On Chebyshev Polynomials of Matrices

Petr Tichý

joint work with

Vance Faber and Jörg Liesen

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

June 6-11, 2010, Dolní Maxov Programy a algoritmy numerické matematiky 15 (PANM 15)

1

Chebyshev polynomials of a compact set

- \bullet Chebyshev polynomials on the interval [-1;1] [Chebyshev 1859].
- Generalized by [Georg Faber 1920] to the idea of the Chebyshev polynomials of Ω , where Ω is a compact set in the complex plane \mathbb{C} : These polynomials $T_m^{\Omega}(z)$ solve the problem

 $\min_{p \in \mathcal{M}_m(\Omega)} \| p(z) \|_{\infty}$

where \mathcal{M}_m is the class of monic polynomials of degree m.

Chebyshev polynomials of a compact set

- \bullet Chebyshev polynomials on the interval [-1;1] [Chebyshev 1859].
- Generalized by [Georg Faber 1920] to the idea of the Chebyshev polynomials of Ω , where Ω is a compact set in the complex plane \mathbb{C} : These polynomials $T_m^{\Omega}(z)$ solve the problem

 $\min_{p \in \mathcal{M}_m(\Omega)} \| p(z) \|_{\infty}$

where \mathcal{M}_m is the class of monic polynomials of degree m.

Example:

 $\boldsymbol{\Omega}$ is an interval, a set of discrete points, the unit circle, etc.

Chebyshev polynomials of normal matrices

Let $\mathbf{A} \in \mathbb{C}^{n imes n}$ be normal, i.e.,

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

Let $\|\cdot\|$ be the spectral norm and consider the problem

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

Chebyshev polynomials of normal matrices

Let $\mathbf{A} \in \mathbb{C}^{n imes n}$ be normal, i.e.,

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

Let $\|\cdot\|$ be the spectral norm and consider the problem

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

Then

$$\begin{split} \min_{p \in \mathcal{M}_m} \| \, p(\mathbf{A}) \| &= \min_{p \in \mathcal{M}_m} \| \, p(\mathbf{A}) \| = \min_{p \in \mathcal{M}_m(\Omega)} \| \, p(z) \|_{\infty} \end{split}$$
where $\Omega = \{\lambda_1, \dots, \lambda_n\}.$

Chebyshev polynomials of normal matrices

Let $\mathbf{A} \in \mathbb{C}^{n imes n}$ be normal, i.e.,

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

Let $\|\cdot\|$ be the spectral norm and consider the problem

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

Then

$$\begin{split} \min_{p\in\mathcal{M}_m}\|\,p(\mathbf{A})\| &= \min_{p\in\mathcal{M}_m}\|\,p(\mathbf{A})\| = \min_{p\in\mathcal{M}_m(\Omega)}\|\,p(z)\|_\infty \end{split}$$
 where $\Omega = \{\lambda_1, \dots, \lambda_n\}.$

The problem for \mathbf{A} is solved by the Chebyshev polynomial of Ω .

Chebyshev polynomials of general matrices

Let $\mathbf{A} \in \mathbb{C}^{n imes n}$ be a general matrix. We consider the problem

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

- Introduced in [Greenbaum, Trefethen 1994].
- Unique solution $T_m^A(z) \in \mathcal{M}_m$ exists if $m < d(\mathbf{A})$, [Greenbaum, Trefethen 1994; Liesen, T. 2009].
- $T_m^A(z)$ is called the *m*th Chebyshev polynomial of **A**, or the *m*th ideal Arnoldi polynomial of **A**.
- Previous work on these polynomials in [Toh PhD thesis 1996], [Toh, Trefethen 1998], [Trefethen, Embree 2005].
- Here: [Faber, Liesen, T. 2010].

Motivation

[Toh, Trefethen 1998] "Chebyshev polynomials of matrices are never far away from any discussion of convergence of Krylov subspace iterations in numerical linear algebra".

Motivation

[Toh, Trefethen 1998] "Chebyshev polynomials of matrices are never far away from any discussion of convergence of Krylov subspace iterations in numerical linear algebra".

GMRES and Arnoldi approximation problems:

[Greenbaum, Trefethen 1994]

 $\min_{p \in \pi_m} \| p(\mathbf{A})b \| \quad (\mathsf{GMRES}),$ $b \approx \{\mathbf{A}b, \dots, \mathbf{A}^mb\},$

$$\begin{split} \min_{q \in \mathcal{M}_m} \| \, q(\mathbf{A}) b \, \| & (\mathsf{Arnoldi}) \,, \\ \mathbf{A}^m b &\approx \{b, \dots, \mathbf{A}^{m-1} b\}. \end{split}$$

Motivation

[Toh, Trefethen 1998] "Chebyshev polynomials of matrices are never far away from any discussion of convergence of Krylov subspace iterations in numerical linear algebra".

