The worst-case GMRES for normal matrices

Petr Tichý

Technical University of Berlin and ICS AS CR

joint work with

Jörg Liesen

Technical University of Berlin

May 28, IMET 2004, Prague

Introduction

A system of linear algebraic equations

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

$\mathbf{A} \in \mathbb{C}^{n\times n}$ is **nonsingular and normal**, $b \in \mathbb{C}^n$.

Eigendecomposition of $\mathbf{A}$:

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H, \qquad \mathbf{Q}^H\mathbf{Q} = \mathbf{I}, \qquad \mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

We will assume that all eigenvalues of $\mathbf{A}$ are **distinct**.

GMRES

Given $x_0 \in \mathbb{C}^n$, $r_0 = b - \mathbf{A}x_0$.

GMRES computes a sequence of iterates x_i ,

$$x_i \in x_0 + \mathcal{K}_i(\mathbf{A}, r_0)$$

so that $r_i = b - \mathbf{A}x_i$ satisfies

$$\|r_i\| \;=\; \min_{p \in \pi_i} \|p(\mathbf{A})\, r_0\| \,,$$

where

$$\mathcal{K}_i(\mathbf{A}, r_0) \quad \equiv \quad \operatorname{span}\{r_0, \cdots, \mathbf{A}^{i-1} r_0\} \,,$$

$$\pi_i \quad \equiv \quad \{\, p \text{ is a polynomial};\ \deg(p) \le i;\ p(0) = 1\} \,.$$

Worst-case GMRES residual norm

Let $\|r_0\| = 1$, $L \equiv \{\lambda_1, \ldots, \lambda_n\}$. Then

$$\begin{aligned} \|r_i\| &= \min_{p \in \pi_i} \|p(\mathbf{A})\, r_0\| \\ &\leq \min_{p \in \pi_i} \|p(\mathbf{A})\| = \min_{p \in \pi_i} \max_{\lambda_j \in L} |p(\lambda_j)|\,, \end{aligned}$$

This "standard" bound **is sharp**. [Greenbaum, Gurvits-94, Joubert-94]

Worst-case GMRES residual norm

Let $\|r_0\| = 1$, $L \equiv \{\lambda_1, \ldots, \lambda_n\}$. Then

$$\begin{aligned} \|r_i\| &= \min_{p \in \pi_i} \|p(\mathbf{A})\, r_0\| \\ &\leq \min_{p \in \pi_i} \|p(\mathbf{A})\| = \min_{p \in \pi_i} \max_{\lambda_j \in L} |p(\lambda_j)|\,, \end{aligned}$$

This "standard" bound **is sharp**. [Greenbaum, Gurvits-94, Joubert-94]

Definition: An ith worst-case GMRES residual r_i^w for $\mathbf{A} \in \mathbb{C}^{n \times n}$ is a GMRES residual that satisfies

$$\|r_i^w\| = \max_{\|r_0\|=1} \min_{p \in \pi_i} \|p(\mathbf{A})\, r_0\|\,,$$

where $i = 1, \ldots, n-1$.

Questions

For every GMRES residual r_i it holds

$$\|r_i\| \;\leq\; \|r_i^w\| \;=\; \min_{p\in\pi_i} \max_{\lambda_j\in L} |p(\lambda_j)|\,.$$

Evaluation of the bound

- Can we describe the standard bound in terms of eigenvalues?

Relevance of the bound

- How good describes the standard bound the GMRES convergence?

Other questions

- Which r_0 yields the worst-case residual norm?
- How does the worst-case polynomial look like?

Outline

1. Worst-case GMRES residual norm
2. Numerical experiments
3. Relevance of the standard bound
4. A model problem with known eigenvalues
5. Conclusions

1. Worst-case GMRES residual norm

Factorization of Krylov matrix

Krylov matrix:

$$\mathbf{K}_{i+1} \equiv [r_0, \mathbf{A}r_0, \ldots, \mathbf{A}^i r_0] \, .$$

We consider $\mathbf{A}$ and r_0 in the form

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H, \qquad r_0 = \mathbf{Q}\, [\varrho_1, \ldots, \varrho_n]^T \, .$$

