On efficient numerical approximation of the bilinear form $c^*\mathbf{A}^{-1}b$

Petr Tichý

joint work with

Zdeněk Strakoš

Institute of Computer Science AS CR

July 19–20, 2010,
French-Czech Workshop
on Krylov Methods for Inverse Problems
Prague, 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 $e_j^T \mathbf{A}^{-1} b$.
- **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**
- **Nuclear physics, quantum mechanics, other disciplines**

Krylov subspace methods approach

Projection of the original problem onto Krylov subspaces

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

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

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

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

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

- The approximation c^*x_n can be highly inefficient!
  How to approximate c^*x without looking for 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,

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

- The approximation c^*x_n can be highly inefficient!
  How to approximate c^*x without looking for x_n?
- We need a theoretical background
  (find the best possible approximation in some sense).

Krylov subspace methods approach

Projection of the original problem onto Krylov subspaces

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

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

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

- The approximation c^*x_n can be highly inefficient!
  How to approximate c^*x without looking for x_n?
- We need a theoretical background
  (find the best possible approximation in some sense).
- Efficient numerical computation and justification
  of the approximation in finite precision arithmetic.

Outline

1. Symmetric, positive definite case
2. Matching moments
3. Approximation of the bilinear form $c^*\mathbf{A}^{-1}b$
4. Numerical experiments

Outline

1. Symmetric, positive definite case
2. Matching moments
3. Approximation of the bilinear form $c^*\mathbf{A}^{-1}b$
4. Numerical experiments

The CG method

Let $\mathbf{A}$ be symmetric, positive definite

Solve $\mathbf{A}x = b$.

input $\mathbf{A}$, b

$$x_0 = 0$$
$$r_0 = p_0 = b$$

for $k = 0, 1, \ldots$

$$\alpha_k = \frac{\|r_k\|^2}{p_k^* \mathbf{A} p_k},$$
$$x_{k+1} = x_k + \alpha_k p_k,$$
$$r_{k+1} = r_k - \alpha_k \mathbf{A} p_k,$$
$$\beta_{k+1} = \frac{\|r_{k+1}\|^2}{\|r_k\|^2},$$
$$p_{k+1} = r_{k+1} + \beta_{k+1} p_k,$$

end

The Lanczos algorithm

Let $\mathbf{A}$ be symmetric

Compute orthonormal basis of $\mathcal{K}_n(\mathbf{A}, b)$.

input $\mathbf{A}$, b

$$v_1 = b/||b||,\ \delta_1 = 0\,,$$

for $k = 1, 2, \ldots$

$$\gamma_k = v_k^T(\mathbf{A}v_k - \delta_k v_{k-1})\,,$$
$$w = \mathbf{A}v_k - \gamma_k v_k - \delta_k v_{k-1}\,,$$
$$\delta_{k+1} = \|w\|\,,$$
$$v_{k+1} = w/\delta_{k+1}\,,$$

end

The Lanczos algorithm

Let $\mathbf{A}$ be symmetric

Compute orthonormal basis of $\mathcal{K}_n(\mathbf{A}, b)$.

input $\mathbf{A}$, b

$v_1 = b/||b||$, $\delta_1 = 0$,

for $k = 1, 2, \ldots$

$$\gamma_k = v_k^T(\mathbf{A}v_k - \delta_k v_{k-1})\,,$$
$$w = \mathbf{A}v_k - \gamma_k v_k - \delta_k v_{k-1}\,,$$
$$\delta_{k+1} = \|w\|\,,$$
$$v_{k+1} = w/\delta_{k+1}\,,$$

end

The Lanczos algorithm is represented by

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

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

CG versus Lanczos

Let $\mathbf{A}$ be symmetric, positive definite

$$\mathbf{T}_n = \begin{bmatrix} \gamma_1 & \delta_2 & & \\ \delta_2 & \ddots & & \\ & & \ddots & \delta_n \\ & & \delta_n & \gamma_n \end{bmatrix} = \mathbf{L}_n \mathbf{L}_n^T$$

where

$$\mathbf{L}_n = \begin{bmatrix} \frac{1}{\sqrt{\alpha_0}} & & & \\ \sqrt{\frac{\beta_1}{\alpha_0}} & \ddots & & \\ & \ddots & \ddots & \\ & & \frac{\sqrt{\beta_{n-1}}}{\alpha_{n-2}} & \frac{1}{\sqrt{\alpha_{n-1}}} \end{bmatrix}.$$

