

On worst-case GMRES

Petr Tichý

joint work with

Vance Faber and Jörg Liesen

Czech Academy of Sciences

June 18, 2012,

SIAM Conference on Applied Linear Algebra,

Valencia, Spain

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

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

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

$$\leq \min_{p \in \pi_k} \|p(\mathbf{A})\| \equiv \varphi_k(\mathbf{A}) \quad (\text{ideal GMRES})$$

$$\leq 1.$$

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

- **Relationship** between **ideal** and **worst case** GMRES?
- **Characterization** of solutions? Understanding?
- Existence and **uniqueness** of the solution?
- How to approximate **ideal/worst-case** quantities?

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?

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

Nonnormal matrices – Toh's example

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

Consider the 4 by 4 matrix

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

Then, for $k = 3$,

$$0 \xleftarrow{\epsilon \rightarrow 0} \psi_k(\mathbf{A}) < \varphi_k(\mathbf{A}) = \frac{4}{5}.$$

[Toh '97; another example in Faber, Joubert, Knill, Manteuffel '96]

Uniqueness

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

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

is unique.

[Greenbaum, Trefethen '94]

Generalization to

$$\min_{p \in \mathcal{P}_m} \| f(\mathbf{A}) - p(\mathbf{A}) \|$$

can be found in [Liesen, T. '09].

Matrix approximation problems in spectral norm

and characterization of Ideal GMRES

Ideal GMRES is a special case of the problem

$$\min_{\mathbf{M} \in \mathbb{A}} \|\mathbf{B} - \mathbf{M}\| = \|\mathbf{B} - \mathbf{A}_*\|$$

\mathbf{A}_* is called a **spectral approximation** of \mathbf{B} from the subspace \mathbb{A} .

In our case,

$$\min_{p \in \pi_k} \|p(\mathbf{A})\| = \min_{\alpha_i \in \mathbb{C}} \left\| \mathbf{I} - \sum_{j=1}^k \alpha_j \mathbf{A}^j \right\|,$$

i.e. $\mathbf{B} = \mathbf{I}$, $\mathbb{A} = \text{span}\{\mathbf{A}, \dots, \mathbf{A}^k\}$.

General characterization by [Lau and Riha, 1981] and [Ziřtak, 1993, 1996]

→ based on the **Singer's theorem** [Singer, 1970].

Singer's theorem

Theorem

[Singer, 1970]

Let $(\mathcal{V}, \|\cdot\|)$ be a **normed vector space**, $\mathcal{X} \subset \mathcal{V}$ a k -dimensional linear subspace of \mathcal{V} , and $y \in \mathcal{V} \setminus \mathcal{X}$. $x_* \in \mathcal{X}$ is a **best approximation** of y **iff** there exists ℓ **extremal points** f_1, \dots, f_ℓ of Ω^* , where $1 \leq \ell \leq k + 1$ if the scalars are real and ℓ numbers $\omega_1, \dots, \omega_\ell > 0$, with $\omega_1 + \dots + \omega_\ell = 1$, such that

$$\begin{aligned} \sum_{i=1}^{\ell} \omega_i f_i(x) &= 0, & \forall x \in \mathcal{X}, \\ f_i(y - x_*) &= \|y - x_*\|, & i = 1, \dots, \ell. \end{aligned}$$

Specialized for matrices by [Lau and Riha, 1981].

Characterization of Ideal GMRES

by Faber, Joubert, Knill, Manteuffel '96

Given a polynomial $q \in \pi_k$ and \mathbf{A} , define the set

$$\Omega_k(q) \equiv \left\{ \left[\begin{array}{c} w^* q(\mathbf{A})^* \mathbf{A} w \\ \vdots \\ w^* q(\mathbf{A})^* \mathbf{A}^k w \end{array} \right] : w \in \Sigma(q(\mathbf{A})), \|w\| = 1 \right\}$$

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

Theorem

[Faber, Joubert, Knill, Manteuffel '96]

$h \in \pi_k$ is the k th ideal GMRES pol. of $\mathbf{A} \iff \mathbf{0} \in \text{cvx}(\Omega_k(h))$.

This characterization was derived without using Singer's theorem.

Worst-case GMRES

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

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

Theorem

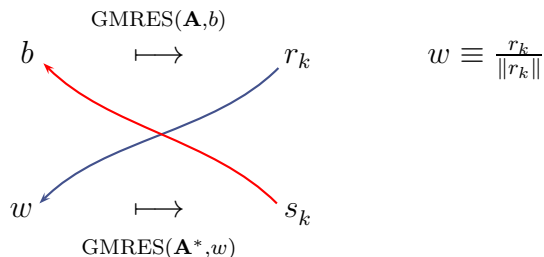
[Zavorin '02; Faber, T., Liesen '12]

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.
- [Faber, T., Liesen '12] \rightarrow for all nonsingular matrices.

