On best approximations of matrix polynomials

Petr Tichý

joint work with

Jörg Liesen

Institute of Computer Science,
Academy of Sciences of the Czech Republic

September 12, 2008
Technische Universität Hamburg-Harburg, Germany
GAMM Workshop on Applied and Numerical Linear Algebra

Ideal Arnoldi approximation problem

$$\min_{p\in\mathcal{M}_{m+1}} \| p(\mathbf{A}) \| = \min_{p\in\mathcal{P}_m} \| \mathbf{A}^{m+1} - p(\mathbf{A}) \| ,$$

where $\mathcal{M}_{m+1}$ is the class of monic polynomials of degree $m+1$,
$\mathcal{P}_m$ is the class of polynomials of degree at most m.

Ideal Arnoldi approximation problem

$$\min_{p \in \mathcal{M}_{m+1}} \| p(\mathbf{A}) \| = \min_{p \in \mathcal{P}_m} \| \mathbf{A}^{m+1} - p(\mathbf{A}) \| ,$$

where $\mathcal{M}_{m+1}$ is the class of monic polynomials of degree $m+1$, $\mathcal{P}_m$ is the class of polynomials of degree at most m.

- Introduced in [Greenbaum and Trefethen, 1994], paper contains uniqueness result ($\rightarrow$ story of the proof).
- The unique polynomial that solves the problem is called the $(m+1)$st ideal Arnoldi polynomial of $\mathbf{A}$, or the $(m+1)$st Chebyshev polynomial of $\mathbf{A}$.
- Some work on these polynomials in [Toh PhD thesis, 1996], [Toh and Trefethen, 1998], [Trefethen and Embree, 2005] .

Matrix function best approximation problem

We consider the matrix approximation problem

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

$\| \cdot \|$ is the spectral norm (matrix 2-norm), $\mathbf{A} \in \mathbb{C}^{n \times n}$, f is analytic in neighborhood of $\mathbf{A}$'s spectrum.

Matrix function best approximation problem

We consider the matrix approximation problem

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

$\| \cdot \|$ is the spectral norm (matrix 2-norm), $\mathbf{A} \in \mathbb{C}^{n \times n}$, f is analytic in neighborhood of $\mathbf{A}$'s spectrum.

Well known: $f(\mathbf{A}) = p_f(\mathbf{A})$ for a polynomial p_f depending on values and possibly derivatives of f on $\mathbf{A}$'s spectrum.

Without loss of generality we assume that f is a given polynomial.

Matrix function best approximation problem

We consider the matrix approximation problem

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

$\| \cdot \|$ is the spectral norm (matrix 2-norm), $\mathbf{A} \in \mathbb{C}^{n \times n}$, f is analytic in neighborhood of $\mathbf{A}$'s spectrum.

Well known: $f(\mathbf{A}) = p_f(\mathbf{A})$ for a polynomial p_f depending on values and possibly derivatives of f on $\mathbf{A}$'s spectrum.

Without loss of generality we assume that f is a given polynomial.

Does this problem have a unique solution $p_* \in \mathcal{P}_m$?

Outline

1. General matrix approximation problems
2. Formulation of matrix polynomial approximation problems
3. Uniqueness results
4. Ideal Arnoldi versus ideal GMRES polynomials

Outline

1. General matrix approximation problems
2. Formulation of matrix polynomial approximation problems
3. Uniqueness results
4. Ideal Arnoldi versus ideal GMRES polynomials

General matrix approximation problems

Given

- m linearly independent matrices $\mathbf{A}_1, \ldots, \mathbf{A}_m \in \mathbb{C}^{n \times n}$,
- $\mathbb{A} \equiv \mathrm{span}\,\{\mathbf{A}_1, \ldots, \mathbf{A}_m\}$,
- $\mathbf{B} \in \mathbb{C}^{n \times n} \backslash \mathbb{A}$,
- $\|\cdot\|$ is a matrix norm.

Consider the best approximation problem

$$\min_{\mathbf{M} \in \mathbb{A}} \|\, \mathbf{B} - \mathbf{M} \,\| \,.$$

General matrix approximation problems

Given

- m linearly independent matrices $\mathbf{A}_1, \ldots, \mathbf{A}_m \in \mathbb{C}^{n \times n}$,
- $\mathbb{A} \equiv \operatorname{span}\{\mathbf{A}_1, \ldots, \mathbf{A}_m\}$,
- $\mathbf{B} \in \mathbb{C}^{n \times n} \backslash \mathbb{A}$,
- $\| \cdot \|$ is a matrix norm.