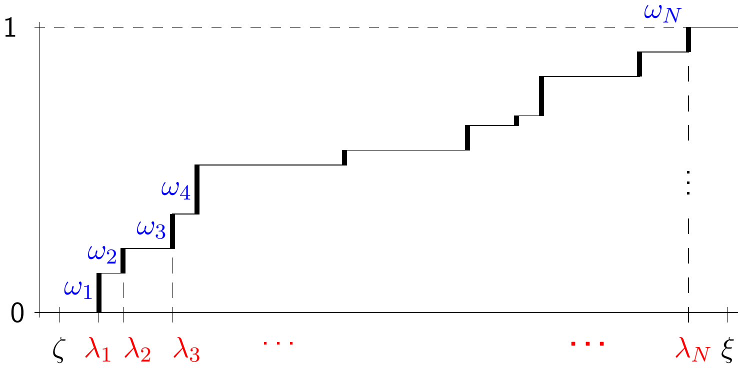
The CG approximation is the given by

$$\mathbf{T}_n y_n = \|b\| e_1, \qquad x_n = x_0 + \mathbf{V}_n y_n \,.$$

Distribution function $\omega(\lambda)$

Without loss of generality $\|b\| = 1$

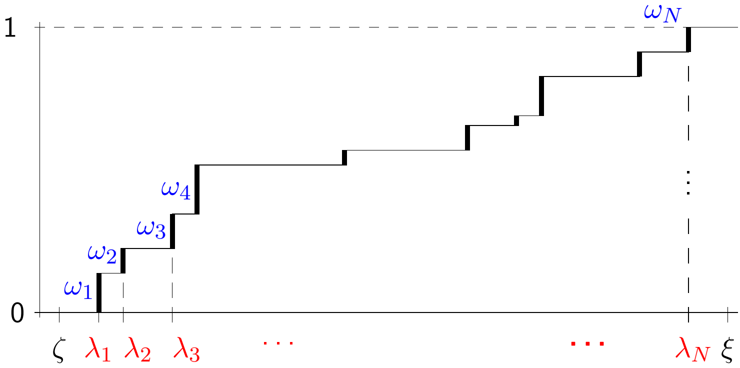
$$(\lambda_i, u_i) \ \ldots \ \text{eigenpair of } \mathbf{A}, \qquad \omega_i = (b^T u_i)^2.$$



Distribution function $\omega(\lambda)$

Without loss of generality $\|b\| = 1$

$$(\lambda_i, u_i) \ \ldots \ \text{eigenpair of } \mathbf{A}, \qquad \omega_i = (b^T u_i)^2.$$



$$\int_{\zeta}^{\xi} f(\lambda)\, d\omega(\lambda) \ = \ \sum_{i=1}^{N} \omega_i \, f(\lambda_i)\,.$$

The Conjugate gradient method and Gauss Quadrature

Symmetric, positive definite case

At any iteration step n, CG (implicitly) determines weights and nodes of the n-point Gauss quadrature

$$\int_{\zeta}^{\xi} f(\lambda)\, d\omega(\lambda) \;=\; \sum_{i=1}^{n} \omega_i^{(n)} f(\theta_i^{(n)}) \;+\; R_n(f)\,.$$

$\mathbf{T}_n$ ... the corresponding Jacobi matrices,
$\theta_i^{(n)}$ ... eigenvalues of $\mathbf{T}_n$, $\omega_i^{(n)}$ ... scaled and squared first components of the normalized eigenvectors of $\mathbf{T}_n$.

At any iteration step n, CG (implicitly) determines weights and nodes of the n-point Gauss quadrature

$$\int_{\zeta}^{\xi} f(\lambda)\, d\omega(\lambda) \;=\; \sum_{i=1}^{n} \omega_i^{(n)} f(\theta_i^{(n)}) \;+\; R_n(f)\,.$$

$\mathbf{T}_n$ ... the corresponding Jacobi matrices,
$\theta_i^{(n)}$ ... eigenvalues of $\mathbf{T}_n$, $\omega_i^{(n)}$ ... scaled and squared first components of the normalized eigenvectors of $\mathbf{T}_n$.

CG matches the first $2n$ moments, $f(\lambda) = \lambda^k$, $k = 0, \ldots, 2n-1$

$$\int_0^{\infty} \lambda^k\, d\omega(\lambda) \;=\; \sum_{i=1}^{n} \omega_i^{(n)}\, (\theta_i^{(n)})^k \;=\; \int_0^{\infty} \lambda^k\, d\omega^{(n)}(\lambda)\,.$$

Moment problem:

$$\omega(\lambda) \quad \rightarrow \quad \omega^{(n)}(\lambda)\,.$$

CG and Gauss Quadrature for $f(\lambda) = \lambda^{-1}$

Symmetric, positive definite case

For $f(\lambda) \equiv \lambda^{-1}$ the formula takes the form

$$\int_{\zeta}^{\xi} \lambda^{-1} \, d\omega(\lambda) \; = \; \sum_{i=1}^{n} \frac{\omega_i^{(n)}}{\theta_i^{(n)}} \; + \; R_n(\lambda^{-1})$$

or, equivalently [Golub & Strakoš '94],

$$\frac{\|x\|_{\mathbf{A}}^2}{\|b\|^2} \; = \; n\text{-th Gauss quadrature} \; + \; \frac{\|x - x_n\|_{\mathbf{A}}^2}{\|b\|^2} \, .$$