GMRES and Arnoldi approximation problems:

[Greenbaum, Trefethen 1994]

$$\begin{split} \min_{p \in \pi_m} \| p(\mathbf{A})b \| \quad \text{(GMRES)}, \qquad \min_{q \in \mathcal{M}_m} \| q(\mathbf{A})b \| \quad \text{(Arnoldi)}, \\ b &\approx \{\mathbf{A}b, \dots, \mathbf{A}^m b\}, \qquad \qquad \mathbf{A}^m b \approx \{b, \dots, \mathbf{A}^{m-1}b\}. \end{split}$$

One may remove \boldsymbol{b} from the discussion and pose the following "ideal" approximation problems:

$$\begin{split} \min_{p \in \pi_m} \| p(\mathbf{A}) \| & (\text{Ideal GMRES}), \\ I &\approx \{\mathbf{A}, \dots, \mathbf{A}^m\}, \end{split} \qquad \begin{aligned} \min_{q \in \mathcal{M}_m} \| q(\mathbf{A}) \| & (\text{Ideal Arnoldi}), \\ \mathbf{A}^m &\approx \{I, \dots, \mathbf{A}^{m-1}\} \\ & (\text{Chebyshev polynomial of } \mathbf{A}) \end{aligned}$$

Motivation Example

Let $\lambda \in \mathbb{C}$. Consider an n by n Jordan block

$$\mathbf{J}_{\lambda} = \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix} \in \mathbb{C}^{n \times n}.$$

Question: How do the ideal GMRES and Chebyshev polynomials of \mathbf{J}_{λ} look like?

Motivation Example

Let $\ \lambda \in \mathbb{C}$. Consider an n by n Jordan block

$$\mathbf{J}_{\lambda} = \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix} \in \mathbb{C}^{n \times n}.$$

Question: How do the ideal GMRES and Chebyshev polynomials of \mathbf{J}_{λ} look like?

- Ideal GMRES polynomial of \mathbf{J}_{λ} a very difficult problem [T., Liesen, Faber 2007].
- Chebyshev polynomial of J_{λ} [Liesen, T. 2009]:

$$T_m^{\mathbf{J}_\lambda}(z) = (z - \lambda)^m$$







3 Matrices and sets in the complex plane







Shifts and scaling

Theorem



For $\mathbf{A} \in \mathbb{C}^{n \times n}$ and $\alpha \in \mathbb{C}$ the following hold: $\min_{p \in \mathcal{M}_m} \|p(\mathbf{A} + \alpha \mathbf{I})\| = \min_{p \in \mathcal{M}_m} \|p(\mathbf{A})\|,$ $\min_{p \in \mathcal{M}_m} \|p(\alpha \mathbf{A})\| = \|\alpha\|^m \min_{p \in \mathcal{M}_m} \|p(\mathbf{A})\|.$

- Shift invariance: Not surprising, because the polynomials are normalized at infinity.
- Paper contains explicit relations between the coefficients of $T_m^{\mathbf{A}}(z)$, $T_m^{\mathbf{A}+\alpha\mathbf{I}}(z)$, and $T_m^{\alpha\mathbf{A}}(z)$.

Example - shift of a matrix

Let $a, b \in \mathbb{R}$ be given. Consider the block-diagonal matrix \mathbf{A} with two $n \times n$ Jordan blocks,

$$\mathbf{A} \equiv \left[\begin{array}{cc} \mathbf{J}_a & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_b \end{array} \right].$$

Define

$$\alpha \equiv \frac{a+b}{2}.$$

Then

$$\mathbf{A} - \alpha \mathbf{I} \;=\; \left[\begin{array}{cc} \mathbf{J}_{\lambda} & 0 \\ 0 & \mathbf{J}_{-\lambda} \end{array} \right] \quad \text{where} \quad \lambda \;\equiv\; \frac{a-b}{2} \,,$$

and the previous theorem implies

$$\min_{p \in \mathcal{M}_m} \|p(\mathbf{A})\| = \min_{p \in \mathcal{M}_m} \|p(\mathbf{A} - \alpha \mathbf{I})\|.$$

Symmetry with respect to the origin

The Chebyshev polynomials of real intervals that are symmetric with respect to the origin are alternating between *even* and *odd*, i.e.

$$T_m^{[-a,a]}(z) = (-1)^m T_m^{[-a,a]}(-z).$$

Analogous result for Chebyshev polynomials of A?

