On the influence of eigenvalues on Bi-CG residual norms

Jurjen Duintjer Tebbens

Institute of Computer Science

Academy of Sciences of the Czech Republic

duintjertebbens@cs.cas.cz

Gérard Meurant

30, rue du sergent Bauchat

75012 Paris, France.

gerard.meurant@gmail.com

4th IMA Conference on Numerical Linear Algebra and Optimisation
University of Birmingham, 4.9.2014

Outline

Introduction: The Problem

We consider the solution of linear systems

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

where $\mathbf{A} \in \mathbb{C}^{n \times n}$ is **non-normal and nonsingular**, by a Krylov subspace method.

Introduction: The Problem

We consider the solution of linear systems

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

where $\mathbf{A} \in \mathbb{C}^{n \times n}$ is **non-normal and nonsingular**, by a Krylov subspace method.

- On the one hand there are methods with relatively nice theoretical properties (orthogonal bases) using long recurrences like the GMRES method [Saad & Schultz 1986] or the FOM method [Arnoldi 1951]. In practice they need to be restarted.

Introduction: The Problem

We consider the solution of linear systems

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

where $\mathbf{A} \in \mathbb{C}^{n \times n}$ is **non-normal and nonsingular**, by a Krylov subspace method.

- On the one hand there are methods with relatively nice theoretical properties (orthogonal bases) using long recurrences like the GMRES method [Saad & Schultz 1986] or the FOM method [Arnoldi 1951]. In practice they need to be restarted.
- On the other hand, there are methods like Bi-CG [Lanczos 1952, Fletcher 1974], QMR [Freund & Nachtigal 1991] and Bi-CGStab [van der Vorst 1992] with constant costs per iteration based on less natural projection processes. They can break down without having found the solution.

Introduction: Eigenvalues govern convergence?

The convergence behavior of both types of methods is in practice often governed by the eigenvalue distribution. For instance, eigenvalues clustered around the origin seem to guarantee fast convergence; this is what is generally used to analyse the quality of preconditioners.

Introduction: Eigenvalues govern convergence?

The convergence behavior of both types of methods is in practice often governed by the eigenvalue distribution. For instance, eigenvalues clustered around the origin seem to guarantee fast convergence; this is what is generally used to analyse the quality of preconditioners.

There is, however, no sound theoretical explanation for the relation between eigenvalue distribution and convergence speed. On the contrary, there are theoretical results which imply there is, in general, no relation at all.

Introduction: Eigenvalues govern convergence?

The convergence behavior of both types of methods is in practice often governed by the eigenvalue distribution. For instance, eigenvalues clustered around the origin seem to guarantee fast convergence; this is what is generally used to analyse the quality of preconditioners.

There is, however, no sound theoretical explanation for the relation between eigenvalue distribution and convergence speed. On the contrary, there are theoretical results which imply there is, in general, no relation at all.

The most convincing results showing that the GMRES method need not be governed by eigenvalues alone appeared in a series of papers by Arioli, Greenbaum, Pták and Strakoš [Greenbaum & Strakoš 1994, Greenbaum & Pták & Strakoš 1996, Arioli & Pták & Strakoš 1998]:

Introduction: Eigenvalues govern convergence?

Theorem 1 **[Greenbaum & Pták & Strakoš 1996]** *Let*

$$\|b\| = f_0 \geq f_1 \geq f_2 \cdots \geq f_{n-1} > 0$$

be **any** *non-increasing sequence of real positive values and let*

$$\lambda_1, \ldots, \lambda_n$$

be **any** *set of nonzero complex numbers. Then there exists a class of matrices $A \in \mathbb{C}^{n \times n}$ and right-hand sides $b \in \mathbb{C}^n$ such that the residual vectors r_k generated by the GMRES method applied to A and b satisfy*

$$\|r_k\| = f_k, \quad 0 \leq k \leq n, \quad \text{and} \quad \text{spectrum}(A) = \{\lambda_1, \ldots, \lambda_n\}.$$

Introduction: Eigenvalues govern convergence?

For assessing the quality of a preconditioner $\mathcal{P}$ when GMRES is applied to

$$\mathcal{P}Ax = \mathcal{P}b, \qquad \mathcal{P}A \quad \text{non-symmetric},$$

this means that analysis of the spectrum of $\mathcal{P}A$ alone is not enough.