Factorization:

$$\mathbf{K}_{i+1} = \mathbf{Q}\mathbf{D}\mathbf{V}_{i+1}$$

where

$$\mathbf{D} \equiv \begin{bmatrix} \varrho_1 & & \\ & \ddots & \\ & & \varrho_n \end{bmatrix}, \qquad \mathbf{V}_{i+1} \equiv \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^i \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^i \end{bmatrix} .$$

GMRES residual

Residual r_i can be written as [LiRoSt-02, Ipsen-00]

$$r_i = \|r_i\|^2 \, (\mathbf{K}_{i+1}^{+})^H \, e_1$$
$$= \|r_i\|^2 \, \mathbf{Q} \, \left[(\mathbf{D}\mathbf{V}_{i+1})^{+} \right]^H \, e_1 \, .$$

and

$$\|r_i\| = \frac{1}{\| [(\mathbf{D}\mathbf{V}_{i+1})^{+}]^H \, e_1 \|} \, .$$

(Assumption: $\mathbf{K}_{i+1}$ has full column rank)

GMRES residual norm (next-to-last step)

Let $\varrho_j \neq 0$ for all j. Then [Liesen, T.-04, Ipsen-00]

$$\|r_{n-1}\| = \frac{1}{\|\mathbf{D}^{-H}\mathbf{V}_n^{-H}e_1\|} = \frac{1}{\left(\sum_{j=1}^{n}\left|\frac{l_j(0)}{\varrho_j}\right|^2\right)^{1/2}},$$

where

$$l_j(\lambda) \equiv \prod_{\substack{k=1\\k\neq j}}^{n} \frac{\lambda_k - \lambda}{\lambda_k - \lambda_j}.$$

GMRES residual norm (next-to-last step)

Let $\varrho_j \neq 0$ for all j. Then [Liesen, T.-04, Ipsen-00]

$$\|r_{n-1}\| = \frac{1}{\|\mathbf{D}^{-H}\mathbf{V}_n^{-H} e_1\|} = \frac{1}{\left(\sum_{j=1}^{n} \left| \frac{l_j(0)}{\varrho_j} \right|^2 \right)^{1/2}},$$

where

$$l_j(\lambda) \equiv \prod_{\substack{k=1 \\ k \neq j}}^{n} \frac{\lambda_k - \lambda}{\lambda_k - \lambda_j}.$$

Let $\|r_0\| = 1$. Using Cauchy's inequality, [Liesen,T.-04]

$$\|r_{n-1}^w\| = \frac{1}{\sum_{j=1}^{n} |l_j(0)|}.$$

Worst-case residual norm in a general step i

For each $S \subseteq L = \{\lambda_1, \ldots, \lambda_n\}$ we denote

$$M_i^S \;\equiv\; \min_{p \in \pi_i} \max_{\lambda_j \in S} |p(\lambda_j)|\,.$$

- **We want** to determine the value $M_i^L \;=\; \|r_i^w\|$.
- **We are able** to determine

$$M_i^S \;=\; \left(\sum_{k=1}^{i+1} |l_k^S(0)| \right)^{-1}, \quad S \subseteq L,\;\; |S| = i+1.$$

Worst-case residual norm in a general step i

For each $S \subseteq L = \{\lambda_1, \ldots, \lambda_n\}$ we denote

$$M_i^S \;\equiv\; \min_{p \in \pi_i} \max_{\lambda_j \in S} |p(\lambda_j)| \,.$$

- **We want** to determine the value $M_i^L \;=\; \|r_i^w\|$.
- **We are able** to determine

$$M_i^S \;=\; \left(\sum_{k=1}^{i+1} |l_k^S(0)| \right)^{-1}, \quad S \subseteq L, \;\; |S| = i+1.$$

For each subset $S \subseteq L$ it holds $M_i^L \;\geq\; M_i^S$, i.e.

$$M_i^L \;\geq\; \max_{\substack{S \subseteq L \\ |S| = i+1}} M_i^S \;\equiv\; B_i^L.$$

lower bound

Tightness of the bound

All eigenvalues are real: [Liesen, T.-04, Greenbaum-79]

$$\|r_i^w\| \;=\; M_i^L \;=\; B_i^L\,.$$

At least one non-real eigenvalue: [Liesen, T.-04]

$$B_i^L \;\leq\; \|r_i^w\| \;\leq\; \sqrt{(i+1)(n-i)}\; B_i^L\,.$$

Numerical Experiments show that B_i^L is very close to $\|r_i^w\|$.

