\mathbf{A}\,x \equiv \begin{bmatrix} -\mathbf{I} & \mathbf{I} & e^{\mathbf{i}\sqrt{\mathbf{C}}\varrho} & 0 \\ \mathbf{Y}_I & \sqrt{\mathbf{C}} & -\sqrt{\mathbf{C}}e^{\mathbf{i}\sqrt{C}\varrho} & 0 \\ 0 & e^{\mathbf{i}\sqrt{\mathbf{C}}\varrho} & I & -\mathbf{I} \\ 0 & \sqrt{\mathbf{C}}e^{\mathbf{i}\sqrt{\mathbf{C}}\varrho} & -\sqrt{\mathbf{C}} & -\mathbf{Y}_{\mathrm{II}} \end{bmatrix} x \; = \; b\,,$$

$\mathbf{Y}_{\mathrm{I}},\ \mathbf{Y}_{\mathrm{II}}, \mathbf{C} \in \mathbb{C}^{(2M+1)\times(2M+1)}$, $\varrho > 0$, M is the discretization parameter representing the number of Fourier nodes used for approximation of the electric and magnetic fields as well as the material properties.

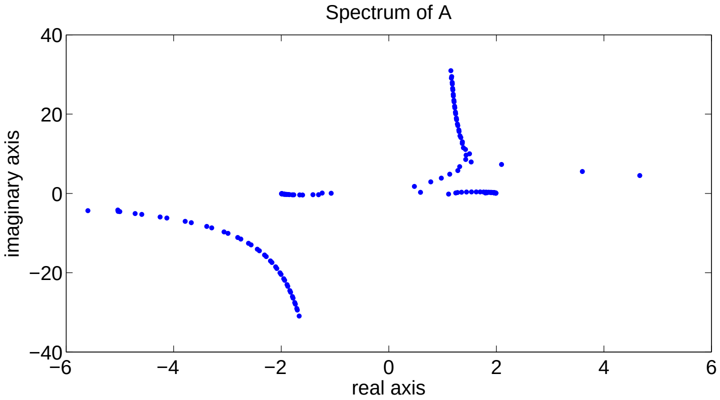
Typically, one needs only the dominant $(M+1)$st component

$$e^*_{M+1}\mathbf{A}^{-1}b.$$

In our experiments $M = 20$, i.e. $\mathbf{A} \in \mathbb{C}^{164\times 164}$. [Strakoš & T. '09]

The matrix $\mathbf{A}$

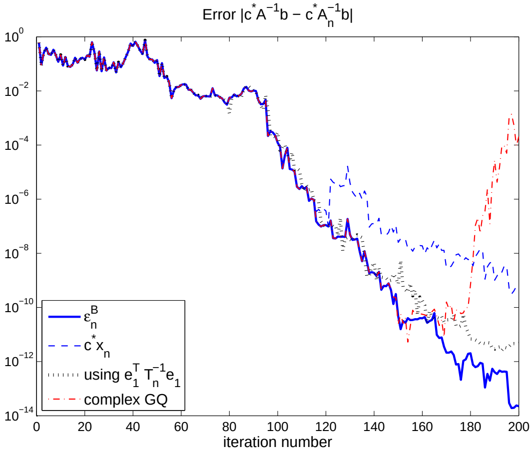
Spectrum of $\mathbf{A}$ computed via the Matlab command `eig`



Some eigenvalues have large imaginary parts
in comparison to the real parts, $\kappa(\mathbf{A}) \approx 104$.

Non-Hermitian Lanczos approach

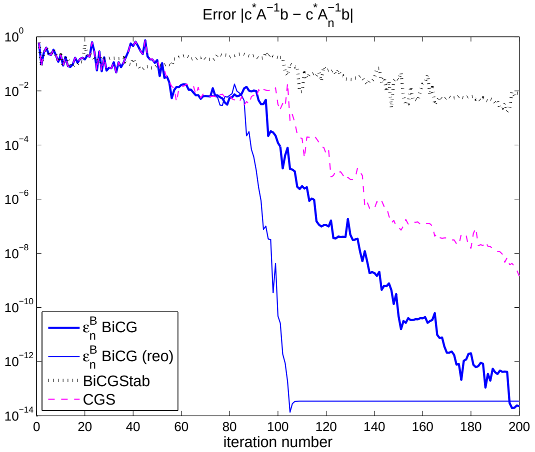
Mathematically equivalent estimates I



Comparison of mathematically equivalent approximations based on BiCG and non-Hermitian Lanczos.