Introduction: Eigenvalues govern convergence?

For assessing the quality of a preconditioner $\mathcal{P}$ when GMRES is applied to

$$\mathcal{P}Ax = \mathcal{P}b, \qquad \mathcal{P}A \quad \text{non-symmetric},$$

this means that analysis of the spectrum of $\mathcal{P}A$ alone is not enough.

For instance in constraint preconditioning, the fact that spec($\mathcal{P}A$) is, say,

$$\text{spec } (\mathcal{P}A) = \{1, \frac{1}{2} \pm \frac{\sqrt{(5)}}{2}\}$$

does not suffice to guarantee fast convergence of GMRES when $\mathcal{P}A$ is non-symmetric. What is needed additionally, is the fact that the eigenvalues have maximal geometric multiplicity (i.e. that $\mathcal{P}A$ is non-derogatory).

Introduction: Eigenvalues govern convergence?

For assessing the quality of a preconditioner $\mathcal{P}$ when GMRES is applied to

$$\mathcal{P}Ax = \mathcal{P}b, \qquad \mathcal{P}A \quad \text{non-symmetric},$$

this means that analysis of the spectrum of $\mathcal{P}A$ alone is not enough.

For instance in constraint preconditioning, the fact that spec($\mathcal{P}A$) is, say,

$$\text{spec } (\mathcal{P}A) = \{1, \frac{1}{2} \pm \frac{\sqrt{(5)}}{2}\}$$

does not suffice to guarantee fast convergence of GMRES when $\mathcal{P}A$ is non-symmetric. What is needed additionally, is the fact that the eigenvalues have maximal geometric multiplicity (i.e. that $\mathcal{P}A$ is non-derogatory).

The main problem with possibly non-normal input matrices is that besides eigenvalues, the eigenspaces can strongly influence residual norms because they are not orthogonal to eachother.

Introduction: Objects governing convergence?

Tools other than eigenvalues used to explain GMRES convergence include:

- the pseudo-spectrum (e.g. [Trefethen & Embree 2005])
- the field of values (e.g. [Eiermann 1993])
- the polynomial numerical hull (e.g. [Greenbaum 2002])
- decomposition in normal plus small rank (e.g. [Huhtanen & Nevanlinna 2000])

Introduction: Objects governing convergence?

Tools other than eigenvalues used to explain GMRES convergence include:

- the pseudo-spectrum (e.g. [Trefethen & Embree 2005])
- the field of values (e.g. [Eiermann 1993])
- the polynomial numerical hull (e.g. [Greenbaum 2002])
- decomposition in normal plus small rank (e.g. [Huhtanen & Nevanlinna 2000])

One may wonder whether the strong potential independence between convergence behavior and eigenvalues is just an artefact of the GMRES method ?

Introduction: Objects governing convergence?

Tools other than eigenvalues used to explain GMRES convergence include:

- the pseudo-spectrum (e.g. [Trefethen & Embree 2005])
- the field of values (e.g. [Eiermann 1993])
- the polynomial numerical hull (e.g. [Greenbaum 2002])
- decomposition in normal plus small rank (e.g. [Huhtanen & Nevanlinna 2000])

One may wonder whether the strong potential independence between convergence behavior and eigenvalues is just an artefact of the GMRES method ?

For *restarted* GMRES and FOM, the same independence has been proved [Vecharinsky & Langou 2011, Schweitzer 2014?].

The aim of the talk

As all Krylov subspace methods for non-symmetric matrices (e.g. GMRES, FOM, Bi-CG, QMR, TFQMR, CGS, Bi-CGStab, IDR) project in different ways onto essentially the same Krylov subspaces, one may expect that similar results are possible for short recurrence Krylov subspace methods. In this talk,

- we concentrate on the theoretically simplest method with short recurrences, the Bi-CG method.
- we will try to show whether arbitrary convergence curves can be combined with arbitrary eigenvalues in the Bi-CG method. We know the answer in nearly all cases.
- if possible, we will try to show how linear systems can be constructed generating prescribed Bi-CG residual norms with input matrices having prescribed spectrum.