Consider the best approximation problem

$$\min_{\mathbf{M} \in \mathbb{A}} \| \mathbf{B} - \mathbf{M} \| .$$

This problem has a unique solution if $\| \cdot \|$ is strictly convex.

[see, e.g., Sreedharan, 1973]

Strictly convex norms

The norm $\| \cdot \|$ is strictly convex if for all $\mathbf{X}$, $\mathbf{Y}$,

$$\|\mathbf{X}\| = \|\mathbf{Y}\| = 1\,, \quad \|\mathbf{X} + \mathbf{Y}\| = 2 \quad \Rightarrow \quad \mathbf{X} = \mathbf{Y}\,.$$

Strictly convex norms

The norm $\|\cdot\|$ is strictly convex if for all $\mathbf{X}$, $\mathbf{Y}$,

$$\|\mathbf{X}\| = \|\mathbf{Y}\| = 1\,, \quad \|\mathbf{X}+\mathbf{Y}\| = 2 \quad \Rightarrow \quad \mathbf{X} = \mathbf{Y}\,.$$

Which matrix norms are strictly convex?
Let $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$ be singular values of $\mathbf{X}$ and $1 \leq p \leq \infty$.

The c_p-norm: $$\|\mathbf{X}\|_p \equiv \left(\sum_{i=1}^{n} \sigma_i^p\right)^{1/p}\,.$$

- $p = 2$ . . . Frobenius norm,
- $p = \infty$ . . . spectral norm, matrix 2-norm, $\|\mathbf{X}\|_\infty = \sigma_1$,
- $p = 1$ . . . trace (nuclear) norm.

Strictly convex norms

The norm $\|\cdot\|$ is strictly convex if for all $\mathbf{X}$, $\mathbf{Y}$,

$$\|\mathbf{X}\| = \|\mathbf{Y}\| = 1\,, \quad \|\mathbf{X}+\mathbf{Y}\| = 2 \quad \Rightarrow \quad \mathbf{X} = \mathbf{Y}\,.$$

Which matrix norms are strictly convex?

Let $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$ be singular values of $\mathbf{X}$ and $1 \leq p \leq \infty$.

The c_p-norm: $$\|\mathbf{X}\|_p \equiv \left(\sum_{i=1}^{n} \sigma_i^p\right)^{1/p} .$$

- $p = 2$ ... Frobenius norm,
- $p = \infty$ ... spectral norm, matrix 2-norm, $\|\mathbf{X}\|_\infty = \sigma_1$,
- $p = 1$ ... trace (nuclear) norm.

Theorem. If $1 < p < \infty$ then the c_p-norm is strictly convex.

[see, e.g., Ziętak, 1988]

Spectral norm (matrix 2-norm)

A useful matrix norm in many applications: spectral norm

$$\|\mathbf{X}\| \equiv \sigma_1 \,.$$

Spectral norm (matrix 2-norm)

A useful matrix norm in many applications: spectral norm

$$\|\mathbf{X}\| \equiv \sigma_1 \, .$$

This norm is not strictly convex:

$$\mathbf{X} = \begin{bmatrix} \mathbf{I} & \\ & \varepsilon \end{bmatrix}, \qquad \mathbf{Y} = \begin{bmatrix} \mathbf{I} & \\ & \delta \end{bmatrix}, \qquad \varepsilon,\ \delta \in \langle 0, 1 \rangle \, .$$

Then we have, for each $\varepsilon,\ \delta \in \langle 0, 1 \rangle$,

$$\|\mathbf{X}\| = \|\mathbf{Y}\| = 1 \qquad \text{and} \qquad \|\mathbf{X} + \mathbf{Y}\| = 2$$

but if $\varepsilon \neq \delta$ then $\mathbf{X} \neq \mathbf{Y}$.

Spectral norm (matrix 2-norm)

A useful matrix norm in many applications: spectral norm

$$\|\mathbf{X}\| \equiv \sigma_1 \,.$$

This norm is not strictly convex:

$$\mathbf{X} = \begin{bmatrix} \mathbf{I} & \\ & \varepsilon \end{bmatrix}, \qquad \mathbf{Y} = \begin{bmatrix} \mathbf{I} & \\ & \delta \end{bmatrix}, \qquad \varepsilon,\ \delta \in \langle 0, 1 \rangle \,.$$

Then we have, for each $\varepsilon,\ \delta \in \langle 0, 1 \rangle$,

$$\|\mathbf{X}\| = \|\mathbf{Y}\| = 1 \qquad \text{and} \qquad \|\mathbf{X} + \mathbf{Y}\| = 2$$

but if $\varepsilon \neq \delta$ then $\mathbf{X} \neq \mathbf{Y}$.