Symmetry with respect to the origin

The Chebyshev polynomials of real intervals that are symmetric with respect to the origin are alternating between *even* and *odd*, i.e.

$$T_m^{[-a,a]}(z) = (-1)^m T_m^{[-a,a]}(-z).$$

Analogous result for Chebyshev polynomials of A?

Theorem[Faber, Liesen, T. 2010]Let $\mathbf{A} \in \mathbb{C}^{n \times n}$ and a positive integer $m < d(\mathbf{A})$ be given.If there exists a unitary matrix \mathbf{P} such that either $\mathbf{P}^* \mathbf{A} \mathbf{P} = -\mathbf{A}$ or $\mathbf{P}^* \mathbf{A} \mathbf{P} = -\mathbf{A}^T$,

then

$$T_m^{\mathbf{A}}(z) = (-1)^m T_m^{\mathbf{A}}(-z).$$

Example

$$\begin{split} \mathbf{A} \; = \; \left[\begin{array}{c} \mathbf{J}_{\lambda} \\ \mathbf{J}_{-\lambda} \end{array} \right], \quad \mathbf{P} \; = \; \left[\begin{array}{c} \mathbf{I}^{\pm} \\ \mathbf{I}^{\pm} \end{array} \right], \\ \text{where } \mathbf{I}^{\pm} \; = \; \mathrm{diag}(1, -1, 1, \dots, (-1)^{n-1}). \text{ Then} \\ \\ \mathbf{J}_{-\lambda} \; = \; -\mathbf{I}^{\pm} \mathbf{J}_{\lambda} \mathbf{I}^{\pm} \quad \Rightarrow \quad \mathbf{P}^{*} \mathbf{A} \mathbf{P} \; = \; -\mathbf{A} \; . \end{split}$$

Example

$$\mathbf{A} = \begin{bmatrix} \mathbf{J}_{\lambda} \\ \mathbf{J}_{-\lambda} \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} \mathbf{I}^{\pm} \\ \mathbf{I}^{\pm} \end{bmatrix},$$

here $\mathbf{I}^{\pm} = \operatorname{diag}(1, -1, 1, \dots, (-1)^{n-1}).$ Then

wł

$$\mathbf{J}_{-\lambda} \; = \; -\mathbf{I}^{\pm}\mathbf{J}_{\lambda}\mathbf{I}^{\pm} \quad \Rightarrow \quad \mathbf{P}^{*}\mathbf{A}\mathbf{P} \; = \; -\mathbf{A} \; .$$

Moreover

$$T_m^{\mathbf{A}}(\mathbf{A}) = \begin{bmatrix} T_m^{\mathbf{A}}(\mathbf{J}_{\lambda}) & \\ & T_m^{\mathbf{A}}(\mathbf{J}_{-\lambda}) \end{bmatrix},$$

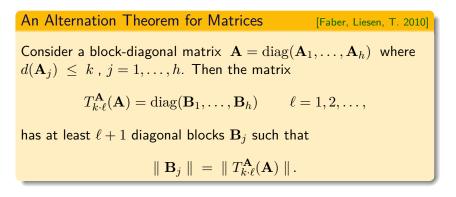
and

$$\| T_m^{\mathbf{A}}(\mathbf{J}_{-\lambda}) \| = \| \mathbf{I}^{\pm} T_m^{\mathbf{A}}(-\mathbf{J}_{\lambda}) \mathbf{I}^{\pm} \| = \| T_m^{\mathbf{A}}(\mathbf{J}_{\lambda}) \|,$$

i.e., the Chebyshev polynomial of ${\bf A}$ attains the same norm on each of the two diagonal blocks.

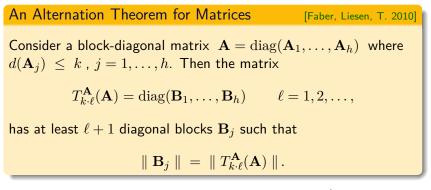
An alternation theorem

- Chebyshev polynomials for compact sets are characterized by alternation properties.
- Example: $T_m(z)$ for $[a,b] \subset \mathbb{R}$ has at least m+1 alternations.



An alternation theorem

- Chebyshev polynomials for compact sets are characterized by alternation properties.
- Example: $T_m(z)$ for $[a,b] \subset \mathbb{R}$ has at least m+1 alternations.



