On Ideal and Worst-case GMRES

Petr Tichý

joint work with

Jörg Liesen and Vance Faber

Institute of Computer Science AS CR

May 15, 2009, Prague

GMRES, Worst-case GMRES and Ideal GMRES

$\mathbf{A}x = b$, $\mathbf{A} \in \mathbb{C}^{n \times n}$ is nonsingular, $b \in \mathbb{C}^n$,

$x_0 = \mathbf{0}$ and $\|b\| = 1$ for simplicity.

GMRES, Worst-case GMRES and Ideal GMRES

$\mathbf{A}x = b$, $\mathbf{A} \in \mathbb{C}^{n\times n}$ is nonsingular, $b \in \mathbb{C}^n$,

$x_0 = \mathbf{0}$ and $\|b\| = 1$ for simplicity.

GMRES computes $x_k \in \mathcal{K}_k(\mathbf{A}, b)$ such that $r_k \equiv b - \mathbf{A}x_k$ satisfies

$$\|r_k\| = \min_{p \in \pi_k} \|p(\mathbf{A})b\| \qquad \text{(GMRES)}$$

GMRES, Worst-case GMRES and Ideal GMRES

$\mathbf{A}x = b$, $\mathbf{A} \in \mathbb{C}^{n\times n}$ is nonsingular, $b \in \mathbb{C}^n$,

$x_0 = \mathbf{0}$ and $\|b\| = 1$ for simplicity.

GMRES computes $x_k \in \mathcal{K}_k(\mathbf{A}, b)$ such that $r_k \equiv b - \mathbf{A}x_k$ satisfies

$$\|r_k\| = \min_{p\in\pi_k} \|p(\mathbf{A})b\| \qquad \text{(GMRES)}$$

$$\leq \max_{\|b\|=1} \min_{p\in\pi_k} \|p(\mathbf{A})b\| \equiv \psi_k(\mathbf{A}) \qquad \text{(worst-case GMRES)}$$

GMRES, Worst-case GMRES and Ideal GMRES

$\mathbf{A}x = b$, $\mathbf{A} \in \mathbb{C}^{n\times n}$ is nonsingular, $b \in \mathbb{C}^n$,

$x_0 = \mathbf{0}$ and $\|b\| = 1$ for simplicity.

GMRES computes $x_k \in \mathcal{K}_k(\mathbf{A}, b)$ such that $r_k \equiv b - \mathbf{A}x_k$ satisfies

$$\begin{aligned}
\|r_k\| &= \min_{p\in\pi_k} \|p(\mathbf{A})b\| && \text{(GMRES)} \\
&\le \max_{\|b\|=1} \min_{p\in\pi_k} \|p(\mathbf{A})b\| \equiv \psi_k(\mathbf{A}) && \text{(worst-case GMRES)} \\
&\le \min_{p\in\pi_k} \|p(\mathbf{A})\| \equiv \varphi_k(\mathbf{A}) && \text{(ideal GMRES)}.
\end{aligned}$$

Questions

$$\|r_k\| \leq \underbrace{\max_{\|b\|=1} \min_{p \in \pi_k} \|p(\mathbf{A})b\|}_{\psi_k(\mathbf{A})} \leq \underbrace{\min_{p \in \pi_k} \|p(\mathbf{A})\|}_{\varphi_k(\mathbf{A})}$$

- Which (known) approximation problem is solved?
- How to approximate ideal/worst-case quantities?
- When does it hold that ideal = worst case GMRES?
- Is the solution unique?

Outline

1. Worst-case GMRES for normal matrices
2. Results for nonnormal matrices
3. Cross equality for worst-case GMRES vectors
4. Results for a Jordan block

Outline

Worst-case GMRES for normal matrices

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^*, \quad \mathbf{Q}^*\mathbf{Q} = \mathbf{I}\,.$$

- [Greenbaum & Gurvits '94, Joubert '94] showed:

$$\max_{\|b\|=1} \min_{p\in\pi_k} \|p(\mathbf{A})b\| \;=\; \min_{p\in\pi_k} \|p(\mathbf{A})\|$$

Worst-case GMRES for normal matrices

$$\mathbf{A} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^*, \quad \mathbf{Q}^*\mathbf{Q} = \mathbf{I}\,.$$

- [Greenbaum & Gurvits '94, Joubert '94] showed:

$$\max_{\|b\|=1} \min_{p\in\pi_k} \|p(\mathbf{A})b\| \;=\; \min_{p\in\pi_k} \|p(\mathbf{A})\|$$

- Which (known) approximation problem is solved?

$$\min_{p\in\pi_k} \|p(\mathbf{A})\| \;=\; \min_{p\in\pi_k} \|\mathbf{Q}p(\boldsymbol{\Lambda})\mathbf{Q}^*\| \;=\; \min_{p\in\pi_k} \max_{\lambda_i} |p(\lambda_i)|\,.$$

Worst-case GMRES for normal matrices

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^*, \quad \mathbf{Q}^*\mathbf{Q} = \mathbf{I}\,.$$

- [Greenbaum & Gurvits '94, Joubert '94] showed:

$$\max_{\|b\|=1} \min_{p\in\pi_k} \|p(\mathbf{A})b\| \;=\; \min_{p\in\pi_k} \|p(\mathbf{A})\|$$

- Which (known) approximation problem is solved?

$$\min_{p\in\pi_k} \|p(\mathbf{A})\| \;=\; \min_{p\in\pi_k} \|\mathbf{Q}p(\mathbf{\Lambda})\mathbf{Q}^*\| \;=\; \min_{p\in\pi_k} \max_{\lambda_i} |p(\lambda_i)|\,.$$

- Is the solution unique? **Yes**

Worst-case GMRES for normal matrices

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^*, \quad \mathbf{Q}^*\mathbf{Q} = \mathbf{I}\,.$$

- [Greenbaum & Gurvits '94, Joubert '94] showed:

$$\max_{\|b\|=1} \min_{p \in \pi_k} \|p(\mathbf{A})b\| \;=\; \min_{p \in \pi_k} \|p(\mathbf{A})\|$$

- Which (known) approximation problem is solved?

$$\min_{p \in \pi_k} \|p(\mathbf{A})\| \;=\; \min_{p \in \pi_k} \|\mathbf{Q}p(\mathbf{\Lambda})\mathbf{Q}^*\| \;=\; \min_{p \in \pi_k} \max_{\lambda_i} |p(\lambda_i)|\,.$$

- Is the solution unique? **Yes**
- How to approximate ideal/worst-case quantities?