Tightness of the bound

All eigenvalues are real: [Liesen, T.-04, Greenbaum-79]

$$\|r_i^w\| \;=\; M_i^L \;=\; B_i^L\,.$$

At least one non-real eigenvalue: [Liesen, T.-04]

$$B_i^L \;\leq\; \|r_i^w\| \;\leq\; \sqrt{(i+1)(n-i)}\;B_i^L\,.$$

Numerical Experiments show that B_i^L is very close to $\|r_i^w\|$.

We conjecture:

$$B_i^L \;\leq\; \|r_i^w\| \;\leq\; C\,B_i^L\,,$$

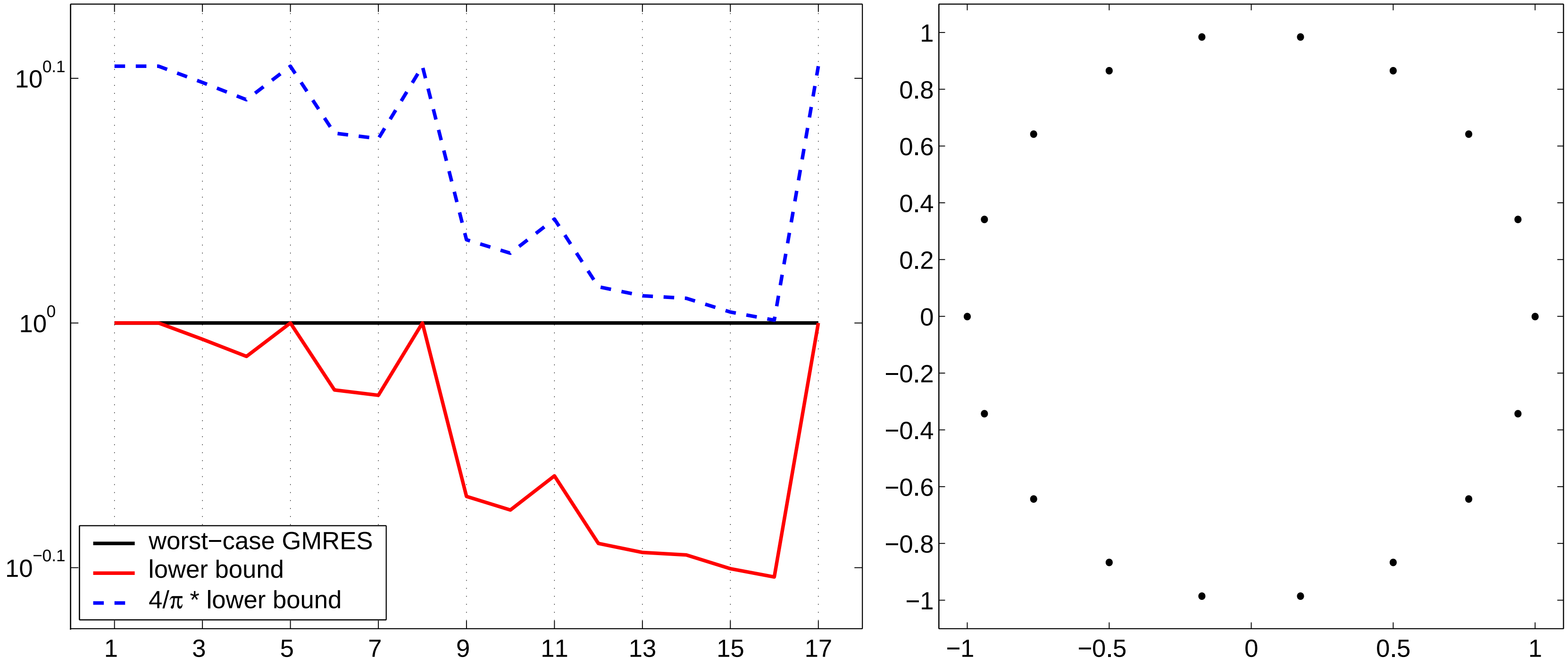
$C \;>\; 1$ is a constant independent on i, n and on eigenvalue distribution

2. Numerical Experiments in MATLAB

Experiment 1: roots of unity

In this case the worst-case GMRES completely stagnates, i.e.