Consequently: Best approximation problems in the spectral norm are not guaranteed to have a unique solution.

Matrix approximation problems in spectral norm

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

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

Open question: When does this problem have a unique solution?

Matrix approximation problems in spectral norm

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

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

Open question: When does this problem have a unique solution?

Ziętak's sufficient condition

Theorem [Ziętak, 1993]. If the residual matrix $\mathbf{B} - \mathbf{A}_*$ has an n-fold maximal singular value, then the spectral approximation $\mathbf{A}_*$ of $\mathbf{B}$ from the subspace $\mathbb{A}$ is unique.

Matrix approximation problems in spectral norm

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

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

Open question: When does this problem have a unique solution?

Ziętak's sufficient condition

Theorem [Ziętak, 1993]. If the residual matrix $\mathbf{B} - \mathbf{A}_*$ has an n-fold maximal singular value, then the spectral approximation $\mathbf{A}_*$ of $\mathbf{B}$ from the subspace $\mathbb{A}$ is unique.

Is this sufficient condition satisfied, e.g., for the ideal Arnoldi approximation problem?

General characterization of spectral approximations

General characterization by [Lau and Riha, 1981] and [Zi̧etak, 1993, 1996]
$\longrightarrow$ based on the Singer's theorem [Singer, 1970].

General characterization of spectral approximations

General characterization by [Lau and Riha, 1981] and [Ziętak, 1993, 1996]
$\longrightarrow$ based on the Singer's theorem [Singer, 1970].

Define $|||\cdot|||$ (trace norm, nuclear norm, c_1-norm) and $\langle\cdot,\cdot\rangle$ by

$$|||\mathbf{X}||| \,=\, \sigma_1 + \cdots + \sigma_n\,, \quad \langle \mathbf{Z}, \mathbf{X}\rangle \,\equiv\, \mathrm{tr}(\mathbf{Z}^*\mathbf{X})\,.$$

General characterization of spectral approximations

General characterization by [Lau and Riha, 1981] and [Ziętak, 1993, 1996]
$\rightarrow$ based on the Singer's theorem [Singer, 1970].

Define $|||\cdot|||$ (trace norm, nuclear norm, c_1-norm) and $\langle\cdot,\cdot\rangle$ by

$$|||\mathbf{X}||| = \sigma_1 + \cdots + \sigma_n\,, \quad \langle \mathbf{Z}, \mathbf{X}\rangle \equiv \operatorname{tr}(\mathbf{Z}^*\mathbf{X})\,.$$

Characterization: [Ziętak, 1996] $\mathbf{A}_* \in \mathbb{A}$ is a spectral approximation of $\mathbf{B}$ from the subspace $\mathbb{A}$ iff there exists $\mathbf{Z} \in \mathbb{C}^{n\times n}$, s.t.

$$|||\,\mathbf{Z}\,||| = 1, \qquad \langle \mathbf{Z}, \mathbf{X}\rangle = 0, \quad \forall \mathbf{X} \in \mathbb{A}\,,$$

and

$$\operatorname{Re}\langle \mathbf{Z}, \mathbf{B} - \mathbf{A}_*\rangle = \|\mathbf{B} - \mathbf{A}_*\|\,.$$

Chebyshev polynomials of Jordan blocks

Theorem. Let $\mathbf{J}_\lambda$ be the $n \times n$ Jordan block. Consider the ideal Arnoldi approximation problem

$$\min_{p \in \mathcal{M}_m} \| p(\mathbf{J}_\lambda) \| = \min_{\mathbf{M} \in \mathbb{A}} \| \mathbf{B} - \mathbf{M} \| ,$$

where $\mathbf{B} = \mathbf{J}_\lambda^m$, $\mathbb{A} = \mathrm{span}\{\mathbf{I}, \mathbf{J}_\lambda, \ldots, \mathbf{J}_\lambda^{m-1}\}$. The minimum is attained by the polynomial $p_* = (z - \lambda)^m$ [Liesen and T., 2008].