Krylov matrix:

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

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

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

GMRES for normal matrices

Factorization of Krylov matrix

Krylov matrix:

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

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

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

Factorization:

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

where

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

GMRES for normal matrices

GMRES residual norm

Residual r_k can be written as [Liesen et al. '02, Ipsen '00]

$$
\begin{aligned}
r_k &= \|r_k\|^2 \, (\mathbf{K}_{k+1}^{\dagger})^H \, e_1 \\
&= \|r_k\|^2 \, \mathbf{Q} \, \left[(\mathbf{D}\mathbf{V}_{k+1})^{\dagger} \right]^H \, e_1 \, .
\end{aligned}
$$

GMRES for normal matrices

GMRES residual norm

Residual r_k can be written as [Liesen et al. '02, Ipsen '00]

$$\begin{aligned} r_k &= \|r_k\|^2 \, (\mathbf{K}_{k+1}^{\dagger})^H \, e_1 \\ &= \|r_k\|^2 \, \mathbf{Q} \, \left[(\mathbf{D}\mathbf{V}_{k+1})^{\dagger} \right]^H \, e_1 \, . \end{aligned}$$

and

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

(Assumption: $\mathbf{K}_{k+1}$ 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]

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

where

$$\ell_j \;\equiv\; \prod_{\substack{i=1\\ i\neq j}}^{n} \frac{\lambda_i}{\lambda_i - \lambda_j}\,.$$

GMRES residual norm (next-to-last step)

Let $\varrho_j \neq 0$ for all j. Then [Liesen & T. '04, Ipsen '00]

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

where

$$\ell_j \;\equiv\; \prod_{\substack{i=1 \\ i\neq j}}^{n} \frac{\lambda_i}{\lambda_i - \lambda_j}\,.$$

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

$$\|r^w_{n-1}\| \;=\; \frac{1}{\sum\limits_{j=1}^{n} |\ell_j|} \;=\; \min_{p\in\pi_{n-1}} \max_{\lambda_i} |p(\lambda_i)|\,.$$

Worst-case residual norm in a general step k

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

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

Worst-case residual norm in a general step k

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

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

- We want to determine the value $M_k^L \;=\; \|r_k^w\|$.

Worst-case residual norm in a general step k

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

$$M_k^S \equiv \min_{p \in \pi_k} \max_{\lambda_j \in S} |p(\lambda_j)| \,.$$

- We want to determine the value $M_k^L = \|r_k^w\|$.
- We are able to determine

$$M_k^S = \left(\sum_{j=1}^{k+1} |\ell_j^S| \right)^{-1}, \quad S \subseteq L, \;\; |S| = k+1.$$

Worst-case residual norm in a general step k

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

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

- We want to determine the value $M_k^L \;=\; \|r_k^w\|$.
- We are able to determine

$$M_k^S \;=\; \left(\sum_{j=1}^{k+1} |\ell_j^S| \right)^{-1}, \quad S \subseteq L, \;\; |S| = k+1.$$

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

$$M_k^L \;\geq\; \max_{\substack{S \subseteq L \\ |S| = k+1}} M_k^S \;\equiv\; B_k^L .$$

lower bound

Tightness of the bound

All eigenvalues are real

Approximation theory:

There exists a set $S \subseteq L$, $|S| = k+1$ such that

$$\underbrace{\min_{p \in \pi_k} \max_{\lambda_j \in S} |p(\lambda_j)|}_{M_k^S} = \underbrace{\min_{p \in \pi_k} \max_{\lambda_j \in L} |p(\lambda_j)|}_{M_k^L}$$

i.e.

$$\|r_k^w\| = M_k^L = M_k^S = \frac{1}{\sum_{j=1}^{k+1} \prod_{\substack{i=1 \\ i \neq j}}^{n} \frac{|\lambda_i^S|}{|\lambda_i^S - \lambda_j^S|}} = B_k^L .$$

[Liesen & T. '04, Greenbaum '79]

Tightness of the bound

Complex eigenvalues

Approximation theory:

The smallest set $S \subseteq L$ for which $M_k^L = M_k^S$ might contain as many as $2k+1$ distinct elements in the general complex case.

We proved that : [Liesen & T. '04]

$$B_k^L \;\leq\; \|r_k^w\| \;\leq\; \sqrt{(k+1)(n-k)}\; B_k^L\,.$$

Tightness of the bound

Complex eigenvalues

Approximation theory:

The smallest set $S \subseteq L$ for which $M_k^L = M_k^S$ might contain as many as $2k+1$ distinct elements in the general complex case.

We proved that : [Liesen & T. '04]

$$B_k^L \;\le\; \|r_k^w\| \;\le\; \sqrt{(k+1)(n-k)}\; B_k^L\,.$$

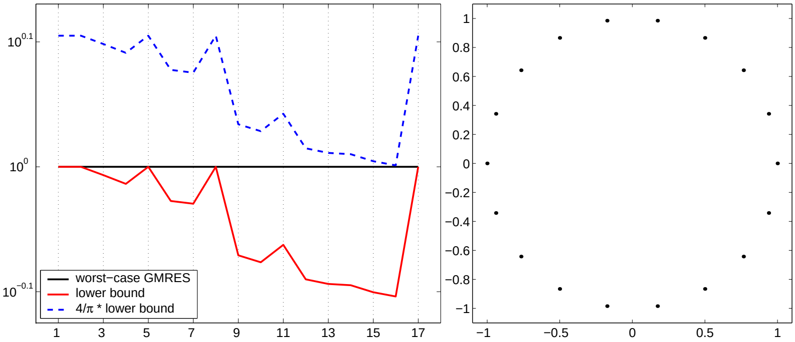
Moreover, we conjecture:

$$B_k^L \;\le\; \|r_k^w\| \;\le\; \frac{4}{\pi}\, B_k^L\,.$$

Experiment 1: roots of unity

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

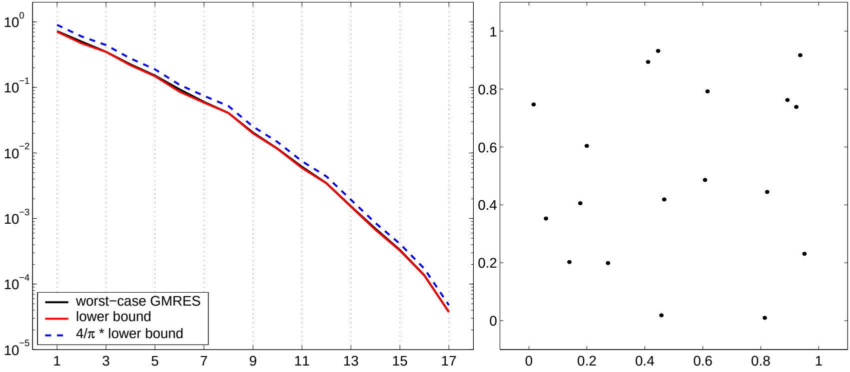
$$1 \;=\; \|r_i^w\|, \quad i = 0, \ldots, n-1.$$



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

Experiment 2: random eigenvalues

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



Interesting open problem

Approximation theory

Conjecture: Let $L = \{\lambda_1, \ldots, \lambda_n\}$ be a set of n distinct points in the complex plane. Then there exists a subset $S \subset L$ containing $k+1$ points such that

$$\underbrace{\min_{p \in \pi_k} \max_{\lambda_j \in S} |p(\lambda_j)|}_{M_k^S} \leq \underbrace{\min_{p \in \pi_k} \max_{\lambda_j \in L} |p(\lambda_j)|}_{M_k^L = \|r_k^w\|} \leq \frac{4}{\pi} \underbrace{\min_{p \in \pi_k} \max_{\lambda_j \in S} |p(\lambda_j)|}_{M_k^S}$$

and M_k^S can be evaluated as

$$M_k^S = \frac{1}{\sum_{j=1}^{k+1} \prod_{\substack{i=1 \\ i \neq j}}^{n} \frac{|\lambda_i^S|}{|\lambda_i^S - \lambda_j^S|}}.$$

Outline

Questions

$$\|r_k\| \;\leq\; \underbrace{\max_{\|b\|=1}\min_{p\in\pi_k} \|p(\mathbf{A})b\|}_{\psi_k(\mathbf{A})} \;\leq\; \underbrace{\min_{p\in\pi_k} \|p(\mathbf{A})\|}_{\varphi_k(\mathbf{A})}$$

- Which (known) approximation problem is solved?
- How to approximate ideal/worst-case quantities?
- When does it hold that ideal $=$ worst-case GMRES?
- Is the solution unique?