Outline

The GMRES and FOM methods

In GMRES and FOM an *orthogonal* basis for $\mathcal{K}_k(\mathbf{A}, b)$ is constructed with the Arnoldi process. In the kth iteration it computes (when there is no breakdown) the decomposition

$$\mathbf{A}V_k = V_k H_k + h_{k+1,k} v_{k+1} e_1^T = V_{k+1} \tilde{H}_k,$$

where the columns of $V_k = [v_1, \ldots, v_k]$ (the Arnoldi vectors) contain an orthogonal basis for the kth Krylov subspace,

$$\mathcal{K}_k(\mathbf{A}, b) \equiv \operatorname{span}\{b, \mathbf{A}b, \ldots, \mathbf{A}^{k-1}b\}\,.$$

The GMRES and FOM methods

In GMRES and FOM an *orthogonal* basis for $\mathcal{K}_k(\mathbf{A}, b)$ is constructed with the Arnoldi process. In the kth iteration it computes (when there is no breakdown) the decomposition

$$\mathbf{A}V_k = V_k H_k + h_{k+1,k} v_{k+1} e_1^T = V_{k+1} \tilde{H}_k,$$

where the columns of $V_k = [v_1, \dots, v_k]$ (the Arnoldi vectors) contain an orthogonal basis for the kth Krylov subspace,

$$\mathcal{K}_k(\mathbf{A}, b) \equiv \operatorname{span}\{b, \mathbf{A}b, \dots, \mathbf{A}^{k-1}b\}\,.$$

H_k (square), resp. $\tilde{H}_k$ (rectangular) are upper Hessenberg matrices containing the coefficients of the long recurrences;

$$\tilde{H}_k = \begin{bmatrix} & H_k \\ 0 & h_{k+1,k} \end{bmatrix} \in \mathbb{C}^{(k+1)\times k}.$$

The GMRES method

With initial guess $x_0 = 0$,

- GMRES iterates are given by

$$x_k^G = V_k y_k, \qquad y_k = \min_{y \in \mathbb{C}^k} \left\| \|b\| e_1 - \tilde{H}_k y \right\|$$

- FOM iterates are given by

$$x_k^F = V_k y_k, \qquad y_k = H_k^{-1} \|b\| e_1$$

  if H_k is nonsingular; otherwise they are not defined.

The GMRES method

With initial guess $x_0 = 0$,

- GMRES iterates are given by

$$x_k^G = V_k y_k, \qquad y_k = \min_{y \in \mathbb{C}^k} \left\| \|b\| e_1 - \tilde{H}_k y \right\|$$

- FOM iterates are given by

$$x_k^F = V_k y_k, \qquad y_k = H_k^{-1} \|b\| e_1$$

  if H_k is nonsingular; otherwise they are not defined.

Thus iterates and residual norms generated by GMRES and FOM are fully determined by the Hessenberg matrices $\tilde{H}_k$, H_k and $\|b\|$. Prescribing GMRES and FOM residual norms amounts to prescribing the entries of these Hessenberg matrices in the right way.

The GMRES method

The kth GMRES residual vector can be characterized through

$$r_k = \min_{x \in \mathcal{K}_k(\mathbf{A},b)} \|b - \mathbf{A}x\|, \quad \text{equivalently,} \quad r_k^G \perp \mathbf{A}\mathcal{K}_k(\mathbf{A},b),$$

whereas the kth FOM residual vector is characterized through

$$r_k^F \perp \mathcal{K}_k(\mathbf{A},b).$$

The GMRES method

The kth GMRES residual vector can be characterized through

$$r_k = \min_{x \in \mathcal{K}_k(\mathbf{A},b)} \|b - \mathbf{A}x\|, \quad \text{equivalently,} \quad r_k^G \perp \mathbf{A}\mathcal{K}_k(\mathbf{A}, b),$$

whereas the kth FOM residual vector is characterized through

$$r_k^F \perp \mathcal{K}_k(\mathbf{A}, b).$$

The corresponding residual norms are related through to formula

$$\frac{1}{\|r_k^F\|} = \sqrt{\frac{1}{\|r_k^G\|^2} - \frac{1}{\|r_{k-1}^G\|^2}}.$$

Note that FOM residual norms need not be non-increasing and are not defined if the corresponding GMRES iterate stagnates.