Chebyshev polynomials of Jordan blocks

Theorem. Let $\mathbf{J}_\lambda$ be the $n \times n$ Jordan block. Consider the ideal Arnoldi approximation problem

$$\min_{p \in \mathcal{M}_m} \| p(\mathbf{J}_\lambda) \| = \min_{\mathbf{M} \in \mathbb{A}} \| \mathbf{B} - \mathbf{M} \| ,$$

where $\mathbf{B} = \mathbf{J}_\lambda^m, \; \mathbb{A} = \operatorname{span}\{\mathbf{I}, \mathbf{J}_\lambda, \ldots, \mathbf{J}_\lambda^{m-1}\}$. The minimum is attained by the polynomial $p_* = (z-\lambda)^m$ [Liesen and T., 2008].

Proof. For $p = (z-\lambda)^m$, the residual matrix $\mathbf{B} - \mathbf{M}$ is given by

$$\mathbf{B} - \mathbf{M} = p(\mathbf{J}_\lambda) = (\mathbf{J}_\lambda - \lambda \mathbf{I})^m = \mathbf{J}_0^m .$$

Chebyshev polynomials of Jordan blocks

Theorem. Let $\mathbf{J}_\lambda$ be the $n \times n$ Jordan block. Consider the ideal Arnoldi approximation problem

$$\min_{p \in \mathcal{M}_m} \| p(\mathbf{J}_\lambda) \| = \min_{\mathbf{M} \in \mathbb{A}} \| \mathbf{B} - \mathbf{M} \|,$$

where $\mathbf{B} = \mathbf{J}_\lambda^m$, $\mathbb{A} = \operatorname{span}\{\mathbf{I}, \mathbf{J}_\lambda, \ldots, \mathbf{J}_\lambda^{m-1}\}$. The minimum is attained by the polynomial $p_* = (z - \lambda)^m$ [Liesen and T., 2008].

Proof. For $p = (z - \lambda)^m$, the residual matrix $\mathbf{B} - \mathbf{M}$ is given by

$$\mathbf{B} - \mathbf{M} = p(\mathbf{J}_\lambda) = (\mathbf{J}_\lambda - \lambda \mathbf{I})^m = \mathbf{J}_0^m.$$

Define $\mathbf{Z} \equiv e_1 e_{m+1}^T$. It holds that

$$|||\mathbf{Z}||| = 1, \qquad \langle \mathbf{Z}, \mathbf{J}_\lambda^k \rangle = 0, \quad k = 0, \ldots, m-1$$

and

$$\langle \mathbf{Z}, \mathbf{B} - \mathbf{M} \rangle = \langle \mathbf{Z}, \mathbf{J}_0^m \rangle = 1 = \|\mathbf{B} - \mathbf{M}\|$$

Chebyshev polynomials of Jordan blocks

Theorem. Let $\mathbf{J}_\lambda$ be the $n \times n$ Jordan block. Consider the ideal Arnoldi approximation problem

$$\min_{p \in \mathcal{M}_m} \| \, p(\mathbf{J}_\lambda) \, \| \; = \; \min_{\mathbf{M} \in \mathbb{A}} \| \, \mathbf{B} - \mathbf{M} \, \| \, ,$$

where $\mathbf{B} = \mathbf{J}_\lambda^m, \; \mathbb{A} = \operatorname{span}\{\mathbf{I}, \mathbf{J}_\lambda, \ldots, \mathbf{J}_\lambda^{m-1}\}$. The minimum is attained by the polynomial $p_* = (z - \lambda)^m$ [Liesen and T., 2008].

Proof. For $p = (z - \lambda)^m$, the residual matrix $\mathbf{B} - \mathbf{M}$ is given by

$$\mathbf{B} - \mathbf{M} = p(\mathbf{J}_\lambda) = (\mathbf{J}_\lambda - \lambda \mathbf{I})^m = \mathbf{J}_0^m .$$

Define $\mathbf{Z} \equiv e_1 e_{m+1}^T$. It holds that

$$|||\mathbf{Z}||| = 1, \qquad \langle \mathbf{Z}, \mathbf{J}_\lambda^k \rangle = 0, \quad k = 0, \ldots, m-1$$

and

$$\langle \mathbf{Z}, \mathbf{B} - \mathbf{M} \rangle = \langle \mathbf{Z}, \mathbf{J}_0^m \rangle = 1 = \|\mathbf{B} - \mathbf{M}\| \quad \Rightarrow \quad \mathbf{M} = \mathbf{A}_* .$$

Ziętak's sufficient condition

Theorem [Ziętak, 1993]. If the residual matrix $\mathbf{B} - \mathbf{A}_*$ has an n-fold maximal singular value, then the spectral approximation $\mathbf{A}_*$ of $\mathbf{B}$ from the subspace $\mathbb{A}$ is unique.