Toh's example

Worst-case GMRES can be very different from ideal GMRES!

Consider the 4 by 4 matrix

$$\mathbf{A} = \begin{bmatrix} 1 & \epsilon & & \\ & -1 & \epsilon^{-1} & \\ & & 1 & \epsilon \\ & & & -1 \end{bmatrix}, \qquad \epsilon > 0\,.$$

Then, for $k = 3$,

$$0 \stackrel{\epsilon \to 0}{\longleftarrow} \psi_k(\mathbf{A}) < \varphi_k(\mathbf{A}) = \frac{4}{5}\,.$$

[Toh '97, another example in Faber et al. '96]

Results concerning $\psi_k(\mathbf{A})$ and $\varphi_k(\mathbf{A})$

Theorem [Joubert '94, Faber et al. '96]

Let $\mathbf{A} \in \mathbb{C}^{n \times n}$ be a matrix with minimal polynomial degree $d(\mathbf{A})$. Then the following statements hold:

1. $\psi_0(\mathbf{A}) = \varphi_0(\mathbf{A}) = 1$.

Results concerning $\psi_k(\mathbf{A})$ and $\varphi_k(\mathbf{A})$

Theorem [Joubert '94, Faber et al. '96]

Let $\mathbf{A} \in \mathbb{C}^{n \times n}$ be a matrix with minimal polynomial degree $d(\mathbf{A})$. Then the following statements hold:

1. $\psi_0(\mathbf{A}) = \varphi_0(\mathbf{A}) = 1$.
2. $\psi_k(\mathbf{A})$ and $\varphi_k(\mathbf{A})$ are both nonincreasing in k .

Results concerning $\psi_k(\mathbf{A})$ and $\varphi_k(\mathbf{A})$

Theorem [Joubert '94, Faber et al. '96]

Let $\mathbf{A} \in \mathbb{C}^{n\times n}$ be a matrix with minimal polynomial degree $d(\mathbf{A})$. Then the following statements hold:

1. $\psi_0(\mathbf{A}) = \varphi_0(\mathbf{A}) = 1$.
2. $\psi_k(\mathbf{A})$ and $\varphi_k(\mathbf{A})$ are both nonincreasing in k .
3. $0 < \psi_k(\mathbf{A}) \leq \varphi_k(\mathbf{A})$ for $0 < k < d(\mathbf{A})$.

Results concerning $\psi_k(\mathbf{A})$ and $\varphi_k(\mathbf{A})$

Theorem [Joubert '94, Faber et al. '96]

Let $\mathbf{A} \in \mathbb{C}^{n \times n}$ be a matrix with minimal polynomial degree $d(\mathbf{A})$. Then the following statements hold:

1. $\psi_0(\mathbf{A}) = \varphi_0(\mathbf{A}) = 1$.
2. $\psi_k(\mathbf{A})$ and $\varphi_k(\mathbf{A})$ are both nonincreasing in k .
3. $0 < \psi_k(\mathbf{A}) \leq \varphi_k(\mathbf{A})$ for $0 < k < d(\mathbf{A})$.
4. If $\mathbf{A}$ is nonsingular, then $\psi_k(\mathbf{A}) = \varphi_k(\mathbf{A}) = 0$ for all $k \geq d(\mathbf{A})$.

Results concerning $\psi_k(\mathbf{A})$ and $\varphi_k(\mathbf{A})$

Theorem [Joubert '94, Faber et al. '96]

Let $\mathbf{A} \in \mathbb{C}^{n\times n}$ be a matrix with minimal polynomial degree $d(\mathbf{A})$. Then the following statements hold:

1. $\psi_0(\mathbf{A}) = \varphi_0(\mathbf{A}) = 1$.
2. $\psi_k(\mathbf{A})$ and $\varphi_k(\mathbf{A})$ are both nonincreasing in k .
3. $0 < \psi_k(\mathbf{A}) \leq \varphi_k(\mathbf{A})$ for $0 < k < d(\mathbf{A})$.
4. If $\mathbf{A}$ is nonsingular, then $\psi_k(\mathbf{A}) = \varphi_k(\mathbf{A}) = 0$ for all $k \geq d(\mathbf{A})$.
5. If $\mathbf{A}$ is singular, then $\psi_k(\mathbf{A}) = \varphi_k(\mathbf{A}) = 1$ for all $k \geq 0$.

Ideal GMRES polynomial and ideal GMRES matrix

Definition

The polynomial $p_* \in \pi_k$ is called the kth ideal GMRES polynomial of $\mathbf{A} \in \mathbb{C}^{n\times n}$, if it satisfies

$$\|p_*(\mathbf{A})\| = \min_{p\in\pi_k} \|p(\mathbf{A})\| .$$

We call the matrix $p_*(\mathbf{A})$ the kth ideal GMRES matrix of $\mathbf{A}$.

Existence and uniqueness of p_* proved by

[Greenbaum & Trefethen '94]

Results concerning $\psi_k(\mathbf{A}) = \varphi_k(\mathbf{A})$

When does it hold that

$$\underbrace{\max_{\|b\|=1} \min_{p\in\pi_k} \|p(\mathbf{A})b\|}_{\psi_k(\mathbf{A})} = \underbrace{\min_{p\in\pi_k} \|p(\mathbf{A})\|}_{\varphi_k(\mathbf{A})} \,?$$

Results concerning $\psi_k(\mathbf{A}) = \varphi_k(\mathbf{A})$

When does it hold that

$$\underbrace{\max_{\|b\|=1} \min_{p \in \pi_k} \|p(\mathbf{A})b\|}_{\psi_k(\mathbf{A})} = \underbrace{\min_{p \in \pi_k} \|p(\mathbf{A})\|}_{\varphi_k(\mathbf{A})} \ ?$$

[Greenbaum & Gurvits '94, Joubert '94]:

- if $\mathbf{A}$ is normal,
- for $k = 1$,
- if $p_*(\mathbf{A})$ has a simple maximal singular value.

Results concerning $\psi_k(\mathbf{A}) = \varphi_k(\mathbf{A})$

When does it hold that

$$\underbrace{\max_{\|b\|=1} \min_{p \in \pi_k} \|p(\mathbf{A})b\|}_{\psi_k(\mathbf{A})} = \underbrace{\min_{p \in \pi_k} \|p(\mathbf{A})\|}_{\varphi_k(\mathbf{A})} \, ?$$

[Greenbaum & Gurvits '94, Joubert '94]:

- if $\mathbf{A}$ is normal,
- for $k = 1$,
- if $p_*(\mathbf{A})$ has a simple maximal singular value.