The parametrization for FOM and GMRES

Here is how one can prescribe residual norms, eigenvalues and Ritz values in FOM and GMRES [DT & Meurant 2013]:

The parametrization for FOM and GMRES

Here is how one can prescribe residual norms, eigenvalues and Ritz values in FOM and GMRES [DT & Meurant 2013]:

- Choose a unitary matrix V and put $b = Ve_1$ and

$$A = VHV^*, \qquad H \text{ upper Hessenberg.}$$

The parametrization for FOM and GMRES

Here is how one can prescribe residual norms, eigenvalues and Ritz values in FOM and GMRES [DT & Meurant 2013]:

- Choose a unitary matrix V and put $b = Ve_1$ and

$$A = VHV^*, \qquad H \text{ upper Hessenberg.}$$

- To force the desired eigenvalues, H is of the form

$$H = U^{-1}CU, \qquad U \text{ nonsingular upper triangular,}$$

where C is the companion matrix for the prescribed spectrum.

The parametrization for FOM and GMRES

Here is how one can prescribe residual norms, eigenvalues and Ritz values in FOM and GMRES [DT & Meurant 2013]:

- Choose a unitary matrix V and put $b = Ve_1$ and

$$A = VHV^*, \qquad H \text{ upper Hessenberg.}$$

- To force the desired eigenvalues, H is of the form

$$H = U^{-1}CU, \qquad U \text{ nonsingular upper triangular,}$$

  where C is the companion matrix for the prescribed spectrum.
- To force FOM residual norms $f(0), \ldots, f(n-1)$, $f(i) > 0$, the first row g^T of U can be chosen as

$$g_k = \frac{1}{f(k-1)}, \quad k = 1, \ldots, n.$$

The parametrization for FOM and GMRES

Let

$$A = V(U^{-1}CU)V^*, \qquad b = Ve_1.$$

The parametrization for FOM and GMRES

Let

$$A = V(U^{-1}CU)V^*, \qquad b = Ve_1.$$

- To force GMRES residual norms $f(0) \geq \cdots \geq f(n-1) > 0$, the first row g^T of U can be chosen as

$$g_1 = \frac{1}{f(0)}, \qquad g_k = \sqrt{\frac{1}{f(k-1)^2} - \frac{1}{f(k-2)^2}}, \qquad k = 2, \ldots, n.$$

The parametrization for FOM and GMRES

Let

$$A = V(U^{-1}CU)V^*, \qquad b = Ve_1.$$

- To force GMRES residual norms $f(0) \geq \cdots \geq f(n-1) > 0$, the first row g^T of U can be chosen as

$$g_1 = \frac{1}{f(0)}, \qquad g_k = \sqrt{\frac{1}{f(k-1)^2} - \frac{1}{f(k-2)^2}}, \qquad k = 2, \ldots, n.$$

- To force Ritz values, i.e. eigenvalues of H_k, the remaining submatrix T of

$$U = \begin{bmatrix} & g^T & \\ 0 & & T \end{bmatrix}$$

can be chosen to have entries satisfying

$$\prod_{i=1}^{k}(\lambda - \rho_i^{(k)}) = \frac{1}{t_{k,k}} \left(g_{k+1} + \sum_{i=1}^{k} t_{i,k}\lambda^i \right).$$

Outline

Bi-Lanczos

The FOM/GMRES pair of methods with long recurrences has an analogue among methods with short recurrences: The Bi-CG/QMR pair.

Bi-Lanczos

The FOM/GMRES pair of methods with long recurrences has an analogue among methods with short recurrences: The Bi-CG/QMR pair.

Bi-CG and QMR use bi-orthogonal bases of Krylov subspaces constructed with the Bi-Lanczos algorithm. In the kth iteration it computes (when there is no breakdown) the decomposition

$$AV_k = V_k T_k + h_{k+1,k} v_{k+1} e_1^T = V_{k+1} \tilde{T}_k,$$

where the columns of $V_k = [v_1, \ldots, v_k]$ span $\mathcal{K}_k(\mathbf{A}, b)$ and satisfy

$$W_k^* V_k = \mathrm{diag}(\omega_1, \ldots, \omega_k), \quad \omega_i \neq 0$$

for a matrix W_k whose columns span $\mathcal{K}_k(\mathbf{A}^*, b)$. The matrix $\tilde{T}_k$ (rectangular), resp. T_k (square) is tridiagonal, thus allowing for short recurrences to generate the bi-orthogonal bases.