CG and Gauss Quadrature for $f(\lambda) = \lambda^{-1}$

Symmetric, positive definite case

For $f(\lambda) \equiv \lambda^{-1}$ the formula takes the form

$$\int_\zeta^\xi \lambda^{-1}\, d\omega(\lambda) \;=\; \sum_{i=1}^{n} \frac{\omega_i^{(n)}}{\theta_i^{(n)}} \;+\; R_n(\lambda^{-1})$$

or, equivalently [Golub & Strakoš '94],

$$\frac{\|x\|_{\mathbf{A}}^2}{\|b\|^2} \;=\; n\text{-th Gauss quadrature} \;+\; \frac{\|x - x_n\|_{\mathbf{A}}^2}{\|b\|^2}\,.$$

We can approximate

$$\|x\|_{\mathbf{A}}^2 \;=\; x^T \mathbf{A} x \;=\; b^T x \;=\; b^T \mathbf{A}^{-1} b$$

using Gauss quadrature.

CG and Gauss Quadrature for $f(\lambda) = \lambda^{-1}$

Mathematically equivalent formulas (multiplied by $\|b\|^2$)

Gauss Quadrature based formula:

$$\|x\|_{\mathbf{A}}^2 \quad = \quad \|b\|^2 \, C_n \; + \; \|x - x_n\|_{\mathbf{A}}^2 \, ,$$

C_n is continued fraction corresponding to $\omega^{(n)}(\lambda)$

[Golub & Strakoš '94, Golub & Meurant '94, '97, '10]

Formulas based on algebraic manipulations

$$\begin{aligned} \|x\|_{\mathbf{A}}^2 &= b^T x_n + \|x - x_n\|_{\mathbf{A}}^2 \\ \|x\|_{\mathbf{A}}^2 &= \sum_{i=0}^{n-1} \alpha_i \|r_i\|^2 + \|x - x_j\|_{\mathbf{A}}^2 . \end{aligned}$$

The first one derived by [Warnick '00], the second one independently by [Hestenes & Stiefel '52, Deufelhard '93, Axelsson & Kaporin '01, Strakoš & T. '02]

CG and the approximation of $b^T \mathbf{A}^{-1} b$

Mathematically equivalent approximations

Approximation based on the formula

$$\|x\|_{\mathbf{A}}^2 \;=\; \|b\|^2 \; n\text{-th Gauss quadrature} \;+\; \|x - x_n\|_{\mathbf{A}}^2 \,.$$

If $\|x - x_n\|_{\mathbf{A}}^2$ is small then

$$b^T \mathbf{A}^{-1} b \;\approx\; \|b\|^2 \; n\text{-th Gauss quadrature}$$

CG and the approximation of $b^T \mathbf{A}^{-1} b$

Mathematically equivalent approximations

Approximation based on the formula

$$\|x\|_{\mathbf{A}}^2 \;=\; \|b\|^2 \; n\text{-th Gauss quadrature} \;+\; \|x - x_n\|_{\mathbf{A}}^2 \,.$$

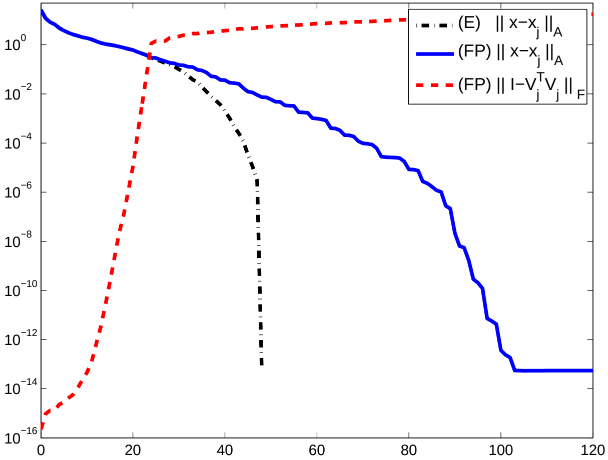
If $\|x - x_n\|_{\mathbf{A}}^2$ is small then

$$b^T \mathbf{A}^{-1} b \;\approx\; \|b\|^2 \; n\text{-th Gauss quadrature}$$

Mathematically equivalent approximations:

$$\|b\|^2 \, C_n, \qquad b^T x_n \qquad \text{and} \qquad \sum_{i=0}^{n-1} \alpha_i \|r_i\|^2 .$$

Orthogonality is lost, convergence is delayed!