[Faber et al. '96]:

Let $\mathbf{A}$ be n by n triangular Toeplitz matrix. Then

$$\max_{\|b\|=1} \min_{p \in \pi_k} \|p(\mathbf{A})b\| = 1 \quad \Longleftrightarrow \quad \min_{p \in \pi_k} \|p(\mathbf{A})\| = 1 \, .$$

Characterization of the situation $\psi_k(\mathbf{A}) = \varphi_k(\mathbf{A})$

Let $\Sigma(\mathbf{B})$ be the span of maximal right singular vectors of $\mathbf{B}$.

Lemma [T & Liesen & Faber '07, Faber et al. '96]

Let $\mathbf{A}$ be nonsingular and $1 < k < d(\mathbf{A})$.

Then $\psi_k(\mathbf{A}) = \varphi_k(\mathbf{A})$ **if and only if** there exist a polynomial $q \in \pi_k$ and a unit norm vector $b \in \Sigma(q(\mathbf{A}))$, such that

$$q(\mathbf{A})b \perp \mathbf{A}\mathcal{K}_k(\mathbf{A}, b)\,.$$

If such q and b exist, then $q = p_*$.

Characterization of the situation $\psi_k(\mathbf{A}) = \varphi_k(\mathbf{A})$

Let $\Sigma(\mathbf{B})$ be the span of maximal right singular vectors of $\mathbf{B}$.

Lemma [T & Liesen & Faber '07, Faber et al. '96]

Let $\mathbf{A}$ be nonsingular and $1 < k < d(\mathbf{A})$.

Then $\psi_k(\mathbf{A}) = \varphi_k(\mathbf{A})$ **if and only if** there exist a polynomial $q \in \pi_k$ and a unit norm vector $b \in \Sigma(q(\mathbf{A}))$, such that

$$q(\mathbf{A})b \perp \mathbf{A}\mathcal{K}_k(\mathbf{A}, b)\,.$$

If such q and b exist, then $q = p_*$.

Consequence: If $\psi_k(\mathbf{A}) = \varphi_k(\mathbf{A})$ then the worst-case GMRES polynomial is unique.

Simple maximal singular value of $p_*(\mathbf{A})$

Lemma [Greenbaum & Gurvits '94]

If $p_*(\mathbf{A})$ has a simple max. singular value then $\psi_k(\mathbf{A}) = \varphi_k(\mathbf{A})$.

Is this situation frequent or rare for nonnormal matrices?

Simple maximal singular value of $p_*(\mathbf{A})$

Lemma [Greenbaum & Gurvits '94]

If $p_*(\mathbf{A})$ has a simple max. singular value then $\psi_k(\mathbf{A}) = \varphi_k(\mathbf{A})$.

Is this situation frequent or rare for nonnormal matrices?

Normal case: $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^*$, $\mathbf{Q}^*\mathbf{Q} = \mathbf{I}$

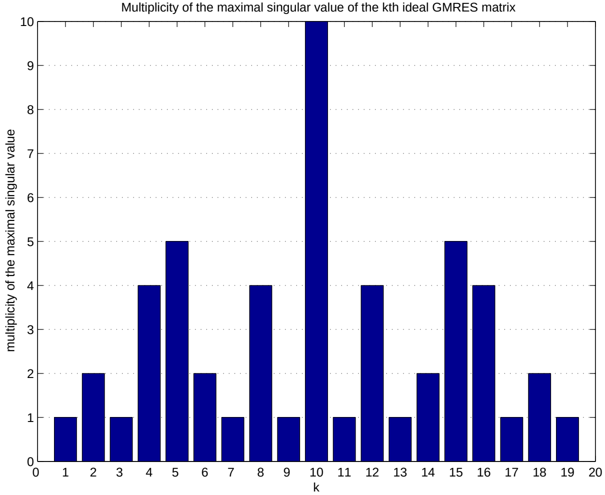
$$\min_{p \in \pi_k} \|p(\mathbf{A})\| = \min_{p \in \pi_k} \|\mathbf{Q} p(\mathbf{\Lambda}) \mathbf{Q}^*\| = \min_{p \in \pi_k} \max_{\lambda_i} |p(\lambda_i)| \,.$$

$p_*(\xi)$ attains its maximum value on at least $k+1$ eigenvalues, i.e. **the multiplicity** of max. sing. value of $p_*(\mathbf{A})$ **is at least** $k+1$.

Multiplicity of the maximal singular value of $p_*(\mathbf{J}_\lambda)$

computed using the software SDPT3 by Toh

Jordan block $\mathbf{J}_\lambda$, $\lambda = 1$, $n = 20$.



k-dimensional generalized field of values of $\mathbf{A}$

$$F_k(\mathbf{A}) \equiv \left\{ \begin{pmatrix} v^*\mathbf{A}v \\ \vdots \\ v^*\mathbf{A}^k v \end{pmatrix} \in \mathbb{C}^k \; : v^*v = 1 \right\}$$

k-dimensional generalized field of values of $\mathbf{A}$

$$F_k(\mathbf{A}) \equiv \left\{ \begin{pmatrix} v^*\mathbf{A}v \\ \vdots \\ v^*\mathbf{A}^k v \end{pmatrix} \in \mathbb{C}^k \ : v^*v = 1 \right\}$$

Theorem [Faber et al. '96]

For a nonsingular matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ the following statements hold:

- $\psi_k(\mathbf{A}) = 1 \iff \mathbf{0} \in F_k(\mathbf{A})$,
- $\varphi_k(\mathbf{A}) = 1 \iff \mathbf{0} \in \text{cvx}[\,F_k(\mathbf{A})\,]$.

k-dimensional generalized field of values of $\mathbf{A}$

$$F_k(\mathbf{A}) \equiv \left\{ \begin{pmatrix} v^* \mathbf{A} v \\ \vdots \\ v^* \mathbf{A}^k v \end{pmatrix} \in \mathbb{C}^k \ : v^* v = 1 \right\}$$

Theorem [Faber et al. '96]

For a nonsingular matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ the following statements hold:

- $\psi_k(\mathbf{A}) = 1 \iff \mathbf{0} \in F_k(\mathbf{A})$,
- $\varphi_k(\mathbf{A}) = 1 \iff \mathbf{0} \in \mathrm{cvx}[\, F_k(\mathbf{A})\,]$.