For the ideal Arnoldi approximation problem and the Jordan block $\mathbf{J}_\lambda$, we have shown that

$$\mathbf{B} - \mathbf{A}_* = \mathbf{J}_0^m.$$

One is $(n-m)$-fold maximal singular value of $\mathbf{B} - \mathbf{A}_*$,
zero is m-fold singular value of $\mathbf{B} - \mathbf{A}_*$.

Ziętak's sufficient condition

Theorem [Ziętak, 1993]. If the residual matrix $\mathbf{B} - \mathbf{A}_*$ has an n-fold maximal singular value, then the spectral approximation $\mathbf{A}_*$ of $\mathbf{B}$ from the subspace $\mathbb{A}$ is unique.

For the ideal Arnoldi approximation problem and the Jordan block $\mathbf{J}_\lambda$, we have shown that

$$\mathbf{B} - \mathbf{A}_* = \mathbf{J}_0^m.$$

One is $(n-m)$-fold maximal singular value of $\mathbf{B} - \mathbf{A}_*$,
zero is m-fold singular value of $\mathbf{B} - \mathbf{A}_*$.

The spectral approximation is unique [Greenbaum and Trefethen. 1994], but, apparently, Ziętak's sufficient condition is not satisfied!

Outline

The problem and known results

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

where $\| \cdot \|$ is the spectral norm and f is a given polynomial.

The problem and known results

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

where $\| \cdot \|$ is the spectral norm and f is a given polynomial.

Known results

- If $\mathbf{A}$ is normal, the problem reduces to the well studied scalar approximation problem on the spectrum of $\mathbf{A}$,

$$\min_{p \in \mathcal{P}_m} \max_{\lambda \in \Lambda} | f(\lambda) - p(\lambda) | .$$

The problem and known results

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

where $\| \cdot \|$ is the spectral norm and f is a given polynomial.

Known results

- If $\mathbf{A}$ is normal, the problem reduces to the well studied scalar approximation problem on the spectrum of $\mathbf{A}$,

$$\min_{p \in \mathcal{P}_m} \max_{\lambda \in \Lambda} | f(\lambda) - p(\lambda) | .$$

- For general $\mathbf{A}$ - only a special case of $f(\mathbf{A}) = \mathbf{A}^{m+1}$ is known to have a unique solution [Greenbaum and Trefethen, 1994].

Reformulation of the problem

Let f be a polynomial of degree $m + \ell + 1$ ($m \geq 0,\ \ell \geq 0$). Then

$$f(z) \ = \ z^{m+1} g(z) + f_m z^m + \cdots + f_1 z + f_0 \,,$$

where g is a polynomial of degree at most ℓ. Approximate f by p

$$p(z) = p_m z^m + \cdots + p_1 z + p_0 \,.$$

Reformulation of the problem

Let f be a polynomial of degree $m + \ell + 1$ ($m \geq 0,\ \ell \geq 0$). Then

$$f(z) \ = \ z^{m+1}g(z) + f_m z^m + \cdots + f_1 z + f_0 \,,$$

where g is a polynomial of degree at most ℓ. Approximate f by p

$$p(z) = p_m z^m + \cdots + p_1 z + p_0 \,.$$

It is easy to show that

$$\min_{p \in \mathcal{P}_m} \|f(\mathbf{A}) - p(\mathbf{A})\| = \min_{h \in \mathcal{P}_m} \|\mathbf{A}^{m+1} g(\mathbf{A}) - h(\mathbf{A})\| \,.$$

Reformulation of the problem

Let f be a polynomial of degree $m+\ell+1$ ($m \geq 0,\ \ell \geq 0$). Then

$$f(z) \ = \ z^{m+1} g(z) + f_m z^m + \cdots + f_1 z + f_0\,,$$

where g is a polynomial of degree at most ℓ. Approximate f by p

$$p(z) = p_m z^m + \cdots + p_1 z + p_0\,.$$

It is easy to show that

$$\min_{p \in \mathcal{P}_m} \|f(\mathbf{A}) - p(\mathbf{A})\| = \min_{h \in \mathcal{P}_m} \|\mathbf{A}^{m+1} g(\mathbf{A}) - h(\mathbf{A})\|\,.$$

Without loss of generality we can consider the problem

$$\min_{h \in \mathcal{P}_m} \|\mathbf{A}^{m+1} g(\mathbf{A}) - h(\mathbf{A})\|\,,$$

where g is a given polynomial of degree at most ℓ.