Example: If $\mathbf{A} = \operatorname{diag}(\lambda_1, \dots, \lambda_n) \in \mathbb{C}^{n \times n}$, then $T_m^{\mathbf{A}}(\mathbf{A})$ has at least m + 1 diagonal entries with the same maximal absolute value.

$$\mathbf{A} = \operatorname{diag}(\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \mathbf{A}_4)$$

where each $\mathbf{A}_j = \mathbf{J}_{\lambda_j}$ is a 3×3 Jordan block. The four eigenvalues are -3, -0.5, 0.5, 0.75, and $k = d(\mathbf{A}_j) = 3$.

$$\mathbf{A} = \operatorname{diag}(\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \mathbf{A}_4)$$

where each $\mathbf{A}_j = \mathbf{J}_{\lambda_j}$ is a 3×3 Jordan block. The four eigenvalues are -3, -0.5, 0.5, 0.75, and $k = d(\mathbf{A}_j) = 3$.

m	$\ T_m^{\mathbf{A}}(\mathbf{A}_1)\ $	$\ T_m^{\mathbf{A}}(\mathbf{A}_2)\ $	$\ T_m^{\mathbf{A}}(\mathbf{A}_3)\ $	$\ T_m^{\mathbf{A}}(\mathbf{A}_4)\ $
1	<u>2.6396</u>	1.4620	2.3970	<u>2.6396</u>
2	<u>4.1555</u>	<u>4.1555</u>	3.6828	<u>4.1555</u>
3	<u>9.0629</u>	5.6303	7.6858	<u>9.0629</u>
4	<u>14.0251</u>	<u>14.0251</u>	11.8397	<u>14.0251</u>
5	<u>22.3872</u>	20.7801	17.6382	<u>22.3872</u>
6	<u>22.6857</u>	<u>22.6857</u>	20.3948	<u>22.6857</u>
7	<u>26.3190</u>	<u>26.3190</u>	<u>26.3190</u>	<u>26.3190</u>







Perturbed Jordan block

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \nu & & & 0 \end{bmatrix} = \nu (\mathbf{J}_0^T)^{n-1} + \mathbf{J}_0 \in \mathbb{C}^{n \times n},$$

 $\nu \in \mathbb{C}$ is a complex parameter (studied by [Greenbaum 2009]). We have $d(\mathbf{A}) = n$ for any $\nu \in \mathbb{C}$.

Perturbed Jordan block

[Faber, Liesen, T. 2010]

For $1 \leq m \leq n-1$ and any $\nu \in \mathbb{C}$:

$$T_m^{\mathbf{A}}(z) = z^m$$

Special bidiagonal matrices

Given are $\lambda_1, \ldots, \lambda_\ell \in \mathbb{C}$ and $n \ge 1$. Consider

$$\mathbf{A} = \begin{bmatrix} \mathbf{D} & \mathbf{E} & & \\ & \mathbf{D} & \ddots & \\ & & \ddots & \mathbf{E} \\ & & & \mathbf{D} \end{bmatrix} \in \mathbb{C}^{\ell \cdot h \times \ell \cdot h},$$
$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 1 & & \\ & \lambda_2 & \ddots & \\ & & \ddots & 1 \\ & & & & \lambda_\ell \end{bmatrix} \in \mathbb{C}^{\ell \times \ell}, \quad \mathbf{E} = \begin{bmatrix} & & \\ & 1 \end{bmatrix} \in \mathbb{R}^{\ell \times \ell},$$

[Reichel, Trefethen 1992] related the pseudospectra of ${\bf A}$ to their symbol $f_{\bf A}(z)={\bf D}+z{\bf E}$.

Special bidiagonal matrices

[Faber, Liesen, T. 2010]

Consider the matrix ${\bf A}$ defined above. Let $\chi_{{\bf D}}(z)$ be the characteristic polynomial of ${\bf D},$

$$\chi_{\mathbf{D}}(z) = (z - \lambda_1) \cdot \ldots \cdot (z - \lambda_\ell).$$

Then

$$T_{k \cdot \ell}^{\mathbf{A}}(z) = (\chi_{\mathbf{D}}(z))^k, \qquad k = 1, 2, \dots, h - 1.$$

Special bidiagonal matrices

[Faber, Liesen, T. 2010]

Consider the matrix ${\bf A}$ defined above. Let $\chi_{{\bf D}}(z)$ be the characteristic polynomial of ${\bf D},$

$$\chi_{\mathbf{D}}(z) = (z - \lambda_1) \cdot \ldots \cdot (z - \lambda_\ell).$$

Then

$$T_{k \cdot \ell}^{\mathbf{A}}(z) = (\chi_{\mathbf{D}}(z))^k, \qquad k = 1, 2, \dots, h - 1.$$

Question: Can we find a set $S \subset \mathbb{C}$ such that

$$T^{\mathbf{A}}_{k \cdot \ell}(z) = T^{S}_{k \cdot \ell}(z)$$

for this matrix A?