Relations need not hold in finite precision arithmetic!

Rounding error analysis
Strakoš & T. 2002

Do the relations hold for computed quantities?

1.
$$\|x\|_{\mathbf{A}}^2 \quad = \quad b^T x_n \;+\; \|x - x_n\|_{\mathbf{A}}^2$$

does not hold for computed quantities - its validity is based on preserving global orthogonality among CG residuals.

Rounding error analysis

Strakoš & T. 2002

Do the relations hold for computed quantities?

1.
$$\|x\|_{\mathbf{A}}^2 \quad = \quad b^T x_n \; + \; \|x - x_n\|_{\mathbf{A}}^2$$

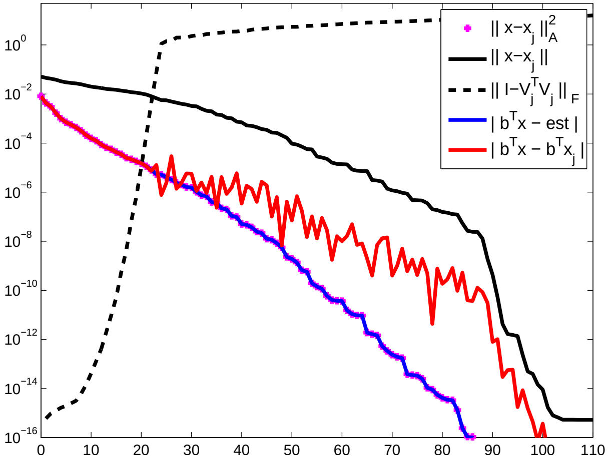
   does not hold for computed quantities - its validity is based on preserving global orthogonality among CG residuals.

2.
$$\|x\|_{\mathbf{A}}^2 \quad = \quad \sum_{i=0}^{n-1} \alpha_i \|r_i\|^2 \; + \; \|x - x_n\|_{\mathbf{A}}^2 .$$

   holds also for computed quantities - it is based on preserving local orthogonality between r_{n+1} and p_n.

Behavior in finite precision arithmetic

$b^T x_n$ versus $\sum_{i=0}^{n-1} \alpha_i \|r_i\|^2$



Theoretical background: Gauss quadrature

$$\frac{b^T \mathbf{A}^{-1} b}{\|b\|^2} \;=\; n\text{-th Gauss quadrature} \;+\; \frac{\|x - x_n\|_{\mathbf{A}}^2}{\|b\|^2}\,.$$

Theoretical background: Gauss quadrature

$$\frac{b^T \mathbf{A}^{-1} b}{\|b\|^2} \; = \; n\text{-th Gauss quadrature} \; + \; \frac{\|x - x_n\|_{\mathbf{A}}^2}{\|b\|^2} \, .$$

If $c = b$, the best way how to approximate $b^T \mathbf{A}^{-1} b$ is to use the Hestenes-Stiefel estimate

$$b^T \mathbf{A}^{-1} b \; \approx \; \sum_{i=0}^{n-1} \alpha_i \|r_i\|^2 \, .$$

Symmetric, positive definite case

Summary

Theoretical background: Gauss quadrature

$$\frac{b^T \mathbf{A}^{-1} b}{\|b\|^2} \; = \; n\text{-th Gauss quadrature} \; + \; \frac{\|x - x_n\|_{\mathbf{A}}^2}{\|b\|^2} \, .$$

If $c = b$, the best way how to approximate $b^T \mathbf{A}^{-1} b$ is to use the Hestenes-Stiefel estimate

$$b^T \mathbf{A}^{-1} b \;\approx\; \sum_{i=0}^{n-1} \alpha_i \|r_i\|^2 \, .$$

- We have seen that due to numerical instabilities, the explicit numerical computation of $c^* x_n$ can be highly inefficient.
  [Strakoš & T. '02, '05]

- How to generalize ideas from the SPD case to a general case?

Outline

CG, Gauss Quadrature and Matching Moments

Overview

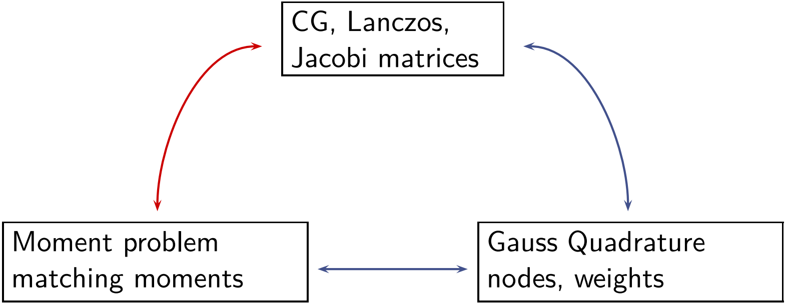
CG, Lanczos,
Jacobi matrices

Moment problem
matching moments

Gauss Quadrature
nodes, weights

CG, Gauss Quadrature and Matching Moments

Overview



Matching moments

Matrix formulation, without loss of generality $\|b\| = 1$

How to express moments in terms of $\mathbf{A}$, b and $\mathbf{T}_n$?