Non-Hermitian Lanczos approach

Mathematically equivalent estimates II



The BiCGStab and CGS approximations are significantly more affected by rounding errors than the BiCG approximations.

Conclusions

- Generalization of the HPD case:
  - Via Vorobyev moment problem $\rightarrow$ very natural and general.
    - no assumptions on $\mathbf{A}$, based on approximation properties
  - Complex Gauss Quadrature approach
    - $\mathbf{A}$ has to be diagonalizable, just a formalism

Conclusions

- Generalization of the HPD case:
  - Via Vorobyev moment problem $\rightarrow$ very natural and general.
    - no assumptions on $\mathbf{A}$, based on approximation properties
  - Complex Gauss Quadrature approach
    - $\mathbf{A}$ has to be diagonalizable, just a formalism
- We proved mathematical equivalence of the existing approximations based on Non-Hermitian Lanczos.

Conclusions

- Generalization of the HPD case:
  - Via Vorobyev moment problem $\rightarrow$ very natural and general.
    - no assumptions on $\mathbf{A}$, based on approximation properties
  - Complex Gauss Quadrature approach
    - $\mathbf{A}$ has to be diagonalizable, just a formalism

- We proved mathematical equivalence of the existing approximations based on Non-Hermitian Lanczos.

- Preferable approximation

$$\varepsilon_n^B \equiv \sum_{j=0}^{n-1} \alpha_j \, s_j^* r_j \, .$$

It is simple and numerically better justified.

Conclusions

- Generalization of the HPD case:
  - Via Vorobyev moment problem $\rightarrow$ very natural and general.
    - no assumptions on $\mathbf{A}$, based on approximation properties
  - Complex Gauss Quadrature approach
    - $\mathbf{A}$ has to be diagonalizable, just a formalism
- We proved mathematical equivalence of the existing approximations based on Non-Hermitian Lanczos.
- Preferable approximation

$$\varepsilon_n^B \equiv \sum_{j=0}^{n-1} \alpha_j \, s_j^* r_j \, .$$

  It is simple and numerically better justified.

- In finite precision arithmetic, the identities need not hold. A justification is needed (e.g. local biorthogonality).

Related papers

- Z. Strakoš and P. Tichý, [On efficient numerical approximation of the scattering amplitude $c^* \mathbf{A}^{-1} b$ via matching moments, submitted to SISC, 2009].
- G. H. Golub, M. Stoll, and A. Wathen, [Approximation of the scattering amplitude and linear systems, Electron. Trans. Numer. Anal., 31 (2008), pp. 178–203].
- Z. Strakoš and P. Tichý, [On error estimation in the conjugate gradient method and why it works in finite precision computations, Electron. Trans. Numer. Anal., 13 (2002), pp. 56–80].
- P. E. Saylor and D. C. Smolarski, [Why Gaussian quadrature in the complex plane?, Numer. Algorithms, 26 (2001), pp. 251–280].
- G. H. Golub and G. Meurant, [Matrices, moments and quadrature, in Numerical analysis 1993 (Dundee, 1993), vol. 303 of Pitman Res. Notes Math. Ser., Longman Sci. Tech., Harlow, 1994, pp. 105–156].

More details

More details can be found at

http://www.cs.cas.cz/~strakos
http://www.cs.cas.cz/~tichy

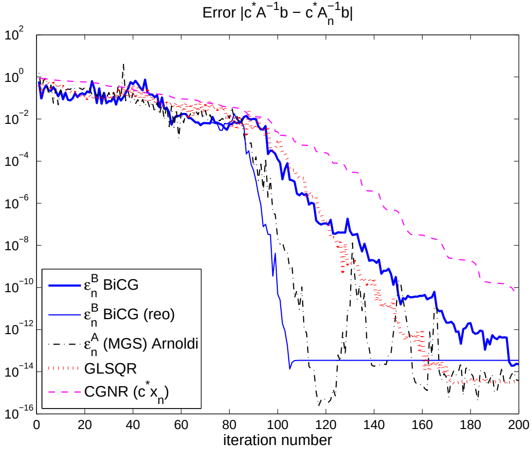
More details

More details can be found at

http://www.cs.cas.cz/~strakos
http://www.cs.cas.cz/~tichy

Thank you for your attention!

Non-Hermitian Lanczos, Arnoldi, GLSQR



GLSQR: [Golub & Stoll & Wathen '08], [Saunders & Simon & Yip '88]

Different approaches with preconditioning

Non-Hermitian Lanczos, Arnoldi, GLSQR