Matrix polynomial approximation problems

We consider the problem

$$\min_{h \in \mathcal{P}_m} \| \mathbf{A}^{m+1} g(\mathbf{A}) - h(\mathbf{A}) \|,$$

where g is a given polynomial of degree at most ℓ.
For $g \equiv 1$ we obtain the ideal Arnoldi approximation problem.

Matrix polynomial approximation problems

We consider the problem

$$\min_{h \in \mathcal{P}_m} \| \mathbf{A}^{m+1} g(\mathbf{A}) - h(\mathbf{A}) \|,$$

where g is a given polynomial of degree at most ℓ.
For $g \equiv 1$ we obtain the ideal Arnoldi approximation problem.

Related problem:

$$\min_{g \in \mathcal{P}_\ell} \| \mathbf{A}^{m+1} g(\mathbf{A}) - h(\mathbf{A}) \|,$$

where h is a given polynomial of degree at most m.
For $h \equiv 1$ we obtain the ideal GMRES approximation problem.

Matrix polynomial approx. problems - general notation

We consider matrix approximation problems of the form

$$\min_{\mathbf{M}\in\mathbb{A}} \| \mathbf{B} - \mathbf{M} \| .$$

1.

$$\begin{aligned} \mathbf{B} &\equiv \mathbf{A}^{m+1} g(\mathbf{A}), \qquad g \in \mathcal{P}_\ell \text{ given} \\ \mathbb{A} &\equiv \operatorname{span}\{\mathbf{I}, \mathbf{A}, \ldots, \mathbf{A}^m\}, \end{aligned}$$

2.

$$\begin{aligned} \mathbf{B} &\equiv h(\mathbf{A}), \qquad h \in \mathcal{P}_m \text{ given} \\ \mathbb{A} &\equiv \operatorname{span}\{\mathbf{A}^{m+1}, \mathbf{A}^{m+2}, \ldots, \mathbf{A}^{m+\ell+1}\}. \end{aligned}$$

$\mathbf{B} \in \mathbb{C}^{n\times n}\backslash\mathbb{A}$ means that the minimum > 0.

Outline

Uniqueness results

Theorem [Liesen and T., 2008].

1. Given $g \in \mathcal{P}_\ell$, the problem

$$\min_{h \in \mathcal{P}_m} \| \mathbf{A}^{m+1} g(\mathbf{A}) - h(\mathbf{A}) \| > 0,$$

has the unique minimizer.

Uniqueness results

Theorem [Liesen and T., 2008].

1. Given $g \in \mathcal{P}_\ell$, the problem

$$\min_{h \in \mathcal{P}_m} \| \mathbf{A}^{m+1} g(\mathbf{A}) - h(\mathbf{A}) \| > 0,$$

has the unique minimizer.

2. Let $\mathbf{A}$ be nonsingular and $h \in \mathcal{P}_m$ given. Then the problem

$$\min_{g \in \mathcal{P}_\ell} \| \mathbf{A}^{m+1} g(\mathbf{A}) - h(\mathbf{A}) \| > 0,$$

has the unique minimizer.

Uniqueness results

Theorem [Liesen and T., 2008].

1. Given $g \in \mathcal{P}_\ell$, the problem

$$\min_{h \in \mathcal{P}_m} \| \mathbf{A}^{m+1} g(\mathbf{A}) - h(\mathbf{A}) \| > 0,$$

has the unique minimizer.

2. Let $\mathbf{A}$ be nonsingular and $h \in \mathcal{P}_m$ given. Then the problem

$$\min_{g \in \mathcal{P}_\ell} \| \mathbf{A}^{m+1} g(\mathbf{A}) - h(\mathbf{A}) \| > 0,$$

has the unique minimizer.

The nonsingularity in (2) cannot be omitted in general.

Idea of the proof (by contradiction)

Based on the proof by [Greenbaum and Trefethen, 1994].
Consider the problem

$$\min_{p \in \mathcal{G}_{\ell,m}^{(g)}} \|p(\mathbf{A})\|$$

where

$$\mathcal{G}_{\ell,m}^{(g)} \equiv \left\{ z^{m+1} g + h \,:\, g \in \mathcal{P}_\ell \text{ is given}, \, h \in \mathcal{P}_m \right\}.$$

Let q_1 and q_2 be two different solutions, $\|q_1(\mathbf{A})\| = \|q_2(\mathbf{A})\| = C$.