$$\int_0^\infty \lambda^k \, d\omega(\lambda) = \sum_{i=1}^{N} \omega_j \, (\lambda_j)^k = b^* \, \mathbf{A}^k \, b \,,$$

$$\int_0^\infty \lambda^k \, d\omega^{(n)}(\lambda) = \sum_{i=1}^{n} \omega_i^{(n)} \, (\theta_i^{(n)})^k = e_1^T \, \mathbf{T}_n^k \, e_1 \,.$$

Matching moments

Matrix formulation, without loss of generality $\|b\| = 1$

How to express moments in terms of $\mathbf{A}$, b and $\mathbf{T}_n$?

$$\begin{aligned} \int_0^\infty \lambda^k \, d\omega(\lambda) &= \sum_{i=1}^{N} \omega_j \, (\lambda_j)^k &= b^* \, \mathbf{A}^k \, b \,, \\ \int_0^\infty \lambda^k \, d\omega^{(n)}(\lambda) &= \sum_{i=1}^{n} \omega_i^{(n)} \, (\theta_i^{(n)})^k &= e_1^T \, \mathbf{T}_n^k \, e_1 \,. \end{aligned}$$

Matching the first $2n$ moments therefore means

$$b^* \, \mathbf{A}^k \, b \equiv e_1^T \, \mathbf{T}_n^k \, e_1 \,, \quad k = 0, 1, \ldots, 2n-1 \,.$$

Let $\|b\| = 1$.

CG (Lanczos) reduces for $\mathbf{A}$ HPD at the step n the original model

$$\mathbf{A}x = b \qquad \text{to} \qquad \mathbf{T}_n y_n = e_1$$

Let $\|b\| = 1$.

CG (Lanczos) reduces for $\mathbf{A}$ HPD at the step n the original model

$$\mathbf{A}x = b \qquad \text{to} \qquad \mathbf{T}_n y_n = e_1$$

such that $2n$ moments are matched,

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

The Vorobyev moment problem

Vorobyev '58, '65, popularized by Brezinski '97, Strakoš '08

Find a linear HPD 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 $\mathbf{Q}_n$ projects onto $\mathcal{K}_n(\mathbf{A}, b)$ orthogonally to $\mathcal{K}_n(\mathbf{A}, b)$.

The Vorobyev moment problem

Vorobyev '58, '65, popularized by Brezinski '97, Strakoš '08

Find a linear HPD 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 $\mathbf{Q}_n$ projects onto $\mathcal{K}_n(\mathbf{A}, b)$ orthogonally to $\mathcal{K}_n(\mathbf{A}, b)$.

Moment problem:

$$\omega(\lambda) \quad \rightarrow \quad \omega^{(n)}(\lambda)\,.$$

Vorobyev moment problem:

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

Lanczos and the Vorobyev moment problem

Model reduction via matching moments

Let $\mathbf{V}_n$ and $\mathbf{T}_n$ are matrices from the Lanczos algorithm. Then

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

We can identify Lanczos with the Vorobyev moment problem.

Lanczos and the Vorobyev moment problem
Model reduction via matching moments

Let $\mathbf{V}_n$ and $\mathbf{T}_n$ are matrices from the Lanczos algorithm. Then

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

We can identify Lanczos with the Vorobyev moment problem.

Using the Vorobyev moment problem one can show [Strakoš '08]

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

The matching moment property of Lanczos (CG) can be shown without using Gauss Quadrature!

Lanczos and the Vorobyev moment problem

Model reduction via matching moments

Let $\mathbf{V}_n$ and $\mathbf{T}_n$ are matrices from the Lanczos algorithm. Then

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

We can identify Lanczos with the Vorobyev moment problem.

Using the Vorobyev moment problem one can show [Strakoš '08]

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

The matching moment property of Lanczos (CG) can be shown without using Gauss Quadrature!

This view of Krylov subspace methods appears to be useful when generalizing the ideas from the HPD case.

Vorobyev moment problem

General case

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

General case

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 identified with 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

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

Theoretical background - general framework, Strakoš & T. '09

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

Define approximation:

$$c^*\mathbf{A}^{-1}b \;\approx\; c^*\mathbf{A}_n^{-1}b$$

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

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 $\mathcal{K}_n(\mathbf{A}, b)$,

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

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

Define approximation:

$$c^*\mathbf{A}^{-1}b \;\approx\; c^*\mathbf{A}_n^{-1}b$$

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

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

Questions:

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

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

Theoretical background - general framework, Strakoš & T. '09

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

Define approximation:

$$c^* \mathbf{A}^{-1} b \;\approx\; c^* \mathbf{A}_n^{-1} b$$

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

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

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

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,

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

We used the global biorthogonality !