If $F_k(\mathbf{A})$ is convex then

$$\psi_k(\mathbf{A}) = 1 \quad \iff \quad \varphi_k(\mathbf{A}) = 1\,.$$

A possible connection

Using $F_k(\mathbf{A})$, it is possible to define two sets

$$\begin{aligned} \mathscr{G}_k(\mathbf{A}) &= \{\xi \in \mathbb{C} \;:\; \mathbf{0} \in F_k(\mathbf{A} - \xi\mathbf{I})\} \\ \mathscr{H}_k(\mathbf{A}) &= \{\xi \in \mathbb{C} \;:\; \mathbf{0} \in \mathrm{cvx}[\, F_k(\mathbf{A} - \xi\mathbf{I})\,]\}. \end{aligned}$$

[Nevanlinna '93, Greenbaum '02]

A possible connection

Using $F_k(\mathbf{A})$, it is possible to define two sets

$$\begin{aligned} \mathscr{G}_k(\mathbf{A}) &= \{\xi \in \mathbb{C} \;:\; \mathbf{0} \in F_k(\mathbf{A} - \xi\mathbf{I})\} \\ \mathscr{H}_k(\mathbf{A}) &= \{\xi \in \mathbb{C} \;:\; \mathbf{0} \in \mathrm{cvx}[\, F_k(\mathbf{A} - \xi\mathbf{I}) \,]\}. \end{aligned}$$

[Nevanlinna '93, Greenbaum '02]

Equivalent definitions:

$$\begin{aligned} \mathscr{G}_k(\mathbf{A}) &= \{\xi \in \mathbb{C} \;:\; \exists b \,\forall p \in \mathcal{P}_k \; |p(\xi)| \;\le\; \|p(\mathbf{A})b\| \,\}, \\ \mathscr{H}_k(\mathbf{A}) &= \{\xi \in \mathbb{C} \;:\; \quad \forall p \in \mathcal{P}_k \; |p(\xi)| \;\le\; \|p(\mathbf{A})\| \,\}, \end{aligned}$$

[Greenbaum '02, T. '??]

where $\mathcal{P}_k$ denotes the set of polynomials of degree k or less.

A possible connection

Using $F_k(\mathbf{A})$, it is possible to define two sets

$$\begin{aligned}\mathscr{G}_k(\mathbf{A}) &= \{\xi \in \mathbb{C} \ : \ \mathbf{0} \in F_k(\mathbf{A} - \xi \mathbf{I})\} \\ \mathscr{H}_k(\mathbf{A}) &= \{\xi \in \mathbb{C} \ : \ \mathbf{0} \in \mathrm{cvx}[\, F_k(\mathbf{A} - \xi \mathbf{I})\,]\}.\end{aligned}$$

[Nevanlinna '93, Greenbaum '02]

Equivalent definitions:

$$\begin{aligned}\mathscr{G}_k(\mathbf{A}) &= \{\xi \in \mathbb{C} \ : \ \exists b \, \forall p \in \mathcal{P}_k \ |p(\xi)| \ \leq \ \|p(\mathbf{A})b\| \,\}\,, \\ \mathscr{H}_k(\mathbf{A}) &= \{\xi \in \mathbb{C} \ : \ \forall p \in \mathcal{P}_k \ |p(\xi)| \ \leq \ \|p(\mathbf{A})\| \,\}\,,\end{aligned}$$

[Greenbaum '02, T. '??]

where $\mathcal{P}_k$ denotes the set of polynomials of degree k or less.

There might be a connection between convexity of $F_k(\mathbf{A})$ and the relation between ideal and worst-case GMRES.

Open problem

When does it hold that

$$\underbrace{\max_{\|b\|=1} \min_{p \in \pi_k} \|p(\mathbf{A})b\|}_{\psi_k(\mathbf{A})} = \underbrace{\min_{p \in \pi_k} \|p(\mathbf{A})\|}_{\varphi_k(\mathbf{A})} \ ?$$

- In both known examples of matrices $\mathbf{A}$ such that $\psi_k(\mathbf{A}) < \varphi_k(\mathbf{A})$, $F_k(\mathbf{A})$ is not convex.

Open problem

When does it hold that

$$\underbrace{\max_{\|b\|=1} \min_{p\in\pi_k} \|p(\mathbf{A})b\|}_{\psi_k(\mathbf{A})} = \underbrace{\min_{p\in\pi_k} \|p(\mathbf{A})\|}_{\varphi_k(\mathbf{A})} \,?$$

- In both known examples of matrices $\mathbf{A}$ such that $\psi_k(\mathbf{A}) < \varphi_k(\mathbf{A})$, $F_k(\mathbf{A})$ is not convex.
- For $k = 1$, $F_k(\mathbf{A})$ is always convex and $\psi_k(\mathbf{A}) = \varphi_k(\mathbf{A})$.

Open problem

When does it hold that

$$\underbrace{\max_{\|b\|=1} \min_{p \in \pi_k} \|p(\mathbf{A})b\|}_{\psi_k(\mathbf{A})} = \underbrace{\min_{p \in \pi_k} \|p(\mathbf{A})\|}_{\varphi_k(\mathbf{A})} \, ?$$

- In both known examples of matrices $\mathbf{A}$ such that $\psi_k(\mathbf{A}) < \varphi_k(\mathbf{A})$, $F_k(\mathbf{A})$ is not convex.
- For $k = 1$, $F_k(\mathbf{A})$ is always convex and $\psi_k(\mathbf{A}) = \varphi_k(\mathbf{A})$.
- Is the convexity of $F_k(\mathbf{A})$ sufficient?

Uniqueness

Let $\mathbf{A}$ be a nonsingular matrix. Then the kth ideal GMRES polynomial $p_* \in \pi_k$ that solves the problem

$$\min_{p \in \pi_k} \|p(\mathbf{A})\|$$

is unique.

[Greenbaum & Trefethen '94]

Uniqueness

Let $\mathbf{A}$ be a nonsingular matrix. Then the kth ideal GMRES polynomial $p_* \in \pi_k$ that solves the problem

$$\min_{p \in \pi_k} \|p(\mathbf{A})\|$$

is unique.

[Greenbaum & Trefethen '94]

Corrected proof and generalization can be found in

[Liesen & T. '09, accepted to SIMAX]

Uniqueness

Let $\mathbf{A}$ be a nonsingular matrix. Then the kth ideal GMRES polynomial $p_* \in \pi_k$ that solves the problem

$$\min_{p \in \pi_k} \|p(\mathbf{A})\|$$

is unique.

[Greenbaum & Trefethen '94]

Corrected proof and generalization can be found in

[Liesen & T. '09, accepted to SIMAX]

What can be said about a polynomial $q_* \in \pi_k$ that solves

$$\max_{\|b\|=1} \min_{p \in \pi_k} \|p(\mathbf{A})b\| \, ?$$

Uniqueness

Let $\mathbf{A}$ be a nonsingular matrix. Then the kth ideal GMRES polynomial $p_* \in \pi_k$ that solves the problem

$$\min_{p \in \pi_k} \|p(\mathbf{A})\|$$

is unique.

[Greenbaum & Trefethen '94]

Corrected proof and generalization can be found in

[Liesen & T. '09, accepted to SIMAX]

What can be said about a polynomial $q_* \in \pi_k$ that solves

$$\max_{\|b\|=1} \min_{p \in \pi_k} \|p(\mathbf{A})b\| \, ?$$

- If $\psi_k(\mathbf{A}) = \varphi_k(\mathbf{A})$ then q_* is unique and $q_* = p_*$.
- If $\psi_k(\mathbf{A}) < \varphi_k(\mathbf{A})$ ... uniqueness in an open problem.