Idea of the proof (by contradiction)

Based on the proof by [Greenbaum and Trefethen, 1994].
Consider the problem

$$\min_{p \in \mathcal{G}_{\ell,m}^{(g)}} \|p(\mathbf{A})\|$$

where

$$\mathcal{G}_{\ell,m}^{(g)} \equiv \left\{ z^{m+1} g + h \,:\, g \in \mathcal{P}_\ell \text{ is given, } h \in \mathcal{P}_m \right\}.$$

Let q_1 and q_2 be two different solutions, $\|q_1(\mathbf{A})\| = \|q_2(\mathbf{A})\| = C$.

Use q_1 and q_2 to construct the polynomial

$$q_\epsilon \;=\; (1-\epsilon)\, q + \epsilon\, \widetilde{q} \;\in\; \mathcal{G}_{\ell,m}^{(g)}$$

and show that, for sufficiently small ϵ,

$$\|q_\epsilon(\mathbf{A})\| < C.$$

Outline

Ideal Arnoldi versus ideal GMRES polynomials

Ideal Arnoldi and ideal GMRES problems

$$\min_{p \in \mathcal{M}_m} \|p(\mathbf{A})\|\,, \qquad \min_{p \in \pi_m} \|p(\mathbf{A})\|\,.$$

The ideal Arnoldi polynomial is $(z-\lambda)^m$ [Liesen and T., 2008].

Ideal Arnoldi versus ideal GMRES polynomials

Ideal Arnoldi and ideal GMRES problems

$$\min_{p \in \mathcal{M}_m} \|p(\mathbf{A})\|, \qquad \min_{p \in \pi_m} \|p(\mathbf{A})\|.$$

The ideal Arnoldi polynomial is $(z-\lambda)^m$ [Liesen and T., 2008].

For $\lambda \neq 0$, we can write

$$(z-\lambda)^m = (-\lambda)^m (1-\lambda^{-1}z)^m .$$

$(1-\lambda^{-1}z)^m$ is a candidate for solving ideal GMRES problem.

Ideal Arnoldi versus ideal GMRES polynomials

Ideal Arnoldi and ideal GMRES problems

$$\min_{p \in \mathcal{M}_m} \|p(\mathbf{A})\|, \qquad \min_{p \in \pi_m} \|p(\mathbf{A})\|.$$

The ideal Arnoldi polynomial is $(z-\lambda)^m$ [Liesen and T., 2008].

For $\lambda \neq 0$, we can write

$$(z-\lambda)^m = (-\lambda)^m (1-\lambda^{-1}z)^m.$$

$(1-\lambda^{-1}z)^m$ is a candidate for solving ideal GMRES problem.

Is the ideal GMRES polynomial a scaled version of the ideal Arnoldi polynomial (at least for $\mathbf{J}_\lambda$)?

Ideal Arnoldi versus ideal GMRES polynomials

Ideal Arnoldi and ideal GMRES problems

$$\min_{p \in \mathcal{M}_m} \|p(\mathbf{A})\|, \qquad \min_{p \in \pi_m} \|p(\mathbf{A})\|.$$

The ideal Arnoldi polynomial is $(z-\lambda)^m$ [Liesen and T., 2008].

For $\lambda \neq 0$, we can write

$$(z-\lambda)^m = (-\lambda)^m (1-\lambda^{-1}z)^m.$$

$(1-\lambda^{-1}z)^m$ is a candidate for solving ideal GMRES problem.

Is the ideal GMRES polynomial a scaled version of the ideal Arnoldi polynomial (at least for $\mathbf{J}_\lambda$)?

No! Determination of ideal GMRES polynomials for $\mathbf{J}_\lambda$ is very complicated and intriguing problem [T., Liesen and Faber, 2007].

Ideal Arnoldi and ideal GMRES polynomials for $\mathbf{J}_\lambda$

Theorem [T., Liesen and Faber, 2007]. The mth ideal GMRES polynomial is $(1-\lambda^{-1}z)^m$ iff $0 \le m < \frac{n}{2}$ and $|\lambda| \ge \varrho_{m,n-m}^{-1}$.

$\varrho_{k,n}$ is the radius of the polynomial numerical hull of degree k of an $n \times n$ Jordan block (independent of λ).

Ideal Arnoldi and ideal GMRES polynomials for $\mathbf{J}_\lambda$

Theorem [T., Liesen and Faber, 2007]. The mth ideal GMRES polynomial is $(1-\lambda^{-1}z)^m$ iff $0 \le m < \frac{n}{2}$ and $|\lambda| \ge \varrho_{m,n-m}^{-1}$.

$\varrho_{k,n}$ is the radius of the polynomial numerical hull of degree k of an $n \times n$ Jordan block (independent of λ).