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 that (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 that (using global biorthogonality) that

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

Finally,

$$c^*\mathbf{A}_n^{-1}b \;=\; c^*x_n \;=\; \sum_{j=0}^{n-1}\alpha_j s_j^* r_j \;\equiv\; \xi_n^B\,.$$

Approximations based on the BiCG method

and possible troubles in finite precision arithmetic

It holds that

$$c^*\mathbf{A}^{-1}b \;=\; \sum_{j=0}^{n-1}\alpha_j s_j^* r_j \;+\; \underbrace{s_n^*\mathbf{A}^{-1}r_n}_{\text{error} \;\sim\; \|y-y_n\|\|r_n\|} \,.$$

It can be shown that

$$c^*\mathbf{A}^{-1}b \;=\; c^* x_n \;+\; \underbrace{y_n^* r_n \;+\; s_n^*\mathbf{A}^{-1}r_n}_{\text{error} \;\sim\; \|y_n\|\|r_n\|} \,.$$

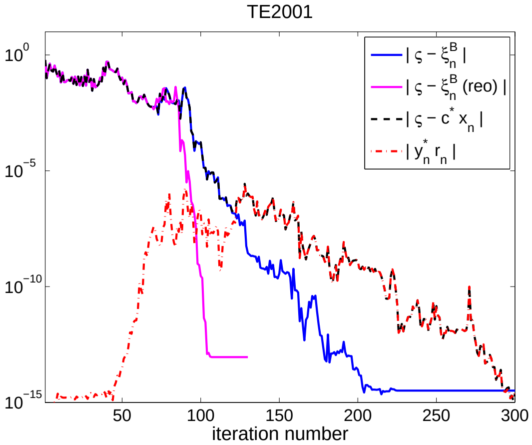
In exact arithmetic $y_n^* r_n = 0$.

If the global biorthogonality is lost, one can expect that

$$|y_n^* r_n| \;\sim\; \|y_n\|\|r_n\| \,.$$

Approximations based on the BiCG method

Mathematically equivalent approximations ξ_n^B and $c^* x_n$, $\varsigma \equiv c^* \mathbf{A}^{-1} b$



$$\begin{aligned} |c^* \mathbf{A}^{-1} b - c^* x_n| &\approxeq \left| y_n^* r_n + s_n^* \mathbf{A}^{-1} r_n \right|, \\ |c^* \mathbf{A}^{-1} b - \xi_n^B| &\approxeq \left| s_n^* \mathbf{A}^{-1} r_n \right|. \end{aligned}$$

Yet another approach

Hybrid BiCG methods

We know that

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

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 $c^* x_n$?

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:

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

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

Let $\mathbf{P}_{L}$ and $\mathbf{P}_{R}$ be a left and a right preconditioner. Then

$$c^*\mathbf{A}^{-1}b = (\underbrace{\mathbf{P}_{R}^{-*}c}_{\hat{c}})^* \underbrace{(\mathbf{P}_{L}^{-1}\mathbf{A}\mathbf{P}_{R}^{-1})^{-1}}_{\hat{\mathbf{A}}^{-1}} (\underbrace{\mathbf{P}_{L}^{-1}b}_{\hat{b}}).$$

The approximation techniques can be applied to the problem

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

It is obvious that $\hat{\mathbf{A}}$ need not be formed explicitly.

Preconditioning
General case

Let $\mathbf{P}_L$ and $\mathbf{P}_R$ be a left and a right preconditioner. Then

$$c^*\mathbf{A}^{-1}b = (\underbrace{\mathbf{P}_R^{-*}c}_{\hat{c}})^* \underbrace{(\mathbf{P}_L^{-1}\mathbf{A}\mathbf{P}_R^{-1})^{-1}}_{\hat{\mathbf{A}}^{-1}} (\underbrace{\mathbf{P}_L^{-1}b}_{\hat{b}}).$$

The approximation techniques can be applied to the problem

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

It is obvious that $\hat{\mathbf{A}}$ need not be formed explicitly.

It is easier to derive the preconditioned algorithm
for approximating the bilinear form $c^*\mathbf{A}^{-1}b$ than
the preconditioned algorithm for solving linear systems.

Theoretical background: Model reduction via matching moments.

Several Krylov subspace methods (Lanczos, Arnoldi) can be identified with the Vorobyev moment problem $\mathbf{A} \quad \rightarrow \quad \mathbf{A}_n$.