Estimating the ideal GMRES approximation

How to estimate the ideal GMRES approximation

$$\min_{p \in \pi_k} \|p(\mathbf{A})\| \ ?$$

Estimating the ideal GMRES approximation

How to estimate the ideal GMRES approximation

$$\min_{p \in \pi_k} \|p(\mathbf{A})\| \ ?$$

Try to determine sets $\Omega \subset \mathbb{C}$ associated with $\mathbf{A}$ such that

$$\min_{p \in \pi_k} \|p(\mathbf{A})\| \sim \min_{p \in \pi_k} \max_{z \in \Omega} |p(z)|.$$

Estimating the ideal GMRES approximation

How to estimate the ideal GMRES approximation

$$\min_{p \in \pi_k} \|p(\mathbf{A})\| \; ?$$

Try to determine sets $\Omega \subset \mathbb{C}$ associated with $\mathbf{A}$ such that

$$\min_{p \in \pi_k} \|p(\mathbf{A})\| \sim \min_{p \in \pi_k} \max_{z \in \Omega} |p(z)|.$$

- If $\mathbf{A}$ is normal then $\Omega = L$.

Estimating the ideal GMRES approximation

How to estimate the ideal GMRES approximation

$$\min_{p \in \pi_k} \|p(\mathbf{A})\| \ ?$$

Try to determine sets $\Omega \subset \mathbb{C}$ associated with $\mathbf{A}$ such that

$$\min_{p \in \pi_k} \|p(\mathbf{A})\| \sim \min_{p \in \pi_k} \max_{z \in \Omega} |p(z)|.$$

- If $\mathbf{A}$ is normal then $\Omega = L$.
- If $\mathbf{A}$ is nonnormal then there are several approaches:

Estimating the ideal GMRES approximation

How to estimate the ideal GMRES approximation

$$\min_{p\in\pi_k} \|p(\mathbf{A})\| \ ?$$

Try to determine sets $\Omega \subset \mathbb{C}$ associated with $\mathbf{A}$ such that

$$\min_{p\in\pi_k} \|p(\mathbf{A})\| \sim \min_{p\in\pi_k} \max_{z\in\Omega} |p(z)|.$$

- If $\mathbf{A}$ is normal then $\Omega = L$.
- If $\mathbf{A}$ is nonnormal then there are several approaches:
  - $\Omega = \varepsilon$-pseudospectrum of $\mathbf{A}$, [Trefethen '90]

Estimating the ideal GMRES approximation

How to estimate the ideal GMRES approximation

$$\min_{p\in\pi_k} \|p(\mathbf{A})\| \ ?$$

Try to determine sets $\Omega \subset \mathbb{C}$ associated with $\mathbf{A}$ such that

$$\min_{p\in\pi_k} \|p(\mathbf{A})\| \sim \min_{p\in\pi_k} \max_{z\in\Omega} |p(z)|.$$

- If $\mathbf{A}$ is normal then $\Omega = L$.
- If $\mathbf{A}$ is nonnormal then there are several approaches:
  - $\Omega = \varepsilon$-pseudospectrum of $\mathbf{A}$, [Trefethen '90]
  - $\Omega =$ polynomial num. hull of $\mathbf{A}$. [Nevanlinna '93, Greenbaum '02]

Polynomial numerical hull

Definition

Let $\mathbf{A}$ be n by n matrix. Polynomial numerical hull of degree k is a set in the complex plane defined by

$$\mathscr{H}_k(\mathbf{A}) \;\equiv\; \{z \in \mathbb{C} : \; |p(z)| \le \|p(\mathbf{A})\| \;\; \forall\, p \in \mathcal{P}_k\},$$

where $\mathcal{P}_k$ denotes the set of polynomials of degree k or less.

Polynomial numerical hull

Definition

Let $\mathbf{A}$ be n by n matrix. Polynomial numerical hull of degree k is a set in the complex plane defined by

$$\mathscr{H}_k(\mathbf{A}) \equiv \{z \in \mathbb{C} : \; |p(z)| \leq \|p(\mathbf{A})\| \;\; \forall \, p \in \mathcal{P}_k\},$$

where $\mathcal{P}_k$ denotes the set of polynomials of degree k or less.

The set $\mathscr{H}_k(\mathbf{A})$ provides a lower bound

$$\min_{p \in \pi_k} \max_{z \in \mathscr{H}_k} |p(z)| \; \leq \; \min_{p \in \pi_k} \|p(\mathbf{A})\| \, .$$

[Greenbaum '02]

Polynomial numerical hull

Definition

Let $\mathbf{A}$ be n by n matrix. Polynomial numerical hull of degree k is a set in the complex plane defined by

$$\mathscr{H}_k(\mathbf{A}) \equiv \{z \in \mathbb{C} : \; |p(z)| \leq \|p(\mathbf{A})\| \;\; \forall \, p \in \mathcal{P}_k\},$$

where $\mathcal{P}_k$ denotes the set of polynomials of degree k or less.

The set $\mathscr{H}_k(\mathbf{A})$ provides a lower bound

$$\min_{p \in \pi_k} \max_{z \in \mathscr{H}_k} |p(z)| \;\leq\; \min_{p \in \pi_k} \|p(\mathbf{A})\|\,.$$

[Greenbaum '02]

How do these sets look like for classes of nonnormal matrices?

$\mathscr{H}_k(\mathbf{J}_\lambda)$ for a Jordan block $\mathbf{J}_\lambda$

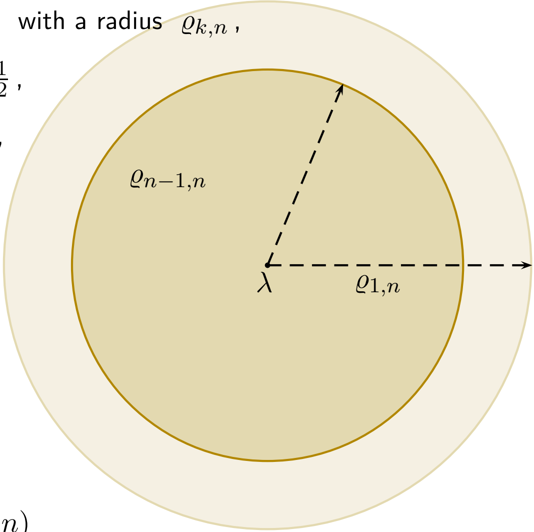
$\mathscr{H}_k(\mathbf{J}_\lambda)$ is a circle around λ with a radius $\varrho_{k,n}$,

$1 > \varrho_{1,n} > \cdots > \varrho_{n-1,n} \geq \frac{1}{2}$,

$\varrho_{1,n}$ and $\varrho_{n-1,n}$ are known,

[Faber et al. '03]

$$\varrho_{1,n} = \cos\left(\frac{\pi}{n+1}\right).$$



If n is even,

$\varrho_{n-1,n}$ is the positive root of

$$2\varrho^n + \varrho - 1 \ = \ 0.$$

$$\varrho_{n-1,n} \ \geq \ 1 - \frac{\log(2n)}{n}.$$

Quality of the bound based on $\mathscr{H}_k(\mathbf{J}_\lambda)$

Jordan block $\mathbf{J}_\lambda \in \mathbb{R}^{n\times n}$, $\lambda = 1$.

- For $k \le n/2$ it holds that

$$\frac{1}{2} \;\le\; \min_{p\in\pi_k} \max_{z\in\mathscr{H}_k} |p(z)| \;\le\; \min_{p\in\pi_k} \|p(\mathbf{A})\| \;\le\; 1\,.$$

[T. & Liesen & Faber ’07]

Quality of the bound based on $\mathscr{H}_k(\mathbf{J}_\lambda)$

Jordan block $\mathbf{J}_\lambda \in \mathbb{R}^{n\times n}$, $\lambda = 1$.