Bi-Lanczos

In analogy with the FOM/GMRES pair, the kth QMR residual norm is

$$\|r_k^{QMR}\| = \|V_{k+1}(\|b\|e_1 - \tilde{T}_k y_k)\|, \quad y_k = \min_{y \in \mathbb{C}^k} \left\| \|b\|e_1 - \tilde{T}_k y \right\|.$$

Bi-Lanczos

In analogy with the FOM/GMRES pair, the kth QMR residual norm is

$$\|r_k^{QMR}\| = \|V_{k+1}(\|b\|e_1 - \tilde{T}_k y_k)\|, \quad y_k = \min_{y \in \mathbb{C}^k} \left\| \|b\|e_1 - \tilde{T}_k y \right\|.$$

The kth Bi-CG residual norm satisfies

$$\|r_k^{BiCG}\| = \|b\| \cdot |t_{k+1,k} e_k^T T_k^{-1} e_1| \cdot \|v_{k+1}\|.$$

and does not exist for singular T_k.

Bi-Lanczos

In analogy with the FOM/GMRES pair, the kth QMR residual norm is

$$\|r_k^{QMR}\| = \|V_{k+1}(\|b\|e_1 - \tilde{T}_k y_k)\|, \quad y_k = \min_{y \in \mathbb{C}^k} \left\| \|b\|e_1 - \tilde{T}_k y \right\|.$$

The kth Bi-CG residual norm satisfies

$$\|r_k^{BiCG}\| = \|b\| \cdot |t_{k+1,k} e_k^T T_k^{-1} e_1| \cdot \|v_{k+1}\|.$$

and does not exist for singular T_k.

With a correct scaling of the columns of V_{k+1}, Bi-CG residual norms are fully determined by the entries of $\tilde{T}_k$. (QMR residual norms also depend upon $\|V_{k+1}\|$). We can therefore try to extend the construction to prescribe FOM convergence behavior.

The Bi-CG parametrization

To force desired eigenvalues and Bi-CG residual norms we can

- Choose a nonsingular matrix V with normalized columns and put $b = Ve_1$ and

$$A = VTV^{-1}, \qquad T \text{ tridiagonal.}$$

The Bi-CG parametrization

To force desired eigenvalues and Bi-CG residual norms we can

- Choose a nonsingular matrix V with normalized columns and put $b = Ve_1$ and

$$A = VTV^{-1}, \qquad T \text{ tridiagonal.}$$

- Try to find a tridiagonal T allowing the decomposition

$$T = U^{-1}CU, \qquad U \text{ nonsingular upper triangular,}$$

where C is the companion matrix for the prescribed spectrum and

The Bi-CG parametrization

To force desired eigenvalues and Bi-CG residual norms we can

- Choose a nonsingular matrix V with normalized columns and put $b = Ve_1$ and

$$A = VTV^{-1}, \qquad T \text{ tridiagonal.}$$

- Try to find a tridiagonal T allowing the decomposition

$$T = U^{-1}CU, \qquad U \text{ nonsingular upper triangular,}$$

where C is the companion matrix for the prescribed spectrum and where the first row g^T of U has entries

$$g_k = \frac{1}{f(k-1)}, \quad k = 1, \ldots, n,$$

if $f(0), \ldots, f(n-1)$, $f(i) > 0$ are the prescribed Bi-CG residual norms.

The Bi-CG parametrization

Let us assume for the moment, that we wish to prescribe only convergence curves where all iterates are defined. That means, we assume that the entries

$$g_k = \frac{1}{\|r_{k-1}^{BiCG}\|}, \quad k = 1, \ldots, n.$$

of the first row of U are all nonzero. Equivalently, all leading principal submatrices T_k of T are nonsingular.