Approximation:

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

Theoretical background: Model reduction via matching moments.

Several Krylov subspace methods (Lanczos, Arnoldi) can be identified with the Vorobyev moment problem $\mathbf{A} \quad \rightarrow \quad \mathbf{A}_n$.

Approximation:

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

Promising approaches:

BiCG and $$c^* \mathbf{A}^{-1} b \,\approx\, \sum_{j=0}^{n-1} \alpha_j s_j^* r_j \,,$$

Arnoldi and $$c^* \mathbf{A}^{-1} b \,\approx\, \|b\| \, t_n^* \mathbf{H}_n^{-1} e_1 \,,$$

where $t_n = \mathbf{V}_n^* c$.

Outline

[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 modes used for approximation of the electric and magnetic fields as well as the material properties.

[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 modes used for approximation of the electric and magnetic fields as well as the material properties.

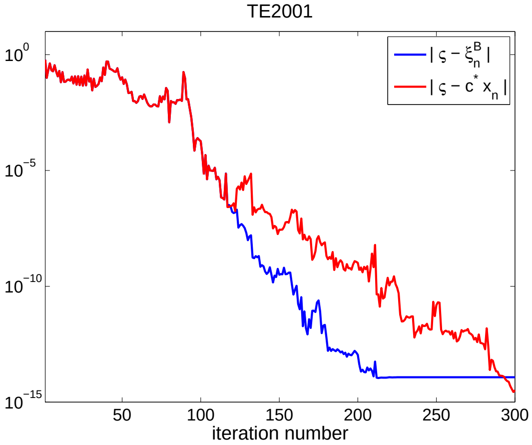
Typically, one needs only the dominant $(M+1)$st component

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

In our experiments $M = 20$, i.e. $\mathbf{A} \in \mathbb{C}^{164\times164}$. [Strakoš & T. '10]

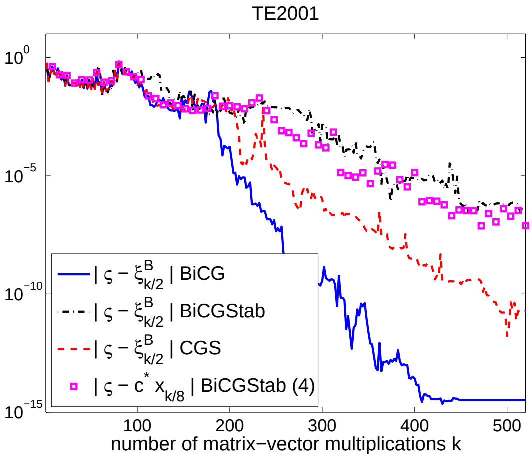
Approximations based on the BiCG method

$$b^T x_n \qquad \text{versus} \qquad \sum_{j=0}^{n-1} \alpha_j s_j^* r_j$$



Non-Hermitian Lanczos approach

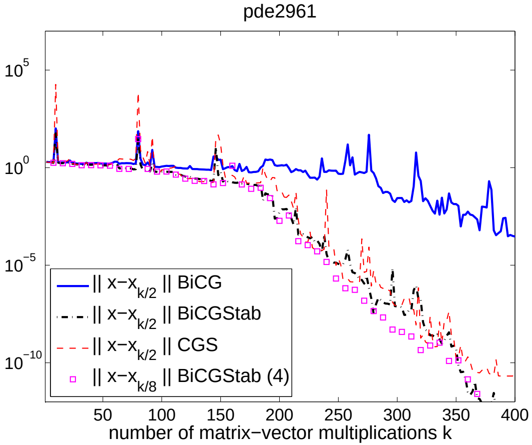
Mathematically equivalent approximations based on hybrid BiCG methods



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

Non-Hermitian Lanczos approach

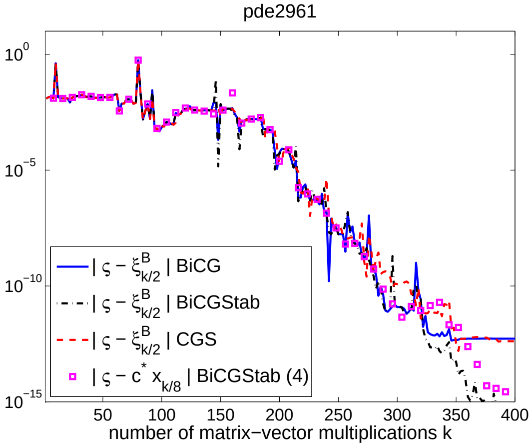
Solving the system $\mathbf{A}x = b$



Hybrid BiCG methods can be more efficient than BiCG when approximating the solution of $\mathbf{A}x = b$.