Example: Let n be even and consider $m = \frac{n}{2}$.

- If $|\lambda| \le 2^{-\frac{2}{n}}$, the ideal GMRES polynomial is 1.
- If $|\lambda| \ge 2^{-\frac{2}{n}}$, the ideal GMRES polynomial is equal to

$$\frac{2}{4\lambda^n+1} + \frac{4\lambda^n-1}{4\lambda^n+1}(1-\lambda^{-1}z)^{\frac{n}{2}}.$$

Ideal Arnoldi and ideal GMRES polynomials for $\mathbf{J}_\lambda$

Theorem [T., Liesen and Faber, 2007]. The mth ideal GMRES polynomial is $(1-\lambda^{-1}z)^m$ iff $0 \le m < \frac{n}{2}$ and $|\lambda| \ge \varrho_{m,n-m}^{-1}$.

$\varrho_{k,n}$ is the radius of the polynomial numerical hull of degree k of an $n \times n$ Jordan block (independent of λ).

Example: Let n be even and consider $m = \frac{n}{2}$.

- If $|\lambda| \le 2^{-\frac{2}{n}}$, the ideal GMRES polynomial is 1.
- If $|\lambda| \ge 2^{-\frac{2}{n}}$, the ideal GMRES polynomial is equal to

$$\frac{2}{4\lambda^n + 1} + \frac{4\lambda^n - 1}{4\lambda^n + 1}\,(1-\lambda^{-1}z)^{\frac{n}{2}}\,.$$

Obviously, neither 1 nor the above polynomial are scalar multiples of the corresponding ideal Arnoldi polynomial.

Summary

- We showed uniqueness of two matrix best approximation problems in spectral norm,

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

and

$$\min_{p \in \mathcal{P}_\ell} \| h(\mathbf{A}) - \mathbf{A}^{m+1} p(\mathbf{A}) \| .$$

Summary

- We showed uniqueness of two matrix best approximation problems in spectral norm,

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

and

$$\min_{p \in \mathcal{P}_\ell} \| h(\mathbf{A}) - \mathbf{A}^{m+1} p(\mathbf{A}) \| .$$

- Generalization of ideal Arnoldi and ideal GMRES problems.

Summary

- We showed uniqueness of two matrix best approximation problems in spectral norm,

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

and

$$\min_{p \in \mathcal{P}_\ell} \| h(\mathbf{A}) - \mathbf{A}^{m+1} p(\mathbf{A}) \| .$$

- Generalization of ideal Arnoldi and ideal GMRES problems.
- Nontrivial problem for nonnormal $\mathbf{A}$.

Summary

- We showed uniqueness of two matrix best approximation problems in spectral norm,

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

and

$$\min_{p \in \mathcal{P}_\ell} \| h(\mathbf{A}) - \mathbf{A}^{m+1} p(\mathbf{A}) \| .$$

- Generalization of ideal Arnoldi and ideal GMRES problems.
- Nontrivial problem for nonnormal $\mathbf{A}$.
- Ideal Arnoldi and Ideal GMRES polynomials can differ.

Summary

- We showed uniqueness of two matrix best approximation problems in spectral norm,

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

and

$$\min_{p \in \mathcal{P}_\ell} \| h(\mathbf{A}) - \mathbf{A}^{m+1} p(\mathbf{A}) \| .$$

- Generalization of ideal Arnoldi and ideal GMRES problems.
- Nontrivial problem for nonnormal $\mathbf{A}$.
- Ideal Arnoldi and Ideal GMRES polynomials can differ.
- Open question: When does the general problem

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

have a unique solution?

Related paper

- J. Liesen and P. Tichý, [On best approximations of polynomials in matrices in the matrix 2-norm, submitted, June 2008.]
- P. Tichý, J. Liesen and V. Faber, [On worst-case GMRES, ideal GMRES, and the polynomial numerical hull of a Jordan block, Electronic Transactions on Numerical Analysis (ETNA), Volume 26, pp. 453-473, published online, 2007.]

More details can be found at

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

Related paper

- J. Liesen and P. Tichý, [On best approximations of polynomials in matrices in the matrix 2-norm, submitted, June 2008.]
- P. Tichý, J. Liesen and V. Faber, [On worst-case GMRES, ideal GMRES, and the polynomial numerical hull of a Jordan block, Electronic Transactions on Numerical Analysis (ETNA), Volume 26, pp. 453-473, published online, 2007.]

More details can be found at

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

Thank you for your attention!