The Bi-CG parametrization

Let us assume for the moment, that we wish to prescribe only convergence curves where all iterates are defined. That means, we assume that the entries

$$g_k = \frac{1}{\|r_{k-1}^{BiCG}\|}, \quad k = 1, \ldots, n.$$

of the first row of U are all nonzero. Equivalently, all leading principal submatrices T_k of T are nonsingular.

Clearly

$$T = U^{-1}CU, \qquad U \text{ nonsingular upper triangular,}$$

where C is a given companion matrix and the first row g^T of U is a given vector is always upper Hessenberg. Note that there holds [Parlett 1967]

$$U^{-1} = [e_1, Te_1, \ldots, T^{n-1}e_1].$$

To obtain a *tridiagonal* T one can apply the Bi-Lanczos algorithm to C with some starting vector z. But the Bi-Lanczos process can break down ...

The Bi-CG parametrization

Theorem [Joubert 1992]: The Bi-Lanczos algorithm applied to a nonderogatory matrix runs till completion and no leading principal submatrix of the generated tridiagonal matrix will be singular for almost every starting vector, i.e. except for a measure zero set of vectors.

The Bi-CG parametrization

Theorem **[Joubert 1992]**: The Bi-Lanczos algorithm applied to a nonderogatory matrix runs till completion and no leading principal submatrix of the generated tridiagonal matrix will be singular for almost every starting vector, i.e. except for a measure zero set of vectors.

Thus let for an appropriate z,

$$CZ = ZT, \quad Ze_1 = z.$$

Then $T = Z^{-1}CZ$ will have the desired spectrum.

The Bi-CG parametrization

Theorem [Joubert 1992]: The Bi-Lanczos algorithm applied to a nonderogatory matrix runs till completion and no leading principal submatrix of the generated tridiagonal matrix will be singular for almost every starting vector, i.e. except for a measure zero set of vectors.

Thus let for an appropriate z,

$$CZ = ZT, \quad Ze_1 = z.$$

Then $T = Z^{-1}CZ$ will have the desired spectrum. If we define $U = [e_1, Te_1, \ldots, T^{n-1}e_1]^{-1}$, T can be decomposed as

$$T = U^{-1}CU, \quad U \text{ nonsingular upper triangular}$$

The Bi-CG parametrization

Theorem [Joubert 1992]: The Bi-Lanczos algorithm applied to a nonderogatory matrix runs till completion and no leading principal submatrix of the generated tridiagonal matrix will be singular for almost every starting vector, i.e. except for a measure zero set of vectors.

Thus let for an appropriate z,

$$CZ = ZT, \quad Ze_1 = z.$$

Then $T = Z^{-1}CZ$ will have the desired spectrum. If we define $U = [e_1, Te_1, \ldots, T^{n-1}e_1]^{-1}$, T can be decomposed as

$$T = U^{-1}CU, \quad U \text{ nonsingular upper triangular}$$

To have the first row of U equal to our given g^T, it suffices to scale T with a diagonal matrix D as

$$\hat{T} = D^{-1}TD = (UD)^{-1}C(UD).$$

Outline

Bi-CG breakdown types

So far, we assumed that we prescribe only convergence curves where all iterates are defined. Let us list all situations where the Bi-CG method can break down:

Bi-CG can break down if the underlying Bi-Lanczos process breaks down:

Bi-CG breakdown types

So far, we assumed that we prescribe only convergence curves where all iterates are defined. Let us list all situations where the Bi-CG method can break down:

Bi-CG can break down if the underlying Bi-Lanczos process breaks down:

- In case of a lucky breakdown, a subdiagonal entry of T is zero and we have found an A- or A^*-invariant subspace. This gives, in exact arithmetics, the solution of the linear system or of the system with A^*.

Bi-CG breakdown types

So far, we assumed that we prescribe only convergence curves where all iterates are defined. Let us list all situations where the Bi-CG method can break down:

Bi-CG can break down if the underlying Bi-Lanczos process breaks down:

- In case of a lucky breakdown, a subdiagonal entry of T is zero and we have found an A- or A^*-invariant subspace. This gives, in exact arithmetics, the solution of the linear system or of the system with A^*.
- In case of a serious breakdown, two basis vectors v_k and w_k are orthogonal to eachother and the bi-orthogonality condition

  $$W_k^* V_k = \mathrm{diag}(\omega_1, \ldots, \omega_k), \quad \omega_i \neq 0$$

  cannot be satisfied. One way to continue the Bi-Lanczos process is to use a look-ahead technique, i.e. to perform further iterations until for some i, $w_{k+i}^* v_{k+i} \neq 0$.