Non-Hermitian Lanczos approach

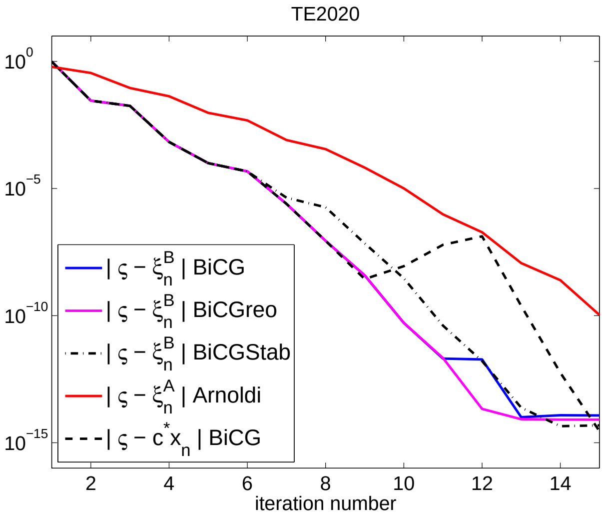
Mathematically equivalent approximations based on hybrid BiCG methods



BiCG is usually more efficient than hybrid BiCG methods when approximation the bilinear form $c^*\mathbf{A}^{-1}b$.

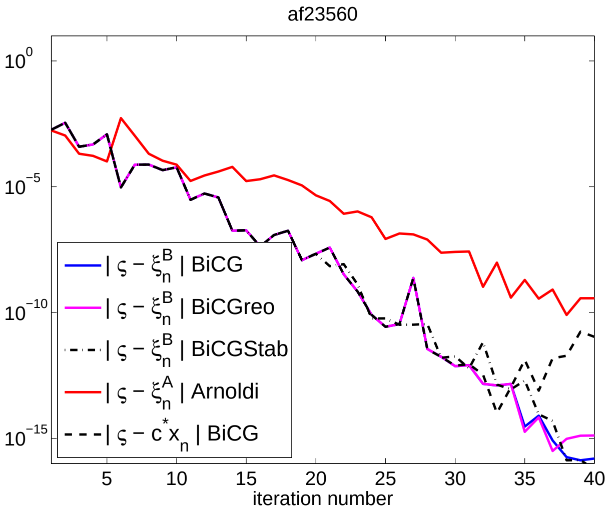
Different approaches with preconditioning

TE polarization, 20 slabs, $A \in \mathbb{C}^{1722\times 1722}$



Different approaches with preconditioning

AF23560: from set AIRFOIL, from the NEP Collection



Conclusions

- Some Krylov subspace methods can be seen as model reduction via matching moments.

Conclusions

- Some Krylov subspace methods can be seen as model reduction via matching moments.
- Generalization of the HPD case: Via Vorobyev moment problem $\rightarrow$ very natural and general.
  - no assumptions on $\mathbf{A}$, based on approximation properties

Conclusions

- Some Krylov subspace methods can be seen as model reduction via matching moments.
- Generalization of the HPD case: Via Vorobyev moment problem $\rightarrow$ very natural and general.
  - no assumptions on $\mathbf{A}$, based on approximation properties
- We proved mathematical equivalence of the existing approximations based on Non-Hermitian Lanczos.

Conclusions

- Some Krylov subspace methods can be seen as model reduction via matching moments.
- Generalization of the HPD case: Via Vorobyev moment problem $\rightarrow$ very natural and general.
  - no assumptions on $\mathbf{A}$, based on approximation properties
- We proved mathematical equivalence of the existing approximations based on Non-Hermitian Lanczos.
- Preferable approximation

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

It is simple and numerically better justified.

Conclusions

- Some Krylov subspace methods can be seen as model reduction via matching moments.
- Generalization of the HPD case: Via Vorobyev moment problem $\rightarrow$ very natural and general.
  - no assumptions on $\mathbf{A}$, based on approximation properties
- We proved mathematical equivalence of the existing approximations based on Non-Hermitian Lanczos.
- Preferable approximation

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

  It is simple and numerically better justified.
- In finite precision arithmetic, the relations need not hold. A justification is necessary (e.g. local biorthogonality).

Related papers

- Z. Strakoš and P. Tichý, [On efficient numerical approximation of the bilinear form $c^* \mathbf{A}^{-1} b$, submitted to SISC, 2010].
- 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].
- 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].

Recent book by G. H. Golub and G. Meurant, [Matrices, Moments and Quadrature With Applications, Princeton University Press, USA, 2010].

More details

More details can be found at

http://www.karlin.mff.cuni.cz/~strakos/
http://www.cs.cas.cz/tichy

More details

More details can be found at

http://www.karlin.mff.cuni.cz/~strakos/
http://www.cs.cas.cz/tichy

Thank you for your attention!