Cross equality for worst-case GMRES vectors

Given: $\mathbf{A} \in \mathbb{C}^{n \times n}$, k , a worst-case starting vector b



It holds that

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

[Zavorin '02; Faber, T., Liesen '12]

A new characterization of worst-case GMRES

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be given. For $v \in \mathbb{R}^n$ and $c \in \mathbb{R}^k$ define

$$F(c, v) \equiv \frac{\|v - K(v)c\|^2}{\|v\|^2},$$

where $K(v) \equiv [\mathbf{A}v, \mathbf{A}^2v, \dots, \mathbf{A}^k v]$. We want to characterize the solution of the problem

$$\max_{v \in \mathbb{R}^n \setminus 0} \min_{c \in \mathbb{R}^k} F(c, v). \quad (1)$$

Theorem

[Faber, T., Liesen '12]

$\tilde{c} \in \mathbb{R}^k$ and $\tilde{v} \in \mathbb{R}^n$ that solve the problem (1) satisfy

$$\frac{\partial F}{\partial c}(\tilde{c}, \tilde{v}) = 0, \quad \frac{\partial F}{\partial v}(\tilde{c}, \tilde{v}) = 0,$$

i.e., (\tilde{c}, \tilde{v}) is a stationary point of the function $F(c, v)$.

Consequences of the new characterization

Let b , $\|b\| = 1$, be a worst-case starting vector and

$$r_k = p_k(\mathbf{A})b, \quad \|r_k\| = \psi_k(\mathbf{A})$$

the corresponding GMRES residual vector. Then

Theorem

[Faber, T., Liesen '12]

- b is a **right singular vector** of $p_k(\mathbf{A})$, i.e.

$$\psi_k(\mathbf{A})^2 b = p_k(\mathbf{A}^T)p_k(\mathbf{A})b.$$

- p_k is also a worst-case GMRES polynomial for \mathbf{A}^T .

Worst-case GMRES polynomials need not be unique

Theorem

[Faber, T., Liesen '12]

A worst-case GMRES polynomial for the Toh matrix

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

and the step $k = 3$ is **not unique**.

Example: If $\epsilon = 0.1$, then both

$$-39.9^{-1}(z - 1.181)(z + 0.939)(z + 35.96)$$

and

$$39.9^{-1}(z + 1.181)(z - 0.939)(z - 35.96)$$

are worst-case GMRES polynomials of \mathbf{A} .

Ideal versus worst-case GMRES

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

- $\psi_k(\mathbf{A}) = \varphi_k(\mathbf{A})$ iff a worst-case starting vector b is a **maximal right singular vector** of $p_k(\mathbf{A})$.

- If $\Omega_k(p_*)$ is **convex** then $\psi_k(\mathbf{A}) = \varphi_k(\mathbf{A})$.

[Faber, Joubert, Knill, Manteuffel '96]

- $\psi_k(\mathbf{A}) = \varphi_k(\mathbf{A})$ iff

$$\max_{v \in \mathbb{R}^n \setminus \{0\}} \min_{c \in \mathbb{R}^k} F(c, v) = \min_{c \in \mathbb{R}^k} \max_{v \in \mathbb{R}^n \setminus \{0\}} F(c, v).$$

Existence of a **saddle point** of $F(c, v)$?

[Faber, T., Liesen '12]

Summary on worst-case GMRES

- 1 Worst-case **starting vectors** satisfy the **cross equality**.
- 2 Worst-case GMRES **polynomials need not be unique**.
- 3 Worst-case GMRES **data** (\tilde{c}, \tilde{v}) are **stationary points** of the function $F(c, v)$.
- 4 Worst-case **starting vector** b is a **right singular vector** of the corresponding GMRES matrix $p_k(\mathbf{A})$.
- 5 $\psi_k(\mathbf{A}) = \varphi_k(\mathbf{A})$ **iff**
 b is a **maximal right singular vector** of $p_k(\mathbf{A})$.
- 6 $\psi_k(\mathbf{A}) = \varphi_k(\mathbf{A})$ **iff** $F(c, v)$ has a **saddle point**.

Related papers

- V. FABER, P. TICHÝ AND J. LIESEN, [Ideal and Worst-case GMRES: Characterization and examples, in preparation, (2012?).]
- J. LIESEN AND P. TICHÝ, [On best approximations of polynomials in matrices in the matrix 2-norm, SIMAX, 31 (2009), pp. 853–863.]
- 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.]
- A. GREENBAUM AND L. N. TREFETHEN, [GMRES/CR and Arnoldi/Lanczos as matrix approximation problems, SISC, 15 (1994), pp. 359–368.]

Thank you for your attention!