Bi-CG breakdown types

So far, we assumed that we prescribe only convergence curves where all iterates are defined. Let us list all situations where the Bi-CG method can break down:

Bi-CG can break down if the underlying Bi-Lanczos process breaks down:

- In case of a lucky breakdown, a subdiagonal entry of T is zero and we have found an A- or A^*-invariant subspace. This gives, in exact arithmetics, the solution of the linear system or of the system with A^*.
- In case of a serious breakdown, two basis vectors v_k and w_k are orthogonal to eachother and the bi-orthogonality condition

  $$W_k^* V_k = \mathrm{diag}(\omega_1, \ldots, \omega_k), \quad \omega_i \neq 0$$

  cannot be satisfied. One way to continue the Bi-Lanczos process is to use a look-ahead technique, i.e. to perform further iterations until for some i, $w_{k+i}^* v_{k+i} \neq 0$.
- In case of an incurable breakdown, no look-ahead strategy will help to generate a pair of bi-orthognal bases ($w_{k+i}^* v_{k+i} = 0$ for all i).

Bi-CG breakdown types

Bi-CG can also break down if the underlying Bi-Lanczos process runs to completion.

Bi-CG breakdown types

Bi-CG can also break down if the underlying Bi-Lanczos process runs to completion.

In this case, the breakdown is due to the fact that the current tridiagonal matrix T_k is singular.

Bi-CG breakdown types

Bi-CG can also break down if the underlying Bi-Lanczos process runs to completion.

In this case, the breakdown is due to the fact that the current tridiagonal matrix T_k is singular.

This corresponds precisely to a zero entry on position k of the first row of U in the decomposition

$$T = U^{-1}CU, \quad U \text{ nonsingular upper triangular}, \quad C \text{ a companion matrix.}$$

Bi-CG breakdown types

Bi-CG can also break down if the underlying Bi-Lanczos process runs to completion.

In this case, the breakdown is due to the fact that the current tridiagonal matrix T_k is singular.

This corresponds precisely to a zero entry on position k of the first row of U in the decomposition

$$T = U^{-1}CU, \quad U \text{ nonsingular upper triangular}, \quad C \text{ a companion matrix.}$$

If we wish to prescribe *any* Bi-CG convergence curve, we must also be able to prescribe curves with this type of breakdown, i.e. with possibly non-defined iterates.

Prescribing Bi-CG breakdowns

Thus the question is:

Can we find *tridiagonal* matrices T of the form

$$T = U^{-1}CU, \qquad U \text{ nonsingular upper triangular,}$$

where C is a given companion matrix and the first row g^T of U is a given vector with zeros on prescribed positions.

Prescribing Bi-CG breakdowns

Thus the question is:

Can we find *tridiagonal* matrices T of the form

$$T = U^{-1}CU, \qquad U \text{ nonsingular upper triangular,}$$

where C is a given companion matrix and the first row g^T of U is a given vector with zeros on prescribed positions.

The answer is no, or not yet ? But first, we have this proposition:

Prescribing Bi-CG breakdowns

Thus the question is:

Can we find *tridiagonal* matrices T of the form

$$T = U^{-1}CU, \qquad U \text{ nonsingular upper triangular,}$$

where C is a given companion matrix and the first row g^T of U is a given vector with zeros on prescribed positions.

The answer is no, or not yet ? But first, we have this proposition:

Theorem **[DT & Meurant 2014?]**: Assume Bi-Lanczos with nonsingular input matrix A with initial vector b runs to completion. Let $x_i \equiv \infty$ if and only of the ith leading principal submatrix of T is singular. Then whenever $x_{k-1} = \infty$, we have $x_k \neq \infty$.

Prescribing Bi-CG breakdowns

Thus it is impossible in Bi-CG to have two subsequent iterations where iterates are not defined.