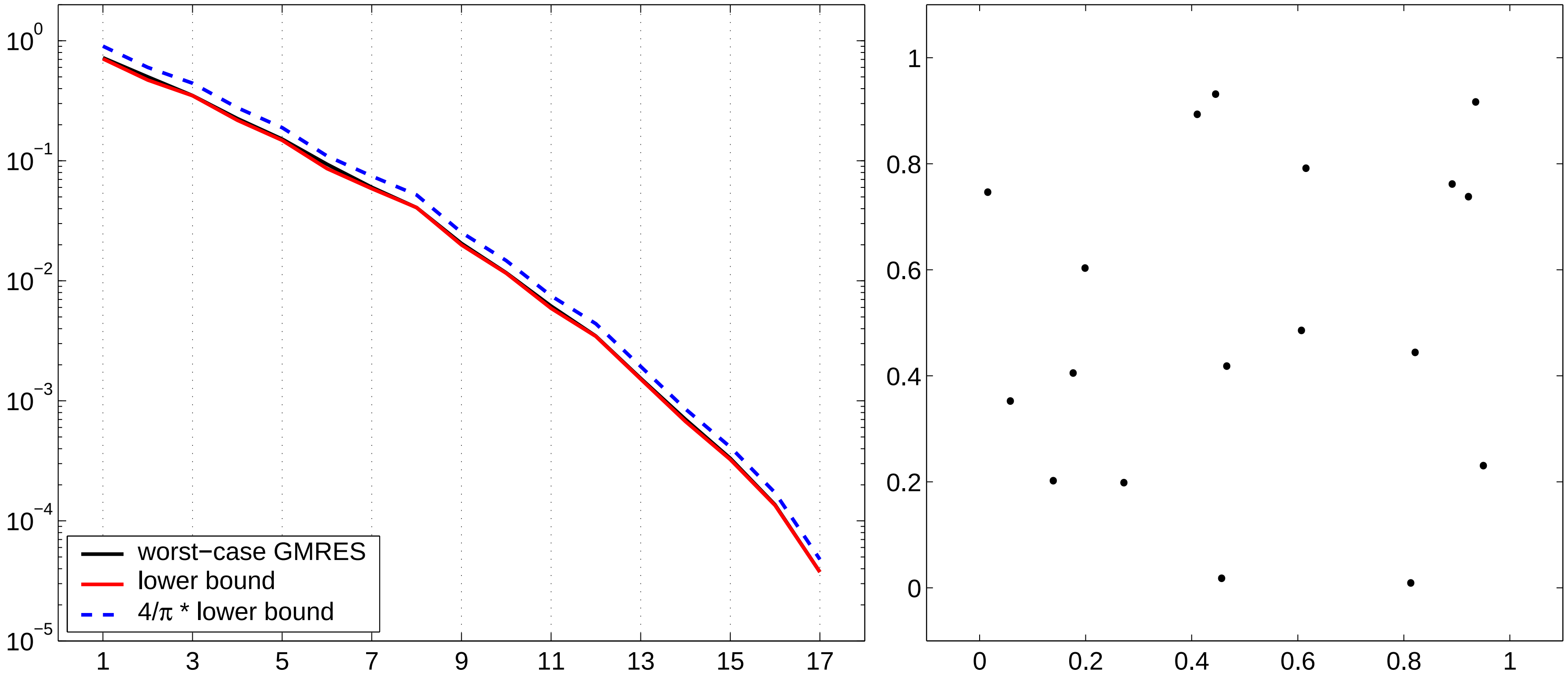
$$1 = \|r_i^w\|, \quad i = 0, \ldots, n-1.$$



We proved: $\|r_{n-2}^w\| < \frac{4}{\pi} B_{n-2}^L, \qquad \lim_{n\to\infty} \left[\frac{4}{\pi} B_{n-2}^L\right] = \|r_{n-2}^w\|,$

Experiment 2: random eigenvalues

Random eigenvalues in the region $[0, 1] \times \mathbf{i}\,[0, 1]$



Constant $4/\pi$

Numerical experiments predict that

$$B_i^L \;\le\; \|r_i^w\| \;\le\; \frac{4}{\pi}\, B_i^L$$

holds for all sets L containing n distinct complex numbers, where

$$B_i^L \;\equiv\; \max_{\substack{S \subseteq L \\ |S| = i+1}} \frac{1}{\displaystyle\sum_{j=1}^{i+1} \prod_{\substack{k=1 \\ k \neq j}}^{n} \frac{|\lambda_k^S|}{|\lambda_k^S - \lambda_j^S|}}\,.$$

3. Relevance of the standard bound

Relevance of the standard bound

How good describes the standard bound the convergence of GMRES?