- For $k \leq n/2$ it holds that

$$\frac{1}{2} \leq \min_{p\in\pi_k} \max_{z\in\mathscr{H}_k} |p(z)| \leq \min_{p\in\pi_k} \|p(\mathbf{A})\| \leq 1\,.$$

- In later iterations, ideal GMRES converges slower that the lower bound predicts. For $k = n-1$ we have

$$\frac{1}{2n} \sim \min_{p\in\pi_k} \max_{z\in\mathscr{H}_k} |p(z)| \leq \min_{p\in\pi_k} \|p(\mathbf{A})\| \sim \frac{1}{1+\log(n)}\,.$$

[T. & Liesen & Faber '07]

Outline

1. Worst-case GMRES for normal matrices
2. Results for nonnormal matrices
3. Cross equality for worst-case GMRES vectors
4. Results for a Jordan block

Worst-case GMRES

For a given k, there exists a right hand side b^w such that

$$\|r_k^w\| \;=\; \min_{p\in\pi_k} \|p(\mathbf{A})b^w\| \;=\; \max_{\|b\|=1} \min_{p\in\pi_k} \|p(\mathbf{A})b\|$$

Theorem [Zavorin '02, T. '??]

Let $\mathbf{A} \in \mathbb{C}^{n\times n}$ be a nonsingular matrix. Then GMRES achieves the same worst-case behavior for $\mathbf{A}$ and $\mathbf{A}^*$ at every iteration.

- Zavorin '02 $\rightarrow$ only for diagonalizable matrices.
- T. '?? $\rightarrow$ for all nonsingular matrices.

Proof (for simplicity we consider everything real)

$\mathbf{A}$ is a given matrix, b is a unit norm starting vector,

$$\|r_k\| = \|p_b(\mathbf{A})b\| = \min_{p \in \pi_k} \|p(\mathbf{A})b\|, \qquad r_k \perp \mathbf{A}\mathcal{K}_k(\mathbf{A}, b).$$

Proof (for simplicity we consider everything real)

$\mathbf{A}$ is a given matrix, b is a unit norm starting vector,

$$\|r_k\| = \|p_b(\mathbf{A})b\| = \min_{p\in\pi_k} \|p(\mathbf{A})b\|, \qquad r_k \perp \mathbf{A}\mathcal{K}_k(\mathbf{A}, b).$$

Then, $\forall q \in \pi_k$ it holds that

$$\langle q(\mathbf{A})b, r_k\rangle = \langle b, r_k\rangle = \langle p_b(\mathbf{A})b, r_k\rangle = \|r_k\|^2.$$

Proof (for simplicity we consider everything real)

$\mathbf{A}$ is a given matrix, b is a unit norm starting vector,

$$\|r_k\| = \|p_b(\mathbf{A})b\| = \min_{p \in \pi_k} \|p(\mathbf{A})b\|, \qquad r_k \perp \mathbf{A}\mathcal{K}_k(\mathbf{A}, b).$$

Then, $\forall q \in \pi_k$ it holds that

$$\langle q(\mathbf{A})b, r_k \rangle = \langle b, r_k \rangle = \langle p_b(\mathbf{A})b, r_k \rangle = \|r_k\|^2.$$

Let b^w be a worst-case starting vector, $c \equiv r_k^w / \|r_k^w\|$, then

$$\psi_k(\mathbf{A}) = \|r_k^w\| = \langle q(\mathbf{A})b^w, c \rangle = \langle b^w, q(\mathbf{A}^T)c \rangle.$$

Proof (for simplicity we consider everything real)

$\mathbf{A}$ is a given matrix, b is a unit norm starting vector,

$$\|r_k\| = \|p_b(\mathbf{A})b\| = \min_{p \in \pi_k} \|p(\mathbf{A})b\|, \qquad r_k \perp \mathbf{A}\mathcal{K}_k(\mathbf{A}, b).$$

Then, $\forall q \in \pi_k$ it holds that

$$\langle q(\mathbf{A})b, r_k \rangle = \langle b, r_k \rangle = \langle p_b(\mathbf{A})b, r_k \rangle = \|r_k\|^2.$$

Let b^w be a worst-case starting vector, $c \equiv r_k^w / \|r_k^w\|$, then

$$\psi_k(\mathbf{A}) = \|r_k^w\| = \langle q(\mathbf{A})b^w, c \rangle = \langle b^w, q(\mathbf{A}^T)c \rangle.$$

Since $q \in \pi_k$ is arbitrary, we can also choose $q = q_c$,

$$\psi_k(\mathbf{A}) = \langle b^w, q_c(\mathbf{A}^T)c \rangle \leq \|q_c(\mathbf{A}^T)c\| \leq \psi_k(\mathbf{A}^T).$$

Proof (for simplicity we consider everything real)

$\mathbf{A}$ is a given matrix, b is a unit norm starting vector,

$$\|r_k\| = \|p_b(\mathbf{A})b\| = \min_{p \in \pi_k} \|p(\mathbf{A})b\|, \qquad r_k \perp \mathbf{A}\mathcal{K}_k(\mathbf{A}, b).$$

Then, $\forall q \in \pi_k$ it holds that

$$\langle q(\mathbf{A})b, r_k \rangle = \langle b, r_k \rangle = \langle p_b(\mathbf{A})b, r_k \rangle = \|r_k\|^2.$$

Let b^w be a worst-case starting vector, $c \equiv r_k^w / \|r_k^w\|$, then

$$\psi_k(\mathbf{A}) = \|r_k^w\| = \langle q(\mathbf{A})b^w, c \rangle = \langle b^w, q(\mathbf{A}^T)c \rangle.$$

Since $q \in \pi_k$ is arbitrary, we can also choose $q = q_c$,

$$\psi_k(\mathbf{A}) = \langle b^w, q_c(\mathbf{A}^T)c \rangle \leq \|q_c(\mathbf{A}^T)c\| \leq \psi_k(\mathbf{A}^T).$$

Switch the role of $\mathbf{A}$ and $\mathbf{A}^T$ to obtain the opposite inequality.