Prescribing Bi-CG breakdowns

Thus it is impossible in Bi-CG to have two subsequent iterations where iterates are not defined.

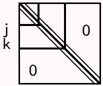
We were unable to find a way to generate a tridiagonal matrix with given spectrum and singular leading principal submatrices of particular (not subsequent) sizes. This seems to be an open or unconsidered problem in the literature.

Prescribing Bi-CG breakdowns

Thus it is impossible in Bi-CG to have two subsequent iterations where iterates are not defined.

We were unable to find a way to generate a tridiagonal matrix with given spectrum and singular leading principal submatrices of particular (not subsequent) sizes. This seems to be an open or unconsidered problem in the literature.

T_n



.

Prescribing Bi-CG breakdowns

The problem amounts to finding a starting vector for Bi-Lanczos applied to a companion matrix C such that the generated tridiagonal matrix has some singular leading principal submatrices.

Prescribing Bi-CG breakdowns

The problem amounts to finding a starting vector for Bi-Lanczos applied to a companion matrix C such that the generated tridiagonal matrix has some singular leading principal submatrices.

To force a leading principal submatrices to be singular, one can modify two entries of a starting vector. In this way we are able to influence the singularity of only the first half of the leading principal submatrices of T.

Prescribing Bi-CG breakdowns

The problem amounts to finding a starting vector for Bi-Lanczos applied to a companion matrix C such that the generated tridiagonal matrix has some singular leading principal submatrices.

To force a leading principal submatrices to be singular, one can modify two entries of a starting vector. In this way we are able to influence the singularity of only the first half of the leading principal submatrices of T.

Similarly, one can use a backward eigenvalue-forcing strategy described in Parlett and Strang [Parlett & Strang 2008] to influence the singularity of only the second half of the leading principal submatrices of T.

Outline

Conclusions

- Any history of positive, finite Bi-CG residual norm is possible with any eigenvalues.

Conclusions

- Any history of positive, finite Bi-CG residual norm is possible with any eigenvalues.
- Additionally, infinite Bi-CG residual norms can be prescribed either in the first or in the second half of the iterations.

Conclusions

- Any history of positive, finite Bi-CG residual norm is possible with any eigenvalues.
- Additionally, infinite Bi-CG residual norms can be prescribed either in the first or in the second half of the iterations.
- Future work: Consequences for the QMR method, prescribing Ritz values of tridiagonal matrices.

Conclusions

- Any history of positive, finite Bi-CG residual norm is possible with any eigenvalues.
- Additionally, infinite Bi-CG residual norms can be prescribed either in the first or in the second half of the iterations.
- Future work: Consequences for the QMR method, prescribing Ritz values of tridiagonal matrices.

For more details see:

J. Duintjer Tebbens, G. Meurant: On the Convergence of QOR and QMR Krylov Methods for Solving Linear Systems, to be submitted.

Last but not least

Thank you for your attention!

Related papers

- A. Greenbaum and Z. Strakoš, Matrices that generate the same Krylov residual spaces, IMA Vol. Math. Appl., 60 (1994), pp. 95–118.
- A. Greenbaum, V. Pták and Z. Strakoš, Any nonincreasing convergence curve is possible for GMRES, SIMAX, 17 (1996), pp. 465–469.
- M. Arioli, V. Pták and Z. Strakoš, Krylov sequences of maximal length and convergence of GMRES, BIT, 38 (1998), pp. 636–643.
- E. Vecharinsky and J. Langou, Any admissible cycle-convergence behavior is possible for restarted GMRES at its initial cycles, 18 (2011), Num. Lin. Alg. Appl., pp. 499–511.
- J. Duintjer Tebbens and G. Meurant, Any Ritz value behavior is possible for Arnoldi and for GMRES, SIMAX, 33 (2012), pp. 958–978.
- J. Duintjer Tebbens and G. Meurant, Prescribing the behavior of early terminating GMRES and Arnoldi iterations, Numer. Algorithms, 65 (2014), pp. 69–90.
- M. Schweitzer, Any cycle-convergence curve is possible for restarted FOM, technical report Univ. Wuppertal, Preprint BUW-IMACM 14/19, 2014.