3 Matrices and sets in the complex plane

Chebyshev polynomials $T_m^{\Omega}(z)$ of compact sets $\Omega \subset \mathbb{C}$

... unique polynomials that solve the problem

 $\min_{p \in \mathcal{M}_m} \max_{z \in \Omega} |p(z)|.$

Chebyshev polynomials of Ω and Ψ

[Kamo, Borodin 1994]

Let T_k^{Ω} be the *k*th Chebyshev polynomial of the infinite compact set $\Omega \subset \mathbb{C}$, let p(z) be a monic polynomial of degree ℓ , and let

$$\Psi \equiv p^{-1}(\Omega) = \{ z \in \mathbb{C} : p(z) \in \Omega \}$$

be the pre-image of Ω under the polynomial map $p. \label{eq:polynomial}$ Then

$$T^{\Psi}_{k \cdot \ell}(z) = T^{\Omega}_k(p(z)) \,.$$

Chebyshev polynomials for lemniscates

$$\mathbf{A} = \begin{bmatrix} \mathbf{D} & \mathbf{E} & & \\ & \mathbf{D} & \ddots & \\ & & \mathbf{D} & \\ & & \ddots & \mathbf{E} \\ & & & \mathbf{D} \end{bmatrix} \in \mathbb{C}^{\ell \cdot h \times \ell \cdot h}.$$

- Let $p(z) = (z \lambda_1) \cdots (z \lambda_\ell)$.
- The lemniscatic region $\mathcal{L}(p) \equiv \{z \in \mathbb{C} : |p(z)| \le 1\}$.
- $\Psi \equiv \mathcal{L}(p)$, $\Omega \equiv$ the unit circle.

Chebyshev polynomials of A and of $\mathcal{L}(p)$ [Faber, Liesen, T. 2010]

$$T_{k \cdot \ell}^{\mathcal{L}(p)}(z) = (p(z))^k = T_{k \cdot \ell}^{\mathbf{A}}(z), \qquad k = 1, 2, \dots, h-1.$$

Moreover,

$$\max_{z \in \mathcal{L}(p)} |T_{k \cdot \ell}^{\mathcal{L}(p)}(z)| = ||T_{k \cdot \ell}^{\mathbf{A}}(\mathbf{A})||.$$

- We considered Chebyshev polynomials of matrices and showed general properties (shifts and scaling, alternation).
- We can relate Chebyshev polynomials for lemniscatic regions to those for certain block-Toeplitz matrices.

- We considered Chebyshev polynomials of matrices and showed general properties (shifts and scaling, alternation).
- We can relate Chebyshev polynomials for lemniscatic regions to those for certain block-Toeplitz matrices.

Open question: Is it possible to translate the problem

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

into the problem

 $\min_{p \in \mathcal{M}_m} \max_{z \in \Omega} |p(z)|$

where Ω is a set in the complex plane associated with $\mathbf{A}?$

Related papers

- V. Faber, J. Liesen and P. Tichý, [On Chebyshev polynomials of matrices, accepted for publication in SIMAX (2010).]
- K-C. Toh, N. L. Trefethen,

[The Chebyshev polynomials of a matrix, SIMAX 20 (1999), no. 2, 400-419]

• A. Greenbaum and N. L. Trefethen, [GMRES/CR and Arnoldi/Lanczos as matrix approximation problems, SISC 15 (1994), no. 2, 359–368]

More details can be found at

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

Related papers

- V. Faber, J. Liesen and P. Tichý, [On Chebyshev polynomials of matrices, accepted for publication in SIMAX (2010).]
- K-C. Toh, N. L. Trefethen,

[The Chebyshev polynomials of a matrix, SIMAX 20 (1999), no. 2, 400-419]

• A. Greenbaum and N. L. Trefethen, [GMRES/CR and Arnoldi/Lanczos as matrix approximation problems, SISC 15 (1994), no. 2, 359–368]

More details can be found at

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

Thank you for your attention!