Another results

Since we know that $\psi_k(\mathbf{A}) = \psi_k(\mathbf{A}^T)$ and

$$\psi_k(\mathbf{A}) = \langle b^w, q_c(\mathbf{A}^T)c \rangle \leq \|q_c(\mathbf{A}^T)c\| \leq \psi_k(\mathbf{A}^T),$$

it holds that

$$\langle b^w, q_c(\mathbf{A}^T)c \rangle = \|q_c(\mathbf{A}^T)c\|.$$

Another results

Since we know that $\psi_k(\mathbf{A}) = \psi_k(\mathbf{A}^T)$ and

$$\psi_k(\mathbf{A}) = \langle b^w, q_c(\mathbf{A}^T)c \rangle \leq \|q_c(\mathbf{A}^T)c\| \leq \psi_k(\mathbf{A}^T),$$

it holds that

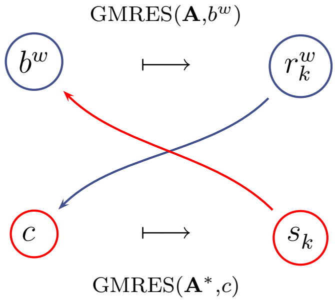
$$\langle b^w, q_c(\mathbf{A}^T)c \rangle = \|q_c(\mathbf{A}^T)c\|.$$

This is true iff

$$b^w = \frac{q_c(\mathbf{A}^T)c}{\|q_c(\mathbf{A}^T)c\|}.$$

Cross equality for worst-case GMRES vectors

Given: $\mathbf{A} \in \mathbb{C}^{n\times n}$, k



$$c \equiv \frac{r_k^w}{\|r_k^w\|}$$

It holds that

$$\|s_k\| = \|r_k^w\| = \psi_k(\mathbf{A}), \qquad b^w = \frac{s_k}{\|s_k\|}.$$

[Zavorin '02, T. '??]

Outline

Results for a Jordan block $\mathbf{J}_\lambda$

Consider an $n \times n$ Jordan block $\mathbf{J}_\lambda$, $\lambda \in \mathbb{C}$,

$\varrho_{k,n}$ ... the radius of the polynomial numerical hull $\mathscr{H}_k(\mathbf{J}_\lambda)$

$$\frac{1}{2} \leq \varrho_{k,n} < 1 .$$

Results for a Jordan block $\mathbf{J}_\lambda$

Consider an $n \times n$ Jordan block $\mathbf{J}_\lambda$, $\lambda \in \mathbb{C}$,

$\varrho_{k,n}$ ... the radius of the polynomial numerical hull $\mathscr{H}_k(\mathbf{J}_\lambda)$

$$\frac{1}{2} \leq \varrho_{k,n} < 1 .$$

$\psi_k(\mathbf{J}_\lambda) = \varphi_k(\mathbf{J}_\lambda)$

- if $|\lambda| \leq \varrho_{k,n}$, [Faber et al. '03, Faber et al. '96]
- if $|\lambda| \geq \varrho_{k,n-k}^{-1}$ and $k < n/2$,
- in steps k such that k divides n,
- in steps $n-k$ such that k divides n and $|\lambda| \geq 1$.

[T. & Liesen & Faber '07]

Ideal GMRES approximation $\varphi_k(\mathbf{J}_\lambda)$

- $|\lambda| \le \varrho_{k,n}$,

$$\varphi_k(\mathbf{J}_\lambda) = 1.$$

[Faber et al. '03, Faber et al. '96]

Ideal GMRES approximation $\varphi_k(\mathbf{J}_\lambda)$

- $|\lambda| \leq \varrho_{k,n}$,

$$\varphi_k(\mathbf{J}_\lambda) = 1.$$

[Faber et al. '03, Faber et al. '96]

- $|\lambda| \geq \varrho_{k,n-k}^{-1}$ and $k < n/2$,

$$\varphi_k(\mathbf{J}_\lambda) = |\lambda|^{-k}.$$

Ideal GMRES approximation $\varphi_k(\mathbf{J}_\lambda)$

- $|\lambda| \leq \varrho_{k,n}$,

$$\varphi_k(\mathbf{J}_\lambda) = 1.$$

[Faber et al. '03, Faber et al. '96]

- $|\lambda| \geq \varrho_{k,n-k}^{-1}$ and $k < n/2$,

$$\varphi_k(\mathbf{J}_\lambda) = |\lambda|^{-k}.$$

- steps $n-k$ such that k divides n and $|\lambda| \geq 1$,

$$\varphi_{n-k}(\mathbf{J}_\lambda) \;=\; \frac{1}{\lambda^{n-k}} \left[\sum_{i=0}^{n/k-1} \lambda^{-2ki} 4^{-2i} \binom{2i}{i}^2 \right]^{-1} .$$

[T. & Liesen & Faber '07]

Radius of polynomial numerical hull for $\mathbf{J}_\lambda$

Theorem

Let d be the greatest common divisor of n and k and define

$$\ell = \frac{k}{d}, \quad m = \frac{n}{d}.$$

Then

$$\varrho_{k,n} = \varrho_{\ell,m}^{1/d}.$$

Radius of polynomial numerical hull for $\mathbf{J}_\lambda$

Theorem

Let d be the greatest common divisor of n and k and define

$$\ell = \frac{k}{d}, \quad m = \frac{n}{d}.$$

Then

$$\varrho_{k,n} = \varrho_{\ell,m}^{1/d}.$$

Consider k such that k divides n. Then

$$\varrho_{k,n} = \left[\cos\left(\frac{\pi}{m+1}\right)\right]^{\frac{1}{k}}, \qquad \varrho_{n-k,n} = \varrho_{m-1,m}^{\frac{1}{k}}.$$

[T. & Liesen & Faber '07]

Conclusions

1. Ideal and worst-case GMRES for nonnormal matrices are not well understood.

Conclusions

1. Ideal and worst-case GMRES for nonnormal matrices are not well understood.

2. There might be a connection between the convexity of the generalized field of values and the relation between ideal and worst-case GMRES.

Conclusions

1. Ideal and worst-case GMRES for nonnormal matrices are not well understood.

2. There might be a connection between the convexity of the generalized field of values and the relation between ideal and worst-case GMRES.

3. Worst-case GMRES achieves the same convergence behavior for $\mathbf{A}$ and $\mathbf{A}^*$. Worst-case GMRES vectors satisfy the cross equality.

Conclusions

1. Ideal and worst-case GMRES for nonnormal matrices are not well understood.

2. There might be a connection between the convexity of the generalized field of values and the relation between ideal and worst-case GMRES.

3. Worst-case GMRES achieves the same convergence behavior for $\mathbf{A}$ and $\mathbf{A}^*$. Worst-case GMRES vectors satisfy the cross equality.

4. Based on numerical observation and theoretical results we conjecture that ideal GMRES $=$ worst-case GMRES for a Jordan block.

References

- J. Liesen and P. Tichý, [On best approximations of polynomials in matrices in the matrix 2-norm, accepted SIMAX, 2009.]
- P. Tichý, J. Liesen, and V. Faber, [On worst-case GMRES, ideal GMRES, and the polynomial numerical hull of a Jordan block, ETNA, 26 (2007), pp. 453–473.]
- J. Liesen and P. Tichý, [The worst-case GMRES for normal matrices, BIT Num. Math., 44 (2004), pp. 79–98.]
- V. Faber, A. Greenbaum, and D. E. Marshall, [The polynomial numerical hulls of Jordan blocks and related matrices, LAA, 374 (2003), pp. 231–246.]
- A. Greenbaum, [Generalizations of the field of values useful in the study of polynomial functions of a matrix, LAA., 347 (2002), pp. 233–249.]
- K. C. Toh, [GMRES vs. ideal GMRES, SIMAX, 18 (1997), pp. 30–36.]
- V. Faber, W. Joubert, E. Knill, and T. Manteuffel, [Minimal residual method stronger than polynomial preconditioning, SIMAX, 17 (1996), pp. 707–729.]

Thank you for your attention!

More details can be found at

http://www.cs.cas.cz/tichy
http://www.math.tu-berlin.de/~liesen