It depends on r_0, in case of $x_0 = 0$ on the right-hand side b:

Relevance of the standard bound

How good describes the standard bound the convergence of GMRES?

It depends on r_0, in case of $x_0 = 0$ on the right-hand side b:

- b is **unbiased** $\rightarrow$ b has components in the matrix eigenvectors of approximately equal size. Example: $\varrho_i \equiv n^{-1/2}$,

$$\|r_i^w\| \;\geq\; \|r_i^u\| \quad\geq\quad n^{-1/2}\,\|r_i^w\|.$$

  The standard bound describes the convergence (up to factor $n^{-1/2}$).

- b is **biased** $\rightarrow$ the standard (worst-case) bound need not describe the convergence (often is an overestimation)

4. Model problem

Model problem: Poisson equation

$$-u''(z) \;=\; f(z), \qquad z \in (0,1)\,, \qquad u(0) = u_0, \;\; u(1) = u_1\,.$$

The central finite difference approximation on the uniform grid kh, $k = 1, \ldots, n$, $h = 1/(n+1)$, leads to a linear algebraic system $\mathbf{A}x = b$

$$\overbrace{\begin{bmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{bmatrix}}^{\mathbf{A}} x \;=\; \overbrace{h^2 \begin{bmatrix} f(h) \\ \vdots \\ \vdots \\ f(nh) \end{bmatrix} + \begin{bmatrix} u_0 \\ \\ \\ u_1 \end{bmatrix}}^{b}\,.$$

The eigenvalues λ_k and the eigenvectors q_k of $\mathbf{A}$ are known.

We can apply MINRES or CG to this system.

Some results for the Poisson equation

Worst-case $\times$ unbiased case (MINRES)

$$\|r_{n-1}^{w}\| = \frac{1}{n}, \qquad \|r_{n-1}^{u}\| > \sqrt{\frac{2}{3}}\,\frac{1}{n}.$$

Some results for the Poisson equation

Worst-case × unbiased case (MINRES)

$$\|r^w_{n-1}\| = \frac{1}{n}, \qquad \|r^u_{n-1}\| > \sqrt{\frac{2}{3}}\,\frac{1}{n}.$$

Worst data for linear solver (CG)

worst ϱ — worst r_0 — data for diff. eq.

$$\begin{bmatrix} \sin(\pi h) \\ \vdots \\ \sin(n\pi h) \end{bmatrix} \leftrightarrow \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftrightarrow \begin{aligned} f(z) &= 0, \\ u(0) &= 1, \\ u(1) &= 0. \end{aligned}$$

CG started with $x_0 = 0$ and $b = e_1$ attains the worst-case relative A-norm of the error in the $(n-1)$st iteration step ($= n^{-1}$).

Conclusions

- We completely characterized (for normal matrices):

$\longrightarrow$ the GMRES residual in the next-to-last step,

$\longrightarrow$ the worst-case GMRES residual in the next-to-last step.

- The worst-case GMRES norm $\|r_i^w\|$ can be estimated by B_i^L :

$\longrightarrow$ $\|r_i^w\| \;=\; B_i^L$ for real eigenvalues,

$\longrightarrow$ $\|r_i^w\| \;\geq\; B_i^L$ in general case.

- Numerical experiments predict that

$$B_i^L \;\leq\; \|r_i^w\| \;\leq\; \frac{4}{\pi}\, B_i^L\,.$$

- The standard bound seems to be reasonable for unbiased b.

- Our results allow to study model problems with known eigenvalues.

- If A is SPD, all results transfer to the $\mathbf{A}$-norm of the error in CG.

About

<u>More details can be found in</u>

Liesen, J. and Tichý, P., **The worst-case GMRES for normal matrices**, BIT Numerical Mathematics, 44 (2004), pp. 79–98.

Liesen, J. and Tichý, P., **Behavior of CG and MINRES for symmetric tridiagonal Toeplitz matrices**, in preparation, (2004).

or at

`http://www.math.tu-berlin.de/~tichy`
`http://www.math.tu-berlin.de/~liesen`